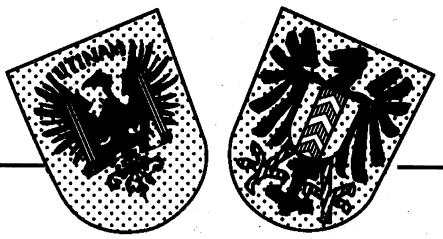


Proceedings of the

1999 Joint Meeting of
The EUROPEAN FREQUENCY AND TIME FORUM

and

The IEEE INTERNATIONAL FREQUENCY CONTROL SYMPOSIUM

13-16 APRIL 1999

99CH36313
87-654207

VOLUME 2

Proceedings of the

1999 Joint Meeting of
The
EUROPEAN FREQUENCY AND TIME FORUM
and The
IEEE INTERNATIONAL FREQUENCY CONTROL SYMPOSIUM

Sponsored by:

- The Institute of Electrical & Electronics Engineers, Inc. (Ultrasonics, Ferroelectrics and Frequency Control Society)
- Région de Franche-Comté
- Conseil Général du Doubs
- Ville de Besançon
- Bureau National de Métrologie (BNM, Paris)
- Services de l'Etat (Préfecture, DRIRE)

IEEE Catalog No. 99CH36313

Library of Congres No. 87-654207

13-16 April 1999

MICROPOLIS
Besançon, FRANCE

Volume 2

IEEE Catalog Number:	99CH36313	
ISBN:	0-7803-5400-1	Softbound Edition
	0-7803-5401-X	Casebound Edition
	0-7803-5402-8	Microfiche Edition
Library of Congress:	87-654207	
ISSN:	1075-6787	

1999 Joint Meeting
EFTF - IEEE IFCS

SYMPOSIUM CHAIRMEN

Co-Chairmen:	Raymond J. BESSON, ENSMM - France Donald B. SULLIVAN, NIST - USA
Technical Program Chairman:	Fred L. WALLS, NIST - USA
Conference Secretariat:	Isabelle BOURGON, ENSMM - France

TECHNICAL PROGRAM COMMITTEE

Charles Adams, Hewlett-Packard Co., USA
Jean-Pierre Aubry, Oscilloquartz SA, Switzerland
Claude Audoin, CNRS/LHA, France
Ivan Avramov, Institute of Solid State Physics, Bulgaria
Arthur Ballato, U.S. Army CECOM RDEC, USA
Andreas Bauch, Physikalisch-Technische Bundesanstalt, Germany
Ewald Benes, Institut fuer Allgemeine Physik, Austria
Raymond Besson, LCEP/ENSMM, France
Martin Bloch, Frequency Electronics Inc., USA
Rafael Boloix, Real Instituto y Observatorio de la Armada, Spain
Jean-Simon Boulanger, National Research Council of Canada, Canada
David Briggs, Frequency and Time Systems, USA
Jan Brown, JB Consulting, USA
Michel Brunet, CNES, France
Giovanni Busca, Observatoire Cantonal, Switzerland
Stephen Cantor, Mitre Corp., USA
James Cloeren, Johns Hopkins University, USA
Leonard Cutler, Hewlett-Packard Lab., USA
Gerrit De Jong, NMI Van Swinden Laboratory, The Netherlands
Philippe Defranould, Thomson Microsonics, France
Andrea DeMarchi, Politecnico di Torino, Italy
Francois Deyzac, ONERA, France
Michele DiFranza, Mitre Corp., USA
Klaus Dorenwendt, Physikalisch-Technische Bundesanstalt, Germany
John Dowsett, C-MAC Quartz Crystals Ltd., UK
Michael Driscoll, Northrop Grummen Corp., USA

Clemens Ruppel, SIEMENS, Germany
Wolfgang Schluter, Institut fur Angewandte Geodasie, Germany
Jesse Searles, Poseidon Scientific Instruments Pty. Ltd., Australia
Tadashi Shiosaki, Kyoto University, JapanHugues Duchaussoy, D.R.E.T., France
Errol Eernisse, Quartztronics Inc., USA
Christopher Ekstrom, U.S. Naval Observatory, USA
Pekka Eskelinen, Kotka Institute of Technology, Finland
Steve Feltham, ESA - European Space Agency ESTEC, The Netherlands
Raymond Filler, US Army CECOM, USA
Marvin Frerking, Rockwell International (Collins), USA
Michael Garvey, Frequency and Time Systems, USA
Aldo Godone, IEN Galileo Ferraris, Italy
Erhard Graf, Switzerland
Michel Granveaud, BNM/LPTF - Observatoire de Paris, France
Yuri Gulyaev, Russian Academy of Sciences, Russia
Jörg Hahn, DLR, Germany
William Hanson, Piezo Crystal Company, USA
Philipp Hartl, Universitat Stuttgart, Germany
Daniel Hauden, CNRS/LPMO, France
James Heighway, Chateau Mon Repos, France
Kurt Hilty, Swiss Federal Office of Metrology, Switzerland
Carl Hruska, York University, Canada
Bernardo Jaduszliwer, The Aerospace Corporation, USA
Steven Jefferts, NIST, USA
Charles Jensik, Piezo Crystal Company, USA
Gary Johnson, Sawyer Research Products, USA
Kalevi Kalliomaki, University Oulu, Finland
Peter Kartaschoff, Switzerland
Hirofumi Kawashima, Seiko Electronics Components Ltd., Japan
Masami Kihara, NTT Optical Network Systems Laboratories., Japan
Dieter Kirchner, Technische Universitat Graz, Austria
William Klepczynski, Innovative Solutions International Inc., USA
Alvin Kong, TRW, USA
John Kosinski, U.S. Army CECOM, USA
Peter Krempl, AVL List GmbH, Austria
Jack Kusters, Hewlett-Packard Co., USA
John Laverty, National Physical Laboratory, UK
Ryszard Lec, University of Maine, USA
Sigfrido Leschiutta, Instituto Electtrotecnico Nazionale, Italy
Lute Maleki, Jet Propulsion Laboratory, USA
Donald Malocha, University of Central Florida, USA
George Maronich, Frequency Management, USA
Steven Martin, Sandia National Labs., USA
Gary Montress, Raytheon Research Division, USA
Koji Nakagiri, Kinky University, Japan
Hartmut Nau, DLR-Deutsche Forschungsanstalt fur Luft-u, Germany
Bernd Neubig, Tele Quarz GmbH, Germany
Jerry Norton, Johns Hopkins Applied Physics Labs, USA
Thomas Parker, NIST, USA
John Prestage, Jet Propulsion Laboratory, USA
Leon Prost, Office Federal de Metrologie, Switzerland
T.J. Quinn, B.I.P.M., France
Andy Reid, British Telecom, UK

Victor Reinhardt, Hughes Space & Comm. Co., USA
Ulrich Rohde, Synergy Microwave Corp., USA
Robert Smythe, Piezo Technology Inc., USA
Waldemar Soluch, Institute of Electronic Materials Tech., Poland
James Steele, National Physical Laboratory, UK
Samuel Stein, Timing Solutions Corp., USA
Dan Stevens, Vectron Technologies, USA
Donald Sullivan, NIST, USA
Pierre Thomann, Observatoire Cantonal, Switzerland
Robert Tjoelker, Jet Propulsion Laboratory, USA
Adolph Uljanov, Kvarz Institute of Electronic Measurements, Russia
Mike Underhill, University of Surrey, UK
Rainer Unverdross, Unverdross Technik, Germany
Jacques Vanier, Canada
John Vig, U.S. Army CECOM, USA
Bruce Vojak, Motorola, Inc., USA
Fred Walls, NIST, USA
Roger Ward, Quartztronics, USA
Werner Weidemann, Datum Efratom, USA
Joseph White, U.S. Naval Research Laboratory, USA
Gernot Winkler, Innovative Solutions International, USA
Kazuhiko Yamanouchi, Tohoku University, Japan
Nicholas Yannoni, USA
Lidia Zhourkina, Russian Committee for the IEC, Russia
Walter Zingg, Micronas Semiconductor S.A., Switzerland

TECHNICAL SESSION CHAIRMEN

PLENARY SESSION
R. Besson, ENSMM, France
D. Sullivan, National Institute of Standards & Technology, USA

OPTICAL SYNTHESIS I
E. Bava, Politecnico di Milano, Italy

TWO WAY TIME TRANSFER
D. Kirchner, Technical University, Austria

SENSOR FUNDAMENTALS
R. Bourquin, ENSMM/LCEP, France

OPTICAL SYNTHESIS II
P. Gill, National Physical Laboratory, United Kingdom

INNOVATIONS ON GPS
D. Allan, Allan's TIME, USA

RESONATOR THEORY
J. Vig, US Army Communications-Electronics Command, USA

POSTER SESSION
H. Bahadur, National Physical Laboratory, India

OPTICAL SYNTHESIS III
L. Hollberg, National Institute of Standards & Technology, USA

TIME SCALES
P. Tavella, IEN, Italy

CHEMICAL SENSORS
J. Vig, US Army Communications-Electronics Command, USA

OPTICAL FREQUENCY STANDARDS
P. Gill, National Physical Laboratory, United Kingdom

TIME COMPARISON AND DISSEMINATION
G. DeJong, NMI Van Swinden Laboratory, The Netherlands

SENSOR INSTRUMENTATION
C. Ruppel, Siemens AG, Germany

OPTICAL FREQUENCY STANDARDS
J. Vanier, University of Montreal, Canada

OSCILLATOR MODELING
F. Walls, NIST, USA

LOW LOSS SAW FILTERS
P. Defranould, Thompson-Microsonics, France

FOUNTAINS
J. Boulanger, National Research Council of Canada

NOISE & RESONATORS
K. Hilty, Swiss P.O. Telecomm R&D, Switzerland

QUARTZ MATERIAL
J. Kosinski, US Army Communications-Electronics Command, USA

CESIUM CLOCKS
A. DeMarchi, Politecnico di Torino, Italy

MICROWAVE OSCILLATORS
P. Eskelinen, Lappeenranta University of Technology, Finland

NEW MATERIALS
C. Ruppel, Siemens AG, Germany

MEASUREMENTS GPS GLONASS
K. Kalliomaki, University of Oulu, Finland

FREQUENCY SYNTHESIS
M. Driscoll, Northrop-Grumman Corporation, USA

RESONATOR MANUFACTURING
J. Kusters, Hewlett-Packard Company, USA

MICROWAVE OSCILLATORS
T. Parker, National Institute of Standards and Technology, USA

REVIEW UNIQUE TECHNIQUES
R. Filler, US Army Communications-Electronics Command, USA

CLOCKS WITH COOLED SPECIES
S. Jefferts, National Institute of Standards and Technology, USA

SAW PROPAGATION
D. Hauden, CNRS-LPMO, France

CLOCKS IN SPACE
L. Maleki, Jet Propulsion Laboratory, USA

PLASTIC PARTS
G. Maronich, Frequency Management Inc., USA
M. DiFranza, The Mitre Corporation, USA

TABLE OF CONTENTS

VOLUME 1

TIME SCALES AND TIME-SCALE ALGORITHMS

TIME TRANSFER, DISSEMINATION, GPS AND GLONASS

CRYSTAL OSCILLATORS, FREQUENCY CONTROL CIRCUITRY AND PLASTIC PARTS

VOLUME 2

MICROWAVE OSCILLATORS AND MICROWAVE FREQUENCY SYNTHESIS

OPTICAL FREQUENCY STANDARDS AND OPTICAL FREQUENCY SYNTHESIS

PIEZOELECTRIC MATERIALS, RESONATOR THEORY, DESIGN AND ANALYSIS

SAW AND STW DEVICES

SENSORS AND ACTUATORS

NOISE AND OTHER STABILITY ISSUES

STABILITY AND PHASE NOISE TESTS OF TWO CRYO-COOLED SAPPHIRE OSCILLATORS*

G. John Dick and Rabi T. Wang

California Institute of Technology
Jet Propulsion Laboratory
4800 Oak Grove Drive
Pasadena, California 91109

Abstract

A cryocooled Compensated Sapphire Oscillator (CSO), developed for the Cassini Ka-band Radio Science experiment, and operating in the 7K - 10K temperature range was previously demonstrated to show ultra-high stability of $\sigma_y = 2.5 \times 10^{-15}$ for measuring times 200 seconds $\leq \tau \leq$ 600 seconds using a hydrogen maser as reference [1]. CSO-1 and CSO-3 are now both operational with new low noise receivers. We have made initial phase noise and Allan Deviation measurements that show more than ten times stability improvement over the hydrogen maser for measuring times 1 second $\leq \tau \leq$ 10 seconds, and indicate performance for the individual units of $\sigma_y \approx 3 \times 10^{-15}$ for measuring times from 10 to 1000 seconds. Phase noise is reduced by 20 to 28 dB over the design offset frequency range from 1 Hz to 40 Hz. Receiver design is also discussed.

1 Background

Cryogenic oscillators operating below about 10K offer the highest possible short term stability of any frequency sources. However, their use has so far been restricted to research environments due to the limited operating periods associated with liquid helium consumption. The cryocooled CSO is being built in support of the Cassini Ka-band Radio Science experiment and is designed to operate continuously for periods of a year or more. Performance targets are a stability of 3-4 $\times 10^{-15}$ (1 second $\leq \tau \leq$ 100 seconds) and phase noise of -73dB/Hz @ 1Hz measured at 34 GHz. Installation in 5 stations of NASA's deep space network (DSN) is planned in the years 2000 - 2002.

In the previous tests, actual stability of the CSO for measuring times $\tau \leq$ 200 seconds could not be directly measured, being masked by short-term fluctuations of the H-maser reference. Excellent short-term performance, however, could be inferred by the success of an application of the CSO as local oscillator (L.O.) to the JPL LITS passive atomic standard, where medium-term stability showed no degradation due to L.O. instabilities at a level of $\sigma_y = 3 \times 10^{-14}/\tau$. A second and third CSO have now been constructed, and all cryogenic aspects have been verified, including resonator turn-over temperatures of 7.907 K and 7.336K. Q's for both resonators are greater than 10^9. These values compare to a turn-over of 8.821 K and Q of 1.0×10^9 for the first resonator. Operation of this second unit provides a capability to directly verify for the first time the short-term (1 second $\leq \tau \leq$ 200 seconds) stability and the phase noise of the CSO units.

The RF receiver used in earlier tests was sufficient to meet Cassini requirements for $\tau \leq$ 10 seconds but had short-term stability limited to $\approx 4 \times 10^{-14}$ at $\tau = 1$ second, a value 10 times too high to meet our requirements. A new low-noise receiver has been designed with noise performance of $\approx 10^{-15}$ performance at 1 second. Short-term performance was degraded in the old receiver due to insufficient tuning bandwidth in a 100MHz quartz VCO that was frequency-locked to the cryogenic sapphire resonator. The new receivers are designed for sufficient bandwidth, loop gain and low noise to achieve the required performance.

2 Design Aspects

The mechanical support system for the cryocooler and dewar are shown in Fig. 1. The interpenetrating box design allows both Dewar and cryocooler supports to be very rigid, with mechanical resonances external to the dewar itself above 70 Hz. A principal point of contact is the helium-confining fiber-reinforced bellows between them. To reduce such coupled vibrations, the cryocooler is mounted as rigidly as possible to the floor. The dewar's box is supported by 8 commercial vibration isolators as can be seen in the figure.

Our experience to date indicates that this mounting system is successful at virtually eliminating vibration-

*This work was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

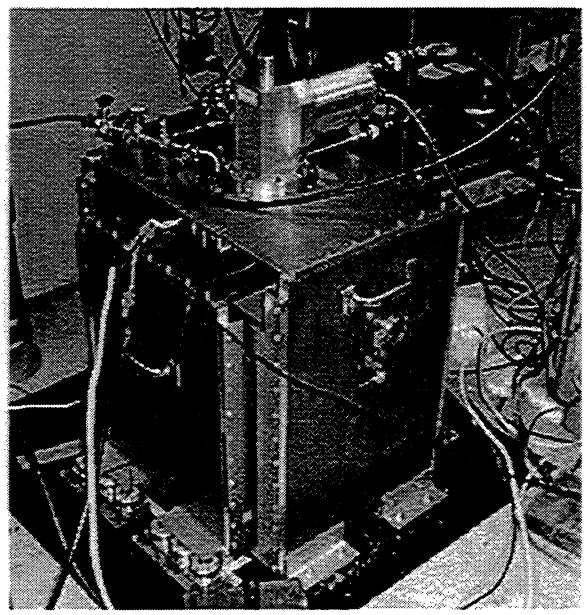


Figure 1: Interpenetrating boxes provide independent rigid mounting structures for the cryocooler and Dewar with mechanical resonant frequencies above 50 Hz. The cryocooler support structure is rigidly attached to the cement floor while the dewar is supported on shock-absorbing mounts.

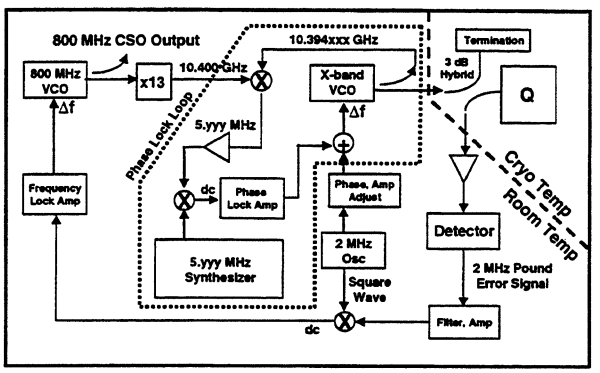


Figure 2: Schematic diagram of the "low noise" CSO receiver. An inner, phase locked, loop effects an offset of the X-band VCO frequency from the 13th harmonic of the 800 MHz primary VCO at 10.400 GHz to match the actual sapphire resonator frequency. In this way the (slower) Pound frequency lock circuit keeps the 800 MHz VCO operating "on frequency".

induced phase noise fluctuations below about 40 Hz. The small remaining fluctuations in this low frequency range appear to be due to helium gas expansion and contraction as the temperature of the cold head fluctuates each cycle. Vibrations above about 40 Hz are significant, limiting the excellent phase noise performance to lower frequencies.

The new "low noise" receiver is shown in Fig. 2. We find that stable operation of the outer "Pound" frequency lock loop requires a very tight inner loop. This in turn restricts the offset frequency in the inner loop to values above about 4 MHz. At lower offset frequencies, the outer loop becomes unstable at the gains desired.

We have ordered sapphire resonators with size constraints that give various frequencies centered on 10.400 GHz. All of the resonators are within 20 MHz as we had designed with most within about 6 very close to MHz. However, one is very close to 10.400 exactly, and will need to be reground to bring it at least 4 MHz away.

The preferred output frequency is 800 MHz, with performance difficult to verify at 100 MHz. All frequency stability comparisons reported here were done by differencing two 800 MHz signals.

3 Experimental

The temperature turn-over for CSO-3 is shown in Fig. 3. The actual value is a little lower than we had originally designed for, but cryocooler and dewar performance have been excellent, and weak spin-coupling

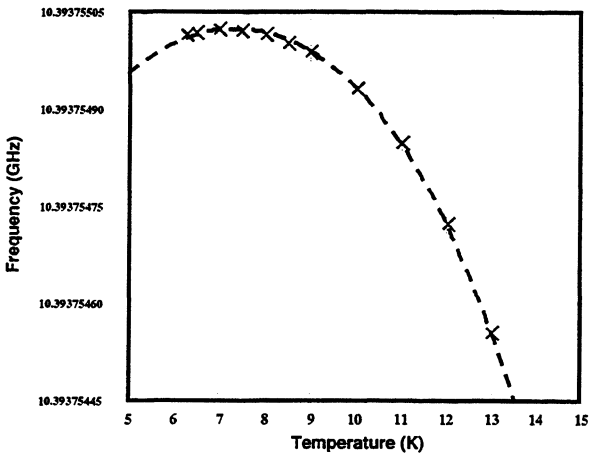


Figure 3: Temperature turnover characteristics for the CSO-3 resonator showing a curve fit with an extremum at 7.38 K. Turnover temperatures of the first two resonators were previously reported as 8.34 K and 7.907 K.

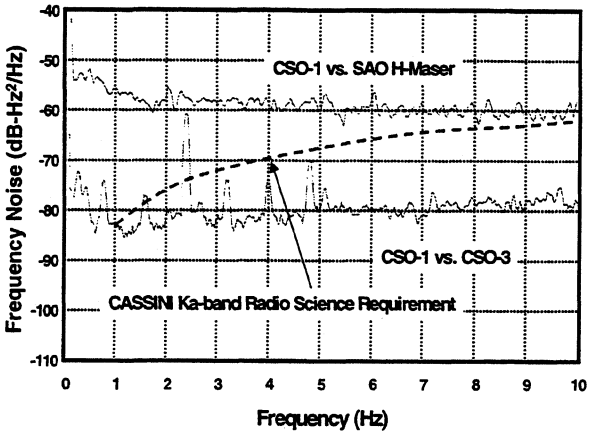


Figure 4: Measured frequency noise at 10.4 cavity frequency. One CSO was used as a frequency discriminator to make each of these measurements. Input frequency to the discriminator setup was 100 MHz for the Maser measurement and 800 MHz for the double CSO test. The CASSINI requirement corresponds to $S_\phi = -83$ dB/Hz at 10.4 GHz.

and low temperature both impact resonator Q. The resonator for CSO-3 is somewhat overcoupled, but still shows a loaded Q above 6×10^8.

CSO-1 and CSO-3 are now both operational with low-noise receivers installed. We have made initial phase noise and Allan Deviation measurements that validate our performance estimates.

Pair stability measurements show an Allan Deviation of 1.0×10^{-14} at 1 second and 5×10^{-15} at 10 seconds. Assuming both units are similar their individual performances are 7×10^{-15} at 1 second and 3.5×10^{-15} at 10 seconds. These results are approximately 10 times better than the hydrogen maser in this time range.

Pair stability measurements show an Allan Deviation of 1.3×10^{-14} at 1 second and $\approx 4 \times 10^{-15}$ for measuring times between 10 and 100 seconds. Assuming both units are similar their individual performances are 8×10^{-15} at 1 second and $\approx 2.5 \times 10^{-15}$ for 10 to 100 seconds. These results are approximately 10 times better than the hydrogen maser in this time range. phase noise measurements show a corresponding 20 to 28 db improvement over the best of the masers in the frequency range 1Hz to 40Hz. These tests verify that the noise performance of the CSO's meet the Cassini Ka-band requirement of -73dBc/Hz (referenced to Ka-band) all the way down to 1 Hz.

Phase noise measurements show a corresponding 24 to 28 db improvement over the best of the masers in the frequency range 1Hz to 40Hz. These tests verify that the noise performance of the CSO's meet the Cassini Ka-band requirement of -73dBc/Hz (referenced to Ka-band) all the way down to 1 Hz.

A small spectral bright line shows in these initial tests at the cryo-cooler cycle frequency of 2.4 Hz This small line could not be seen in previous tests, being masked by maser noise, and we expect to eliminate it with further optimization of the mechanical configuration. Based on our past experience it is likely due to helium pressure fluctuations, especially since the helium vents constructed for this purpose were not in use. Further improvement in the 1 second Allan Deviation is also expected.

4 Conclusions

The first tests of a pair of long-running cryocooled ultra-high stability short-term frequency standards demonstrated the short-term stability and low phase noise hoped for.

The 10K Compensated Sapphire Oscillator continues as the first continuously operable frequency standard with ultra–high short term stability. The first unit is expected to be installed at DSS-25 at NASA's Goldstone antenna complex in California in early '00. Passive atomic standards such as the LITS and Cesium Fountain can be operated continuously while realizing their inherent capabilities.

References

[1] G. J. Dick, R. T. Wang, and R. L. Tjoelker, "Cryocooled Sapphire Oscillator with Ultra-High Stability", *Proc. 1998 International IEEE Frequency Control Symposium*, pp. 528–533, 1998.

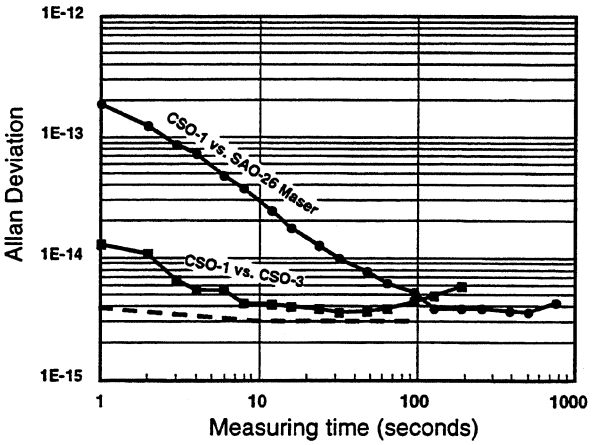


Figure 5: Preliminary stability measurements for two CSO units with the new receivers. Short term stability is much higher than for the hydrogen maser reference, and, on a per unit basis, is below 1×10^{-14} for all measuring times between 1 and 1000 seconds. The CASSINI requirement for a single unit is shown by the dashed line.

[2] J. Dick and R. T. Wang, "Cryocooled sapphire oscillator for the CASSINI Ka-band Experiment", *Proc. 1997 International IEEE Frequency Control Symposium*, pp. 1009–1014, 1997.

[3] R. T. Wang and G. J. Dick, "Improved Performance of the Superconducting Cavity Maser At Short Measuring Times", *Proceedings of the 44th Annual Frequency Control Symposium*, pp. 89–93, 1990.

[4] A. N. Luiten, A. G. Mann, N. J. McDonald and D. G. Blair, "Latest Results of the U. W. A. Cryogenic Sapphire Oscillator," *Proc. 1995 International Frequency Control Symposium*, pp. 433–437, 1995.

[5] Balzers KelCool 4.2GM Cryocooler from Leybold Vacuum Inc., Cryogenics Div., 8 Sagamore Park Road, Hudson, NH 03051-4914.

[6] A. G. Mann, A. N. Luiten, D. G. Blair, and M. J. Buckingham, "Ultra-stable Cryogenic Sapphire Dielectric Microwave Resonators", *Proceedings of the 46th Annual Frequency Control Symposium*, pp. 167-171, 1992.

[7] HEMEX sapphire from Crystal Systems, 27 Congress street, Salem, MA 01970 USA.

FUTURE TRENDS IN THE DEVELOPMENT OF ULTRA-LOW NOISE MICROWAVE OSCILLATORS WITH INTERFEROMETRIC SIGNAL PROCESSING

E. N. Ivanov and M. E. Tobar

Physics Department, The University of Western Australia, Nedlands, 6907, WA
Fax: 61 8 9380 1014, phone: 61 8 9380 3443, e-mail: eugene@physics.uwa.edu.au

1. Introduction

Application of microwave circuit interferometry to noise reduction in electromagnetic oscillators has enabled significant improvements in the performance of signal sources at microwave frequencies. For instance, the phase noise of X-band transistor oscillators based on the room temperature Sapphire Loaded Cavity (SLC) resonators was reduced to $-150\ dBc/Hz$ at Fourier frequency $f = 1\ kHz$ [1,2]. Such noise performance was primarily due to thermal fluctuations, which originate both inside the microwave interferometer as well as are coupled from the external circuitry.

In this paper we consider some problems associated with the design of ultra-low noise oscillators with interferometric signal processing. This includes discussion on (i) amplitude to frequency conversion in SLC resonators, (ii) novel amplitude noise suppression technique, and (iii) means of improving the long term frequency stability of microwave oscillators.

2. Amplitude to Frequency Conversion in Sapphire Loaded Cavity Resonators.

The amplitude to frequency conversion (AM to FM) in the SLC resonators results from the dependence of sapphire dielectric permittivity on temperature. The latter is altered by fluctuations of the dissipated microwave power giving rise to fluctuations of SLC resonant frequency.

The experimental setup for studying the power induced frequency fluctuations in the sapphire dielectric resonators included the SLC stabilised oscillator and an external frequency discriminator based on a hollow metal cavity tuned at the SLC resonant frequency. The tight frequency locking of the oscillator to the SLC resonance was implemented to ensure that oscillator frequency closely followed the SLC resonant frequency. To vary power dissipated in the SLC resonator the frequency stabilised oscillator was complemented by power control system based on a voltage controlled attenuator (VCA) placed in front of the power amplifier in the loop oscillator.

The dependence of the SLC resonant frequency on power was measured by introducing a pilot signal into the power control system and monitoring both power incident of the resonator and voltage at the output of external frequency discriminator. These measurements were performed in the range of Fourier frequencies from 15 mHz to 1 Hz and it was found that power to frequency transfer function of the room temperature SLC resonator behaves as that of the first order low-pass filters and is given by

$$\frac{df_{res}}{dP} = \frac{\alpha}{1 + j2\pi f\tau_{th}}. \qquad (1)$$

For a room temperature SLC resonator operating at frequencies around 9 GHz the values of α and τ_{th} were measured to be equal to 0.4 kHz/mW and 34 s, respectively. These results are reasonably consistent with numerical estimates of α and τ_{th} which could be obtained from knowing the thermal conductivity and specific heat of sapphire as well as the geometry of the dielectric resonator.

The above mechanism of power to frequency conversion gives rise to oscillator phase noise with spectral density

$$S_{\varphi}^{osc}(f) = 4P_{diss}^2 \frac{\alpha^2}{1 + (2\pi f\tau_{th})^2} \frac{S_{AM}^{osc}(f)}{f^2}, \qquad (2)$$

where P_{diss} is a power dissipated in the resonator and $S_{AM}^{osc}(f)$ is the spectral density of oscillator amplitude fluctuations. As $S_{AM}^{osc}(f) \propto 1/f$ for a typical microwave transistor oscillator, this results in the phase noise spectral density varying at a rate of 50 $dB/decade$ at low Fourier frequencies above $f \approx 1/\tau_{th} \approx 20\ mHz$.

Such an instability of SLC resonant frequency induced by power fluctuations has serious implications for the design of ultra-low phase noise oscillators making the goal of improving their phase noise performance a difficult task. For instance, from the results of simulations it folows that at $P_{diss} \approx 500\ mW$ and $f \approx 1\ Hz$ the spectral density of the excess phase noise due to AM to FM conversion is almost 70 dB higher than the thermal noise floor at.

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

Remembering that the thermal noise floor scales inversely proportional with power P_{diss} [1], while the intensity of phase fluctuations resulting from the AM to FM conversion in the SLC resonator is proportional to P_{diss}^2 (2), one can arrive at rather pessimistic conclusion that the medium term frequency stability of the high power microwave oscillator based on the SLC resonator is always going to be sacrificed if the thermal noise limited performance at high Fourier frequencies is to be achieved. Fortunately, there are ways of controlling the AM to FM conversion in the SLC resonators. They include:

(i) frequency-temperature compensation of sapphire dielectric resonators [3];
(ii) suppression of oscillator amplitude noise;
(iii) controlling the SLC operating temperature by making use of the difference in frequency-temperature coefficients of different modes of the SLC resonator.

As far as the frequency-temperature compensation of SLC resonators is concerned, at room temperatures it is usually associated with relatively large loss in the resonator Q-factor. For this reason we will discuss the principles of amplitude noise reduction in microwave oscillators and the concept of a dual mode oscillator where two resonant modes with orthogonal polarisations are excited in the same SLC resonator.

3. AM Noise Suppression in Oscillators with Interferometric Signal Processing

Considering the oscillator with automatically balanced microwave interferometer [1], the carrier in such oscillator is suppressed due to the joint operation of two feedback control systems. The frequency control system maintains the phase balance of the interferometer, while the amplitude control system keeps the amplitude mismatch between the interferometer arms at minimum. Provided that the carrier suppression point, f_{cs}, is close to the SLC resonant frequency, f_{res}, the sensor of the amplitude mismatch control system appears to be sensitive to oscillator amplitude fluctuations. The efficiency of amplitude to voltage conversion, $dU/d\alpha$, calculated as a function of frequency offset, $f_{cs} - f_{res}$, at Fourier frequency $f = 50\ Hz$ is shown in Figure 1.

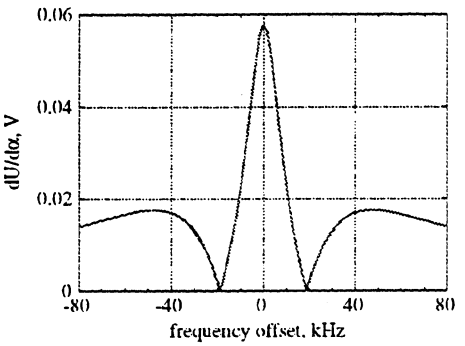


Figure 1. Sensitivity of amplitude sensor as a function of frequency offset, $f_{cs} - f_{res}$

It should be noted that such a sensor has a drawback which consists in the reduced sensitivity to slow fluctuations of oscillator amplitude. For instance, at $f_{cs} = f_{res}$, the value of $dU/d\alpha$ is proportional to

$$\frac{f/\Delta f_{0.5}}{\sqrt{1+(f/\Delta f_{0.5})^2}}, \qquad (3)$$

where $\Delta f_{0.5}$ is the half-loaded bandwidth of the SLC resonator. From (3) it is clear that implementing an oscillator amplitude noise reduction system based on the interferometric AM discriminator is not going to significantly affect the oscillator low frequency phase noise caused by AM to FM conversion in the SLC resonator. Nevertheless, the development of interferometric AM noise reduction system can still be well justified because of its high efficiency at high Fourier frequencies. This is illustrated by the results of modeling presented in Figure 2.

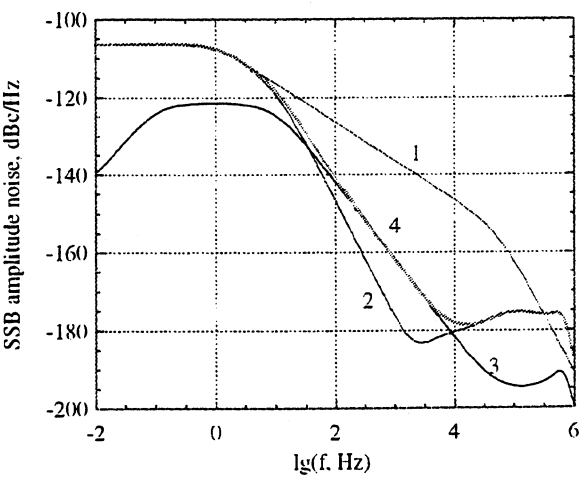


Figure 2. Oscillator amplitude noise budget ($f_{res} = 9\ GHz$, $P_{diss} \approx 500\ mW$)

Curve 1 in Figure 2 corresponds to the amplitude noise of a free-running oscillator. Curve 2 shows the noise floor, which could be potentially achieved, if the sensor of AM noise reduction system was noiseless. Curve 3 represents the noise floor of the

interferometric amplitude discriminator. This noise floor is due to the amplitude discriminator effective noise temperature and depends on the quality of carrier suppression. The spectral density of overall amplitude fluctuations of such oscillator is given by curve 4. It varies as $1/f^2$ in the frequency range 10 Hz...10 kHz and is equal to -160 dBc/Hz at $f = 1$ kHz. This result, apart from being at least 25 dB below the free-running oscillator AM noise, also means that in the high power oscillators with $P_{diss} \approx 500$ mW the total noise power in the above range of Fourier frequencies is going to be almost equally divided between phase and amplitude fluctuations.

4. Improving the Medium Term Frequency Stability of SLC Based Microwave Oscillators

The idea of having two electromagnetic oscillators sharing the same resonator and locked to its different modes in order to improve the oscillators immunity to ambient temperature fluctuations has been earlier applied to the development of precision oscillators at RF and microwave frequencies [4, 5]. In this section we will consider how the concept of such a dual mode oscillator can benefit the SLC based oscillators with interferometric signal processing.

In simplified terms the principle of operation of a dual mode oscillator requires that frequencies of two oscillators depend differently of some common parameter, for instance the temperature of the resonator. In terms of rms frequency fluctuations this can be expressed as

$$\delta f_{H,E} = \frac{df_{res}^{H,E}}{dT}\delta T_s \qquad (4)$$

where δT_s is the rms amplitude of temperature fluctuations of the SLC active region (where the most of electromagnetic energy is concentrated), $df_{res}^{H,E}/dT$ are the resonator frequency temperature coefficients for transverse electric (H) and transverse magnetic (E) modes, respectively. At room temperatures: $df_{res}^{H}/dT \approx 500\,kHz/K$ and $df_{res}^{E}/dT \approx 700\,kHz/K$.

Assuming that some feedback control system makes the beat frequency between 'H' and 'E' oscillators exactly equal to the frequency of an external reference oscillator, the SLC temperature fluctuations become entirely determined by frequency fluctuations of the reference oscillator, δf_{osc}. In other words, the following relationship between temperature and frequency fluctuations is satisfied:

$$\delta T_s = \frac{\delta f_{osc}}{v_{HE}}, \qquad (5)$$

where $v_{HE} = \left| df_{es}^{H}/dT - df_{res}^{E}/dT \right|$.

Substituting (5) into (4) yields

$$\delta f_{H,E} = \frac{df_{res}^{H,E}/dT}{v_{HE}}\delta f_{osc} \qquad (6)$$

Making transition from frequency to time domain, the analytical expression for the Square Root Allan Variance (SRAV) of H oscillator is obtained

$$\sigma_y^H = \frac{df_{res}^{H}/dT}{v_{HE}}\frac{f_{osc}}{f_H}\sigma_y^{osc}, \qquad (7)$$

where f_H is the H-oscillator operating frequency, f_{osc} and σ_y^{osc} are the operating frequency and SRAV of reference oscillator respectively. A similar expression for the SRAV of E-oscillator is obtained from (7) by simply replacing index H with E.

Assuming that separation between H and E modes in the SLC resonator is equal to 100 MHz and a quartz oscillator with $\sigma_y^{osc} \approx 10^{-11}$ is used to bridge this frequency gap, the estimate for the SRAV of a 10 GHz H-oscillator is $2.5 \cdot 10^{-13}$.

There are various means of phase locking two microwave oscillators, which share the same SLC resonator. The natural way of doing this involves varying parameters, which affect both SLC resonant frequencies simultaneously, for instance, by varying the SLC operating temperature or dissipated microwave power.

The phase locking can also be obtained by coupling the varactor diode to SLC resonator and tuning its frequencies electronically. The drawbacks of this technique are associated with the high level of varactor self noise and the necessity to match voltage tuning coefficients of SLC resonant frequencies with those due to temperature variations.

Another way of achieving the phase locking of two microwave oscillators relies on the concept of electronically tunable oscillators described in [1]. Once again, the critical problem here would be to match the oscillator frequency tuning coefficients with those of SLC resonant frequencies.

Below we will summarise the results of noise analysis of a dual mode oscillator in which the constant frequency offset between H and E oscillators is maintained by varying the level of dissipated microwave power. We assume that both oscillators are almost identical except for their output power: the H-oscillator dissipates 500 mW in the resonator,

while the P_{diss} due to E-oscillator is 50 mW. We also assume that the correction signal produced by the phase lock loop is applied to the voltage controlled attenuator in the low power E-oscillator. This is because such a feedback increases AM noise of E-oscillator (pilot oscillator) but leaves the H-oscillator (master oscillator) AM noise unchanged

Linearising the characteristic equation of equivalent power control system, the analytical expressions for the components of master oscillator phase noise floor are found. For instance, sapphire temperature fluctuations, δT_s, impose the following limit on the master oscillator rms frequency fluctuations:

$$\delta f^H_{rms} = \delta T_s \frac{df^H_{res}}{dT}\left|\frac{1}{1+\gamma}\right|, \qquad (8.1)$$

where γ is a phase lock loop gain. Eq. 8.1. implies that temperature induced fluctuations of the SLC resonant frequency are suppressed by the amount of phase lock loop gain.
The noise floor due to reference oscillator frequency fluctuations is given by

$$\delta f^H_{rms} = \frac{df^H_{res}/dT}{\nu_{HE}}\delta f_{osc}\left|\frac{\gamma}{1+\gamma}\right| \qquad (8.2)$$

Eq. 8.2. gives a rigorous relationship between δf_{osc} and df^H_{rms} which takes into account the dynamic properties of the phase lock loop in contrast to almost phenomenological eq. 6.
Intrinsic fluctuations in the frequency discriminator of master oscillator (caused by such factors as circulator phase noise, frequency discriminator effective noise temperature etc.) result in the following noise floor

$$\delta f^H_{rms} = \delta f^H_{FD}\left\{1-\frac{df^H_{res}/dT}{\nu_{HE}}\left|\frac{\gamma}{1+\gamma}\right|\right\} \qquad (8.3)$$

A similar expressions for H-oscillator rms frequency fluctuations caused by the noise sources in the frequency discriminator of pilot oscillator is given by

$$\delta f^H_{rms} = \delta f^E_{FD}\frac{df^H_{res}/dT}{\nu_{HE}}\left|\frac{\gamma}{1+\gamma}\right| \qquad (8.4)$$

Eq. 8.3 and 8.4 indicate that frequency fluctuations due to δf^H_{FD} and δf^E_{FD} (as well as due to frequency noise of reference oscillator δf_{osc}) are not suppressed by the phase lock loop. This is an extra argument in favour of choosing the frequency discriminators with interferometric signal processing and low noise quartz oscillator as a source of reference signal.

Equations 8.1-8.4 (a more detailed noise analysis will be published elsewhere) have been used for modeling the noise performance of a dual mode oscillator. The results of such modeling are presented in Figure 3.

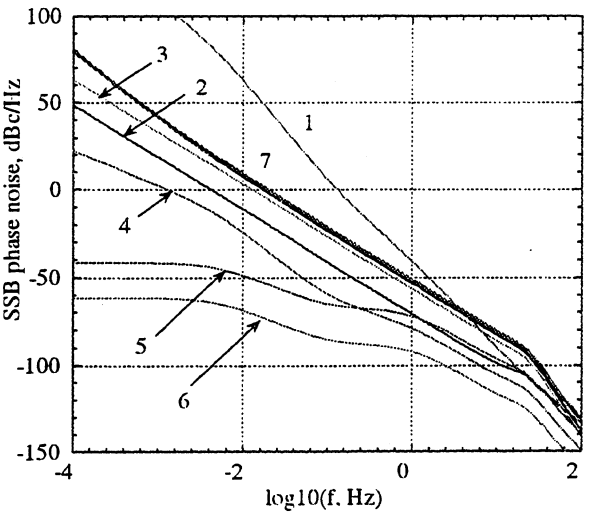


Figure 3. Components of phase noise floor of master-oscillator due to various noise mechanisms

Curve 1 shows the master oscillator phase noise when phase lock loop is open. Curve 2 gives the phase noise floor due to intrinsic fluctuations in the frequency discriminator of master oscillator (see eq. 8.3). Fluctuations in the frequency discriminator of the pilot oscillator are responsible for the noise floor given by curve 3 (see eq.8.4). The 10 dB offset between curves 2 and 3 is mostly due to assumed 10 dB difference in oscillator powers.

The effect of the phase lock loop on the master oscillator phase noise caused by temperature and dissipated power fluctuations is given by curves 4 and 5, 6 respectively.
The total oscillator phase noise is given by curve 7 which closely coincides with the limit imposed by the reference oscillator phase fluctuations. This is a clear illustration of the effectiveness of the phase lock loop which makes contribution of such noise sources as temperature and power fluctuations in a dual mode oscillator almost negligible as compared with that in a single mode oscillator.

The above calculations were performed assuming the reference oscillator with SRAV $\sigma_y^{osc} = 10^{-11} + 1.5\cdot 10^{-12}\sqrt{\tau}$. Apart from that, the SRAV of relative ambient temperature fluctuations was measured to be equal to $\sigma_y^T = 3\cdot 10^{-5}\left(1+\sqrt{\tau}\right)$. This dependence along with the knowledge of the SLC resonator thermal time constant, τ_{th}, and its

residual frequency-temperature coefficient has enabled the spectral density of sapphire crystal temperature fluctuations, δT_s (eq. 8.1) to be evaluated.

Knowing the spectrum density of oscillator phase noise allows its frequency stability in the time domain to be calculated. Such calculations were carried out for three different states of master oscillator. First, a free-running master oscillator with frequency control loop off was analysed.

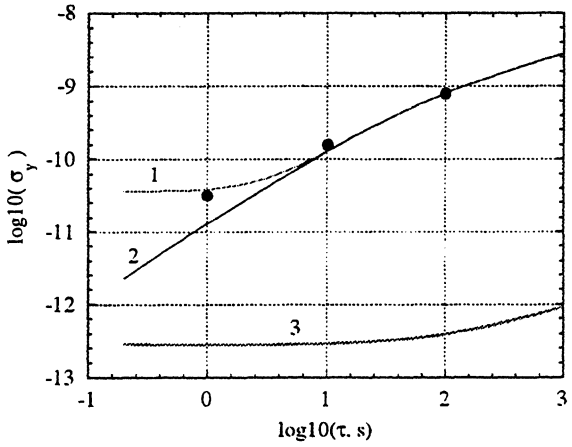


Figure 4. Calculated dependencies of SRAV of master oscillator as functions of integration time

In this case the dependence of SRAV on integration time, τ, is shown in Figure 4 (curve 1). Secondly, the SRAV of a frequency stabilised master oscillator was calculated (curve 2 in Figure 4). Curves 1 and 2 converge at $\tau \geq 10$ s. This is because at such values of τ, ambient temperature fluctuations represent a common mechanism of frequency instability of both oscillators. From Figure 4 it is also clear that the SRAV of frequency stabilised oscillator rapidly improves at short integration times as compared with that of a free-running oscillator. This is a result of the phase noise suppresson by interferometric frequency control system.

Dark circles in Figure 4 correspond to the experimental results obtained at Poseidon Scientific Instruments Pty. Ltd. with two frequency stabilised oscillators ($P_{diss} \approx 300$ mW) with automatically balanced interferometers. The excessively high value of SRAV at $\tau = 1$ s was likely due to the measurement system noise floor.

Closing the phase lock loop results in almost constant value of $\sigma_y^H \approx 3 \cdot 10^{-13}$ for integration times up to 100 s (curve 3). This limit is more than 3 orders of magnitude below than that of a frequency stabilised master oscillator (curve 2) at $10\ s \leq \tau \leq 100\ s$. Such a limit is imposed by frequency instability of the reference oscillator (see eq. 7), which dominates over other noise sources at low Fourier frequencies (see Figure 3).

5. Conclusion

This work has outlined two possible trends in the field of ultra-low noise microwave oscillators. First, it has been shown that applying principles of interferometric signal processing enables the design of microwave oscillators, which combine excellent phase and amplitude noise performance.

Secondly, the means of improving the medium term frequency stability of microwave oscillators have been investigated. It was shown that the frequency stability of a dual mode microwave oscillator can be significantly better than that of a flywheel quartz oscillator which is used for the phase locking of its two spectral components.

6. Acknowledgement

This work is a collaboration between the University of WA and Poseidon Scientific Instruments Pty. Ltd. (PSI) supported by Australian Research Council. The authors are very grateful to Mr. J. H. Searls, Managing Director of PSI and Mr. C. McNeilage, Senior Staff Member of PSI, for the assistance in conducting experiments and useful discussions.

References

1. E. N. Ivanov, M. E. Tobar and R. A. Woode, 'Applications of interferometric signal processing to phase-noise reduction in microwave oscillators', IEEE Trans. On MTT, vol. 46, N10, pp. 1537-1545, 1998.

2. Phase noise detector', international patent (PCT/WO95/32435 priority 25.05.94)
US patent 5,841,322 on 24 November 1998.

3. J. G. Hartnett, M. E. Tobar, A. G. Mann et al, 'Frequency-temperature compensation in Ti3+ and Ti4+ doped sapphire whispering gallery mode resonators', in Proc. of 1998 IEEE Frequency Control Simposium, pp. 512-518, 1998

4. F. L. Walls, 'Fundamental limits on Frequency Stabilities of Crystal Oscillators', IEEE Trans. On UFFC, Vol. 42, N4, pp. 576-589, 1995

5. D. P. Tsarapkin and S. L. Abramov, Private publication.

1999 Joint Meeting EFTF - IEEE IFCS

A NEW METHOD FOR THE DESIGN OF ULTRA LOW NOISE OSCILLATORS

E. VAURY* - V. GIORDANO** - J.C. NALLATAMBY*- J. GROSLAMBERT**-
M. CAMIADE*** - M. PRIGENT* - O. LLOPIS **** - E.GONZALEZ*****-
M. CHAUBET***** - J. OBREGON*

*IRCOM – 7, rue Jules Vallès – 19100 BRIVE – France
**LPMO – 32 , Av. de l'Observatoire – 25044 BESANCON – France
***UMS – Domaine de Corbeville – 91404 ORSAY – France
****LAAS – Av. du Colonel Roche – 31077 TOULOUSE – France
*****CNES – 18 , Av. Edouard Belin – 31055 TOULOUSE – France

ABSTRACT

We describe a newly developed design method of free running oscillator circuits leading to the minimum phase noise, for given transistor and resonator. This method has allowed us to design a 9.2 GHz oscillator using a PHEMT transistor and a sapphire resonator stabilized at 330 K, presenting a measured phase noise of –80 dBc/Hz at 100Hz offset from carrier with an $1/f^3$ slope. This result represents the state of the art of HEMTs free running oscillators at room temperature.

1. PRINCIPLE OF OPERATION

It has been shown recently [1], [2], [3],[4], that the phase noise of microwave oscillators may be drastically improved by adding complementary stabilizing functions to the free-running circuits. Nevertheless, first at all, it is necessary to design a very low phase-noise fundamental free-running oscillator circuit [5], [6] As it was pointed out by Everard the well known equation (1) used to describe the noise performance in free running oscillator [1] is generally misunderstood.

$$L(f_m) = \frac{G.F.K.T}{8P_{osc}}\left(1+\frac{f_c}{f_m}\right)\frac{f_0^2}{f_m^2}\frac{1}{Q_L^2} \qquad (1)$$

where :

$L(f_m)$ is the ratio of sideband power in a 1 Hz bandwidth at f_m to the total power, f_m is the frequency offset, f_0 is the oscillation frequency, f_c is the flicker corner frequency, Q_L is the loaded quality factor, F is the noise figure, K is the Boltzmann's constant, T is the Kelvin temperature, P_{osc} is the average power at the active device output, and G is the amplifier gain.

To be consistent this equation must describe a transistor oscillator as well as a negative conductance oscillator [7].

Figure 1 represents a power comparative assessment of the two oscillator types using a one or two port representation.

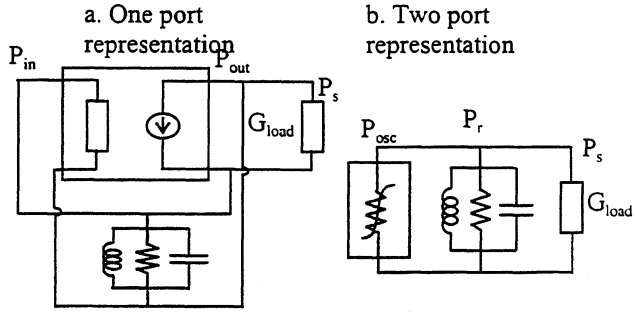


Figure 1 : power comparative assessment between the two oscillator types

where P_{add} is the amplifier power added, P_r is the resonator dissipated power, P_s is the output power, P_{osc} is the diode delivered power, P_{in} and P_{out} are amplifier input and output powers.

In other words, P_{add} is the available power at the oscillator output, considered now as an active dipole with the resonator and external load connected in a reflection type configuration such as figure 1b. In this case equation (1) becomes :

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

$$L(f_m) = \frac{F.G.K.T.B}{8P_{add}}\left(1+\frac{f_c}{f_m}\right)\frac{f_o^2}{f_m^2}\frac{1}{\left(Q_R\dfrac{G_C}{G_C+G_{load}}\right)^2} \quad (2)$$

where Q_R is the intrinsic quality factor of the resonator, G_C is the conductance brought back through the resonator, G_{load} is the conductance brought back through the load circuit.

So to minimize the phase noise, the quantity to be maximized is $P_{add}\left(\dfrac{G_C}{G_C+G_{load}}\right)^2$, for a given Q_R.

It cannot be directly applied to microwave oscillator circuits where the feedback circuit is distributed.

It is the purpose of this letter to describe a new design method of free running oscillator circuits leading to the minimum phase-noise, for given transistor and resonator. For sake of clarity, we will use ideal elements for the oscillator equivalent circuit at the fundamental frequency. Non linear simulations, using accurate models for the transistor and passive-elements fully corroborate the deductions found with this simplified model and will be published elsewhere.

Fundamentally, a transistor is a voltage controlled current-source. The internal controlling voltage is taken at a diode terminals, forward (BJT, HBT) or reverse (MESFET, PHEMT) biased. Due to the non linear capacitance of this diode, the phase noise conversion takes place principally at this controlling voltage port [8]. So, to minimize the phase noise, the energy stored by the oscillator circuit must be maximized [9] and wholly transferred to this controlling voltage port. If load resistance is not considered, the oscillator can be represented by the simplified schematic shown in figure 2. G_{in}, C_{in}, G_{m0} represent the describing functions of the non linear input admittance and current source of the transistor, G_{out} and C_{out} those of the non linear output admittance. L_{in}, L_{out}, n_{in} and n_{out} are matching elements to $G_0 = 1/50\Omega^{-1}$ at the operating frequency ω_0. The lines θ_1 and θ_2 are equivalent to ideal phase shifters.

Finally, the resonator, coupled by ideal transformers n_1 and n_2, is defined by its unloaded Q_R at its resonant frequency ω_0 (figure 2) :

$$Q_R = \omega_0\frac{\text{Reactive energy stored}}{\text{Energy dissipated per cycle}} = \omega_0\frac{\frac{1}{4}\frac{dB_R}{d\omega}V_R^2}{\frac{1}{2}G_R V_R^2} \quad (3)$$

where $\dfrac{dB_R}{d\omega}$ is the resonator susceptance slope, V_R is the voltage amplitude at the resonator port, and G_R is the conductance of the resonator.

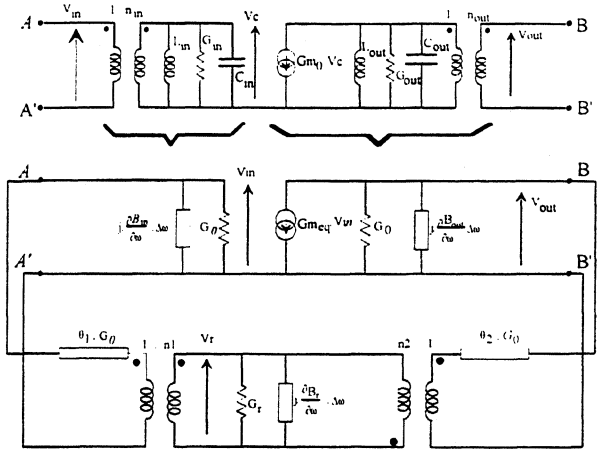


Figure 2 :oscillator equivalent circuit around oscillation frequency

One key point to minimize the phase noise is the determination of the large signal matching conditions to define on the $P_{out} = f(P_{in})$ curves of the amplifier, for a given bias point. Then, once these matching conditions are established, the transistor and its matching elements may be replaced, around the frequency ω_0, by the synthetic model of figure 2.

Let us first consider equation (3) : since Q_R is given, it follows that the stored energy in the resonator will be maximized if the dissipated power is maximized. This power is the difference between the output and input power of the transistor : it is its power added. Then, in oscillating operation, the transistor power added must be maximized and dissipated in the resonator. It follows that the transistor must be large signal matched at its maximum power added operating point. When this optimum is reached :

$$P_{add_{max}} = \frac{1}{2}.G_r.\left|V_{r_{opt}}\right|^2 = P_{in_{opt}}.(G_{opt}-1)$$

$$P_{add_{max}} = \frac{1}{2}.G_0.\left|V_{in_{opt}}\right|^2(G_{opt}-1) \quad (4)$$

where $G_{opt} = \dfrac{G_{meq}^2}{4.G_0^2}$ is the associated power gain of the transistor at this operating point.

Practically, in order to find $P_{add\,max}$, G_{opt} and the corresponding matching elements to 50Ω, a source and load-pull measurement, or a non linear simulation of the transistor in amplifier configuration, may be performed. From equation (4), the resonator coupling coefficients must satisfy the relation :

$$n_{1_{opt}}^2 = \frac{\left|V_{r_{opt}}\right|^2}{\left|V_{in_{opt}}\right|^2} = \frac{G_0}{G_r}\left(G_{opt}-1\right) \; ; \; n_2^2 = \frac{n_{1_{opt}}^2}{G_{opt}} \quad (5)$$

To maximize the transfer of this stored energy toward the internal controlling voltage port of the transistor, the electrical length θ_1 must be tuned while remaining the phase oscillation condition $\theta_1 + \theta_2 = 2\pi$ and the transistor maximum added power operating condition. A straightforward calculation shows that θ_1 must be equal to $k.\pi$. The stored energy in the resonator ε_r is then fully transferred to the controlling voltage port : $\varepsilon_{in} = \varepsilon_r$.

Equations (5) and the phase condition $\theta_1 = k\pi$ are the two key points of the proposed design method. To be useful, the oscillator circuit must be coupled to an external load. The power dissipated in the resonator is reduced by an amount equal to that delivered to the load. So, the phase noise will be deteriorated in the same proportions.

In this quasilinear theory, the signal harmonics have been neglected. In an experimental set-up, we must to make sure that the harmonics are always loaded in the same manner whatever the phase shifter values.

2. PRACTICAL EXAMPLES

Practical examples of optimized oscillator circuits shown in figure 2, operating at 9.2 GHz, have been built.

An isolator at the transistor output eliminates spurious oscillations. It is followed by a 20 dB coupler to measure the spectrum. It is important to note that two phase shifters are used to both obtain oscillator phase conditions and maximum stored energy transfer between resonator and transistor input.

To validate our design method, two oscillators have been built with two different dielectric resonators and the same PHEMT transistor amplifier.

2.1 Influence of phase shifters on phase noise with a low Q DRO

Figure 3 shows the phase shift influence, between the resonator and the transistor input, on phase noise. This measurement has been done with a low Q dielectric resonator and nevertheless the result shows clearly the importance of the phase shift tuning.

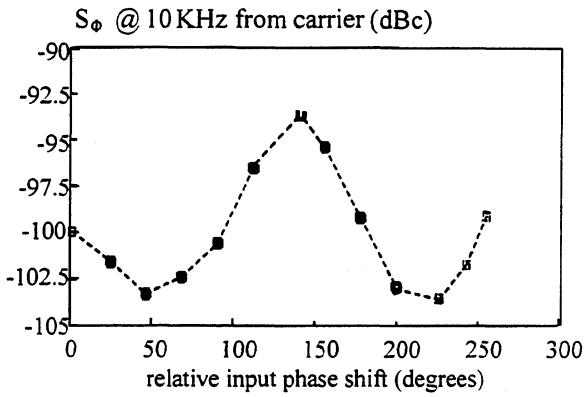


Figure 3 : Phase shift influence on phase noise with low Q DRO.

2.2 Experimental results with the high Q sapphire resonator oscillator.

In the second step, the high Q resonator (Q_0=105000) is a sapphire resonator working in a WGM7 mode stabilized at 330 K [10]. The built up oscillator is presented in figure 4.

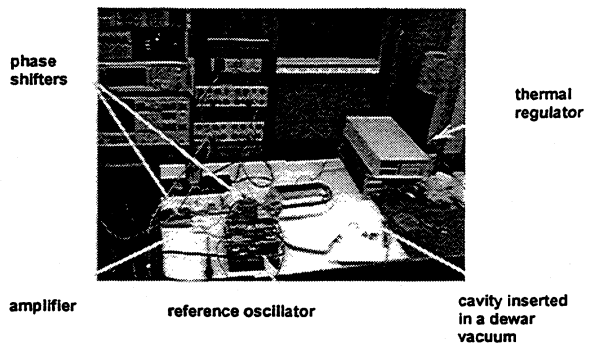


Figure 4 : high Q sapphire resonator oscillator

The transistor is a packaged 0.25 µm × 200 µm PHEMT chosen for its low flicker noise. The measured characteristics of the amplifier matched to 50Ω at its maximum power added operating point are : V_{g0}=0 V, V_{ds0}= 3 V, P_{in}=4 dBm, P_{out}=12 dBm, P_{add}=11.3 dBm. The losses of the feedback circuit (excluding the resonator) are 3 dB. Following equation (5), the resonator coupling coefficients are tuned to β_1 = 0.5 and β_2 = 1.3 giving 9.2 dBm cavity dissipated power. Two phase shifters allow to adjust the phase difference between the resonator and the internal controlling voltage port while maintaining the previous transistor power conditions and 9.2 GHz oscillation frequency.

The noise measurements are carried out using the HP3848 phase noise set-up and a reference synthesized oscillator. A phase noise level of −80 dBc/Hz at 100 Hz offset from carrier with a −30 dBc/dec slope has been measured (figure 5). The measurement frequency range is limited by thermal fluctuations below 10 Hz and by the reference oscillator

noise above 200 Hz. To our knowledge, this result is the best published result for a HEMT free running microwave oscillator at room temperature.

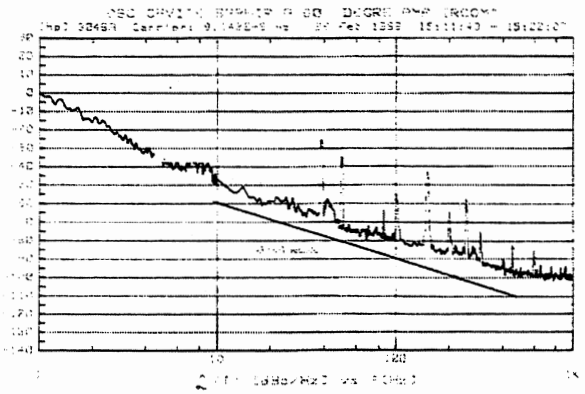


Figure 5 : Measured phase noise spectrum

3. CONCLUSION

A new design-method of very low noise oscillator circuits has been presented leading to the design and build up of HEMT low noise band X oscillator producing a –80 dBc/Hz @ 100 Hz carrier offset phase noise with a –30 dB/dec slope at room temperature. The key point of the method is the maximization of the energy stored in the resonator and its transfer to the controlling voltage port of the transistor. The design method efficiency, independent of the transistor technology, represents in our opinion a real breakthrough in the field of low noise transistor oscillator circuit design.

4. REFERENCES

[1] *J.K.A. Everard*, «A review of low noise oscillator : Theory and design». Proceedings of the IEEE Frequency Control Symposium, 1997, pp 909-918.

[2] *J.K.A. Everard and M.A. Page-Jones*, «Ultra noise microwave oscillators with low residual flicker noise». IEEE International microwave symposium, Orlando, 1995, pp 693-696.

[3] *M.M. Driscoll and R.W. Weinert*, «Spectral performance of sapphire dielectric resonator-controlled oscillators operating in the 80 K to 275 K temperature range». IEEE FCS, 1995, pp 401-412.

[4] *G. J. Dick, R.T. Wang and R.L. Tjoelker* , « Cryo-cooled sapphire oscillator with ultra-high stability », IEEE Frequency Control Symposium, 1998 pp 528-533

[5] *C.A. Flory and H.L. Ko*, « Microwave oscillators incorporating high performance distributed BRAGG reflector microwave resonator».Proceedings of the IEEE Frequency Control Symposium, 1997, pp 994-999.

[6] M.E. *Tobar, E.N. Ivanov, R.A. Woode, J.H. Searls, A.G. Mann*, «Low noise 9 GHz sapphire resonator-oscillator with thermoelectric temperature stabilization at 300 Kelvin». IEEE Microwave and Guided Wave Letters, April 1995, Vol. 5, n° 4, pp 108-112,.

[7] *H. Haus, H. Statz, R. Pucel*, « Optimum noise measure of impatt diodes ». IEEE MTT, vol. 19 n°10, October 1971, pp 808-813

[8] *H.J. Siweris, B. Schiek*, «Analysis of noise upconversion microwave FET oscillators». IEEE Trans. On Microwave Theory and Techniques, 1985, Vol. MTT-33, n° 3, pp 233-238.

[9] *E.J. Shelton*, «Stabilization of microwave oscillators». IRE Trans. , 1954 ED-1, pp 30-40.

[10] *O. Di Monaco, W. Daniau, I. Lajoie, Y. Gruson, M. Chaubet, V. Giordano*, «Mode selection for whispering gallery mode resonator». Electronics Letters, March 28th 1996, Vol. 32, n° 7, pp 669-670.

1999 Joint Meeting EFTF - IEEE IFCS

PERFORMANCE EVALUATION OF OPTOELECTRONIC OSCILLATORS.

S. Römisch, J. Kitching, E. Ferrè-Pikal[#], L. Hollberg and F. L. Walls

National Institute of Standard and Technology, 325 Broadway, Boulder, CO 80303
[#]University of Wyoming, Laramie, WY

ABSTRACT

In the present work, we establish the importance of amplifier phase noise and system environmental fluctuations in determining the low-frequency ($f < 100$ Hz) noise of the oscillator. The system, designed to be a low phase noise microwave oscillator, has a measured single-side-band (SSB) phase noise of -123 dB/Hz (relative to 1 rad^2/Hz) at 10 kHz from the 10.6 GHz carrier. The fractional frequency stability required for advanced atomic frequency standards is on the order of $10^{-14}/\sqrt{\tau}$. This requirement might be met with an ideal version of this optoelectronic oscillator if we could reach the noise limit determined by the optical shot noise.

1. INTRODUCTION

The performance of a hybrid optical/electronic oscillator which uses a fiber-optic delay line as the frequency discriminator has been investigated and an evaluation of the possible use of this oscillator as a frequency reference for atomic standards has been done.

The traditional methods used to obtain spectrally pure microwave signals are based either on crystal oscillators (bulk-acoustic wave, BAW and surface acoustic wave, SAW) or on various schemes that utilize some sort of high-Q resonator (like whispering-gallery mode sapphire resonators). In the first case the output frequency needs to be multiplied, introducing signal degradation. In the second case the resonator requires more sophisticated technology than can usually be justified. A feature common to all such oscillators is the small tunability. This can be increased only at the expense of stability.

A promising alternative is the optoelectronic oscillator (OEO) that has been studied in several laboratories [1], [2], [3]. These hybrid opto-electronic systems use a long optical fiber as the selective element that permits high tunability and almost no limitation on the range of possible oscillation frequencies.

Furthermore, it is possible to define for the fiber a quality factor that defines its capability to select an oscillation frequency. This is similar to the Q factor used for resonators. In the fiber case, this factor is proportional to the product of the time delay introduced by the fiber and the oscillator's frequency. Recognition of this feature invites consideration of higher oscillation frequencies to improve the spectral purity the oscillation. The long delay also generates a high mode density and hence almost any oscillation frequency can be chosen.

Two undesirable features are the thermal dependence of fiber length, which limits the medium and long-term stability, and the non-negligible loss incurred in the conversion from microwave to optical and back again to microwave. To produce oscillation, this loss needs to be compensated, requiring microwave amplifiers that limit the medium-term stability due to their flicker phase noise.

A possible solution to the problem of amplifier phase noise is the use of two stages of amplification. The first amplifier would have low-power, low-flicker noise and the second would provide the remaining gain with its flicker noise canceled using the carrier suppression technique [4].

2. THE OPTOELECTRONIC OSCILLATOR

2.1 Basic scheme and equations

The general configuration for an OEO, shown in Fig. Figure 1, includes a CW laser, an electro-optic amplitude modulator (EOM), and a photodetector at the end of the optical fiber. The loop is then closed by an amplification stage, in order to reach the oscillation threshold, and some kind of filter which selects among the possible modes of this oscillator.

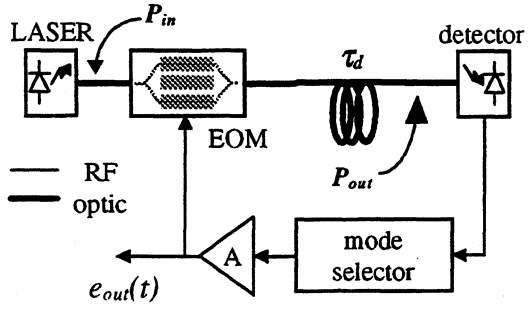


Figure 1. Basic scheme for an OEO.

In our system the EOM is a Mach-Zehnder type modulator which has a cosine-shaped transmittance versus drive voltage,

$$P_{out}(t) = P_{in}(t-\tau_d)\gamma\left[1+\varepsilon\cos\left(\pi\frac{V_{bias}+e_{out}(t-\tau_d)}{V_\pi}\right)\right], \quad (1)$$

where P_{in} and P_{out} are the optical powers incident on the modulator and detected at the end of the fiber. The modulator's parameters are γ, a factor related to the insertion loss; ε, a factor related to the extinction ratio and V_π, the bias voltage needed to move from a maximum to a minimum of the optical power transmitance.

The EOM is biased at the half-transmittance point so that the fundamental of the RF signal is transmitted to the detector. This signal is amplified and then fed back to the modulator. If the RF signal is written as

$$e_{out}(t) = V_0 \sin(\omega_0 t), \quad (2)$$

the oscillation condition will be set by

$$P_{in}\gamma\left[1-2\varepsilon J_1\left(\frac{V_0\pi}{V_\pi}\right)\sin(\omega_0 t-\omega_0\tau_d)\right]=\frac{V_0}{A\rho}\sin(\omega_0 t), \quad (3)$$

where A is the gain of the RF amplifiers, ρ includes the detector responsivity and the fiber coupling loss, and $\omega_0=2\pi\nu_0$ is the oscillation angular frequency. The total delay along the loop is τ, the sum of the fiber delay τ_d and the much smaller delays of the others elements of the system (typically neglected here). The Bessel function comes from the series expansion of the transmittance the modulator.

The solutions of the oscillating term of Eq. (3) can be written as

$$\begin{cases} 2\gamma P_{in}\varepsilon J_1\left(\frac{V_0\pi}{V_\pi}\right)=\frac{V_0}{A\rho}, \\ \omega_0=\frac{(2K+1)\pi}{\tau_d}, \quad K=1,2,3\ldots \end{cases} \quad (4)$$

The *mode selector* in Fig. Figure 1 selects a particular value of K (K=33 455 in our case) among all the solutions of Eq. (4).

In Eq. (3) and Eq. (4) the mode selector and the amplifiers are assumed to have a large enough bandwidth that the dynamics of the system are unaffected by them. A more detailed analysis of the oscillator can be found in [5].

2.2 Experimental arrangement

The experimental setup is shown in Fig. 2. The laser is an InGaAsP DFB device with a 3 mA threshold current and a maximum output optical power of 70 mW at a current of 200 mA. In our present experiments, a 100 mA current is used with an optical output power of about 35 mW. After passing through an isolator, the beam is coupled into the input fiber of the EOM. The inferred coupling efficiency is about 28 %.

The EOM has an insertion loss of about 3 dB and a V_π of about 6 V measured at the operating condition of the oscillator, that is, with a 10.6 GHz signal at the RF port. A servo system is needed to keep the bias point of the EOM stable. At this point (after the EOM) the remaining optical power is 7 % of the laser output power; the fiber coupling efficiency is 28 %, the passive insertion loss of the modulator is 50 % and another 50 % is lost due to the half-power bias point of the EOM.

The optical delay between the EOM and the detector is provided by a 1.2 km, single-mode, temperature-compensated optical fiber. This fiber has a temperature dependence of optical delay of 3 ps/(km·K) (compared to the typical fiber coefficient of about 30 ps/(km·K)).

In our present system the losses are significant and the photodetector sees only 5 % of the optical power delivered by the laser.

The detector is a commercial device with a bandwidth of 25 GHz and an inferred responsivity of 0.17 A/W. Although we are able to obtain (from the EOM) the maximum modulation depth (according to Eq. (3)) with the EOM input RF power near 22 dBm [22 dB (re 1 mW)], the detected RF signal is only -37 dBm as a result of losses primarily in the trasmission of the optical power. The total conversion loss from microwave to optical and back to microwave is about 62 dB.

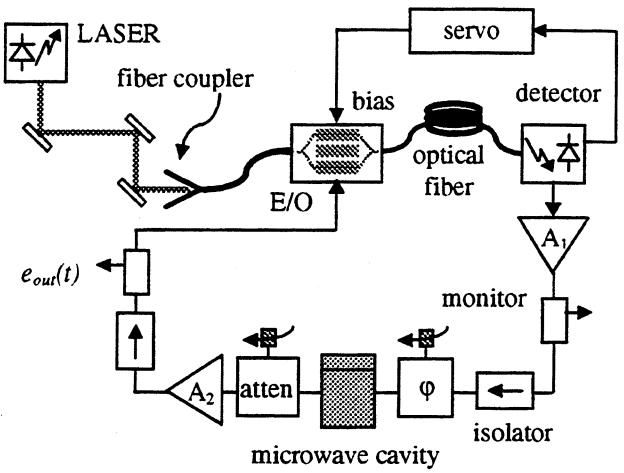


***Figure 2.** Experimental setup for the opto-electronic oscillator (OEO).*

This number sets the minimum gain which needs to be provided by the RF amplifiers. A low-noise, high-gain amplifier provides the first 45 dB of gain and a low-noise, high-power amplifier provides the rest of the required gain and delivers the 22 dBm signal (3 dB compressed) to the RF modulator port.

The mode selector is between the two amplifier stages and is a critically coupled microwave cavity which introduces 6 dB of loss. Finally, the variable attenuator permits us to control the amount of compression in the system.

The resulting microwave signal shows a side-mode suppression of about 73 dB, due to the filter cavity with a loaded quality factor Q_L of about 8300.

3. THE NOISE

The phase noise of this oscillator can be at least partly predicted through a simple model described by Leeson, which has been adapted to this particular case [5]. In order to apply this model, we need to know the phase noise associated with each element of the system.

We therefore need to know the noise contribution of the detection process and the noise introduced by the amplification stages.

In our case the photodetector is simply terminated with a 50 Ω resistor, called R in the equivalent circuit of Fig. 3, where I_d is the noiseless photocurrent, containing both a DC and a RF component.

With an optical power of 1.75 mW (the 5 % of 35 mW delivered by the laser) incident on a detector with responsivity 0.17 A/W, the shot-noise is $\overline{i_{shot}^2}=2eI_d\equiv -220$ dB/Hz. Then, the white phase noise delivered to the load, calculated from the circuit in Fig. 3 will be

$$S_\varphi(f)\Big|_{shot} = \frac{\overline{i_{shot}^2}R^2R_L}{2P_{rf}(R+R_L)^2} = -145 \ \frac{\mathrm{dB_{rad}}}{\mathrm{Hz}},$$
$$S_\varphi(f)\Big|_{therm} = \frac{\overline{i_{therm}^2}R^2R_L}{2P_{rf}(R+R_L)^2} = \frac{kT}{P_{rf}} = -140 \ \frac{\mathrm{dB_{rad}}}{\mathrm{Hz}}, \quad (5)$$

where P_{rf} is the signal power measured at the detector output (-37 dBm) and the dB_{rad}/Hz are intended relative to one squared radians. The term associated with the resistor's thermal noise is clearly dominant.

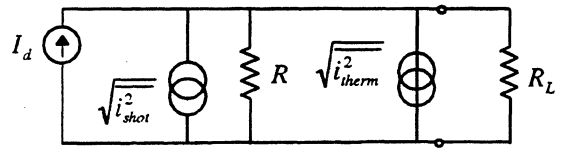


Figure 3. Photodetector equivalent circuit with shot noise and thermal noise sources. The load resistor is the input impedance of the first amplifier and is assumed to be 50 Ω.

The equivalent input phase noise of the amplifiers has been evaluated and includes all other RF devices except the cavity. This ensemble, called the *RF chain*, is displayed in Fig. 4, with the measurement results shown in Fig. 5.

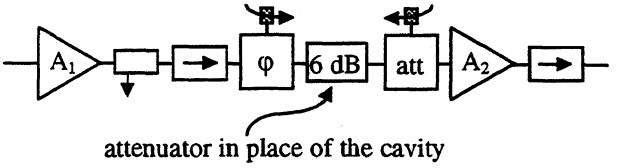


Figure 4. Elements considered in the RF chain.

The total phase noise at the input of the RF chain is the sum of the input equivalent noise of the RF chain and the white noise coming from the detection.

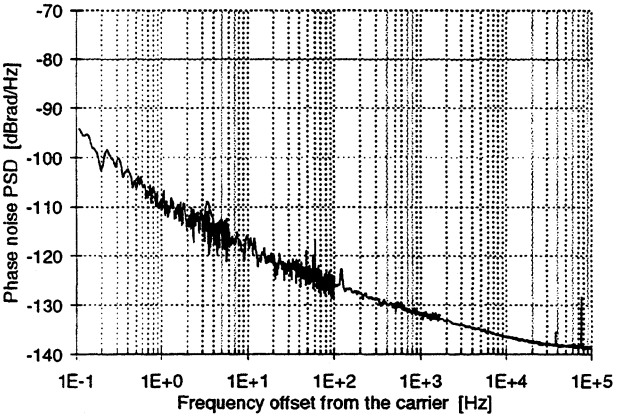


Figure 5. Phase noise power spectral density of the RF chain, *measured with an input signal power of -40 dBm at 10.571 GHz.*

These quantities can be used to predict the oscillator noise with the noise transfer function calculated according to the Leeson's model [5]:

$$S_\varphi(f)\Big|_{osc} = \frac{1}{(1-\cos(F(\Omega)))^2+\sin^2(F(\Omega))} S_\varphi(f)\Big|_{tot}, \quad (6)$$

with

$$F(\Omega) = \Omega\left(\tau_d + \frac{2Q_L}{\omega_0}\right), \quad (7)$$

where τ_d is the fiber delay, Q_L is the loaded quality factor of the microwave cavity, and $\Omega=2\pi f$ is the frequency offset from the carrier.

Fig. 6 compares the phase noise predicted by the model is compared to that measured with a frequency-discriminator noise-measurement system.

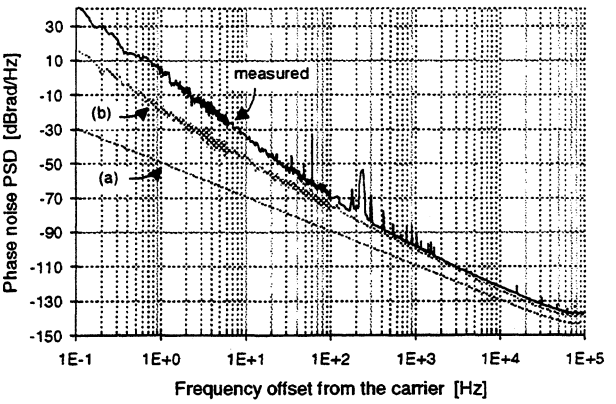


Figure 6. Measured and predicted phase noise PSD of the OPO. The noise predictions are made according to Leeson's model on the basis of the only white input noise (a) and the total equivalent input noise (b).

The measured data from Fig. 6 are replotted in Fig. 7 and analyzed in terms of power-law spectral densities. The small portion of white phase noise (slope f^0) around 100 kHz from the carrier is due to the presence of the first side-mode of the oscillator. For our 1.2 km fiber the free-spectral range is about 150 kHz and the side-mode suppression (with the present cavity) is only about 73 dB.

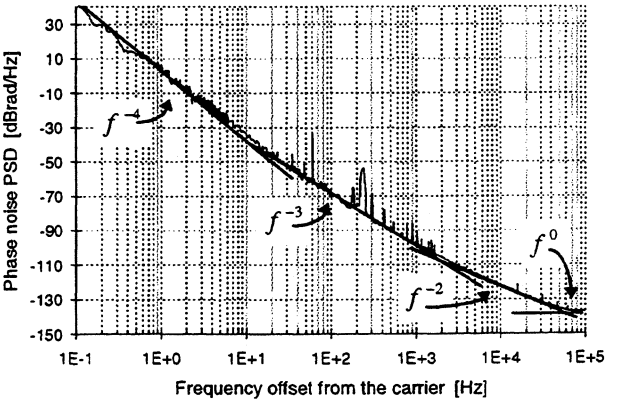


Figure 7. Measured PSD of the OPO with superposed power law model.

The next two segments with slope of 20 dB/decade and 30 dB/decade are due (as is clear from Fig. 6) to the white and flicker phase noise of the amplifiers.

The low-frequency part of the spectrum is random-walk frequency noise (slope f^{-4}) and is known to be mainly related to environmental factors such as temperature fluctuations and vibration.

4. PERFORMANCE EVALUATION AND FUTURE IMPROVEMENTS

One useful figure of merit for oscillators (even at different frequencies) is the Q-frequency product, which gives a general evaluation of the potential performance of the device. It is also possible to define a quality factor for a fiber that permits us to calculate our Q-frequency product [5] as

$$Q_{fiber} = \pi \nu_0 \tau_d , \qquad (8)$$

where ν_0 is the oscillator frequency and τ_d is the optical delay of the fiber. For this OEO we obtain about $2.2\ 10^{15}$ Hz while for the best quartz oscillators we have $1.6\ 10^{13}$ Hz in the case of a BAW and $1.05\ 10^{13}$ Hz for a SAW.

The important difference with respect to resonator-based oscillators is that in this case, the quality factor is proportional to the oscillation frequency. Since for this oscillator the Q-frequency product is proportional to ν_0^2, it is advantageous to have as high a working frequency as possible.

The actual measured frequency stability is shown in Fig. 8.

The flat part of the curve is due to the phase flicker noise in the loop (slope f^{-3} in Fig. 7) while the part with slope τ^{+1} is a direct consequence of the thermal drift of the fiber length.

If we calculate the Allan variance related to the asintote with slope f^{-2} of Fig. 7, we find $\sigma_y = 4.7\ 10^{-13}\tau^{-1/2}$ and, although this is an estimate that assumes the ability to remove all the excess low frequency noise, the oscillator is still not stable enough for use as a local oscillator for the new generation of atomic standards. For this OEO (ν_0=10.571 GHz), the white-frequency noise level (slope f^{-2} in Fig. 7) corresponding to a frequency stability of $\sigma_y = 10^{-14}\tau^{-1/2}$ is about -156 dBm/Hz at 10 kHz from the carrier. Since this portion of the noise spectrum comes from the open-loop white-phase noise, the simplest way to improve it is to obtain a higher RF signal at the detector output. That requires a higher detector responsivity and/or a higher laser power and a transimpedence amplifier photodetector.

However the low-frequency excess noise remains. In particular, the stability between 100 Hz and 10 kHz offset from the carrier is limited by amplifier flicker phase noise; a possible solution is to use a quieter first stage and carrier suppression techniques applied to the second stage. Finally, the medium and long-term stability of this oscillator is limited by environmental factors, and improvement will require temperature stabilization of the fiber.

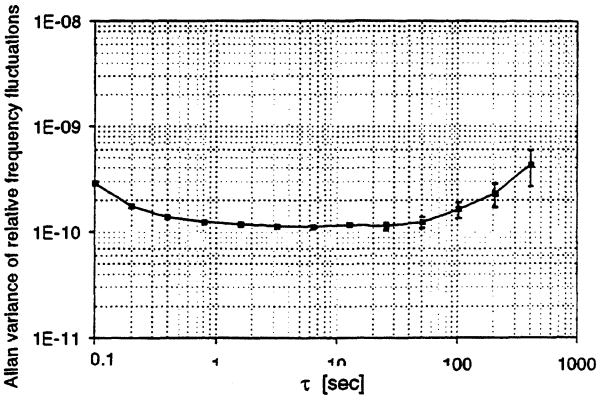


Figure 8. Allan variance of the relative frequency fluctuations versus integration time τ.

We thank S. R. Jefferts and F. Ascarrunz for many helpful and illuminating discussions. We are also grateful to R. Mirin and M. Young for a carefully reading of the manuscript.

5. REFERENCES

[1] X. S. Yao and L. Maleki, "High frequency optical subcarrier generator," Electron. Lett., Vol. 30, pp. 1525-1526, 1994.

[2] X. S. Yao and L. Maleki, "Optoelectronic oscillator for photonic system," IEEE J. Quant. Electron., Vol. 32, no. 7, pp. 1141-1149, 1996.

[3] J. Kitching, L. Hollberg, and F. L. Walls, "A 1 GHz optical-delay-line oscillator driven by a diode laser," in Proceedings of the IEEE International Frequency Control Symposium, 1996, pp. 807-814.

[4] E. N. Ivanov, M. E. Tobar and R. A. Woode, "Advanced phase noise suppression technique for next generation of ultra low noise microwave oscillator," in Proceedings of IEEE International Frequency Control Symposium, 1995, pp. 314-320.

[5] S. Römisch and A. De Marchi, "Noise predictions for the optoelectronic oscillator using different models," in these proceedings.

[6] X. S. Yao and L. Maleki, "New results with the opto-electronic oscillators (OEO)," in Proceedings of IEEE International Frequency Control Symposium, 1996, pp. 1219-1222.

OPTO-ELECTRONIC OSCILLATOR INCORPORATING CARRIER SUPPRESSION NOISE REDUCTION TECHNIQUE

X. Steve Yao, Lute Maleki, and John Dick, Frequency and Time Sciences and Technology Group
Jet Propulsion Laboratory California Institute of Technology, Pasadena, CA 91109

ABSTRACT

The noise performance of the Opto-Electronic Oscillator (OEO) has been improved by about 20 dB. This improvement is obtained with a novel implementation of the carrier suppression technique.

Recent results with the Opto-Electronic Oscillator have led to the realization of very high spectral purity. Experimental results have produced a performance characterized by noise as low as -60 dBc at 10 Hz for a 10 GHz oscillator. This performance is significant because the oscillator is free-running; and since the noise in an OEO is independent of the oscillation frequency, the same performance may also be obtained at higher frequency. We have recently developed a novel carrier suppression technique to reduce the 1/f phase noise of the oscillator even further. The technique is based on the use of a long fiber

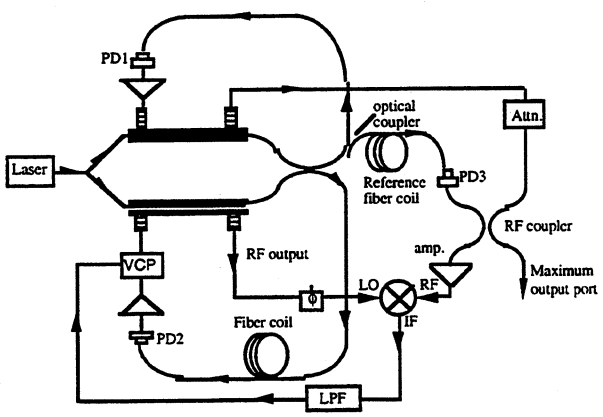


Figure 1.

delay, in place of the high Q cavity to implement the carrier suppression scheme. Our preliminary experimental results indicate an extra 10 to 20 dB phase noise reduction of the OEO with this novel technique. Further noise reduction beyond this value is expected with improved circuit design and longer reference fiber.The carrier suppression technique pioneered at JPL [1,2] has been successfully used in microwave oscillators to reduce close-to-carrier phase noise resulting mainly from the 1/f noise of an amplifier in the oscillator loop.[3] Here we report the first application of the carrier suppression technique in an opto-electronic oscillator (OEO) to reduce the phase noise which result not only from the amplifier, but also from the laser relative intensity noise (RIN).

A double loop OEO incorporating the carrier suppression technique is shown in Fig. 1. The long loop consists of a polarization beam splitter (PBS), a fiber coil having a long length of fiber, a photodetector (PD2), an RF amplifier, a bandpass filter, and a voltage controlled phase shifter (VCP). The short loop consists of a short length of fiber, a photodetector (PD1), an RF amplifier, and a bandpass filter. In one implementation of the fiber optic carrier suppression scheme, part of the light is coupled out from the short loop and is delayed by a reference fiber coil of 2 km length. The delayed light signal is then received at a third photodetector (PD3) and converted to RF signal. This signal is then made to interfere with the RF output of the OEO at an RF bridge consisting of a 3 dB coupler, a variable attenuator, and a variable phase shifter. The variable attenuator and the phase shifter are adjusted such that one of the output ports (port 1) of the bridge has a minimum output power, while the other port (port 2) has the maximum. The signals from the minimum output port is then amplified and mixed with the signal from the RF output port of the oscillator. The relative phase between these two signals is adjusted by another variable phase shifter to be at 0 or π. The error signal from the mixer is then amplified, filtered, and fedback to the VCP to control the frequency of the OEO.

In this configuration, the fiber coil acts both as a high Q phase storage component for the oscillator, and as a frequency discriminator for the carrier suppression bridge. The frequency discriminator converts the frequency jitter of the OEO into amplitude jitter out of the bridge. This amplitude jitter is then detected by the mixer, and then amplified, filtered, and fedback to the VCP to suppress the frequency noise of the OEO. Because the carrier suppression scheme is insensitive to the amplitude noise of the signal source, the relative intensity noise (RIN) will not affect the frequency noise measurement with the

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

RF bridge, and thus will not mask the phase noise contribution from the RIN.

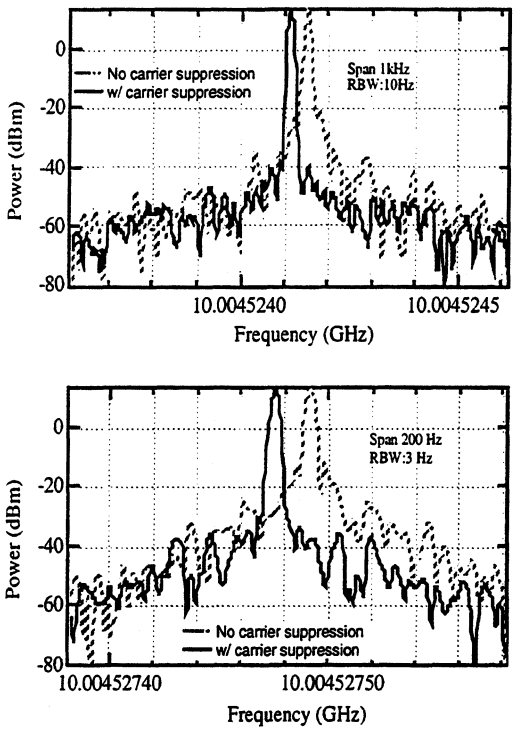


Figure 2.

The RF spectra of the 10 GHz OEO with and without carrier suppression are shown in Fig. 2a and Fig. 2b. The spectra were taken with an HP8563E spectrum analyzer. The span and resolution bandwidth of the spectrum analyzer were set at 1 kHz and 10 Hz respectively for Fig. 2a and 200 Hz and 3 Hz for Fig. 2b.

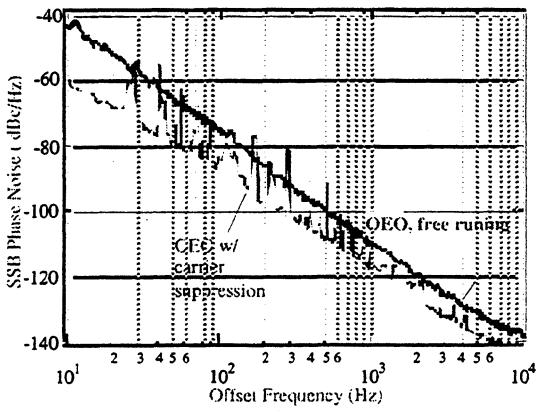


Figure 3.

It is evident that the spectral purity of the OEO was improved significantly (~ 20 dB) with the carrier suppression circuit active.
Fig. 3 shows the phase noises of an OEO with and without carrier suppression noise reduction. A phase noise reduction close to 20 dB at 10 Hz is evident.

In summary, we have successfully implemented a carrier suppression noise reduction technique with an optical fiber delay line, and applied the technique to reduce the close-to-carrier 1/f noise in a double loop OEO. About 20 dB phase noise reduction was achieved with the technique at 10 Hz away from the 10 GHz carrier. Further improvement of the performance of the OEO is expected with improved circuits that are under construction.

The research described in this paper was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under contracts with the National Aeronautics and Space Administration, and US Air Force Rome Laboratories.

References

[1] J. Dick and D. Santiago, "Microwave frequency discriminator with a cryogenic sapphire resonator for ultra-low phase noise," Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17-19 March 1992, ESA SP-340, June 1992, pp. 35-39.

[2] E. N. Ivannov, M. E. Tobar, and R. A. Woode, "Advanced phase noise suppression technique for next generation of ultra low0noise microwave oscillator," Proceedings 1995 IEEE International Frequency Control Symposium, 1995, pp. 314-320.

[3] X. S. Yao and L. Maleki, "Dual-loop Opto-Electronic Oscillator," Proceedings of the 1998 IEEE International Frequency Control Symposium, 1988, pp. 545-549.

SUGGESTIONS FOR FINDING OPTIMUM MATERIALS AND MANUFACTURING METHODS FOR METALLIC MICROWAVE CAVITY RESONATORS

Harri Eskelinen
Lappeenranta University of Technology
PL 20, 53851 Lappeenranta, Finland
phone:358-5-6212442, fax:358-5-6212499, email:harri.eskelinen@lut.fi

1. ABSTRACT

The selection and combination of appropriate materials and manufacturing methods for metrology-grade microwave resonators can be improved by tuned design methodology and manufacturability analysis which will enhance performance and long-term reliability and are often able to reduce cost as well. Typical interest areas include steel resonators both for Cs- or H-maser atomic standards and those operating as stand-alone devices in the millimeter region. New performance levels are feasible due to improved surface quality, stability, and the reduction of seam widths by e.g. laser processing. A stepwise procedure points out solutions particularly for e.g. challenging airborne, sea or space platforms in the higher microwave bands where prototype specifications exceeding those of conventional copper cavities have been achieved. The loaded quality factor of a 15 GHz rectangular steel cavity approaching 400 in a 50 ohm system with a simultaneous tuning error below 0.1 MHz seems feasible.

2. INTRODUCTION

The initial electronic design of a military grade 15 GHz microwave cavity resonator assembly, used in our tests, is documented in [1], [2] and [3]. For a specific frequency control application a very stable, well conducting and dimensionally precise structure was needed, capable of surviving also in a harsh environment like [4] or [5]. Further design wishes were e.g. minimal attenuation in critical areas, magnetic shielding and the ability to withstand excess thermal power. Sub-audio vibration and shocks were known to be encountered in this particular application.

The general requirements for a microwave resonator are based on electromagnetic wave propagation, transmission line theory and the electromagnetic properties of applied construction materials. The aim here is to point out that ***a tuned design methodology*** for microwave mechanics, specially with a new ***tuned manufacturability analysis***, will have a positive effect on microwave performance as well and even enable previously unobtainable features. Focus areas include impedance matching of the resonator, 3-D field pattern and polarization characteristics, maximum Q and its stability and the electric shielding performance of the selected construction - a very important issue in a military application. Imperfections in the cross sectional structure of the resonator itself, missalignments, material dents, oxidation and non-perfect joints between sections can be reduced without adding the design or production time or cost.

3. TUNED MANUFACTURABILITY ANALYSIS

The main reason for developing new design methodologies is the aim to improve the efficiency of engineering design. In practice, there are six different point of views to handle this subject : **1)** making organisational changes (team work, relationships between designers etc.), **2)** knowledge level (necessary education, knowledge about modern manufacturing technologies etc.), **3)** improving the design environment and ergonomy, **4)** personal characters and skills (abilities for creative design and team work), **5)** possibilities to use additional design methodies and tools (for example computer aided tools) and **6)** possibilities to tune the design methodologies or combine the best stages from already known methodologies for a specific design area .

In this paper the main attention is given to the search for new possibilities to tune the design methodologies or combine the best stages from alredy know methodologies for a specific design area by analysing the design procedure of the laser welded resonator. To get an impression of the problems of applying design methodologies to microwave mechanics let's examine two typical examples: For example in [6] Bubka and Flynt present "design methodology" for a typical microwave

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

component (stripline directional coupler) and their result is that the misalingment between adjacent layers is a critical parameter. They present also values for the tolerances to ensure the required performance. However, no connection is given to the manufacturing technology or production. Besides the typical "methodology" for microwave mechanics consists of three stages which are e.g. in the case of designing another common microwave component (multilayer coupled line filter), as follows [7, p.785] : 1) evaluation of normal mode parameters for various coupled line sections, 2) determination of physical dimensions to obtain the required normal mode parameters as computed in step (1) and 3) simulation of physical structure obtained in step (2) to verify the design. It is important to notice that no further steps are presented. These two illustrative examples show clearly that there is a lack of succesful applications of an efficient design methodology for microwave mechanics.

The tunded design methodology for laser processing consists of six basic elemets as follows (see also Fig. 1) :

1. To meet the the special requirements of microwave mechanics the tuned requirement list is needed.
2. Team work according to the cross-tecnological approach should start from the beginning of the design process. Experts of microwave mechanics, engineering design, manufacturing technology and laser processing are the most important ones. In the beginning of the design process other members are not necessary.
3. To avoid useless redesign cycles a mathematical model is used to recognize the products to be laser processed as soon as possible after the first product ideas are found.
4. Tuned questionares are used to establish the special DFM – requirements of laser processing simultaneously with the functional requirements of microwave mechanics.
5. Computer aided enviroment is utilized from the beginning of the design process. In practise it is essential to avoid modelling the same CAD-geometry several times for different purposes and to ensure that data of the CAD-geometry can be transformed without any additional editing for example in DXF-format for laser processing.
6. The design process is carried out inside the optimisation approach starting from the top level (design) and ending at the bottom (production and manufacturing).

The design process is carried out inside the optimisation approach. The design task presented by the customer is an input to the process. Both the cross-technological team work and the use of computer aided means are carried out during the whole design process in the same design environment. The tuned requirement list and questionare for DFM, mathematical model to recognize the suitable manufacturing process during the early stages of the design chain and the creative means are the individual main stages which are necessary for succesful design work. During the process the scope of computer aided means and team work must be evaluated – and if needed there are good possibilities to utilize for example AI-, WWW-applications and Concurrent Engineering (CE). The output of the process is the readiness for a laser processed product.

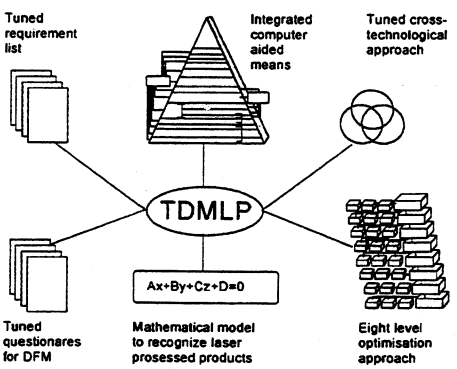


Fig. 1. Basic elements of the tuned design methodology for a laser processed product.

A lot of attention should be given to check the common requirements of design methodology which obey the following eight instructions [8,p.10]: **1)** the methodology must be applicable to every type of design activity, no matter in which specialist field, **2)** the methodology should facilitate the search for optimum solutions, **3)** the methodology should be compatible with the concepts, methods and findings of other disciplines, **4)** the methodology should not rely on finding solutions by chance, **5)** the methodology should facilitate the application of known solutions to related tasks, **6)** the methodology should be compatible with electric data processing, **7)** the methodology should be easily taught and learned, and **8)** the methdology should reduce workload, save time, prevent human errors, and help to maintain active interest.

However, according to our experiences a methodology which is too universal will cause useless iterations or redesign in the beginning of the process and will make it difficult to recognize the most suitable manufacturig method soon enough to start an effective DFM. To avoid these disadvantages the tuned methodology is needed: With the help of the cross-technological approach it is possible to create new questionares which assist to form tuned requirement lists for any design activity and tuned DFM-questionares for any manufacturing technology. If the characteristics for different manufacturing methods (laser processing in this case) are clarified in detail, it is possible to form a mathematical formula to recognize a particular manufacturing method which could be used in the production. These three facts mean that to follow the first instruction extra work and research is needed. It is also important to notice that the members of the design team are selected related to the design task to ensure that the solution could be found as quickly as possible. If the design task is known, it is obvious to choose the most appropriate members of the design team and the tuned design methodology as well and not to just satisfy the universal goals.

4. TEST SET-UP

Both the capabilities of the tuned manufacturability chain, laser processing in general and the suitability of steel for microwave work can be judged by e.g. manufacturing cavity resonators for various center frequencies and testing their radio frequency characteristics, particularly their structural losses. Meanwhile, also the design methodology and its milestones can be verified. In our case, a 15.0000 GHz rectangular cavity design operating in the TE_{mno} -mode was used, because its feed arrangement, i.e. the coupling to a coaxial test system, was readily available from a waveguide horn antenna project [9]. Two test samples were assembled; one as a conventional soldered copper construction and another from steel with laser welded joints. The basic structure is shown schematically in Fig. 2 (3D-modelled with the Genius Mechanical Desktop-software) and a photograph of the built prototype in Fig. 3.

Various initial imperfections in the raw material and possible burrs or dents caused by the welding process will lower the resonator quality factor and easily generate unwanted modes depending on their location, and dimensional uncertainties particularly along the z-axis, change the center frequency.

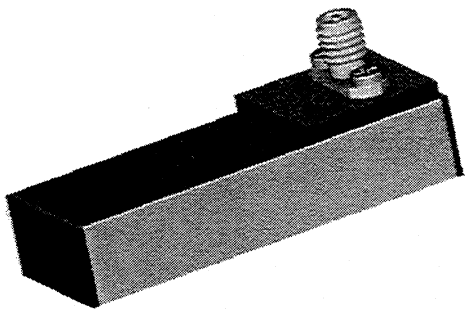


Fig.2. The rectangular cavity resonator has six conducting walls and a coaxial feed. The wave propagates and resonates along the z-axis.

Fig. 3. The laser welded steel prototype is completely sealed except for the small coaxial SMA-type connector. The outside appearance of the test unit is of no technical importance.

Ideally, the cavity should have four walls along its longitude axis (z-axis), each of them perfectly conducting, perfectly flat and adjacent panels should be at right angles against each other. This assures a correct propagation mode (usually of TE-type) in the cavity. E.g. welding joints proturding inside the cavity will spoil the corners and induce unwanted modes which usually decrease the quality factor Q. The effective surface conductivity might be changed as well. Another possible problem is the oxidation of the inner surface of the cavity due to welding - a drawback hardly correctable later due to the completely closed construction. Further on, the length of the resonator must be an integer multiple of the half-wavelength or

$\lambda/2$ with conductive plates at both ends if we want to produce a series resonant behaviour which shows an impedance minimum at the center frequency. A mechanical deviation will cause an immediate change in the resonance frequency.

5. SOME MEASUREMENT RESULTS

The unloaded quality factor Q of the test resonators was calculated according to [3] and was found to be about 24600 for a copper cavity and 8600 for a stainless steel version. The laser welded steel prototype was measured with a microwave VNA (Vector Network Analyzer). A typical result, presented as a Smith diagram, is shown in Fig. 4. Based on this, we can compute the realistic Q value either looking at the frequency difference between those points were the series resonant input impedance grows to 1.414.... of its minimum or from (1)

$$Q = \frac{\omega}{2 \cdot R} \cdot \frac{\partial X}{\partial \omega} \qquad (1)$$

shown originally in [3], where ω is the angular frequency and the resonator is assumed to have an input impedance $\mathbf{Z}_{in} = R + j\,X$. A narrow 100 MHz sweep with as many points as possible is needed due to the sharp resonance. The measured, loaded Q varies between 290 and 360, depending on the frequency span and evaluation procedure. It seems thus, that the material selection was succesful and steel, particularly if laser welded, can be used effectively for small-sized microwave components.

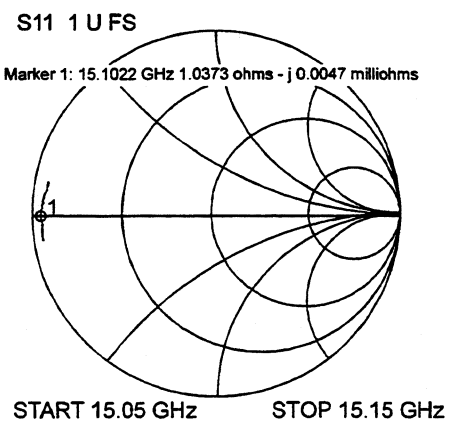


Fig. 4. The steel resonator's impedance plot at the vicinity of the center frequency. The loss resistance is about 1.2 ohms and the 3 dB loaded bandwidth 62 MHz.

The prototype resonator length was somewhat, about 0.5 mm shorter than designed. This is readily visible in Fig. 5 where the resonance frequency is plotted as a function of z-axis dimension d. The original mathematical formula is to be found in e.g. [3]. Unfortunately, the coupling arrangement to the resonator ***E***-field is dimensioned for 15.0000 GHz and thus optimum performance can not be achieved. The effects of the coupling arrangement are discussed further in Chapter 6 below.

Actually it turned out that a feasible and practical, non-destructive method for the accurate measurement, other than the applied microwave technique, of the resonator's internal length does not exist. With another 50 µm longer feed rod, the center frequency was found to be 15.0710 GHz. The resonance frequency could be measured (using the VNA as such) with an uncertainty smaller than 0.1 MHz which equals a dimensional error of 0.5 µm, see Fig. 6. Further tests with an elevated temperature showed that the resonance frequency of the steel prototype decreases 23 kHz / degree.

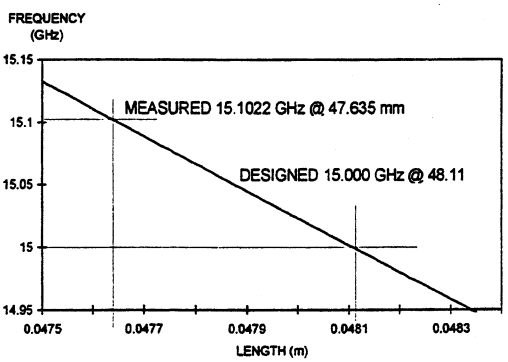


Fig. 5. The steel prototype resonator turned out to be 475 µm shorter than designed. This shifts the center frequency nearly 102.2 MHz upwards.

6. EFFECTS OF THE COUPLING PINS'S SHAPE

Most microwave resonators are very often connected with a coaxial coupling structure feeding the electric field of the hollow rectangular waveguide.

From the manufacturability point of view, the tuning of the transition is complicated due to the need to cut the rod to a suitable length l ***after*** it has been installed in the resonator.

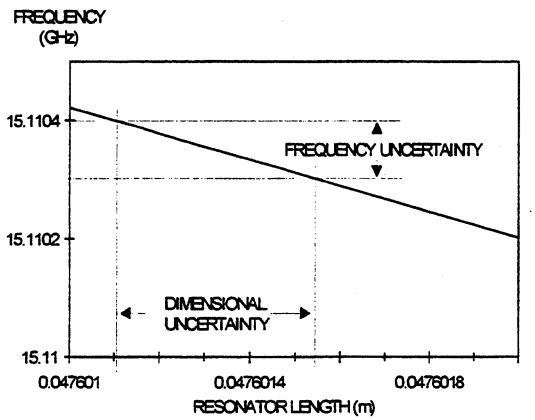


Fig. 6. The microwave network analyzer provides an accurate way for the dimensional measurement. At 15 GHz the average frequency uncertainty of 100 kHz (span 100 MHz) yields a dimensional error of 0.5 μm.

If the guide is open from one end, as is the case with e.g. horn antennas and waveguide adapters, a post-mounting cut would be highly attractive. However, as the transition is usually deep inside the structure, no mechanical means exist to perform this adjustment.

Cutting with a laser beam would provide an interesting alternative but the actual shape of the cut depends on the selected laser process as well and very little has been published about the possible electrical effects of a non-planar end surface.

Four different coaxial transitions were manufactured in order to simulate the effects of the coupling rod's actual shape. They were all tested with a rectangular waveguide.

Fig. 7 shows the microscopic photographs of the end shapes utilized with a straight-edged one as a reference and a tilted, grooved, and ball-shaped one simulating the possible manufacturing defects. The mean length was kept the same in order to avoid disturbing this way the tuning. The differences in performance exceeded the expectations.

Fig. 8 clarifies how the best performance is obtained with the "perfect" straight-ended rod having a Z_{in} = (36.7+j 8.9) Ω and the worst with a ball-shaped end where the Z_{in} = (33.4+ j 17) Ω . The tilted and the grooved feed rods behave about the same and fit in between the two extremes. The test bandwidth was 1 GHz.

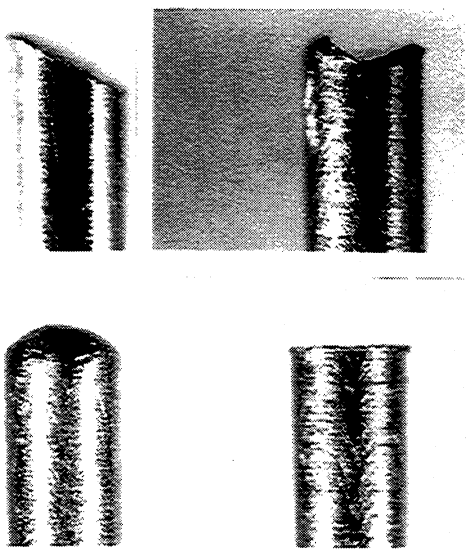


Fig. 7. Microscopic photographs of the four tested coaxial transition rods.

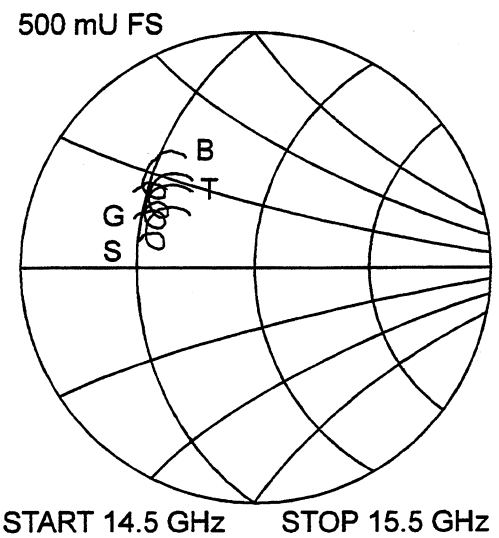


Fig. 8. The zoomed Smith chart presentation (0.5 U full scale deflection) of the effects of the end shape found in coaxial to waveguide - transitions. The best electrical performance is obtained with a straight-edged rod (S) whereas a ball-shaped version (B) provides the worst impedance match.

7. REFERENCES

[1] Ramo, Whinnery & van Duzer: Fields and Waves in Communication Electronics, John Wiley & Sons, New York 1984

[2] Collin R.: Foundations for Microwave Engineering, McGraw-Hill, New York 1992

[3] Matthaei, Young and Jones: Microwave filters, impedance matching networks and coupling structures, Artech House, Norwood 1985

[4] G. Busca et al, " The SHM Compact Hydrogen Maser for Space Applications - Report on the PEM Physics Package Design Verification", in the *Proceedings of the 12th EFTF'98,* pp. 121 - 125, March 1998

[5] Eskelinen, H."Observations on The Manufacturability of Mechanical Subassemblies for Remote Sensing Transducers", in *Proceedings of the URSI/IEEE XXIII Convention on Radio Science*, August 1998

[6] Bubka,T.P.;Flynt,R.A; Alignment tolerant stripline directional couplers, *Proceedings of the 1997 IEEE MTT-5 International Microwave Symposium,* Part 2 (of 3), Denver, USA, 8.-13.6,1997, ISSN:0149-645X, p.773-776

[7] Cho, C., Gupta, K.C.; "Design methodology for multilayer coupled line filters" ,*Proceedings of the 1997 IEEE MTT-S International Microwave Symposium,* Part 2 (of 3), Denver,USA,8.-13.6.1997,ISSN: 0-149-645-X, p. 785-788

[8] Pahl,G.,Beitz,W.; Engineering Design, A Systematic Approach, Springer-Verlag, Berlin, 1996,ISBN 3-540-19917-9

[9] Eskelinen, H. " Using Manufacturability Analysis for Efficient Design of Microwave Mechanics", *Engineering Mechanics* vol 6, No 1, 1999, pp. 71 – 74

MEASUREMENTS OF LOW-LOSS CRYSTALLINE MATERIALS FOR HIGH-Q TEMPERATURE STABLE RESONATOR APPLICATIONS

Michael E. Tobar[*], *Jerzy Krupka*[#], *John G. Hartnett*[*], *Richard G. Geyer*[&], *Eugene N. Ivanov*[*]
*Department of Physics, University of Western Australia, Nedlands, WA, Australia
Instytut Mikroelektroniki i Optoelektroniki PW, Koszykowa 75, 00-662 Warszawa, Poland
& National Institute of Standards and Technology, Electromagnetic Fields Division, Boulder, CO 80303, USA

ABSTRACT

Whispering gallery modes were used for very accurate permittivity and dielectric loss tangent measurements for low loss isotropic and uniaxially anisotropic materials. We present the measurements of several specimens including sapphire, YAG, quartz, rutile and $SrLaAlO_4$. The total absolute uncertainty in real part of the permittivity tensor was estimated to be less than 0.1% and was limited by the uncertainty in the dimensions of the samples. Imaginary parts of the permittivity tensor were measured to about 10% accuracy, limited by the accuracy of Q-factor measurements in whispering gallery modes. The anisotropy ratio of the measured materials varied from 1 (isotropic YAG) to 2.2 (rutile). All anisotropic materials exhibited anisotropy in the imaginary part of the permittivity tensor as well as the real part. For most crystals dielectric losses can be approximated by a power function of absolute temperature in only a limited temperature range. At very low temperatures (4-50 K) properties of both the real and imaginary permittivity tensor are often affected by impurities which are always present in real crystals.

1. INTRODUCTION

Low loss single crystals have become important materials for constructing high-Q temperature compensated resonators for frequency stabilised microwave oscillators. This paper reports on a collaborative effort to assess and characterise several single crystals using the Whispering Gallery Mode (WGM) method, which has proved to be the most accurate method for measurements of the complex permittivity of extremely low loss dielectrics [1-3].

1I. MEASUREMENT PROCEDURE

The most effective way to eliminate conductor losses for loss tangent measurements is to excite WGMs in a cylindrical specimen of the material. To distinguish the anisotropy for complex permittivity measurements the crystal axis must be aligned along the cylindrical axis, then the Q-factor and frequency of two modes must be measured. The modes must be significantly different and exhibit quasi TE and TM electromagnetic field structure. To solve for the tensor components of permittivity a system of two non-linear determinant equations are evaluated.

$$\left.\begin{aligned} F_1(f^{(H)}, \varepsilon_\perp, \varepsilon_{//}) &= 0 \\ F_2(f^{(E)}, \varepsilon_\perp, \varepsilon_{//}) &= 0 \end{aligned}\right\} \quad (1)$$

Here $f^{(H)}$ and $f^{(E)}$ are the measured frequencies of the quasi-TE (H) and the quasi-TM (E) WGMs and $\varepsilon_{//}$ and $\varepsilon_\perp$ are the real parts of the permittivity tensor perpendicular and parallel to the anisotropy axis respectively. The eigenvalue equations F_1 and F_2 result from the application of rigorous method of analysis of the resonant structure (we used the mode matching technique). Once the permittivities are evaluated from (1), the loss tangents are evaluated from (2),

$$\left.\begin{aligned} Q_{(E)}^{-1} &= p_{e\perp}^{(E)} \tan\delta_\perp + p_{e//}^{(E)} \tan\delta_{//} + R_S / G^{(E)} \\ Q_{(H)}^{-1} &= p_{e\perp}^{(H)} \tan\delta_\perp + p_{e//}^{(H)} \tan\delta_{//} + R_S / G^{(H)} \end{aligned}\right\} \quad (2)$$

Here $\tan\delta_\perp$ and $\tan\delta_{//}$ are the dielectric loss tangents perpendicular and parallel to the anisotropy axis: $p_{e\perp}^{(E)}$, $p_{e//}^{(E)}$, $p_{e\perp}^{(H)}$ and $p_{e//}^{(H)}$ are the electric energy filling factors perpendicular and parallel for E and H modes, $G^{(E)}$ and $G^{(H)}$ are the geometric factors and Rs is the surface resistance. The electric energy filling factors were determined from the incremental frequency rule [4].

$$\left.\begin{aligned} p_{e\perp}^{(E)} &= 2\frac{\partial f^{(E)}}{\partial \varepsilon_\perp}\frac{\varepsilon_\perp}{f^{(E)}} \\ p_{e//}^{(E)} &= 2\frac{\partial f^{(E)}}{\partial \varepsilon_{//}}\frac{\varepsilon_{//}}{f^{(E)}} \\ p_{e\perp}^{(H)} &= 2\frac{\partial f^{(H)}}{\partial \varepsilon_\perp}\frac{\varepsilon_\perp}{f^{(H)}} \\ p_{e//}^{(H)} &= 2\frac{\partial f^{(H)}}{\partial \varepsilon_{//}}\frac{\varepsilon_{//}}{f^{(H)}} \end{aligned}\right\} \quad (3)$$

The geometric factors can be calculated from,

$$G^{(E)(H)} = \omega \frac{\iiint_V \mu |\bar{H}|^2 dV}{\iint_S |\bar{H}_t|^2 dS} \quad (4)$$

In practice for a properly chosen azimuthal mode number and a sufficiently large shield, the geometric factor is large enough to ignore the effect of R_S.

III RESULTS

Results of measurements of the real permittivities versus temperature for several single crystals are shown in Figs. 1-6. Absolute accuracy was limited principally by uncertainty in the dimensions of the samples and estimated to be better than 0.1%. Relative accuracy versus temperature depends on inaccuracies in the

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

thermal expansion coefficient [5-8]. Influence of thermal expansion is especially important at low temperatures where permittivity changes are small. At temperatures in the range of 4-20 K. The presence of paramagnetic impurities can affect the measurement uncertainty depending on the kind and amount [9]. Such variations are visible in quartz Figs.2-3 at temperatures 4-20 K. For most materials, except rutile, the permittivity increases with respect to temperature above 100 K. For this reason rutile is an ideal material to combine with other materials and construct thermally compensated resonators [10-13]. All materials measured by us, except YAG, exhibit dielectric uniaxial anisotropy. The anisotropy ratio, $\varepsilon_{//}/\varepsilon_{\perp}$, varies with temperature and it is different for various materials, as shown in Table1. At 20 K, rutile had the largest anisotropy ratio, 2.23, while quartz had the smallest of 1.044.

Table 1. Anisotropy ratio for various materials

Material	Anisotropy ratio at 20 K	Anisotropy ratio at 300 K
Sapphire	1.224	1.232
Rutile	2.23	1.90
Quartz	1.044	1.045
$SrLaAlO_4$	1.16	1.18

Table 2. Dielectric loss factors and dielectric loss-temperature exponents at 100 K

GHz	Material	$A_\perp$	$A_{\|\|}$	$K_\perp$	$K_{\|\|}$
22	Sapphire	$5\ 10^{-16}$	$2\ 10^{-16}$	4.5	4.5
17	YAG	$1\ 10^{-12}$		3.0	
4.5	Rutile	$3.2\ 10^{-11}$	$1.2\ 10^{-10}$	2.7	2.7
18	Quartz	$8\ 10^{-6}$	$5\ 10^{-6}$	≈0	≈0
12	$SrLaAlO_4$	$1.6\ 10^{-9}$	$4\ 10^{-9}$	1.7	1.7
10	Sapphire[6]	$2\ 10^{-16}$	$8\ 10^{-17}$	4.5	4.5
5	Rutile [11]	$2.8\ 10^{-11}$		2.7	

Table 3. Dielectric losses at 100 K

GHz	Material	$\tan\delta_\perp$	$\tan\delta_{\|\|}$
22	Sapphire	$5\ 10^{-7}$	$2\ 10^{-7}$
17	YAG	$1\ 10^{-6}$	
4.5	Rutile	$8\ 10^{-6}$	$3\ 10^{-5}$
18	Quartz	$8\ 10^{-6}$	$5\ 10^{-6}$
12	$SrLaAlO_4$	$4\ 10^{-6}$	$1\ 10^{-5}$
10	Sapphire [6]	$2\ 10^{-7}$	$8\ 10^{-8}$
5	Rutile [11]	$7\ 10^{-6}$	

In Figs.7-11 loss tangents measurements are shown. For all materials the dielectric loss tangent increases with temperature. The temperature where the slope of dielectric loss tangent versus temperature approaches maximum on is about 100 K. Near 100 K, dielectric losses can be approximated by the following power functions of absolute temperature:

$$\tan\delta_\perp = A_\perp T^{K_\perp}$$
$$\tan\delta_{||} = A_{||} T^{K_{||}} \quad (4)$$

Table 2 shows the evaluation of the parameters at 100 K given from (4), while Table 3 shows the loss tangents at 100 K. Data in the last two rows of Table 2 and Table 3 were taken from literature. When comparing data one should notice that dielectric losses are frequency dependent. For example, for sapphire losses increase approximately linearly with frequency. This means that the data measured at 22 GHz (first row) and 10 GHz (sixth row) are in good agreement.

There are theoretical papers on dielectric losses in pure single crystals [10] that predict temperature and frequency dependence for various crystal lattice structures. Measurements show that losses in real crystals exhibit more complex behaviour, even in limited temperature ranges. For most crystals at very low temperatures losses are affected by impurities, in other crystals like quartz piezoelectric effects dominate the loss mechanism.

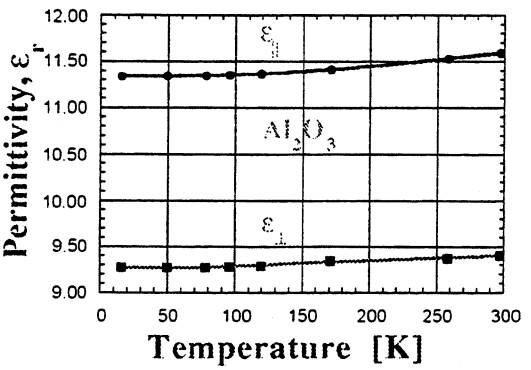


Fig.1. Permittivity versus temperature for sapphire.

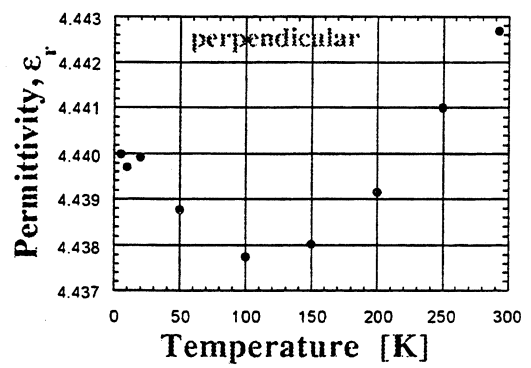


Fig.2. Perpendicular permittivity versus temperature for quartz.

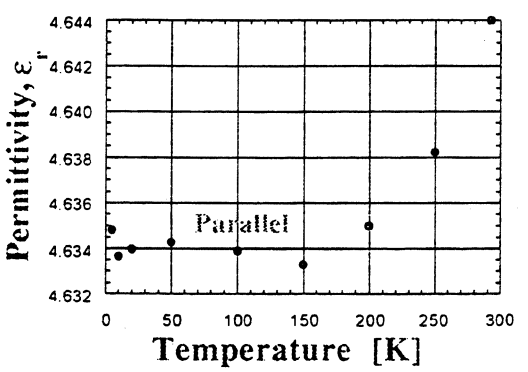


Fig.3. Parallel permittivity versus temperature for quartz.

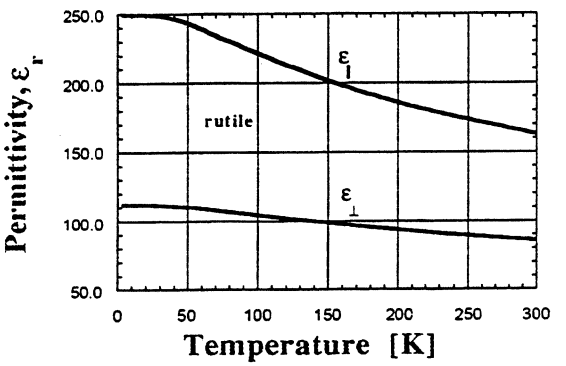


Fig.4. Permittivity versus temperature for rutile.

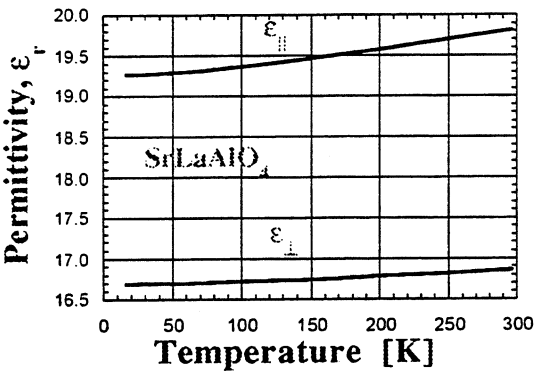


Fig.5. Permittivity versus temperature for strontium lanthanum aluminate.

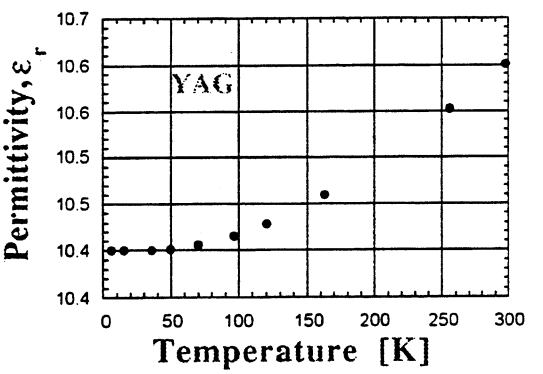


Fig.6. Permittivity versus temperature for YAG

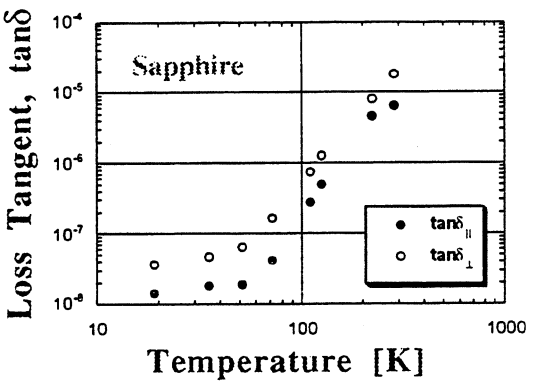


Fig.7. Dielectric loss tangent versus temperature for sapphire.

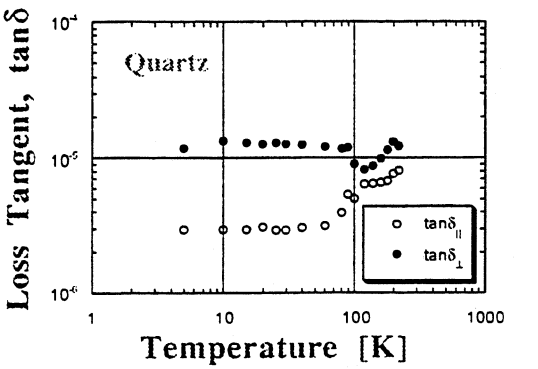


Fig. 8. Loss tangent versus temperature for quartz.

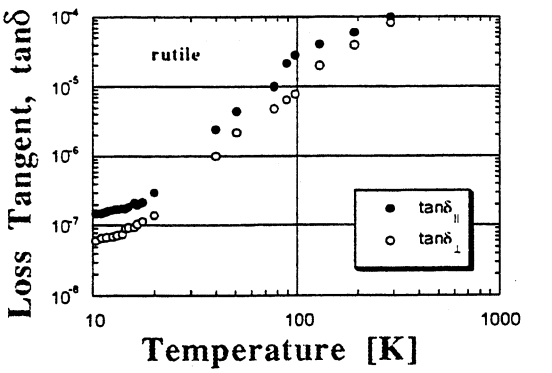


Fig. 9. Loss tangent versus temperature for rutile.

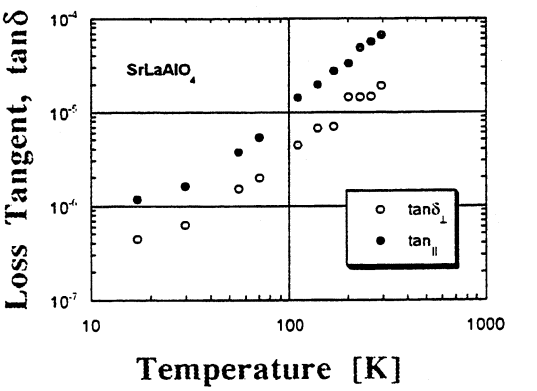


Fig. 10. Loss tangent versus temperature for strontium lanthanum aluminate.

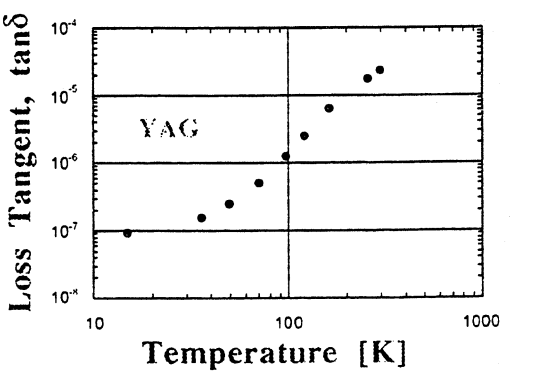


Fig. 11. Loss tangent versus temperature for YAG.

DISCUSSION

Results presented in this paper create an accurate permittivity data basis for a few ultra-low loss crystals. Also, the dielectric loss tangent data presented here can be treated as reference values, however only within a limited temperature range not affected by impurities (which are always present in real materials). Losses in real crystals exhibit more complicated behaviour than that predicted by simplified theoretical models presented in literature. Power loss-temperature formulae describing dielectric loss can be applied only in limited temperature ranges. Further studies of more materials at various frequencies using the WGM technique are necessary. This data could act as an experimental basis for new theories, which may predict accurately the complex permittivity of materials.

The technique of combining low-loss dielectrics has proven to be excellent for building high-Q temperature compensated microwave resonators [10-13]. Specifically we have concentrated on combining sapphire and rutile [11]. The sapphire-rutile resonator consists of a large piece of sapphire with small pieces of rutile acting to perturb the resonant frequency of the sapphire resonator. Rutile has low loss and opposite Temperature Coefficient of Permittivity (TCP) to sapphire and thus can annul the frequency-temperature characteristic of the sapphire resonator with minimum degradation in Q-factor. Because the field is very small in the rutile (only acts as a perturbation) there are no high field gap regions, and also because it is physically very small, the thermal time constant is very fast and insignificant when compared to sapphire. Thus, the rutile temperature should follow that of sapphire. Also, the structure can be held rigidly fixed together, which is necessary to eliminate g-sensitivity.

The disadvantage of such a device is that it is very hard to design the structure to have a specific frequency-temperature annulment point. However, this problem has been solved due to collaboration with the Institut de Recherche en Communications Optiques et Microondes (IRCOM) at the University of Limoges. Recent preliminary work proves that finite element software developed at IRCOM is capable of accurately designing such structures [14,15]. This gives us the power to design resonators with specific annulment temperatures which would not be otherwise possible.

8. ACKNOWLEDGMENT

This work was supported by the Australian Research Council and the Department of Industry Science and Tourism.

9. REFERENCES

[1] J. Krupka, K. Derzakowski, A. Abramowicz, M. Tobar and R. G. Geyer, "Measurements of the complex permittivity of extremely low loss dielectric materials using whispering gallery modes", IEEE MTT-Symp. Dig., Denver, June 8-14, pp.1347-1350, 1997.

[2] M.E. Tobar, J. Krupka, E.N. Ivanov, R.A. Woode, "Measurement of the complex permittivity of rutile between 10 to 300 Kelvin using whispering gallery modes", J.Appl. Phys., vol. 83, no. 3, pp.1604-9, 1998.

[3] J. Krupka, K. Derzakowski, A. Abramowicz, M. Tobar and R. G. Geyer, "Use of whispering gallery modes for complex permittivity determinations of ultra-low loss dielectric materials", IEEE Trans. MTT, June 1999.

[4] Y. Kobayashi and T. Senju, "Resonant modes in shielded uniaxial-anisotropic dielectric rod resonators", IEEE Trans. Microwave Theory Tech, vol. 41, 2198-2205, December 1993.

[5] Y. Kobayashi, Y. Aoki, and Y. Kabe,"Influence of conductor shields on the Q-factors of a TE_0 dielectric resonator", IEEE MTT-S Int. Microwave Symp. Dig., pp.281-284, St. Louis, 1985.

[6] C.A. Swenson, R.B. Roberts, and G.K. White, *Thermophysical Properties of Some Key Solids*, Ch.4, Editors G.K. White M.L. Minges, Oxford Pergamon Press, CODATA Bulletin 59, 1985.

[7] G.K. White, "Reference materials for thermal expansion: certified or not ?" Thermochimica Acta, vol. 218, pp.83-99, 1993.
G.K. White, "Thermal expansion of silica at low temperatures", Cryogenics, pp.1-7, Feb. 1964.

[8] N. Klein, C. Zuccaro, U. Dahne, H. Schultz, and N. Tellmann, "Dielectric properties of rutile and its use in high temperature superconducting resonators", J. Appl. Phys., vol.78, pp.6683-6686, Dec. 1995.

[9] V.L. Gurevich and A.K. Tagantsev, "Intrinsic dielectric loss in crystals: low temperatures", Sov.Phys. JEPT, vol. 64, pp.142-151, July 1986.

[10]. M.E. Tobar, J. Krupka, E.N. Ivanov, R.A. Woode, "Dielectric frequency-temperature compensated whispering gallery mode resonators", J. Appl. Phys D, vol. 30, no. 19, pp. 2770-2775, 1997.

[11] M.E. Tobar, J. Krupka, J.G. Hartnett, E.N. Ivanov, R.A. Woode, "High-Q sapphire-rutile frequency-temperature compensated microwave dielectric resonators," IEEE Trans. on UFFC, vol. 45, no. 3, pp. 830-836, 1998.

[12] N. Klein, A. Scholen, N. Tellmann, C. Zuccaro, K.W. Urban, "Properties and applications of HTS-shielded dielectric resonators: a state-of-the-art report," IEEE Trans. MTT, vol. 44, no. 7, pp. 1369-73, 1996.

[13] I.S. Gosh, D. Schemion, N. Klein, "Temperature compensated high-Q dielectric resonators for long term stable low phase noise oscillators." 1997 IEEE Freq. Contr. Symp., pp. 1024-1029.

[14] M.E. Tobar, J.G. Hartnett, D. Cros, P. Blondy, J. Krupka, E.N. Ivanov, P. Guillon, "Design of high-Q frequency-temperature compensated dielectric resonators, Eletron. Lett., vol. 35, no. 4, pp. 303-305 1999.

[15] M.E. Tobar, P. Blondy, D. Cros, J.G. Hartnett, P. Guillon, "Finite Element Realisation of Ultra-High Quality Factor Frequency-Temperature Compensated Sapphire-Rutile Whispering Gallery Mode Resonators," in IEEE MTT-Symp. Dig., 1999.

1999 Joint Meeting EFTF - IEEE IFCS

APPLICATIONS OF COUPLED DIELECTRIC RESONATORS USING $SrTiO_3$ PUCKS: TUNEABLE RESONATOR AND NOVEL THERMOMETRY

J C Gallop & L Hao
National Physical Laboratory, Teddington, TW11 0LW, UK

ABSTRACT

The combination of very low loss dielectric single crystal materials and cryogenic temperatures (Below about 80 K) has led to the achievement of high Q resonators which show great promise for frequency standard applications. A number of problems remain to be solved. We have already reported how high frequency stability with minimum temperature control may be achieved by means of composite dielectric pucks, the components having opposite signs for their temperature coefficients of permittivity. A second requirement for a frequency standard is that it should possess a specified frequency. Here we report implementation of electronic tuning of a sapphire dielectric puck resonator by using a $SrTiO_3$ (STO) tuning element situated in the evanescent field region outside the sapphire puck. In addition the same structure may be used when the sapphire and STO elements act as weakly coupled resonators so that the resonant frequency becomes a very sensitive function of temperature, allowing the possibility of very high resolution thermometry.

1. INTRODUCTION

The availability of high purity single crystal dielectrics (such as sapphire) which exhibit very low loss tangents at microwave frequencies for temperatures below about 80K, combined with high temperature superconducting thin films, has led to the production of very high Q compact microwave resonators with many potential applications in microwave communications (see, for example, [1],[2]). As well as the utility of such devices for filter applications in mobile and satellite communications systems we have already shown elsewhere that very stable oscillators with long term stability and very low phase noise performance may be built [3], especially if frequency temperature compensation is employed [4].

The resonant frequency of a dielectric resonator is determined by its relative permittivity and physical dimensions. To produce a resonator with a specific resonant frequency one must have well characterised and homogeneous single crystal dielectrics and good electromagnetic modelling capabilities. However we estimate that, at best, the machining tolerances (including crystallographic axes alignment), material parameters and modelling accuracy limit the ability to produce a required frequency to an accuracy of no better than ~ 0.01 %. Much greater accuracy is required for local oscillator or frequency standard applications and tuning to an accuracy of perhaps 1 in 10^{10} (1 Hz at 10 GHz) is desirable (1 in 10^6 of the tuning range of 0.01 %).

Various tuning techniques may be attempted and we list three possible methods with their main disadvantages below. First temperature tuning may be implemented, where small controlled temperature changes affect the oscillator frequency via the slowly varying function df/dT. However typically the resonator is operated at a temperature turning point where df/dT=0, to realise the highest stability [4] which thus eliminates the possibility of temperature tuning. Mechanically adjustable elements within the resonator housing have also been used to provide tuning (see, for example, [5]). When the highest stability is required mechanical adjustment has the disadvantage that the mechanism will almost certainly be prone to vibrational instabilities. Also automated operation is bulky, complex and slow.

The most appropriate form of tuning, from the point of view of simplicity, speed and noise isolation could be provided by voltage controlled permittivity of a dielectric element in the resonator structure, such as is demonstrated by some paraelectric or ferroelectric high permittivity materials. In the remainder of this paper we explore this latter approach to fine tuning.

2. TEMPERATURE AND VOLTAGE DEPENDENCE OF RELATIVE PERMITTIVITY OF $SrTiO_3$

Strontium Titanate ($SrTiO_3$) is a perovskite structured insulator with an exceptionally high relative permittivity. It is structurally and chemically highly compatible with the cuprate high temperature superconductors. Its relative permittivity is very high even at room temperature and shows considerable crystalline anisotropy. As the temperature is lowered the permittivity rises strongly and bulk crystals of $SrTiO_3$ enter a paraelectric state before becoming ferroelectric below about 35 K. In both of these states the permittivity can be controlled by an applied electric field.

3. EXPERIMENTAL REALISATION OF TUNING

Figure 1 shows our measurements of the temperature dependence of the relative permittivity of a bulk single crystal of $SrTiO_3$ (10x10x0.5 mm^3), using a metallised parallel plate resonator technique [6] whereas Fig. 2 shows the effect of an applied d.c. electric field to the same structure. Both figures indicate measurements made using different resonant modes in the parallel plate

0-7803-5400-1/99/ $ 10.00

structure. Note that the agreement between measurements from the different modes is quite good. Discrepancies arise mainly from edge effects since the fringing field in the region outside the dielectric of a parallel plate resonator with open side boundary conditions is not negligible, whereas the calculated values for permittivity and change of permittivity with electric field neglect the fringing fields. The values for $\varepsilon_r(T)$ and $d\varepsilon_r(T)/dT$ are in quite good agreement with other published data for single crystal STO [7,8].

4. RESONATOR TUNING RESULTS

The loss tangent $\tan\delta(T)$ of STO at cryogenic temperatures is not as low as for sapphire, being typically in the range 3×10^{-5} to 10^{-3}. Microwave losses in the STO tuning element must not compromise the overall Q of the resonator (since this is essential for high oscillator stability) so the STO tuning element is situated in the fringing field of the main resonator so that its electromagnetic field filling factor is small. It is coated with thick film YBCO on both sides providing electrodes to apply a d.c. electric field. Various patterns for the HTS electrodes are being tested, chosen to maximise the coupling of the fringing field to the tuning element for a variety of resonant modes.

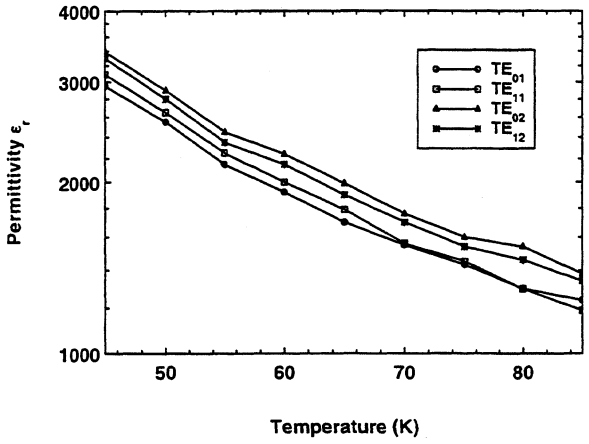


Fig. 1. Temperature dependence of the relative permittivity ε_r of a bulk single crystal of $SrTiO_3$.

The pattern is applied to one HTS coated face of the disc shaped STO tuning element, the other side is coated with a plain HTS thick film ground plane. Wires are attached to the HTS films by a small amount of silver loaded paint, applied at a point where the fringing field should be minimal.

The STO disc is positioned close to the sapphire puck, with the HTS ground plane attached to a copper plate, the axial position of which can be adjusted by a simple screw mechanism. The whole composite structure is enclosed in a copper box to reduce radiation losses and

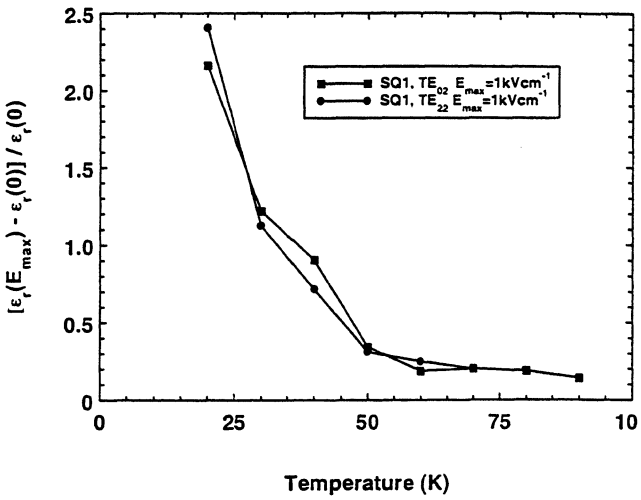


Fig. 2. Temperature dependence of the total percentage change in permittivity ε_r due to the application of a dc electric field.

provide an isothermal enclosure. Microwave power may be coupled in and out through two adjustable loops soldered to the ends of coaxial cables on opposite sides of the copper housing. The resonator is incorporated into a loop oscillator geometry. When the phase shifter is adjusted so that the total phase shift around the external microwave circuit is equal to an integer times 2π the gain in the amplifier is sufficient to cause the circuit to oscillate at the centre frequency of the sapphire resonant mode. This frequency is counted and recorded by a frequency counter as the direct voltage applied to the STO tuning element is varied. Figure 3 shows a schematic layout of the loop oscillator.

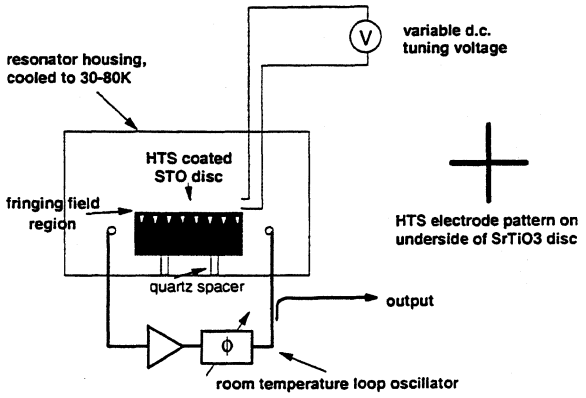


Fig. 3. Schematic of oscillator with HTS/STO fine tuning.

Figure 4 shows results for the effect of a d.c. voltage V applied across the STO tuning element on the frequency of the loop oscillator referenced to the TE_{011} mode of the sapphire puck. Note that a modest voltage change of ~50 V can produce a frequency change of ~0.03%. This should be compared with the design target suggested above of 0.01% so that a tuning voltage of less than 20V

should be adequate to achieve this. At T = 35 K slight hysteresis is visible. At a higher temperature (T > 40 K) this is undetectable though the tuning range is reduced.

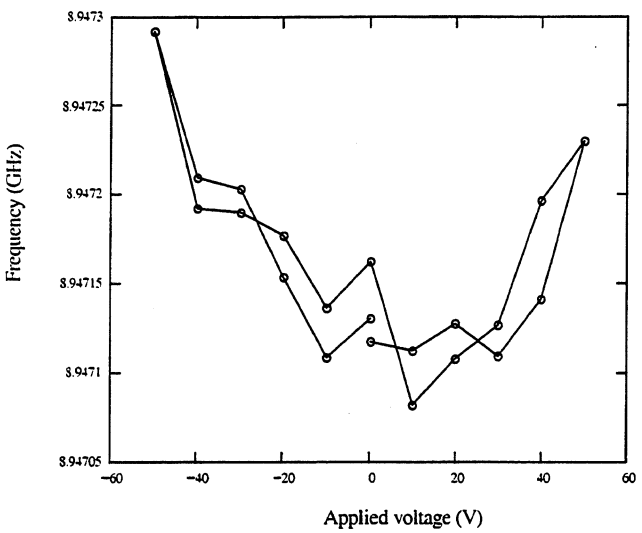


Fig. 4. Tuning effect of STO applied voltage, measured at 35 K.

5. COUPLED RESONATORS

The work described in the previous section concerns a configuration where the sapphire puck and the STO puck are rather tightly coupled together. The temperature and voltage change of the composite puck resonant frequency can be reasonably accurately modelled in terms of a voltage or temperature dependent perturbation through $\varepsilon_r(V,T)$. A different situation arises when the microwave fields of the sapphire and STO pucks are only extremely weakly coupled. Then the resonant frequency of a particular mode of the sapphire puck f_0 (assumed temperature independent to a first approximation) is quite unaffected by the STO except for a temperature selected resonance condition when the resonant frequency $g_0(T)$ of a mode in the STO comes into close coincidence with that of the selected sapphire mode. In this situation there is a measurable interaction between the two modes so that they may be treated as coupled independent resonators [9]. The following four equations describe the resulting temperature dependent frequencies and linewidths:

$$f(T) = f_0(T) + \mathrm{Re}\left[\frac{A}{(f_0 - g(T)) + iW_{STO}}\right]$$

$$g(T) = g_0(T) + \mathrm{Re}\left[\frac{A}{(g_0(T) - f_0) + iW_{sap}}\right]$$

$$W_{sap}(T) = W_{sap} + \mathrm{Im}\left[\frac{A}{(f_0 - g(T)) + iW_{STO}}\right]$$

$$W_{STO}(T) = W_{STO} + \mathrm{Im}\left[\frac{A}{(g_0(T) - f_0) + iW_{sap}}\right]$$

Here A is the coupling strength between the two modes and is essentially proportional to the overlap of the electromagnetic standing wave patterns of the stored energy of the two field distributions of the modes, integrated throughout the housing. W_{sap} and W_{STO} are the unperturbed linewidths of the sapphire and $SrTiO_3$ resonances respectively. Which of the two coupled modes is observed in any experiment depends on the nature of the input and output coupling structures, especially their positions. If these are situated closer to the sapphire puck then the mode observed is the one which has the dominant stored energy within the sapphire puck.

Figure 5 shows some experimental results for the temperature variation of such a coupled mode (predominantly the TE_{011} mode in sapphire) as the temperature is changed over a small range. Note that two STO resonances in turn come in to coincidence with the sapphire resonance, each producing similar frequency and width shifts. The solid curves represent fits to the experimental data for frequency shift and linewidth, using the above equations and treating ε_r as a linear function of T, with coefficient as derived from Fig. 4 (this is an excellent approximation in view of the very limited temperature range). The agreement between experiment and the model is seen to be very good. Note that the coupled resonant frequency varies rapidly over a temperature range of around 50 mK.

The coupled resonators provide the possibility of sensitive temperature change measurements to be made. Thus the rate of change of resonant frequency with temperature can be at least as high as 75 MHz/K (see Fig. 5). Since the output frequency of a microwave loop oscillator based on a high Q dielectric resonator can be stable to at least 1 in 10^{11} for an averaging time of 1s [10] this thermometer has a potential temperature resolution as low as 1.5 nK, comparable with the best high resolution susceptibility or superconducting transition edge thermometers. It has an added advantage that it is non contacting since the temperature change of interest would be occurring in the STO element which is coupled through the fringing electromagnetic field to the main resonator. The latter would require to be separately temperature compensated using, for example a combination of sapphire and rutile elements [11].

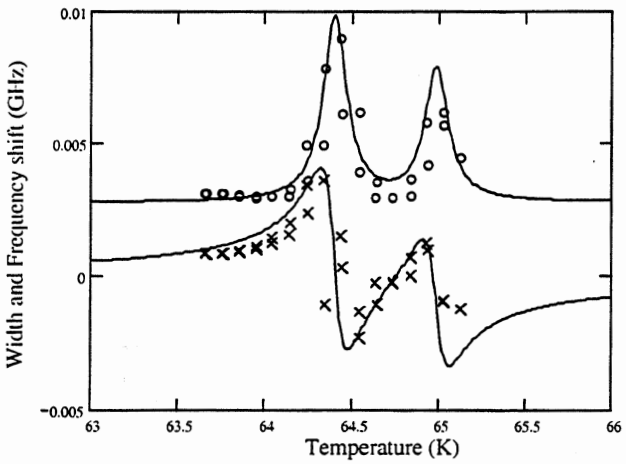


Fig. 5. Experimental results (circles and crosses) for the temperature variation of a coupled mode (predominantly the TE_{011} mode in sapphire) as the temperature is changed over a small range. The solid curves represent fits to the experimental data for frequency shift (bottom) and linewidth (top).

6. CONCLUSIONS AND FURTHER WORK

We have demonstrated in principle that electronic fine tuning of HTS shielded dielectric resonators may be carried out using low loss tuning elements of single crystal STO to which small voltages may be applied to adjust the relative permittivity and hence the combined dielectric resonator resonant frequency. Next we plan to maximise the tuning by positioning the tuning element closer to the main dielectric resonator. A resonator housing with a screw adjustment has been constructed which will allow smooth variation of the tuning element position. Different electrode structures have been fabricated to maximise the tuning effect on the oscillator frequency. The loaded Q values currently attained with the tuneable resonators are relatively low, ~$3x10^3$. It is believed this results from the presence of silver paint electrodes and normal metal wiring to the STO tuning elements.

In the near future we plan to position the tuning elements in future so that the stray fields are minimal in the region of the non-superconducting connection to the STO electrodes. Even at cryogenic temperatures the loss tangent in single crystal samples of STO is typically ~ 10^{-4}. However for the fine tuning application described here the filling factor (that is the fraction of the total stored electromagnetic energy) in the STO element will be only of order 10^{-3} so that overall loaded Q values as high as 10^7 may be expected. Further work is planned to evaluate the potential of the coupled resonator geometry as an ultra high resolution non-contacting thermometer. Issues of mechanical stability and microwave power dissipation in the STO element will need to be addressed.

REFERENCES

[1] M A Hein, "Progress, properties and prospects of passive high-temperature superconductive microwave devices in Europe," *Supercond. Sci. Technol.*, vol. 10, pp. 867-71, 1997.

[2] N. Newman and W.G. Lyons, "HTS microwave devices: fundamental issues in materials, physics and engineering," *J. Supercond.*, Vol. 6, pp. 119-60, 1993.

[3] J C Gallop, L Hao, F Abbas and C D Langham, "Frequency stability of dielectric loaded HTS microwave resonators," *IEEE Trans. Appl. Supercond.*, Vol. 7, pp. 3504-7, 1997.

[4] L Hao, J Gallop and F. Abbas, "High Q HTS Shielded composite puck resonators," Proc. 3rd European Conference on Applied Superconductivity, IOP Conference Publication vol. 158, pp. 275-8, 1997.

[5] I S Ghosh, N Tellmann, D Schemion, A Scholen and N Klein , "Low phase noise microwave oscillators based on HTS shielded dielectric resonators," *IEEE Trans. Appl. Supercond.*, Vol. 7, pp. 3071-4, 1997.

[6] D Lacey, J C Gallop and L E Davis, "The effects of an air gap in the measurement of the dielectric constant of $SrTiO_3$ at cryogenic temperatures," *Meas. Sci. Technol.*, Vol. 9, pp. 536-9, 1998.

[7] J Krupka, R G Geyer, M Kuhn and J H Hinken, "Dielectric properties of single crystals of Al_2O_3, $LaAlO_3$, $NdGaO_3$, $SrTiO_3$ and MgO at cryogenic temperatures," *IEEE Trans MTT*, vol. 42, pp. 1886-9, 1994.

[8] S Gevorgian, E Carlsson, P Linner, E Kollberg, O Vendik and E Wikborg, "Lower order modes of YBCO/STO/YBCO circular disc resonators," *IEEE Trans MTT*, vol. 44, pp. 1738-41, 1996.

[9] See for example A B Pippard, "The physics of vibration" Cambridge UP, 1978.

[10] L Hao, N Klein, W J Radcliffe, J C Gallop and I S Ghosh "Temperature compensated cryogenic whispering gallery mode resonator for microwave frequency standard applications," Proc.12th European Frequency and Time Forum, pp,112-4, 1998.

[11] J C Gallop, C D Langham, L Hao and Farhat Abbas, "Dielectric loaded HTS resonators as frequency standards and low-phase noise oscillators," *IEEE Trans. Instrum. Meas.*, Vol.46, pp.122-5 , 1997.

1999 Joint Meeting EFTF - IEEE IFCS

COMPACT TEMPERATURE-COMPENSATED CRYOGENIC WHISPERING GALLERY MODE RESONATOR OPERATED AT 63 K IN A CLOSED CYCLE COOLER

L. Hao[+], N. Klein[*], J. C. Gallop[+] and M. Winter[*]

NPL[+], Teddington, TW11 0LW, UK
FZ Jülich [*], D - 52425 Jülich, Germany

ABSTRACT

Single crystal dielectric materials such as sapphire demonstrate remarkably low values for the dielectric loss tangent, even at rather modest cryogenic temperatures (40-80 K) which are readily achieved using closed cycle coolers. We have taken our previously reported novel design [1,2] for a composite dielectric resonator incorporating both sapphire and rutile single crystal elements and further improved the Q value while raising the frequency versus temperature compensation point to greater than 60 K. At the same time the geometrical size of the dielectric pucks has been increased while still using the same closed cycle cooler so that a compact loop-oscillator system is realised at a frequency close to the Cs atomic clock hyperfine frequency (~10 GHz) and an operating temperature of 63.5 K. Unloaded Q values as high as 10^7 have been achieved under these operating conditions for this composite resonator.

In addition we have investigated a compensated sapphire/rutile whispering gallery mode system at 17 GHz with a view to reducing the phase noise of oscillator based on this technology. Such oscillators can be used both as low phase noise sources and as long term stable frequency standards. The following sections describe some of the issues to be addressed to optimise performance.

1. CRYOGENIC DIELECTRICS & RESONATOR DESIGN

The loss tangent tanδ of dielectric single crystalline materials such as sapphire (Al_2O_3), magnesium oxide (MgO) and rutile (TiO_2) [3] is strongly temperature dependent, falling rapidly as the temperature is reduced below room temperature. For even the highest purity materials this rapid reduction with temperature is replaced by a low temperature plateau region where the losses are dominated by extrinsic contributions such as impurities and a variety of point defects.

The phase noise of a low phase noise oscillator based on a simple feedback circuit can be expressed by Leeson's formula [4]:

$$S_\phi(f_m) = 10 \cdot \log\left[\left(1+\frac{f_0^2}{4Q_l^2 f_m^2}\right)\left(\frac{\alpha}{f_m}+\beta\right)\right] \quad (1)$$

For a given amplifier with phase noise (1/f and white contributions) $10.\log(\alpha/f_m+\beta)$ the oscillator phase noise scales approximately like $1/Q_L^2$. Typically, commercial amplifiers based on FET transistors have amplifier noise of -135 dB_c/Hz at f_m= 1kHz. In order to meet a specification of -125 dB_c/Hz at 1 kHz with a simple loop oscillator, Q_L of about 10^7 is required.

A frequency standard consists of a stable oscillator, in which a high Q resonator can be used as the frequency discriminator. Assuming an infinitely stable resonator, frequency deviations of the oscillator are proportional to the ability of the stabilisation circuit to keep the oscillator frequency within a fraction χ of the half power resonance linewidth $\Delta f_{1/2} = f / Q_L$. The resulting fractional frequency stability is:

$$\delta f / f = \chi / Q_L \quad (2)$$

For $\chi=10^{-6}$ a stability of about 10^{-13} results for $Q_L=10^7$. Assuming such a value to be independent of the integration time τ, the stability of state-of-the-art frequency standards could be exceeded for certain ranges of τ:

Table 1

Type of Standard	Measuring time
hydrogen maser (JPL)	$\tau < 0.1$ sec
passive hydrogen maser:	$\tau < 10$ sec
1[ary] caesium standard:	$\tau < 10^2$ sec
Commercial Cs standard:	$\tau < 10^5$ sec
Commercial Rb standards	for all τ
helium cooled SCMO (JPL)	$\tau > 10^5$ sec

The resonator type which has the best potential for Q_L of the order of 10^7 at 10 GHz and temperatures achievable with one-stage compact cryocoolers (T>40K) is whispering-gallery-mode resonators (WGRs) using sapphire as the dominant dielectric material. Temperature compensation may be achieved by combining a sapphire element with one or more rutile elements for which the temperature dependence of the permittivity has the opposite sign to that of sapphire [3]. The frequency versus temperature relationship may be calculated from Eq. 3 where κ is the *rutile filling factor* (fractional electromagnetic energy stored in the rutile).

$$f(T) = f(0)\left[(1-\kappa)\sqrt{\frac{\varepsilon_s(0)}{\varepsilon_s(T)}} + \kappa\sqrt{\frac{\varepsilon_r(0)}{\varepsilon_r(T)}}\right] \quad (3)$$

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

Polynomial fits to experimentally determined temperature dependences of the permittivity $\varepsilon(T)$ of sapphire and rutile are used in Eq. (3). There exists a maximum in the f(T) curve for all values of κ. From the T-derivative of Eq.3 the relative frequency deviation for a 1 mK temperature deviation from this turning point has also been calculated for a range of κ values. The ultimate quality factor of the composite resonator can be calculated from the loss tangent $\tan\delta$ of sapphire and rutile:

$$Q(T) = \frac{1}{\kappa \tan\delta_r(T) + (1-\kappa)\tan\delta_s(T)} \qquad (4)$$

For $\tan\delta(T)$ measured temperature dependences were used [3,5]. For $\tan\delta(T \to 0K)$ we assumed a value of 10^{-8} both for rutile and sapphire. Table 2 summarises these results. These calculations show that it is possible to achieve Q's of 10^7 for a turning point temperature attainable on a low-power cryocooler. Therefore the design parameters for the resonator were selected to achieve the specifications shown by the shaded line in the table.

Table 2

κ	T_{TP}	$Q_{0,max}$	$[f(T_{TP}+10^{-3})-f(T_{TP})]/f(T_{TP})$
0.001	41.3	$8.5\cdot10^7$	$8\cdot10^{-15}$
0.002	58.9	$1.9\cdot10^7$	$2\cdot10^{-14}$
0.0023	63.0	$1.22\cdot10^7$	$2.4\cdot10^{-14}$
0.003	71.4	$8.5\cdot10^6$	$3.2\cdot10^{-14}$
0.004	81.3	$2.8\cdot10^6$	$4\cdot10^{-14}$

The construction of the composite resonator is as follows: a stack consisting of 2 quartz disks (radius = 1.5mm), two thin (≈ 100μm) rutile disk and a sapphire puck (radius = height = 14.8 mm) is enclosed in a cylindrical housing machined from OFHC-copper (see Fig. 3). In order to get the designed κ-values with a rutile plate thickness above 100μm the diameter of the plate has to be smaller than the diameter of the sapphire puck. The plate diameter was optimised with respect to azimuthal mode number of the WG mode ($7<n<12$) using MAFIA software. These calculations show that the loss contributions from the quartz support and copper shielding cavity may be made negligible by suitable choice of dimensions. The lower and upper quartz plate are glued to the copper housing and end plate, respectively, the stack being held together by a spring arrangement. It was felt important to avoid glue close to the sapphire and rutile pucks to achieve maximum Q values, the essential feature for the low phase noise oscillator. Coupling to the whispering gallery mode is provided by semi-rigid coaxial loops.

Figs. 1 and 2 show the experimental results obtained for frequency and Q_L value of a WG700 resonance (at a frequency of 9.975GHz) demonstrating that the turning point and maximum Q_L are close to the values predicted. Note that an unloaded Q_L value of $1.2\text{x}10^7$ at 63 K and 10 GHz has been achieved.

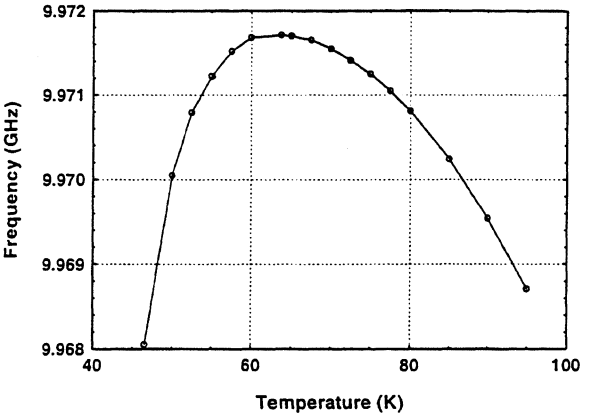


Fig. 1. Frequency vs. temperature for sapphire/rutile resonator. Solid line represents calculated dependence whereas the data points are experimental measurements.

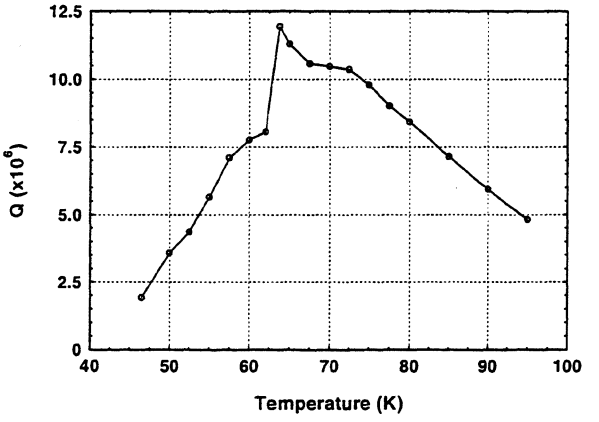


Fig. 2. Q_L vs. temperature for sapphire/rutile resonator.

2. CRYOGENIC DESIGN

The resonator based frequency standard is designed to operate continuously using a closed cycle cryocooler. With coolers of the Giffard-McMahon or Stirling configurations mechanical vibrations are significant and may limit the phase noise performance or long term stability of the standard. The key steps in reducing the effects of mechanical vibration consisted of attaching the coaxial lines connected to the resonator housing to the second stage of the cooler cold finger, the part which provides cooling for the radiation shield. This is much less subject to mechanical vibration from the displacer movement than is the cold finger itself. Cooling (and essentially the only mechanical support) from the cold finger to the resonator housing is provided by a bundle of copper braids which appear to be sufficiently flexible, even when cold, to provide mechanical vibration isolation. In addition three compression springs are included in the anti-vibration stage (as shown in Fig. 3) so that the axis of the housing was maintained vertical. Measurements of the vibration amplitude spectrum

were made at various points in the cryostat volume at room temperature and it has been shown that over most of the frequency range the vibration amplitude is reduced by greater than a factor of 10 between the cold finger and the resonator housing. The vibration spectra are shown in Fig. 4. The upper curve is the vibration spectrum on the cold head of GM refrigerator. Lower curve is from the top of copper housing with vibration isolation. It is clear that the vibrational amplitude has been reduced by more than 10 dB at all frequency between 100 Hz and 10 kHz. Further improvements are being investigated.

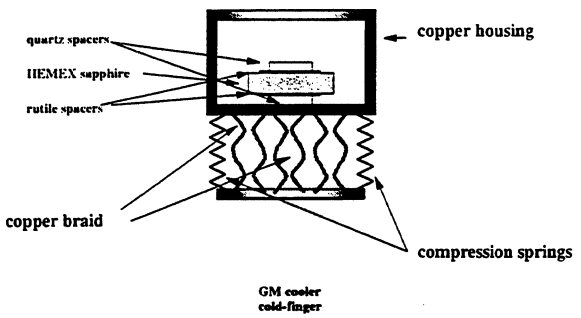


Fig. 3. Schematic of composite rutile/sapphire puck in copper housing, coupled to GM cooler cold stage via anti-vibration copper braid and compression springs.

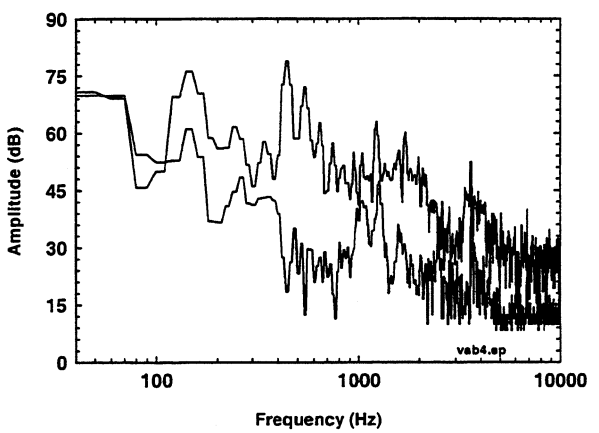


Fig. 4. Test of vibration isolation.

3. TEMPERATURE CONTROL

Recently two stage temperature control has been introduced into the Giffard-McMahon closed cycle cooler cold stage. Thus the cooler cold stage has a Si diode temperature sensor and 50 Ω heater which is controlled by a Lakeshore 93C temperature controller at a temperature a few kelvin below the turning point of the resonator. The resonator housing has a Pt resistance temperature sensor attached, together with a non-inductively wound phosphor bronze wire heater wrapped around its curved surface, with a resistance of around 30 Ω. A second temperature controller (Lakeshore model 340) is used to control the housing temperature close to the resonator turning point. This controller has a set point resolution of 1mK and the short term noise at a temperature of around 60 K is also ~ ±1 mK. Temperature measurements made over a long period of the housing Pt sensor deviations from the control point have been analysed using the two-sample variance concept similar to that used for frequency fluctuation analysis. A representative example of the square root two-sample variance of the temperature $\sigma_T(\tau)$ as a function of sampling time τ is shown in Fig.5. Note that the long term temperature stability is better than 100 μK, representing a frequency stability better than 10^{-15} if the resonator can be stabilised with 1 mK of the f(T) turning point temperature, well beyond the design specifications for this project.

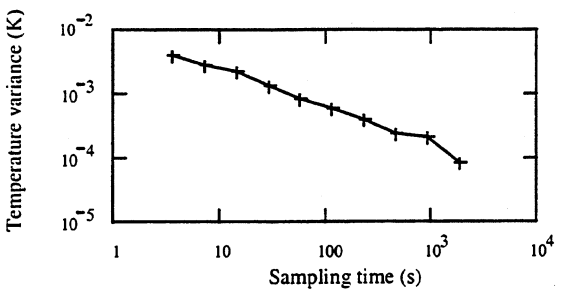


Fig. 5. Two-sample variance of the temperature stability of the G-M closed-cycle cooler cold head.

4. PHASE NOISE PERFORMANCE

A feedback oscillator has been assembled, based on the WGR plus a multi-stage low-noise FET amplifier with high amplification of 30-40dB, adjustable phase shifter and variable attenuator. The output is attenuated by 10-20 dB to achieve a loop gain of about 20dB so that the resonator operates with weak coupling (in particular at the output port). Low resonator power is also required to avoid drift of the resonance due to heating from microwave dissipation.

Because of the high Q value the phase noise of such an oscillator should be very low (see Eq. 1) and this means that unconventional methods are required to measure the spectral density of phase noise as a function of offset frequency $S_\varphi(f_m)$. A technique has been used in which the output of the oscillator is mixed with suitably amplified input to the resonator in phase quadrature [6]. This obviates the need for a tuneable reference oscillator with a lower phase noise than the oscillator under test.

Note that a performance of -120 dB_c/Hz is achieved at an offset of 1 kHz and at 10 kHz offset and greater the phase noise performance is limited by the white noise of the amplifier, at a level of -128dBc/Hz (see Fig. 6). Further significant improvement are possible with noise reduction techniques which have

been described elsewhere (see for example Galani et al. [7]). In this situation the mixer output from in phase and quadrature signals is amplified, appropriately low-pass filtered and fed back to a voltage controlled phase shifter included in the feedback oscillator circuit (shown schematically in Fig. 7).

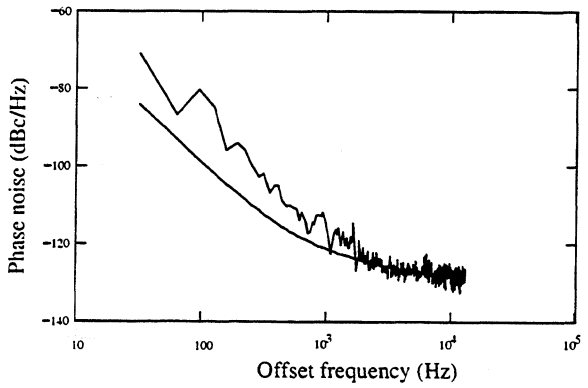


Fig. 6. Spectral density of phase noise from WG oscillator. Smooth curve represents fit to Eq. 1

With optimised amplifier gain the reduction in S_φ can be quite significant. This technique has so far only been applied to a 17 GHz whispering gallery mode of the rutile/sapphire system and the phase noise performance with and without the feedback applied are shown in Fig. 8. An equivalent reduction should be attainable with the 10 GHz mode also.

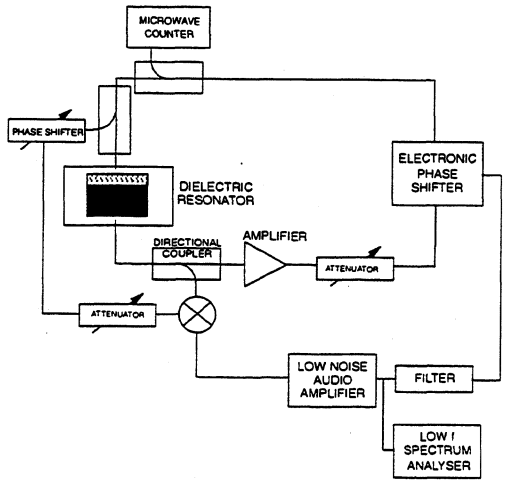


Fig. 7. Phase noise measurement circuit showing Galani noise reduction feedback element.

5. CONCLUSIONS AND FURTHER WORK

The work reported here shows for the first time that whispering gallery modes in temperature compensated composite dielectric resonators can exhibit loaded Q values above 10^7 at 10 GHz, that excellent long term temperature control is possible at around 60 K in a closed cycle cryocooler and that the high Q of the dielectric resonators can be used to produce low-phase noise oscillators limited only by the performance of amplifiers and mixers. Effective mechanical vibration isolation has also been demonstrated in a Giffard-McMahon cooler.

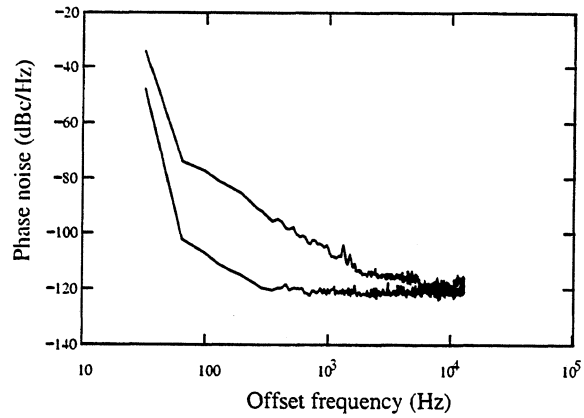


Fig. 8. Upper curve: phase noise at 17GHz as measured using system of Fig. 6. Lower curve: phase noise with Galani noise reduction applied.

Future work will concentrate on demonstrating the potential long term frequency stability of such an oscillator as well as making use of state-of-the-art microwave components to further improve the phase noise performance.

REFERENCES

[1] J. C. Gallop, C. D. Langham, L. Hao and F. Abbas, 'Dielectric loaded HTS resonators as frequency standards and low phase noise oscillators'", *IEEE Trans. Instrumentation and Measurement.* 46 pp.122-125 (1997)
L Hao, J Gallop and F. Abbas, 'High Q HTS Shielded composite puck resonators', Applied Superconductivity (IOP Conference Publication 158), pp. 275-8 (1997).
[2] L Hao, N. Klein, W. J. Radcliffe, J. C. Gallop and I. S. Ghosh, 'Temperature compensated cryogenic whispering gallery mode resonator for microwave frequency standard applications' Proc.12th EFTF (1998, Warsaw) pp112-114.
[3] N Klein et al. 'Rutile dielectric resonators shielded by YBCO films' J. Appl. Phys. 78 pp. 6683-7 (1995)
[4] D B Leeson, 'A simple model of feedback oscillator noise spectrum' IEEE Proceedings 54 pp.329-30 (1966)
[5] M E Tobar, J Krupka, J G Hartnett E N Ivanov & R A Woode, 'High Q sapphire-rutile frequency temperature compensated microwave dielectric resonators', IEEE Trans. Ultrasonics 45 pp. 830-5 (1998)
[6] F L Walls, A J D Clements, C M Felton, M A Lombardi & M D Vansk, 'Extending the range and accuracy of phase noise measurements', IEEE Int. Freq. Control Symp. 42 pp.432-41 (1988)
[7] Z Galani, M J Bianchini, R C Waterman, R Dibiase, R W Laton & J B Cole, 'Analysis & design of a single resonator GaAs FET oscillator with noise degeneration', IEEE Trans. MTT 32 pp.1556-64 (1984)

1999 Joint Meeting EFTF - IEEE IFCS

THERMAL STABILIZATION OF MICROWAVE SAPPHIRE RESONATOR REFERENCES.

Y.KERSALE, V.GIORDANO, F.LARDET VIEUDRIN,I.LAJOIE, M CHAUBET#
Laboratoire de Physique et Métrologie des Oscillateurs du CNRS associée à l'Université de Franche Comté
32, Avenue de l'observatoire - 25044 Besançon cedex – France
E-mail:kersale@lpmo.univ-fcomte.fr
#Centre National d'Etudes Spatiales 18 Avenue Edouard Belin-31055 Toulouse cedex – France

ABSTRACT

Sapphire single crystal, associated with a special mode configuration (Whispering Gallery Mode), is an ideal material for the realisation of high Q microwave resonator. Unfortunately, the frequency sensitivity to temperature fluctuations is relatively high. With such a sensitivity it is impossible to reach high frequency stability over a long time interval without efficient temperature regulation. We implemented original thermal regulation scheme on different sapphire resonators. We report, in this paper, frequency stability measurements of temperature controlled sapphire resonator oscillator at room temperature and at liquid nitrogen temperature.

1 INTRODUCTION

Traditional microwave oscillators using low Q dielectric resonators suffer from a high level flicker frequency noise. For some metrological applications, better spectral purity microwave oscillator are needed. Today, an emerging technology based on sapphire dielectric resonator enables to build microwave source presenting ultra low phase noise. Indeed the low dielectric losses of the sapphire monocrystal occurs to achieved Q factor of the order of 200000 at room temperature and better than 10^7 at the liquid nitrogen temperature [1]. Nevertheless, such a high Q factor can be only obtained with the use of a special mode configuration ,i.e. Whispering Gallery Mode Resonator (WGMR).

The principal difficulty to build a frequency reference based on a sapphire WGMR is the temperature sensitivity of the dielectric permittivity of the sapphire. Indeed the resonator frequency sensitivity is about –70ppm/K at room temperature and –10ppm/K at 77K. With such a sensitivity it is not possible to achieve long term frequency stability without efficient temperature regulation.

We have developed a temperature compensated resonator inspired by the work of J.Dick from the JPL, and a thermal regulation, using quartz resonator oscillator as sensor, to stabilise the sapphire resonator just near its temperature turning point. In this paper we report the configuration of our temperature compensated sapphire resonator oscillator and the details of our thermal regulator. Performances of the thermal regulator are given and frequency stability measurement of the temperature compensated resonator oscillator is also given.

2 SAPPHIRE RESONATOR PRINCIPLE

Whispering Gallery Modes are divided in quasi-TM (WGH) and quasi-TE (WGE) mode families. They are characterised by three integers m,n,l representing the electromagnetic field component variations along the azimutal, radial and axial directions respectively. In the following we only consider the WGH_{700} resonant mode.

In Whispering Gallery Mode configuration there is no field in the central part of the crystal so the later can be drilled along its axis. A simple metallic screw is then used to ensure a rigid mechanical mounting in the cavity. The metallic enclosure is made of gold plated OFHC copper 5mm thick. Two electrical probes parallel to the cavity axis and fixed on each flat of the cavity ensure the coupling with the external circuit.

Mode selection

One of the major problem with WGM resonator is its high density of mode. The main resonance is generally perturbed by a large number of low-Q spurious modes corresponding to the other high order hybrid modes of the structure. We have developed an original mode selection. In the case of a $WGH_{m,0,0}$ mode we have, in the dielectric, a succession of 2m meridian planes acting as perfect electrical conductor (PEC) surfaces. Thin metallic radial lines on the top of the sapphire disk deposited along the direction of the PEC have negligible influence on the main resonance. Conversely, the neighbouring spurious modes presenting a different azimutal variation of the electric field components are strongly affected by the metallic lines. We have in a preceding paper demonstrated the efficiency of this technology for a WGMR operating at ambient temperature. At lower temperature, the intrinsic resonator Q-factor increases rapidly, and the influence of the deposited metallic lines can no longer be neglected. Any departure from the ideal mode configuration induces extra-losses that could be detrimental especially when the intrinsic Q-factor is higher than 1.10^6 [2]. Hence a compromise between efficient mode selection and the Q-factor has been found by decreasing the number of the metallic lines. The best configuration consists in only two radial lines. In that case, the resonator can be inserted in a classical oscillator loop without the

0-7803-5400-1/99/ $ 10.00

need of additional filter. Moreover the Q-factor remains unchanged.

3 TEMPERATURE COMPENSATION

3.1Principle

As previously mentioned, the resonator frequency instabilities are the result of the intrinsic temperature sensibility of the sapphire permittivity. We opted for a configuration depicted in figure 1 initially proposed by J.Dick [3]. The sapphire resonator is now constituted by two disks separated by a small gap whose dimension is fixed by the length of an intermediate copper piece. When the temperature decreases, the gap is reduced due to the thermal contraction of the copper. Hence the permittivity mean value seen by the electrical field increases and that is just the opposite of the natural permittivity variation.
At a given temperature that can be adjusted by changing the copper piece length the two effects compensated their-self.

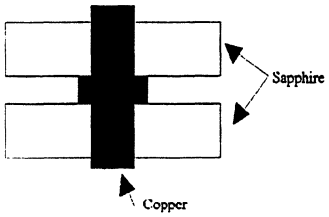


Figure1: Structure of the temperature compensated resonator.

3.2 Configuration of the temperature compensated resonator.

Two disks of HEMEX sapphire rod with 35 mm diameter and 8.75 mm height each constitute the temperature compensated sapphire resonator. The resonator operates on the $WGH_{7,0,0}$ mode at 9 GHz . On one disk we have deposited two metallic lines. The spurious modes are eliminated for a frequency range of 300 MHz and the Q-factor is about 2 millions at 87K. The length of the copper piece was calculated to obtain the inversion point near 87K. On figure 2 we can see the evolution of the temperature compensated resonator frequency versus temperature. This plot shows a turn-over temperature of T_0=87.724K. A quadratic approximation in the vicinity of the peak gives:

$$\frac{\Delta f}{f} = 1.04\times10^{-7}(T-T_0)^2 \qquad (1)$$

A residual linear thermal coefficient due to imperfect temperature adjustment $\partial T=T-T_0$ can be derived from the slope of the curve as

$$\frac{1}{f}\frac{\Delta f}{\Delta T} = 2.08\times10^{-7}\partial T \qquad (2)$$

So if we stabilise the resonator just near the turning point (∂T=1mK), with a temperature stability of 100µK, a relative frequency stability of 1.10^{-13} could be obtain, if we assume that the frequency instability is only due to the temperature fluctuations.

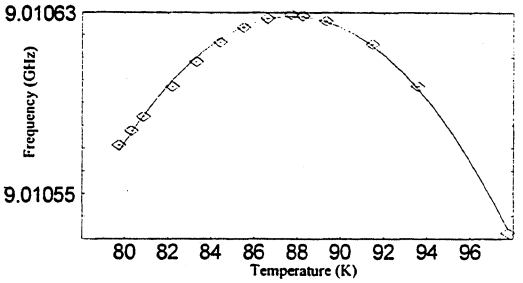


Figure 2: Frequency variation of the compensated resonator versus temperature.

4 TEMPERATURE REGULATION

4.1 Sensor

In order to obtain long term frequency stability we have developed a thermal regulator based on a thermosensitive quartz sensor [4]. The temperature sensitive element is a 28 MHz LC cut quartz resonator available in TO5 package. The 28 MHz oscillator circuit fixed directly on the resonator package was specially designed to allow operation at room temperature as well as 77K. The measured temperature sensitivity of the quartz oscillator frequency is about 1200 Hz/K at room temperature and 925 Hz/K at 77 K.

4.2 Thermal regulator.

The experimental device is shown in figure 3. It consists of a 28 MHz quartz oscillator as sensor, a high sensitivity counter with 4.10^{-9} of relative frequency resolution and a DC power supply. The counter is used to measure the beat note between a 25 MHz reference and the 28 MHz sensor signal. Combining the counter and the sensor sensibility, we find that the temperature resolution measurement is in the order of 12µK. The digital information is treated by a numeric PID that commands the electrical power delivered to the heaters.

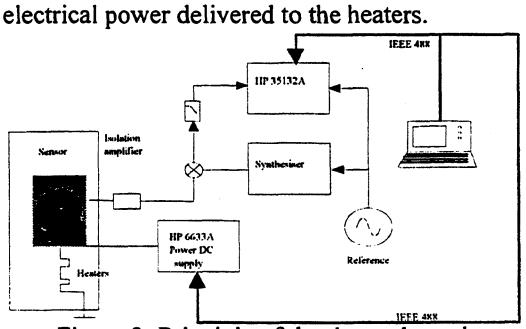


Figure 3: Principle of the thermal regulator

4.3 Performances.

To test our thermal regulator we have controlled the temperature of another LC cut quartz resonator oscillator, at room temperature and at the liquid nitrogen temperature. Phase noise measurements have been carried out with this oscillator. Figure 4 and 5 show the phase noise

measurement respectively at room temperature and at 105 K. The phase noise measurement present a f^{-3} dependence for carrier frequency between 0.01 Hz and 10 kHz. We obtain $S_\varphi(f=1Hz)$=-45 dB.rad^2/Hz at room temperature (figure 4) and $S_\varphi(f=1Hz)$=-62 dB.rad^2/Hz at 105 K (figure 5). In term of frequency stability, these results lead to flicker floor equal to 2.10^{-10} at the ambient temperature and 3.10^{-11} at 105 K. Taking into account of the oscillator frequency sensitivities the weakest temperature fluctuation that can be detected is of the order of 3 µK at room temperature and 0.7µK at 105 K.

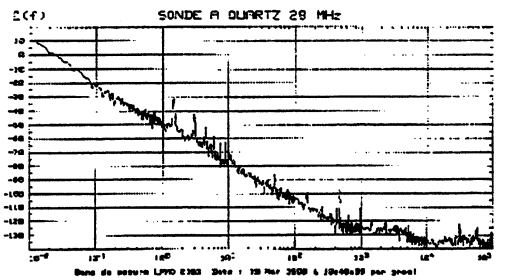


Figure 4: Phase noise of a LC cut quartz oscillator stabilised at 300K

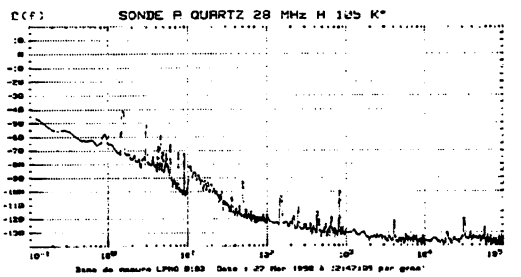


Figure 5: phase noise of a LC cut quartz oscillator stabilised at 105K

5 TEMPERATURE CONTROLLED SAPPHIRE OSCILLATOR AT ROOM TEMPERATURE.

The quartz sensor and the heaters are anchored in an intermediate flange supporting the sapphire resonator and linked to the cold source which can be the ambient environment or a liquid nitrogen filled reservoir. The ensemble is mounted in a vacuum chamber to insure a proper thermal insulation. Two 20 cm long coaxial cables enable the connection with a classical microwave oscillator loop circuit which is placed outside the vacuum enclosure (Figure 6). We measured the frequency stability of the microwave oscillator with a microwave synthesiser referenced to a rubidium frequency standard as reference. Figure 7 shows the standard deviation $\sigma_y(\tau)$ of the relative frequency fluctuations evaluated with different integration times, i.e 0.1, 1, and 10s.

For low integration times ($\tau \leq 2s$) we reached the flicker floor of 4.10^{-11}, which is consistent with phase noise measurements carried out with this oscillator: $S_\varphi(f_m)$= 10-30 Log(f) dBrad^2/Hz. For long integration times the frequency of the oscillator signal is comparable with the results obtained with another type of temperature controller [5]. Moreover, if we assume that the observed frequency fluctuations are only due to temperature variations we could deduced that the sapphire resonator temperature instability is less than 14µK for τ<1000s, which is consistent with the limitation that we have calculated for our thermal regulator.

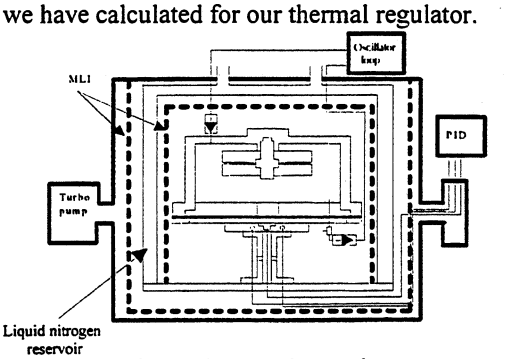


Figure 6: Experimental set up.

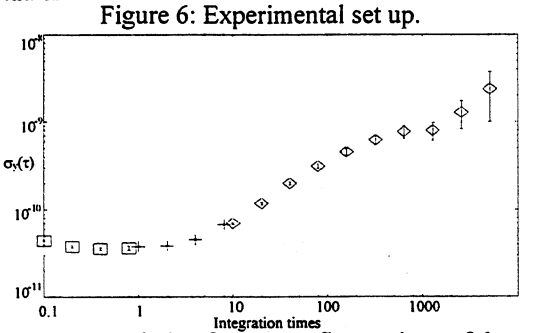


Figure 7: relative frequency fluctuations of the oscillator at 9.4GHz

6 TEMPERATURE CONTROLLED CONPENSATED SAPPHIRE RESONATOR OSCILLATOR AT 87 K.

6.1 Oscillator circuit loop.

The loop is shown in figure 8.

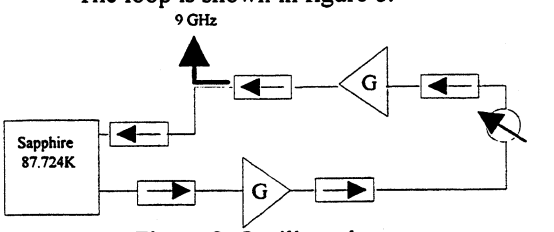


Figure 8: Oscillator loop

It consists in five isolators, two amplifiers (to obtain sufficient gain in the loop), one 10dB coupler and one mechanical phase shifter. The first amplifier was developed at the IRCOM [6]. It is centred at 9 GHz with a power gain of 13dB. The second one is a commercial amplifier 16 dB power gain. Isolators insure the adaptation at each amplifiers ports and minimise the reflection of the signal in the loop.

The external loop circuit is temperature controlled with a commercial regulator to eliminate the effect of the ambient temperature fluctuations.

6.3 Experimental result

We measured the frequency stability of the

microwave oscillator with a microwave synthesiser referenced to a cesium clock. This reference oscillator was mixed with the sapphire resonator oscillator creating a 2 kHz beat frequency. A frequency counter interfaced to a computer was used to calculate the frequency deviation of the 2 kHz beat note. The figure 9 shows the standard deviation $\sigma_y(\tau)$ of the relative frequency fluctuations evaluated with different integration times, i.e 1 and 10s.

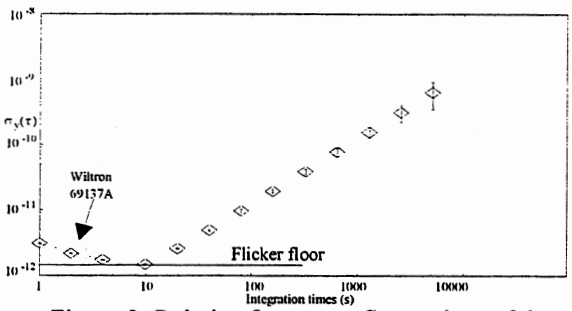


Figure 9: Relative frequency fluctuations of the temperature compensated sapphire oscillator at 9 GHz

For low integration times the measurements are limited by the synthesiser instabilities. The obtained results at 10 s, i.e $\sigma_y(\tau)=1{,}6.10^{-12}$, is consistent with the predicted flicker floor with a loaded Q-factor of 2.10^6 and an amplifier flicker phase noise equal to – 110 dB.rad^2/Hz. Nevertheless at long integration times the oscillator frequency shows a large drift of the order of 1.10^{-8} per day.

Such a frequency drift can not be attributed to the temperature regulator. Indeed a platinum sensor with typical drift of 10mK/year was placed just near the quartz oscillator sensor. During a period of two weeks any temperature variation has been detected within the resolution measurement of 1mk.

The observed frequency drift could be due to:

- *Long term phase variation* along the coaxial cables linking the cryogenic resonator to the oscillator loop. In the next month a Pound discriminator will be implemented to control the phase lag along the oscillator loop between the two-resonator ports [7].

- *Long term resonator frequency ageing.* Such a resonator ageing could be due to residual stress existing in the resonator structure itself or to long relaxing temperature gradients between the different parts of the structure.

7 CONCLUSION

We have implemented a thermal regulation, based on a quartz resonator oscillator sensor, which can be used at room temperature and also in a cryogenic environment. With this thermal regulator we have controlled a non-compensated sapphire resonator oscillator at room temperature and obtained good frequency stability. We have also developed a temperature compensated sapphire resonator oscillator at the liquid nitrogen temperature. The flicker floor of this oscillator is $1{,}6.10^{-12}$ consistent with the resonator Q-factor and the loop amplifier phase noise. Nevertheless, a large frequency drift has been observed. The exact cause of this detrimental linear frequency variation is not today well understood. Additional experiments are under development to determine the process responsible of this frequency drift.

8 ACKNOLEDGEMENTS

This work is supported by the Centre National d'Etudes Spatiales within the framework of the PHARAO project. We also thank M.M.Chaubet, C.Maréchal and F.Gonzales from CNES.

9 REFERENCES

[1] V.B Braginski, V.S.H Chenko; K.S Bagdassarov. 'Experimental observation of fundamental microwave absorption in high-quality dielectric crystal'. Physics. Letters A, Vol. 120, N°6, March 1987.

[2] O.DiMonaco, W.Daniau, H. Merigoux, Y. Kersalé, V. Giordano. 'Effect of C-axis misalignment on sapphire disk resonator performances'. Proc. 1997. 11 Th European. Frequency and Time Forum, Neuchâtel, Switzerland, 4 – 6 March 1997, pp 35- 38.

[3] D.G Santiago; R.T Wang and J.Dick. 'Improved performance of a temperature compensated LN_2 cooled sapphire oscillator.' Proc. 1995 IEEE Int. Frequency Control Symposium, pp 397-400.

[4]Y.Kersalé; F.Lardet-Vieudrin; M.Chaubet and V.Giordano. 'Thermal stabilisation of high-Q sapphire microwave resonator using thermosensitive quartz sensor.' Elect. Lett., 1998, 34, (8), pp783-784.

[5] M.E Tobar; E.N Ivanov; R.A Wood; J.H Searls and A.G Mann: 'Low noise 9 GHz sapphire resonator for ultra-low noise phase noise.' Proc. 1992 IEEE Int. Frequency Control Symposium, Hersley, PA, USA, 27-29 May 1992, pp176-182.

[6] E.Vaury, J.C Nallatamby, M.Pringent, J.Obregon, V.Giordano, J.Groslambert, M.Camiade, O.Loppis, F.Gonzales, M.Chaubet. 'A new method for the design of ultra low noise oscillators' Proc. 1999. 13 Th European. Frequency and Time Forum, Besançon, France, 13-16 May 1999.

[7] A.N Luiten, A.G.Mann, N.J.McDonald and D.G Blair. 'Latest results of the U.W.A cryogenic sapphire oscillator'. Proc. 1995 IEEE Int. Frequency Control Symposium, pp 433-437.

HIGH-Q WHISPERING GALLERY TRAVELLING WAVE RESONATORS FOR OSCILLATOR FREQUENCY STABILISATION

Michael E. Tobar, Eugene N. Ivanov
Department of Physics, University of Western Australia, Nedlands, WA, Australia

Pierre Blondy, Dominique Cros, Pierre Guillon
IRCOM, Faculte Des Sciences, 123 Avenue A. Thomas, 87060 Limoges Cedex, France

ABSTRACT

Usually a frequency stabilised standing wave resonator-oscillator that incorporates a resonator as a frequency discriminator, requires a circulator to separate the input and reflected wave. A ferrite circulator is a noisy device and can limit the phase noise or frequency stability. Moreover, we show that the noise in a circulator varies, and detailed low noise measurements are necessary to choose an appropriate quite circulator. Thus, by realising a configuration that does not require a circulator an improvement in performance and reliability can be obtained. A solution to this problem is to design a high-Q Whispering Gallery Travelling Wave WGTW resonator. This device naturally separates the input and reflected wave in the same way as a ring cavity at optical frequencies without degrading the frequency discrimination. Q-factor measurements of a WGTW sapphire resonator are presented, along with a derivation of critical parameters to maximise the frequency discrimination. Measurements of noise in ferrite circulators and isolators have also been made which we follow with a discussion on oscillator design.

1. INTRODUCTION

A Whispering Gallery (WG) resonance in a cylindrical dielectric is a higher order mode of large azimuthal mode number, m. It is informative to describes these modes as $WGE_{m,n,p}$ (quasi-TE) or $WGH_{m,n,p}$ (quasi-TM), where n and p are the radial and axial mode number respectively. The mode energy density exists near the cylindrical boundary and can be further decomposed into two travelling waves of clockwise and counter clockwise direction. When utilising a standard magnetic or electric field probe both directions are excited and a standing wave resonance is formed. If the excitation method is made unidirectional, it is possible to excite only one direction and form a travelling wave resonance.

At the University of Limoges IRCOM a free running oscillator based on a WGTW dielectric resonator excited by parallel microstrip has been developed [1]. This work solved the basic equations for the WGTW resonator. However, the equivalent standing wave resonator parameters with respect to the travelling wave parameters were not derived. In this paper we equate the reflection coefficients from the standing wave and travelling wave resonators and show how the coupling and unloaded and loaded Q-factor may be defined with respect to the travelling wave parameters. Based on these results the necessary design parameters are defined to maximise the frequency discrimination of the WGTW resonator. Following this we report on results of measurements of a high-Q sapphire cavity excited by parallel microstrip line.

Because of the unidirectionality of the WGTW resonator, a ferrite circulator is not required to separate the input and reflected waves to and from the resonator. We present phase noise measurements of some circulators under varying input power conditions, and show that the reliability and performance of frequency stabilised oscillators can be improved by utilising WGTW resonators.

2. WHISPERING GALLERY TRAVELLING WAVE RESONATORS

2.1 Single Port Cavity

A single microstrip line placed adjacent to a dielectric resonator will excite WG travelling waves in the same direction as the excitation. In contrast a WG standing wave resonator has both clockwise and anti-clockwise directions excited which are superimposed to create a single standing wave. The standing wave resonator has the same port as the input and output and to separate these waves a circulator is required as shown in figure 1. This is not true for the travelling wave resonator. In this case part of the input wave is reflected in the on going direction of the microstrip, while the other part couples into the resonator with a transmission coefficient of k, and continues to travels around the resonator in a similar way to an optical ring cavity. After one revolution, part of the travelling wave is transmitted out of the resonator with transmission coefficient k, while the remaining wave continues for another revolution, etc.

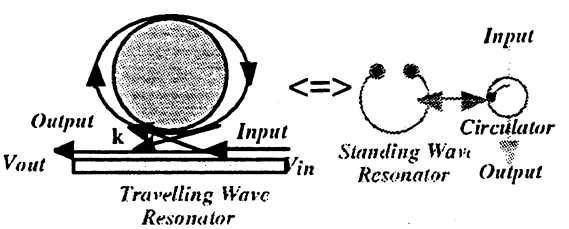


Figure 1. *A travelling wave resonator excited by an adjacent microstrip line is schematically equivalent to a single port standing wave resonator with a circulator at the input. The resulting resonators are two port devices that separate the input and output wave.*

Assuming that the transmission coefficient into and out of the resonator is k, and in one revolution of the resonator there is α attenuation and ϕ phase shift, then the complex reflection coefficient (output complex amplitude with respect to the input) can be calculated by

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

summing the series of travelling wave revolutions (which is convergent). Following this it may be shown that the complex reflection coefficient is given by;

$$\Gamma = \frac{V_{out}}{V_{in}} = \frac{\sqrt{1-k^2} - e^{-(\alpha+j\phi)}}{1-\sqrt{1-k^2}\,e^{-(\alpha+j\phi)}} \quad (1)$$

Also, the complex amplitude of the wave circulating in the resonator may be written as;

$$T_c = \frac{k\,e^{-(\alpha+j\phi)}}{1-\sqrt{1-k^2}\,e^{-(\alpha+j\phi)}} \quad (2)$$

By taking the absolute value squared and calculating the half power points, the loaded bandwidth in radians of the resonator can be calculated to be;

$$\theta_{BW} = \frac{2\left(1-\sqrt{1-k^2}\,e^{-\alpha}\right)}{\sqrt{\sqrt{1-k^2}\,e^{-\alpha}}} \quad (3)$$

The quality factor may then be calculated from the ratio of the bandwidth to the value of the resonance in radians. A resonance occurs every cyclic value of 2π, ie. $\phi=2\pi m$ where m is an integer and equal to the order of the resonance (or azimuthal mode number since the wave travels around the perimeter of the cylinder). Thus the loaded Q-factor may be written as;

$$Q_L = \frac{m\pi\sqrt{\sqrt{1-k^2}\,e^{-\alpha}}}{\left(1-\sqrt{1-k^2}\,e^{-\alpha}\right)} \quad (4)$$

The unloaded Q-factor may be calculated by setting k=0. If we then assume α is small, to first order we calculate the unloaded Q to be;

$$Q_0 = \frac{m\pi}{\alpha} = \frac{\pi}{\gamma\lambda} \quad (5)$$

Here γ is the attenuation in nepers/m and λ is the wavelength. The second part of (5) is the same equation for a microstrip resonator as expected. If the loss tangent is constant, then (5) implies $\gamma\lambda$ should be constant. This means that as λ decreases (frequency increases) γ will increase as the energy of the WG mode is pushed into a smaller effective volume of space near the perimeter of the resonator. Because of the smaller volume the energy density will increase and thus the effective cross sectional area normal to the travelling wave direction will be reduced causing an increase in attenuation per metre.

By combining (4) and (5) then by taking the series expansion with respect to k and α (both should be small in a high-Q system), the relation between Q_L and Q_0 can be calculated to be;

$$Q_L = \frac{Q_0}{\left(1+k^2/2\alpha\right)} \quad (6)$$

Thus, the equivalent coupling coefficient to the cavity is given by;

$$\beta = k^2/2\alpha \quad (7)$$

From (7), to perfectly impedance match a cavity the transmission coefficient must be set as $k=(2\alpha)^{1/2}$. Figure 2 shows the reflection coefficient given by (1) when m=6, Q=200,000 and k=0.0137. These are the conditions for a unity coupled cavity in a $WGH_{6,00}$ 10 GHz mode in a 3.1 cm diameter sapphire resonator cooled to 0^0 Celsius.

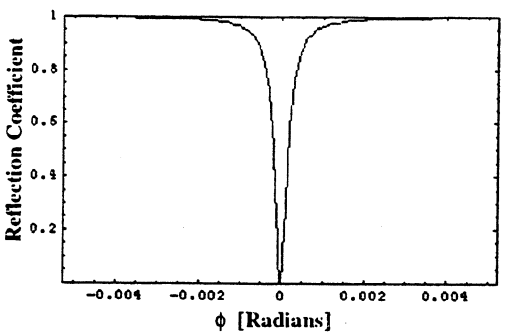


Figure 2. Γ versus ϕ, for a unity coupled WGTW sapphire resonator operating on the $WGH_{6,00}$ mode.

2.2 Two Port Cavity

To realise the travelling wave equivalent of a two port standing wave resonator, two transmission lines may be placed parallel with the resonator placed in between. In this case two transmission coefficients may be defined as shown in figure 3.

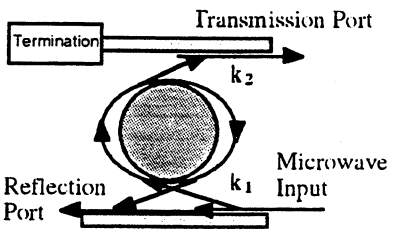


Figure 3. Schematic of a WGTW resonator with reflection and transmission port.

The complex value of the reflection and transmission port can be shown to be given by;

$$\Gamma = \frac{\sqrt{1-k_1^2} - \sqrt{1-k_2^2}\,e^{-(\alpha+j\phi)}}{1-\sqrt{1-k_1^2}\sqrt{1-k_2^2}\,e^{-(\alpha+j\phi)}} \quad (8)$$

$$T = \frac{k_1 k_2\, e^{-\left(\alpha+j\phi/2\right)}}{1-\sqrt{1-k_1^2}\sqrt{1-k_2^2}\,e^{-(\alpha+j\phi)}} \quad (9)$$

By following a similar procedure as before it may be shown that the relation between coupling and loaded and unloaded Q is given by:

$$Q_L = \frac{Q_0}{\left(1 + k_1^2/2\alpha + k_2^2/2\alpha\right)} \qquad (10)$$

where:

$$\beta_i = k_i^2/2\alpha \; ; i = 1 \text{ or } 2 \qquad (11)$$

The preceding equations present a method to maximise the frequency discrimination of a WGTW Resonator. For example, a single port resonator has a maximum discrimination when β=1 (impedance matched), if the unloaded Q is known, then from (5) α may be calculated. Substituting β and α into (7) gives the value of k necessary to impedance match the cavity, and may be designed for using 3D finite element software[1].

3. HIGH-Q SAPPHIRE WGTW RESONATOR

A sapphire resonator of hight 7.7 mm and radius 15.5 mm was excited by microstrip lines mounted in a teflon substrate inside a Cu cavity as shown in figure 4.

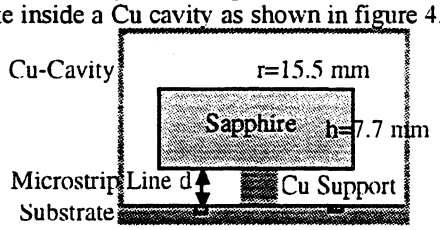


Figure 4. Sapphire cavity excited by microstrip lines.

To determine the Q-factor, the cavity was measured in reflection from one port as shown in figure 5. Two different swept frequency methods were used. The first method implemented a fast sweep, which excites the cavity with an effective impulse. The signal detected by the microwave detector rings down at the rate governed by the amplitude decay time. By measuring the ring down time the Q-factor can be calculated. A typical wave form is shown in figure 6.

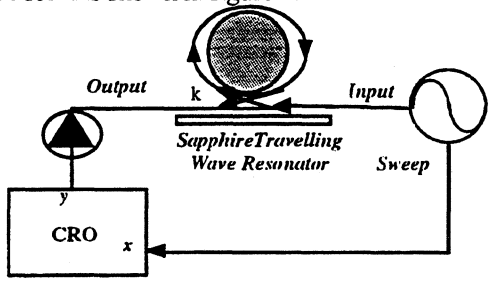


Figure 5. Measurement set up for determining the Q-factor using swept frequency methods.

The second method utilises a slow sweep, which is slow enough that the signal source traces the frequency response with no significant transients. In this case the reflection coefficient is traced out similar to that shown in figure 2. From this plot the frequency and half power bandwidth may be measured and the Q-factor calculated.

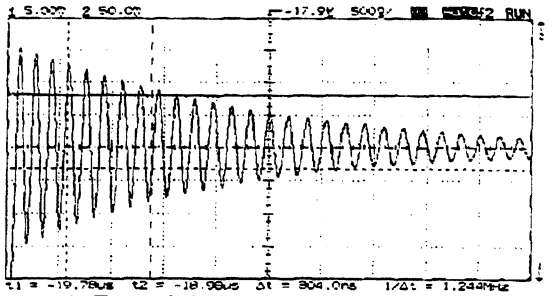


Figure 6. Typical fast swept transient response of the high-Q sapphire WGTW resonator.

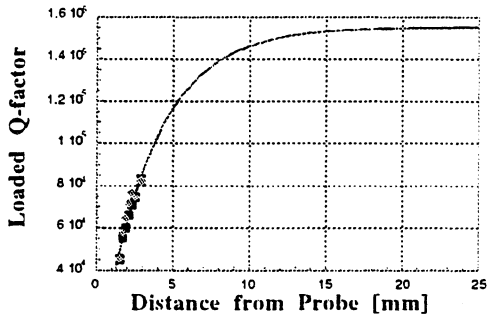


Figure 7. Loaded Q-factor versus probe distance of the $WGH_{6,0,0}$ mode. The results fit well to an exponential. At large distances the coupling is small and the loaded Q value approaches the unloaded Q-value of 155,000. Dark squares are the values calculated using the slow sweep method, while the light diamonds are the values calculated with the fast sweep method.

The waveforms observed were very symmetric and no feed through or VSWR effects could be noticed. This suggests that the technique works well without a circulator. To use these methods with a standing wave resonator a circulator is essential.

5. NOISE IN FERRITE CIRCULATORS AND ISOLATORS

To understand the noise from a ferrite circulator in a standing wave frequency discriminator, we carried out detailed measurements utilising the interferometric noise measurement system[2], where the DUT was a ferrite circulator/isolator. The phase noise model of a typical isolator was shown to be $\mathcal{L}_{\varphi}^{isol}(f) \approx -150\text{-}12\log_{10}(f)$ dBc/Hz. This corresponds to -186 dBc/Hz at 1 kHz.

Figure 8 show the dependence on power of a noisy circulator. It is evident that close to a square law power is exhibited in the noisy devices. We are not sure of the mechanism of noise in the circulators. Also, we are not sure why some devices exhibit excess noise and others do not. However, we have established that the excess noise is not due to AM to PM conversions in the circulators. This excess noise is intrinsic and induced by the incident signal. Further investigations are warranted to understand these effects. Especially since

even the quietest circulators still limit the spectral performance of an interferometric standing wave resonator-oscillator below 200 Hz Fourier frequency[3]

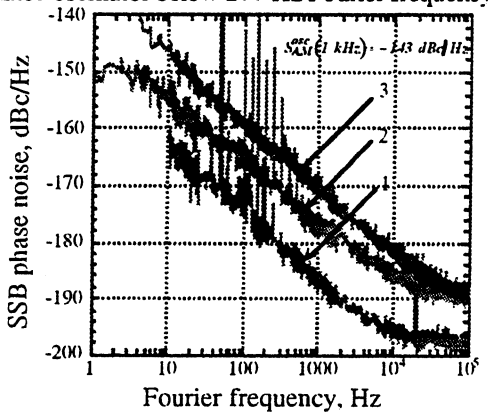


Figure 8. The spectral density of phase fluctuations in a noisy microwave circulator, at different power levels. 1. The Intrinsic phase noise of a quiet microwave circulator. 2. Noise of a noisy circulator with 190 mW input power. 3. Noise of a noisy circulator with 480 mW input power.

The alternative approach is to design an oscillator without a circulator utilising the WGTW resonator. In the following section some designs based on the WGTW resonator are presented.

6. LOW-NOISE HIGH-STABILITY WGTW RESONATOR-OSCILLATORS

A realisation of an oscillator (patented) with an interferometric WGTW frequency discriminator is shown in figure 9. The loop oscillator is maintained by positive feedback from the transmission port (port 2) to the input port (port 1). Signals from the reflection (port 3) and transmission port are combined in a hybrid coupler to form an interferometer. The transmission port also drives the LO of the mixer while the signal from the Dark Port (DP) is amplified (flicker free) and fed to the mixer RF port to create a low noise phase detector. The Bright Port (BP) may be used as another output port, or as an alternative LO drive for the mixer. If the cavity was perfectly unity coupled, the reflected port could be directly fed to the RF port to provide the DP without the need for an interferometer. However, in practice it is not possible to set the coupling precisely to unity to maintain the amplifier in the small signal flicker free regime. Furthermore, the optimum condition to maximise the sensitivity of the frequency/phase detection is to unity couple the input/reflection ports. Consequently, we are implementing 3D FE software to design a coupling of unity for the WGTW sapphire resonator. The coupling of the transmission port need not be large due to the loop amplifier in the circuit.

In fig. 10, a possible variation of a Pound stabilised system that makes use of WGTW resonator is illustrated. This system is commonly used for highly stable secondary frequency standards. Likewise, the circulator noise and unity coupling are very important for the performance of this type of oscillator.

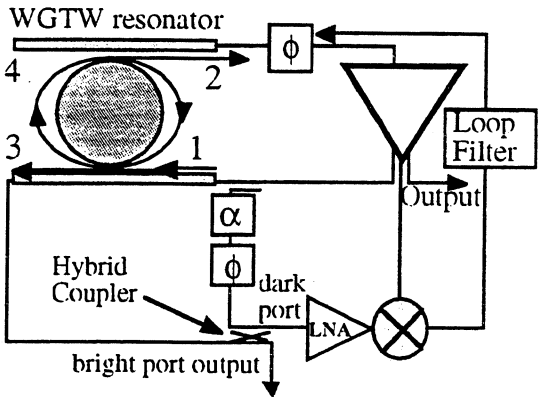


Figure 9. WGTW loop oscillator with interferometric noise cancellation. (Patented)

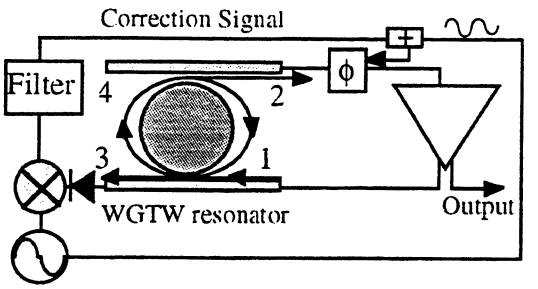


Figure 10. WGTW loop oscillator with Pound frequency stabilisation.

7. CONCLUSION

We have shown that a high-Q sapphire resonator maybe excited by microstrip line to create WGTW resonances. These resonances are unidirectional and do not require a ferrite circulator to separate the input and reflected waves. To maximise the noise cancellation the necessary parameters to optimise the sensitivity to frequency fluctuations of the WGTW resonator were defined with respect to the effective standing wave parameters. Low noise oscillator circuits based on WGTW sapphire resonators were presented.

8. ACKNOWLEDGMENT

This work was supported by the Australian Research Council and the Visiting Professorship scheme at the University of Limoges.

9. REFERENCES

[1] F. Nigon, "Conception de circuits planaires microondes à modes de galerie des résonateurs diélectriques," in *IRCOM*. Limoges: University of Limoges, 1997.

[2] E. N. Ivanov, M. E. Tobar, and R. A. Woode, "Microwave interferometry: Application to precision measurements and noise reduction techniques," *IEEE Trans. on UFFC*, vol. 45, pp. 1526-1536, 1998.

[3] E. N. Ivanov, M. E. Tobar, and R. A. Woode, "Applications of Interferometric Signal Processing to Phase Noise Reduction in Microwave Oscillators," *IEEE Trans. on MTT*, vol. 46, pp. 1537-45, 1998.

1999 Joint Meeting EFTF - IEEE IFCS

High – Q SiO_2 Whispering Gallery Mode Resonator

V. Giordano , R. Barhaila, D. Cros*, G. Duchiron*

Laboratoire de physique et métrologie des oscillateurs
32, Av. de l'Observatoire F-25044 Besançon – France
e-mail giordano@lpmo.univ-fcomte.fr

*Institut de recherche en communication Optiques et Microondes
123, rue A.Thomas F-87060 Limoges-France
e-mail cros@ircom.unilim.fr

1- Introduction

During the last five years, low-loss single crystals have been extensively studied in the aim to develop high-Q microwave resonators. Attention has been paid principally to Sapphire (Al_2O_3) single crystal that presents extremely low losses and leads to the achievement of high spectral purity X-band sources. In the case of sapphire, High-Q resonance are obtained by using a special mode configuration - i.e. Whispering gallery mode – allowing to confine the electromagnetic energy inside the single crystal itself despite the relatively low value of the sapphire permittivity.
Pure quartz (SiO_2) single crystal is known as a good microwave dielectric presenting low dielectric losses and a slight permittivity anisotropy. Moreover, due to the extensive use of quartz in piezoelectric transducers and resonators, high purity, well orientated and low cost single crystals are easily available. Then it was tempting to apply the Whispering Gallery Mode technology to design a SiO_2 microwave resonator. We present in this paper preliminary results obtained with a quartz single crystal microwave resonator in whispering gallery mode configuration. Q factor of the order of 50 000 has been obtained at 16GHz at 300K. Theoretical predictions based on 3D Finite Elements model are compared with experimental results at ambient and liquid nitrogen temperatures.

2- Quartz permittivity Measurement

A Courtney type cavity mounted in a closed cycle cryocooler has been used to measure the evolution of the two permittivity components of the SiO_2 crystal. The resonance frequencies of the TE_{011} and the TM_{011} modes were recorded between 300K and 12K. The dispersion equations for these modes enable then to extract ε_t and ε_z. The figure 1 shows the variation of ε_t and ε_z versus the temperature.

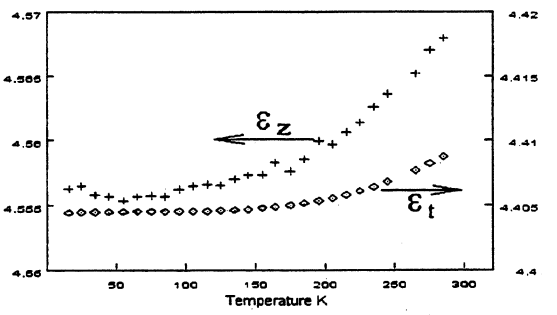


Fig 1 : SiO_2 permittivity components vs the temperature

From the resonator material characteristics, we can simply evaluate the temperature sensitivity of the different modes of the structure. Considering whispering gallery modes with high azimutal number, we have [1].

$$\text{WGH-mode} \quad \frac{1}{f}\frac{\Delta f}{\Delta T} \approx -\frac{1}{2}\frac{1}{\varepsilon_z}\frac{\Delta \varepsilon_z}{\Delta T} - \alpha_\perp$$

$$\text{WGE-mode:} \quad \frac{1}{f}\frac{\Delta f}{\Delta T} \approx -\frac{1}{2}\frac{1}{\varepsilon_t}\frac{\Delta \varepsilon_t}{\Delta T} - \alpha_\perp$$

Where $\alpha_\perp$ is the thermal expansion coefficient in the radial direction.
The following table gives the material characteristics deduced from the experimental data as well as the calculated temperature frequency sensitivities.

T	300K	77K
$\frac{1}{\varepsilon_z}\frac{\Delta \varepsilon_z}{\Delta T}$	$22.\ 10^{-6}$	$<1.10^{-6}$
$\frac{1}{\varepsilon_t}\frac{\Delta \varepsilon_t}{\Delta T}$	$14.\ 10^{-6}$	$<1.\ 10^{-6}$
$\alpha_\perp$[2]	$13{,}6.\ 10^{-6}$	$5{,}6.\ 10^{-6}$
$\frac{1}{f}\frac{\Delta f}{\Delta T}$ (WGH)	$-25.\ 10^{-6}$	$-6.\ 10^{-6}$
$\frac{1}{f}\frac{\Delta f}{\Delta T}$ (WGE)	$-21.\ 10^{-6}$	$-6.\ 10^{-6}$

Table1. Material characteristics and temperature frequency sensitivities.

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

It should be noted that conversely to the sapphire resonator the temperature frequency sensitivities are mainly due to the thermal expansion of the resonator material.

3-Theoritical prediction

The electromagnetic tool used to determine resonant frequencies of whispering gallery modes resonators is based on finite element method developed at IRCOM. Nedelec ' second order polynomial interpolation is used and does not generate any spurious solutions, which exist for the finite elements methods with Lagrange's polynomial elements [3].
In this application, the model of resonant structure under analysis is only composed of a resonator placed in a metallic cavity. We don't take into account the system used to hold on the resonator in the cavity as shown on figure 2 and 3. In facts for whispering gallery modes, energy is confined at the periphery of the resonator and resonant frequencies is not affected by the fixation screws. To discretize the structure we have used a mesh with 600 elements for both cavities. The permittivity anisotropic effects of the material is take into account. The material losses can be considered but for this application there are a little influence on resonant frequencies. For the structures analysed in this application, computation time to obtain resonant frequencies is very short and less than 15 sec.

4- Technical features of the SiO_2 resonators

The dielectric resonator consists in a SiO_2 disk 40mm diameter and 10mm thick. The crystal C- axis and the geometrical axis are aligned within 0,5°. A 5mm brass screw passing through the SiO_2 resonator enables a rigid mechanical mounting in the metallic cavity. The latter is made with gold plated OFHC copper. Two types of metallic cavity have been tested differing from their coupling geometry.

a- Cavity 1

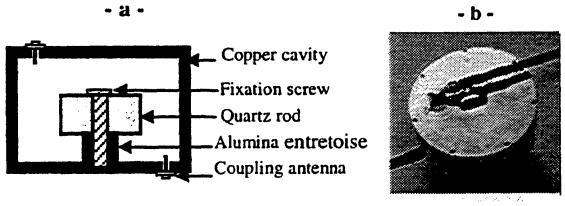


Fig 2 : The cavity used to excite WGH modes

The inner dimensions are 72mm in diameter and 18mm high. The coupling with the external circuit is insured by two electric probes (antenna) parallel to the cylindrical axis.
This cavity is suitable to excite preferentially the quasi TM or WGH modes. Moreover, the relative angular position of the probes could be easily adjusted by rotating one of the cavity plate.
Such a mechanical adjustment offers a supplementary degree of freedom to optimize the main resonance coupling and the spurious mode rejection.

b-Cavity 2

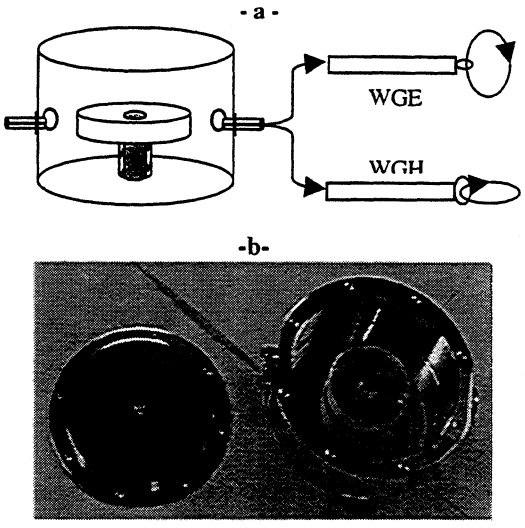


Fig 3 : The cavity 2 enables to observe both the WGH and the WGE mode families

The inner dimensions are 80mm diameter and 20mm height.
In this cavity, two radial magnetic probes (loops) are used to excite both the quasi-TE or quasi- TM modes. The mode family is determined by the position of the loop plane which is for example perpendicular to the cavity axis for WGE mode excitation.
Contrary to the previous one, this configuration does not enable angular adjustment of the coupling probes. In these condition, the spurious resonance affecting the main mode are difficult to attenuate.

5- Experimental result at the ambient temperature.

Table 1 compares the theorical frequencies with the measured one's for the cavity 1 excited on the WGH modes.

m	f th (GHz)	f exp (GHz)
6	11,301	11,19
7	12,559	12,436
8	13,779	13,640
9	14,976	14,821
10	16,161	15,99

Table 2. Theoretical and experimental frequencies of $WGH_{n,0,0}$ modes

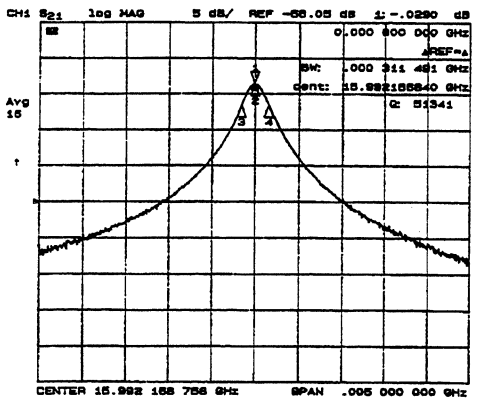


Fig 4 : Transmission spectrum around the $WGH_{10.0.0}$ resonance at 300K.

As an example, the figure 4 shows the transmission spectrum of the resonator around the $WGH_{10.0.0}$ resonance at 16GHz.
The coupling coefficient at each port are adjusted to a low value ($\beta<0,1$).
The unloaded quality factor is assumed equal to the measured value, i.e. 50 000 Such a value is about two times lower than the Q-factor obtained with a sapphire resonator[4].

Some measurements have been conducted with the second cavity for which WGH and WGE have been observed.
Experimental results and theoretical predictions are given in table2.

F_{WGH} theo (GHz)	F_{WGH} exp (GHz)	F_{WGE} theo (GHz)	F_{WGE} exp (GHz)
12,493	12,423	13,116	13,040
13,751	13,643	14,356	14,301
14,966	14,834	15,609	15,558
16,159	16,008	16,859	16,796

Table. 3 . Theorical and experimental frequencies of $WGH_{n.0.0}$ and $WGE_{n.0.0}$ modes ($10 \geq n \geq 7$)

6- Experimental result at low temperature

The resonator cooling has been achieved thanks to the experimental set-up schematized in the figure (5).
The cavity 1 was placed in a stainless steel vacuum can that can be immersed in a large Dewar filled with liquid nitrogen. The SiO_2 resonator in thermal contact with the bottom flange of the containing can was then rapidly cooled down to about 77K. The resonator temperature was measured thanks to a platinum sensor placed on the top of the copper cavity.

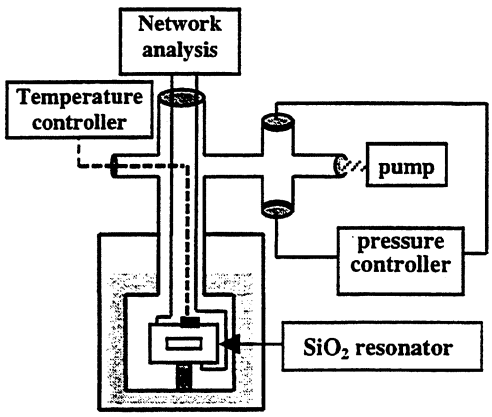


Fig 5 : Sketch of the experimental set-up

The variation of the $WGH_{10.0.0}$ mode frequency was recorded during the slow temperature rising of the resonator when the liquid nitrogen evaporated figure (6)

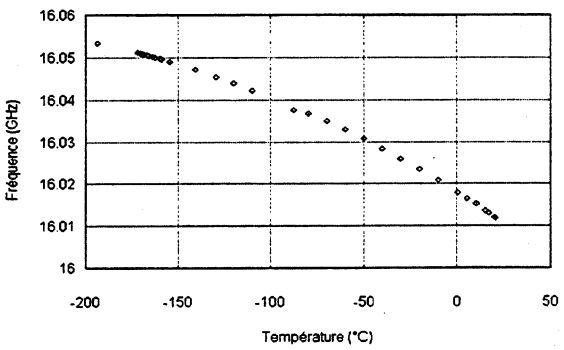


Fig 6 : $WGH_{10.0.0}$ frequency versus temperature

From these experimental data, we deduced the resonator frequency sensitivity as a function of the temperature.
We obtained :

	$\frac{1}{f}\frac{\Delta f}{\Delta T}$ (ppm/K)	
Temperature	**Quartz**	**Sapphire**
300K	-26	-70
77K	-5.7	-10

Table 4. Measured temperature frequency sensitivities

The observed frequency sensitivities agree well with the preceding predictions (see section 2).
The quality factor increases as the temperature is lowered but the improvement is weak.
Q_0 is only multiplied by 2 when the temperature reaches 77K.

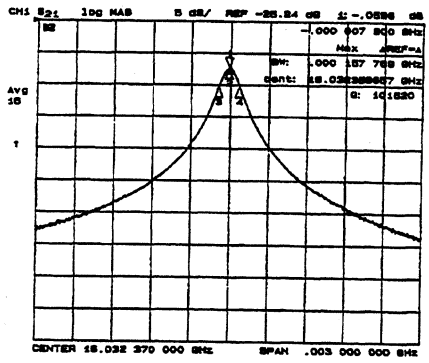


Fig 7 : Transmission spectrum around the $WGH_{10.0.0}$ resonance at 77K

7- Mode selection

As mentioned before, the whispering gallery mode is generally perturbed by low Q spurious resonances which have to be eliminated. In a preceding paper, we described an original mode selection that we have implemented on sapphire resonator [5]. The spurious mode rejection is insured by thin metallic lines deposited on top of the dielectric resonator along the 2.m radial direction corresponding to perfect electric conductor plane. We use the same technique for the SiO_2 resonator where 20 aluminium lines were deposited The line dimensions are 20µm width and 0,2µm thick.
The following figures show the effect produced by the mode selection.

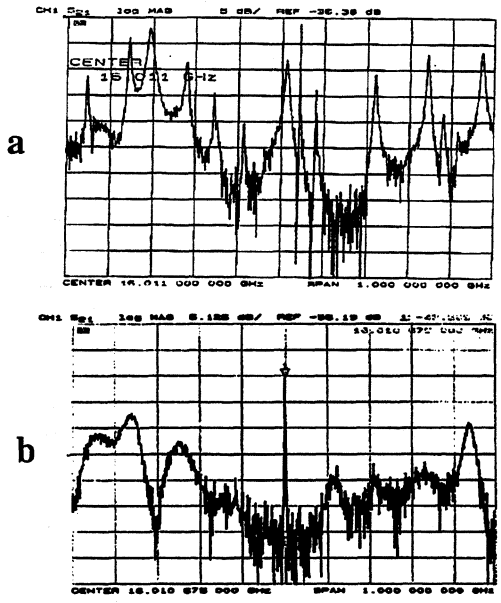


Fig 8. Transmission spectrum at 300K, **a** - without mode selection **b**- with mode selection
$Q_0 = 50\ 000$.

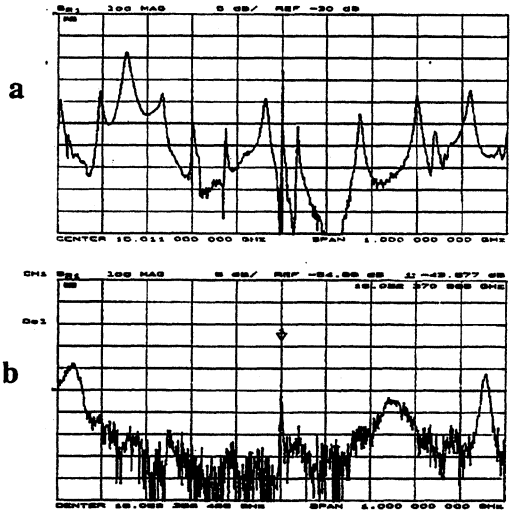


Fig 9. Transmission spectrum at 77K, **a** - without mode selection, **b** - with mode selection.$Q_0 = 100\ 000$

8- Conclusion

We have implemented two whispering gallery mode resonators in order to test the potentialities of such a structure using SiO_2 monocrystals as dielectric resonator. Permittivity measurements and a 3D finite elements model enable to calculate the resonance frequencies of the WGH and WGE modes with a good accuracy.
The SiO_2 whispering gallery mode resonator presents at 16 GHz a Q-factor of the order of 50 000 at the ambient temperature with a temperature frequency stability of – 26ppm/K. These results are comparable to these obtained with a sapphire dielectric resonator. Attempt have been mode in order to reach higher performances by cooling the resonator down to 77K, but the obtained results show only an weak increase in Q-factor.
We can conclude that the main advantages of the SiO_2 resonator are its low cost and its availability. It could be used as frequency reference operating at the ambient temperature when no extreme performances are required.

Bibliography

[1] M.E.Tobar, J.Krupka, E.N.Ivano, R.A.Woode, Dielectric frequency – temperature compensation of high quality sapphire dielectric resonators, IEEE International frequency control symposium, pp 799 – 806. 1996.
[2] American physics handbook.
[3] M. Aubourg and P. Guillon. "A mixed finite element formulation for microwave devices problems. Application to MIS structure." J. Electromagn. Waves Appl., vol. 5, n° 415, pp. 371-386, 1991.
[4] O.DiMonaco, W.Daniau, I.Lajoie, Y.Gruson, J.Groslambert and V.Giordano, Mode selection for whispering gallery mode resonator, Electronics letters 28 march 1996 vol. 32 n° 7, pp 669 – 670.

MICROWAVE WHISPERING GALLERY MODE DIELECTRIC RESONATOR OSCILLATOR

S.L.Badnikar, N.Shanmugam and V.R.K. Murthy*
Defence Electronics Research Laboratory Hyderabad, India 500 005
*Dept. of physics, IIT, Madras, Chennai,India 600 036
e-mail: root@drderl.ren.nic.in, Tel: 91-040-4440364, Fax: 91-040-4440525/4442976

ABSTRACT

This paper presents the design of a Whispering Gallery dielectric resonator and the design of a Ku band dielectric resonator oscillator using commercial packaged HEMT as an active device. A novel feedback method is employed in the structural design of DRO. A lumped element circuit model is used to represent coupling between cylindrical DR and a microstripline. As an illustrative example, the design of a 17.5 GHz DRO is discussed. The oscillator realised with this technique exhibits excellent phase noise performance, good power output and high temperature stability.

1. INTRODUCTION

Low power, fixed frequency microwave oscillators form an essential part of communication, radar, and electronic warfare systems. With rapid advancements there has been an ever increasing need for oscillators with an improved performance. This has shifted focus from the traditional fixed frequency oscillators to a new class of dielectric resonator oscillators. The use of temperature stable, high dielectric, low loss ceramic material in the form of a dielectric disc operating in TE or TM modes offers advantages of high temperature stability, low phase noise, easy integrability in Microwave Integrated Circuit (MIC) as well as that of a smaller size[1].

At higher microwave frequencies, however, Whispering Gallery (WG) mode DRs are employed in preference to conventional volume mode DRs[2,3]. When used in WG mode these cylindrical DRs become oversized. They offer distinct advantages such as good suppression of spurious mode, high quality factor and quasi insensitivity to conducting material boundaries.

In this paper a WGDR is used as a frequency selective element in the oscillator where negative resistance is induced in the active i.e. packaged HEMT by the device structure. Good oscillator power output and very good phase noise characteristics have been achieved. The temperature stability study over -20°C to +55°C has been carried out and the measured results are reported.

2.THEORY

2.1 Description of whispering gallery modes

The whispering gallery modes move essentially in the plane of the circular cross-section and most of the modal energy is confined between the resonator boundary and the inner modal caustic as well as to a small region in an axial direction. The whispering gallery mode disc dielectric resonator is thus capable of achieving high quality factor due to circumferential modes isolating electromagnetic energy into a narrow dielectric element itself and away from lossy environment.

2.2 Eigen frequencies of WGE DR modes

The general configuration of WGDR in MIC is depicted in Fig.1.

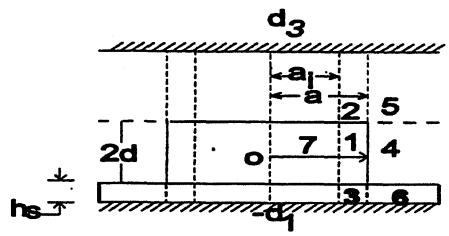


Fig 1. Composite structure of WGDR in MIC environment

As shown in this figure, a cylindrical DR of radius **a** and thickness **2d** is mounted on a grounded dielectric substrate of thickness **hs**. The relative dielectric constant of the resonator is much higher than that of the substrate.

The WG modes are classified as either $WGE_{n,m,l}$ or $WGH_{n,m,l}$. For WGE modes, the electric field is essentially transverse while for WGH modes, the electric field is essentially axial. Here integer n denotes the azimuthal variation of modes, m the radial variations and l the axial ones. The WGE modes are of interest as they couple with quasi TEM mode of microstrip. The structure under analysis for resonant frequency computation is divided into various subregions.

Most of the energy is confined in region 1. In the cross section, the field of WG modes is oscillatory between modal caustic of radius a_i and the resonator boundary. The field decays exponentially in region 2, 3 and 4. For WGE modes the longitudinal components in each subregion of the structure can be obtained by solving appropriate Helmholtz equation.

$$\nabla^2 H_{zi} + k^2 H_{zi} = 0, \qquad (1)$$

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

$i=1,2,\ldots\ldots,7$

The method used to compute em field components is derived from theory described in [1,4]. The mathematical expressions have been quoted in [5] but misprinted due to their complexity. The correct expression of the longitudinal magnetic field component in each zone of the structure is given as

$$H_{z1}=[AJ_n(k_1r)+BY_n(k_1r)]\ * \\ [C\cos\beta z+D\sin\beta z]\exp(jn\theta) \quad (2)$$

$$H_{z2}=-N[AJ_n(k_1r)+BY_n(k_1r)][\exp\{-\alpha_a(z-d)\}]\exp(jn\theta) \quad (3)$$

$$H_{z3}=-S[AJ_n(k_1r)+BY_n(k_1r)]\sinh[\alpha_s(z+d_1)]\exp(-\alpha_s d_1) \\ * \exp(jn\theta) \quad (4)$$

$$H_{z4} = GK_n(k_3r)[C\cos\beta z+D\sin\beta z]\exp(jn\theta) \quad (5)$$

$$H_{z5} = -PK_n(k_3r)\exp[-\alpha_a(z-d_3)]\exp(jn\theta) \quad (6)$$

$$H_{z6} = -T\exp(-\alpha_s d_1)K_n(k_3r)\sinh[\alpha_s(z+d_1)]\exp(jn\theta) \quad (7)$$

$$H_{z7} = EI_n(k_2r)[C\cos\beta z+D\sin\beta z]\exp(jn\theta) \quad (8)$$

$$k_o = \frac{\omega}{c} \quad (9)$$

where

$$k_1 = \sqrt{k_o^2\epsilon_{r1}-\beta^2} \quad (10)$$

$$k_2 = \sqrt{k_o^2\epsilon_{r7}-\beta^2} \quad (11)$$

$$k_3 = \sqrt{\beta^2-k_o^2\epsilon_{r4}} \quad (12)$$

$$\alpha_a^2 = k_1^2-k_o^2\epsilon_{r2} \quad (13)$$

$$\alpha_s^2 = k_1^2-k_o^2\epsilon_{r3} \quad (14)$$

$$a_i = \frac{n}{k_1} \quad (15)$$

Here J_n, Y_n are Bessel functions of the first kind and order n while I_n and K_n are modified Bessel functions of the second kind and order n. ϵ_{ri} is the dielectric constant in dielectric region i. k_0 is the free space wave number. k_i is the em mode separation constant in dielectric region i. α_a and α_s are attenuation constants in air and substrate medium. β is axial propagation constant. a_i is the radius of modal caustic. The transverse field components can be deduced from Maxwell equations.

The solution of the Maxwell equations permits us to obtain transverse components E_θ, H_r and H_θ as functions of longitudinal components H_z.

Using the above expressions for H longitudinal field components, we have a complete set of em field expressions first the tangential components at r=a and r=a_i are matched so that $E_{\theta1}=E_{\theta4}$; $H_{z1}=H_{z4}$; $E_{\theta1}=E_{\theta7}$; $H_{z1}=H_{z7}$.

We obtain a set of four equations. For a non trivial solution, the determinant of the resulting 4 X 4 matrix should vanish. This condition is equivalent to

$$\begin{vmatrix} J_n(k_1a_i) & Y_n(k_1a_i) & -I_n(k_2a_i) & 0 \\ k_2J'_n(k_1a_i) & k_2Y'_n(k_1a_i) & -k_1I'_n(k_2a_i) & 0 \\ J_n(k_1a) & Y_n(k_1a) & 0 & -K_n(k_3a) \\ k_3J'_n(k_1a) & k_3Y'_n(k_1a) & 0 & -k_1K'_n(k_3a) \end{vmatrix} = 0 \quad (16)$$

The ′ denotes the derivative with respect to the argument. In order to obtain correct resonant frequency, another set of equations is obtained by forcing axial confinement at dielectric air/substrate interface on fields through the boundary condition of continuity and their first order derivatives so that,

For z= d and -d

$$E_{r1}=E_{r2} \ ; \ \frac{\partial E_{r1}}{\partial z} = \frac{\partial E_{r2}}{\partial z}$$

$$E_{r1}=E_{r3} \ ; \ \frac{\partial E_{r1}}{\partial z} = \frac{\partial E_{r3}}{\partial z}$$

$$\begin{vmatrix} \cos(\beta d) & \sin(\beta d) & 1 & 0 \\ -\beta\sin(\beta d) & \beta\cos(\beta d) & -\alpha_a & 0 \\ \cos(-\beta d) & \sin(-\beta d) & 0 & \sinh(\alpha_s hs) \\ -\beta\sin(-\beta d) & \beta\cos(-\beta d) & 0 & \alpha_s\cosh(\alpha_s hs) \end{vmatrix} = 0 \quad (17)$$

The resonant frequencies are obtained by simultaneous solution of equations 16 and 17. For the case under consideration DR from Trans Tech 87 series was chosen with D=10.3 mm, 2d=2.0 mm ϵ_r=30 operating in $WGE_{5,0,0}$ mode. The resonant frequency was calculated as 17.29 GHz from the simultaneous solution of the above equations using special root search algorithms developed[6].

2.3 Estimation of WGDR microstripline coupling parameters

A cylindrical DR coupled to microstripline in a MIC environment is a fundamental electromagnetic field problem. There are various ways of modeling this coupling. One such coupling is parallel resonant RLC circuit whose resistance is related to the strength of coupling and susceptance is calculated from the centre frequency and the resonating circuit Q. The parallel tuned circuit parameters are given in [1].

The coupling degenerates to pure resistance at the resonant frequency. The value of R,L,C can be found out numerically or from practical measurements. In the present case investigation of coupling is done by straight forward measurement technique. The network analyzer and sweep oscillator was used as a test setup for DR microstripline coupling characterisation. The DR

coupling was measured by keeping DR by the side of 50 ohm microstripline mounted in a test jig and noting S_{21}. DR coupled to the single transmission line acts as a band stop filter. Loaded quality factor was calculated for a particular lateral distance of DR to microstripline. The expression for loaded factor is given as

$$Q_l=(f_2-f_1)/f_0 \quad (18)$$

where (f_2-f_1) is the 3 dB bandwidth and f_0 is the resonant frequency.

Measurements were repeated for various DR microstripline lateral distances and the corresponding loaded quality factors were computed.

The unloaded quality factor Q_o at a desired frequency is calculated by curve fitting the manufacturer's data . The measured quality factor for the present case is $Q_l=1489$. Using these values lumped element parameters of DR were found to be R=384 ohm, L=0.611pH and C=135pF

2.4 Active device consideration

The selection of active device is an important step. It is the device instability inherent or incorporated through external means that is essential for successful design of an oscillator. There are a number of ways in which this can be achieved. Some of them are reverse channel operation, introduction of inductance in gate port or the most commonly used one of capacitance in source port. However, in the present case gate to drain capacitance C_f has been employed to generate negative resistance. The characteristic feed forward capacitance was selected because of realizability consideration. The simulated data for magnitude of reflection coefficient of the device over frequency range is given in Table I. Data provided by manufacturer was used to decide the operating point and to achieve reasonable power output.The circuit was characterised using commercial circuit simulator[7]. Equivalent circuit of WGDRO is shown in fig.2.

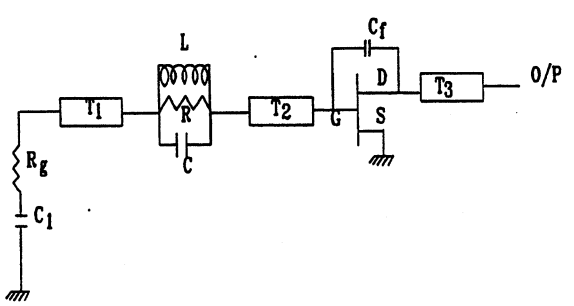


Fig 2. Equivalent Circuit of WGDRO

3. EXPERIMENTAL PROCEDURE

The schematic diagram of WGDRO is shown in fig.3.The circuit was fabricated on alumina substrate of thickness 0.635 mm. The substrate was die attached to the carrier which in turn was fixed to the aluminium housing. The active device of M/s. HP-Avantek ATF-35376 HEMT was mounted on the alumina substrate. Input and output lines were simply 50 ohms. C_1 was used as a DC block. R_g was used to suppress any spurious oscillation . A driver card was used for providing dual bias V_{ds}=2.2 V at I_{ds}= 22mA V_{gs}= -0.3 V and to ensure

Table I. Magnitude of reflection coefficient

Freq. (GHz)	|S[1,1]| for various capacitance(pf)C_f				
	0.04	0.05	0.07	0.09	0.10
10.00	0.657	0.661	0.671	0.683	0.690
11.00	0.676	0.692	0.723	0.753	0.768
13.00	0.712	0.744	0.810	0.874	0.904
13.00	0.764	0.823	0.950	1.085	1.151
14.00	0.831	0.927	1.167	1.474	1.648
15.00	0.905	1.051	1.480	2.243	2.859
16.00	0.973	1.167	1.813	3.315	4.994
17.00	1.017	1.232	1.890	2.789	3.053
18.00	1.024	1.217	1.652	1.885	1.863
19.00	0.995	1.140	1.363	1.396	1.358
20.00	0.944	1.039	1.140	1.124	1.094

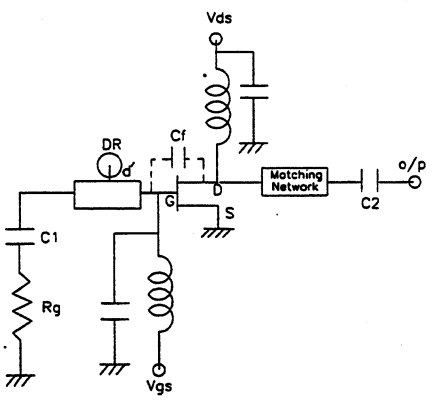


Fig 3. A schematic diagram of WGDRO

safe operation. Dielectric resonator was moved along the gate line and its lateral separation d' from the transmission line was adjusted for stable frequency and reasonable power output. Measurements were done using spectrum analyzer, frequency counter and phase noise evaluation setup.

4. RESULTS

Power output was measured on a HP8563A spectrum analyser. The output spectrum of WGDRO is shown in fig.4. Phase noise measurements were carried out and performance is shown in fig 5. Temperature stability study was carried out over -20°C to + 55°C. The measured results of WGDRO are given in Table II.

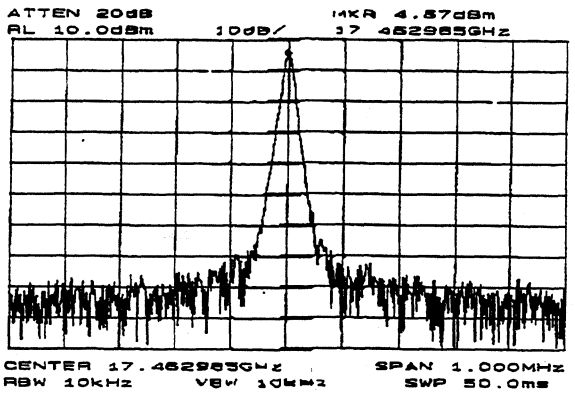


Fig 4. Spectral response of WGDRO

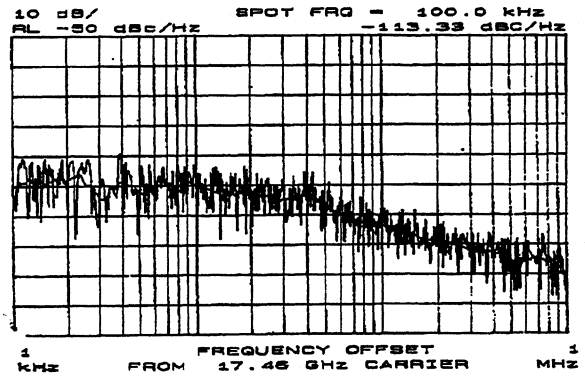


Fig 5. Phase noise response of WGDRO.

Table II. Experimental Results

S.No.	Parameter	Measured results
1	Frequency(GHz)	17.463
2	Power output(dBm)	4.7
3	Phase Noise(dBc/Hz @ 100 KHz away)	-113
4	Temp. stability(ppm/°C over -20°C to 55°C)	3

5. DISCUSSION AND CONCLUSION

A whispering gallery mode dielectric resonator oscillator employing novel feedback method is presented. The circuit offers a very selective feedback over a specific range of frequency band. It gives a reasonable power output and excellent phase noise. The phase noise of WGDRO is definitely better than the conventional $TE_{0.1.\delta}$ mode DRO at the designed frequency of operation. As envisaged the top plate of WGDRO does not play a crucial role for the onset of desired oscillation and frequency shift as it does in $TE_{0.1.\delta}$ DR oscillator. This is evident from WG mode pattern and energy confinement. Measured results also validate computational accuracy of eigen frequencies.

6. ACKNOWLEDGEMENTS

The authors wish to thank Mr. N. Divakar, Director, DLRL, Hyderabad for his interest and continuous encouragement. They wish to acknowledge the help rendered by Dept. of Physics IIT, Madras for carrying out part of this work. The authors specially like to thank Mr.T. Madhava Rao and Mr.V. Ramasankaram for many valuable ideas and helpful discussions. Thanks are also due to Mrs.Leelamadhuri and Mrs.Kireetamma for help during testing and evaluation of the component.

7. REFERENCES

[1] D.Kajfez and P. Guillon (Eds.), Dielectric Resonators. Dedham, MA, Artech House, 1986.

[2] X.H. Jiao, P. Guillon and L. Bermudez, "Resonant frequencies of Whispering gallery dielectric resonator modes", IEE Proc.H, vol. 134, pp.497-501, December 1987.

[3] A.E. Centeno and G.B. Morgan, "Design of millimeter wave dielectric resonators for Integrated Circuits",IEE Proc.H, vol. 139, pp.307-308, June 1992.

[4] P.Guillon,B.Byzery and M. Chaubet,"Coupling parameters between a dielectric resonator and microstripline", IEEE Trans., MTT, vol.33, pp.222- 226 ,March 1985 .

[5] I. U. Khairuddin and I.C. Hunter,"Computation of coupling between whispering gallery mode dielectric resonators and a microstrip transmission line", IEE Proc.Microwave and Antenna propagation, vol. 142, pp.265-268, June 1995.

[6] S.L. Badnikar, N.Das, N.Shanmugam and V.R.K. Murthy," Resonant frequencies of whispering gallery dielectric resonator in MIC environment" , Tenth national seminar on ferroelectrics and dielectrics,NSFD-X,IIT Madras, Dec 16-18, 1998.

[7] Series IV HP EEsof High Frequency Design Solutions, July 1995.

A MINIATURIZED MICROWAVE RESONATOR FOR RUBIDIUM FREQUENCY STANDARDS

G.H. Mei and J.T. Liu
Wuhan Institute of Physics and Mathematics, Chinese Academy of Sciences,
P.O. Box 71010, Wuhan 430071, The People's Republic of China

ABSTRACT

A new type of miniaturized resonator based on the magnetron resonator theory has been developed. The resonator is of small size and high quality factor, and has been used in a rubidium frequency standard.

1. INTRODUCTION

Based on the magnetron resonator theory[1], various microwave resonators[2-7] have been designed for different purposes. This type of resonators has obvious advantages of small size and high quality factor. Among them, the magnetron resonator developed by Schweda and Mileti et al has been successfully used in a compact rubidium frequency standard[6,7]. Recently we developed a new type of miniaturized resonator for rubidium frequency standards. The resonator is also based on the magnetron resonator theory[1], but in structure is different from those mentioned above.

2. MAIN FEATURES

A slotted tube was used in the resonator. The resonance frequency is mainly determined by the geometry parameters of the slotted tube and the cylinder cavity surrounding it. For a reasonable set of the parameters, a resonance frequency close to 6.8 GHz, which is the hyperfine splitting of the ground state of 87-Rb, can be easily obtained. Fine tuning of resonance frequency of the resonator can be realized by changing the position of the rubidium absorption cell in the slotted tube, but we found that it can be realized also by using a tuning screw more conveniently.

Experiments show the resonator has nice performances. A return loss of microwave from the resonator larger than 20dB, and a loaded quality factor approximately 500 have been observed. The design of the resonator allows to load dielectric material into the cavity to reduce the cavity volume though it may be unnecessary if a slightly lager cavity volume is acceptable. Experiments also show that the loaded dielectric material has no much influence on the cavity quality factor. Typical dimension of the inner wall of the cavity is 20mm in diameter and 28mm in height for the cavity with dielectric material loaded. We believe the resonator can be made even smaller.

Using the resonator we have built a physical package and a rubidium frequency standard with a volume slightly larger than one liter, but the resonator itself allows to reduce the volumes of both the physical package and the frequency standard.

3.REFERENCES

[1] G.B. Collins in "Microwave Magnetrons", Mc Graw Hills, N. Y., 1948, pp. 49-50, pp. 59-62

[2] T.M. Wang, J.B. Lewis and S.B. Crampton, Proc. 33rd Annual Symposium on Frequency Control, Philadelphia, 1979, pp.543-548

[3] W. N. Hardy and L. A. Whitehead, Rev. Sci. Instrum. 52, 213 (1981)

[4] W. Froncisz and James S. Hyde, J. Mag. Reson. 47, 515 (1982)

[5] T. Sphicopoulos and F. Gardiol, IEE Proc., Vol. 134, Pt. H, No. 5 , 405 (1987)

[6] H. Schweda, L. Landeron, G. Busca and P. Rochat, US. Pat. No. 5,387,881

[7] G. Mileti, I. Ruedi and H. Schweda, Proc. 8th EFTF, 515 (1992)

0-7803-5400-1/99/ $ 10.00

CONTINUOUS FREQUENCY CONTROL FROM 5 TO 10 GHZ OF AN OPTICAL OSCILLATOR

O.Frazão[1], P.Tavares[1], A.F.Cunha[2], L.Ribeiro[3], J.Ferreira da Rocha[1,4]

[1] Institute of Telecommunications, Portugal, +351 34 377900, ofrazao@av.it.pt, ptavares@av.it.pt
[2] Dept. of Physics, University of Aveiro, Portugal, +351 34 370818, acunha@fis.ua.pt
[3] Dept. of Industrial Electronics, University of Minho, Portugal, +351 53 510190, lfbr@dei.uminho.pt
[4] Dept. of Electronics and Telecommunications, University of Aveiro, Portugal, +351 34 370200, frocha@ua.pt

ABSTRACT

Continuous frequency control from 5 to 10 GHz of an optical oscillator, can be induced in a single cavity DFB-MQW laser by simple change the laser bias current using the injection locking mechanism. This system is very attractive for optical communications, namely for clock recovery in high bit rate optical transmission systems.

1. INTRODUCTION

The use self-pulsating lasers is becoming very attractive, namely in transparent optical networks (TON) specified to operate exclusively in the optical domain. Recently, self-pulsating lasers have been selected to operate in conjunction with TOAD demultiplexers to build optical add/drop nodes.

This paper reports recent laboratory results of a tunable self-pulsation prototype. The physical process responsible for the laser self-pulsation is the injection locking mechanism. An anti-reflection (AR) coated facet of the laser chip is optically coupled to a single mode fiber through an optical ferrule. In this coupling process some light is reflected from the ferrule and injected back into the laser cavity. This enables the generation of ultrashort pulses at higher repetition rates compared to the traditional configuration [1]. This experiment is inspired on a theoretical study [2] using optical resonant cavity model.

2. THEORY

The conceptual model of the semiconductor LD with external light injection is illustrated in Fig.1. A coherent optical beam from an external source with frequency ν is injected into the laser diode resonator through one of the facets.

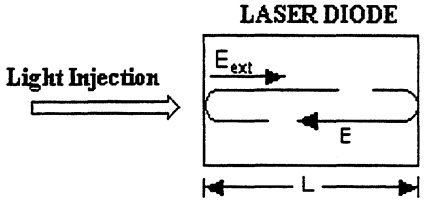


Fig. 1 - Diode laser model with external light injection

According to Lang [2], the injection locking mechanism is governed by the following equations.

$$\frac{dE}{dt} = \frac{1}{2}[G(n) - \alpha]E - i[\omega(n) - \nu]E + \kappa E_{ext} \quad (1)$$

$$\frac{dS}{dt} = [G(n) - \alpha]S + C_{sp}\frac{n}{\tau_s} \quad (2)$$

$$\frac{dn}{dt} = P - \frac{n}{\tau_s} - G(n)\left[|E|^2 + S\right] \quad (3)$$

The first equation describes the electric field, E, behavior, where α is the cavity loss, $G(n)$ represents the modal gain, $\omega(n)$ is the resonant frequency of the cavity mode to which the light is injected, both are dependent on n, the carrier density. The electric field amplitude of the light coupled into the cavity, modified in such a way that it has units of $(\text{photons/m}^3)^{1/2}$, is represented by E_{ext}. For the calculation, it is assumed that $G(n)$ and $\omega(n)$ are approximated by,

$$G(n) = \alpha + g(n - n_{th}) \quad (4)$$

$$\omega(n,\Omega) = \omega_{th} + \frac{\partial\omega}{\partial n}(n - n_{th}) - (\frac{\eta_{eff}}{\eta} - 1)(\Omega - \omega_{th}) \quad (5)$$

where g is the differential gain, Ω is the frequency of the light injected mode which has been assumed equal to the injected light frequency ν and ω_{th} is the threshold cavity resonance frequency. The possible excitation of the neighboring axial mode with detuning increase can be accounted for by the second equation for the photon density, S, in that mode. Parameter C_{sp} is the spontaneous emission coefficient and τ_s is the carrier lifetime. The last equation describes the excited carrier density, where P is the carrier injection rate per unit of volume. In carrying out the numerical solution, it was convenient to define a new set of normalized variables [2]. The parameters are normalized with respect to cavity loss, α, equal to the inverse of photon lifetime. The model above was used in the assessment of the instability region for a 1320nm semiconductor laser diode. This device has been characterized and its main features are summarized in table 1.

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

Table 1 - Laser diode parameters

Name	Description	Value
V	Active layer volume	1.0×10^{-16} m^3
τ_p	Photon lifetime	2 ps
τ_s	Carrier lifetime	2 ns
A	Gain constant	2×10^{-20} m^2
I_{th}	Threshold current	25mA
η	Refractive index	3.5
η_{eff}	Effective refractive index	3.8
Γ	Confinement factor	0.28
C_{sp}	Spontaneous emission coef.	5.5×10^{-5}

We have studied the instability conditions in a single cavity semiconductor laser. The influence of the external light injection level, detuning and increasing step height of excitation bias current were analysed [3].

3. EXPERIMENTAL RESULTS

The experimental set-up (see Fig. 2) was originally built upon using one Multiple Quantum-Well (MQW) distributed feedback (DFB) semiconductor laser. The device operates at 1310nm, having an active length of 250μm and a threshold current of I_{th}=11.2 mA. One facet is AR-coated and the other has 30% reflectivity. The light is collected from the laser by a lensed fiber tapper. It was verified that the device is very sensible to the light polarization.

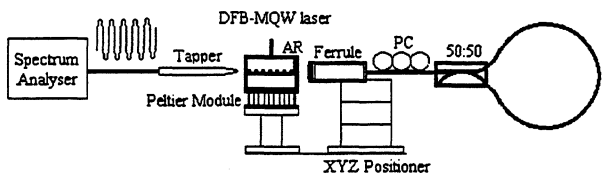


Fig. 2 - Experimental Set-up

Adjusting the distance between the AR laser facet and the fiber ferrule for achieving maximum optical coupling, the frequency of self pulsation can be modified simply adjusting the bias current (see Fig. 3).

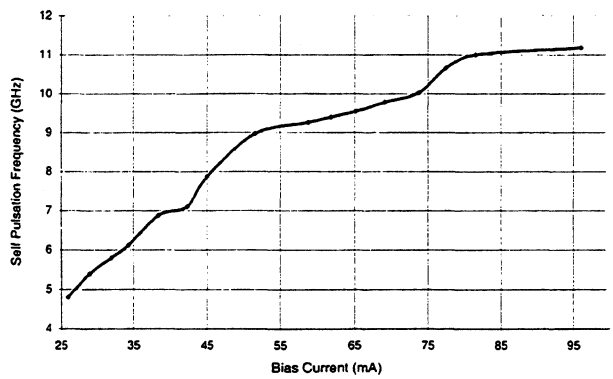


Fig. 3 - Pulsation frequency as function of the current.

For lower currents it was verified that the self pulsation frequency is near 6GHz and the device is stable. For higher currents the self pulsation frequency increases to near 11GHz. The device shows some jitter for high frequencies as a consequence of the spontaneous emission.

The RF spectrum of the self pulsation behavior for different bias current is shown in Fig. 4. It can be seen that for 30mA the frequency is approximately 5 GHz and for 90mA it goes near 10 GHz. It can also be seen the second harmonic near 10 GHz for 30 mA and near 20 GHz for 90mA. The resolution bandwidth is 1MHz.

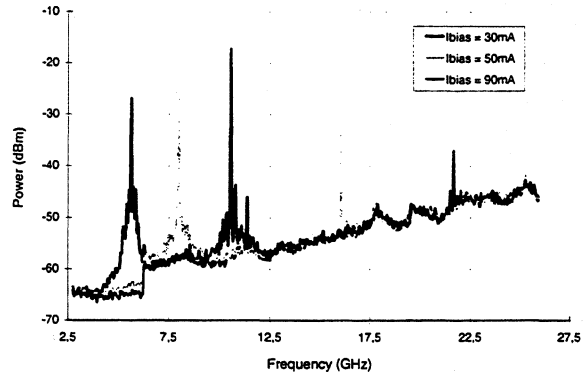


Fig. 4 - RF spectrum for different bias currents.

3. CONCLUSIONS

An experimental set-up configuration was projected and built. With this configuration the frequency of self pulsation can be modified simply adjusting the bias current between 25 and 100mA. For lower currents it was verified that the self pulsation frequency is near 6GHz and the device is stable. For higher currents the self pulsation frequency increases to near 11GHz. This work is sponsored by the UPGRADE project from the ACTS program.

5. REFERENCES

[1] Peter Vasil'ev "Ultrafast Diode Lasers - Fundamentals and Applications", Artech House Publishers, London, 1995.

[2] Roy Lang, "Injection Locking Proprieties of a Semiconductor Laser", IEEE Journal of Quantum Electronics, Vol. QE-18, N°6, pp. 976-983, Jun. 1982.

[3] O.Frazão, A.F.Cunha, P.Tavares, L.Ribeiro, J.Ferreira da Rocha, "Modelling of a 4x10Gbps All-Optical Clock Extraction System", in proceedings of the conference IEEE/ICECS'98, 1998, pp 541-544.

1.5 - 40 GHz MULTIFUNCTIONAL MICROWAVE SAMPLING CONVERTER

Arcady M. Schitov
Institute of Electronic Measurements (IEM) KVARZ, 176 Prospekt Gagarina, Nizhni Novgorod, 603009, Russia

ABSTRACT

A two-channel superbroadband Frequency Converter was developed to convert 1.5-40 GHz microwave range signals into 0.4-500 MHz intermediate frequency (IF) range signals. The Converter is a modular-type instrument manufactured to meet VXI standard requirements. The Converter is conjunction with the corresponding LF devices effectively extends the frequency range up to 40 GHz of the following radiomeasuring instruments: spectrum analyzers, frequency meters, sampling oscilloscopes, amplitude/phase meters, network analyzers, power meters and so on.

1. INTRODUCTION

A great variety of radiomeasuring instruments for microwave signal analysis in time and frequency domain (oscilloscopes, amplitude/phase meters, network analyzers, frequency meters, deviation meters and etc.) use the sampling frequency conversion principle: a conversion of a microwave signal frequency into relatively low intermediate frequency in accordance with the relation

$$f_{IF} = f_{RF} - nf_{LO} .$$

where n = 1, 2, ... – local oscillator (LO) harmonic number. In spite of some peculiarities of each of these instruments, their input circuits - samplers are identical.

We tried to manufacture a small size microwave sampler for the max. allowable 1.5-40 GHz frequency range with main specifications (sensitivity, dynamic range and others) and input circuits as good as in individual instruments.

A two-channel sampler with a built-in 2-4 GHz LO locked by 200-400 MHz synthesizer external signal is the heart of the Converter.

2. BLOCK DIAGRAM AND DESIGN

The Converter is manufactured in accordance with the block-diagram on Fig. 1 and is composed of:

1) a two-channel sampling mixer with a built-in short pulse former based on a step-recovery diode (SRD) and a low noise 10-500 MHz IF amplifier in one channel and a 0.4-50 MHz IF amplifier in the other;
2) a local oscillator - a miniature 2-4 GHz JIG-tuned transistor oscillator with ~ 10 mW output;
3) a LO power amplifier (P_{out} ~ 200 mW) with a microwave signal separation circuit;

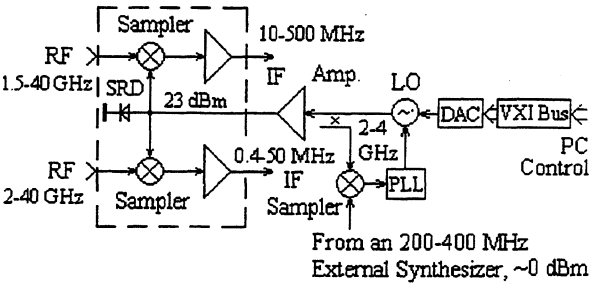


Fig. 1.The block-diagram of the Two-Channel Sampling Mixer.

4) a printed circuit (p.c.) board with control and phase-locked loop (PLL) circuits;
5) a VXI interface p.c. board.

The Sampling Converter is manufactured as a VXI standard instrument. It is a single width C-size module designed to be plugged into a slot of a mainframe. The module is composed in accordance with functional-assembly principle and comprises two p.c. assemblies and four hybrid integrated circuit (HIC) protected by a cover.

The Converter front panel comprises:

1) two microwave inputs (coaxial connectors 2.4/1.04 mm) for application of the investigated signals;
2) two IF outputs (coaxial connectors 3.5/1.5 mm) to connect LF measuring instruments;
3) input connector (3.5/1.5 mm) to apply a 200 - 400 MHz external synthesizer signal;
4) output connector (3.5/1.5 mm) to output a 1 mW synthesized signal of a 2 - 4 GHz LO (not shown on Fig. 1);
5) output connector (3.5/1.5 mm) to apply a 10 MHz reference signal to a PLL circuit (not shown on Fig. 1);
6) light emitting diode (LED) indicating PLL locking.

3. CONVERTER OPERATION

To provide measurement task solution the microwave Converter is aggregated into an automatic measurement system with 200-400 MHz frequency synthesizer, 10 MHz reference signal source and the required LF signal loggers: an oscilloscope, a frequency meter, a spectrum analyzer, an amplitude and phase meter and other LF instrumentation. For network analysis the automatic measurement system must comprise microwave synthesized signal source and external directional

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

couplers. Each measurement task is solved in accordance with the developed algorithms of an automatic measurement system operation, which is controlled by a personal computer (PC).

The Converter control amounts to setting the required LO frequency (2-4 GHz) and its switching in accordance with the predefined algorithm.

The initial LO frequency presetting is made by the voltage of the (digital-to-analog converter) DAC controlled by PC through VXI bus. PLL circuit tunes LO frequency to the 10^{th} harmonic of a 200-400 MHz signal of a certain frequency, which is applied from an external synthesizer.

4. EXPERIMENTAL RESULTS

The main specifications of the Converter have been experimentally investigated.

4.1 Frequency range

Table 1 presents the values of IF output voltage (V) at P_{in} = 0.1 mW, f_{IF} ~ 1-10 MHz, R_{IF} =50 Ohm for LO various frequencies - 2, 3 and 4 GHz when tuning input signal frequency from 1.5 to 40 GHz.

Table 1

f_{RF}, GHz	1.5-40 GHz Channel (1)			2-40 GHz Channel (2)		
	LO Frequency, GHz			LO Frequency, GHz		
	2	3	4	2	3	4
1.5	0.2					
2	0.2			0.22		
3		0.24			0.28	
4	0.2		0.3	0.28		0.36
6	0.25	0.3		0.3	0.32	
8	0.28		0.36	0.35		0.43
10	0.28			0.35		
12	0.26	0.3	0.36	0.35	0.4	0.44
14	0.22			0.3		
16	0.2		0.3	0.3		0.4
18	0.18	0.18		0.25	0.3	
20	0.18		0.25	0.22		0.3
22	0.2			0.25		
24	0.2	0.22	0.26	0.25	0.3	0.35
26	0.22			0.26		
28	0.22		0.35	0.25		0.35
30	0.22	0.3		0.25	0.3	
32	0.2		0.3	0.25		0.32
34	0.16			0.22		
36	0.14	0.16	0.18	0.2	0.24	0.26
38	0.14			0.17		
40	0.15		0.18	0.17		0.22
ΔK_p, dB	±3	±3	±3	±3	±2.2	±3

The measurement results show that the conversion ratio flatness K_p=10·lg P_{out} /P_{in} in 1.5(2) - 40 GHz range at LO fixed frequencies is not more than ±3 dB. K_p dependence on LO frequency at fixed frequencies of a microwave signal and IF lies within 2-3 dB. Conversion ratio (K_p) non-identity of the Converter microwave channels is not more than 2 dB in a full frequency range at equal IF amplifier gain. The measured IF channel flatness at f_{LO}= 4 GHz, $f_{RF} \approx$ 4- 4.5 GHz is not more than 1.5 dB in 0.4-50 MHz range (2) and not more than 3 dB in 10-500 MHz range (1).

Microwave input VSWR values are given in Table 2.

Table 2

Δf, GHz	1.5-18	18-26	26-40
Channel 1	2.15	2.05	2.25
Channel 2	2.45	2.2	2.45

4.2 Dynamic range

The IF signal output level was measured at 10 dB input signal change (P_{in}= 0.1 mW and 0.01 mW) and f_{RF} = 36.05 GHz, f_{IF} = 50 MHz. The conversion error caused by nonlinearity was not more than 0.8 dB (Table 3).

Table 3

f_{LO}, GHz	2	3	4
Channel 1	0.6	0.15	0.8
Channel 2	0.1	0.3	0.3

This means that compression level $(\Delta K_p = 1\,\text{dB})$ corresponds to input power level of more than 0.1 mW. Dynamic range estimate results (signal-to-noise-ratio (dB) measurement) at f_{RF}=36.05 GHz, P_{RF}= 0.1 mW, f_{IF} = 50 MHz in 10 kHz bandwidth are given in Table 4.

Table 4

f_{LO}, GHz	2	3	4
Channel 1	84	87	89
Channel 2	81	84	84

The obtained results allow to realize dynamic range in spectrum analysis mode 80-85 dB (in 10 kHz bandwidth) and in oscilloscope mode 35-50 dB in 0.4-50 MHz, 10-500 MHz bandwidths.

4.3 Absolute power level (voltage) measurement

The IF signal power level time stability was investigated. After the Converter 30 min wormup IF signal level during 30 min was changed by not more than 1%. This testifies that the Converter may be used to measure absolute levels of power and voltage of microwave signals.

5. CALIBRATION. MEASUREMENT RESULTS CORRECTION

Conversion errors may be caused by:

1) conversion mode non-linearity;
2) amplitude and phase-frequency ripple of microwave, LO, IF channels;
3) non-sufficient channel isolation;

which causes a distortion during absolute measurements of microwave signal voltage and power.

To prevent a non-linear and intermodulation distortion one has to limit a microwave input signal level to 0.1 mW. The error caused by non-ideal frequency response of channels may be taken into account and excluded from a measurement result by a preliminary calibration. This is favoured by the Converter building concept, which sets the unique relation

$$f_{IF} = f_{RF} - nf_{LO} = [f_{RF} - n(10 \cdot f_{syn} + 10\ \text{MHz})]$$

between the known f_{syn} and the measured n, f_{IF} and f_{RF}.

The calibration embodies the preliminary measurement and storage of transfer ratios of all the Converter channels (signal, LO, IF) in a full frequency range.

6. CONCLUSION

The two-channel microwave Converter was developed. The manufacturer guaranties the following specifications:

1. Microwave input frequency range 1.5(2) - 40 GHz.
2. Microwave input VSWR not more than 2.5 (1.5-26 GHz); 3 (26-40 GHz).
3. Output frequency range 10-500 MHz (1); 0.4-50 MHz (2);
4. Microwave signal max level at 1 dB compression is not less than 0.1 mW.
5. Dynamic range (signal-noise-ratio) at P_{in} = 0.1 mW in 10 kHz bandwidth is not less 70 dB (1) and 80 dB (2).
6. IF output signal level at P_{in} = 0.1 mW not less than 0.1 mW at 50 Ohm load.

The high performance is achieved thanks to unique components developed at IEM KVARZ and used in the Converter: especially monolithic two diode mixer IC with low parasitic parameters and identical diode specifications, high speed SRDs (switching time is not more than 20 ps), special HIC, and also thanks to the selection of relatively high frequencies of the Sampling Converter LO (2-4 GHz) [1].

The selection of LO high frequencies, on the one hand, provides low conversion loss in the Sampling Converter (20-25 dB) and as a result the Converter high sensitivity and, on the other hand, limits the low frequency end of the converted frequencies

$$f_{in} \geq f_{LO} - f_{IF}.$$

The technical solution proposed in work [2] and later realized in the RK4-55 Network Analyzer developed at IEM KVARZ extended the Converter frequency range. A microwave signal stabilized in power and frequency with $f_0 \gtrsim f_{LO\ max}$ is additionally applied to the sampling Converter LO. In so doing the Converter output signal frequency is

$$f_{IF} = f_{RF} - (f_0 - f_{LO}),$$

that lowers the low end of the investigated frequency range to

$$f_{RF\ min} \approx f_0 - f_{LO\ max}.$$

One has to note that during two frequency conversion the conversion loss is increased by 6 dB.

In the investigated Converter we preclude the possibility of two frequency conversion realization for 0.01-2 GHz range.

Besides we envisage the Converter realization as a small size self-contained instrument comprising an additional 10 MHz reference signal source, a 200-400 MHz synthesizer and a ADS used to log IF signals with display of measurement results on PC screen in oscilloscope and spectrum analyzer mode.

7. REFERENCES

[1] A.M. Schitov, J.S. Zaitsev, A.M. Lomakin «DC to 120 GHz Harmonic Mixers for Millimeter Wave Instrumentation».- International Conference on Millimeter Waves and Far-Infrared Technology. Beijing, China, 1992, pp.217-220.

[2] А.М. Щитов. Стробоскопический преобразователь амплифазометра СВЧ в режиме двойного преобразования частоты. - Техника средств связи, серия РИТ, 1991, вып.3, с.72-76.

A HIGHLY STABLE, PHASE COHERENT, WIDEBAND MICROWAVE SYNTHESIZER FOR RADIO ASTRONOMY APPLICATIONS

G.H. Tan
Netherlands Foundation for Research in Astronomy

Postbox 2
7990 AA Dwingeloo
The Netherlands

ABSTRACT

A description is given of a wideband microwave synthesizer that is used as a local oscillator in a novel front-end design. This front-end design is used in a synthesis radio telescope operating in the frequency range up to 8.6 GHz. The aperture synthesis principle on which this radio telescope is based, as well as Very Long Baseline Interferometry (VLBI) techniques require exceptional phase and frequency stability which are fulfilled by the presented oscillator design.
A phase stability of better than 5° rms, averaged over 60 seconds, is obtained at 5 GHz over periods of 12 hours. Allan-variance values approaching 10^{-14} are measured for τ = 1000 seconds. The output signals among two synthesizers tuned to the same frequency are always phase coherent with a fixed phase offset.

1. APPLICATION AREA

The described synthesizer design is used as a tuneable first local oscillator in a microwave receiver. This receiver covers eight frequency bands in the range form 250 MHz up to 8600 MHz. All receiver systems together with feeds are housed in one unit, which is normally addressed as the front end. More details on this front-end design can be found in references [1] and [2].

Fourteen of these front ends are currently in use in one of the largest radio telescopes in the world, the Westerbork Synthesis Radio Telescope (WSRT) [3]. The WSRT consists of 14 prime focus, parabolic reflector antennas, each having a diameter of 25 metres, on an approximately 3 kilometres long east-west baseline.

The primary mode of operation for the WSRT uses the earth rotation synthesis technique [4]. Within 12 hours the instrument rotates by 180 degrees relative to the sky. The 14 telescopes sample a virtual aperture with a diameter equal to the baseline. Cross-correlation and Fourier transformation of this information gives the brightness distribution across the sky. With this technique sky maps with a high angular resolution can be obtained. At a commonly observed wavelength of 6 centimetres a resolution of 3.7 arcseconds is achieved. Figure 1 gives a schematic view of this mode of operation. The same central reference oscillator drives each local oscillator in one of the 14 front ends. This central reference oscillator is a hydrogen maser and delivers two fixed frequencies, 1 MHz and 180 MHz.

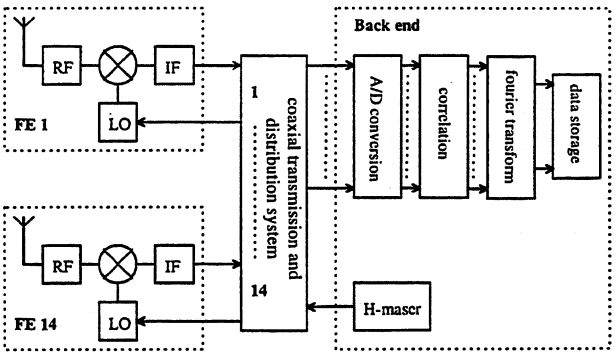


Figure 1. WSRT local interferometry configuration

Another observational technique used at the WSRT is Very Long Baseline Interferometry (VLBI) [5]. In this mode the WSRT is one part of an interferometer pair while the other part is a radio telescope elsewhere in the world. The baseline between the two radio telescopes forming the interferometer can range from a few 100 kilometres up to more then 10.000 kilometres. With space based radio telescopes this distance can be even much longer. The main aim of this observational technique is to provide a very high angular resolution. The data is stored on tape and a later moment correlated with each other.

For VLBI measurements the WSRT is arranged as a phased array. The signals from all telescopes are added with a correction for the proper delay of each telescope. Figure 2 shows a schematic diagram of this VLBI mode.

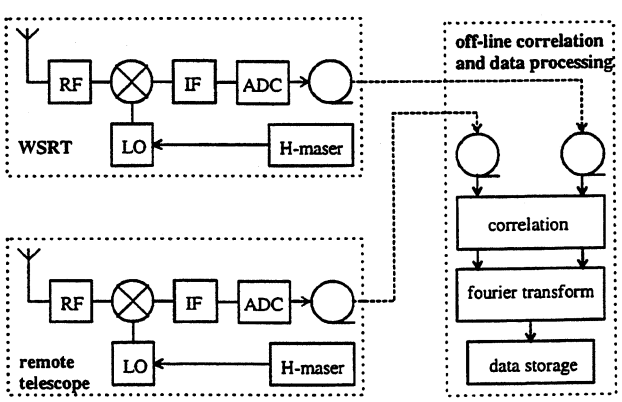


Figure 2. VLBI configuration

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

The oscillators of the radio telescopes forming the interferometer pair are not directly locked to each other during the measurement. Instead, the individual oscillators should have sufficient frequency stability while the measurement lasts. Synchronisation is done off-line in the correlation process.

Both modes of operation call for specific requirements of the local oscillator system. For local synthesis measurements, differential, phase stability is of importance. While for VLBI observations an additional requirement is needed for the frequency stability.

2. MAIN SPECIFICATIONS

From the overall receiver requirements the following synthesizer specifications have been deduced:

- Frequency range: 1200 MHz–2200 MHz (low)
  2200 MHz–9600 MHz (high)
- Step size: 1 MHz
- Spectral purity: -90 dBc/Hz at 10 kHz offset
- long term stability: $<10^{-14}$ for $\tau = 1000$ sec.
- Spurious: <-80 dBc
  <-110 dBm below 1100 MHz
- Harmonics: < - 10 dBc
- Phase coherence between synthesizers

Note that two versions, having different frequency ranges, have been developed. Design of both versions is identical and only minor hardware differences exist for some of the microwave components like eg. the YTO.

3. SYNTHESIZER CONFIGURATION

Due to the already existing frequency reference infrastructure at the WSRT, the H-maser with output frequencies of 1 MHz and 180 MHz, and the wide tuning range needed a design consisting of two loops was chosen. Furthermore each loop has a mixer to convert the oscillator signal to a lower frequency which is then applied to the programmable digital dividers. Figure 3 shows the basic lay out of the synthesizer.

Due to the high spectral purity of the signal driving the mixer a much better phase noise performance is obtained compared to a phase locked loop where only dividers are used to convert the oscillator signal to the input of the phase detector.

In the main loop a sampling mixer is used instead of a fundamental one to cover the wide frequency range of the YTO with the limited tuning range, 177 MHz to 183 MHz, of the control loop. The harmonic number N is somewhere in the range from 6 to 54 depending on the selected output frequency.

The choice for this sampling mixer has played a major role in the architecture of the synthesizer. It was known from other applications, like network analysers, that these sampling mixers could have a high degree of long-term phase stability. Especially this property has been used in the design to obtain the necessary long-term phase stability.

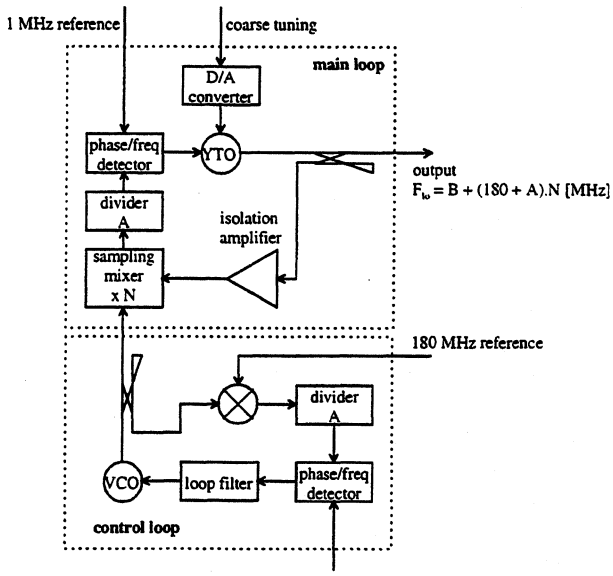


Figure 3. Synthesizer lay out

The synthesizer lay out is such that the phase offset between two different units tuned to the same frequency is always constant. This even holds when one or both synthesizers are temporarily tuned to another frequency and later on return to the original frequency.

4. DEVICE TECHNOLOGIES

In the realisation of the synthesizer a number of device technology choices have been made. The most important choices are the following.

For the oscillator of the main loop a YIG Tuned Oscillator (YTO) has been chosen. A YTO can be tuned over a very large frequency range as is especially needed for the high band synthesizer. Spectral purity is good and fulfils the requirements of the application. The active component of the oscillator is a silicon bipolar transistor. It has considerably less 1/f noise compared to GaAs transistors. Tuning is done by varying the magnetic field strength around the YIG sphere. For this purpose two coils are present. One is used for coarse frequency control while a second coil with less tuning sensitivity is used to phase lock the YTO. The fine tuning control of the YTO is nearly constant over the whole frequency range and very repeatable from unit to unit. These features, tight tolerances, make the design of the phase locked loop less complicated. The relatively slow tuning speed of the YTO coarse tuning compared to e.g. varactor-tuned oscillators is of no concern in this application.

All digital circuits, which are applied inside both phased locked loops, are based on standard ECL technology. GaAs circuits have been evaluated at an earlier stage, but showed considerably less performance in terms of phase noise and long term stability compared to silicon circuits. Although new TTL

families like ACT might show similar or even better phase noise performance then ECL [6], the choice has been made for ECL. Main reason for this is that the balanced nature of ECL circuits has important benefits from an EMC point of view. The power supply rail is not contaminated by the switching digital circuits, while the balanced signal lines, assuming a proper lay out on the PCB, radiate little energy and vice versa are less susceptible to interfering EM fields.

Phase/frequency detectors have been used for all loops. Due to the nature of these detectors they have a very large pull-in range. This avoids the necessity for additional acquisition circuitry to obtain phase lock. The AD9901 from Analog Devices has been chosen since it does not have the dead-zone problem encountered in older type phase/frequency detectors. This dead-zone feature might give loop stability problems because of the unpredictical detector gain around the crossover point. In addition it might introduce unwanted phase drift.

The sampling mixer has a lay out which is very common for microwave applications and can be found in e.g. network analysers and spectrum analysers [7]. It consists basically of two parts, an impulse generator, which drives a single balanced schottky diode mixer. The impulse generator uses a Step Recovery Diode (SRD) to generate an impulse with a width of approximately 50 picoseconds. The SRD has an external DC bias, which is used to compensate for various interfering effects, as is described in the next chapter.

5. SYSTEM DETAILS

5.1 PLL parameters

Both phased locked loop synthesizers are basically second order systems. The loop bandwidth. Has been chosen such that no noise peaking occurs [8]. With the phase noise performances of the reference frequencies and the phase noise of the control loop VCO respectively that of the main loop YTO this resulted in bandwidths of around 20 kHz.

Due to the large division ranges for both A and B programmable dividers some form of compensation was necessary to ensure proper, stable loop operation. This has been achieved by making the gain of the loop filter variable. According to the selected division ratio the loop filter gain is adjusted in such a way that the loop gain remains constant.

5.2 Tuning procedure

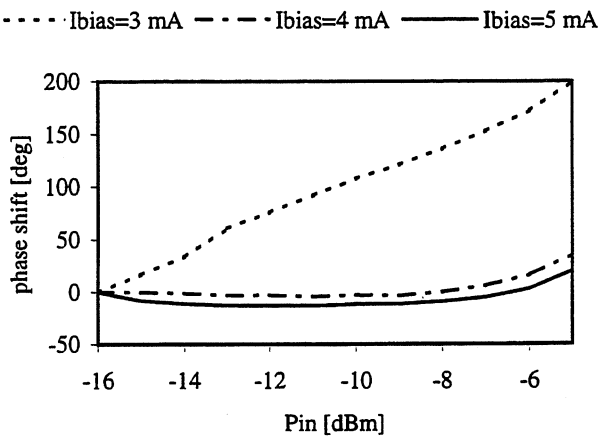
The programmable dividers, course tuning of the YTO through a D to A converter and other functions of the synthesizers are controlled by a separate embedded microcomputer in the front end.

The normal procedure is that the wanted frequency can be directly passed on to this microcomputer. From the entered frequency appropriate values for A, B and N are calculated. The YTO is tuned to the approximate output frequency by the coarse tuning coil. The A and B values are stored in the programmable dividers. After this initialisation the loops lock automatically to the proper frequencies.

5.3 Biasing of SRD pulse generator

The SRD used in the pulse generator of the sampling mixer has an external DC bias from a current source. The current delivered by this source is temperature dependent in such a way that it compensates phase shifts in the SRD pulse generator. Without this compensation the temperature dependent recombination time of the charge carriers in the PIN junction of the SRD would cause a considerable phase shift.

Furthermore the DC bias is selected such that it minimises the AM to PM conversion in the pulse generator.



Graph 1. Phase shift as a function of pulse generator input power at 3610 MHz

Graph 1 illustrates this effect. Normally a variation of the input power introduces a phase shift. By selecting a proper bias current, in this case approximately 4 mA, this effect is almost cancelled.

5.4 Construction

Two similar frequency synthesizers, one for the frequency range from 1.2 GHz to 2.2 GHz while the other one covers the range from 2.2 GHz to 9.6 GHz, are housed in a single 19" Eurocard rack.

Each synthesizer consists of three small plug-in modules, one for the control loop, one for the low frequency and digital parts of the main loop and a third one for the microwave components of the main loop. The first two units have solid aluminium housing. Removable cover plates are mounted with conductive rubber gaskets, while all power supply and control lines entering the unit are filtered. This construction ensures a mechanically stable environment with very good EMC shielding.

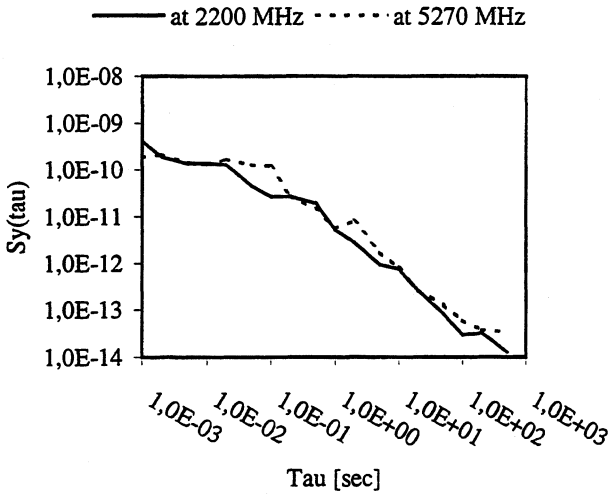
The microwave module consists of an open chassis on which a.o. the YTO and sampling mixer are mounted. Interconnections between the microwave components are made of semi rigid cable.

The interconnections between modules and backplane of the 19" rack use floating mounted D-connectors for supply and control lines. RF connections are made through floating, coaxial connectors of the BMA type (equivalent to OSP).

The 19" rack has an active heating system that keeps all modules at a temperature of approximately 45° Celsius. Temperature variation is kept within 2° Celsius over an outside temperature range of 15 to 40° Celsius.

6. PERFORMANCE

Limited long term data is yet available on the performance of the single synthesizer separately.



Graph 2. Allan standard deviation of synthesizers

Graph 2 shows the Allan standard deviation for low and high band synthesizers at respectively 2200 MHz and 5270 MHz.

From interferometric measurements at the WSRT it is concluded that the phase drift over a twelve hour period at 5 GHz is less then 5° rms, averaged over 60 seconds.

Other specifications are in agreement with the values given in paragraph 2.

7. CONCLUSIONS

A description has been given of a microwave synthesizer system that is used in an interferometric radio astronomical receiver system. The application calls for specific requirements with respect to long term phase and frequency stability. The presented system operates for more then a year and has shown satisfactory performance. Limited data of the measured performance on single synthesizers, which support the proper operation, have been presented.

In the near future more work will be done to characterise the long-term stability of the synthesizer and that of the individual components in it.

Furthermore plans exist to investigate the applicability of the design to a new generation of radio telescopes [9]. Among others for this a further miniaturisation and low power consumption is of importance.

8. ACKNOWLEDGMENT

The author would like to express his gratitude to Mr. J. Roosjen and Mr. H. Heutink, both of NFRA. Mr. Roosjen played an active roll in the development of the synthesizer while Mr. Heutink participated in the design and construction of the system and did most of the time consuming Allan-variance measurements.

The Dwingeloo Radio Observatory is operated by the Netherlands Foundation for Research in Astronomy (NFRA) with the financial support of the Netherlands Organisation for Scientific Research (NWO).

9. REFERENCES

[1] G.H. Tan, "Upgrade of the Westerbork Synthesis Radio Telescope: The Multi Frequency Front End", in Proceedings of the 26th European Microwave Conference, 1996, pp. 59-64.

[2] G.H. Tan, "A novel front end for the Westerbork Synthesis Radio Telescope", in Proceedings of the 9th Military Microwave Conference, 1992, pp. 329-335.

[3] J.W.M. Baars et al, "The Synthesis Radio Telescope at Westerbork", Proceedings of the IEEE, vol. 61, pp. 1258-1266, September 1973.

[4] E.B. Fomalont, "Earth-Rotation Aperture Synthesis", Proceedings of the IEEE, vol. 61, pp. 1211-1218, September 1973.

[5] A.R. Thompson, J.M. Moran, G.W. Swenson Jr., Interferometry and Synthesis in Radio Astronomy, New York: John Wiley & Sons Inc., 1986, ch. 9, pp. 247-313.

[6] J.A. Crawford, Frequency Synthesizer Design Handbook, Norwood: Artech House Inc., 1994, ch. 3.6.1, pp. 71-81.

[7] J. Merkelo, R.D. Hall, "Broad-band thin-film signal sampler", IEEE Journal of Solid-State Circuits, vol. SC-7, no. 1, pp. 50-54.

[8] J.A. Crawford, Frequency Synthesizer Design Handbook, Norwood: Artech House Inc., 1994, ch. 3.6.1, pp. 71-81.

[9] A. van Ardenne, F.M.A. Smits, "Technical aspects for the Square Kilometer Interferometer", in Proceedings of the ESA/ESTEC Workshop on Large Antennas in Radio Astronomy WPP-110, 1996, pp. 117-128.

THE ADIABATIC ANTI- JITTER CIRCUIT

M.J. Underhill
School of Electronic Engineering, Information Technology and Mathematics, University of Surrey, Guildford, Surrey, GU2 5XH, UK

ABSTRACT

The Anti-Jitter Circuit (AJC) [1,2] uniquely is able to reduce phase noise of any frequency source at sideband frequencies above a defined cut-off frequency. By contrast a Phase Lock Loop (PLL) reduces phase noise closer to carrier lower than a defined frequency. The purpose of this paper is to report several further improvements in the implementation of the AJC in respect of power consumption, frequency range and maximum frequency of operation. The improvements stem mainly from the invention of the Adiabatic AJC [6], which affords significantly enhanced performance over the AJC previously reported at the EFTF[3,4,5]. The term "adiabatic" is adopted to indicate that the core part of the new circuit does not require a power supply. It takes power from the input source directly to create the sawtooth waveform that has considerably reduced time jitter on the longer of its two ramp waveforms. This paper also reports work that FET (CMOS) technology in general is to be preferred to bipolar technology for the AJC. Discrete models are now operational at 30Mhz, which is twice the 15Mhz operation previously reported. The cut-off frequency of suppression has been maintained at a few kHz. Noise analysis now shows performance comparable to an LC oscillator is possible. SPICE simulations show potential operation up to 5 GHz The AAJC is also cascadable up to the intrinsic shot noise limit. Shot noise can be reduced by feedback.

1. INTRODUCTION

The Anti-Jitter Circuit is a unique way of reducing wider band phase noise or time jitter on any frequency source. The basic principle of the Anti-Jitter circuit (AJC) was first announced in the 1996 EFTF [1]. Subsequent EFTF papers[2-4] have described AJC performance obtained using discrete components and the application of these to obtain improvements in Direct Digital Synthesis (DDS) and PLL (Phase Lock Loop) fractional-n frequency synthesisers.

The AJC is ideally suited for application to DDS. Uniquely it can suppress the wideband spurs that are inevitable with DDS with a limited bit precision for phase or waveform amplitude definition. Below the AJC cut off frequency the spectrum is unaffected andthe good close-to-carrier performance of DDS can be retained.

In the feedback path of a fractional-n PLL synthesiser the fractional n-spurs can be suppressed above the AJC DC removal cut-off frequency whilst retaining a much higher PLL loop cut-off and reference frequency. In this way the requirements on the fractional n-spur cancellation process can be much reduced if not altogether eliminated. It has been found that the AJC apparently suppresses component at frequencies above the loop bandwidth. This is because higher order Bessel function phase modulation components decay at 12 dB per octave away from carrier. Widespread application of the AJC requires a higher frequency of operation and manufacturability in integrated circuit form as well as the best possible noise reduction. It is also highly desirable that the AJC should operate over as wide a frequency range as possible with little or no adjustment of circuit component values.

The first AJC implementations used bipolar transistors for the active devices. However for time constants long enough to give a low enough frequency of sideband suppression with minimal on-chip capacitor area, FET IC technology is to be preferred.

2. THE AJC PRINCIPLE

The Anti-Jitter Circuit combines several operations into one simple circuit arrangement. It first frequency demodulates an input waveform to baseband. Then an integrator converts the baseband signal to give a phase demodulation signal. Next a time delay modulation process operates using the phase-demodulated signal to cancel the time jitter on one of the transition edges of the original input signal. The comparator that is part of the time delay modulator then delivers transitions of one sense with substantially no phase jitter and with the phase jitter of the opposite transitions approximately doubled. An edge triggered monostable or divide- by-two flip-flop can then operate on the low jitter transitions to produce the low phase noise output at the original, or half the original frequency.

Fig. 1 shows the operation of any Anti-Jitter Circuit. The input frequency waveform is converted into short pulses of constant area. The DC component is removed usually by DC feedback around the integrator rather than by DC blocking in a forward path as shown. The integrator then converts the input pulse train into a sawtooth waveform with a constant mean DC level. The comparator is arranged to switch approximately at this mean DC level (op3). Fig 1(c) shows that time jitter on a pulse as shown in Fig 1(b) does not affect the comparator switching time (at the mean DC level) of the long downslope ramp section of the sawtooth waveform, (op2). Thus an output monostable triggered at the time of this intersection will have much reduced time jitter and hence much reduced phase noise.

0-7803-5400-1/99/ $ 10.00

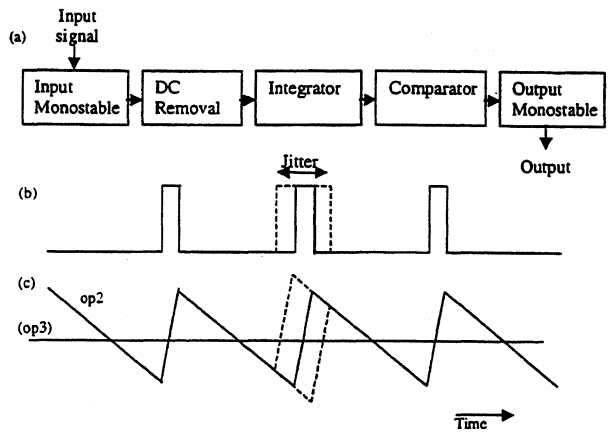


Figure 1. Anti Jitter Circuit Principle: -
(a) Basic Block Diagram
(b) Input monostable with jitter on central pulse
(c) Integrator output (op2) and Comparator switching level (op3)

Conveniently the output monostable can be arranged to provide approximately an equal mark - space ratio at the output frequency and if necessary can be made to maintain this over the operating frequency range. Alternatively pulses can be formed on both edges of the input frequency waveform. The AJC then operates at double the input frequency and a divide-by two circuit in place of the output monostable can then provide an equal mark-space ratio output waveform over the whole frequency range of operation.

The lowest sideband frequency of suppression of jitter or phase noise is determined by the dominant time constant or cut-off frequency of the DC removal part of the circuit.

3. THE ADIABATIC AJC PRINCIPLE

The Adiabatic AJC (AAJC) is a considerable simplification over the original AJC. This leads to a much higher potential frequency of operation and a lower power consumption. In fact the AAJC does not require a DC power source at all. Because it takes what power is required from the RF input signal the term "adiabatic" is used to signify the power conservation taking place.

3.1 AAJC basic operation

Fig. 2 shows a basic Adiabatic AJC circuit. This simple circuit combines the functions of the input pulse monostable, integrator, and DC removal circuit, but does not include the comparator or output monostable. Also shown are optional fast acquisition diodes, which only come into play at switch on or after a very large frequency jump.

In this circuit the input sinusoid or square waveform is converted into short almost constant charge pulses by what is effectively a half wave rectifier circuit made up from the input capacitor *C1* and the two diodes D1 and D2. The integrator storage capacitor becomes charged to nearly the peak amplitude of the input waveform. The n-channel FET Q1 discharges the integrator capacitor *C3* at a constant rate to create a sawtooth voltage waveform on the capacitor.

The charge and discharge action on the integrated capacitor produces the sawtooth voltage waveform required for AJC operation.

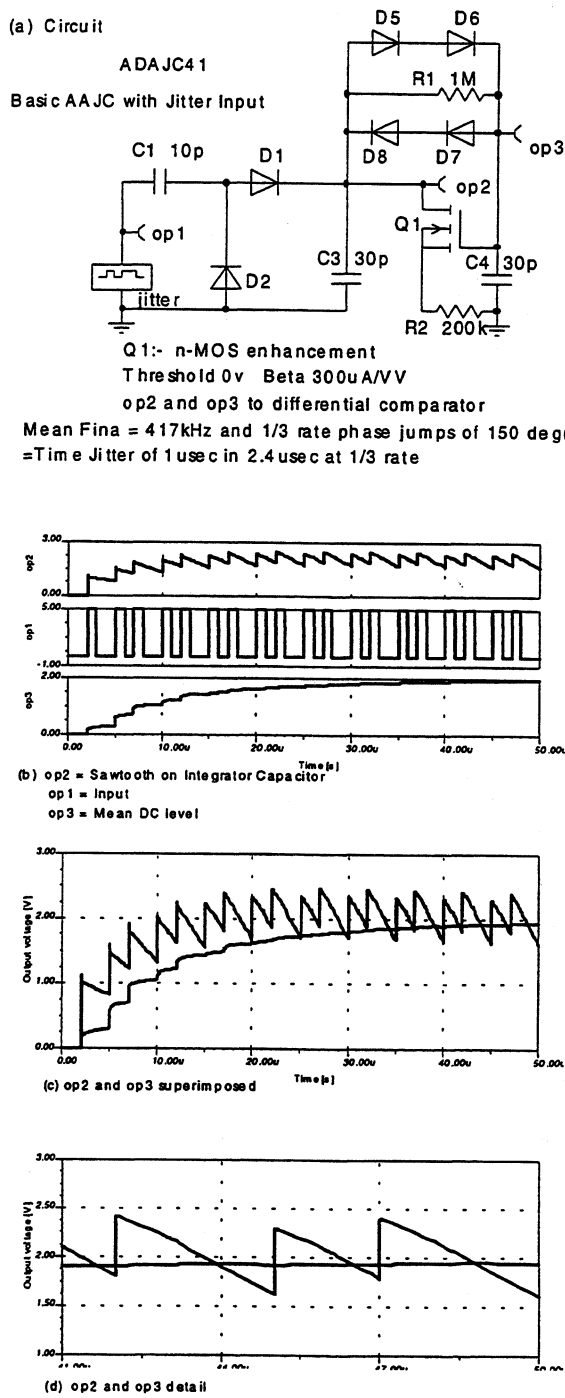


Fig 2:- (a) Basic Adiabatic AJC (AAJC) and Simulation Results (b) to (d)

It can be observed that the correct switching level on the longer slope (downslope) part of the sawtooth

waveform is in fact the mean DC level. This means that the mean DC level on the gate of Q1 is the correct comparator switching reference level. Thus a differential comparator can be connected to compare the outputs op2 and op3 to give correct AJC comparator operation.

A further advantage of this connection is that correct comparator switching is obtained for any input frequency up to a maximum after a short settling time.

The four optional "speed up" diodes shown across the resistor R1 provide a low impedance path from input to the output shunt capacitor C4 of the low pass filter if the positive or negative voltage exceeds Vbe per diode. This option lowers the time constant of the low pass filter by orders of magnitude during initial acquisition of lock of output signal to input signal, or if large frequency or phase jump deviations occur in the input signal. The time for initial acquisition is thus much reduced. The input to output lock is maintained (with no input pulses missed) over a wider range of input deviations of phase or frequency. The presence of the diodes allows phase jitter sideband components much closer to carrier to be better suppressed after a full settling has occurred.

3.2 DC removal feedback

Negative feedback of the mean DC level of the sawtooth waveform to the gate of the FET occurs through the low pass action of the time constant $R1C3C4/(C3+C4)$. The time constant determines the sideband frequency below which the jitter suppression starts to degrade at a 6dB per octave rate.

The optimum loop gain for the negative feedback is $g_mR1 = (C3 + C4)^2/C3C4$. But the loop gain is not critical and can vary over at least a four to one range without serious degradation of the transient response.

3.3 AAJC frequency range

For a FET we have $g_m = \sqrt{(2I_{dis}\beta)}$ where I_{dis} is the discharge current through the FET. β depends on the n-channel FET width to length ratio w/l and is typically $75w/l\mu A/V^2$ But we have $I_{dis} = f_{ina}\ V_{ppst}C3$, where f_{ina} is the input frequency and V_{ppst} is the peak to peak sawtooth voltage on the storage/integrator capacitor $C3$. I_{dis} is therefore proportional to input frequency. The consequence is that the loop gain varies as the square root of input frequency. For such a control loop the loop gain can typically be allowed to vary by up to four to one with little variation in overall settling time or loop bandwidth, giving a working frequency range of sixteen to one with no changes in component values.

To extend the frequency range further, simulation shows that variation of $R1$ only can give operation over a thousand to one frequency range. To obtain the highest frequency of overall operation the input capacitor $C1$ and the storage capacitor $C3$ should also be reduced as far as possible.

3.4 AAJC suppression

A disadvantage of the simple diode charge pump as shown is that the charge packets depend on the voltage on the integrator storage capacitor. To obtain the best jitter reduction it is advisable to keep the peak to peak sawtooth voltage as a small percentage of the mean voltage.

Also for best jitter suppression the input capacitor $C1$ should be smaller than the storage capacitor $C3$.Stray capacitance at this point can prove to be the limiting factor for the highest frequency of operation in practice.

However it is possible to insert a grounded gate series p-channel FET [7] to improve the constancy of the charge packets going into the integrator/storage capacitor. In simulation this improves the suppression per AAJC stage to the 25dB level from typically 20dB. The series FET is a major limiting factor on frequency response.

3.5 Frequency doubling AAJC

A frequency doubling circuit can be implemented in a very simple way with the Adiabatic AJC, as shown in Fig.3. Two input charge pumps operate alternatively on the rising and falling edges of the input waveform. Essentially it is a full wave rectifier circuit. The transformer XMR would normally be replaced by some transformerless push pull active circuit operating on the input signal. Advantages of frequency doubling and then dividing to obtain the final output are a further small improvement jitter and an equal output mark space ratio over the whole frequency range of operation.

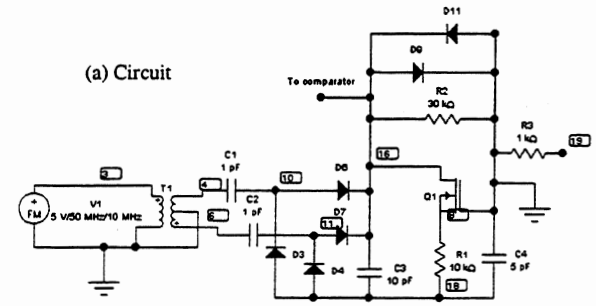
3.6 AAJC Noise

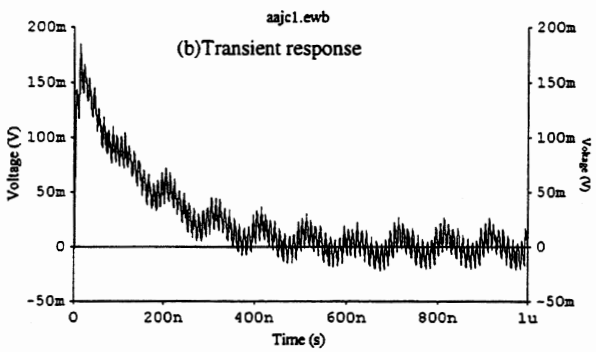
For the AAJC the only dominant source of noise is shot noise in the FET current source (sink). Previous analysis [3] has shown that this gives a phase noise about 20-30dB worse than a varactor tuned LC oscillator (Q=100).

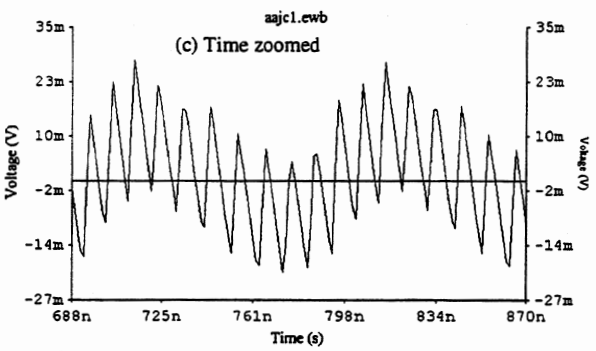
However the FET source resistor $R1$ can reduce FET noise. $R1$ provides negative feedback that reduces the noise approximately by the loop gain g_mR1. $R1$ can be made larger for a given discharge current provided that a larger amplitude input signal is used. A large width to length ratio FET provides a high g_m but at the cost of high capacitance and reduced frequency response. Further investigations are taking place with the objective of achieving a loop gain of 100 for a 40dB reduction of shot noise. This should give the AAJC a theoretical performance of an oscillator of a Q of at least 100 for the same power requirement

4. SIMULATIONS AND RESULTS

Available discrete FETs do not allow optimised AJC designs; capacitances, gains and offset voltages are too high. In practice it is not possible to design a (mixed signal) integrated circuit without extensive simulation for testing and optimisation of the design. For these reasons the AAJC has been simulated in two different SPICE based packages, TINA and Electronics Workbench.







Figs 3(a to c) Frequency-doubling AAJC and transient responses.

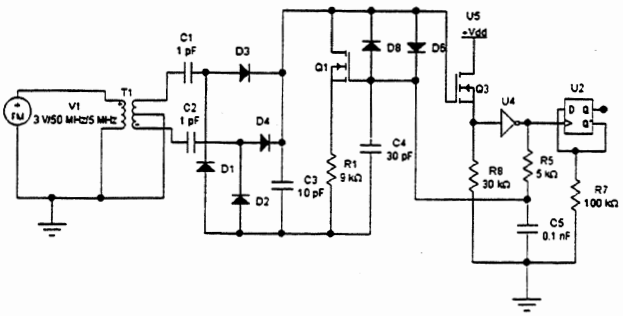
Figs 2(b to d) show TINA simulated time responses for the basic AAJC with an input that has alternating 150° phase jumps. The downslope intersection shows that the jitter has been reduced (≈ 15dB in this case).

The frequency doubling AAJC in Fig 3 and the doubler/divider AAJC in Fig 4, both have sinusoidal FM of a 50MHz input, with m=0.5 and 1.5 and are simulated with Electronics Workbench.

Other TINA simulations have shown AAJC operation at 50GHz [7].

5. CONCLUSIONS

In conclusion the Adiabatic Anti-Jitter Circuit is a considerable improvement over the original AJC implementations. It will therefore increase the range of applications for this unique type of circuit.



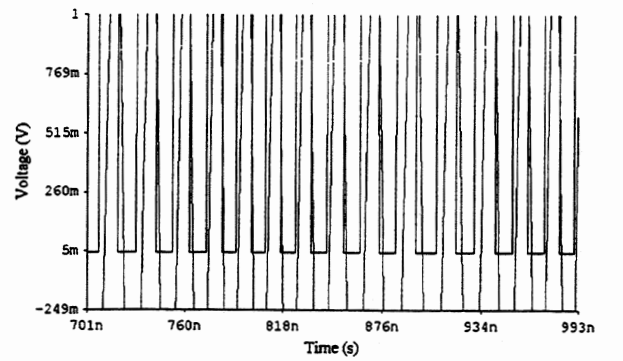


Fig3 Doubler/divider AAJC with input (upper) and output (lower)waveforms compared for FM index of m = 1.5.

6. REFERENCES

[1] M.J.Underhill, "Phase Noise Reduction Circuits". European Pat App No. 97903456.8.

[2] M.J.Underhill and N.Downie – 3 European Pat Applns on AJC improvements.

[3] M.J.Underhill and M.J.Blewett, "Spectral improvement of direct digital frequency synthesisers and other frequency sources". Proc. 10th EFTF, Brighton, March 1996, pp. 452-460 (IEE Conf. Pub. 418).

[4] M.J.Underhill and M.J.Blewett, "Performance of a delay compensation phase noise and time jitter reduction method". Proc. 11th EFTF, Neuchatel, March 1997, pp. 364-368.

[5] M.J.Underhill, S.Stavrou, M.J.Blewett and N.Downie "The Anti Jitter Circuit for low spurious DDS square waves and low cost fractional n synthesis". Proc 12th EFTF, Warsaw, March 1998, pp. 292-297.

[6] M.J.Underhill, 2 Provisional Patent Applications for the Adiabatic Anti-Jitter Circuit.

[7] M.J.Underhill, "The anti-jitter circuit for the suppression of wideband phase noise", IEE colloquium 98/480 on Microwave & Millimetre Wave Oscillators and Mixers, 1st Dec 1998

Cs FREQUENCY SYNTHESIS: A NEW APPROACH

A. Sen Gupta+, D. Popovic#, and F. L. Walls*

+National Physical Laboratory, New Delhi, India
#University of Belgrade, Belgrade, Yugoslavia
*National Institute of Standards and Technology, Boulder CO, USA

ABSTRACT

This paper describes a new approach to synthesizing the Cs hyperfine frequency of 9.192 GHz that is designed to be sufficiently rugged for use in space, specifically for the primary atomic reference clock in space (PARCS) planned for the international space station, as well as ground applications. This new approach requires no narrow band filters or frequency multiplication and the primary source of cooling is conduction. Instead of frequency multiplication, it uses a custom regenerative divider stage followed by two commercial binary dividers and several mixing stages. A fractional frequency step of 2×10^{-17} is achieved by mixing the output of a 48 bit numerically controlled oscillator with the microwave signal. Preliminary tests on the new synthesizer design indicates an internal fractional frequency stability of 1×10^{-15} at 10 s and 1×10^{-18} at 1 day, dominated by the daily room temperature variations. The phase and amplitude noise are similar to our previous designs that used frequency multiplication and narrow band filters. The temperature coefficient is less than 0.2 ps/ ^{o}C.

1. INTRODUCTION

Within the last decade new generations of frequency standards based on laser cooled atoms and ions with stabilities approaching 10^{-16} at 1 day have been developed [1-3]. At theses stabilities, the performance of the local oscillator and the microwave synthesis chain appear to be limiting factors. To support NIST's new standards, we have designed a new generation of microwave frequency synthesizers [4-6]. The basic criteria for these designs have been high resolution, low phase modulation (PM) and amplitude modulation (AM) noise, low spurs near the carrier, and very high phase stability with respect to environmental effects.

The aim of the present work was to design a microwave synthesizer for a laser cooled Cs fountain compatible for use in space, specifically for a slow beam Cs primary atomic reference clock in space (PARCS) scheduled for flight the international space station [7]. This required the use of components that are or can be space-qualified. Other major design goals were that the synthesizer be compact, robust, consume low power, and very importantly, that all dissipated heat be removed by conduction and radiation as there is minimal convection in space. The performance goal was a frequency resolution of at least 10^{-15} and a fractional frequency stability at 1 day of lower than 1×10^{-16}.

In the following section we describe our design. The approach yields a simple synthesis scheme that avoids frequency multiplication, narrow band filters, and phase-locked loops while still producing an output that is settable with high resolution. We also discuss some initial performance results of the two units that were constructed. All performance goals were met or exceeded.

2. DESIGN DETAILS

Figure 1 shows a schematic block diagram of the synthesizer. The output of a 6.4 GHz voltage-controlled dielectric resonant oscillator (DRO) drives a custom regenerative divide-by-2 circuit [8], comprised of a mixer, a 3.2 GHz low-pass filter, and an amplifier in a closed-loop configuration. This regenerative divider provides low PM noise [9], a very low coefficient of phase shift with temperature, and simultaneous outputs at 0.5 and 1.5 times the input frequency, i.e., at 3.2 GHz and 9.6 GHz. The 3.2 GHz and 9.6 GHz outputs are separated using a diplexer.

The 3.2 GHz output is divided to 100 MHz using a commercial divide-by-8 followed by a commercial divide-by-4 [10]. The 100 MHz output of the dividers is low-pass filtered and buffered in a two-stage amplifier to provide the output reference frequency.

Another key constituent of the synthesizer is a numerically controlled oscillator using a direct digital synthesis (DDS) chip with 48 bit resolution [11]. This unit is clocked by a 50 MHz signal obtained by dividing the 100 MHz by 2. This DDS chip not only provides the high resolution and low spurs, but it is also available in a space-qualified version. The DDS output is a sine wave, which can be set to a resolution of 0.2 μHz and a rate up to 20 MHz using 14 data and control lines. Another feature of this DDS is that its frequency

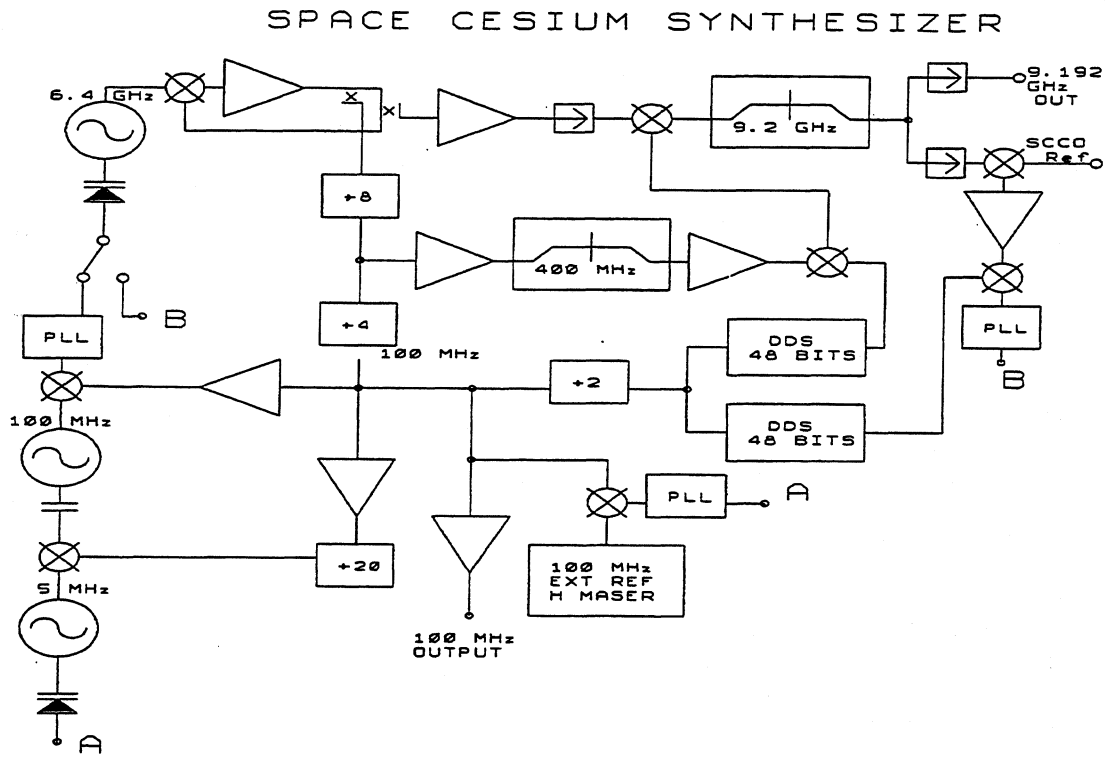


Figure 1. Block diagram of the new Cs frequency synthesizer.

can be switched to a new value without phase discontinuity in less than 1 μs. The frequency is nominally set at 7.368 MHz and mixed with the 400 MHz from the binary divider using a simple upper sideband (USB) mixer, which gives more than 25 dB rejection at the carrier and lower sidebands. Additional suppression of these spurs could be obtained with further design refinements. The 407.368 MHz signal is mixed with the 9.6 GHz output of the regenerative divider to produce the 9.192 GHz signal, which is amplified and passed through a broadband filter and isolator to provide the final microwave output of the synthesizer.

As in our earlier synthesizer designs [4, 5], a 5 MHz oscillator controls the PM noise for Fourier frequency offsets up to approximately 50 Hz while a 100 MHz oscillator controls the range from 50 Hz to 50 kHz.

The synthesizer output can be steered by voltage control of the 5 MHz oscillator using a DC error signal. This voltage control could be derived from the physics package of a Cs clock. The 100 MHz output from the synthesizer also drives a mixer, which can be used to detect the phase difference between the 100 MHz signal from a hydrogen maser or other high stability reference. This mixer output can be used to phase-lock the 5 MHz to this reference with a loop bandwidth of approximately 1 Hz. The low bandwidth was chosen to exclude spurs due to ground loops and pickup of electromagnetic interference. There is also a provision to phase lock the synthesizer to a microwave reference such as a super conducting cavity oscillator (SCCO) as long as the frequency is within ± 20 MHz of one of the internal reference frequencies (3.2 GHz, 6.4 GHz, 9.192 GHz, 9.6 GHz, 10.007 GHz). To achieve this we use another phase-locked-loop (PLL) with a second DDS set to the difference between the internal reference frequency and that of the SCCO. The output error of this PLL could correct the 6.4 GHz DRO with a bandwidth of up to approximately 300 kHz. It would also be possible to add an additional mixing stage to use a stable reference at the Rb hyperfine frequency of 6.834 GHz [12].

We constructed two prototype synthesizers of the above design. The mechanical assembly consisted of five interconnected modules with machined aluminum cases of 6 mm wall thickness. These were stacked tightly together with heat sink fins on two outer sides of the assembly. All circuitry, including the DDS, used surface mounted components that were heat sunk to the printed circuit boards. The printed circuit boards in turn were heat sunk to the aluminum cases by mounting them flush against the bottom with thermal conductive film. In addition, metal straps connected the top of the few integrated circuits that dissipated significant heat to the sidewalls. The entire microwave portion of the circuit was assembled in one module with the components mounted flush against the walls or the bottom. Thus most of the heat dissipated in the synthesizer had a direct conduction path to the external heat sinks. In

some applications the heat sinks could be removed and the assembly mounted to a thermally stabilized base plate. The unit is very compact and quite robust. The total size of the each unit, excluding the external power supply and the removable heat sinks was 22 cm x 13 cm x 18 cm. The volume could be reduced by roughly 40 % by simply repackaging the present subassemblies. The total power consumption was approximately 22 W at 24°C. Although no quantitative vibration tests have yet been performed, many strong bumps, including the customary 10 cm drop test, produced no noticeable effects.

3. RESULTS AND DISCUSSION

In this section we describe results of tests made on the two units described above. Extensive studies were made of the spurs and of the PM and AM noise.

Figure 2 shows that there are no spurs within ±2.5 MHz of the 9.192 GHz output greater than -75 dBc. PM noise measurements were performed on the 100 MHz output using a three-cornered-hat method involving two of the present synthesizers and one of the earlier types [4]. The results are summarized in Fig 3. Also shown for comparison are the results from of earlier designs [4, 5]. The results of measurements of AM noise and spurs at the microwave output are also very similar to those obtained with the earlier NIST synthesizers. These results indicate that frequency pulling by spectral impurities is of little consequence to present slow beam or fountain frequency standards [4].

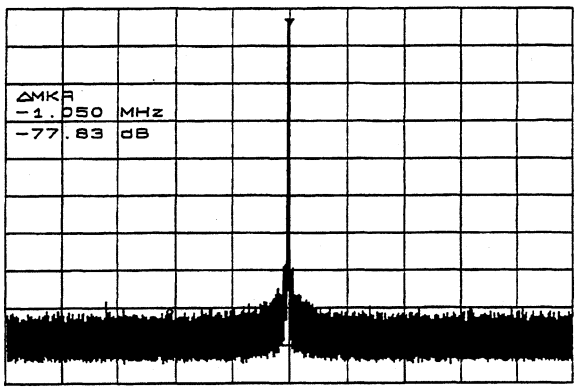


Figure 2. Microwave power spectrum of the 9.192 GHz output from the new Cs synthesizer. The span is 5 MHz and the resolution bandwidth is 1 kHz.

The set up shown in Fig 4 was used to study the phase stability with respect to environmental temperature variation between the microwave output and the user output at 100 MHz. The microwave outputs of the two synthesizers were phase locked and the 100 MHz outputs were used to drive a phase comparator and data

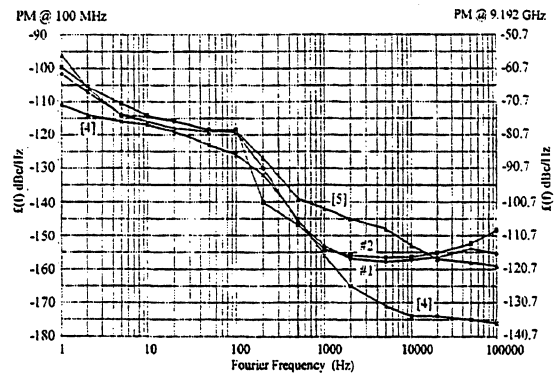


Figure 3. Comparison of the PM noise spectrum of the new Cs synthesizer #1 and #2 to previous designs [4] and [5].

logger. The temperature of one of the synthesizers was kept constant while that of the other was varied. Figures 5a and 5b show typical initial result of such temperature cycling from which we compute a temperature coefficient of 2.2 ps/°C. An important

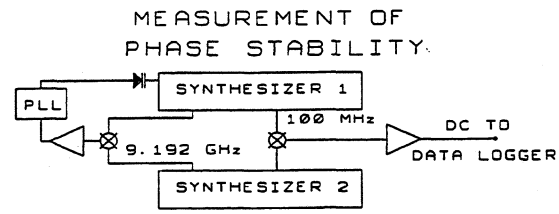


Figure 4. Block diagram of setup used to measure the phase stability.

point was that the temperature coefficient was almost identical for the two units and that there was less than a minute delay and almost no hysteresis between the temperature and phase variations. The reason for this is that all the components of are thermally very well connected to the case. Further investigations revealed that the major part of the temperature coefficient arose from the last frequency divider and the output isolation amplifier. A very simple circuit comprised of a thermistor (heat sunk to the inside of the case containing the divider and the output amplifier) and a varactor to change the phase delay at the output of the 100 MHz in response to the temperature variations was successfully used to compensate most of the temperature coefficient. With several trials of scaling the variation of the voltage across the varactor, we were able to obtain a temperature coefficient of less than 0.2 ps/K as shown in Fig. 5c. In subsequent tests we obtained a very highly phase stable 100 MHz output as shown for a 4 day stretch during which room temperature varied by approximately ± 0.5°C (see Fig. 6). Finally, we show in Fig 7 a typical plot of internal

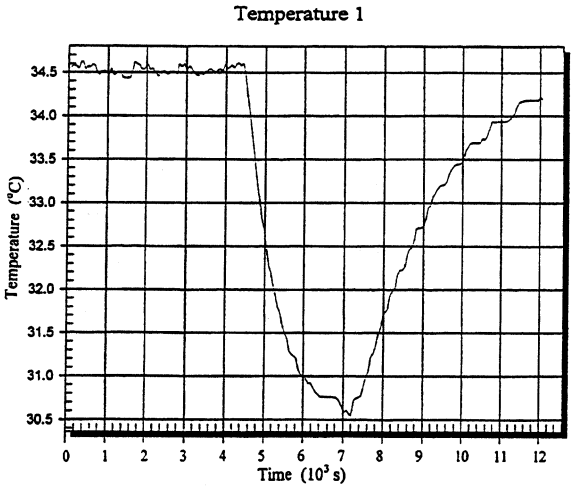


Figure 5a. Temperature change of synthesizer # 1.

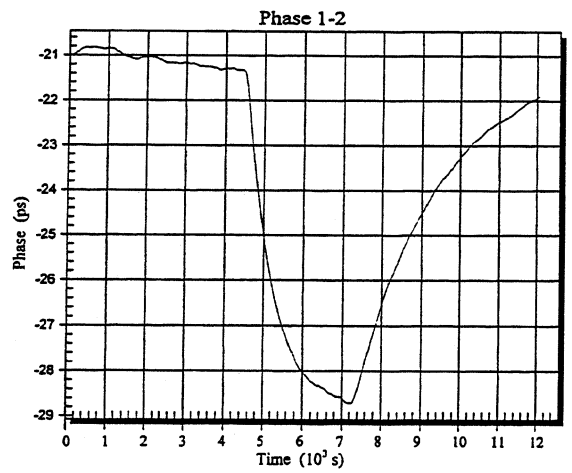


Figure 5b. Initial uncompensated phase response of synthesizer #1.

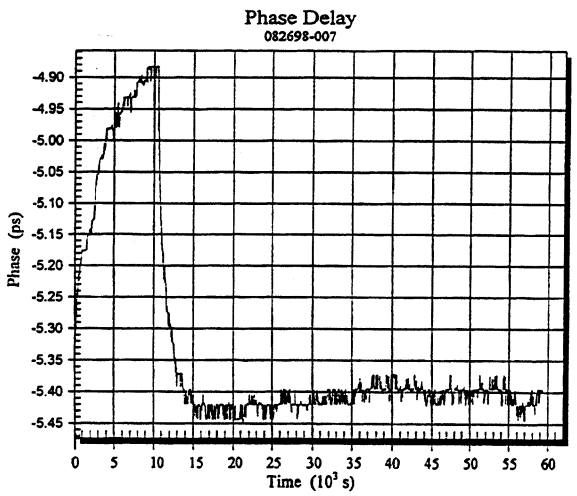


Figure 5c. Compensated phase response of synthesizer #2 for a temperature stepup and back of 3.5 °C.

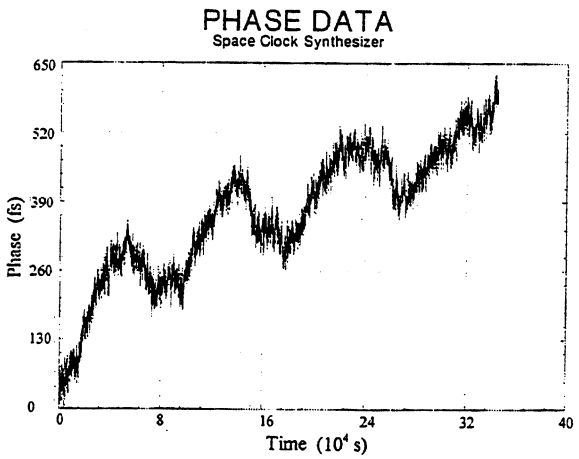


Figure 6. Phase change between two synthesizers over 4 days.

frequency stability, obtained by processing the phase data with sampling times from 5 s up to 1 d. We obtain a fractional frequency stability Allan deviation of 1 x 10^{-15} at 10 s and 1 x 10^{-18} at 1 d. The 1 d frequency stability is a significant improvement over stability results reported for previous designs.

4. CONCLUSION

We have described a novel approach for synthesizing the microwave signal for a Cs standard. The realization of the design is quite simple, requiring no frequency multiplication or narrow band filters. Two prototype synthesizers were fabricated keeping in mind the proposed use in space and hence taking special care to ensure that all the components could eventually be obtained in space qualified versions. We also ensured that most components were heat sunk by thermal conduction to the case. This led to a very simple but effective means of compensating the temperature coefficient to lower than 0.2 ps/ ^{o}C, and consequently obtaining internal frequency stability that reached 1 x 10^{-18} at 1 d. Another feature of the present design, the ability to set the microwave output with a resolution of 2 x 10^{-17} is due to the use of a 48 bit DDS. Since there is no PLL in the final microwave synthesis, the phase excursions after switching frequency are small and the settling time is less than 1 μs. Yet another feature of the synthesizer is the provision for phase locking to an SCCO which is trimmed within 20 MHz of one of the internal reference frequencies (3.2 GHz, 6.4 GHz, 9.192 GHz, 9.6 GHz, 10.007 GHz), providing for a much superior PM noise and short-term frequency stability than is possible using quartz oscillators.

Although the present synthesizers are already very robust, compact, and consume low power, we

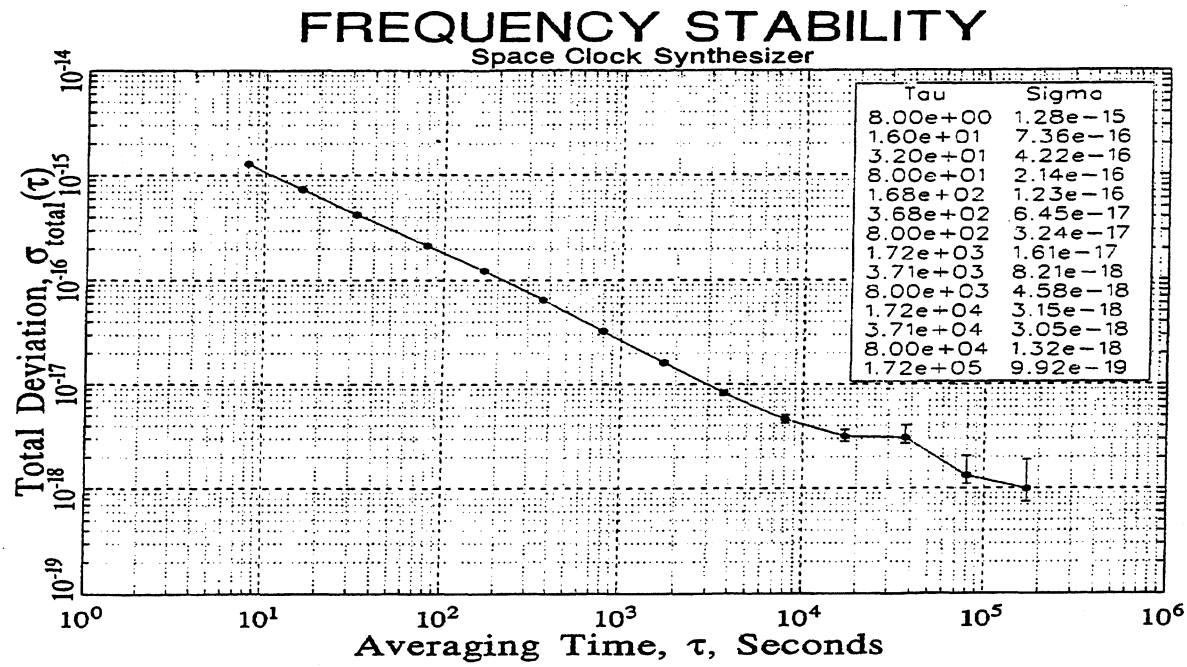


Figure 7. Allan deviation for a pair of Cs synthesizers computed using total deviation and the data of Figure 6.

estimate that further optimization should lead to a 70 % reduction in size and a 50 % reduction in power consumption. Finally, although our present focus has been to synthesize the Cs hyperfine frequency, this approach could be adapted to produce similar performance at many other frequencies of interest (for example, the Rb frequency).

5. ACKNOWLEDGEMENTS

We gratefully acknowledge helpful discussions with J.F. Garcia Nava, C. Nelson, S. R. Jefferts, and D. B. Sullivan.

6. REFERENCES

1. A. Clairon, P. Laurent, G. Santarelli, S. Ghezali,, S. N. Lea and M. Bahoura, " A Cesium Fountain Frequency Standard: Preliminary results," IEEE Trans I&M, pp. 128-131, 1995.
2. D.J. Wineland, J.C.Bergquist, J. Bollinger, W.M. Itano, D. J. Heinzen, S. L. Gilbert, C. H. Maney, and M. G. Raizen, "Progress at NIST towards Absolute Frequency Standards Using Stored Ions," IEEE Trans. UFFC, 37, pp. 515-523, 1990.
3. J. D. Prestage, R. L. Tjoelker, R. T. Wang, G. J. Dick and L. Maleki, "Hg^+ Trapped Ion Standard with Superconducting Cavity Maser Oscillator," IEEE Trans. UFFC, 42, pp. 200-205, 1993.
4. J. F. Garcia Nava, F. L. Walls, J. H. Shirley, W. D. Lee and M. C. Delgado Aramburo, "Environmental Effects in Frequency Synthesizers for Passive Frequency Standards," Proc. of 1996 IEEE Freq. Control Symp., pp. 973-979, 1993.
5. G. Rovera, G. Santarelli and A. Clairon, "Frequency Synthesis Chain for the Atomic Fountain Primary Frequency Standard," IEEE Trans. UFFC, 43, No 3, pp. 354-358, 1996.
6. R. F. Karlquist, "A New RF Architecture for Cesium Frequency Standards," Proc. 1992 IEEE Freq. Control Symp., pp. 134-142, 1992.
7. N. Ashby, A. Demarchi, T. Heavner, L. Hollberg, S. Jefferts, J. Kitching, T. Parker, W. Phillips, S. Rolston, H. Robinson, D. Sullivan and F. Walls, "PARCS: A Primary Atomic Reference Clock in Space," these proceedings.
8. R. G. Harrison, "Theory of Regenerative Frequency Divider using Double Balanced Mixers," 1989 IEEE MTT-S Int. Microwave Symp. Digest, pp. 459-462, 1989.
9. E. S. Ferre-Pikal and F. L. Walls, "Microwave Regenerative Frequency Dividers with Low Phase Noise," Proc. IEEE Trans. UFFC, 46, No 1, pp. 216-219, 1999.
10. The Dividers were a Sciteq 1208 Divide-by-8 and a Plessey 8402 divide-by-4. These are identified only for completeness and do not represent an endorsement by NIST.
11. The DDS was a Stanford Telecom STEL 1173. This part is identified only for completeness and does not represent an endorsement by NIST.
12. J.Q. Deng, A. De Marchi, F.L. Walls, and R.E. Drullinger, "Reducing the Effects of Local Oscillator Noise on the Frequency Stability of Cell-Based-Passive-Frequency Standards," Proc. 1998 Intl. Freq. Cont. Symp. pp. 95-98, 1998.

THE MEASUREMENT OF LARGE OPTICAL FREQUENCY DIFFERENCES AND THE DESIGN OF A NEW TYPE OF FREQUENCY CHAIN

Thomas Udem*, Jörg Reichert, Ronald Holzwarth, and Theodor Hänsch
Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Str.1 85740 Garching, Germany and
Motonobu Kourogi
Tokyo Institute of Technology, Midori-ku, Yokohama, Kanagawa 226, Japan

ABSTRACT

We have used optical frequency interval dividers, optical frequency comb generators and mode-locked lasers to measure large optical frequency differences of up to 45.2 THz between laser frequencies. We have shown that the modes of a mode-locked laser are distributed uniformly in frequency space within the experimental limit of 3.0 parts in 10^{17} and that the mode separation equals the pulse repetition rate within the experimental limit of 6.0 parts in 10^{16}. We applied these techniques to determine the absolute frequencies of the cesium D_1 line at 335 THz (895 nm), the hydrogen 1S–2S transition at 2466 THz (121 nm) and the frequency of a methane stabilized He–Ne Laser at 88.4 THz (3.39μm) as well as the isotope shift of the 1S–2S frequency.

1. INTRODUCTION

Following the dramatic progress made in recent years in the development of optical frequency standards based on trapped ions [1] or narrow atomic resonances such as the hydrogen 1S–2S transition [2], precise and reliable optical to radio frequency conversion has become the missing 'clockwork' for the construction of a future optical clock. These clocks possess a potential accuracy far beyond the accuracy of the current state of the art cesium clocks. In the past, phase coherent comparisons between optical and radio frequencies have been performed with harmonic frequency chains which create successive harmonics from a well known radio frequency provided for example by a cesium atomic clock. Only a few frequency chains that reach all the way up to the visible or the UV have been built so far [2, 3, 4]. The use of broad frequency combs as provided by optical frequency comb generators [5, 6] or mode-locked lasers in combination with optical frequency interval dividers [7, 8] promises to provide a compact and portable optical clockwork [7, 8, 9] that has the capability to operate reliably for an extended period of time. A phase-locked optical frequency interval divider can reduce an arbitrarily large frequency difference between two input lasers by locking a third laser to the precise center of the gap. If n divider stages are used in cascade a large optical frequency gap can be reduced by a factor of 2^n until it becomes accessible to radio frequency counting techniques. As a novel concept such a chain of divider stages may be used to measure an absolute optical frequency by determining the frequency gap between laser harmonics. In the simplest form this would be the gap between a laser at the frequency f and its second harmonic $2f$: $2f - f = f$. Such a new type of frequency chain allows one to choose a path from the optical region to a beat note in the radio frequency domain without ever leaving the optical region [7, 8]. The frequencies of the laser oscillators may be chosen such that compact and reliable diode lasers can be used exclusively. However, to reach the radio frequency domain in this way a large number of interval divider stages would be necessary so that the reduction in complexity is only due to the fact that it becomes possible to avoid expensive or delicate lasers usually present in harmonic frequency chains. The number of necessary divider stages is reduced substantially if the largest directly measurable beat frequency is increased by the use of an optical frequency comb. To check the accuracy of this new approach we have verified that the modes of an optical frequency comb generator (OFCG) or a mode-locked laser are distributed uniformly in frequency space and that their separation is easily measured with radio frequency precision. We then applied these techniques to measure the absolute frequency of the hydrogen 1S–2S transition, its isotope shift, the cesium D_1 and D_2 line and the frequency of our methane stabilized He–Ne laser.

*Correspondence should be addressed to:
Thomas.Udem@mpq.mpg.de

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

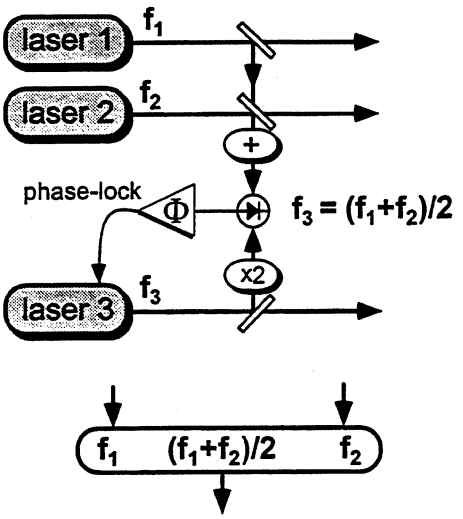


Figure 1: Principle of an optical frequency interval divider. The second harmonic of the third laser $2f_3$ is phase locked to the sum frequency $f_1 + f_2$. The frequency f_3 thus divides the interval $f_1 - f_2$ into two equal sections. The symbol is also used in Fig. 2, 3 and 4.

2. OPTICAL FREQUENCY INTERVAL DIVIDER

As illustrated in Fig. 1, an optical frequency interval divider receives two input laser frequencies f_1 and f_2. The sum frequency $f_1 + f_2$ and the second harmonic of a third laser $2f_3$ are created in a non–linear crystal. The radio frequency beat signal between them at $2f_3 - (f_1 + f_2)$ is either used to phase–lock the third laser at the midpoint $f_3 = (f_1 + f_2)/2$, or to verify that the third laser actually oscillates at the center frequency. Phase–locking of two optical frequencies is achieved electronically by locking the phase of their beat signal to zero or, to reduce $1/f$ noise, to a given offset radio frequency provided by a local oscillator [10]. Techniques of conventional radio frequency phase–locked loops can be applied. Optical frequency interval dividers can be cascaded to divide an arbitrary large frequency gap successively by 2. A four stage interval divider chain, that divides a frequency gap of $\Delta f \approx 1$ THz by 16 has been used in a previous determination of the hydrogen 1S–2S transition frequency [2]. In this set–up diode lasers and single pass $KNbO_3$ crystals have been used to generate the second harmonic and the sum frequencies.

3. OPTICAL FREQUENCY COMB GENERATORS (OFCG)

Our optical frequency comb generator (OFCG) consists of a monolithic $LiNbO_3$ optical resonator that acts at the same time as an electro–optic modulator. To create a spectrally broad comb of side bands a large modulation index and/or a high modulation frequency is required. For efficient high frequency modulation one has to match the phase velocity of the modulating microwave to the group velocity c_g of the light traveling through the crystal by choosing the proper transverse dimensions of the crystal and the proper shape of the microwave cavity in which it is placed [11]. If an integer multiple of half the microwave wavelength equals the length of the crystal and if the velocities are matched, the modulation frequency is an integer multiple of the optical free spectral range $c_g/2L$. In this case the side bands are resonantly enhanced and efficiently create further side bands if one of the modes is locked to the input laser frequency. In our experiments the modulation frequency was set to $f_{OFC} = 6.3418$ GHz, which is twice the free spectral range of the $L = 21$ mm long monolithic resonator. To operate the OFCG in a stable condition one of its optical resonances has to be locked to the input laser frequency. We use a scheme, similar to the Pound–Drever–Hall lock [12], that is discussed along with some variations in Ref. [5]. Because crystal dispersion leads to a wavelength dependence of the free spectral range, the frequency combs generated in this way are limited to less than 8 THz [5]. By the use of self–phase modulation in an optical fiber the width of these frequency combs can be increased externally. Frequency combs as wide as 30 THz have been created this way [13] and beat signals of lasers separated by 12 THz have been observed [14].

4. FREQUENCY COMBS FROM MODE–LOCKED LASERS

The use of mode–locked lasers for the measurement of optical frequency intervals has been demonstrated already 20 years ago with picosecond pulses [15]. As the spectral width of these lasers scales inversely with the pulse duration (for Fourier limited pulses) their application was limited to rather small frequency differences. With the development of Kerr–lens mode–locking [16] femtosecond pulses became readily available. Meanwhile pulses as short as 6.5 fs have been created directly from a Ti:Sapphire laser oscillator [17]. The spectrum of such a laser consists of a comb of laser frequencies. The modes may be thought of being the longitudinal mutually injection locked cw–modes of the laser cavity or as being due to a strong amplitude modulation of some carrier fre-

quency. The pulse repetition rate $f_r = v_g/2L$, which is the inverse round trip time of the pulse circulating in the cavity, is determined from the length of the cavity L and the group velocity $v_g = \partial\omega/\partial k$. Due to dispersive elements in the cavity the group velocity of the circulating pulse does in general not agree with the phase velocity of its carrier. This causes a pulse to pulse phase shift by say $\Delta\varphi$ of the carrier with respect to the envelope [18]. Unlike the pulse envelope the electric field at the output coupler is therefore not a periodic function that repeats itself with f_r. It can not be expressed as a Fourier series containing harmonics of f_r. However, if $\Delta\varphi$ stays constant, we still expect a comb of equidistant modes, spaced by the repetition frequency f_r, but with each frequency displaced from a multiple nf_r by an offset frequency $f_o = \Delta\varphi/2\pi T$, where $T = f_r^{-1}$ is the pulse repetition time. Such a comb of laser frequencies can be used to measure large optical frequency differences. In a recent measurement of the absolute frequency of the hydrogen 1S–2S transition we have used a mode–locked laser to phase–coherently bridge a 45 THz frequency gap [22].

5. THE ACCURACY OF OPTICAL FREQUENCY COMBS

In order to verify that the frequency comb emitted by a mode locked laser can satisfy the exceptional accuracy demands of an all new optical clock, we have performed a series of experiments. In the first experiment [23] we have connected a four stage interval divider chain to two modes of a frequency comb produced by an optical OFCG that were separated by 1.059 THz. The 4 stage interval divider chain divided this gap by 2^4 into a measurable 66 GHz beat signal. We found that both methods to bridge this 1 THz gap agree within the experimental limit of 6.8×10^{-15}. In a second experiment we have phase–locked two laser diodes at 822.8 nm and 870.9 nm to two modes of a mode locked laser (Coherent model Mira 900, measured pulse length 73 fs) separated by more than 20 THz [19]. If the number of modes in between the two diode laser frequencies happens to be an odd number, we expect another mode of the frequency comb right at the center between the two laser diodes. As shown in Fig. 2, this is verified with an optical frequency interval divider. We have measured possible deviations with a radio frequency counter (Hewlett Packard model 53132A). The averages of the frequency readings calculates to -0.59 ± 0.48 mHz. This verifies the uniform distribution of the modes within a relative precision of 3.0×10^{-17}. A slightly modified version of this experiment, which is sketched in Fig. 3, was performed after reducing the frequency difference between the

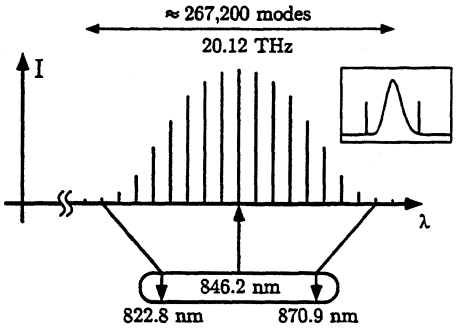


Figure 2: The uniform distribution of the modes of a mode–locked laser is verified by comparison with an optical frequency interval divider. The inset shows a measured spectrum of the frequency comb together with the 822.8 nm and the 870.9 nm laser diodes drawn into it (linear scale).

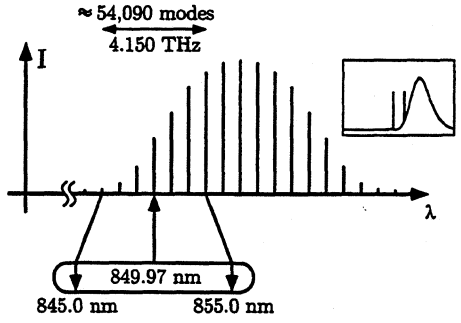


Figure 3: Verifying the uniform distribution of the modes of a mode–locked laser on one side of its spectrum. The inset shows a measured spectrum of the frequency comb together with the 845 nm and the 855 nm laser diodes drawn into it (linear scale).

laser diodes to 4.15 THz and locking them asymmetrically with respect to the spectrum of the mode–locked laser. In this case we found a frequency deviation of -0.70 ± 0.61 mHz. Taking into account the different total number of readings the statistical uncertainty in both cases seem to be identical (see Ref.[19] for further details).

To experimentally demonstrate that optical frequency combs emitted by mode locked lasers are useful tools for the precise determination of large optical frequency differences it is not sufficient to verify the comb spacing constancy. In addition, one has to show that the mode separation equals the pulse repetition rate. Therefore we have performed an actual frequency difference measurement. The set up used for this purpose is similar to the one shown in Fig. 2, but with a reduced frequency gap of 4.1 THz. Two laser diodes, at 845 nm and 855 nm were phase locked to the frequency comb. We used an OFCG

to phase lock the centered laser diode to 849.974 nm and observed the beat note of the divider stage. The pulse repetition rate was locked to 76.5 MHz. We find a frequency deviation from the expected value of 2.2 ± 2.5 mHz. This confirms that the pulse repetition rate equals the mode separation with an accuracy of at least 6.0×10^{-16} [19].

6. OPTICAL FREQUENCY MEASUREMENTS

To demonstrate the new type of optical to microwave frequency chain for the first time we have recently redesigned our existing harmonic frequency chain (see Fig. 4). In the previous set-up only the $\Delta f \approx 1.059$ THz gap was measured with a 4 stage optical frequency interval divider chain [2]. In this way the 1S–2S transition frequency at $28f - 8\Delta f$ was referenced to the absolute frequency of a methane stabilized He–Ne laser [20] at 3.39 μm ($f = 88.4$ THz). By replacing the optical divider stages with an optical frequency comb generator we could switch within a few seconds between the 1S–2S transition frequencies of hydrogen and deuterium. In this case the absolute frequency of the He–Ne standard cancels and the most precise value for the isotope shift and the deuteron structure radius was derived [21].

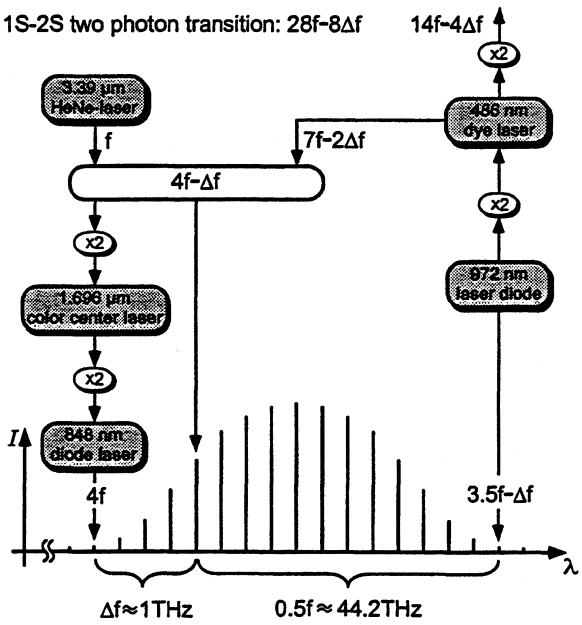


Figure 4: The first frequency chain based on the measurement of large optical frequency gaps. The mode–locked laser determines the frequency differences Δf and $0.5f$. Once these frequencies are known all other frequencies in the chain, including the hydrogen 1S–2S transition frequency $28f - 8\Delta f$, are determined.

In the redesigned version we use a commercial mode–locked laser (Coherent model Mira 900), whose spectrum is externally broadened by coupling 225 mW average power through a 40 cm long single mode quartz fiber. With this laser we measure both, the frequency gap $\Delta f \approx 1.059$ THz and the absolute frequency of the He–Ne standard f, which appears as a frequency interval of the size $0.5f$. As shown in Fig. 4 these gaps are created with several non–linear steps. One of the modes of the mode–locked laser is phase–locked to the 972 nm laser diode which is phase–locked to the dye laser. At the same time the pulse repetition rate of the mode–locked laser is phase–locked to a GPS–referenced Cs clock [24]. Once Δf and f are known all other frequencies in the chain are determined including the hydrogen 1S–2S transition frequency $28f - 8\Delta f$ and the absolute frequency of the methane stabilized He–Ne laser that does not serve as a frequency reference any more. Our new chain is providing the first phase coherent link from the vacuum UV (121 nm) to the radio frequency domain [22]. A preliminary value of the 1S–2S transition frequency, with a considerably smaller uncertainty, derived from a few days of data taking already reproduces our previous measurement [2], that was performed within $2\frac{1}{2}$ months. We expect to have an improved value in the near future after collecting more data. To achieve an even higher accuracy a measurement with a Cs fountain clock [25] is planned. Together with other absolute frequency measurements performed by F. Biraben and coworkers [4], precise values for the 1S Lamb shift and the Rydberg constant can be derived from this measurement. From our early data we find a value for the absolute frequency of the He–Ne standard that is significantly larger by 65 Hz (almost 3 σ) than the value obtained from the last calibration 3 years ago [2]. Because the value of the 1S–2S transition frequency obtained then is consistent with the current value we believe that the frequency of the He–Ne standard did change in the meantime.

Besides the measured frequencies of the hydrogen 1S–2S transition and the He–Ne standard, the new frequency chain of Fig. 4 also provides us with about 600,000 usable reference frequencies, i.e. the modes of the mode–locked laser. This enables us to measure every optical frequency between 848 nm and 972 nm. With a few additional non–linear steps every frequency from DC to UV becomes accessible.

Previously we have used frequency combs to determine the Cs D_1 and D_2 lines [26, 27]. However, these experiments were performed with the 3.39 μm, the 1.696 μm and the 848 nm lasers only (as shown in Fig. 4) so that the He–Ne standard was used as the reference. To bridge the 18.4 THz frequency gap between the fourth harmonic of the standard at 848 nm and the Cs D_1 line at 895 nm a mode–locked

laser without external spectral broadening in a fiber was used. The D_2 line at 852 nm was reached in the same way but with an optical frequency comb generator. With a similar set-up we also measured the $5s^2\ ^1S_0 - 5s5p\ ^3P_0$ transition of a single trapped $^{115}In^+$ ion [28].

7. OUTLOOK

In the future we hope to be able to realize a compact optical frequency counter based on a chain of optical frequency interval dividers and a mode–locked laser. Fig. 5 shows a possible realization. In principle this system can be made transportable since the second harmonic of 972 nm can be produced with sufficient efficiency to drive further non–linear processes [29]. In this case the dye laser at 486 nm is not needed and besides the femtosecond laser only diode lasers are employed. The non–linear processes are highly efficient in periodically poled lithium tantalate [30]. This new type of frequency chain is promises to become a precise and reliable optical clockwork still missing for the construction of a future optical clock.

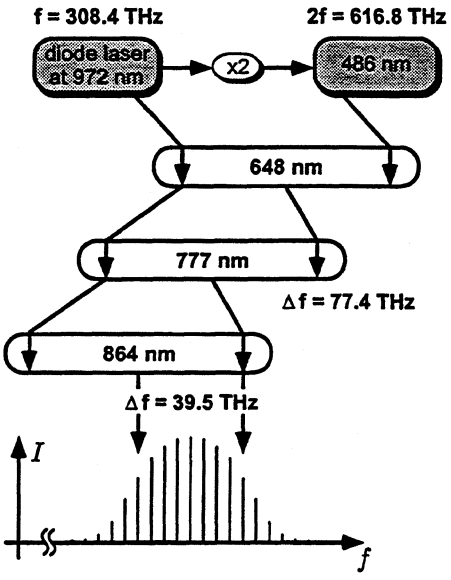


Figure 5: An example for a novel frequency chain, that allows the measurement of the hydrogen 1S–2S (243 nm), 2S–4S (972 nm) and the 2S–8S (777 nm) two–photon transition frequencies, and a quasi continuous range of frequencies between 777 nm and 864 nm, where the frequency comb operates. For this chain 4 diode lasers and one femtosecond laser is needed.

REFERENCES

[1] See for example: M. Roberts, P. Taylor, G. P. Barwood, P. Gill, H. A. Klein, and W. R. Rowley, "Observation of an Electric Octupole Transition in a single Ion", Phys. Rev. Lett., vol. 78, pp. 1876–1879, March 1997.

[2] Th. Udem, A. Huber, B. Gross, J. Reichert, M. Prevedelli, M. Weitz, and T. W. Hänsch, "Phase–coherent Measurement of the Hydrogen 1S–2S Transition Frequency with an optical Frequency Interval Divider Chain", Phys. Rev. Lett., vol. 79, pp. 2646–2649, October 1997.

[3] H. Schnatz, B. Lipphardt, J. Helmcke, F. Riehle, and G. Zinner, "First phase–coherent Frequency Measurement of visible Radiation", Phys. Rev. Lett., vol. 76, pp. 18–21, January 1996.

[4] B. De Beauvoir, F. Nez, L. Julien, B. Cagnac, F. Biraben, D. Touahri, L. Hilico, O. Acef, A. Clairon, and J. J. Zondy, "Absolute Frequency Measurement of the 2S–8S/D Transitions in Hydrogen and Deuterium: New Determination of the Rydberg Constant", Phys. Rev. Lett., vol. 78, pp. 440–443, January 1997.

[5] M. Kourogi, B. Widiyatomoko, Y. Takeuchi, and M. Ohtsu, "Limit of Optical–Frequency Comb Generation Due to Material Dispersion", IEEE J. Quant. Electron., vol. 31, pp. 2120–2126, December 1995.

[6] L. Brothers, D. Lee, and N. Wong, "Terahertz Optical Frequency Comb Generation and Phase–Locking of an Optical Parametric Oscillator at 665 GHz", Opt. Lett., vol. 19, pp. 245–247, February 1994.

[7] D. McIntyre, and T. W. Hänsch,"Novel Optical Frequency Divider and Synthesizer", presented at the Annual Meeting of the Optical Society of America, Washington D.C., USA 1988.

[8] H. R. Telle, D. Meschede, and T. W. Hänsch, "Realization of a new Concept for visible Frequency Division: Phase–locking of Harmonic and Sum Frequencies", Opt. Lett., vol. 15, pp. 532–534, May 1990.

[9] N. Nakagawa, M. Kourogi, and M. Ohtsu,"Proposal of a Frequency–Synthesis Chain between the Microwave and optical Frequencies of the Ca Intercombination Line at 657 nm using Diode Lasers", Appl. Phys. B, vol. 57, pp. 425–430, December 1993.

[10] M. Prevedelli, T. Freegarde, and T. W. Hänsch,"Phase locking of Grating–tuned Diode Lasers", Appl. Phys., vol. B 60, pp. S241–S248, December 1995.

[11] M. Kourogi, T. Enami, and M. Ohtsu, "A Monolithic Optical Frequency Comb Generator", IEEE Photonics Technol. Letters, vol. 6, pp 214–217, February 1994.

[12] R. W. P. Drever, J. L. Hall, F. V. Kowalski, J. Hough, G. M. Ford, A. J. Munley, and H. Ward, "Laser Phase and Frequency Stabilization using an Optical Resonator", Appl. Phys. B, vol. 31, pp 97–105, June 1983.

[13] K. Imai, M. Kourogi, and M. Ohtsu, "30–THz Span Optical Frequency Comb Generation by Self–Phase Modulation in an Optical Fiber", IEEE J. Quant. Electron., vol. 34, pp. 54–60, January 1998.

[14] K. Imai, M. Kourogi, B. Widiyatmoko, and M. Ohtsu, "Beat Signal Measurement of 12–THz separated Lasers by using Self–Phase–Modulation in Optical Fiber", presented at the Conference on Lasers and Electro-Optics, San Francisco, USA, May 3-8, 1998.

[15] J. N. Eckstein, A. I. Ferguson, and T. W. Hänsch, "High–Resolution Two–Photon Spectroscopy with Picosecond Light Pulses", Phys. Rev. Lett., vol. 40, pp. 847–850, March 1978.

[16] F. Krausz, M. E. Fermann, T. Brabec, P. F. Curley, M. Hofer, M. H. Ober, C. Spielmann, E. Wintner, A. J. Schmidt, "Femtosecond Solid–State Lasers", IEEE J. Quant. Electron., vol. 28, pp. 2097–2122, October 1992.

[17] Jung I. D., Kärtner F. X., Matuschek N., Sutter D. H., Morier–Genoud F.,Zhang G., Keller U., Scheuer V., Tilsch M., Tschudi T., "Self–Starting 6.5–fs Pulses from a Ti:Sapphire Laser", Opt. Lett., vol. 22, pp. 1009–1011, July 1997.

[18] L. Xu, Ch. Spielmann, A. Poppe, T. Brabec, F. Krausz, T. W. Hänsch, "Route to Phase Control of ultrashort light Pulses", Opt. Lett., vol. 21, pp. 2008–2010, December 1996.

[19] Th. Udem, J. Reichert, R. Holzwarth, T. W. Hänsch, "Accurate Measurement of Large Optical Frequency Differences with a Mode–Locked Laser", Opt. Lett., in press.

[20] S. N. Bagayev, A. K. Dmitriyev, and P. V. Pokasov, "Transportable He–Ne/CH_4 Frequency Standard for Precision Measurements", Laser Physics, vol. 7, pp. 989–992, July 1997.

[21] A. Huber, Th. Udem, B. Gross, J. Reichert, M. Kourogi, K. Pachucki, M. Weitz, and T. W. Hänsch, "Hydrogen–Deuterium 1S–2S Isotope Shift and the Structure of the Deuteron", Pys. Rev. Lett., vol. 80, pp. 468–471, January 1998.

[22] J. Reichert, M. Niering, R. Holzwarth, M. Weitz, Th. Udem, and T. W. Hänsch, to be published.

[23] Th. Udem, J. Reichert, T. W. Hänsch, and M. Kourogi, "Accuracy of Optical Frequency Comb Generators and Optical Frequency Interval Divider Chains", Opt. Lett., vol. 23, pp. 1387–1389, September 1998.

[24] J. Reichert, R. Holzwarth, Th. Udem, T. W. Hänsch, to be published (german patent pending).

[25] S. Ghezali, Ph. Laurent, S. N. Lea, and A. Clairon, "An Experimental Study of Spin–Exchange Frequency Shift in a Laser–Cooled Cesium Fountain Frequency Standard", Europhys. Lett., vol. 36, pp. 25–30, October 1996.

[26] Th. Udem, J. Reichert, R. Holzwarth, and T. Hänsch, "Absolute Optical Frequency Measurement of the Cesium D_1 Line with a Mode Locked Laser", Phys. Rev. Lett., in press.

[27] J. Reichert et al, to be published.

[28] J. von Zanthier, J. Abel, Th. Becker, M. Fries, E. Peik, H. Walter, R. Holzwarth, J. Reichert, Th. Udem, T. W. Hänsch, A. Yu. Nevsky, M. N. Skvortsov, and S. N. Bagaev, "Absolute Frequency Measurement of the ^{115}Indium$^+$ $5s^2\ {}^1S_0 - 5s5p\ {}^3P_0$ transition", submitted to Opt. Comm. See also the contribution of E. Peik in this volume.

[29] C. Zimmermann, V. Vuletić, A. Hemmerich, T. W. Hänsch, "All Solid State Laser Source for Tunable Blue and Ultraviolet Radiation", Appl. Phys. Lett., vol. 66, pp. 2318–2320, May 1995.

[30] J. P. Meyn, and M. M. Fejer, "Tunable Ultraviolet Radiation by Second–Harmonic Generation in Periodically Poled Lithium Tantalate", Opt. Lett., vol. 22, pp. 1214–1216, August 1997.

FREQUENCY STABILITY LIMITS OF OPTICAL FREQUENCY INTERVALS IN NEW GENERATION OPTICAL-TO-MICROWAVE FREQUENCY CHAINS

R P Kovacich, and A N Luiten
Physics Department, University of Western Australia (UWA), Nedlands 6907 WA, Australia

ABSTRACT

The limit of the frequency stability that can be transferred by new generation optical-to-microwave frequency chains is discussed. This limit is imposed by the quality of the optical frequency intervals which are formed by electronic phase locked extended cavity laser diodes. The major noise sources associated with this were identified and independently measured, in particular the residual phase noise, cycle slipping and path length fluctuations.

1. INTRODUCTION

High resolution optical frequency measurements, and optical frequency standards will require optical-to-microwave frequency chains in order to convert the optical signal to a microwave one that can be calibrated against known standards. To ensure the frequency chain is not the limit to this process its frequency stability must be superior to the optical frequency standard. New generation optical-to-microwave frequency chains use the Telle-Hänsch optical frequency interval bisection technique in combination with an optical frequency comb generator (OFCG)[1]. An example is shown in figure 1. It consists of a cascade of optical frequency interval divider stages (OFID) that from successively smaller optical frequency intervals, until it can be spanned by an OFCG. An OFID is made by phase locking a slave laser to the mean frequency of the two optical signals that form the preceding optical frequency interval. Therefore the slave laser bisects the optical frequency interval and provides the signal which is passed onto the next OFID. With a cascade of N OFIDs the initial optical frequency interval is divided by 2^N. In order to implement the chain conveniently and economically, extended cavity laser diodes (ECLD) are used which can be electronically phase locked. To provide the necessary optical signals for phase locking, highly efficient non-linear optical crystals i.e. optical mixers are used in each OFID.

The frequency stability that can be transferred by this chain design is limited by the frequency stability of the optical frequency intervals which is determined by the following noise sources:

1. Residual phase noise of the phase locked ECLDs.
2. Cycle slipping of the phase locking
3. Length fluctuations of optical signal paths.

We have independently measured these parameters, and have deduced that in order to achieve a stability of 10^{-15}, one must use optical frequency intervals greater than 10THz or filter the noise introduced by the chain.

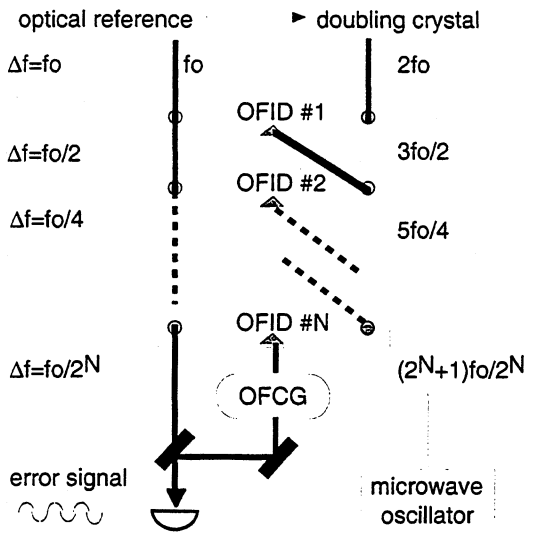


Figure 1. New generation optical-to-microwave frequency chain.

2.PHASE LOCKING OF EXTENDED CAVITY LASER DIODES

It has been demonstrated that an OFID can be simulated by phase locking two lasers through a non-linear optical crystal[2]. The optical beat signal is between the sum frequency and second harmonic of the lasers which is generated in the crystal. However the experiment is not a perfect model since the sum frequency is not independent of the slave laser as it is in an OFID, therefore this must be taken into account. Figure 2 shows such an experiment that we have used to measure the noise parameters, and hence determine the frequency stability limit imposed by the OFIDs. We used two commercial ECLDs at 860nm which had linewidths of 40kHz. One was electronically phase locked to the other which was free running. As the quality of the phase lock depends greatly on the signal-to-noise ratio of the optical beat, it is important to duplicate the rather weak non-linear signal that is expected in an OFID, typically a few 100nW. Using a potassium niobate crystal we could generate 1.5μW of blue light with 20mW of infra-red into it, but this was deliberately reduced by misaligning the polarization using a half-wave plate before the crystal.

The optical beat signal was detected by an avalanche photodiode, and this was followed by a second order heterodyne phase locked loop (PLL) circuit, consisting of

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

a phase detector followed by an active filter which adjusted the laser diode bias current. However to correct for the phase characteristic of the current modulation a lead compensator was included to improve the phase margin. Low frequency drift was corrected by a piezoelectric actuator inside the ECLD which had extra gain provided by a slow integrator.

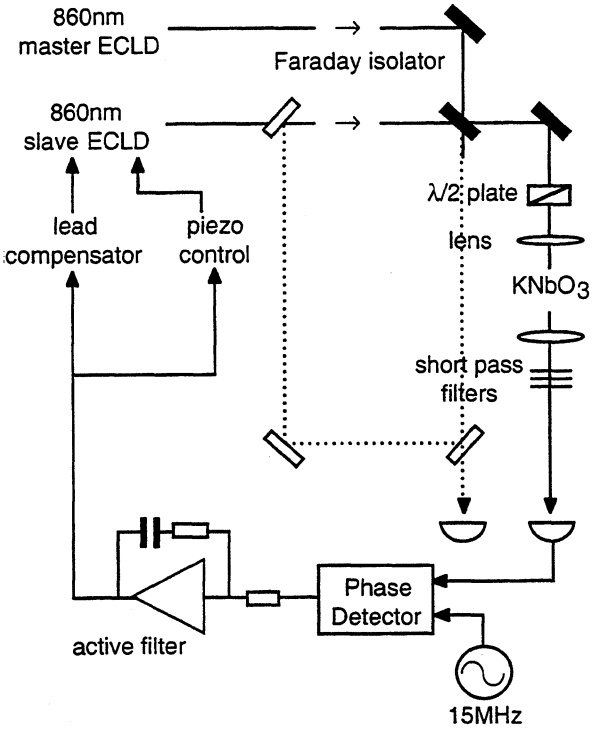


Figure 2. Electronic phase locking of extended cavity laser diodes. Dotted line shows interferometer used for measurement of optical path length fluctuation.

Two types of phase detector were tried, an analog mixer and a digital phase-frequency comparator (PFC) based on a MC12040 integrated circuit. The digital PFC made a more robust PLL due to superior relocking which was aided by its frequency discrimination ability. The analog mixer does not have this ability, and therefore it was unable to regain lock if the beat signal was outside the PLL bandwidth of 1MHz.

The residual phase noise was measured by demodulating the beat signal detected on a separate photodetector. Note that the infra-red beat signal between the two ECLDs is equal to the non-linear beat signal i.e. $f_1+f_2=2f_1-(f_1+f_2)$. The results shown in figure 3 are of the PLL using an analog mixer. Curve a is the residual phase noise which is above the shotnoise level (curve b), and is therefore limited by the PLL gain. The hump which occurs at approximately the noise bandwidth is due to a poor phase margin, and the bright line at 1.7MHz is an oscillation caused by the phase crossover. For large phase modulations, such as the mains signal at 50Hz, the residual phase noise is limited by the relative intensity noise (RIN - curve c). This arises since the bias current on the laser diode changes both the amplitude and phase, which cannot be distinguished on the photodetector.

The Allan variance of the frequency fluctuations can be deduced by transforming the phase noise[3], as shown by curve a (dotted) in figure 4. However an analytic expression which approximates this is more useful. If the residual phase noise is assumed to be white then the Allan variance is given by equation 1, where $S_{\phi ave}$ is the average residual phase noise level and B_N is the PLL noise bandwidth. The solid curve in figure 4 shows the approximation is valid, assuming $S_{\phi ave}=10^{-8}Rad^2/Hz$.

$$\sigma_\phi^2(\tau)=\frac{3B_N S_{\phi ave}}{(2\pi)^2\tau^2} \qquad (1)$$

This is intended to represent the output noise of an OFID, which is correct if $S_{\phi ave}$ is limited by the gain of the PLL. However, if $S_{\phi ave}$ is the shotnoise level then the action of the doubling crystal inside the OFID must be considered, in which case equation 1 must be divided by 4.

2.1 Optical Path Length Fluctuation

The error signal for each PLL in the frequency chain is generated by comparing two physically separate optical signals of the same frequency. Any fluctuations between the two optical path lengths will result in a phase error being transferred to the slave laser or microwave oscillator[1]. To estimate the noise that would be introduced by this effect, an interferometer (dotted line in

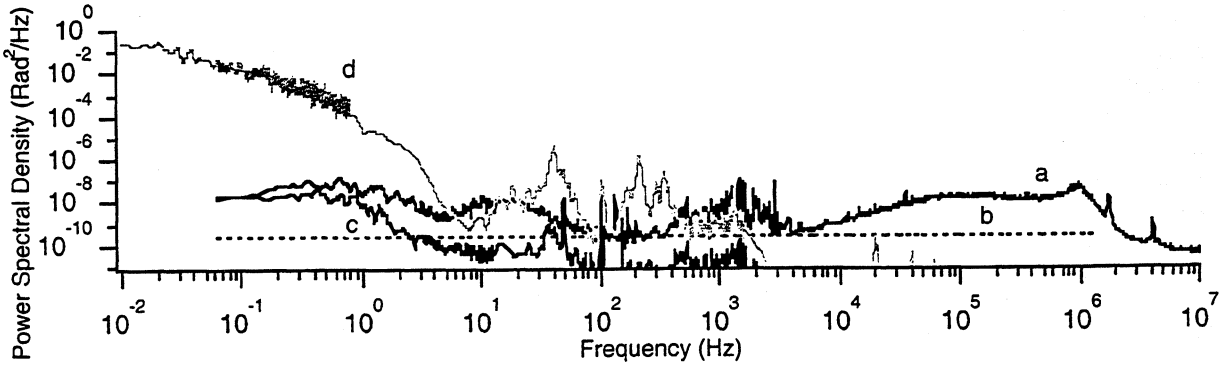


Figure 3. Phase noise measurements (a- residual, b- shotnoise, c- RIN, d- path length fluctuation).

figure 2) using only the master ECLD was used to measure the phase fluctuations introduced by a one meter open optical path.

Curve d in figure 2 shows the phase noise due to the optical path length fluctuation. Compared to the residual phase noise of the PLL, it dominates at low frequencies where it follows a $1/f^2$ frequency dependence. We have confirmed that this is mainly due to air currents. The Allan variance of the corresponding frequency fluctuation is shown in figure 4 (curve b), and this can be approximated by equation 2 where S_p is the phase noise at 1Hz.

$$\sigma_p^2(\tau) = \frac{S_p}{2\tau} \qquad (2)$$

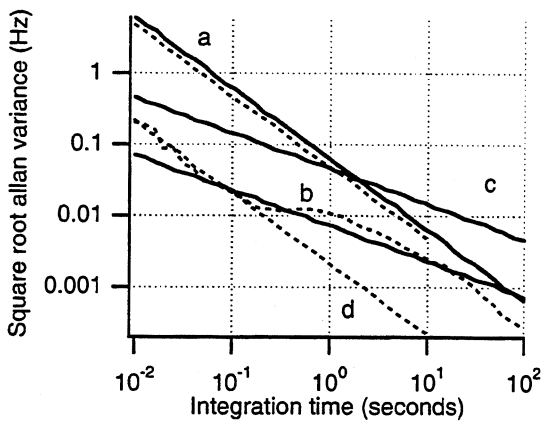


Figure 4. Square root allan variance of frequency fluctuations (a- residual phase noise, b- path length fluctuation, c- cycle slipping, d- filtered residual phase noise). Solid lines are approximations.

2.2 Cycle Slipping

The presence of noise in a PLL and a phase detector with finite range leads to the phenomenon of cycle slipping. For an instant, when the noise exceeds the range of the phase detector, there is a false error signal in which time the laser can accumulate an integer multiple of 2π radian phase error (cycle slips) before it regains lock. Cycle slips can be detected and measured, and the information can be used to correct the error. This is usually done in post-processing of data [1], but in applications where the frequency chain must provide a real-time signal it should still be possible to apply corrections in real-time. However the rate at which the real-time signal can be provided will be determined by the time needed for the system to provide a correction, which is ultimately limited by the time required for the PLL to recover from a cycle slip (roughly $10/B_N$ seconds). In this situation it is useful to know how often corrections are needed, if at all.

The mean time between cycle slipping for a second order PLL is described by equation 3 in a low phase noise approximation[4], where σ^2 is the total integrated phase noise and R is the range of the phase detector in π multiples. For an analog mixer R=1, and for the digital PFC R=2.

$$T_{av} \propto \frac{1}{B_N}\exp\left(\frac{0.6R}{\sigma^2}\right) \qquad (3)$$

Cycle slips were measured by comparing the phase of the beat signal with the heterodyne reference after frequency division of both by a digital counter. Figure 5 shows the cycle slipping rate for different values of phase noise, which was obtained by reducing the PLL gain (and hence B_N). This was done for both the analog mixer and digital PFC phase detectors. The relationship between the gradient of both lines-of-best-fit in figure 5, which should equal 0.6R, confirms equation 3.

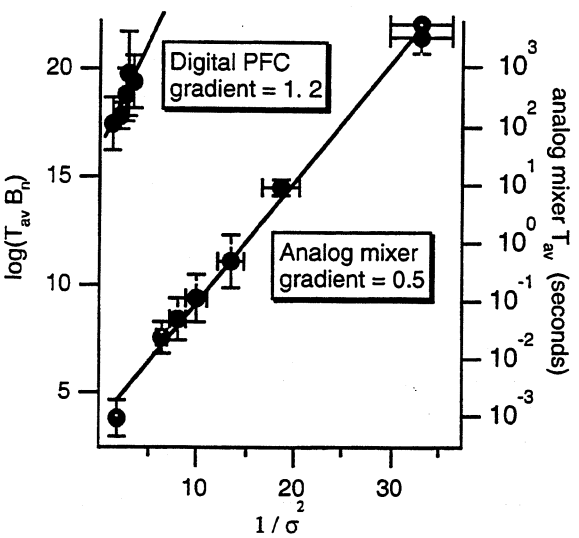


Figure 5. Cycle slipping dependence on phase noise.

The digital PFC is a noisier device than an analog mixer due to aliasing of noise outside the PLL bandwidth, such as radio interference and shotnoise. So although the digital PFC offers a lower cycle slipping rate and superior relocking than an analog mixer, it requires an optical beat signal with a larger signal-to-noise ratio. This is particularly serious if there is strong radio interference, which we have observed to increase the cycle slipping rate significantly. The analog mixer was unaffected in the same situation. A phase detector which combines the advantages of both the analog mixer and digital PFC would be ideal. This has been achieved by Prevedelli et al[7] using an analog mixer in parallel with a digital PFC to provide the high frequency and low frequency error signals respectively. We have demonstrated a similar system with success. This gives the advantage of superior relocking with low noise, but the cycle slipping follows that of the analog mixer.

The effect of cycle slipping, as would be seen by a frequency counter, is to produce a white frequency noise. If left uncorrected then the Allan variance of the frequency fluctuations caused by it is given by equation 4, where m is the average phase error (in 2π multiples) accumulated during the cycle slipping.

$$\sigma_c^2(\tau) = \frac{m^2}{4T_{av}\tau} \quad (4)$$

This is shown as curve c in figure 4, using T_{av}=3000 and m=5 which have been determined by our measurements. The factor of 4 takes into account the doubling crystal in the OFID. Compared to other noise sources, cycle slipping is a serious limit at long integration times.

3.CHAIN FREQUENCY STABILITY

The frequency stability that can be transferred by a new generation optical-to-microwave frequency chain is mainly determined by the stability of the smallest optical frequency interval created by the OFIDs. This interval will accumulate the frequency error of all the OFIDs in the frequency chain cascade. The output frequency of each OFID obeys equation 5, where df_n is the noise introduced by the OFID and f_{n-1},f_0 are the inputs to the OFID.

$$f_n = \frac{f_{n-1} + f_0}{2} + df_n \quad (5)$$

The noise introduced by the OFID is the sum of the frequency fluctuations that have been discussed, and this can be expressed by equation 6.

$$df_n = \sqrt{\sigma_\phi^2(\tau) + \sigma_p^2(\tau) + \sigma_c^2(\tau)} \quad (6)$$

By using equation 5 recursively the frequency fluctuation in the final optical frequency interval can be found. This is given by equation 7, where N is the number of OFIDs.

$$\frac{df}{f} = \frac{2^N}{f_0} \sum_{i=1}^{N} \frac{df_i}{2^{i-1}} \quad (7)$$

The noise of the last OFID is the most dominant term in the summation, since the noise of the preceding OFIDs are progressively reduced going up the chain. So noisier OFIDs can be tolerated at the top of the chain if necessary. If the noise in each OFID is nearly identical then a simpler expression can be formed:

$$\frac{df}{f} = \frac{(2^{N+1} - 2)df_n}{f_0} \quad (8)$$

Optical and microwave oscillators with excellent short term stability (high 10^{-15} at one second) have been demonstrated[6, 7], and these provide an excellent means of testing the precision of new-generation optical-to-microwave frequency chains. By substituting equation 6 into equation 8 and using the measured values, we note that in order to transfer a frequency stability at this level an optical frequency interval greater than 10THz is required due to the residual phase noise, and cycle slipping should be corrected. The microwave oscillator driving the OFCG is phase locked to this optical frequency interval. It provides the final output of the frequency chain. The PLL controlling the microwave oscillator acts to filter the residual phase noise of the optical frequency interval, and if the bandwidth of this PLL is much less than the laser PLL bandwidth then a smaller optical frequency interval can be tolerated. Curve d in figure 4 shows the Allan variance of the residual phase noise after a 10kHz PLL bandwidth. The limit imposed by path length fluctuations is more difficult to overcome. Although this can be reduced by enclosing the frequency chain to restrict air currents.

From this investigation it appears that larger optical frequency intervals are advantageous since the OFID noise remains fixed. This can be realized using mode-locked laser OFCGs which have demonstrated spans of up to 20THz[8]. However the microwave oscillator then requires extremely low phase noise, which can be met by state-of-the-art sapphire dielectric microwave oscillators[7].

4.REFERENCES

[1] T. Udem, J. Reichert, T. W. Hänsch and M. Kourogi, "Accuracy of optical frequency comb generators and optical frequency interval divider chains", Optics Letters, vol. 23, pp. 1387-9, September 1998.

[2] H. R. Telle, D. Meschede and T. W. Hänsch, "Realization of a new concept for visible frequency division: phase locking of harmonic and sum frequencies", Optics Letters, vol. 15, pp. 532-4, May 1990.

[3] M. Ohtsu, H. Fukada, T. Tako and H. Tsuchida, "Estimation of the Ultimate Frequency Stability of Semiconductor Lasers", Japanese Journal of Applied Physics, vol. 22, pp. 1157-66, July 1983.

[4] G. Ascheid, H. Meyr, "Cycle Slips in Phase-Locked Loops: A Tutorial Survey", IEEE transactions on communications, vol. 30, pp. 2228-41, October 1982.

[5] M. Prevedelli, P. Cancio, G. Giusfredi, F. S. Pavone and M. Inguscio, "Frequency control of DBR diode lasers at 1.08 micrometer and precision spectroscopy of helium", Optics Communications, vol. 125, pp. 231-6, April 1996.

[6] S. Seel, R. Storz, G. Ruoso, J. Mlynek and S. Schiller, "Cryogenic Optical Resonators: A New Tool for Laser Frequency Stabilization at the 1Hz level.", Physical Review Letters, vol. 78, pp. 4741-4, June 1997.

[7] A. N. Luiten, A. G. Mann, N. J. McDonald and D. G. Blair, "Latest Results of the UWA Cryogenic Sapphire Oscillator", in the proceedings of the 49th Annual Frequency Control Symposium, 1995, pp. 433-7.

[8] T. Udem, J. Reichert, R. Holzwarth and T. W. Hänsch, "Accurate Measurement of Large Optical Frequency Differences with a Mode Locked Laser", submitted to Physical Review Letters, 1999.

SUB-SYSTEMS FOR OPTICAL FREQUENCY MEASUREMENTS: APPLICATION TO THE 282 NM $^{199}HG^{+}$ TRANSITION AND THE 657 NM CA LINE.

B. Frech, J. S. Wells, C. W. Oates, J. Mitchell, Yu-Ping Lan[a], T. Kurosu[b], L. Zink, L. Hollberg, T. Zibrova[c], B. C. Young, and J.C. Bergquist

Time and Frequency Division 847
National Institute of Standards and Technology
325 Broadway, Boulder, CO 80303

ABSTRACT

We are developing laser frequency measurement technologies that should allow us to construct an optical frequency synthesis system capable of measuring optical frequencies with a precision limited by the atomic frequency standards. The system will be used to interconnect and compare new advanced optical-frequency references (such as Ca, Hg^{+} and others) and eventually to connect these references to the Cs primary frequency standard. The approach we are taking is to subdivide optical frequency intervals into smaller and smaller pieces until we are able to use standard electronic-frequency-measurement technology to measure the smallest interval.

1. DOMAIN-ENGINEERED LITHIUM-NIOBATE FOR OPTICAL SYNTHESIS

Ideally the optical-frequency-measurement system would be constructed from self-similar frequency dividing units that use compact lasers, simple nonlinear optical-mixing elements that are easy to assemble, and then run reliably. With this goal in mind we have constructed and evaluated the performance of a number of diode-laser systems and optical-mixing stages. We have made progress in fabricating and using periodically poled lithium-niobate (PPLN) nonlinear crystals for most of the mixing stages in our optical synthesis system. With extended-cavity diode lasers (ECDLs) and a CO $\Delta v=2$ laser we have tested these crystals for practicality and efficiency in second-harmonic generation (SHG), sum-frequency generation (SFG), difference-frequency generation (DFG), and even third-harmonic generation (THG) [1]. By designing and fabricating the crystals for specific interactions we can achieve efficient nonlinear mixing (see Fig. 1).

Optical-grade lithium niobate (transparency window from 330-5500 nm) is readily available, and periodically poled crystals are now being produced commercially with various poling periods. Since not all of the poling periods that we needed were available, we developed our own poling capability at NIST. The poling is done by the electric- field method, and we have now poled a number of samples with poling-periods ranging from 8 to 36 μm.

[a,b,c] see References

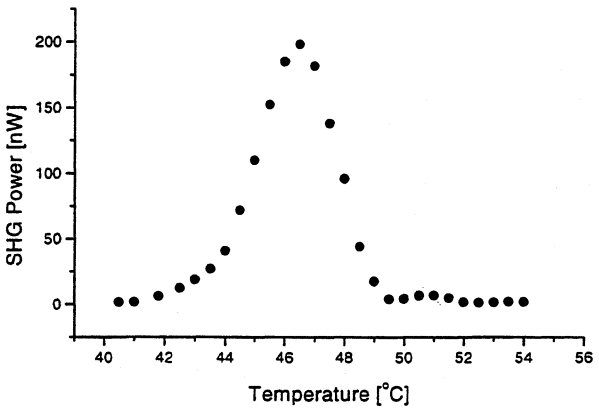


Figure 1. Second harmonic power verses crystal temperature using 5mW of 1315 nm diode laser radiation to produce 200 nW of 657 nm light in a single pass through a 20 mm long PPLN crystal. This is adequate power to phase-lock the 1315 nm laser to the 657 nm reference laser.

Domain-engineered materials (of which PPLN is by far the most common) are a powerful and practical tool for the general task of optical-frequency synthesis [2]. We have used PPLN for a wide range of nonlinear mixing applications, some of which are summarized in Table 1. For example we have explored all possible combinations of mixing between three important wavelengths (563 nm, 657 nm and 3925 nm) used in our present synthesis scheme. By pure luck, these mixings require the same poling period (12.6 μm) that we need for SHG of 1314 nm to reach 657 nm. Further evidence for the usefulness of this material as an optical mixer were weak blue and green beams that we observed coming out of a crystal designed for first order DFG using 800 nm minus 1064 nm to produce 3.4 μm. These short-wavelength beams were the two SHG signals and the SFG signal that were phase-matched at slightly different angles and high orders of the poling period [1]. In a special case, it was even possible to achieve third harmonic generation of a cw laser at 3561 nm as a result of cascaded second order nonlinearities, that simultaneously phase-matched second-harmonic generation and sum frequency mixing.

Table 1. Characteristics of the different PPLN stages involved in the studies

PPLN stage	SHG#1	SHG#2	DFG#1	DFG#2	SHG Test	DFG Test	DFG Test	DFG Test	DFG Test
Period (μm)	29.9	12.7	12.6	30.9	32.7	12.6	22.2	30.75	31.0
Temp (°C)	87	24	31	34	87	31	61	83	42
λ in (nm)	3925	1308	563 3925	1308 3925	3378	657 563	810 1064	1064 2902	1064 2860
Power in (mW)	100	5	15 50	4 80	200	4 20	50 250	300 65	300 48
λ out (μm)	1962	654	657	1962	1689	3925	3393	1680	1694
Power out (nW)	2500	200	4000	150	10000	10	4000	2500	2500

Domain-engineered materials such as PPLN might begin to play the role of a universal optical mixer. Unfortunately, optical nonlinearities are tiny in comparison to what can be achieved at lower frequencies, with microwave harmonic-mixer diodes, for example. Using low-power cw lasers we are basically limited to second order optical nonlinearities which do not create high-order mixing products efficiently. In addition, some technical challenges still limit the usefulness of PPLN in the short-wavelength regions, for example SHG to the UV. Some of these cases require poling periods that are so short that they are difficult to fabricate with the present state of poling technology. As mentioned above, going to higher order in the poling period (3,5,7...) is an alternative in some cases, but the efficiency drops rapidly, as $(1/\text{order})^2$. Also, the optical power levels that can be generated in the visible/UV regions can be limited by photo-refractive effects. This, however, is not a serious problem for use in optical synthesis since we don't require high powers. Future use of waveguide mixing devices could be attractive, although waveguide technology is more challenging when the wavelengths are very different.

2. CONNECTION TO CO_2 LASERS

With a dependable connection between visible wavelengths using the CO $\Delta v=2$ laser (~90 THz) and PPLN mixing crystals, we wanted to explore the possibility of stepping between diode laser wavelengths separated by a CO_2 laser frequency (~30 THz). Three CO_2 laser steps would then be able to span the CO $\Delta v=2$ laser frequency interval. Thus, we set up an experiment to sum an ECDL at 844 nm wavelength with a CO_2 laser running on the $R_I(38)$ transition in the 10 μm band of the normal CO_2 isotopomer at a 10.134 μm wavelength.

The nonlinear crystal used for this experiment was Ag_3AsS_3 with a transparency window between 600 nm and 13 μm. Phase matching occurred at an angle of 22.6° with both input beams in ordinary polarization. Input powers of 250 mW at 10 μm and 1 mW at 844 nm led to an output power of 25 nW (efficiency: $\eta = 1\times 10^{-4}\ W^{-1}$) at 780 nm at extraordinary polarization. A beatnote with 25 dB signal-to-noise ratio (SNR) in a 10 kHz resolution bandwidth was sufficient to phase lock a second ECDL at 780 nm to the sum signal.

This experiment demonstrates the possibility to divide the coarser steps of the CO $\Delta v=2$ laser mixing stages phase coherently with a CO_2 laser. The CO_2 laser frequency can then be measured against Cs with a frequency chain [3,4].

3. PRELIMINARY MEASUREMENT OF THE HG^+ TRANSITION

Mixing of optical frequencies in PPLN opens the possibility for an accurate measurement of the Hg^+ clock transiton at 282 nm wavelength [5] using the Ca standard [6] as a frequency reference. The Ca frequency of the $^1S_0(m=0)$-$^3P_1(m=0)$ transition at $\nu(Ca)$=455 986 240.494 MHz has been measured by PTB against Cs with an uncertainty of ±135 Hz [7].

Even though the Hg^+ standard holds great promise for a future optical-frequency standard due to the very narrow optical clock transition $^2S_{1/2}$-$^2D_{5/2}$ (F=0-F=2, $\Delta m_f=0$) with a natural linewidth of $\Delta\nu\approx 1.7$ Hz, its absolute frequency is realtively poorly known.

Our scheme to measure the Hg^+ transition is based on two fortuitous 'coincidences', namely half the Hg^+ frequency $\nu(½Hg^+)\approx\nu(Ca)+\nu(Ca)/6$, and $\nu(Ca)/6\approx\nu(CO)$, where $\nu(CO)$ is the frequency of a CO $\Delta v=2$ laser operated on the $P_{33}(15)$ transition. Hence, we can use the CO laser as a transfer oscillator between the two optical standards. Since the 282 nm transition in Hg^+ can be interrogated with a frequency-doubled dye laser oscillating at 563 nm wavelength, the CO $\Delta v=2$ laser can connect between the Ca reference and the dye laser frequency.

Using a PPLN crystal with a 12.6 μm poling period we generated a difference signal between the dye laser frequency ($\nu(½Hg^+)$) and the CO $\Delta v=2$ laser frequency

ν(CO). With input powers of 15 mW and 70 mW, respectively, we generate 5μW in the difference signal. A beatnote between this signal and the ECDL from the Ca reference was achieved with a 3.6 GHz offset and a 20 dB SNR in a 100 kHz bandwidth. Careful alignment of the CO laser resonator and centering the CO Δν=2 laser on its gain curve lead to a preliminary measurement of the Hg^+ transition frequency, ν(Hg^+)= 1 064 721 600.4 ± 20 MHz. Knowledge of the CO Δν=2 laser frequency was obtained by generating another beat signal with a CO Δν=1 laser running on the $P_{33}(12)$ transition and a 20 GHz microwave source in seventh mixing order on a MIM diode. By centering both lasers carefully on their gain curves, the CO Δν=2 laser frequency could be estimated from the beat note frequency and the previously measured value of the CO Δν=1 laser frequency [8] with an uncertainty of 10 MHz.

This measurement of the Hg^+ optical clock transition frequency relative to Ca is a proof-of-principle demonstration of our measurement technique along our path toward a more precise measurement of the Hg^+ standard. In order to achieve an uncertainty of less than 1 kHz, we will also measure the CO Δν=2 laser frequency relative to the Ca frequency (Fig. 2).

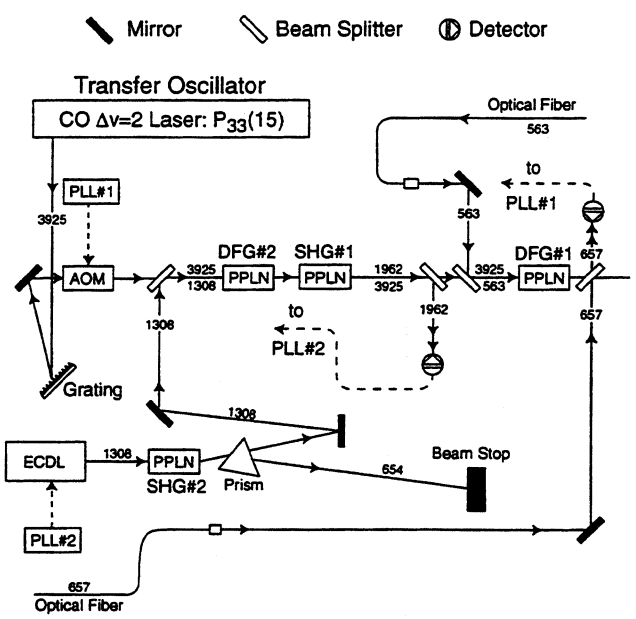


Figure 2. Block diagram of system for frequency measurement of Hg^+ with respect to Ca (AOM: Acousto-optic modulator; PLL: phase locked loop)

For this purpose the CO laser is stabilized by phase locking it (with a frequency offset) to the frequency difference, ν(½Hg^+)-ν(Ca). The use of different nonlinear mixing stages in PPLN permits an effective multiplication of 6 times the CO laser frequency for comparision with ν(Ca). An ECDL with a frequency ν(DL) at 1.3 μm wavelength is phase locked to 3×ν(CO) and at the same time frequency-doubled to 2×ν(DL), which is near ν(Ca). A frequency comb generator will provide the necessary connection between ν(Ca) and 6×ν(CO). A value for the CO laser transition frequency will then give a new accurate value for the frequency of the Hg^+ ion clock transition.

A difference signal (DFG#1) between ν(CO) and ν-(½Hg^+) with a power of 3.5 μW is generated in PPLN and a beat signal with a 3.6 GHz offset to ν(Ca) can be used for phase locking the CO laser. We have demonstrated phase locking of the CO laser by an acousto-optic modulator with a SNR of 40 dB in a 300 Hz bandwidth.

Another PPLN crystal allows doubling of the CO laser frequency ν(CO), delivering 2.5 μW of power in the second harmonic (SHG#1) at 1962 nm wavelength. The SHG#1 radiation can be used to generate a beat signal with the difference frequency (DFG#2) between the ECDL and the CO laser at 1962 nm wavelength that is generated in a third PPLN for a divide-by-three scheme of the ECDL. We achieve a 150 nW DFG#2 signal with input powers of 4 mW and 80 mW for the ECDL and the CO laser, respectively, leading to a beat note of 55 dB SNR in a 30 kHz bandwidth between the SHG#1 and DFG#2. We have shown that this beat signal can be used for phase locking the ECDL to the third-harmonic of the CO laser (Fig. 3).

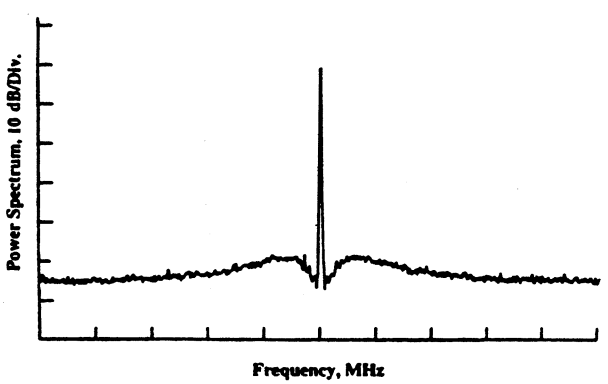


Figure 3: Beat signal between the doubled CO laser (SHG#1) and the difference signal between the ECDL and the CO laser (DFG#2) used for phaselocking the ECDL. (Center frequency: 21.7 MHz; Span: 5 MHz/div; Power: 10 dB/div; Resolution Bandwidth: 30 kHz)

A last PPLN crystal is used for doubling the ECDL (SHG#2) to 2×ν(DL), close to the Ca reference frequency ν(Ca). A power of 200 nW in the SHG#2 was achieved using 5 mW input power of the ECDL.

All of the required nonlinear mixings, beat notes, and phase locks have now been demonstrated. The last remaining step is to operate the whole system simultan-

eously and to measure the 2.24 THz frequency gap between the SHG#2 at 654 nm and the Ca reference at 657 nm. This interval will be bridged by an optical frequency comb generator [9-10] that is currently being built in our group.

4. ROUTE FROM THE VISIBLE TO THE CESIUM CLOCK

Our next major step is to measure the frequencies of Hg^+, Ca, and other important optical references relative to the primary Cs atomic frequency standard. The scheme we are developing for this purpose is diagramed below.

Three separate, but interconnected, frequency bisections are indicated in Fig. 4, and these require making the difference-frequency between each end point and the mid point equal. Alternatively, these could be implemented by the SHG plus SFG bisection method described by Telle et al. [11]. Whichever approach is most appropriate can be chosen for the particular wavelengths involved. Until recently, we had planned to continue the frequency chain shown in Fig. 4 by using CO_2 lasers as described in section 2. However, our present plan is to take advantage of the revolutionary new results from Udem and Hänsch et al. who have shown that it is possible to measure frequency intervals as large as 20 THz using femtosecond mode-locked lasers [12]. Using their method we plan to measure the interval between the 788 and 750 nm (~19 THz) lines shown in Fig. 4 with a pulsed Ti:sapphire laser. Much work remains to be done, however, with these new improved tools, such as diode lasers, PPLN and mode-locked lasers, and with some of the suggested schemes discussed, we believe that the path to precise optical frequency measurements is clear. We also feel that these new measurement systems are coming together just as the optical references are also reaching interesting performance levels. The hope is these optical standards will be able to compete favorably with microwave standards in terms of stability and accuracy.

ACKNOWLEDGEMENTS

We thank S. Waltman, O. Pfister, A. Aust, M. Mürtz, N. Mackie, and F. Walls for significant contributions to our optical synthesis effort. We especially thank Norman Sanford and his collegues for assistance in fabricating the PPLN mixers. This work has been supported in part by Air Force Office of Scientific Research and the German Academic Exchange Office (DAAD).

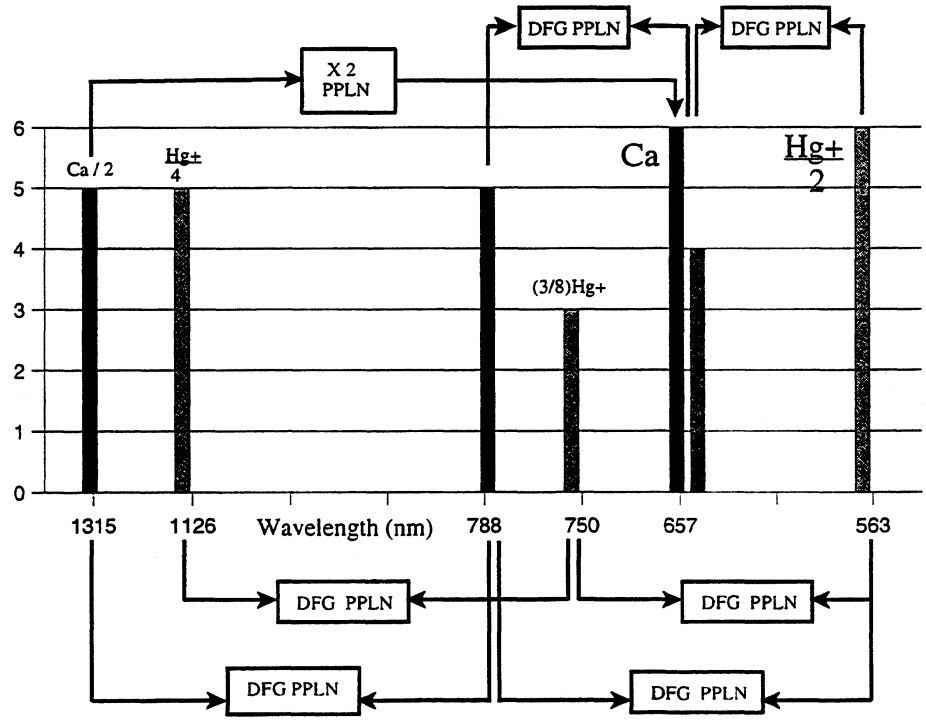


Figure 4. Diagram of proposed optical frequency measurement system to connect the Hg+ and Ca references to the Cs clock. This relies on the sub-units described above, and we have added two new lasers: one at 788 nm that bisects the frequency interval between 563 nm and 1315 nm, and one at 750 nm that bisects the frequency interval between 563 nm and 1126 nm (½ the 563 nm laser frequency). The boxes labeled "DFG PPLN" imply difference-frequency generation in periodically-poled lithium niobate etc.

REFERENCES

permanent addresses:

[a] Center for Measurement Standards, 321 Kuang Fu Road, Hsinchu, Taiwan 30042

[b] National Research Laboratory of Metrology, 1-1-4 Umezono, Tsukuba, Ibaraki 305, Japan

[c] P.N. Lebedev Institute, Russian Academy of Science, 117924 Moscow Leminsky pr. 53, Russia

[1] O. Pfister, J.S. Wells, L. Hollberg, L. Zink, D.A. Van Baak, M.D. Levenson, and W.R. Bosenberg, " Continous-wave frequency tripling and quadrupling by simultaneous three-wave mixings in periodically poled crystals: application to a two-step 1.19-10.71 μm frequency bridge," Opt. Lett. Vol. 22, No. 16, pp. 1211, 1997

[2] L.E. Myers, R.C. Eckhardt, M.M. Fejer, R.L. Byer, W.R. Bosenberg, J. Pierce, "Quasi-phase-matched optical parametric oscillator in bulk periodically poled $LiNbO_3$," J. Opt. Soc. Am.. B Vol. 12, pp. 2102, 1995

[3] D. Touahri, O. Acef, and J.J. Zondy, "30-THz upconversion of an AlGaAs diode laser with $AgGaS_2$: bridging the several-terahertz frequency gap in the near infrared," Opt. Lett. Vol. 21, No. 3, pp.213, 1996

[4] A.A. Madej and K.J. Siemsen, "Absolute heterodyne frequency measurement of the $^{88}Sr^+$ 455-THz S-D single ion transition," Opt. Lett. Vol. 21, No. 11, pp. 824, 1996

[5] B.C. Young, R.J. Rafac, F.C. Cruz, J.C. Bergquist, W.M. Itano, and D.J. Wineland, "$^{199}HG^+$ Optical Frequency Standard," presented in these proceedings

[6] C.W. Oates, F. Bondu, and L. Hollberg, "An All-Diode-Laser Optical Frequency Standard Based on Laser-Trapped Ca Atoms," presented in these proceedings

[7] H. Schnatz, B. Lipphardt, J. Helmcke, F. Riehle, G. Zinner, "First phase-coherent frequency measurement of visible radiation," Phys. Rev. Lett. Vol. 76, pp. 18, 1996, (according to F. Riehle the accuracy of the Ca measurement now reached 135 Hz at PTB)

[8] A. Hinz, J.S. Wells, A.G. Maki, "Heterodyne Measurements of Hot Bands and Isotopic Transitions of N2O Near 7.8 μm," Z. Phys. D-Atoms, Molecules and Clusters Vol. 5, pp. 351, 1987

[9] M. Kourogi, T. Enami and M. Ohtsu, "A Monolitic Optical Frequency Comb Generator," IEEE Phot. Technol. Lett. Vol. 6, No. 2, pp.214, 1994

[10] J. Ye, L-S. Ma, T. Day, and J.L. Hall, "Highly selective terahertz optical frequency comb generator," Opt. Lett. Vol. 22, No. 5, pp. 301, 1997

[11] H.R. Telle, D. Meschede, T.W. Hänsch, "Realization of a new concept for visible frequency division: phase locking of harmonic and sum frequencies," Opt. Lett. Vol. 15, pp. 532, 1990

[12] Th. Udem, J. Reichert, R. Holzwarth, B. Gross, T.W. Hänsch, "The Measurement of Large Optical Frequency Differences and the Design of a New Type of Frequency Chain," presented in these proceedings

Author to whom should be correspondence sent:

Leo Hollberg
Time and Frequency Division 847
National Institute of Standards and Technology
325 Broadway, Boulder, CO, 80303
Tel.: (303) 497-5770
Fax: (303) 497-7845
e-mail: hollberg@boulder.nist.gov

1999 Joint Meeting EFTF - IEEE IFCS

OPTICAL FREQUENCY SYNTHESIS

A N Luiten, J J McFerran and R P Kovacich,
University of Western Australia (UWA), Nedlands, Australia
Nedlands 6907 WA, Australia

ABSTRACT

This paper reports on progress towards the creation of an optical synthesizer. We consider the noise contributions of the two key elements of the new generation of frequency synthesisers: frequency interval dividers and frequency comb generators. We demonstrate that a synthesis resolution of 1 part in 10^{15} over a few seconds of integration time is feasible using these elements.

1. Introduction

We are constructing an optical synthesizer based upon two recent developments: optical frequency-interval dividers (OFID)[1], and optical frequency comb generators (OFCG)[2]. An initial optical frequency interval (Δf), set by the frequency difference between a signal (f_0) and its second harmonic ($2f_0$), is bisected seven times by a cascade of seven optical frequency-interval dividers. The output frequency interval ($\Delta f/2^7 \equiv f_0/2^7$), can then be measured in terms of a known microwave or radio-frequency signal using an optical frequency comb. This project is a collaborative effort between the Physikalisch-Technische Bundesanstalt (OFCG research) and the UWA (OFID research).

Our goal is the synthesis of an optical signal with 1 Hz ($\Delta f/f \sim 3 \times 10^{-15}$) resolution over integration times of a few seconds. The secondary characteristics that we wish to impose on our solution is that the chain be reliable, compact and relatively cheap to build.

2. Frequency Chain Design

The UWA-PTB frequency chain is schematically shown in figure 1. The initial frequency interval for the chain (282 THz) is formed between the fundamental and second harmonic outputs of an intra-cavity doubled Nd:YVO_4 laser. The laser simultaneously produces 30mW of 1064 nm and 20mW of 532 nm light. The particular sequence of OFID output wavelengths is determined by two factors: (i) a desire to commence at 1064 nm as this coincides with the output of a UWA optical frequency standard that has demonstrated a fractional frequency stability at the 10^{-15} level, (ii) a desire to converge on the 860 nm wavelength range where one has available high quality diode lasers, and phase matching capability in the highly non-linear material, $KNbO_3$.

The interval seen between the output of the seventh OFID, and either of its inputs, is of the order of 2 THz. This can be coherently spanned using an optical frequency comb generator (OFCG)[2]. An OFCG produces a broad comb of optical signals, each nominally spaced apart by a frequency equal to the OFCG driving oscillator (typically ~10 GHz).

The 851 nm output from OFID S2, and the output of the OFCG interfere on a photodiode, giving rise to a difference frequency signal between the nearest mode of the comb and the 851 nm signal. We can hold this difference frequency to be constant by adjustment of either the output frequency of the Nd:YVO_4 laser, or, of the microwave oscillator frequency. In the first case we have transferred the frequency stability of the microwave source into the optical domain, while in the second case we have transferred the stability of the optical source into the electronically-useful microwave domain. For simplicity, in the rest of this paper, we will assume that we control the optical standard, thus achieving our goal of synthesizing an optical signal from the microwave signal.

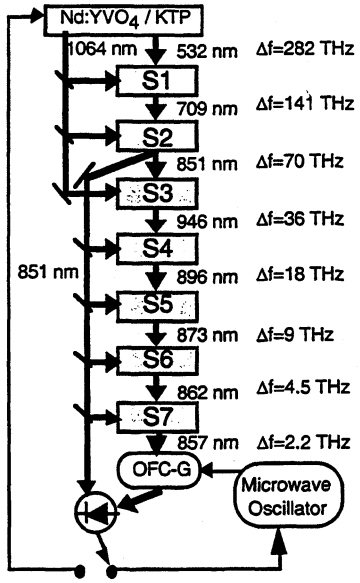


Figure 1. UWA-PTB frequency chain schematic. Each shaded box is an OFID.

3. Optical Frequency-Interval Divider

The basic building block of the UWA-PTB frequency chain is an OFID[1]. An OFID creates an output signal at the mean frequency of its two input signals (thus bisecting the frequency interval defined by the two inputs), by electronically phase-locking the second harmonic of a slave laser, to the sum frequency of its two input signals. Both the second harmonic and sum frequency signals are generated in appropriately selected non-linear crystals. For a successful implementation, an OFID requires the following features:

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

(i) a method for generating single-mode radiation at the mid-point of the input interval of the stage

(ii) generation of sufficient power in the non-linear processes, so that the phase difference between the sum, and second harmonic signals can be determined,

(iii) that the phase-locking circuit has sufficient gain to suppress the intrinsic fluctuations of the slave laser.

(i) Slave-Laser Radiation Generation

We plan to use Litmann-Metcalf (L-M) extended cavity DLs (ECDL) for all radiation sources in the frequency chain because of their compactness, narrow linewidth and tunability. The ECDLs can have their output frequency rapidly tuned by adjusting the diode junction current, and more slowly tuned, but over a larger frequency range, by piezoelectric adjustment of the length of the extended cavity. Since the free running linewidth of the extended cavity laser diode is around 100kHz, a phase-locked loop (PLL) bandwidth of around 1MHz is needed to reduce the slave DL phase fluctuations to below 1 rad. Unfortunately, commercial lasers are not available close to all the wavelengths called for in the frequency chain design. Table 1 presents our progress in producing single-mode radiation at all wavelengths in the chain to date.

λ (nm)	Nominal	Achieved	Comments
709	690	No	703 nm at 40°C
709	758	No	723 nm at 100K
709	1418	No	Resonant SHG of 200mW Multimode 1440nm DL
851	850	Yes	
946	980	No	Need -35°, 953 nm now
896	905	No	Not single-mode
873	870	Yes	
862	860	Yes	
857	860	Yes	

Table 1: Radiation generation at the wavelengths required in the synthesizer design

The output of the first OFID falls at 709nm. Unfortunately this wavelength lies in the gap between the ranges of two available semiconductor materials: InGaAlP (635-690 nm) and AlGaAs (750-850 nm). By cooling, and optically tuning a nominal 750 nm DL (actually 758 nm at room temperature), we have generated radiation down to 723 nm. By heating, and optically tuning, a 690 nm DL we have generated 703 nm radiation. We expect that we could tune these DLs further if we were to place an anti-reflection layer upon their output facet. We have recently obtained a vacuum system that can do this, and will undertake these measurements in the near future.

We are also researching an alternative approach for generating 709nm that makes use of resonant second harmonic conversion of a 250mW multi-mode DL in a LiB_3O_5 crystal. A 1440 nm DL is placed in an L-M extended cavity for mode control, and for tuning of its output to 1418 nm. We calculate that we can produce 5-10mW of single-mode 709nm radiation with this technique.

To produce 946nm radiation we have placed a 50mW 980nm DL in an EC. This has successfully produced a single-mode output, although our present thermoelectric cooler can only reduce the diode's temperature to –18°C. We calculate that cooling to –35°C will produce 946nm radiation. Finally, we have been able to produce 896 nm radiation by cooling a 905nm DL, although the 896 nm output radiation exhibited very poor coherence. Commercial single-mode 905 nm DLs have since become available, however, so we will replace the original DL with one of these models.

Although ECDLs exhibit many desirable characteristics for synthesis applications, they have several serious disadvantages: a high sensitivity to mechanical excitation, large drifts in output frequency (~200 MHz/day) and a tendency to hop between several adjacent longitudinal modes because the net round-trip gain for nearby laser modes is almost identical. These aspects do not bode well for a reliable frequency synthesis chain. Together with this problem, analog PLL circuits exhibit a very poor ability to phase-lock two signals that initially differ in frequency by much more than the bandwidth of the circuit. Thus, we find that manual intervention is generally required to initiate a phase-locked state, and to regain it, if it is lost through a large mechanical excitation. If one of the two phase-locked ECDLs experiences a mode-hop, induced by either mechanical or thermal changes, a phase-locked state can only be regained by manual intervention. We observe a mode-hop on free-running ECDLs, on average, once a day. To overcome these disadvantages, and to provide the system with some degree of automatism, we have implemented an additional frequency stabilisation system.

Initially, we have only locked the master ECDL to a mode of an Invar Fabry-Pérot (FP) cavity (Finesse ~ 200, FSR ~ 2 GHz). A small fraction of the output of the laser (~10%) is directed towards the cavity. We stabilize the laser using the standard Pound-Drever-Hall technique, and with a low bandwidth loop, have stabilised its frequency to better than 3 kHz, with respect to the mode, for integration times longer than 1s. This stabilisation of the master ECDL prevents large common-mode drifts of all of the phase-locked ECDLs, which leads to mode-hops. The locking system is very reliable and will automatically lock the ECDL to the same FP mode, even after several days. Eventually we plan to lock all of the ECDLs to the same Invar cavity, which has itself been stabilised to the optical frequency standard at the top of the synthesis frequency chain. Each ECDL will be locked at a tunable offset from the modes of the FP cavity by a double modulation

technique. By phase-modulating the laser diode output, with a signal that is itself phase-modulated, we create two pairs of sidebands: $f_c+f_{m1}\pm f_{m2}$, $f_c-f_{m1}\pm f_{m2}$ where $f_{m1}>>f_{m2}$, f_c is the carrier frequency and $f_{m1,2}$ are the modulating signals. By detecting the reflected cavity signal synchronously with f_{m2}, we can derive a dispersive error signal as the $\pm f_{m1}$ sidebands cross the FP cavity modes. Thus we are able to lock the ECDL to a point offset from the FP modes by f_{m1}. To phase-lock two ECDLs that have been pre-stabilised to the cavity, we feedback the phase-error signal to two points: for fluctuations at Fourier frequencies lower than the bandwidth of the pre-stabilisation loop, we adjust f_{m1}. For higher frequency components we feedback directly to the diode junction current.

(ii) Non-linear Optical Processes

Although cw laser diodes offer reliability and efficiency, their low output power places a great burden on the required efficiency of the non-linear optical crystals. Non-critically phase-matched (NCPM) potassium niobate ($KNbO_3$) is the crystal of choice in this application because of its unusually large non-linear coefficient, large angular acceptance and negligible walk-off. However NCPM (ϕ=0°) can only be achieved for wavelengths between 840nm and 960nm for this crystal, and so it is not possible to use $KNbO_3$ for all non-linear processes in the frequency chain. In particular, the two non-linear process in the first OFID: second harmonic generation (SHG) at 709nm and third harmonic generation (THG) of 1064nm, cannot make use of $KNbO_3$. For SHG of 709nm radiation we use non-critically phase matched dueterated rubidium dihydrogen arsenate (RD*A). Since the non-linear coefficient of RD*A is significantly less than that of $KNbO_3$, we note that the achievable output power is significantly lower. The THG process is performed in a critically phase matched LiB_3O_5 crystal. We have performed a numerical calculation to determine the optimal focussing condition for the two input beams in this process[3]. Unfortunately, as this process is not NCPM it suffers from limited angular acceptance and walk-off. We display our experimental results for all non-linear processes on figure 2. The solid curves show the theoretically calculated output powers based on the BK approach[4] for the SHG processes and a numerical method for the THG process.

From the known noise levels of the photodiodes that are used as phase detectors in our OFIDs, we calculate that we need to generate at least 10nW of non-linear power to ensure that the PLL performance is not limited by the detection process. We note that all non-linear processes shown on figure 2 produce at least this much radiation.

(iii) Phase-locked Diode lasers

We have made a detailed study of the mechanisms which cause the output frequency of an OFID to differ from the exact mid-point of its input interval, and they have been presented in another paper in this issue[5]. In summary, the main causes for these deviations are: that the gain of the PLL circuit is insufficient to suppress all of the intrinsic fluctuations of the slave laser, cycle-slip events in the PLL circuit, and phase fluctuations induced in the output of the OFID, by fluctuations in the optical path length between the OFIDs (caused by thermal fluctuations or air density fluctuations). We present the effect of these fluctuations on the synthesizer stability on Figure 3.

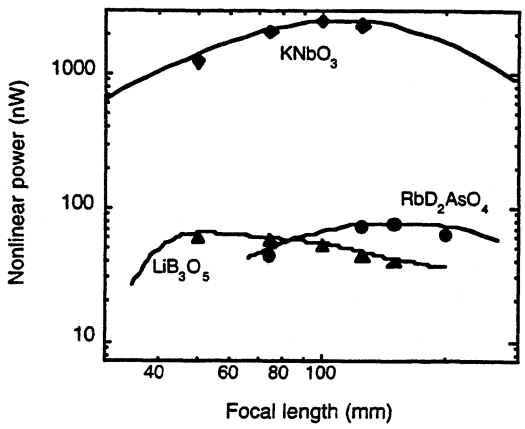


Figure 2. Non-linear power created in the non-linear processes used in this synthesizer design.

4. Optical Frequency Comb Generation

The final element of the synthesis process is provided by a Fabry-Pérot optical phase modulator that generates a frequency comb with a span of several terahertz. The modulator is driven by the high stability UWA sapphire microwave oscillator[6]. Our collaborators have developed techniques to ensure that the comb generation process is reliable[7]. We will incorporate the OFCG into the frequency chain during the second part of this year.

OFCGs suffer from a number of fluctuations that lead to either a non-uniform comb spacing, or to a comb spacing that differs from the driving frequency. In either case this will introduce errors into the synthesis process. The most critical of these fluctuations, from the point of view of short term stability, are fluctuations in the cavity length. The kth order sideband of the comb has been created by light which has cycled around the cavity ~$k/2\beta$ times where β is the single pass phase modulation. Any motion of the mirrors will impose a Doppler shift on the kth sideband that is $k/2\beta$ times larger than that applying to the carrier[2]. To gauge the severity of this effect, let us consider a monolithic OFCG that has previously had its comb uniformity tested[8]. We assume that the majority of the longitudinal mirror motion in this sort of cavity is due to excitation of the fundamental longitudinal acoustic mode (~ 10^5 Hz for the monolithic OFCG in [8]). We find that the excitation amplitude only needs to be 3nm to be consistent with the reported frequency stability. We have

plotted the effect of OFCG frequency fluctuations, of this magnitude, on the optical synthesiser on figure 3 (curve c). It is clear that without some form of control system, it will not be possible to make use of such an OFCG and achieve our performance goal.

We have calculated the noise floor of an OFCG stabilisation technique proposed by Telle et al[7]. In this scheme, the relative phase of the carrier transmitted through the OFCG in the two polarisations is compared to derive an error signal that can be used to stabilise the frequency offset between the input laser field and the modes of the OFCG cavity. We calculate a noise floor for this stabilisation technique as shown in figure 3 (curve d). To actually reduce the fluctuations to this level would require a wide bandwidth control servo as the majority of the phase noise lies near to the fundamental acoustic mode (~10^5 Hz).

5. Frequency Chain Performance

We display on figure 3 the possible contributors to fluctuations in the complete synthesizer. We have calculated these Allan Frequency Deviation curves from the measured phase noise of the OFIDs (curve a) and the path length fluctuations (curve b)[5]. In addition we have displayed the effect of OFCG fluctuations when free running (c), and when it has been stabilised to the incident laser field (d) as calculated in the last section. The phase noise of all of these sources has been filtered with a low pass filter corresponding to the bandwidth of the synthesizer feedback loop (~10kHz). Only those fluctuations that fall within the bandwidth of the feedback loop can affect the optical source. Finally we have displayed the effect on the synthesiser of residual phase noise (curve (e)) of the Nd:YVO_4 laser. This noise arises because of the limited gain and bandwidth of the synthesis feedback loop. We compare these noise sources to the frequency stability of the cryogenic sapphire microwave oscillator, which acts as the frequency reference for the synthesizer (dashed curve). In the short term, the frequency synthesis is limited by residual frequency instability of the Nd:YVO_4 laser, in the medium term by path length variations, and in the long term, the only limit is the performance of the frequency reference. We note that by using a stabilised OFCG, the predicted frequency stability of the synthesis process is of the order of 1 Hz over a few seconds of integration time.

7. Conclusion

A frequency synthesizer based upon frequency-interval division and frequency comb generation has been examined and appears to offer the promise of optical frequency synthesis at the 1 Hz level. Using the optical and microwave frequency standards available at UWA we will be able to independently verify this performance. A stable, synthesized, optical source could be a very particularly convenient interrogating oscillator for the new generation of laser cooled optical frequency standards.

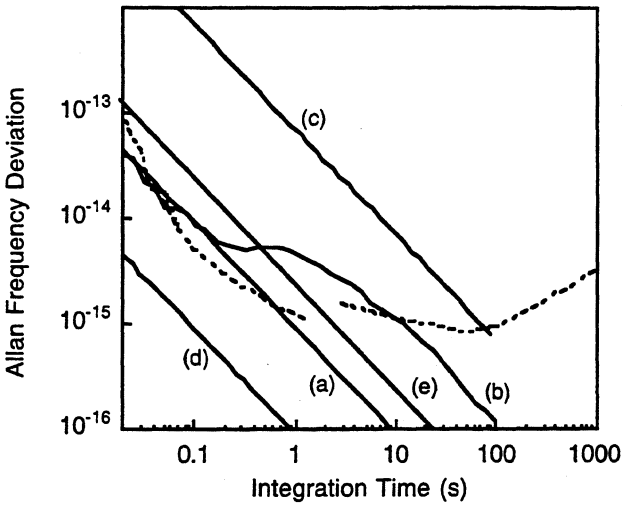


Figure 3. Allan Deviation of Noise Sources in Optical Synthesizer. Dashed Line – Frequency Stability of Sapphire Oscillator. See text for nomenclature.

6. References

1. HR Telle, D Meschede, and TW Hänsch, "Realization of a new concept for visible frequency division: phase locking of harmonic and sum frequencies", Opt. Lett. **15**, pp. 532-534 (1990).
2. M Kourogi, B Widiyatomoko, Y Takeuchi, and M Ohtsu, "Limit of Optical-Frequency Comb Generation Due to Material Dispersion", IEEE J. Quantum Electron. **31**, pp. 2120-2125 (1995).
3. AV Smith, SNLO Non-linear Optics, (Sandia National Laboratories, Albuquerque, 1996).
4. GD Boyd and DA Kleinman, "Parametric Interaction of Focused Gaussian Light Beams", J. Appl. Phys. **39**, pp. 3597-3639 (1968).
5. RP Kovacich and AN Luiten, "Frequency Stability Limits of Optical Frequency Intervals in New Generation Optical-to-Microwave Frequency Chains", in *13th European Frequency and Time Forum and 1999 IEEE International Frequency Control Symposium*, (1999), this volume.
6. AN Luiten, AG Mann, NJ McDonald, and DG Blair, "Latest Results of the UWA Cryogenic Sapphire Oscillator", in *Proc. 49th Freq. Control Symp. (FCS)* pp. 433-437 (1995).
7. HR Telle and U Sterr, "Generation and Metrological Application of Optical Frequency Combs" in *Optical to Microwave Frequency Synthesis and Frequency Standards:*, Topics in Applied Physics, Ed. AN Luiten (Springer-Verlag, Berlin, 1999), in press.
8. T Udem, J Reichert, TW Hänsch, and M Kourogi, "Accuracy of optical frequency comb generators and optical frequency interval divider chains", Opt. Lett. **23**, pp. 1387-1389 (1998).

GENERATION OF COHERENT OPTICAL RADIATION BY ELECTRONIC MEANS: THE ELECTRO-OPTICAL PARAMETRIC OSCILLATOR

A. Wolf and H. R. Telle;
Physikalisch-Technische Bundesanstalt Project 4.301, P.O. Box 3345, D-38023 Braunschweig, Germany;
email: andreas.wolf@ptb.de

ABSTRACT

We describe a novel type of optical oscillator whose gain element is located outside the resonator. Although this is a common technique for microwave oscillators, separated-gain-schemes are rarely used in the optical regime However, the Q-factor of an empty resonator may be orders of magnitude higher than its counterpart with internal gain element, leading to substantial reduction of the quantum-noise limited emission line width. Thus, oscillators of this kind are highly suited for demanding applications like interrogation oscillators for optical frequency standards or flywheel oscillators in synthesis chains.
As a further advantage of our approach, its gain can be chosen almost arbitrarily high since it is provided by electronic amplifiers.
We demonstrate, that the oscillator is capable of generating highly coherent light, starting from incoherent pump light. The incoherent optical pump field used during these initial experiments has been prepared by filtering the coherent emission of a laser diode. However, estimations indicate, that commercially available light emitting diodes have the potential to act as pump sources as well.

INTRODUCTION

Oscillators for the optical frequency synthesis should generate highly coherent light fields with extremely low frequency and intensity noise levels. These requirements are rarely met by free running lasers. Hence, additional active control is required. External control schemes [1,2,3,4] are desirable for this purpose since suitable servo means are rarely available from normal lasers.
Simple dual servo loops suffer from cross-talk problems: intensity noise is reduced, but reenters as excess phase noise and vice versa. Thus, delicate multiple nested loops are required which can cause severe instability problems.
The novel approach outlined here, is much less sensitive to cross-talk between multiplicative intensity and frequency noise contributions. It even tolerates filtered radiation from a thermal light source as input which displays a zero-centered Gaussian field amplitude distribution. As shown below, such inputs can be converted into coherent, unambiguously phase-trackable output radiation with doughnut-shaped field amplitude distribution.
The term 'coherent' is used in the following sense throughout this paper: Coherent radiation displays, after appropriate filtering, a field amplitude distribution which vanishes for small amplitude values, e. g. for values smaller than 50% of the rms value so that the phase of the field is always defined. For 'incoherent' light on the other hand no such filtering does exist and the field amplitude has a certain probability to vanish. In this situation the phase of the field might be not defined. The term 'coherent' must be clearly distinguished from the term 'narrowband' as the spectral width of light might be small even when it is incoherent.

EXPERIMENTAL SETUP

The novel type of oscillator, referred to as electro-optical parametric oscillator (EOPO) in the following, consists of 4 essential components, see Fig. 1:
i) an optical high-Q resonator (HQR) which stores the phase of the optical field,
ii) a photo diode (PD2) which converts the optical into an electronic signal,
iii) a rf amplifier chain (AMP) which provides the parametric gain and
iv) an electro-optic modulator (EOM) which reconverts the amplified electronic signal into the optical domain.

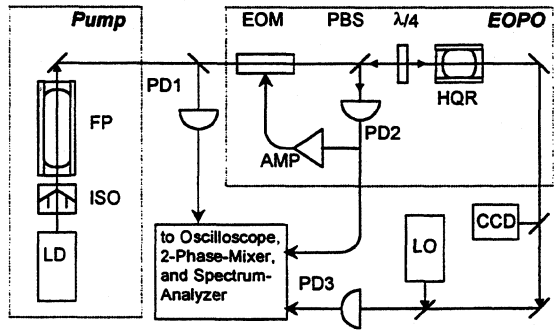


Fig. 1: Experimental set-up. Free-running GaAlAs laser diode LD, isolator ISO (120 dB isolation ratio), optical filter FP, $LiNbO_3$ electro-optic modulator EOM, polarizing beam splitter PBS, high-Q-resonator HQR, optical local oscillator LO.

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

The scheme has close similarities to a single-resonant four-photon optical-parametric oscillator (4P-OPO)[5]. In the 4P-OPO the components ii) to iv) are combined in one building block, i. e. a nonlinear element with large $\chi^{(3)}$-susceptibility which accomplishes strong cross-phase-modulation.

Assuming only one resonance of the HQR which is detuned by $\Delta\nu$ from the center frequency ν_P of the incident pump field, the EOPO operation can be understood as follows: PD2 detects the signal reflected from HQR and thus, besides the dc components, the beat-note at $\Delta\nu$ which arises from the superposition of the out-coupled internal field of HQR (e. g. noise) and the input field component which is directly retro-reflected by the cavity entrance mirror.

This beat-note is amplified and fed to the EOM yielding symmetric FM side bands at $\nu_P \pm \Delta\nu$. Hence, the frequency of one of these, the resonant side-band (RS), matches the cavity resonance at ν_{HQR}. As a result, the excitation of HQR increases for sufficiently large rf gain of AMP and properly chosen round-trip phase of this opto-electronic feedback loop. Consequently, the rf photo current signal of PD2 grows, leading to stronger FM side bands and vice versa. Hence, a sustained oscillation sets in and the optical frequency of RS becomes almost independent of fluctuations of the pump frequency.

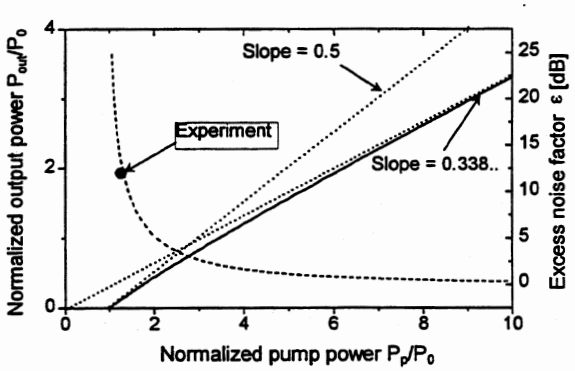


Fig. 2: Calculated pump-to-output power transfer function (solid line, left vertical axis) and excess noise factor ε (dashed line, right vertical axis), P_0=pump power at threshold.

Once the oscillation threshold (at pump power level P_0) is exceeded, the FM side band amplitude grows with increasing pump power, see Fig. 2. Notably, the asymptotic slope efficiencies at threshold

$$\lim_{P_p \to P_0} \frac{dP_{out}}{dP_p} = 0.5$$

and for large pump levels

$$\lim_{P_p \to \infty} \frac{dP_{out}}{dP_p} = J_1^2(1.84...) = 0.338...$$

are given by mathematical properties of the Bessel functions and independent of system parameters. This slope efficiency of 0.338.. (at the maximum of J_1) is equivalent to an insertion loss of about -4.7 dB.

The corresponding pump-to-output intensity noise transfer function,

$$\varepsilon(P_p) = \left(\frac{dP_{out}}{dP_p} \frac{P_p}{P_{out}} \right)^2,$$

named excess noise in the following, is plotted as a dashed line in Fig. 2.

The experimental set-up is depicted in Fig. 1. The EOPO pump light is prepared with the help of a Fabry-Perot resonator FP (l = 60 cm , resonance width $\delta\nu_{FP}$ = 1 MHz (HWHM)). It filters the emission of a GaAlAs laser diode (P_{out}= 20 mW, λ=855 nm, line width = 25 MHz), yielding a pump linewidth $\delta\nu_p$ of 1 MHz. In this filtering process the coherence of the light is destroyed due to the fast phase diffusion of the input light and the storage effect of the FP on a much slower time scale. The center frequency of the LD emission is stabilized with respect to the center of a resonance of FP using a slow (<10 Hz) Pound-Drever-Hall servo loop.

The EOPO setup is shown in the upper right part of Fig 1. The four key components mentioned above are the high finesse Fabry-Perot resonator HQR (l=7,5 cm, resonance width $\delta\nu_{HQR}$= 15 kHz (HWHM)), the fast photo diode PD2 (200 MHz bandwidth), the rf amplifier chain AMP (gain = 22 V/μA at 130-180 MHz, $P_{out,\ max}$ =12 W) and the electro-optical modulator EOM ($LiNbO_3$, dimensions: 25×1×1 mm^3). The overall loop delay time was estimated to be τ_L=40 ns which leads to a 25 MHz mode spacing of the oscillation loop and an effective control bandwidth of $f_c = 1/2\pi\tau_L \cong 4$ MHz.

As a result of this finite round trip delay time, the phase condition mentioned above can be fulfilled for several frequency offsets $\Delta\nu$ yielding a spectrum of equidistant longitudinal modes of the oscillation loop with a mode spacing of $1/\tau_L$. Thus, the spectral width of the pump has to be sufficiently narrow in order to avoid multi-mode oscillation giving rise to un-

desired noise mixing products, in particular in the case of incoherent pumping.
Furthermore, the storage time of HQR has to be sufficiently long in this case in order to bridge power drop-outs of the pump, which typically last for time intervals of the order of the inverse pump line width.
Combining both conditions, the line width $\delta\nu_p$ of the incoherent pump light has to fulfill the relation $\delta\nu_{HQR}<\delta\nu_p<f_c$ in order to generate coherent output signals with a normalized rms intensity modulation index $\sigma_{I,EOPO}$ well below unity. In our case, $f_c/\delta\nu_{HQR}\cong 270$, minimum intensity noise is found for a pump line width of about 1 MHz, see simulation shown in Fig. 3.

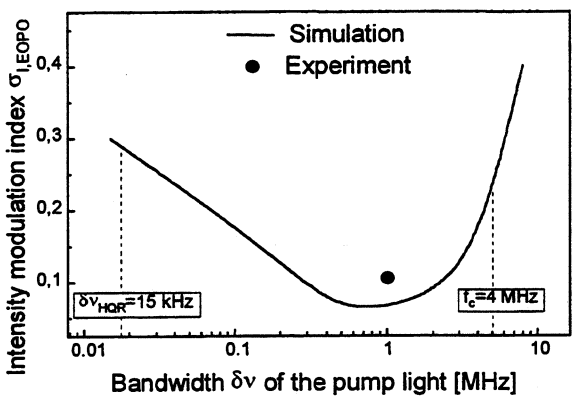


Fig 3: Normalized rms intensity modulation index of the generated signal light for different bandwidths of the pump light. Line: Result of computer simulation; point: experimental result. Measuring time: 400 µs.

RESULTS

Monitoring the output signal of HQR with a CCD camera and a photo diode (PD3), we observed sustained electro-optic-parametric oscillation with a threshold of about 80 µW of average input power at frequency offsets $\Delta\nu\approx 150$ MHz.
The average EOPO output power was 0.5 µW for a typical input power of 100 µW and oscillation in a fundamental transverse mode. Oscillation in higher transverse modes could be observed too, as verified by the CCD camera. Fig. 4 depicts rf-spectrum analyzer traces of the spectral densities of relative intensity noise S_I, i. e. the spectral density of optical power fluctuations normalized to the square of the average power. Curve a) and b) show the spectra of the pump and EOPO output, as detected by PD1 and PD3, respectively (Si PIN diodes, BW=3 MHz). The intensity noise density spectrum of the pump, curve a), is almost flat up to a corner frequency of $\delta\nu_{FP}\cong 1$ MHz, whereas a low-pass roll-off due to the storage effect of FP is observed above 1 MHz. The measured value of $S_{I,P}\cong 10^{-6}$ Hz^{-1} results in a normalized rms intensity modulation index of $\sigma_{I,P}\cong 1$ as expected for a Gaussian process.
For the EOPO output signal, curve b, we find a ε=12 dB higher noise level for Fourier frequencies below $\delta\nu_{HQR}$ and an average pump power of 100 µW in close agreement with expected transfer function of Fig. 2.

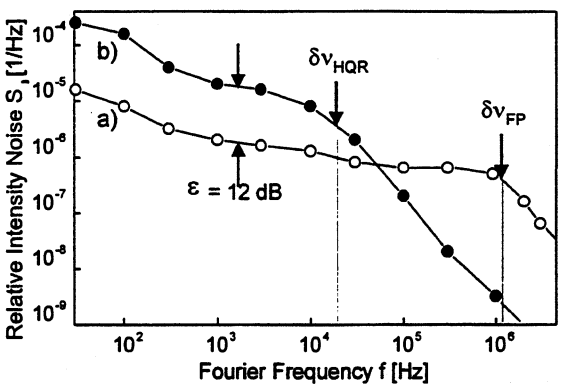


Fig. 4: Measured spectral densities of relative intensity noise of EOPO pump a) and EOPO output b). BW=10 Hz

Curve b) shows the expected low pass characteristic of HQR. One estimates from this curve a rms intensity modulation index $\sigma_{I,EOPO}$ of well below 0.5. Thus the properties of the EOPO output signal cannot be explained in terms of simple filtering of the pump signal. The different properties of both fields become even more obvious if the intensity distribution functions are compared. These functions have been measured with identical photo diodes PD1 and PD3, as above, and using a digital oscilloscope. One clearly finds an exponential distributed pump intensity, curve a), which appears as a straight line in the linear-logarithmic presentation used in Fig. 5.
The intensity distribution corresponding to the EOPO output, curve b), however, shows a remarkably different behavior. Here, the intensity always remains at a non-zero level. The distribution function even remains well separated from curve (c) which represents the distribution of the total electronic noise (light blocked). According to the above frequency domain analysis, one estimates a rms intensity modulation index of about 0.11 (see Fig. 3). The homodyne-detected intensity distributions of Fig. 5 give already hints to the corresponding distributions of the complex field amplitude: a Gaus-

sian bell for the pump field and a doughnut for the EOPO output field.
To confirm this expectations, we have measured both distributions of the complex field amplitude using two-phase heterodyne techniques. In this scheme, the beat-note between the field under consideration and a reference field from an optical local oscillator were detected by a photo diode. Then, both quadrature components of this beat-note were independently down-converted to the base band with the help of two rf mixers

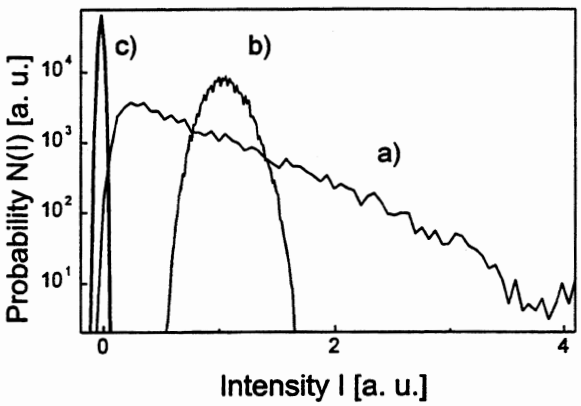


Fig. 5: Measured intensity distributions of EOPO pump a) and EOPO output b). BW= 3 MHz, sampling rate = 125 Msamples/s, record length = 50 ksamples. Electronic noise c) causes negative intensity values in curve a)

One of these mixers was directly driven by a rf local oscillator (LO) oscillating at the beat-note frequency whereas a 90°-phase-shifter was inserted in between LO and the other mixer.

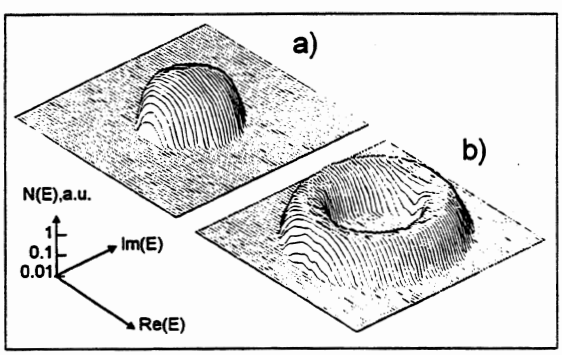


Fig. 6: Measured distributions of the complex amplitude of pump a) and EOPO output b). Local oscillators used for heterodyning: EOPO, curve a), and an independent extended-cavity laser diode, curve b). BW = 3MHz

Both mixer output signals were low-pass filtered (f_c=3MHz) and recorded with a two-channel digital oscilloscope. Distributions of the complex field amplitude were calculated from these data and plotted as a 3D plot, see Fig. 6.
The distribution of the pump field, Fig. 6a, was deduced from the 150 MHz in-loop signal as detected by PD2.
It corresponds to the beat-note between the pump field after passing EOM and the coherent, narrow-band field from HQR.
Other contributions to this signal are negligible. Hence, this 150 MHz in-loop signal represents the properties of the unperturbed EOPO pump field. The Gaussian-like, zero-centered distribution found in Fig. 6a, clearly indicates its lack of coherence.
The field distribution of the EOPO output, Fig. 6b, was measured with the same rf set up as above at a beat-note frequency of 40 MHz. Here, an additional extended-cavity GaAlAs laser diode was employed as optical LO. The observed doughnut shape with vanishing probability at the center underlines the coherence properties of the EOPO output signal.

CONCLUSION

We have demonstrated an optical oscillator scheme that separates the gain medium from the resonator. It promises substantial reduction of the quantum-noise limited emission line-width making it highly suitably as fly-wheel- or interrogation oscillator in optical frequency standards. It has been demonstrated, that the EOPO is capable of converting incoherent, non-phase-trackable input radiation into narrow band, coherent output radiation.
Extrapolation of the experimental results to lower pump power density shows, that the EOPO scheme has the potential to convert the incoherent emission of a light emitting diode into laser-like, coherent output radiation.

Acknowledgement: This research has been supported by the Deutsche Forschungsgemeinschaft through SFB 407.

REFERENCES

1. J. L. Hall and T. W. Hänsch, Opt. Lett. **9**, 502 (1984)
2. M. Zhu and J. L. Hall, JOSA **B10**, 802 (1993)
3. H. R. Telle, Proc. EFTF 1997 Neuchâtel, 109 (1997)
4. C. C. Harb, M. B. Gray, H.-A. Bachor, R. Schilling, P. Rottengatter, I. Freitag, and H. Welling, IEEE J. **QE-30**, 2907 (1994)
5. J. Frahm, Ann. Physik **30**, 173 (1973)

ABSOLUTE FREQUENCY MEASUREMENT IN MID INFRARED.

G. D. Rovera, O. Acef, and A. Clairon
BNM-LPTF Observatoire de Paris, France
e-mail: daniele.rovera@obspm.fr

abstract

A brief description of the measurement chain linking the CO_2/OsO_4 secondary frequency standard to the primary frequency standard is given. The measured frequency stability shows that the measurement system reaches the flicker floor of the secondary frequency standard, below the 10^{-14} level in 100 s. Results of a year of absolute frequency measurements of a CO_2/OsO_4 mid infrared secondary frequency standard are reported.

1 Introduction

The frequency of a CO_2 laser stabilized on the absorption of OsO_4, developed in our laboratory to serve as secondary frequency standard, was measured more than fifteen years ago with an uncertainty of $2 \cdot 10^{-12}$ [1]. This laser was subsequently used as frequency standard for the measurement of several frequencies in the near infrared [2], in the visible [3] and in the ultraviolet [4], without further measurement of its frequency.

During the last decade the frequency stability and the reproducibility of the CO_2/OsO_4 frequency standard have been continuously improved [5], and the CIPM has recently recommended this secondary frequency standard for the *mise en pratique for the definition of the metre* [6].

To measure this standard with increased accuracy, the frequency multiplication chain described in [1] was renewed, and we can now report the result of one year of measurements with uncertainty in the 10^{-13} range.

2 System description

A simplified schematic of the measurement chain is shown in Fig. 1.

The microwave synthesizer was originally developed for the BNM-LPTF atomic fountain primary frequency standard [7]. The output frequency is

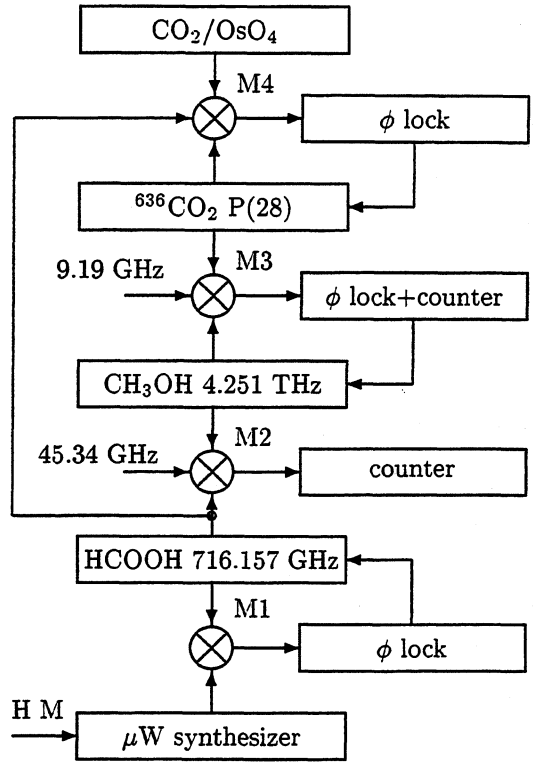


Figure 1: Simplified schematic of the measurement chain

fixed at 9.190 GHz by phase locking the master quartz oscillator to a hydrogen maser. The frequency of the maser is periodically compared against the BNM-LPTF primary frequency standard, which has a claimed uncertainty of $2 \cdot 10^{-15}$ [8]. The microwave synthesizer also delivers 10 MHz and 100 MHz signals which, via distribution amplifiers, are used as frequency references throughout the measurement chain.

The formic acid optically pumped far infrared laser (OPFIRL), operating at 716.157 GHz, is phase locked to the 78^{th} harmonic of the microwave syn-

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

thesizer by means of Schottky diode mixer M1. A detailed description of this part was previously published [9]. The phase noise spectral density S_ϕ of the phase locked OPFIRL was measured by beating the submillimetric signal with the 16^{th} harmonic of a Gunn oscillator phase locked to a second microwave synthesizer. From the measured S_ϕ reported in Fig.2 a fast linewidth of about 0.3 Hz can be estimated.

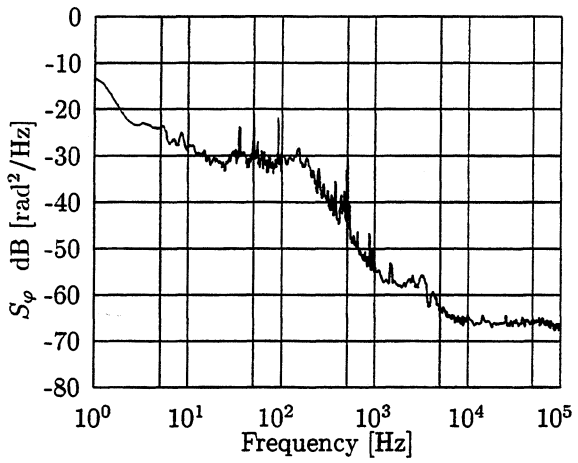


Figure 2: Measured phase noise spectral density S_φ in dB [rad^2/Hz]of the OPFIRL at 716 GHz.

The Schottky diode mixer M2 generates the difference between the 6^{th} harmonic of the 716.157 GHz, the 45.34 GHz signal of from a Gunn oscillator, and the radiation of a CH_3OH OPFIRL at 4.25 THz. The Gunn oscillator is phase locked to the 5^{th} harmonic of 9.19 GHz. The beat note at this point exhibits a signal to noise ratio of at least 20 dB in a 10 kHz bandwidth which is not enough for direct counting; however it is sufficient to steer a tracking oscillator.

The MIM diode M3 generates the beat signal between the 7^{th} harmonic of 4.25 THz, the radiation of the $^{636}CO_2$ transfer laser and the 9.19 GHz from the microwave synthesizer. This 9^{th} order mixing, where a signal to noise ratio of 20 dB in a 10 kHz bandwidth is required, is the most critical point of the chain. Ninety percent of the radiation of the CH_3OH OPFIRL is focused on M3 by an off-axis parabolic mirror; the risk of thermal damage to the whisker during the set up procedure is quite high, but when a stable contact is found, it will last for several hours. The beat note from M3, filtered by a tracking oscillator, is frequency divided in order to phase lock the CH_3OH laser. In this way the virtual dynamics of the phase comparator of the phase lock is increased by the division factor, ensuring phase locking without cycle slips even in case of poor control bandwidth. Actually the useful bandwidth is limited at 200 Hz by some mechanical resonance of the laser structure. The proper operation of this phase lock is checked by permanently counting the output of the tracking oscillator.

The MIM diode M4 generates a beat signal between the CO_2/OsO_4 frequency standard, the radiation from the formic acid OPFIRL and the $^{636}CO_2$ transfer laser is then used to phase lock the latter to the reference. As the mixing order is 3, the power of the laser radiation can be kept low, ensuring very stable behavior of the MIM diode, with a signal to noise ratio higher than 30 dB in a 10 kHz bandwidth. Also for this phase lock the beat signal is filtered by a tracking oscillator and frequency divided.

The described system was used to measure directly the frequency of standard L1 or L2 during the last year. When both the CO_2/OsO_4 frequency standard was operating simultaneously, the system devoted to routine comparisons between the two frequency standards was used to measure the relative frequency of the two standards. By synchronizing all measurements frequency instability of intermediate standard can be removed.

3 Results

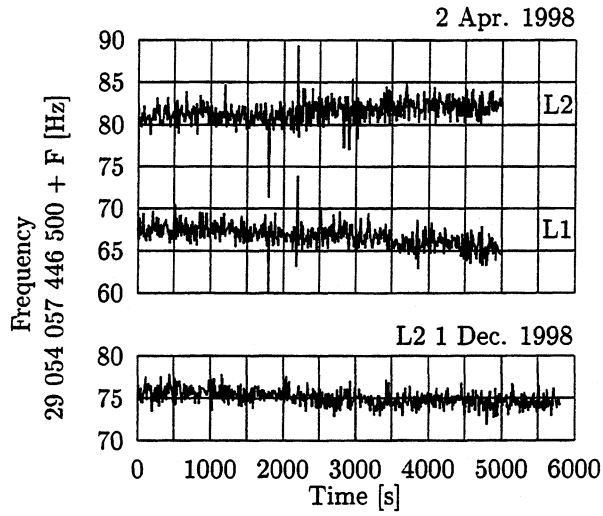


Figure 3: Frequency of CO_2/OsO_4 frequency standard measured against the hydrogen maser with 10 s gate time.

Fig. 3 reports some significant record of the frequency of the CO_2/OsO_4 standard measured with the new chain. The upper part shows the frequency of L1 and L2 recorded on April 2 1998 in a 5000 s continuous run. The lower record shows the frequency of standard L2 measured during the absolute measurement of the He – Ne/CH_4 optical standard [10]. Due

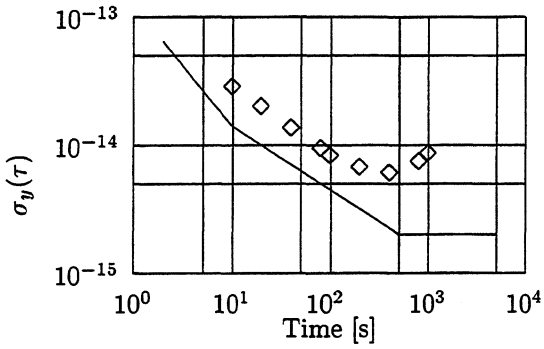


Figure 4: Allan standard deviation of the CO_2/OsO_4 frequency standard measured against the hydrogen maser. The line represent the expected stability of the maser.

to the intense human activity in the laboratory at this time the formic acid OPFIRL went out of lock a couple of time and the corresponding frequency point are removed from the plot. From the Allan standard deviation relative to this run, reported in Fig. 4 it is evident that the measuring system contributes with a white frequency noise for integration time lower than 100 s, while for longer time the flicker floor of the CO_2/OsO_4 frequency standard is reached.

All reported measurements are obtained by synchronously counting the beat notes on mixers M2 and M3, even when the CH_3OH OPFIRL is phase locked. The frequency is obtained by adding the frequency of the beat note on M3 with 7 times the frequency of the beat note in M2. This procedure is not ideal because the blanking time necessary for the syncronization of counters transforms quantization noise and white phase noise in white frequency noise by an aliasing phenomena. Nevertheless this is the procedure that give the best short term stability because the linewidth of the the methanol OPFIRL is not narrowed by the phase lock. By increasing the control bandwidth of this laser we expect to reduce the linewidth at the theoretical value of one seventh the linewidth of the CO_2 laser and therefore increase the short term stability of the measuring system.

To verify that the phase lock of the formic acid OPFIRL does not limit the stability we have measured the beat note between the radiation of this laser and the 16^{th} harmonic of the Gunn oscillator, both phase locked on the same microwave synthesizer. The measured Allan standard deviation of this beat is $\sigma_y(\tau) = 2 \cdot 10^{-14} \tau^{1/2}$, for τ varying between 1 and 1000 s guarantees that the system is not limited by the phase lock of this laser.

The absence of cycle slips in the PLL of the tracking oscillators was checked with the usual technique of comparing the frequency of two oscillators with different bandwidths [3]. A further verification comes from the fact that the measured frequency does not change when the signal to noise ratio is reduced by decreasing the power of the CH_3OH laser.

By comparing the phase, over several days, of a round trip of the signal on the double coaxial line carrying the signal of the maser we can conclude that the maximum frequency shift due to the temperature cycling of the cable is less than $2 \cdot 10^{-15}$. This is the major source of type B uncertainty in the measurement system.

Up to now the frequency of the CO_2/OsO_4 standard L2 was measured 20 days distributed over one year for the CO_2/OsO_4 standard L2. The standard L1 was operating only during the first period, because during the fall 1998 it was used as transfer oscillator for the measurement of the methane stabilized He – Ne laser at 3.39 μm [10].

All possible way of averaging data (i.e equal weight for each day, or each file, or each 100 s sample) gives for the standard L2 the value:

$$\nu_{L2} = 29\,054\,057\,446\,579\,(4)\,\text{Hz} \qquad (1)$$

We recall that this value holds for BNM-LPTF CO_2/OsO_4 frequency standard L2 operating on the (R10) line in standard conditions [5].

An evaluation of the frequency of the corresponding line of OsO_4 and more informations about the reproducibility of the standard can be found in these proceedings [11].

References

[1] A. Clairon, B. Dahmani, A. Filimon, and J. Rutman, "Precise frequency measurements of CO_2/OsO_4 and He-Ne/CH_4 stabilized lasers," *IEEE Trans. Instrum. Meas.*, vol. 34, pp. 265–268, 1985.

[2] D. Touahri, O. Acef, A. Clairon, J. J. Zondy, R. Felder, B. de Beauvoir, F. Biraben, and F. Nez, "Frequency measurement of the $5S_{1/2}(F = 3) - 5D_{5/2}(F = 5)$ two-photon transition in rubidium," *Optics Comm.*, vol. 133, pp. 471–478, 1996.

[3] O. Acef, J. J. Zondy, M. Abed, G. D. Rovera, A. H. Gerard, A. Clairon, P. Laurent, Y. Millerioux, and P. Juncar, "A CO_2 to visible optical frequency synthesis chain: Accurate measurement of th 473 THz He-Ne/I_2 laser," *Optics Comm.*, vol. 97, pp. 29–34, 1993.

[4] B. de Beauvoir, F. Nez, L. Julien, B. Cagnac, F. Biraben, D. Thouari, L. Hilico, O. Acef, A. Clairon, and J. J. Zondy, "Absolute frequency measurement of the 2S-8S/D transitions

in hydrogen and deuterium: New determination of the Ridberg constant," *Phys. Rev. Lett.*, vol. 78, no. 3, pp. 440–443, 1997.

[5] O. Acef, "Metrological proprieties of CO_2/OsO_4 optical frequency standard," *Optics Comm.*, vol. 134, pp. 479–486, Jan. 1997.

[6] "Révision de la mise en pratique de la definition du metre," *CIPM/97*, 1997. 86 session Rec 1-C11997.

[7] G. D. Rovera, G. Santarelli, , and A. Clairon, "Frequency synthesis chain for the atomic fountain primary frequency standard," *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, vol. UFFC 43, pp. 354–358, May 1996.

[8] E. Simon, P. Laurent, and A. Clairon, "Measurement of the Stark shift of the Cs hyperfine splitting in an atomic fountain," *Phys. Rev. A*, vol. 57, Jan. 1998.

[9] G. D. Rovera, "Low frequency noise optically pumped FIR laser with frequency stabilized pump," *Optics Comm.*, vol. 143, pp. 247–251, 1997.

[10] O. Acef, A. Clairon, G. D. Rovera, F. Ducos, L. Hilico, G. Kramer, B. Lipphardt, A. Shelkovnikov, E. Koval'chuk, E. Petrukhin, D. Tyurikov, M. Petrovskiy, M. Gubin, R. Felder, and S. Lea, "1988 absolute frequency measurements with a set of transportable He-Ne/CH_4 optical frequency standards," in *Proc. 13^{th} European Frequency and Time Forum, Besançon*, 1999 (this issue).

[11] O. Acef, G. D. Rovera, L. Hilico, and F. Ducos, "Performances of OsO_4-stabilized CO_2 lasers as optical frequency standards near 29 THz," in *Proc. 13^{th} European Frequency and Time Forum, Besançon*, 1999 (this issue).

OPTICAL FREQUENCY SYNTHESIS IN THE NEAR IR WITH APPLICATION TO TRAPPED ION STANDARDS

S.N. LEA, G. HUANG, H.S. MARGOLIS, G.M. MACFARLANE, P. TAYLOR, AND P. GILL

National Physical Laboratory, Teddington, Middlesex TW11 0LW, UK

ABSTRACT

An optical frequency chain is under construction at NPL, with the object of making improved absolute frequency measurements of forbidden transitions in cold, trapped single ions of strontium and ytterbium, using as reference a methane-stabilised helium-neon laser frequency standard.

1. INTRODUCTION

The quadrupole $^2S_{1/2} - {}^2D_{5/2}$ transition at 444.8 THz (674 nm) in the $^{88}Sr^+$ ion is the first optical transition in a trapped ion to be recommended by the CIPM as a radiation for the realisation of the metre [1]. This recommendation was reached as a result of absolute frequency metrology of laser-cooled, single trapped strontium ions at NRC, Canada [2] and in our laboratory [3]. The relative uncertainty of the frequency measurements at NPL was limited at the level of 1×10^{-10} by the Fabry-Perot wavelength comparator used to relate the ion trap probe laser frequency to the iodine-stabilized helium-neon laser ($HeNe/I_2$) reference at 473 THz (633 nm). Ideally, a fully phase-coherent frequency chain is required to enable the ion trap frequency to be directly intercompared with the caesium primary microwave frequency standard. However, our initial goal is to make improved absolute frequency measurements relative to a well-characterised mid-infrared frequency standard, a methane-stabilised helium-neon laser system ($HeNe/CH_4$) at 88.376 THz (3.39 μm). Our $HeNe/CH_4$ standard is a TOFS-60 device, constructed by the group of M.A. Gubin at the P.N. Lebedev Physical Laboratory [4]. The absolute frequency of this standard is periodically measured through particpation in a series of intercomparisons [5].

2. PRELIMINARY CHAIN TO STRONTIUM

The $HeNe/CH_4$ standard is a particularly convenient choice as a frequency reference for the Sr^+ ion standard as the frequency of the latter is within 3 THz of the fifth harmonic of the former. The scheme of the chain is illustrated in fig. 1. The first step is to synthesise the fourth harmonic of the $HeNe/CH_4$ standard, in the near-infrared at 353.5 THz (848 nm). This frequency lies in a spectral region where high-power, narrow-linewidth oscillators such as Ti:sapphire lasers and extended-cavity diode lasers can be used to drive further non-linear linkages. The Sr^+ frequency is to be reached by difference frequency mixing (DFM) of radiation at 674 nm and 841 nm, generating 3.39 μm radiation which can be referenced back to the $HeNe/CH_4$ frequency. The remaining interval of 2.9 THz between 841 nm and 848 nm will be bridged by an optical frequency comb (OFC) [6].

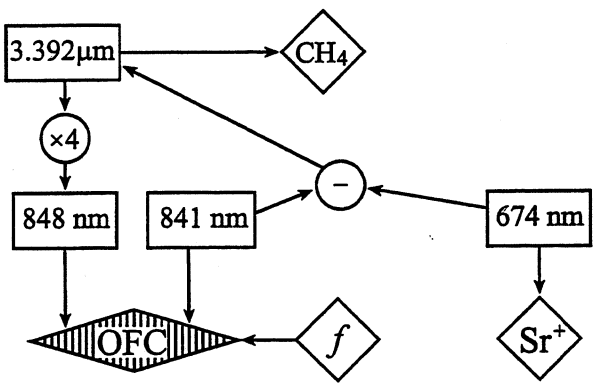


Fig. 1. Scheme of preliminary frequency chain intercomparing the methane-stabilised HeNe standard at 3.392 μm with the strontium ion standard at 674 nm. The OFC is driven at microwave frequency *f*.

2.1 Nd:YLF-Pumped CW SROPO

The output power of the TOFS-60 is insufficient to directly drive a second harmonic generation (SHG) process. As a high-power, single-frequency, 3.39 μm source we have constructed a singly-resonant cw optical parametric oscillator (SROPO), pumped by a Nd:YLF ring laser at 1053 nm [7]. Fig. 2 is a schematic of this device. The Nd:YLF ring laser is end-pumped by two diode lasers, each of nominally 15 W output power at 805 nm. A Brewster plate (BP) in the laser cavity forces lasing on the 1053 nm line in preference to the orthogonally-polarised 1047 nm line, to reduce thermal effects in the laser crystal. An optical diode (OD) imposes uni-directional operation. Single-frequency operation is confirmed by monitoring the laser output with an optical spectrum analyser (OSA). Mode-hop-free operation is maintained for periods of several minutes, without any cavity stabilization. A maximum of 7 W of single frequency output at 1053 nm has been obtained with 30 W of diode pump power.

0-7803-5400-1/99/ $ 10.00

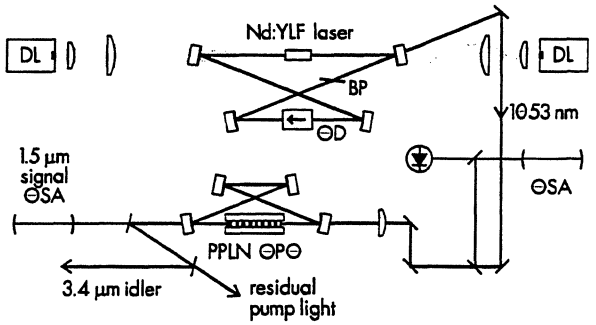


Fig. 2. The Nd:YLF laser and PPLN SROPO.

The SROPO is based around a 50 mm long crystal of periodically-poled lithium niobate (PPLN) in a ring cavity with a finesse F = 300 resonant at the signal wavelength λ_s. Stable operation is sustained for several minutes on a single signal-idler mode pair without any servoing of the signal cavity, as illustrated by fig. 3, which shows optical spectrum analyser fringes of the signal light leaking out of the resonant cavity. Feedback to the Nd:YLF laser is avoided by slightly tilting the PPLN crystal. The wavelength of the signal leakage is also monitored on a wavemeter. By temperature-tuning the poling period of the PPLN over five adjacent stripes with room-temperature poling periods Λ ranging from 28.5 to 29.3 μm, coarse tuning of the OPO action over a range λ_s = 1465 – 1540 nm can be obtained, corresponding to idler output over the range λ_i = 3750 – 3330 nm (fig. 4). Idler output at λ_i = 3392 nm corresponds to λ_s = 1527 nm, which can be obtained using two adjacent PPLN stripes, either Λ = 29.1 μm at 190°C or Λ = 29.3 μm at 151°C.

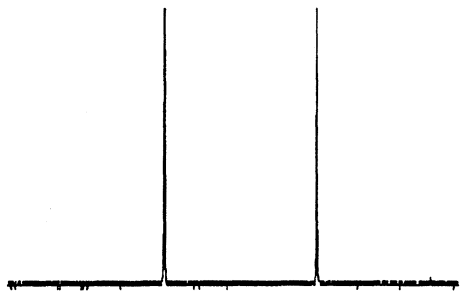


Fig. 3. Spectrum analyser fringes obtained with SROPO signal output, illustrating single mode-pair operation.

The threshold for OPO operation is at a pump power of 3.8 W. Idler output power of up to 200 mW is obtained; it is found that the power on a particular signal-idler mode-pair remains constant as the pump power is slowly changed. Myers and Bosenberg [8] give an expression for the theoretical pump threshold for cw SROPO operation with a single pass pump. Taking the value they give for the effective non-linear coefficient, d_{eff} = 14 pm/V, together with the parameters quoted above, and assuming optimum mode-matching of the pump and signal beams, the theoretical pump threshold for this system is 1.8 W, approximately a factor of 2 below that obtained. This difference may be accounted for by non-optimum mode matching and also by varying quasi-phase-matching conditions along the length of the crystal owing to non-uniform heating of the crystal.

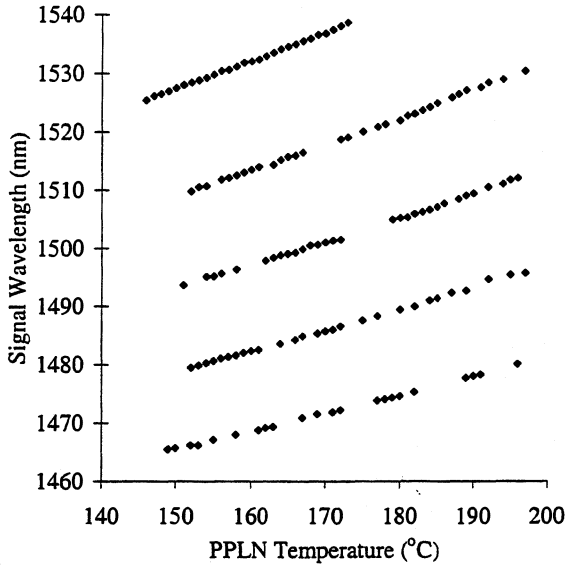


Fig. 4. SROPO signal wavelength as a function of PPLN temperature for five stripes with poling periods between 28.5 μm (shortest wavelengths) and 29.3 μm.

2.2 SHG of 3.39 μm

Light at 1.696 μm, the second harmonic (SH) of the HeNe/CH_4 standard, will be generated by resonantly-enhanced frequency doubling of the SROPO output. Frequency doubling of 3.39 μm radiation has been studied using both silver gallium selenide ($AgGaSe_2$) and PPLN, with a 1 mW infra-red HeNe laser as the source of fundamental radiation. $AgGaSe_2$ is phase-matched for SHG at this wavelength by angle-tuning at room temperature. With resonant enhancement of the fundamental radiation in a bow-tie configuration build-up cavity, an overall conversion efficiency of $\eta = 3 \times 10^{-3}\ W^{-1}$ has been obtained, as reported previously [9]. At the current SROPO output level, this will yield 30 μW of SH power. The FM sideband technique is used to lock this cavity to the laser frequency.

$AgGaSe_2$ suffers from absorption at 3.39 μm which limits the build up in the enhancement cavity; whereas PPLN has better transparency at this wavelength and similar single-pass conversion efficiency. SHG has been performed by quasi phase-matching in a 19 mm long sample of PPLN with room temperature poling periods of 32.6 to 33.5 μm. The enhancement cavity configuration was similar to that used for $AgGaSe_2$.

SHG was obtained by temperature-tuning at 158°C of a stripe with room temperature period 32.74 μm. However, the nanowatt-level 1.7 μm output is detected against a background of thermal radiation from the crystal oven, making an accurate measurement of the conversion efficiency difficult. At the higher fundamental power level available from the SROPO, the thermal background should not present such a problem. The temperature bandwidth for phase-matching is found to be comparable with the temperature stability of the crystal oven.

2.3. SHG of 1.7 μm

It is anticipated that frequency doubling of the SROPO idler output will generate sufficient power at 1.7 μm to directly drive a second SHG stage to 848 nm. However, to enable this frequency doubling to be set up independently of the SROPO and first SHG stage, DFB laser diodes at 1.696 μm have been obtained. These fibre-pigtailed diodes provide 3 mW with a free-running linewidth of 300 kHz and can be temperature-tuned at 0.08 nm/°C.

Either potassium niobate ($KNbO_3$) or PPLN can be used for frequency doubling to 848 nm. Type I phase matching is calculated to be obtained in $KNbO_3$ at crystal angles of $\Theta = 49.5°$, $\Phi = 90°$ [10]. The expected conversion efficiency is a few 10^{-2} W^{-1} for a build-up cavity of enhancement factor 15 to 20. Around 10 nW of output from this stage is expected to be adequate to phase-lock an extended-cavity diode laser (ECDL) as a local oscillator at 848 nm.

2.4 Difference Frequency Mixing to Strontium

From 848 nm, the Sr^+ ion quadrupole transition at 674 nm can be reached by difference frequency mixing of sources at 674 nm and 841 nm, to generate a 3.39 μm signal which can be referenced back to the HeNe/CH_4 standard. Candidate non-linear materials for this mixing include PPLN and $KNbO_3$. The thermal background experienced when using temperature-tuned PPLN for the 3.39 μm SHG would also be a problem in detecting the DFM signal; in addition, the PPLN poling period required is around 16.8 μm which is relatively short. Unfortunately, 674 nm lies outside the transparency range of silver gallium sulphide which has been used to generate radiation in the 3.4 μm region by DFM [11,12]. However, type I phase-matching can be obtained in $KNbO_3$ by angle-tuning at room temperature; the phase matching angle is within 1° of that required for SHG of 1.7 μm [10]. Adequate output is anticipated in a single-pass interaction mixing 40 mW at 841 nm from an ECDL and 500 mW at 674 nm from an injection-seeded tapered amplifier diode (MOPA configuration) currently under construction. The MOPA system will also supply light via a single-mode optical fibre link to the Sr^+ laboratory for intercomparison with the probe laser for the Sr^+ ion traps.

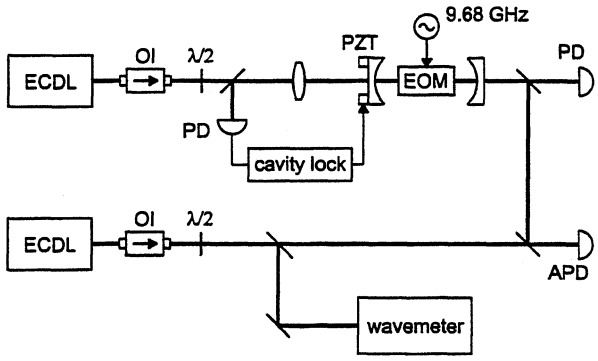


Fig. 5. Schematic of the OFC set-up. OI: optical isolator; λ/2: half-wave-plate; PZT: piezo element; PD: photodiode; APD: avalanche photo-diode.

2.5 Optical Frequency Comb Generator

The residual 2.9 THz gap between the sources at 841 nm and 848 nm will be bridged using the optical frequency comb generator. An OFC is currently being constructed in our laboratory. Initially it is a simple single cavity design driven by an ECDL operating at 845 nm (fig. 5). The OFC uses a commercial EOM consisting of an anti-reflection coated MgO-doped $LiNbO_3$ crystal embedded in a resonant microwave cavity. The resonant frequency can be tuned from 9.2 GHz to 9.7 GHz. The optimum modulation index is obtained near 9.7 GHz. A modulation index of 0.5 at 845 nm has been measured using an rf power of 1 W. Figure 6 shows the transmitted OFC cavity fringes with and without rf power applied to the EOM. An error signal for locking the OFC cavity to the laser frequency with the EOM drive on is generated by dithering the cavity length at 1.2 kHz. The amplitude-modulated reflected light is demodulated at the dither frequency and integrated to provide feedback to the cavity piezo.

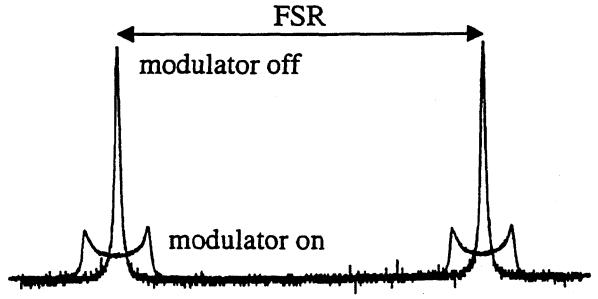


Fig. 6. Transmitted fringes from the single-cavity comb generator with the EOM drive off ('modulator off') and driven by 1 W of rf power at 9.68 GHz ('modulator on').

Fig. 7 shows beat signals between the OFC output and the output from a second ECDL. A beat signal is only seen between the second ECDL and the OFC mode m nearest in frequency. In this preliminary data, beatnotes can only be detected out to m = 32, corresponding to a comb with a span of 0.6 THz. The span is currently limited by the relatively weak cavity lock. However, to achieve the required 2.9 THz span it will also be necessary to increase the finesse of the optical cavity and to increase the rf power used to drive the modulator. It is also intended to use a coupling cavity to increase the efficiency with which the laser power is coupled into the comb [6], and a filter cavity to give efficient output coupling of the desired sideband from the comb [13].

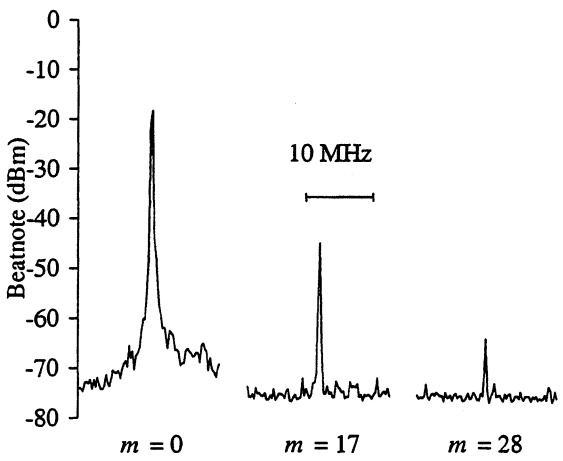


Fig. 7. Beat signals between modes m = 0, 17, and 28 of the comb and a second ECDL.

3. LINKAGE TO YTTERBIUM

In addition to the development of the Sr^+ optical frequency standard, the Yb^+ ion has also been extensively studied in our laboratory [14]. Work at NPL has concentrated on three transitions which may be suitable for use as optical frequency standards: the $^2S_{1/2}$ – $^2D_{5/2}$ quadrupole transition at 729 THz (411 nm) [15], the ultra-narrow $^2S_{1/2}$ – $^2F_{7/2}$ octupole transition at 642 THz (467 nm) [16], and the $^2D_{5/2}$ – $^2F_{7/2}$ transition in the infra-red at 87.4 THz (3.43 μm) [12].

The 729 THz quadrupole transition is probed by a frequency-doubled diode laser at 365 THz (822 nm). This frequency is 11.2 THz away from the fourth harmonic of the HeNe/CH_4 frequency. This offset is too large to be measured directly using our design of OFC generator, which is expected to be limited by dispersion effects at around 5 THz span. However, it can be reduced by a factor of four using two Hänsch-Telle optical frequency divider (OFD) stages [17]. The resulting 2.8 THz interval can be measured using the OFC. Our longer-term intention is to extend the OFD technique to produce a chain having sufficient flexibility to reach the other ytterbium transitions and, ultimately, to provide a direct linkage to the caesium primary standard as has been proposed by Hänsch and co-workers [9,18].

REFERENCES

[1] T.J. Quinn, *Mise en Pratique* of the Definition of the Metre (1997), Metrologia (to be published).

[2] L. Marmet et al., IEEE Trans. Instrum. Meas. 46, 169-173, 1997.

[3] G.P. Barwood et al., IEEE Trans. Instrum. Meas. 46, 133-136, 1997, and paper A4-3 of these proceedings.

[4] M.A. Gubin et al., IEEE J. Quant. Electron. 31, 2177-2182, 1995, and paper A5-3 of these proceedings.

[5] O. Acef et al., paper P2-68 of these proceedings.

[6] A.S. Bell et al., Opt. Lett. 20, 1435-1437, 1995.

[7] P.J. Hardman et al., Opt. Comm. 156, 49-52, 1998.

[8] L.E. Myers and W.R. Bosenberg, IEE J. Quant. Electron. 33, 1663-1672, 1997.

[9] S.N. Lea et al., IEEE Trans. Instrum. Meas. (to be published).

[10] I. Biaggio et al., J. Opt. Soc. Am. B 9, 507-517, 1992.

[11] U. Simon et al., Opt. Lett. 18, 1931-1933, 1993.

[12] P. Taylor et al., Opt. Lett. 23, 298-300, 1998.

[13] J. Ye et al., Opt. Lett. 22, 301-303, 1997.

[14] M. Roberts et al., paper A4-4 of these proceedings.

[15] P. Taylor et al., Phys. Rev. A 56, 2699-2704, 1997.

[16] M. Roberts et al., Phys. Rev. Lett. 78, 1876-1879, 1997.

[17] H.R. Telle et al., Opt. Lett. 15, 532-534, 1990.

[18] T. Udem et al., in "Proceedings of the Workshop: Frequency standards based on laser-manipulated atoms and ions", PTB (Braunschweig), 77-81, 1996.

PHASE COHERENCE IN OPTICAL FREQUENCY MEASUREMENT AND SYNTHESIS

G. KRAMER, B. LIPPHARDT, U. HUEBNER

Physikalisch-Technische Bundesanstalt, Bundesallee 100, D-38116 Braunschweig, Germany

ABSTRACT

We report a new electronic system for processing and recording the harmonic beat signals in our infrared frequency chain as well as its application for stabilizing a laser spectrometer.

1. INTRODUCTION

During the past years at PTB we have developed a frequency multiplication chain and applied it for phase coherent comparison of various optical frequency standards in the infrared and visible part of the spectrum with a hydrogen maser.

The natural objects for precision coherent measurement have been the atomic, molecular standards developed and operated at PTB. So the first configuration of the chain was for measuring the frequency of the 3.39 µm HeNe-laser stabilized to the resonance in methane. In 1988 already we were able to measure the frequency of the non-perturbed central component of the F-line by linear optical Ramsey resonance [1]. At FCS 92 we reported it with 200 Hz uncertainty.

This branch, or better: trunk of the chain is now used for regular and repeated calibration of the stable transportable standards, built by our Russian colleagues [2]. These have been taken to Garching, Paris, Boulder, and Teddington, where they have been applied for absolute frequency measurement of other optical standards. At PTB in Braunschweig we were able to extend the chain for coherent measurement of the Ca standard [3] in the visible and to the OsO_4 fundamental line at 10.5 µm.

2. MULTIPLICATION CHAIN

What must now be called a conventional chain consists of a collection of oscillators connected and referenced to each other by harmonic mixers. Fig.1 is a schematic of our trunk chain, which starts from a low phase noise Xtal oscillator at 100 MHz and reaches up to 88 THz.

2.1 Electronic oscillators

386 GHz is the highest electronically generated frequency in our chain and chosen to be a subharmonic of the 4.2 THz methanol laser. In stabilizing this frequency, quite some care had to be taken to avoid any excessive phase noise, which would have caused carrier collapse at multiplication to 4 THz. The diagram (Fig.2) illustrates, that at 4 THz with the special Xtal oscillator there is enough 'coherence margin' in a wide range of Fourier frequencies.

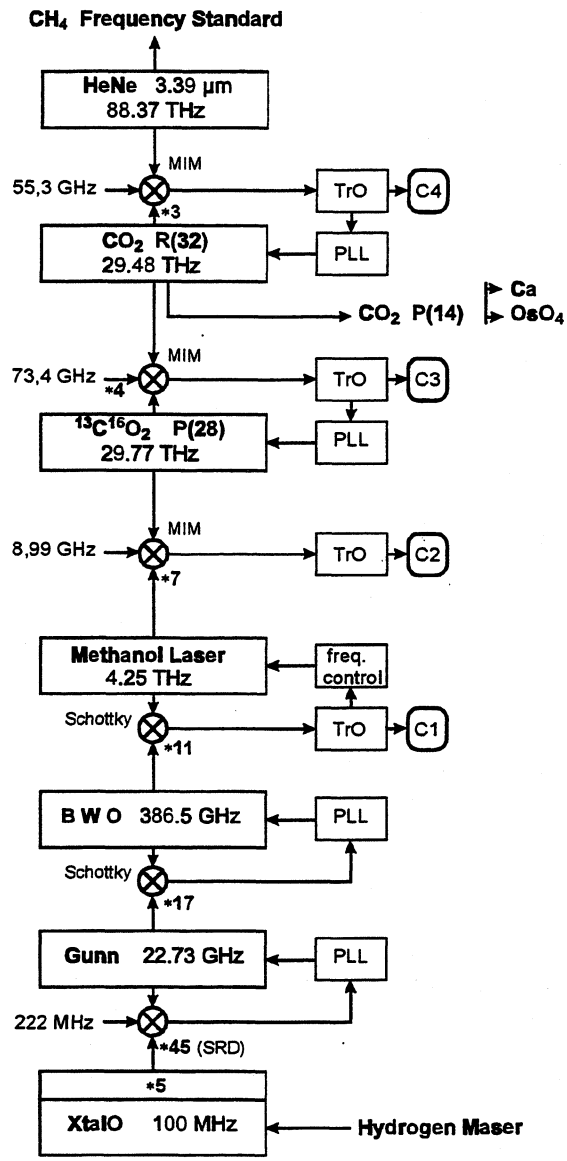


Fig. 1 Schematic of the main (trunk) chain to 88 THz

A microwave source is also needed at practically every multiplication stage in order to bring the harmonic beat frequency down to below a few 100 MHz. All these Gunn and klystron oscillators are phase-locked to the H-maser reference.

2.2 Gas Lasers

CO_2 lasers have many lines and ample power. So, for overcoming problems with the output power of the methanol laser, we have meanwhile introduced another

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

CO_2 laser. The P(28) line $^{13}CO_2$ is closer to the 7th harmonic of the methanol than the R(32) line, which reduces the mixing order necessary at this step from 12th to 9th.

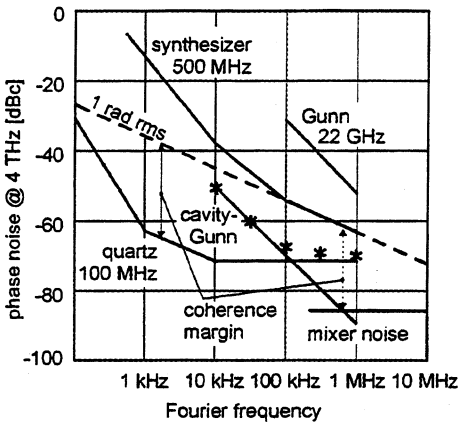


Fig. 2 Phase noise after multiplication to 4 THz

And we have connected - in what we call the 'side-step' - the P(14), which we use for our OsO_4 optical frequency standard work and which, as a lucky coincidence, is also the starting point for reaching the Ca-standard developed by F. Riehle et al.[3].

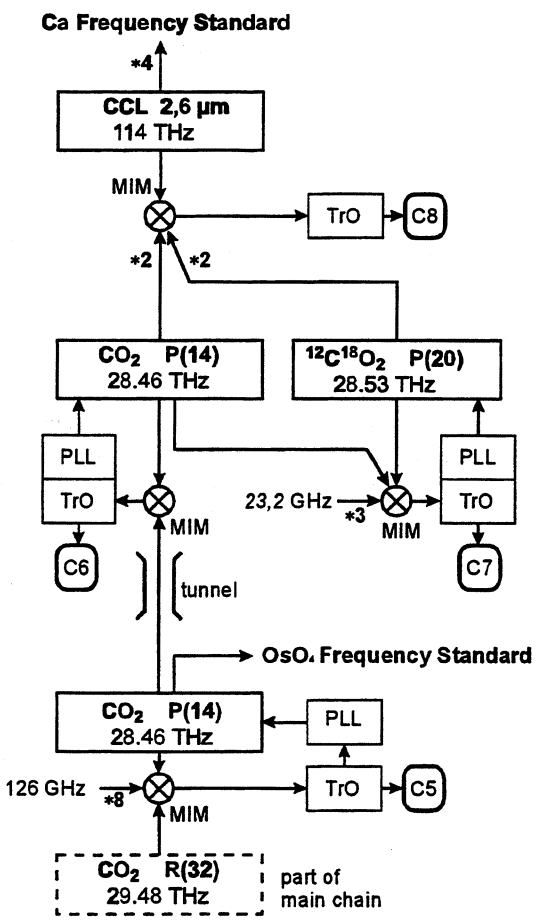


Fig.3 'side step' connecting to OsO_4 and Ca

A P(14) beam is passed through a 60 m long plastic tube ('tunnel') to the neighboring building, where there are two more CO_2 lasers and a color center laser (FCL, at 1/4 the Ca frequency), all phase-locked to the chain.

2.3 Harmonic mixers

MIM diodes (W-Ni and W-Co) can be used to quite high frequencies. At 88 THz (HeNe) and at 114 THz (FCL) they still generate beat signals of sufficient S/N ratio in 5th mixing order. In the 30 THz range, with optimum power available from all mixing partners, even 15th order mixing would be feasible (S/N ≅15 dB/100kHz).

2.4 Mixer noise

Wherever the beat signals are to be processed by digital counting, the input signal to such counter must have a very high S/N ratio, which can be obtained by reduction of the bandwidth only. While often a low pass or a band pass filter may already reduce the probability of false counts to a negligible level, the ideal filter for that is a tracking oscillator (Fig.4). If correctly designed, it works like a miracle.

2.5 Tracking oscillators

The electronic tracking oscillator (Fig.4) consists of a voltage controlled oscillator (VCO), a double balanced mixer (DBM), and a loop filter, that closes the phase lock loop to lock the VCO frequency to that of the noisy input. In many cases the phase noise characteristics of the VCO are not critical and, instead of the normal LC type oscillator with varactor tuning, a multivibrator type may be chosen for its frequency agility and wide tuning range.

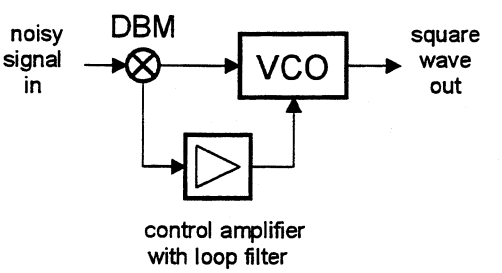


Fig.4 Tracking oscillator schematic

It is essential that the frequency control works up to quite high Fourier (modulation) frequencies. If the loop filter allows for a wide tracking range by providing extra gain at low frequency and dc, the modulation bandwidth must be much higher than the intended locking bandwidth. With low frequency gain enhancement by a factor of 100 for example and the unity gain frequency of the PLL set at 100 kHz, the controller must work up to 10 MHz without excessive phase delay, which would cause pushing instead of pulling-in.

3. MEASUREMENT

Originally we had thought that coherent measurement meant phase-locking all oscillators in the chain. Later we realized that particularly with the FIR laser, this was not only difficult (heavy mirrors, technical vibrations) but also unnecessary.

3.1 Locking from the base and from the top

Using tracking oscillators, prescalers and digital phase locks we are able to reliably lock the CO_2 lasers to the methane. The BWO is tightly locked to the 100 MHz reference.

3.2 Intermediate oscillator free-running

The diagram shows how the two remaining beats were processed to give a direct frequency readout of the chain. With such beat signal processing, transfer lasers need only slow frequency control or may even be left free-running. The essential thing for conserving coherence is that the frequency fluctuations of all beat signals are being tracked without cycle slips.

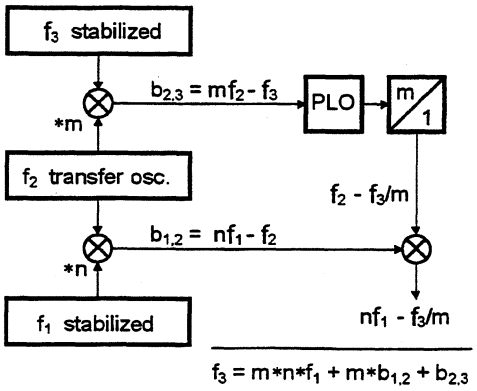


Fig.5 Transfer oscillator free-running

3.3 Synchronously strobed counter readout

The recording system, that we have been using in recent years, is a generalization of what we had done to circumvent the FIR laser locking problem: It counts the periods of all the laser beats in separate accumulating counters, the contents of which are simultaneously transferred ('strobed') into latches. As the counters are never reset, this is equivalent to recording a multidimensional phase. As latches we use shift registers, all connected in series, by this way concatenating the data of all the counters to form one long telegram, which is quasi 'broadcast' to the COM ports of several PCs.

While this 'strobed' read-out is simple conceptually, it requires an electronic trick: Only in a completely synchronous counter would the outputs of all counter stages represent the total pulse count. A normal 'ripple' type counter has to be stopped, until the outputs have stabilized, before it can be read out. The trick is that we interrupt the counting for a fixed number of input pulses, actually 16, and then each time advance the counter by that number feeding one pulse into the appropriate stage of the counter. The strobe command, which is common for all channels, has to be carefully synchronized to the individual input signal of each channel in order not to cut a pulse in two pieces when interrupting the counter (the fragments could be counted as two or not at all). Synchronization consist in delaying execution of the interrupt by the right fraction of the input signal period.

3.4 Resolution

So far the lowest frequency beat signal, that we record by period counters only, is with the alcohol laser at 4 THz. At this level the ±1 count ambiguity of the counters corresponds to a resolution of 240 fs (=1/4.2 THz), which is better than the 'white phase noise' level of the hydrogen maser. When measuring the OsO_4 stabilized CO_2 directly against the CH_4 stabilized He-Ne laser however, the resolution constitutes the short term noise floor. (Fig.6)

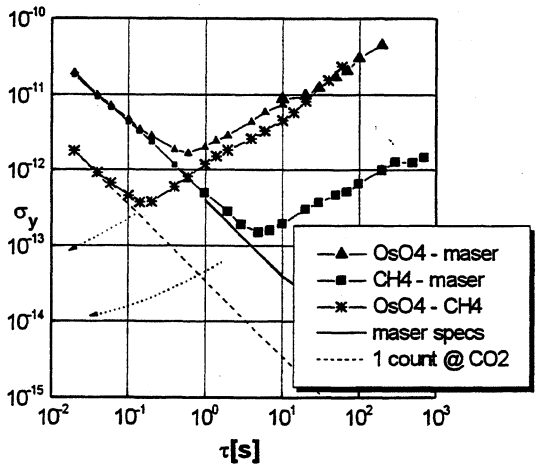


Fig.6 Recording a frequency ratio

So in the course of renovating and modernizing the data recording, we are also enhancing the resolution of the system by applying an old analog interpolation scheme [5], which, in addition to the accumulated full periods, determines the fractions of a period, resolving 0.001 period. As Fig.7 shows, the delay necessary to synchronize the interrupt with the input signal is expanded by a factor of 1000 by means of an analog pulse length multiplier.

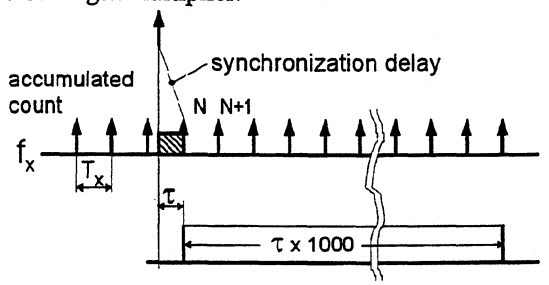


Fig.7 Determining the period fractions

The length of the expanded pulse is then measured in terms of the unknown period. Built with standard HC type logic circuits this digital phase recorder works from 5 to 40 MHz. At 10 MHz the precision really is 100 ps. So far we have converted 2 channels only, so we could not yet run triangle type tests. An FPGA version containing 4 channels in one package is ready now.

3.5 Counter data utilization

The normal use that is made of the counter data in the PC, will be to calculate the frequency of one of the optical inputs to the chain in terms of the standard frequency (Fig.8). Accumulation of rounding errors is completely avoided by using integer arithmetic.

We can also simultaneously solve for OsO_4 and methane f. ex. and calculate the frequency ratio of the two. As the frequencies deviate only very little from their nominal values, this division does not have to be carried out each time, but can be replaced by a subtraction using $1/(1+\delta) \approx 1-\delta$.

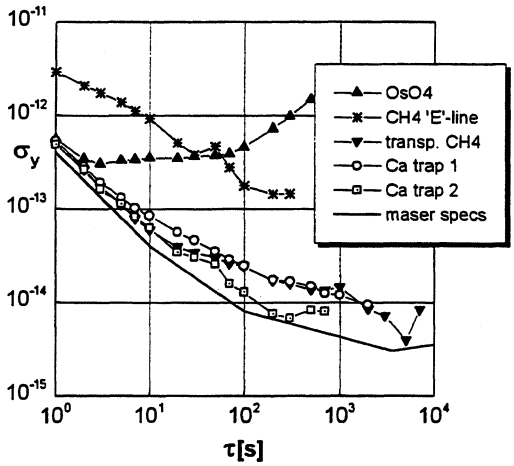


Fig. 8 Allan std. deviation of stable optical standards.

4. FREQUENCY SYNTHESIS

In another application the computer generates a correction voltage to servo phase-lock either a crystal oscillator (optical clock) or a laser.

4.1 Laser spectrometer with absolute frequency scale

Whereas we have little use for an optical clock, we have already (as a test) phase-locked a methane stabilized laser to the hydrogen maser. This technique should provide the ideal reference oscillator for measuring precisely the line shapes of the narrow, though noisy, optical Ramsey resonances.

For the OsO_4 line we have already made the scheme work: The P(14) CO_2 laser was pre-stabilized to an OsO_4 saturated absorption line with the help of an acousto-optical modulator/frequency shifter (AOM), and then computer phase-locked to the hydrogen maser.(Analog pre-stabilization or passive stability is necessary, because the computer servo works at 50 measurements/control actions per second only). Another AOM, driven by a programmable generator, completed the laser spectrometer. It allowed D.Tyurikov (as a guest from FIAN) to precisely record the narrow, noisy, and slightly asymmetric saturated absorption line at varied alignment and power settings. From these recorded data he determined the true center frequency of the resonance:

(28464788929370 ±80) Hz

Another effort should probably be undertaken to improve on the accuracy of the CH_4 F-line.

4.2 Optical clock

As said before, there appears to be little demand for the opposite lock direction, i. e. for phase locking a crystal oscillator to an optical standard. It must be mentioned however, that in a certain Fourier frequency range optical frequency standards probably exhibit the highest stability of all. If this short term stability were transferred to the 9.2 GHz range for example, it might make the stable interrogation oscillator for the cesium fountain. I am not suggesting that existing molecular standards would suffice for that. The extrapolated Allan variance curves were drawn only to illustrate that these lasers are more stable than the hydrogen maser. At times τ shorter than 10 s, at least.

5. CONCLUSION

Our conventional frequency multiplication chain has proven its precision and functionality. With the new data processing it works reliably for many hours, even unattended. Its frequency synthesis capabilities open new applications in precision laser spectrometry. The next candidate for an extension of the chain is now the 871 nm radiation used with the single trapped Ytterbium ion standard [4] of Chr. Tamm in our department. H. Telle has suggested and plans to apply a combination of interval division and a comb generator for a coherent connection with the Ca frequency.

[1] G. Kramer et al., "Coherent Frequency Measurement of the Hfs-Resolved Methane Line" in Proc. of the conference on Frequency Standards & Metrology, Springer(1989)
[2] Gubin et al. "Transporable He-Ne/CH_4 .." J.QE, vol. 31,pp. 2177-2182; and this conference, session A5
[3] F. Riehle et al,"The Optical Ca Frequency Standard", this conference, session A5
[4] Chr. Tamm, V. Bühner, Abstract QFD5, EQEC'98, Glasgow
[5] G. Kramer, "Ein rechnender digitaler Phasenmesser", Frequenz, vol. 27(1973)

1999 Joint Meeting EFTF - IEEE IFCS

MEASUREMENT METHODS OF FREQUENCY NOISE IN OPTICAL SOURCES BASED ON FABRY-PEROT DISCRIMINATORS

E. Bava, G. Galzerano, and C. Svelto

Dipartimento di Elettronica e Informazione del Politecnico di Milano,
Piazza Leonardo da Vinci 32, 20133 Milano, Italy
E-mail: Elio.Bava@polimi.it Phone: +39 02 2399 3609 Fax: +39 02 2399 3413

ABSTRACT

Optical frequency discriminators based on Fabry-Perot cavities are analyzed with particular attention to their amplitude-noise response and to the sensitivity limits imposed by detection shot-noise level. Comparison between different schemes are carried out.

1. INTRODUCTION

The characterization of frequency noise of cw laser oscillators is of great importance both in basic metrology and in many scientific applications as well. For this purpose, time- and frequency-domain measurement methods have been developed and used; these latter, based on a frequency discriminator, are more general and do not require the availability of two equal sources (or of a reference source). However the frequency response to amplitude noise of a discriminator has to be evaluated and the independently measured amplitude noise level, suitably shaped, must be then subtracted from the output of the particular discriminator system.

The use of the Michelson interferometer with balaced detection as a frequency discriminator has been thoroughly studied in [1], also considering quantum noise sources, conditions for minimum amplitude noise detection in non ideal operation, and distortion. This last parameter sets a limit to the maximum frequency deviation. This setup turns out to be very agile allowing frequency noise measurements in an extremely large bandwidth, a feature of great interest in the characterization of semiconductor lasers. This occurs however at the expenses of the sensitivity as it is usual also for other discriminators.

Diode-laser pumped solid-state lasers, like Nd:YAG, Er-Yb:glass, Tm-Ho:YAG etc., show a low-level frequency noise in a narrow bandwidth around the carrier and, to achieve an adequate discrimination sensitivity, large unbalances between the two arms of a Michelson interferometer are necessaries, which could make unpratical the use of this scheme.

On the other hand Fabry-Perot resonators with an adequate linewidth compared to the oscillator noise bandwidth are suitable devices to carry out such measurements. Two modes of operation of the resonator can be devised and are used to build a frequency discriminator by exploiting either the slope of a resonance fringe side or the response to the frequency-modulated light of the laser centered on the Fabry-Perot resonance.

The former scheme is rather simple and widespread. The use of a reference optical path yields partial amplitude noise rejection, which can be increased by exploiting transmitted and reflected beams improving frequency sensitivity as well. Fringe-side slope transients have been analyzed in [2] and [3], whereas the frequency-noise responses have been discussed in [4]. However the unwanted detection of amplitude noise seems not to have been taken fully into account.

The discrimination properties of the FM-based scheme (Pound-Drever) [5] have been analyzed in [6], [7] and [8] considering the amplitude noise immunity as well. Applications of this technique are usually concerned with frequency stabilization of laser oscillators, whereas an example of frequency noise measurement is found in [9].

In this work evaluations of amplitude-noise frequency responses have been performed for both types of discriminators, and a comparison of sensitivity limits imposed by the detector shot-noise has been made.

2. THE FABRY-PEROT FRINGE SIDE DISCRIMINATOR

The set up for measuring laser frequency fluctuations by using the discriminating effect of a Fabry-Perot is shown in Fig. 1 in two possible configurations, the former (a) is quite simpler, uses the fringe side slope in transmission (but could also exploit reflection), however the latter (b), by using both reflection and transmission, has a higher sensitivity to frequency variations and a better rejection of amplitude noise. The resonator is tuned so that the mean laser frequency ω_0 is located on the side of a fringe; the source frequency fluctuations are converted into amplitude fluctuations by the opposite slopes of the resonator transmission or reflection coefficients.

Let us assume the laser beam incident onto the resonator to be expressed as

$$E_i(t) = E_0[1 + a(t)]\exp[\varphi(t)]\exp(j\omega_0 t) \qquad (1)$$

where $a(t)$ and $\varphi(t)$ represent the oscillator amplitude and phase noise, respectively. Let us consider a noise

0-7803-5400-1/99/ $ 10.00

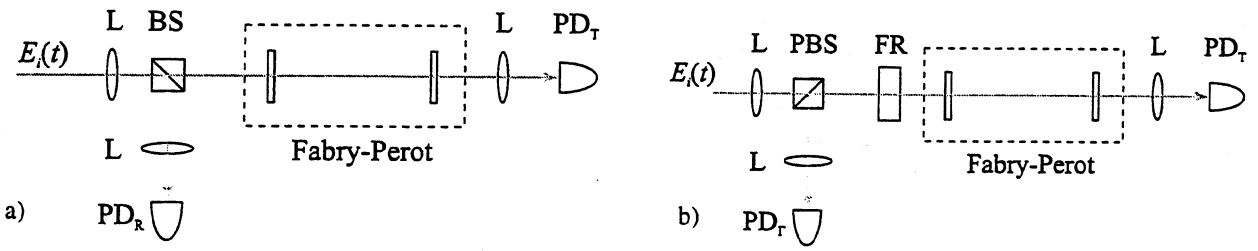


Figure 1. Fringe-side discriminator set-up with a Fabry-Perot used in trasmission with a reference arm **a)**., and in reflection and trasmission **b)** both with a differential output. L lens, (P)BS (polarizing) beam splitter. FR Faraday rotator, PD photodetector.

modulation at an angular frequency $\omega_n = 2\pi f_n$, with amplitude $a(t) = a_n \sin(\omega_n t+\alpha)$ and phase $\varphi(t) = \varphi_n \sin\omega_n t$. The spatial evolution of the beam is not considered here; the effects of non ideal mode matching with the resonator and with the detector area are supposed to be overcome by suitable lens systems and, on the other hand, these problems are common with the other schemes examined. According to the transmission or the the reflection path considered, the electric field in front of the corresponding detector is

$$E_o(t) \cong E_0 \exp(j\omega_0 t)\left\{\begin{array}{l} J_0 H(j\omega_0)+\left[J_0\frac{a}{2}(\sin\alpha - j\cos\alpha)+J_1\right]H(j(\omega_0+\omega_n))\exp(j\omega_n t) \\ +\left[J_0\frac{a}{2}(\sin\alpha + j\cos\alpha)-J_1\right]H(j(\omega_0-\omega_n))\exp(-j\omega_n t)\end{array}\right\} \quad (2)$$

where low amplitude- and phase-noise levels have been assumed, J_0 and J_1 are Bessel's functions of the first kind with argument φ_n, and $H(j\omega)$ stands either for the field reflection $\Gamma(j\omega)=\Gamma'(j\omega)+j\Gamma''(j\omega)$ or transmission $T(j\omega)= T'(j\omega)+jT''(j\omega)$. Suitable expressions for these two coefficients are

$$\Gamma(j\omega)=\frac{\sqrt{R}[1-(1-A)\exp(-j2\pi\nu\tau)]}{1-R\exp(-j2\pi\nu\tau)} \quad (3)$$

$$T(j\omega)=-\frac{(1-R-A)\exp(-j\pi\nu\tau)}{1-R\exp(-j2\pi\nu\tau)} \quad (4)$$

where τ is the resonator round trip time, R and A are power reflection and losses of each mirror, respectively. Moreover, if ω_r is a resonance frequency, for $(\omega-\omega_r)\tau << 1$ the following useful relationships are valid

$$1\pm T'(j\omega) \cong \Gamma'(j\omega) \quad (5)$$
$$T''(j\omega) \cong \pm\Gamma''(j\omega) \quad (6)$$

where the signs + or - correspond to an even or odd number of wavelengths inside the resonator, respectively.

A radiation field with an electric component E_o incident into a photodetector with quantum efficiency η yields a photocurrent

$$i(t) = \eta S\left(\frac{c\varepsilon_0 e}{2h\nu_0}\right)E_o E_o^* = \eta S B E_o E_o^* \quad (7)$$

where S represents an equivalent detector area, which in principle is different for each detector.

Following only one of the two optical beams, e.g. the transmitted one, the field in front of the detector is the sum of three components at different frequencies. After substitution of (2) in (7) with $H(j\omega) = T(j\omega)$, it is possible to extract the output current separated in a dc component and in components at ω_n dependent substantially on φ_n and a_n respectively. By labelling with the subscripts 0, +, and - the field transmission coefficients at $\omega_0, \omega_0+\omega_n, \omega_0-\omega_n$, respectively, the three current components are

$$i_{Tdc} = \eta_T S B E_0^2 J_0^2 |T_0^2| \quad (8)$$

$$i_{T\varphi} = \eta_T S B E_0^2 2J_0 J_1 \{[T_0'(T_+'-T_-')+T_0''(T_+''-T_-'')]\cos\omega_n t + [T_0''(T_+'+T_-')-T_0'(T_+''+T_-'')]\sin\omega_n t\} \quad (9)$$

$$i_{Ta} = \eta_T S B E_0^2 J_0^2 a_n \{[T_0'(T_+'+T_-')+T_0''(T_+''+T_-'')]\sin\alpha\cos\omega_n t + [T_0'(T_+''-T_-'')-T_0''(T_+'-T_-')]\cos\alpha\cos\omega_n t + [T_0'(T_+''+T_-'')+T_0''(T_+'+T_-')]\cos\alpha\sin\omega_n t + [-T_0'(T_+''-T_-''')+T_0''(T_+'-T_-')]\sin\alpha\sin\omega_n t\} \quad (10)$$

The field intensity in the reference arm of Fig. 1a is a fraction k of that incident into the Fabry-Perot. The k value can be chosen according to other requirements of optimization. The reference-arm detected current does not depend on phase modulation

$$i_R = i_{Rdc} + i_{ra} = kS\eta_R B E_0^2\{1+2a_n[\sin\alpha\cos\omega_n t+\cos\alpha\sin\omega_n t]\} \quad (11)$$

As regards the reflected beam of Fig. 1b, by

including the losses of the optical paths (Faraday rotator, polarising beam splitter and focussing lenses) into the detector efficiency, the general expression of the output current is obtained by substitution of the real and imaginary components of T at the proper frequencies with the corresponding ones of Γ.

By using the relationships (5) and (6) it is easy to compare the output voltages of the two differential-scheme detections

$$\Delta v_a = k_T i_T - k_R i_R \qquad (12a)$$

$$\Delta v_b = k_T i_T - k_\Gamma i_\Gamma \qquad (12b)$$

where $k_{T,R,\Gamma}$ are the detector transimpedence gains of the transmission, reference, and reflection arms, respectively. From the expressions for $\Delta v_{a,\varphi}$ obtained from (12a), after computation of the amplitude divided by f_n, the discrimination characteristics for reflection and transmission, already reported in [4], are obtained. They show a 3 dB bandwidth equal to the resonance HWHM and the roll-off away from the cutoff frequency is -20 dB/decade or -40 dB/decade for the reflection or transmission case, respectively. On the other hand, from the expression of $\Delta v_{T-\Gamma}$ a discrimination sensitivity is recognized twice the previous one, as expected, and the roll-off far from the cutoff frequency is o -20 dB/decade. Moreover it appears that, with the assumption of low noise levels, distortion is taken into account in the non-linear behavior of the Bessel function product, whereas the bandwidth of the system depends on the dispersion profile of T and Γ, a situation similar the case of the discriminator based on the Michelson interferometer discussed in [1].

As regards the amplitude noise, if the condition $k_T i_{Tdc} - k_R i_{Rdc} = 0$ is imposed in (12a), the frequency response goes to zero as $f_n \to 0$, as expected. This behavior is linear as can be shown by using expressions (5) and (6). The output voltage is readily obtained from (8), (9), (10) and (11) when substituted in (12a). A few cases have been evaluated and reported in Fig. 2, where frequency responses to amplitude noise are shown, correspondent to schemes using either reflection or transmission (75% and 50% of this last parameter). The values of these curves, multiplied by the rms amplitude-modulation spectrum times the detected voltage due to the carrier, give the rms spectral components at the differential output. These computations have been performed by keeping constant the dc detected output voltages; different power levels are then required at the beam-splitter input. The well known difference in the slopes -20 dB/decade and -40 dB/decade away from the cutoff frequency in discriminators operating in reflection and in transmission, respectively, appears in the output amplitude noise, but the slope changes in this filtering process are reversed.

When the discriminator of Fig. 1b is considered, the output voltage is again separated into three components $\Delta V_{dc,T-\Gamma}$, $\Delta V_{\varphi,T-\Gamma}$, $\Delta V_{a,T-\Gamma}$. The

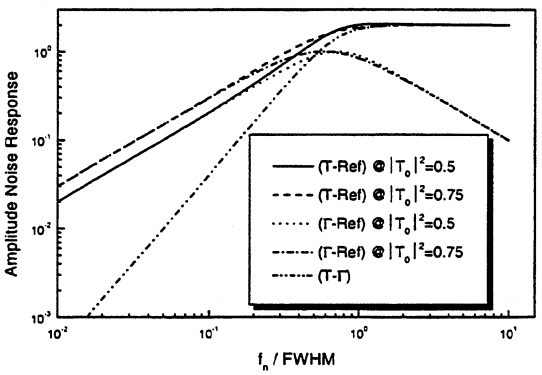


Figure 2. Normalized amplitude-noise response for differential schemes operating in reflection or in transmission.

last two expressions, even if straighforward, are rather complicated and are not reported directly. The first one

$$\Delta V_{dc,T-\Gamma} = SBE_0^2 J_0^2 \left[\eta_T \kappa_T \left|T_0^2\right| - \eta_\Gamma \kappa_\Gamma \left|\Gamma_0^2\right|\right] \qquad (13)$$

shows for $\Delta V_{dc,T-\Gamma} = 0$ the condition for a minimum sensitivity to amplitude noise. Therefore, the choice ω_0-ω_c = HWHM (50% transmission, $\left|T_0^2\right| = \left|\Gamma_0^2\right| = 1/2$ for a lossless Fabry-Perot) is mandatory if the product efficiency × amplification is kept equal in each photoreceiver. This situation is depicted in Fig. 2 as well. From analytical evaluations, confirmed by numerical computations, this curve goes to zero as f_n^2 for $f_n \to 0$, moreover it shows a general advantage for practical values of f_n with respect to the other schemes.

3. THE POUND-DREVER DISCRIMINATOR

A rather general optical and electronic set-up, which includes a few possible realizations of a frequency discriminator following the Pound-Drever-Hall method is shown in Fig. 3.

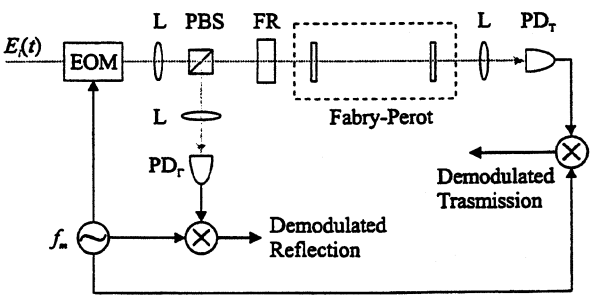


Figure 3. Pound-Drever-Hall discriminator with demodulation of the reflected and trasmitted field. EOM electro-optic modulator.

The noisy laser source is phase modulated at ω_m and its carrier frequency is kept in coincidence with a Fabry-Perot resonance. Either the reflected or the transmitted signal may be synchronously demodulated, the former topology however is much more efficient than the latter and a differential scheme does not seem convenient because of noise effects. The electric field

incident onto the cavity is obtained by adding phase modulation to the expression (1), that is

$$E_i(t) = E_0[1 + a(t)]\exp[j(\omega_0 t + \beta \sin\omega_m t + \varphi(t))] \quad (14)$$

The noise current at ω_m, under the hypothesis that the Fabry-Perot linewidth $\Delta\omega_c \langle\langle \omega_m$, can be obtained following the procedure outlined in [6],[7] and [8] with the condition $\Delta\omega = \omega_c - \omega_0 = 0$. The discrimination curves already reported in [3],[6] and [8] are shaped as a single-pole low-pass filter with a cutoff frequency at the resonance HWHM. According to this model the amplitude-noise rejection is complete. But if a frequency offset exists between ω_c and ω_0, the sensitivity to the amplitude noise is strongly increased (see Fig. 4). The response is constant inside the discriminator bandwidth and appears linearly related to the frequency offset.

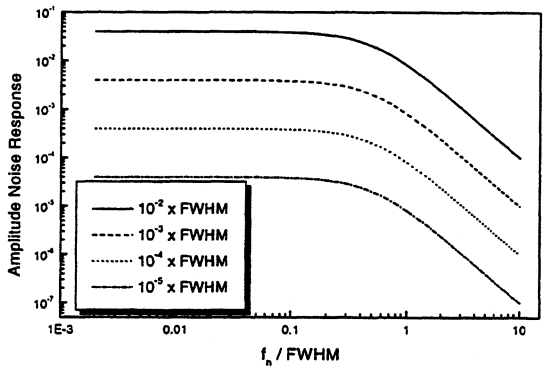


Figure 4. Normalized amplitude-noise response as a function of frequency offset (f_c- f_0) for a Pound-Drever-Hall scheme.

4. DISCRIMINATOR NOISE LIMITS

Limits to the discrimination sensitivity are set by the detector shot-noise levels. In the case of Fig. 1a, being equal the photodetector dc currents, we obtain, for the discriminator in transmission, the following expression of the power spectral density of the output voltage produced by shot-noise

$$N_{\Delta v,T}(f) = 2eBSE_0^2\eta_T\kappa_T^2\left|T_0^2\right|\left(1+\frac{\kappa_R}{\kappa_T}\right) \quad (15)$$

where the sum in the round brackets is $\equiv 2$. The corresponding expression for the reflection discriminator is straightforward. Note that if the reference arm is not used the shot-noise level is reduced of 1/2, but there is no rejection of amplitude noise. The expression (15) holds in the case of Fig. 1b, as well, provided the factor 2 is used in the round brackets in (15), but the discriminator sensitivity is twice and the amplitude noise rejection is higher (see Fig. 2).

The analysis of the Pound-Drever-Hall discriminator leads to

$$N_{\Delta v.PDH}(f) = 2eBSE_0^2\eta_\Gamma\kappa_\Gamma^2 J_0(\beta)J_1(\beta)M\left[1-J_0^2\left(1-\left|\Gamma_0^2\right|\right)\right] \quad (16)$$

where M is the mixer conversion loss and Γ_0 is the reflection coefficient at resonance.

The ratio of the two signal-to-noise ratios with equal power incident onto the resonator is approximately

$$\frac{(S/N_{shot})_{T-\Gamma}}{(S/N_{shot})_{PDH}} \cong \frac{1-J_0^2(\beta)\left(1-\left|\Gamma_0\right|^2\right)}{4M\ J_0(\beta)J_1(\beta)} \quad (17)$$

The numerical value of expression (17), fixed $M = 0.5$, $\beta = 1.08$, has the limiting values 0.7 (for Γ_0=0) and 1.5 (for Γ_0=1). In general it shows a better performance for the fringe-side discriminator, however in this analysis the low-frequency noise contribution from the photodetector and the electronics has not been taken into account. Moreover also the laser frequency locking to the established working point is affected by these noise sources. From this point of view the Pound-Drever method is far less sensitive to these effects.

REFERENCES

[1] E. Goobar, "A Michelson interferometer with balanced detection for the characterization of modulation and noise properties of semiconductor lasers," *IEEE J. Quantum Electron.*, vol. QE-29, pp. 1116-1130, Apr. 1993.

[2] J. Helmcke, S.A. Lee and J.L. Hall, "Dye laser spectrometer for ultrahigh spectral resolution: design and performance," *Appl. Optics,* vol. 21, pp. 1686-1694, 1 May 1982.

[3] M. Houssin, M. Jardino and M. Desaintfuscien, "Comparison of the calculated transient responses of a Fabry-Perot used in reflection and in transmission", *Rev. Sci. Instrum.*, vol. 61, pp. 3348-3352, Nov. 1990.

[4] P. Tremblay and R. Ouellet, "Frequency response of a Fabry-Perot interferometer used as a frequency discriminator," *IEEE Trans. Instrum. and Meas.,* vol. IM-40, pp. 204-207, Apr. 1991.

[5] R.W.P. Drever, J.L.Hall,F.V. Kowalsky, J. Hough, G.M. Ford, A. J. Munley, and H. Ward,"Laser Phase and frequency stabilization using an optical resonator," *Appl. Phys. B,* vol. 31, pp. 97-105, 1983.

[6] D. Hils and J.L. Hall, "Response of a Fabry-Perot cavity to phase modulated light," *Rev. Sci. Instrum.*, vol. 58, pp. 1406-14, 1987.

[7] E. Bava and F. Massari,"Phase sensitive detection of light reflected from a Fabry-Perot interferometer," *Rev. Sci. Instrum.*, vol. 67, pp. 1714-1720, May 1996.

[8] E. Bava and F. Massari,"Comparison of calculated phase-sensitive recoverable signals using a Fabry-Perot cavity in reflection and in transmission," *Rev. Sci. Instrum.*, vol. 68, pp. 1927-1934, May 1997.

[9] Y. J. Cheng and P.L. Mussche,"Measurement of laser quantum frequency fluctuations using a Pound-Drever stabilization system," *IEEE J. Quantum Electron.*,vol. QE-30, pp. 1498-1504, June 1994.

1999 Joint Meeting EFTF - IEEE IFCS

TOWARD A 3:1 FREQUENCY DIVIDER BASED ON PARAMETRIC OSCILLATION USING $AgGaS_2$ AND PPLN CRYSTALS

A. DOUILLET and J.-J. ZONDY
Laboratoire Primaire du Temps et des Fréquences (BNM-LPTF)
Bureau National de Métrologie / Observatoire de Paris
61, Avenue de l'observatoire, F-75014, Paris (France)

A. YELISSEYEV, S. LOBANOV and L. ISAENKO
Design & Tecnological Institute of Monocrystals, Siberian Branch of Russian Academy of Sciences, 43 Russakaya Str., 630058 Novosibirsk (Russia)

ABSTRACT

Frequency divide-by-two (2:1) and divice-by-three (3:1) optical parametric oscillators (OPOs) are basic devices for the implementation of future accurate optical frequency division chains. We report our latest development toward a phase-locked 3:1 frequency division of a radiation at $\lambda_p \approx 843$nm (355.9THz), using doubly resonant oscillators (DROs) based on silver gallium sulfide ($AgGaS_2$ or AGS) and multi-grating periodically poled lithium niobate (PPLN). The characteristics of both devices are reported. The major limitation to their practical phase-locking comes from the too weak output coupling efficiency of the idler-wave, with a maximum of 600µW at 2.53µm. Only ~25pW of second harmonic is generated, which avoids efficient beatnote detection.

INTRODUCTION

To measure the frequency of promising cooled atom or trapped-ion frequency standards in the V-UV range it is necessary to develop a new generation of accurate frequency chains based on optical frequency division down to the microwave domain. One of the proposed solutions is based on optical parametric oscillators (OPO) network [1,2]. In this method the absolute frequency value is obtained through difference frequency measurements in the terahertz domain. The advantage of this approach is that an OPO device, based on a second-order nonlinear downconversion process, is a genuine frequency divider. The phase noise of these devices can be very low. Divide-by-two (2:1) and divide-by-three (3:1) OPOs are basic tools for the implementation of such a chain. Until now 2:1 dividers based on quasi-degenerate oscillation using KTP have been demonstrated [3, 4], but 3:1 OPOs are still being investigated [5]. For our 3:1 divider the diode pump laser operates in the range 840-850nm, which corresponds to the 12th harmonic of the CO_2/O_sO_4 frequency standard [6]. The signal and idler output are then λ_s=1260-1275nm and λ_i=2520-2550nm. To implement a 3:1 divider the idler wave must be doubled ($\omega-\delta \rightarrow 2\omega-2\delta$) and the beatenote 3δ between the signal and the doubled-idler is used to phase-lock the pump laser. This device would also replace advantageously ill-behaved transfer laser oscillators in our current frequency multiplication chain.

EXPERIMENTAL SET-UP

The pump source is a master-oscillator power amplifier (MOPA) AlGaAs diode laser setup (Fig.1), which delivers as much as 0.5W of output power. The master oscillator is a home-made AlGaAs extented-cavity diode laser (ECDL). The typical short-term linewidth is ~100kHz, with a maximum short-term drift of a few MHz per minute. Its frequency can be continuously tuned over ~1.5GHz. Its wavelength is controlled by a 500 MHz resolution wavemeter.

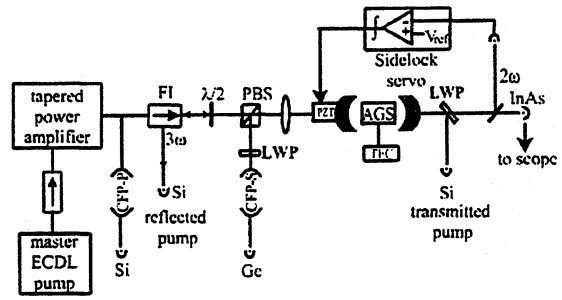


Fig. 1. Type I ($3\omega \rightarrow 2\omega-\delta,\omega+\delta$) DRO setup. FI : 40dB faraday isolator, PBS : polarizing beamsplitter, LWP : long-wave pass filter, CFP : confocal Fabry Pérot, TEC : thermo-electric cooler. The pump beam is polarized in the plane of the figure and the subharmonics in the vertical plane.

The OPO cavity consits of two ZnSe meniscuses with a 10 or 50mm radius of curvature mirrors. Two sets of mirrors with different optical coatings were used for the experiments. The reflectivities are R_s=99.5% and R_i>99.9% for the first set and R_s=99% and R_i~99.8% for the second. In both sets of mirrors

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

the reflectivity at the pump wavelength is R_p~0.6. In the case of AGS-DROs, short cavities with a mirror spacing ranging from 23mm to 30mm are used for resonator stability criteria which will be explained in the following. Only the first set of mirrors were available by the time the AGS experiments were performed. The PPLN-DRO cavity uses a nearly spherical resonator, made of the R=50mm curvature mirrors (mirror spacing= 107mm), in order to satisfy the small waist requirement ($w_{s;i}$<90μm) imposed by the crystal aperture (0.5mm). The input mirror is the most reflecting one while the output coupling belongs to the second set. Two AGS samples differing in their origins and dimensions ($5\times5\times22mm^3$ for AGS-1 and $3\times3\times15mm^3$ for AGS-2) are used. Both samples are triple-band anti-reflection coated with $R_{s,i}$<0.5% at the subharmonics and $R_p\approx5\%$ at the pump and their residual absorption is as low as $0.01cm^{-1}$ at the signal. The multi-grating PPLN crystal is 19mm long (aperture $0.5\times11mm^2$) with eight different gratings whose periods range from 22.4 to 23.1μm. It is also triple-band anti-reflection coated and placed in an oven whose temperature is regulated to less than 50mK. The operating temperature for the 3:1 division of λ_p=842.82nm using the Λ=22.8μm grating is T~125°C. To achieve a single mode pair operation we implement a standard side-of-fringe locking servo.

EXPERIMENTAL RESULTS WITH AGS

The internal pump threshold of these devices is about 100mW. The maximum subharmonic output power does not exceed 2mW.

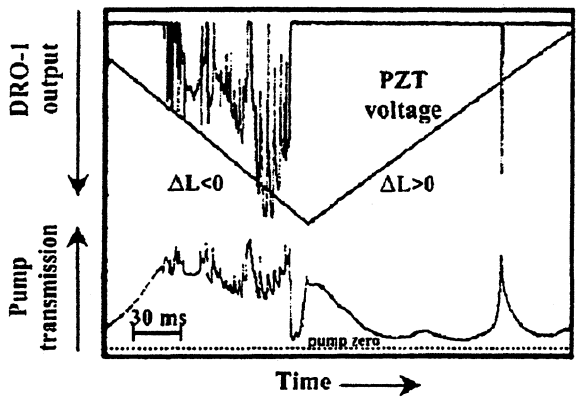


Fig. 2. A trace of the pump and output subharmonic output pattern of DRO-1 with a 29mm mirror spacing, a pump input power of 350mW and a sweeping cavity frequency of 5Hz. The mode pairs are unresolved and the pump depletion is about 50%

Strong thermal loading effects are observed which limit the performance of these DROs. This thermal loading orginates from the residual power absorption, which is converted into heat and from the large figure-of-merit of $AgGaS_2$ $\eta=(dn/dT)/K_c$, where $dn/dT = 1.8\times10^{-4}$/°C is the thermo-optic coefficient [7] and K_c=1.5W/m/°C the thermal conductivity [8]. Compared to commonly used DRO materials such as KTP, LBO or $LiNbO_3$ [9, 10], this thermal figure-of-merit is one order of magnitude larger. Fig.2 shows the typical output patterns of the pump and signal/idler waves under sweeping cavity length operation. It shows an hysteresis on the increasing length scan. The oscillation pattern differs from the usual sharp-spike cluster of modes commonly observed with DRO's. The overall cluster envelope has a triangular shape and each oscillating mode pairs displays a trapezoidal broadened shape. When the cavity length is set in resonance with a subharmonic resonance fringe, we observe self-frequency locking to the single mode pair. The natural tendency of the DRO to mode hop is then strongly inhibited, and the DRO can remain self-frequency locked to the same pair for up to half an hour without the use of an active servo feedback to control the cavity length [5]. This passive thermal-lock is due to the material's ability to correct external cavity length perturbations by increasing or decreasing the optical path length via the thermo-optic effect trigged by the noise-induced intra-cavity signal-idler power fluctuations. The gain of this passive feedback loop [11] depends on the locking point inside the lorentzian fringe transmission peak :

$$G=\frac{\alpha P^{cm}}{4\pi K_c}\times\left[0.57+\ln\left(\frac{2r_0^2}{w^2}\right)\right]\frac{Fl_c}{\lambda}\left(\frac{dn}{dT}\right)\frac{d}{d\delta}\left(\frac{1}{1+\delta^2}\right)$$

where P^{cm} is the maximum stored power, F the resonator finesse, α the power absorption coefficient at frequency ν, r_0 the aperture radius considering the crystal as a rod of length l_c, w the cavity waist, δ the cold cavity detuning $\delta=(\nu-\nu_0)/\Gamma$, ν_0 the cold cavity frequency and Γ the HWHM cold cavity linewith. G is positive (positive feedback loop) on one side of the fringe and negative on the other side, which explains the bistable feature of Fig. 2. The maximum gain is obtained at half fringe maximum where its value is as high as G=188 for λ=2.5μm, $F_i\approx400$, α=0.5 %/cm, w=100 μm, P^{cm}=3 W, G_2=98 W^{-1}, $G_{1m}(\delta=-1/2)$=1.92 W. The thermal loading also modifies the usual linear cavity stability condition, through the thermal lensing effect. Because $(dn/dT)>0$ the thermally loaded medium acts as a power dependent lens and the signal-idler waves must accommodate their growth within the spherical resonator in the presence of this dynamic lens, whose power can be expressed as [11] :

$$p=\frac{1}{f_{th}}=\frac{\alpha(dn/dT)P^c}{\pi K_c}\int_{-l_c/2}^{l_c/2}\frac{dz}{w^2(z)}$$

The study of the stability condition of the symmetric hot resonator leads to the following intra-cavity power stability range :

$$0 \le p \le 2/d$$

where $d=2(L-l_c+l_c/n_0)$ represents the optical distance between one mirror and the middle of the crystal. Hence, the power stability range of the thermally loaded resonators decreases with increasing mirror spacing.

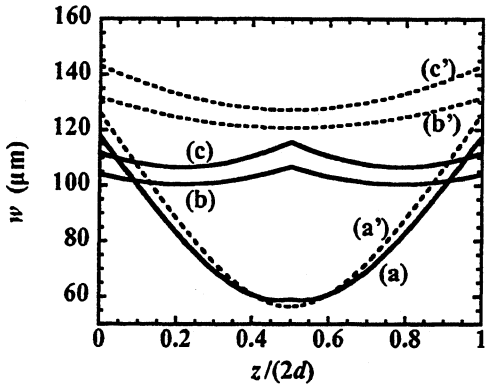


Fig. 3. Idler beam radius versus normalized axial position inside the cavity. Dashed lines correspond to a beam radius without thermal lens (p= 0), and the solid lines to a lens power p = 50dpt. Curves (a) and (a') are plotted for R = 10mm, d = 8mm, curves (b) and (b') for R = 50mm, d = 8mm, curves (c) and (c') for R = 50mm, d = 10.5mm. All the curves are computed with λ=2.5μm.

The computation of the beam radius along the cavity axis evidences two distinct beam waists located symmetrically from the cavity center. For the R=50mm curvature, the two foci lie near the crystal ends and the effect of thermal lensing is severe. When the mirror spacing is increased from 25mm to 33mm, we observe a corresponding pump threshold increase, until oscillation ceases. Close to the stability boundary for the mirror spacing, an instability develops on increasing number of mode pairs in the form of sustained self-pulsing output [11]. With the shorter curvature mirrors (R=10mm) the two waists in fig. 3 are almost merged and located near the center of the crystal. and the thermal load modifies only slightly the beam radius compared to the cold cavity. The OPO is then more stable and the threshold is lower, reflecting an extended power stability range for this DRO geometry.

We have investigated the frequency tuning characteristics under thermal and active cavity length servos. The free spectral ranges ($FSR_{s,i}$) of the subharmonics are $\delta\nu_s = 3243.75\,\text{MHz}$ and $\delta\nu_i = 3154.34\,\text{MHz}$ [11, 12]. Fig. 4 shows the tuning rates measured under thermal-lock. The best linear fit gives $\Delta\nu_s$=0.454$\Delta\nu_p$. A non symmetric signal tuning range is evidenced (200MHz excursion for $\Delta\nu_p$>0 and –700MHz for $\Delta\nu_p$<0). For an increasing pump frequency, the passive servo gain increases whereas it decreases in the opposite direction. The self-locked DRO shows a tendency to resist the tuning process. This effect is inhibited when one operates the DRO with the active servo which maintains the intra-cavity power constant. The suppressed power fluctuations disable the passive thermal lock. The theoretical tuning rate for a DRO under active servo is [11, 13]

$$(\Delta\nu_s)_{lock} = \left(\frac{D(\nu_s)}{\nu_s} + \frac{D(\nu_i)}{\nu_i}\right)^{-1}\left[\frac{D(\nu_i)}{\nu_i}\Delta\nu_p - l_c\left(\frac{\partial n_s}{\partial T} - \frac{\partial n_i}{\partial T}\right)\Delta T\right]$$

yielding a pump tuning rate $(\Delta\nu_s)_{lock}/\Delta\nu_p = 0.66\,\text{Hz/Hz}$ and a temperature tuning rate $(\Delta\nu_s)_{lock}/\Delta T = 306.6\,MHz/°C$. In the above equation $D(\nu)=L-l_c[1-n-\nu(\partial n/\partial\nu)]$ The measured active lock tuning rates agree with the theoretical ones since we obtained $(\Delta\nu_s)_{lock} = (0.66\,\text{Hz/Hz})\Delta\nu_p - (250\,MHz/°C)\Delta T$.

With these AGS-DROs, the ratio of the idler-to-signal power does not exceed 5% due to the mirror transmission loss mismatch. With only 100μW output power, external second-harmonic idler generation is impractical. Higher output coupling of the order of 1% at the idler would be required to generate a mW range idler output.

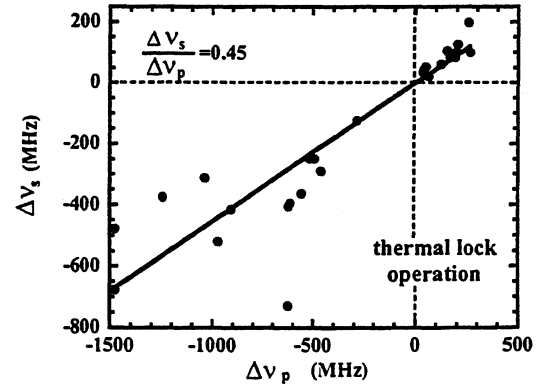


Fig. 4. Mode-hop free signal tuning excursions versus manual pump frequency tuning, under thermal-lock operation. Each point is a measured ($\Delta\nu_s$, $\Delta\nu_p$) until a mode hop occurs.

DRO BASED ON PERIODICALLY POLED LITHIUM NIOBATE

We implement a second 3:1 divider based on PPLN to overcome the thermal power limiting effects. The experimental set-up is sketched in Fig. 4. In this new device we have obtained a pump threshold of 35mW which is comparable with the one reported by Lindsay *et al.* [14] for similar wavelengths. However the output power is limited to 7-10mW at maximum.

With the new set of mirrors the idler-to-signal power ratio is now 10%. The mirror transmission is still not optimized. The experimental set-up is sketched in Fig. 5.

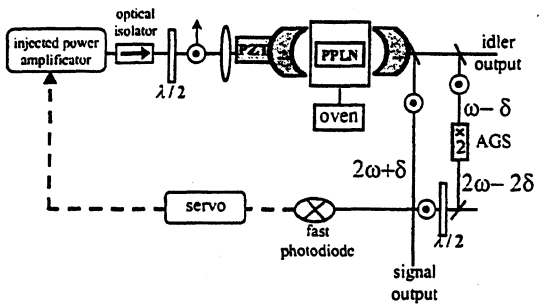


Fig. 5. Experimental device used to phase-lock the DRO. The fast photodiode detects the beatnote between the signal (2ω+δ) and the doubled-idler (2ω-2δ).

With ~600μW of idler output power from the PPLN-DRO, the frequency-doubled idler is limited to ~25pW using a 1 cm long AGS sample (type I cut, θ=43°, φ=45). We have tried unsuccessfully to detect the beatnote. To increase the efficiency of the doubling stage we tried to double the idler in a cavity formed with the OPO mirrors. Due to poor impedance matching with the low transmitting mirrors we could not couple more than 200μW of idler into the cavity which was even lower than the incoming idler power.

CONCLUSION

In this work we have evidenced thermal effects which limit AGS-DROs output performance. However the self-frequency locking is a very attractive phenomenon for a doubly resonant OPO. To phase-lock the 3:1 divider different possibilities are now investigated. The first one is to use new mirrors with higher idler transmission. The second proposed solution is to implement the 3:1 division by inserting in the PPLN-OPO cavity an AGS sample phase matched to generate the second harmonic of the idler. Such a cascaded OPO/SHG interaction device may lead to a self-phase locked operation of the DRO by mutual injection locking of the doubled idler and the signal. Furthermore, the self frequency stabilizing effects of the AGS crystal would help to prevent the numerous mode hops of the PPLN DRO.

REFERENCES

[1] N.C Wong, "Optical frequency division using an optical parametric oscillator", Opt. Lett. vol. 15 (20) , pp. 1129-1131 (1990)

[2] N.C Wong, "Optical-to-microwave frequency chain utilizing a two-laser-based optical parametric network", Appl. Phys. Vol. B 61, pp. 143-149 (1995)

[3] D. Lee and N. C. Wong, " Tunable optical frequency division using a phase-locked optical parametric oscillator", Opt. Lett. Vol. 17 (1), pp. 13-15 (1992)

[4] T. Ikegami, S. Slyusarev, S. Ohshima, and E. Sakuma, "Accuracy of an optical parametric oscillator as an optical frequency divider", Opt.Commun. vol. 127, pp. 69-72 (1996).

[5] A. Douillet and J.- J. Zondy, "Low-threshold, self-frequency-stabilized $AgGaS_2$ continuous-wave subharmonic optical parametric oscillator" Opt. Lett. Vol. 23 (16), pp. 1259-1261 (1998)

[6] O.Acef, "Metrological properties of CO_2/O_sO_4 optical frequency standard" Opt. Comm. Vol. 134, pp. 479-486 (1997)

[7] J.-J. Zondy and D. Touahri, "Updated thermo-optic coefficients of $AgGaS_2$ from temperature-tuned non critical 3ω-2ω→ω infrared parametric amplification", J. Opt. Soc. Am. Vol. B14, pp. 1331 (1997).

[8] J. D. Beasley,"Thermal conductivities of some novel nonlinear optical materials", Appl. Opt. Vol. 33, pp. 1000-1003 (1994)

[9] D. H. Jundt, "Temperature-dependent Sellmeier equation for the index of refraction, n_e, in congruent lithium niobate", Opt. Lett. Vol. 22, pp. 1553-1555 (1997)

[10] V. G Dmitriev, G. G. Ggurzadyan and D.N. Nikogosyan, in Handbook of Nonlinear Optical Crystals, Springer-Verlag, Berlin-Heideberg, pp. 76 (1991)

[11] A. Douillet, J.- J. Zondy, A. Yelisseyev,S. Lobanov and L. Isaenko, "Stability and frequency tuning of thermally loaded continuous-wave $AgGaS_2$ optical parametric oscillators", submitted to J. Opt. Soc. Am. , special feature on optical parametric devices (sept. 1999)

[12] R.G. Smith, "A study of factors affecting the performance of a continuously pumped doubly resonant optical parametric oscillator", IEEE J. Quantum Electron. Vol. QE-9, pp. 530-541 (1973).

[13] R. Al-Tahtamouni, K. Bencheikh, R. Stortz, K. Schneider, M. Lang, J. Mlynek, and S. Schiller, "Long-term stable operation and absolute frequency stabilization of a doubly resonant parametric oscillator", Appl. Phys. Vol. B66, pp. 733-739 (1998).

[14] J. D. Lindsay, G. A. Turnbull, M. H. Dunn and M. Ebrahimzadeh, "Doubly resonant continuous-wave optical parametric oscillator pumped by a single-mode diode laser", Opt. Lett. Vol. 23, pp. 1889-1891 (1998).

1999 Joint Meeting EFTF - IEEE IFCS

ULTRA-STABLE OPTICAL FREQUENCIES FOR SPACE

[1]Clifford C. Hodge, [1]Hugh A. Klein, [2]David J.E. Knight and [3]Lute Maleki

[1]National Physical Laboratory (NPL), Teddington, Middlesex, UK TW11 0LW
[2]DK Research, 110 Strawberry Vale, Twickenham, Middlesex, UK TW1 4SH
[3]Jet Propulsion Laboratory (JPL), Pasadena, CA 91109-8099, USA

ABSTRACT

Attracted by the prospect of vastly increased bandwidth, a move from the microwave to the optical for space-based communications is now underway. It is important to consider the associated optical frequency source requirements. "Direct detection" systems have been considered for the first phase of commercial inter-satellite optical-communication links. "Coherent communication", which offers at least an order of magnitude improvement in capacity, requires spectrally narrow sources, and is more difficult to implement. Ultra-narrow lasers will be needed for low data-rate applications such as the deep-space network (DSN). The required stability and accuracy of the lasers will also depend on the additional demands of any navigation, tracking and "radio/light science" functions served. Cavity-stabilised narrow lasers are expected to be the basis of the best laser oscillators operating in the optical for space applications.

1. INTRODUCTION

Terrestrial telecommunications moved from the microwave to optical carriers for long-distance transmission with the advent of compact diode lasers and lightweight low-loss fibre: these optical systems offer much higher bandwidths (potentially up to 25 THz) and cheaper transmission. There are many reasons why optical communication systems will supplement microwave-based systems for inter-satellite and other space applications. With much higher frequency (about 100,000 times greater) the optical, when compared to the microwave, has the potential for a commensurate increase in both frequency accuracy [1,2] and bandwidth or information-carrying capability. Coupled with factors such as improved directionality for transmission (for example, the Voyager spacecraft mission to Saturn would have a beam over 2 million times more concentrated using optical systems [3]), lower equipment mass and dimensions, lower power consumption and higher detector efficiencies, with secure communication links, optical technology for space application appears very attractive.

In addition to communication requirements, there are other compelling arguments for developing ultra-stable optical frequency sources and detectors for use in space. At present, the microwave radiation is not only used for communication and navigation/tracking, but also sometimes for scientific experiments collectively referred to as "radio science" [3]. As the switch from microwave to optical for communications proceeds, the component technology will find other uses in navigation, earth observation, inter-satellite synchronisation, and potentially for a whole gambit of aptly-termed "light-science" experiments. Although direct intensity modulation or "digital" (direct detection) has been used for early flight experiments and operational missions, coherent systems will be required for the more demanding applications. Following an outline of the space application of optical frequencies, this paper highlights some of the laser-frequency control technology which will provide the next-generation of optical sources, frequency standards and ultimately even "optical clocks".

2. APPLICATION AREAS OF OPTICAL FREQUENCIES IN SPACE

2.1 Communications

Reports on a variety of projects studying different aspects of space-based optical or laser communications which have been underway over the last two decades in Europe, Japan and the USA are reported in a 1998 issue of Space Communications[4]. The organisations interested in this field range widely from strongly commercial (e.g. driven by multimedia and Internet applications), to the military, to Deep-Space Network (DSN) users. Active project examples include:

(i) The European Space Agency (ESA), SILEX (Semiconductor laser Intersatellite Link EXperiment) which will transmit 50 Mbps from LEO to GEO using direct detection[5];
(ii) The US-Japanese "Ground-to-Orbit Lasercom Demonstration" (GOLD) - the first successful bi-directional optical communication with a satellite with rates of up to a Mbps, and;
(iii) A programme funded by the US Ballistic Missile Defence Organisation will test intersatellite optical communication links[6].

2.2 Navigation, Tracking and T2L2

Tracking system frequency requirements for the DSN are currently 3×10^{-13} at 1000 seconds[7]. Examples might include: (i) satellite missions which pass close to the Sun, for the study of gravitational effects, and checking for variation in the fundamental constants[8]; (ii) occultation-type experiments; or (iii) missions where

0-7803-5400-1/99/ $ 10.00

precise one- or two-way Doppler tracking of the spacecraft signal is used. The Satellite Laser Ranging 2000 (SLR 2000) mission looks to develop high-bandwidth real-time communication links to Mars in the future. An example of time-transfer by laser-link (T2L2) is the MIR-based experiment which should make it possible to measure the performance of ground-based clocks having a relative stability of about $3x10^{-15}$[9], and for the ACES ensemble on-board the ISS against ground-based clocks using T2L2[10]. An important advantage of optical over the microwave for ranging and time-transfer is the immunity to refractive index effects in the ionosphere and solar wind. Dispersion effects from the atmosphere are an active area of research, particularly through the GOLD programme using multi-frequency transmission with up to four carriers, to ameliorate scintillation effects. Anderson *et al.* recently reported an apparent anomalous weak long-range acceleration[11] from Galileo and Voyager satellite tracking data. They suggest that "the Pluto express mission could provide an excellent opportunity for high quality data from very deep-space, especially if optical tracking is used". The GNSS system requires very stable clocks for best performance. These may eventually be optical clocks, related to the microwave via highly accurate phase-locked links[1].

2.3 "Radio" and "Light" Science

A move to the optical for space-based communications opens up the vista of "light" rather than "radio" science experiments[3]. The most stringent requirements for DSN frequency source stability may be experiments which call for 1.5×10^{-15} over 1000 to 3600 seconds for the Cassini mission[12]. A wealth of space-based scientific experiments may benefit from the use of ultra-stable laser sources and better clocks[13]. Two examples of research which would benefit from very well controlled optical frequency sources are given below:

A "Pre-Phase A" Study has already been conducted into the feasibility of launching the Laser Interferometer in Space Antenna (LISA) platform[14], as part of ESA's Horizon 2000 Programme. The LISA objective of detecting gravitational waves will test the laser and interferometer physics package to the limit (typical interferometer stability of 10^{-19} at 10^3 seconds averaging): mHz-wide space-qualified lasers are required[15]. A precursor to an all-optical LISA system might be the development of optical hardware on the International Space Station (ISS): a first step to realising this goal is the ACES (Atomic Clock Ensemble in Space) which includes a laser link for accurate time- and frequency-transfer.

Very-Long Baseline Interferometry (VLBI) systems currently using microwave frequencies and microwave atomic frequency standards for synchronisation are limited by atmospheric propagation of the radiation through the ionosphere and the troposphere, and the ability to resolve the carrier phase. Two developments are currently taking place. The first proposes to operate an array of satellites in a new Space VLBI network - the Space Interferometry Mission (SIM), bringing better solid-angle resolution and atmosphere-free performance. The second proposal for next-generation European (or global) VLBI is to steer the network with optical clocks either terrestrially (through optical fibres), or via a space link. Either way, ultra-stable optical clocks rather than microwave frequency standards have an important role to play in the future to furthering our understanding of pulsars, and for geodesy.

3. OPTICAL TECHNOLOGY FOR SPACE APPLICATIONS

The time-scale of the required laser stability depends on the space application (or combination of applications). For purely coherent communications, only linewidth is important, i.e., short-term stability over less than a fraction of a second. The DSN frequency source stability-time requirements are generally, (for say Doppler ranging) dominated by the time taken for the electromagnetic radiation to travel across the solar system - typically thousands of seconds. Certain scientific experiments (and other applications) may demand even longer stability timescales. In addition, the absolute value of the frequency source - its accuracy - may become important.

Coherent communications requires narrow lasers with linewidths ranging from sub-Hz to > MHz, depending on the data-rate. Typically, the required linewidth is a few orders of magnitude below the data bit-rate. Two types of cavity can be used for narrowing and stabilisation. Such ultra-stable optical frequency sources are based on two complementary technologies to stabilise the laser radiation: (i) high Finesse ultra-low expansion (ULE) cavities[16], and (ii) cryogenic sapphire dielectric cavity resonators[17,18]. Sub-Hz ULE-stabilised laser linewidths[16] have been observed on Earth, but these may become dominated by vibration at the highest levels of stability. Vibrations can be a serious problem on satellites[19], but it may ultimately prove possible to provide better vibration isolation in space than on Earth.

Whilst the SILEX and GOLD programs are based on direct detection, subsequent ground-based research and development of optical terminals in Europe has focused on coherent systems such as the SROIL and SOLACOS systems mentioned above[4]. The issue of direct detection versus coherent (heterodyne or homodyne) systems for communications has been examined by numerous authors over the past 15 years (see for

example, Chan[20]). However, in some situations where power is at a premium, where signal-to-noise is bad or data rates are pushed to their limits (> 1 GHz), coherent detection becomes attractive. For low-data-rate applications, where only low power levels can be achieved, coherent systems will require state-of-the-art optical frequency sources - linewidths as narrow as 25 Hz have been mentioned for a DSN relay station application for the Pluto Express mission[21].

The Institute for Communications Technology of the German Aerospace Centre (DLR) has developed an "absolute frequency reference at 1064nm for coherent transmission and laser drift measurement"[22]. The doubled (green 532 nm) output is stabilised to the flank of a Doppler-broadened iodine absorption line to 7 MHz (2.5×10^{-8}). Whilst this is modest compared with the ultimate potential of Doppler-free and ion-trap-based optical standards currently being developed[1,16,23], it may be adequate for short-term requirements. It is important to look further ahead and consider more demanding applications. For example, it has been suggested that long-term laser stability of better than one part in 10^{16} over 1000 seconds would be required to provide a performance comparable to that which is projected in the Ka-band for Doppler tracking [12,24], say for the Cassini mission.

Accurate standards will be required for optical frequency space applications, where sub-kHz or even sub-Hz levels may be required. Iodine and other molecular-stabilised lasers can have stabilities of better than parts in 10^{12} at 10 seconds[2]. Recent results in iodine-stabilised Nd:YAG lasers[25,26,27] have been impressive, with stabilities below 10^{-13} for timescales greater than 1 second. However, such molecular-stabilised laser standards are inherently limited by transit-time broadening and second-order Doppler shifts. Such effects may be avoided by using ultra-narrow (less than 1 Hz level) atomic or ionic transitions in species which are trapped and laser-cooled. Laser sources referenced to transitions with natural linewidths close to or below 1 Hz are the basis of the latest generation of optical frequency standards[1,27], and may lead to the development of optical clocks with superior performance compared to existing microwave clocks. To take advantage of such narrow features requires extremely narrow "flywheel" cavity-stabilised lasers, based on the ULE and cryogenic-cavity techniques described above.

Atomic clocks and frequency standards have been an area of intense interest for more than fifty years, with at least an order of magnitude improvement in stability achieved every decade. There is no foreseeable reason why this progress should not continue[28]. An Allan variance analysis shows that the stability of any frequency standard depends on the line $Q = \upsilon/\Delta\upsilon$, where υ is the transition frequency and $\Delta\upsilon$ is the experimentally observed linewidth (limited by interaction time or the natural width). The inherent potential stability of optical (or laser) standards with frequencies near 10^{15} Hz are superior to those in the microwave (10^{10} Hz), since $\Delta\upsilon$ is similar in both regions of the electromagnetic spectrum: they are widely expected to become the next generation of ultra-stable accurate frequency sources. Eventually, the timing community could redefine the second in terms of a forbidden optical transition in a trapped-ion[1,16,23]. The principle components of such devices are compact, solid-state, and amenable to space qualification. Some solid-state lasers have withstood Shuttle vibration tests, whilst the space environment is certainly accessible, and offers potential performance benefits.

The wide range of potential space applications demanding stable lasers can probably only be matched by a range of devices and technologies. However, there are significant areas of overlap and already the import of certain common elements has been recognised. Today, diode and other solid-state lasers are most attractive. For example, Nd:YAG lasers (1064 nm) and frequency-doubled Nd:YAG (532 nm) lasers should be able to serve a number of the more demanding application areas mentioned above. Nd:YAG is the laser source proposed for LISA[14], and they are already used in the "SROIL" and "SOLACOS" optical terminals[4]. Also, Nd:YAG lasers are used in satellite laser ranging (SLR) and time-transfer by laser-link (T2L2) experiments. As well as the laser sources themselves, related technologies have to be developed. Pointing, acquisition and tracking capabilities are far more demanding in the optical than the microwave, and there has been much work in the last decade on improving this capability, but together with laser stabilisation, it is one of the most challenging research areas.

4. CONCLUSION

Research in the area of optical sources, frequency standards and clocks is needed to establish the best choices for lasers and associated equipment for future space applications. The demands for optical technology are likely to be applications-driven, starting with narrow-linewidth lasers for commercial (and other) communications. The functions of navigation, tracking and "radio/light science" will push the stabilities and accuracies required from sources. The shift from all-microwave technology to a space-based scenario where optical technology augments microwave technology is already underway: the ultimate system could be all-optical, but that is likely to be well into the next millennium!

ACKNOWLEDGEMENTS

CCH, HAK and DJEK would like to thank ESA and the UK NMSPU for supporting studies of optical frequency control applications at NPL. Discussions with Franco Emma, Stephen Feltham, Mike Geer, Hans Fromm and numerous others are gratefully acknowledged.

REFERENCES

[1] Proceedings of the Fifth Symposium on Frequency Standards and Metrology, ed. James C. Bergquist, World Scientific, Singapore (1996) and references therein.

[2] D. J. E. Knight, "Present Frequency Standards for Lasers", Proc. 11th EFTF, Neuchâtel pp. 623-627 (4-6 March 1997).

[3] James R. Lesh and Marc D. Rayman, "Deep-space missions look to laser communications", Laser Focus/Electro-Optics, pp. 81-86 (October 1988).

[4] Special Issue: CRL International Technical Workshop on Space Laser Communications, (10-11 March 1997, Tokyo, Japan), Space Communications 15, No. 2, pp. 65-122 (1998).

[5] G. Oppenhauser, M. Wittig, and A. Popescu, "The European SILEX project and other advanced concepts for space communications", Proc. SPIE (Optical Space Communication II) 1522 pp. 2-13 (1991).

[6] "Laser Transceiver to Launch", Laser Report (1998).

[7] "Frequency systems and timing" (1993-1997) JPL Doc. No D-10145) (15 November 1993).

[8] Stephen Hawking, "A Brief History of Time" p. 138, Bantum Press (1995 edition, reprinted 1998).

[9] E. Samain and P. Fridelance, "Time Transfer by Laser Link (T2L2) experiment on MIR", Metrologia 35, pp. 151-159 (1998).

[10] P. Fridelance, E. Samain & T. Melliti, "Scientific applications of the Time Transfer by Laser Link (T2L2)", Proc. 11th EFTF, Neuchâtel (WORKSHOP: "New Generation of Space Clocks) pp. 659-662 (7 March 1997).

[11] John D. Anderson, Philip A. Laing, Eunice L. Lau, Anthony S. Liu, Michael Martin Nieto and Slava G. Turyshev, Physical Review Letters 81 pp. 2858-2861 (1998).

[12] G. J. Dick and R. T. Wang, TMO Progress Report 42-134 (15 August 1998).

[13] "Proceedings of the Workshop on the Scientific Applications of Clocks in Space", ed. Lute Maleki, NASA JPL Publication 97-15, August 1 1997, and references therein.

[14] LISA Pre-Phase A Report 2nd edition, (July1998).

[15] ESA Intended Invitation to Tender 98.1XA.07 "High stability laser for space interferometry" (1999).

[16] Brenton C. Young et al. "Lasers for an Optical Frequency Standard using Trapped Hg+ Ions" Conf. Proc. 457 (Trapped Charged Particles and Fundamental Physics) pp. 337-342, ed. Daniel H.E. Dubin and Dieter Schneider, AIP (1999); also B. C. Young et al. "Visible lasers for sub-hertz linewidths", Physical Review Letters (accepted for publication in April 1999).

[17] Stefan Seel, Rafael Storz, Giuseppe Ruoso, Juergen Mlynek and, Stephan Schiller, "Cryogenic Optical Resonators: A New Tool for Laser Frequency Stabilisation at the 1 Hz Level", Physical Review Letters, 78 pp. 4741-4744 (23 June 1997).

[18] Farhat Abbas, David W. Allan, Neil Ashby, Clifford C. Hodge, et al. "A novel concept for the direct intercomparison of the Allan variance of frequency standards at the 10^{-18} level with application to metrology, global communication and navigation", Proc. 11th EFTF, Neuchâtel p. 473, (4-6 March 1997).

[19] S. Arnon and N. S. Kopeika, "Laser satellite communication network-vibration effect and possible solutions", Proc. IEEE 85 pp. 1646-1661 (1997).

[20] Vincent W. S. Chan, "Space Coherent Optical Communication Systems-An Introduction", J. Lightwave Technol. LT-5 pp. 633-637 (April 1987).

[21] G. Stephen Mecherle, Wade Akle, C. John Starkus, James E. Klein and Gerald W. Holleman, "Coherent detection optical relay station satellite for deep-space communication", (Free-Space Laser Communication Technologies VI) Proc. SPIE 2123 pp. 218-236 (1994).

[22] M. Schulthess and D. Giggenbach, "Development of absolute frequency reference at 1064 nm for coherent transmission and laser drift measurements", Electronics Letters 34 pp. 1854-1855 (1998).

[23] M. Roberts, P. Taylor, G. P. Barwood, P. Gill, H. A. Klein, W. R. C. Rowley, "Observation of an electric octupole transition in a single ion" Physical Review Letters 78, pp. 1876-1879 (10 March 1997); and G. P. Barwood, P. Gill, G. Huang, H. A. Klein, E. Riis, "Sub-kHz 'clock' transition linewidths in a cold trapped $^{88}Sr^{+}$ ion in low magnetic fields using 1092-nm polarisation switching", Opt. Comm. 151 pp. 50-55 (1998).

[24] C. -C. Chen and M. Z. Win "Laser stability requirements for coherent space communications", Proc. 45th Annual Symposium on Frequency Control pp. 500-507 (1991).

[25] G. M. Macfarlane, et al. "Interferometric Frequency Measurements of an Iodine-stabilised Nd:YAG laser", IEEE Trans. Instrum. Measure. to be published (1999).

[26] John L. Hall, Ye Jun, Ma Long-Sheng, S. Swartz, P. Jungner, S. Waltman, "Optical frequency standards, some improvements, some measurements, and some dreams" in [1], pp. 267-276 (1996).

[27] Revision of the *mise-en-pratique* of the definition of the metre 1997, to be published. See also Le System international d'unités 7th edition (1998).

[28] David W. Allan, Neil Ashby and Clifford C. Hodge, "The Science of Timekeeping", Hewlett-Packard Application Note 1289 pp. 1-88 (June 1997); alternatively: David W. Allan, Neil Ashby & Cliff Hodge, "A brief history of precise time and GPS", Precise Timing, GPS World Supplement, pp. 1-40 (December 1998).

FABRY-PEROT RESONATOR WITH INTERFEROMETRIC READ-OUT FOR LOW NOISE APPLICATIONS

Michael E. Tobar
Department of Physics, University of Western Australia, Nedlands, WA, Australia

ABSTRACT

New readout configurations for a Fabry-Perot resonator that combine the reflected and transmitted waves are analyzed. The waves maybe combined in a Sagnac or Mach-Zehnder (MZ) configuration and tuned to a dark port to suppress the carrier frequency. Both the Sagnac and MZ schemes are analyzed in detail. Suppressing the carrier minimizes the Shot noise floor of a Pound-Drever-Hall (PDH) frequency stabilized laser. Moreover, it is shown that a large improvement in frequency noise can be achieved for high power PDH stabilized lasers utilizing only a small phase modulation index.

1. INTRODUCTION

Various methods of stabilizing an oscillator to a high-Q cavity exist at microwave and laser frequencies. A high-Q cavity combined with a detector at the output comprises a frequency discriminator. Typically the frequency discriminator detects and cancels the oscillator frequency noise as part of a frequency control servo system. The sensitivity and noise characteristics of the frequency discriminator ultimately determine the performance of a well-stabilized oscillator. Obviously, the characteristics of the frequency discriminator are very important and any improvements will directly influence the potential of the stabilized frequency noise.

The implementation of frequency stabilization schemes at microwave frequencies goes back to the 1940's to the work of Pound. He implemented DC stabilization techniques and pioneered the now well known AC modulation technique [1]. It was not until the 1980's that these techniques were brought into the optical domain. The subtle difference at optical frequencies arises due to the lack of sensitive low noise double balanced mixers. Nevertheless, Hansch and Couillaud proposed an analogous DC stabilization technique. In contrast to microwave frequencies, at optical frequencies the frequency discriminator must separate the different light polarization to create an error signal [2]. Around the same time, the equivalent Pound technique was proposed by Drever and Hall [3], and is now referred to the Pound-Drever-Hall (PDH) technique.

Galani et. al. introduced a further advance in DC frequency stabilization at microwave frequencies [4]. They combined the transmitted and reflected signals of the high-Q resonator to improve the frequency discrimination. This technique was also bought to the optical domain by Kourogi and Ohtsu [5]. They adapted the DC polarization technique of Hansch and Collard in a scheme that combined the reflected and transmitted waves in a Mach-Zehnder interferometer. They achieved a factor of 10 larger frequency discrimination over the Hansch Couillaud scheme.

It has also been realized over the years that suppressing the carrier with a microwave interferometer read-out can improve the frequency discrimination of a high-Q microwave cavity [6-8]. However, it was only recently that this method reached its full potential to accomplish a major advance in microwave frequency noise reduction and measurement techniques (patented) [9] [10].

If resonant cavities could be impedance matched perfectly there would be no need to use an interferometric readout for low noise and high power applications. The necessity arises because even in a well matched cavity the reflected power from the cavity will still be significant. A high level of reflected power could destroy the low noise operation of the detection system by causing excess power dependent flicker or Shot noise, or even damaging the device itself. In contrast, a high level of power is necessary to achieve a large phase sensitivity of detection. In the microwave domain an effective impedance match to 1 part in 10^{12} accuracy has been achieved by introducing the interferometric readout along with a control element that maintains a dark port on reflection [10].

In this paper some equivalent optical circuits, such as a high finesse Fabry-Perot resonator embedded in a Sagnac or a MZ Interferometer are analyzed. It is shown that the realization of such devices should likewise advance noise reduction techniques at optical frequencies.

2. MACH-ZEHNDER READ-OUT

The MZ read-out for a FP resonator is shown schematically in fig. 1. Here ρ_1 and ρ_2 are the amplitude reflection coefficients of the two FP mirrors (assume $\rho_1 \leq \rho_2$), θ is the return phase shift, α is the return loss, θ_1 and θ_2 are the phase shifts in the interferometer arms and α_1 and α_2 are the losses in the interferometer arms. The beam splitter reflection coefficient that combines the reflected and transmitted waves is assume to be equal to $1/\sqrt{2}$, with a π phase shift on the left side and a 0 phase shift on the right. In general, the output at the dark (DP) and bright ports (BP) in fig. 1 can be shown to be;

$$DP = \frac{\left(\rho_1 e^{\alpha-\alpha_1+j(\theta-\theta_1)} - \rho_2 e^{-\alpha_1-j\theta_1} + \sqrt{1-\rho_1^2}\sqrt{1-\rho_2^2}\, e^{(\alpha-2\alpha_2+j(\theta-2\theta_2))/2}\right)}{\sqrt{2}\left(e^{\alpha+j\theta} - \rho_1\rho_2\right)} \quad (1)$$

$$BP = \frac{\left(-\rho_1 e^{\alpha-\alpha_1+j(\theta-\theta_1)} + \rho_2 e^{-\alpha_1-j\theta_1} + \sqrt{1-\rho_1^2}\sqrt{1-\rho_2^2}\, e^{(\alpha-2\alpha_2+j(\theta-2\theta_2))/2}\right)}{\sqrt{2}\left(e^{\alpha+j\theta} - \rho_1\rho_2\right)} \quad (2)$$

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

If the transmitted FP power is greater than the reflected, an alternative output can be created at the right mirror by substituting it for a beam splitter. The required reflectivity of the beam splitter will be determined by the required loss, α_2-α_1.

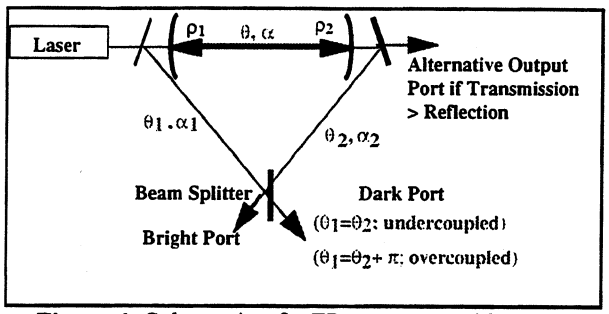


Figure 1. Schematic of a FP resonator with a MZ readout.

A dark port is achieved by adding attenuation and phase in the interferometer arms. This may be calculated by setting (1) to zero and solving for the real (α_1-α_2) and imaginary parts (θ_1-θ_2).

$$e^{(\alpha_2-\alpha_1)} = \frac{e^{\alpha/2}\sqrt{1-\rho_1^2}\sqrt{1-\rho_2^2}}{\rho_2-\rho_1 e^{\alpha}}(\rho_1 < \rho_2 e^{-\alpha}) \quad or \quad \frac{e^{\alpha/2}\sqrt{1-\rho_1^2}\sqrt{1-\rho_2^2}}{\rho_1 e^{\alpha}-\rho_2}(\rho_1 > \rho_2 e^{-\alpha}) \tag{3a}$$

$$\theta_1 - \theta_2 = 0\,(\rho_1 < \rho_2 e^{-\alpha})\; or\; \pi\,(\rho_1 > \rho_2 e^{-\alpha}) \tag{3b}$$

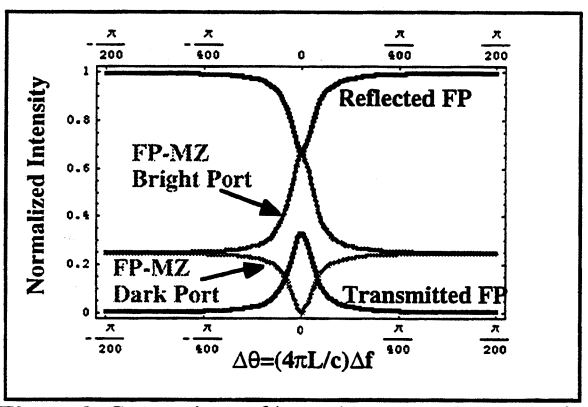


Figure 2. Comparison of intensity versus detuning of a FP and FP-MZ interferometer, with α=0, ρ_1=0.999, ρ_2=0.9999.

The dark and bright ports are compared to the normal FP response, as a function of detuning from the resonant frequency in fig. 2. Here it has been assumed that α=0 (no loss), ρ_1=0.999 and ρ_2=0.9999. The attenuation and phase conditions to obtain a dark port were calculated from (3) to be; θ_1=θ_2 and α_1=α_2+0.353. Also, the following is defined, θ=θ_{res}+$\Delta\theta$, where θ_{res}=2nπ is the on resonance cavity phase length, n is the order of the resonance, $\Delta\theta$ = (4πL/c)Δf is the detuning from resonance and L is the cavity length. Fig. 3 shows the dark port response from (1), for finite values of loss given by condition (3).

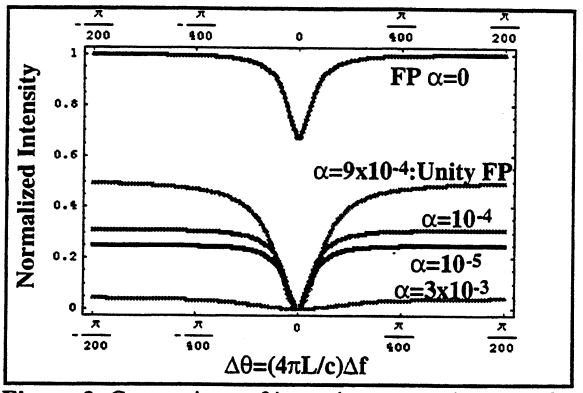


Figure 3. Comparison of intensity versus detuning for the dark port of the FP-MZ interferometer, with finite loss, ρ_1=0.999 and ρ_2=0.9999.

3. SAGNAC READ-OUT

The Sagnac read-out for a FP resonator (FP-S) is shown schematically in fig. 4. The read-out looks similar to the FP-MZ configuration except the laser is input at the beam splitter. Similar assumptions stated in section 2 are made for the FP-S structure. In general, the output at the dark (DP) and bright ports (BP) in fig. 4 can be shown to be;

$$DP = \frac{\left(\begin{array}{l}\rho_2\left(e^{\alpha-2\alpha_2+j(\theta-2\theta_2)}+e^{-2(\alpha_1+j\theta_1)}\right)\\ -\rho_1\left(e^{\alpha-2\alpha_1+j(\theta-2\theta_1)}+e^{-2(\alpha_2+j\theta_2)}\right)\end{array}\right)}{2\left(\rho_1\rho_2-e^{\alpha+j\theta}\right)} \tag{4}$$

$$BP = \frac{\left(\begin{array}{l}\rho_2\left(e^{\alpha-2\alpha_2+j(\theta-2\theta_2)}-e^{-2(\alpha_1+j\theta_1)}\right)+\\ \rho_1\left(e^{\alpha-2\alpha_1+j(\theta-2\theta_1)}-e^{-2(\alpha_2+j\theta_2)}\right)+\\ 2\sqrt{1-\rho_1^2}\sqrt{1-\rho_2^2}\,e^{\alpha+j\theta-2(\alpha_1+\alpha_2+j\theta_1+j\theta_2)}\end{array}\right)}{2\left(\rho_1\rho_2-e^{\alpha+j\theta}\right)} \tag{5}$$

A dark port is achieved by adding attenuation and phase in the interferometer arms. This may be calculated by setting (4) to zero and solving for the real (α_2-α_1) and imaginary parts (θ_2-θ_1).

$$e^{2(\alpha_2-\alpha_1)} = \frac{\rho_1-\rho_2 e^{\alpha}}{\rho_1 e^{\alpha}-\rho_2}(\rho_1 < \rho_2 e^{-\alpha}) \quad or \quad \frac{\rho_2 e^{\alpha}-\rho_1}{\rho_1 e^{\alpha}-\rho_2}(\rho_1 > \rho_2 e^{-\alpha}) \tag{6a}$$

$$\theta_2 - \theta_1 = \pi/2\,(\rho_1 < \rho_2 e^{-\alpha})\; or\; 0\,(\rho_1 > \rho_2 e^{-\alpha}) \tag{6b}$$

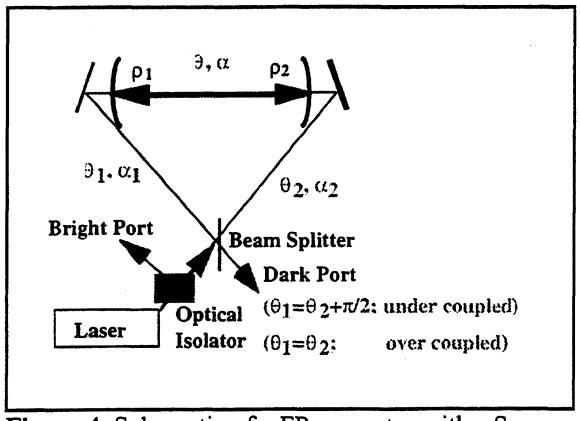


Figure 4. Schematic of a FP resonator with a Sagnac readout.

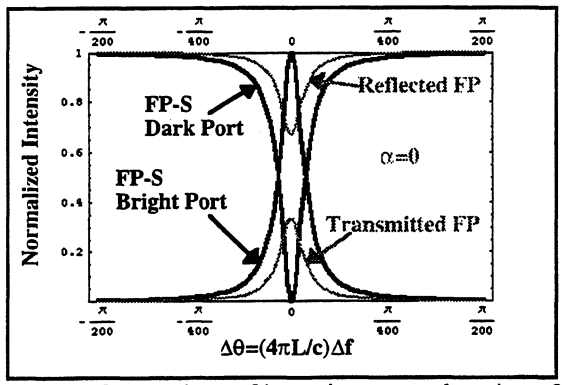


Figure 5. Comparison of intensity versus detuning of a FP and FP-S interferometer, with $\alpha=0$, $\rho_1=0.999$, $\rho_2=0.9999$.

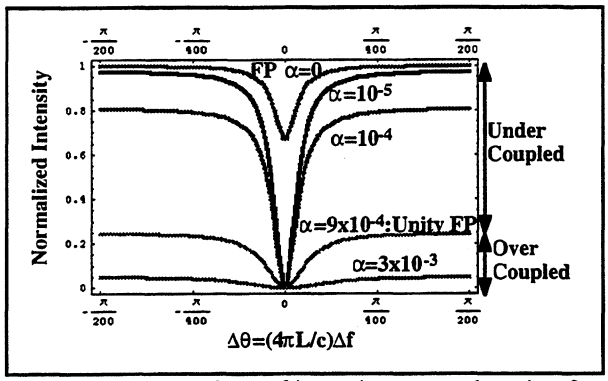


Figure 6. Comparison of intensity versus detuning for the FP-S interferometer, with finite loss, $\rho_1=0.999$ and $\rho_2=0.9999$.

The dark and bright ports are compared to the normal FP response in fig. 5, assuming, $\alpha=\alpha_1=\alpha_2=0$ (no loss), $\rho_1=0.999$ and $\rho_2=0.9999$ and $\theta_2-\theta_1=\pi/2$. Note that for a FP-S system with no loss the dark port is always exact and no attenuation need be added. Fig. 6 shows the dark port response from (4), for finite values of loss given by condition (6).

4. COMPARISON OF THE MACH-ZEHDNER AND SAGNAC READ-OUT

In this section the frequency discrimination sensitivity of the FP-MZ and FP-S interferometers are compared to the usual FP cavity. The exact value of the discriminator slope is given later in equation (7) and (8). To compare the sensitivity we assume identical FP cavities, phase modulation indices and photodetectors. The sensitivity can be compared by only considering how the slope of the imaginary response varies with frequency or phase length, i.e., $\partial\left(\mathrm{Im}[\Gamma_{eff}(\theta)]\right)/\partial\theta$ at $\theta=0$. Here $\Gamma_{eff}(\theta=0)$ is the response of the discriminator at the resonant frequency. For the FP cavity $\Gamma_{eff}(\theta)$ is the reflection coefficient. For the FP-MZ and FP-S interferometers $\Gamma_{eff}(\theta)$ are given by (1) and (4) subject to conditions (3) and (6) respectively. These values are plotted below in fig. 7.

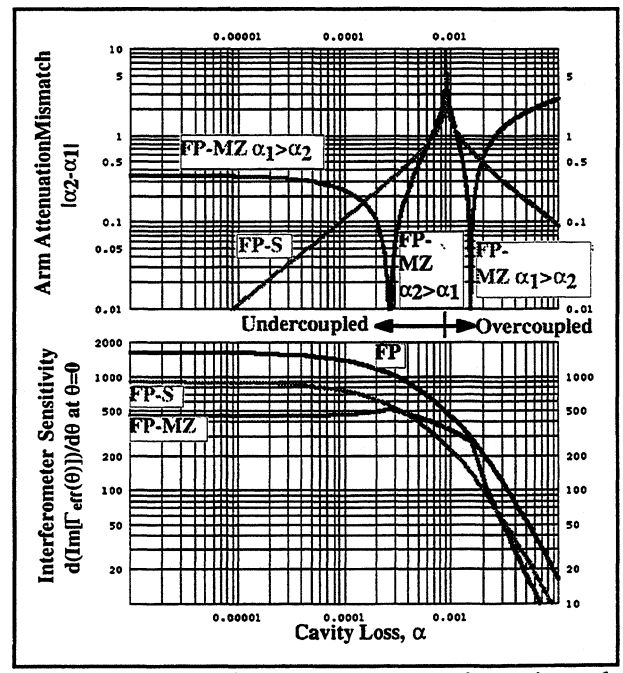


Figure 7. Above, the necessary attenuation mismatch added to the interferometer arms to create a dark port, for both the FP-S and FP-MZ interferometers. Below, discriminator sensitivity as a function of cavity loss. (Plotted results assume $\rho_1=0.999$ and $\rho_2=0.9999$)

In general the FP sensitivity is greater than the FP-S and FP-MZ interferometers. If the contrast of the FP cavity is bad, (i.e. it is well under or over coupled) then the Sagnac read-out is more sensitive than the MZ readout. However at values close to unity coupling the MZ readout is more sensitive than the Sagnac. In general, the MZ readout is more sensitive when the transmitted power from the FP cavity is greater than the reflected power.

5. LASER FREQUENCY STABILISATION WITH THE INTERFEROMETRIC READOUT

It is common to use a resonant optical interferometer as a frequency discriminating element in a frequency stabilization scheme. The Pound-Drever-Hall (PDH) scheme is most common [3], and we limit our discussions to this configuration. Expanding on the derivation by Day et. al. [11], the normalized discriminator sensitivity may be calculated to be;

$$D_N = \frac{4J_o(\beta)J_1(\beta)}{\Delta f_{BW}}\left(\Delta\theta_{BW}\left.\frac{\partial\left(\mathrm{Im}[\Gamma_{eff}(\theta)]\right)}{\partial\theta}\right|_{\theta=0}\right) \quad [1/\mathrm{Hz}] \tag{7}$$

Here Δf_{BW} is the bandwidth of the resonator in Hz, $\Delta\theta_{BW}$ is the bandwidth in radians and Γ_{eff} is the response of the interferometer at the photodetector. For a FP cavity in reflection this is simply the reflection coefficient. If the FP cavity is impedance matched then it is easy to show that the second term in brackets in (7) reduces to two, and D_N reduces to the value derived in Day et al. [11].

Assuming the PDH servo is locked with very large gain, then the limit to the laser frequency noise will solely depend on the discriminator sensitivity and noise. To calculate the discriminator sensitivity in A/Hz, (7) is multiplied by the current induced at the photodiode $e\eta P_{inc}/h\nu$. Here e is the charge of an electron, η is the detector efficiency, P_{inc} is the power incident on the resonator, h is planks constant and ν is the laser frequency. Also given that $\Delta\theta_{BW}/\Delta f_{BW} = 4\pi L/c$, the discriminator sensitivity may be written as;

$$D_A = J_o(\beta)J_1(\beta)\left(\frac{16\pi L}{c}\right)\left(\frac{e\eta P_{inc}}{h\nu}\right)\left(\left.\frac{\partial\left(\mathrm{Im}[\Gamma_{eff}(\theta)]\right)}{\partial\theta}\right|_{\theta=0}\right) \quad [\mathrm{A/Hz}] \tag{8}$$

The next step is to calculate the Shot noise in the frequency discriminator. The spectral density of current noise at the photodetector is given by;

$$S_A = \sqrt{2}\sqrt{2eI} = 2\sqrt{e\left(\gamma^2 J_0^2(\beta) + 2J_1^2(\beta)\right)\frac{e\eta P_{inc}}{h\nu}} \quad [\mathrm{A/\sqrt{Hz}}] \tag{9}$$

Here $2J_1^2(\beta)P_{inc}$ is the power in the modulation sidebands at the photodetector and $\gamma^2\ J_0^2(\beta)P_{inc}$ is the carrier power at the photodetector. The value, γ^2, is the carrier suppression factor supplied by the frequency discriminator. For example, for an exactly impedance matched FP cavity there is no reflected carrier and $\gamma^2=0$, and (9) reduces to the Shot noise calculated in Day et. al. [11].

By combining (8) and (9) the closed loop performance of the frequency stabilized laser may be calculated to be;

$$S_{f.cl} = \frac{\left(\frac{\sqrt{\gamma^2 J_0^2(\beta) + 2J_1^2(\beta)}}{\sqrt{2}J_o(\beta)J_1(\beta)}\right)}{\left(\frac{4\pi L}{c}\sqrt{\frac{2e\eta P_{inc}}{h\nu}}\left.\frac{\partial\left(\mathrm{Im}[\Gamma_{eff}(\theta)]\right)}{\partial\theta}\right|_{\theta=0}\right)} \quad [\mathrm{Hz/\sqrt{Hz}}] \tag{10}$$

The denominator in (10) is dependent on the laser power and cavity and detector properties. To investigate the effects of various degrees of carrier suppression, γ, and modulation index, β, the denominator is normalized to unity. The normalized frequency noise floor (numerator in (10)) is plotted in figure 8 for various levels of carrier suppression.

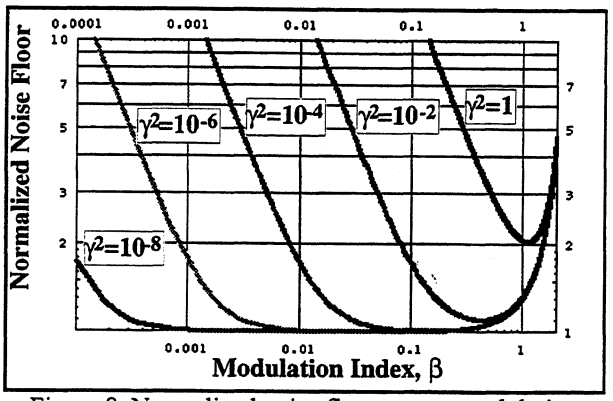


Figure 8. Normalized noise floor versus modulation index for various levels of carrier suppression.

The discriminator sensitivity given by (7) and (8) is maximized when β=1.08 radians. However, this value is not necessarily the best choice for a Shot noise limited system. The large modulation index also creates large sideband power, which creates extra Shot noise. It is more prudent to minimize the noise floor. If no carrier suppression is supplied by the frequency discriminator (i.e. γ^2=1), then the optimum modulation index to minimize the noise floor is 1.1 radians. As the level of carrier is reduced, the optimum modulation index is also reduced. The more exact the carrier suppression the more independent the noise floor is on modulation index. Also, fig. 8 shows that the noise floor is reduced by a factor of two from the optimum non-carrier suppressed response.

The factor of two reduction in noise floor is not the major benefit. The major benefit arises from the fact that in a well suppressed system the noise floor becomes independent of modulation index. This allows the discriminator to be operated with a small modulation index. The small modulation index means that there is only a small amount of power at the photodetector for a large incident power on the discriminator. Given that a photodetector can only handle a power of order 100 mW before it is destroyed, the carrier suppression and small modulation index enables the incident power on the

discriminator to be increased substantially above 100 mW. Very high power frequency stabilized lasers (10-100 W) are now under construction for the Laser Interferometer gravitational wave detectors [12]. If this power can be utilized with this discrimination technique, a large improvement of several orders of magnitude in frequency noise should be possible.

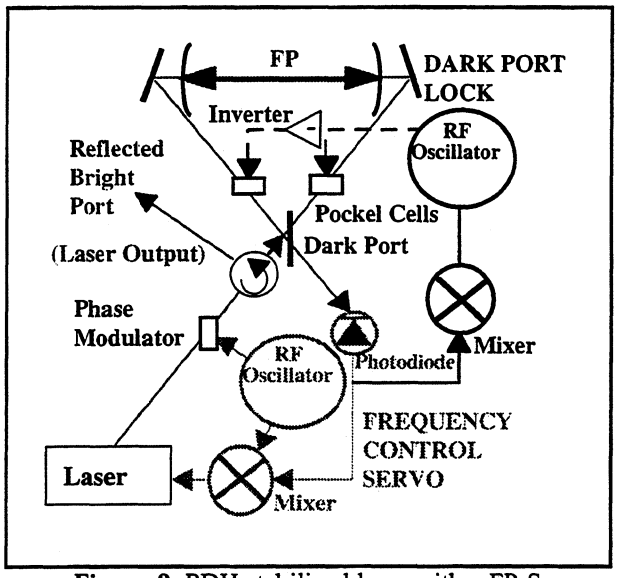


Figure 9. PDH stabilized laser with a FP-S interferometer as a frequency discriminator and a dark port locking circuit.

From the results plotted in fig. 8 it is clear that a well matched FP cavity will not be sufficient to obtain several orders of improvement. Even with only 1% of the carrier power ($\gamma^2=10^{-2}$) the modulation index could only be reduced by at most an order of magnitude. Correspondingly, the power could be increase by one order, resulting in an order of magnitude improvement. At microwave frequencies a γ^2 of 10^{-12} has been obtained with a dark port locking system [10]. To achieve a similar amount in the laser system we need to also introduce a dark port locking system. This can be done in a similar way to a dark port lock on a laser interferometer gravitational wave detector [12]. A schematic of a PDH stabilized laser with a FP-S discriminator and a dark port lock is shown in fig. 9.

6. CONCLUSION

New laser stabilization techniques that implement a carrier suppression interferometer in the read-out of the frequency discriminator have been proposed. The considered discriminators consist of a Fabry-Perot cavity imbedded in a Mach-Zehnder (FP-MZ) or Sagnac (FP-S) interferometer. The sensitivity of the FP-MZ and FP-S interferometers was compared to the standard FP interferometer. It was shown that a large improvement in laser frequency noise could be achieved for high power lasers utilizing the new interferometric discriminators. This improvement in sensitivity is necessarily achieved at low phase modulation indices.

7. ACKNOWLEDGMENT

The author would like to thank Andre Luiten, Eugene Ivanov and John Hartnett for many interesting discussions. This work was supported by the Australian Research Council.

8. REFERENCES

[1] R. V. Pound, "Electronic Frequency Stabilization of Microwave Oscillators," *Rev. Sci. Intrum.*, vol. 17, 1946.

[2] T. W. Hansch and B. Couillaud, "Laser frequency stabilization by polarization spectroscopy of a reflecting refemce cavity," *Optics Comm.*, vol. 35, pp. 441-444, 1980.

[3] R. W. P. Drever, J. L. Hall, F. V. Kowalski, J. Hough, G. M. Ford, A. J. Munley, and H. Ward, "Laser phase and frequency stabilization using an optical resonator," *Appl. Phys. B*, vol. 31, pp. 97-105, 1983.

[4] Z. V. I. Galani, M. J. Bianchini, J. Raymond C. Waterman, R. Dibiase, R. W. Layton, and J. B. Cole, "Analysis and design of a single-resonator GaAs FET oscillator with noise degradation," *IEEE Trans. on Microw. Theory and Tech.*, vol. 32, pp. 1556-1565, 1984.

[5] M. Kourogi and M. Ohtsu, "Novel optical frequency discriminator for FM noise reduction of semiconductor lasers," *Optics Comm.*, vol. 81, pp. 204-208, 1991.

[6] A. L. Whitwell and N. Williams, "A new microwave technique for determining noise spectra at frequencies close to the carrier," *The Microwave Journal*, pp. 27-32, 1959.

[7] F. Labaar, "New discriminator boosts phase-noise testing," in *Microwaves*, 1982, pp. 65-69.

[8] G. J. Dick and D. G. Santiago, "Microwave frequency discriminator with a cryogenic sapphire resonator for ultra-low phase noise," in Proc. IEEE Frequency Control Symposium, 1992.

[9] E. N. Ivanov, M. E. Tobar, and R. A. Woode, "Microwave interferometry: Application to precision measurements and noise reduction techniques," *IEEE Trans. on Ultrason. Ferroelec. Freq. Contr.*, vol. 45, pp. 1526-1536, 1998.

[10] E. N. Ivanov, M. E. Tobar, and R. A. Woode, "Applications of Interferometric Signal Processing to Phase Noise Reduction in Microwave Oscillators," *IEEE Trans. on Microw. Theory and Tech*, vol. 46, pp. 1537-45, 1998.

[11] T. Day, E. K. Gustafson, and R. L. Beyer, "Sub-hertz relative frequency stabilization of two-diode laser-pumped Nd: YAG Lasers locked to a Fabry-Perot Interferometer," *IEEE Journal of Quantum Elec.*, vol. 28, pp. 1106-1117, 1992.

[12] P. R. Saulson, *Fundamentals of Interferometric Gravitational Wave Detectors*: World Scientific, 1994.

DEVELOPMENT TOWARDS A SPACE QUALIFIED LASER STABILIZATION SYSTEM IN SUPPORT OF SPACE-BASED OPTICAL INTERFEROMETERS*

DAVID J. SEIDEL, SERGE DUBOVITSKY
Time and Frequency Systems Sciences and Technology Group
Jet Propulsion Laboratory, California Institute of Technology
4800 Oak Grove Drive, Pasadena, CA 91109 USA

ABSTRACT

At JPL we are developing a space qualified laser stabilization system to support NASA's effort for space-based stellar interferometers. We report on the development, functional performance and space-qualification status of a such a system.

1. INTRODUCTION

The Space Interferometry Mission (SIM) and Deep Space 3 (DS-3) are two missions currently funded by the National Aeronautics and Space Administration (NASA) that are space-based optical interferometers. In order to properly recombine the starlight received at each telescope of the interferometer, it is necessary to perform high resolution laser metrology. A potentially significant error source in performing high resolution length measurements is the potential for fluctuations in the laser "gauge" itself. If the laser frequency or wavelength is changing over time it will be misinterpreted as a length change in one of the legs of the interferometer. An analysis of the frequency stability requirement for SIM resulted in a fractional frequency stability requirement of $\sqrt{S_y(f)} \le 2\times10^{-12}/\sqrt{Hz}$ at Fourier frequencies between 10 Hz and 1000 Hz[1]. The DS-3 mission stability requirement is further increased to $\sqrt{S_y(f)} \le 5\times10^{-14}/\sqrt{Hz}$ at Fourier frequencies between 0.2 Hz and 10 kHz with a goal of extending the low frequency range to 0.05 Hz[2]. The baseline laser for both SIM and DS-3 is the Lightwave Electronics NPRO laser operating at 1.3 μm. The free running performance of these lasers vary in stability and we have measured them to perform as follows

$$\frac{9\times10^{-11}}{f(Hz)}\frac{Hz}{\sqrt{Hz}} \le \sqrt{S_y(f)} \le \frac{1.3\times10^{-8}}{f(Hz)}\frac{Hz}{\sqrt{Hz}}.$$

In order to improve the performance of the laser we stabilize the laser to a high finesse optical cavity by locking the optical frequency of the laser to one of the transmission modes of the cavity. At JPL we have built a prototype space-qualifiable system meeting the stability requirements of SIM, which has been integrated into a metrology source for high-resolution hetrodyne interferometer laser gauges[3] and delivered to the JPL Micro-arcsecond metrology testbed[4]. We have also started on the development of a system to meet the stability needs of DS-3.

2. REQUIREMENTS VS. FREE RUNNING LASER PERFORMANCE

The frequency stability required depends on the configuration of the interferometer and on the differences in optical path length between the collector telescopes and the optical beam combiner. An analysis of the stability requirement was performed for both the SIM and DS-3 missions based on assumptions of update rates and the degree of overlap of all the internal pathlengths of the interferometer. For the SIM mission the frequency stability requirement, expressed in terms of the square root of the power spectral density of frequency fluctuations, is $\sqrt{S_y(f)} \le 455Hz/\sqrt{Hz}$ at Fourier frequencies (f) between 10 and 1000 Hz. For the DS-3 mission the frequency stability requirement is further increased to $\sqrt{S_y(f)} \le 10Hz/\sqrt{Hz}$ at Fourier frequencies between 0.2 Hz and 10 kHz. The performance was measured by optically mixing the light from two different lasers and monitoring the beat frequency using an HP5371A Frequency and Time Interval Analyzer to obtain a time series of frequency data. The data obtained is then Fourier transformed resulting in a frequency domain laser performance measurement. The requirements along with a summary of the measured free running performance curves are summarized in figure 1. As can be seen from the figure the free running performance of the laser is not good enough to meet the requirements of either SIM or DS-3.

* This work was performed at the Jet Propulsion Laboratory, California Institute of Technology under a contract with the National Aeronautics and Space Administration.

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

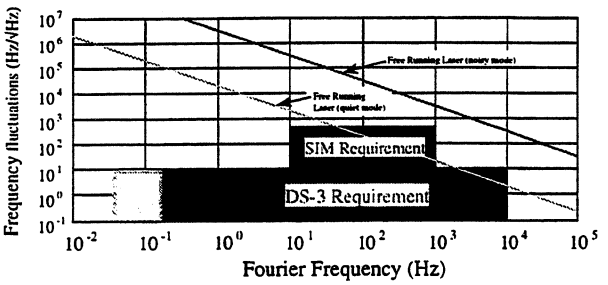


Figure 1: Free running Lightwave NPRO Laser measurements.

3. STABILIZATION SYSTEM DESIGN AND DEVELOPMENT

In order to improve the frequency stability of the laser we stabilize the laser to a high finesse optical cavity by locking the optical frequency of the laser to one of the transmission modes of the cavity. The optical cavity is a semi-monolithic design in which two 10cm radius of curvature mirrors are optically contacted to a 5cm hollow cylindrical spacer. Both the mirror substrates and cylindrical spacer are constructed from Corning Ultra-Low Expansion (ULE) Glass. Preliminary measurements yield a coefficient of thermal expansion (CTE) on the order of 3 - 5 x 10^{-8}/°C for our optical cavity. The free spectral range of the cavity is 3GHz and the finesse is 10,000 which results in a cavity linewidth of 300kHz. The laser is locked to the cavity using the Pound-Drever-Hall stabilization scheme [3]. A diagram of the setup is shown in Fig. 3.2.

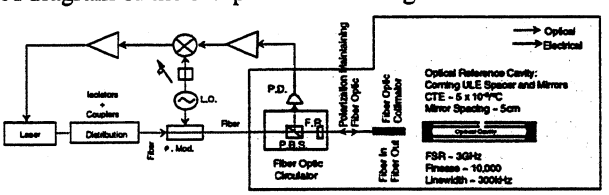


Figure 2: Schematic layout of the laser stabilization scheme.

The laser radiation is first phase-modulated and then coupled to the optical cavity via an optical fiber and a fiberoptic collimator. The optical signal reflected from the cavity is coupled back into the same optical fiber and directed onto a photo-detector, via a fiberoptic circulator, and it is monitored at the frequency of the phase modulation. This signal is then demodulated to produce a DC error signal with amplitude and sign linearly proportional to the frequency mismatch between the cavity resonance and the laser frequency (Fig. 3b). Figure 3.3 shows both the cavity reflection response and the error signal generated as the lasers frequency is swept through resonance.

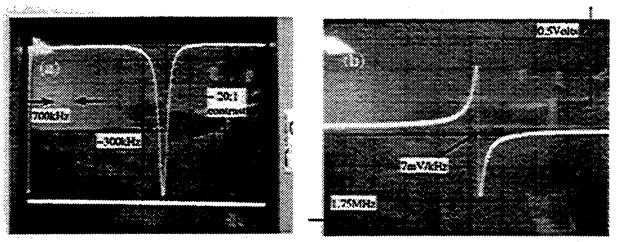


Figure 3: (a) Cavity Reflection Response. (b) Error signal proportional to the detuning between the laser frequency and the cavity resonance.

In addition to measuring the characteristics of the optical cavity and the Pound-Drever-Hall locking system, it was necessary to fully characterize the laser tuning parameters. The laser has two main tuning elements: i) a piezo-electric transducer (PZT) mounted to the laser crystal that can modulate the laser frequency at reasonably high frequency (~100kHz) but with limited range; ii) a thermal electric cooler which can be used to adjust the laser frequency by adjusted the temperature of the crystal. Temperature tuning, while having a much slower response, has a much greater range. Both of these methods cause physical changes in the crystal dimension, resulting in a change in the operating frequency of the laser. The philosophy adopted for the feedback circuitry was to have a fast analog integrator circuit to control the PZT tuning input with a computer controlled digital crossover to the temperature tuning input. The computer monitors the output voltage from the analog circuit and, at predetermined set points, applies small changes in the voltage on the thermal input of the laser controller. Essentially a slow digital loop de-saturates the fast analog loop. We found that the PZT response was ~ 3.5MHz/Volt with less than 2dB of fluctuation out to 100 kHz. Past 100kHz the response starts rolling off quite rapidly and there exist strong resonances in the system at frequencies greater than 200kHz (fig. 4). The response of the thermal tuning is about 2.4 GHz/°C (3.8 GHz/Volt at the laser controller) over the linear regions between the mode hops.

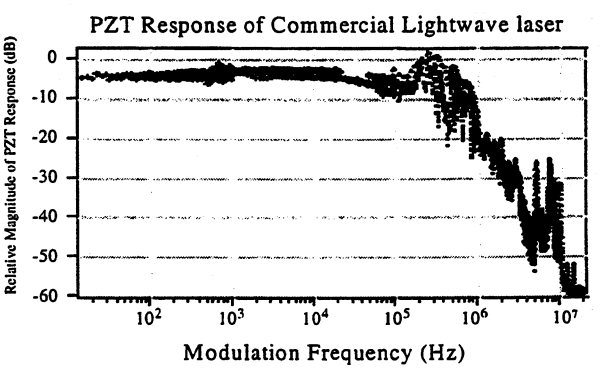
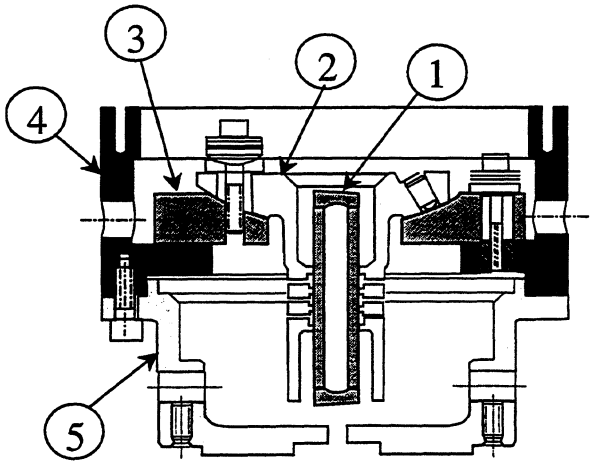


Figure 4: Laser PZT frequency modulation response

In order to stabilize the frequency of the laser to one of the cavity resonances it is necessary to properly couple the laser radiation into TEM_{00} mode of the cavity. The critical parameters involved in this are the translation and angular orientation of the cavity with respect to the collimated laser beam and mode matching the beam waist of the cavity with the beam waist of the input. Mode matching the beam waists is typically done with the use of a separate lens and setting the appropriate spacing between the optical elements. In our case we were able to specify an integrated fiberoptic collimator with a specific beam waist and position(Optic for Research, Inc.). At JPL we designed a four-axis mount (Fig. 5) together with an assembly jig allowing translational and tip/tilt adjustment of the cavity relative to the optical axis. It was not necessary to design and manufacture a full five axis mount since the spacing between the fiberoptic collimator and the optical cavity can be specified to within acceptable tolerances without seriously compromising the optical coupling efficiency. The mount holds both the cavity and the optical collimator; the assembly jig is used to align the cavity with respect to the optical collimator in order to achieve optimum coupling. This system worked extremely well and we achieved an optical coupling efficiency ~95% (Fig. 3a). Once alignment is completed the cavity is firmly locked in place resulting in an extremely rigid structure. We have experienced no degradation in the optical coupling efficiency since the original alignment.



ITEM	TITLE	MATERIAL
1	Optical Cavity	ULE Glass
2	Gimbal Cavity Holder	Titanium
3	XY Stage/Gimbal Table	Al 7075T7351
4	Gimbal/Translation Stage Housing	Al 7075T7351
5	Collimator Support	Al 7075T7351

Figure 5: Optical Cavity Mount Assembly.

One of the key features of this system is that a single optical fiber is used to couple light both in and out of the cavity. There was a concern that back reflections from the fiberoptic collimator would add noise due to phase difference fluctuations between the undesired light back-reflected from the collimator and the desired light reflected from the cavity. For the current specifications these back reflections did not impact the performance of the system.

A major perturbation to the stability of the cavity system is the presence of air. The change in the index of refraction due to temperature and pressure fluctuations at standard temperature and pressure is on the order of -8×10^{-7}/°C and 3×10^{-7}/mmHg respectively [4]. This is equivalent to a change in the physical dimension of the cavity and is substantially worse than the CTE ($\sim 5 \times 10^{-8}$/°C) of the optical cavity. In order to avoid this problem we have mounted the system inside a vacuum chamber and pipe light into and out of the vacuum system through a JPL built polarization maintaining fiberoptic feedthrough. The optical cavity, vacuum system and additional optical components are packaged into a single 10.5" high standard rack mount box (Fig. 6).

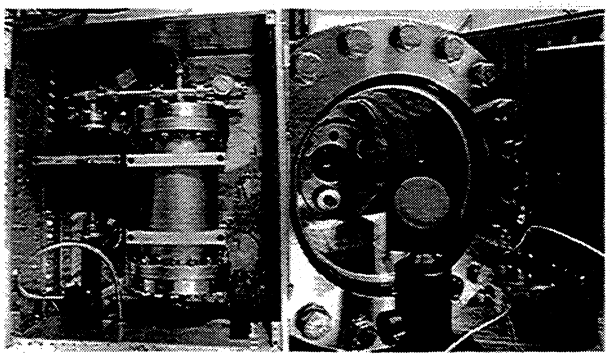


Figure 6: Optical Cavity, associated Vacuum system, and additional optical components.

4. PERFORMANCE RESULTS

Two complete systems have been built and we have measured the performance of the stabilized lasers in an identical manner as we measure the performance of the free running lasers. We found that we were consistently able to achieve frequency fluctuations below $10Hz/\sqrt{Hz}$ at Fourier frequencies between 10Hz and 1kHz (Fig. 7). This greatly exceeds the SIM design goal and improves upon the free running laser performance by 40 to 80dB.

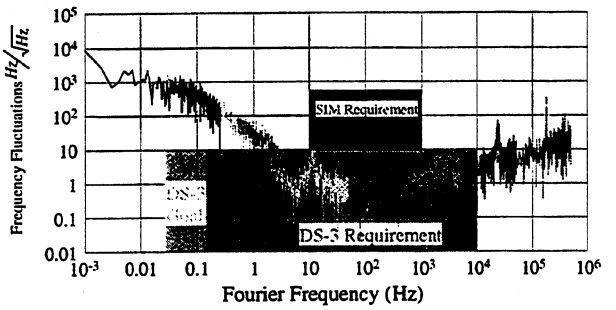


Figure 7: Square root of the Power Spectral Density of frequency fluctuations between 2 stabilized Lightwave lasers. The increase in noise at high frequencies is the measurement system noise floor.

5. REFERENCES:

1. S. Dubovitsky and D.J. Seidel, "Metrology source frequency stability requirements", JPL IOM #346-96-001, 1996.
2. S. Dubovitsky, Private Communications
3. S. Dubovitsky, D.J. Seidel, D. Liu, and R.C. Gutierrez, "Metrology source for high-resolution Hetrodyne Interferometer laser gauges", *Proceeding of SPIE Conf. On Astronomical interferometry,* ed. R. Reasenberg, vol. **3350**, pp. 973-984, 1998.
4. S. Shaklan, S. Azevedo, R. Bartos, A. Carlson, Y. Gursel, P. Halverson, A. Kuhnert, Y. Lin, R. Savedra, E. Schmidtlin, "The micro-arcsecond metrology testbed (MAM)", *Proceeding of SPIE Conf. On Astronomical interferometry,* ed. R. Reasenberg, vol. **3350**, pp. 1009-1019, 1998.
5. R.W.P. Drever, J.L. Hall, F.V. Kowalski, J. Hough, G.M. Ford, A.J. Munley and H. Ward, "Laser phase and frequency stabilization using an optical resonator", *Appl. Phys. B,* vol. **31,** pp. 97-105, 1983.
6. D.E. Gray, *American Institute of Physics Handbook,* McGraw-Hill, New York, 1972.

$^{199}Hg^{+}$ OPTICAL FREQUENCY STANDARD: PROGRESS REPORT*

R. J. Rafac, B. C. Young, F. C. Cruz, J. A. Beall, J. C. Bergquist, W. M. Itano, and D. J. Wineland

National Institute of Standards and Technology, Boulder, CO 80303 USA

ABSTRACT

We are developing an optical frequency standard based on the narrow $^2S_{1/2}-^2D_{5/2}$ electric-quadrupole transition of a single trapped $^{199}Hg^{+}$ ion. Small linear traps designed to operate at liquid helium temperatures have been constructed which will provide confinement in the Lamb-Dicke regime for the optical transition. Cryogenic operation yields long ion storage times and significantly suppresses frequency shifts due to collisions and blackbody radiation. A major step toward the completion of this standard has been the construction and evaluation of a well isolated, high-finesse, Fabry-Pérot cavity to stabilize the frequency of the interrogating laser. We report our cavity results and progress toward locking the laser frequency to the transition in the stored ion.

1 INTRODUCTION

Atomic ions stored in ion traps experience only very small perturbations of their internal energy levels, and hence make good candidates for stable and accurate frequency standards. Neutral atom and trapped ion standards based on microwave transitions have been developed that demonstrate fractional stabilities near or better than $10^{-13}\tau^{-1/2}$, where τ is the measurement averaging time. However, certain proposals such as the timing of millisecond pulsars and low-frequency gravitational wave detection require performance beyond that of present day standards [1, 2, 3]. The quantum-projection-noise-limited Allan variance of an atomic standard interrogated by the time-domain Ramsey method is given by the expression

$$\sigma_y(\tau) = \sqrt{\frac{\langle(\langle\omega_k\rangle_\tau - \langle\omega_{k+1}\rangle_\tau)^2\rangle_\tau}{2\omega_0^2}} = \frac{1}{\omega_0\sqrt{NT_R\tau}}. \quad (1)$$

In this expression for the two-sample Allan variance $\sigma_y(\tau)$, $\langle\omega_k\rangle_\tau$ corresponds to the kth measurement of the frequency of an oscillator locked to a reference consisting of an ensemble of N atoms each having transition frequency ω_0. The interrogation period is of length $T_R < \tau$. Although much performance improvement has been realized in microwave standards by increasing the number of atoms N in the probed sample and by extending the interrogation time T_R, gains in stability may be offset by losses in accuracy and *vice versa*. Such a degradation may arise, for example, from the introduction of uncontrolled perturbations such as collisional or Doppler shifts in dense or large samples, or reduction in signal when evolving for long T_R due to loss of atomic beam intensity. Increasing the frequency ω_0 by making use of long-lived optical transitions is perhaps the most direct way to improve the overall performance, but significant technical obstacles must be overcome. In particular, optimum performance requires a spectrally pure optical local oscillator possessing a linewidth less than the natural width of the atomic transition for times comparable to the interrogation time T_R.

We report on progress made toward the realization of such an optical frequency standard based on the narrow $5d^{10}6s\ ^2S_{1/2} - 5d^9 6s^2\ ^2D_{5/2}$ electric-quadrupole transition of a single laser-cooled $^{199}Hg^{+}$ ion stored in a radiofrequency (Paul) trap [4]. The single-photon transition wavelength is 282 nm with a natural width of 1.8 Hz, corresponding to a line Q of greater than 5×10^{14}. We expect that it should be possible to stabilize the frequency of a laser oscillating in resonance with the atomic transition (or a subharmonic thereof) within $10^{-15}\tau^{-1/2}$, with an accuracy ultimately approaching 10^{-18}.

2 REFERENCE CAVITY

A two-step scheme is employed to stabilize a ring-dye laser operating at 563 nm, which lies at the heart of the local oscillator. The laser is first prestabilized by locking to a low-finesse ($\mathcal{F} \approx 800$) Fabry-Pérot cavity using the Pound-Drever-Hall reflected sideband technique [5]. High-frequency correction of the laser frequency noise is accomplished with an intra-

cavity electro-optic modulator, and long term drifts are eliminated by steering a piezoelectric transducer (PZT) mounted behind one of the dye laser cavity mirrors. The prestabilization stage narrows the laser linewidth to ≈1 kHz for averaging times of about 1 s. The light is then transported through an optical fiber to a high-finesse ($\mathcal{F} > 150\ 000$) cavity, and the reflected sideband technique is again used in conjunction with an acousto-optic modulator (AOM) to lock the laser to the cavity resonance. The short-term corrections are written directly onto the AOM drive frequency, while long-term drifts are fed back to a PZT on the prestabilization cavity to maintain frequency alignment with the high-finesse cavity.

Our requirement that the spectral purity of the laser be less than the 1.8 Hz natural linewidth of the $^2S_{1/2}$–$^2D_{5/2}$ optical clock transition in $^{199}Hg^+$ places stringent demands on the physical stability of the high-finesse cavity. For this reason, the cavity spacer must have intrinsically low sensitivity to temperature variations and must be well isolated from the environment. The present apparatus uses mirrors optically contacted to the ends of a tapered cylindrical spacer 15 cm in diameter and 24 cm in length; a 1 cm diameter open bore forms the intracavity region. Both the spacer and mirror substrates are made of ULE glass [6], and the temperature is regulated near the point at which the temperature-dependent coefficient of thermal expansion is zero (≈30 °C). The cavity is isolated from the environment by mounting it in an evacuated chamber on an optical table which is passively isolated from seismic noise using a support system consisting of vertical strands of surgical tubing stretched to approximately 3 m. In the vertical, the fundamental vibrational mode of the suspended table has a frequency of ≈0.3 Hz, which provides an isolation from floor noise that exceeds a factor of 50 in noise amplitude at frequencies greater than 3 Hz. Some viscous damping is also employed, and the optical table is further enclosed in a wooden box lined internally with lead foam to reduce the coupling of acoustic noise into the cavity.

Dissipation of intracavity light in the dielectric coatings and radiation pressure on the mirrors shifts the cavity resonance. To stabilize this shift and hold it to an acceptable value, we couple only approximately 100 μW of 563 nm light into the cavity, and a servo actively controlling the radiofrequency power driving the AOM stabilizes the optical power exiting the cavity to ≈0.1 %.

To characterize the short term stability of the cavity without reference to the atomic ion, we utilize a second similar high-finesse cavity on a separate isolated table. Figure 1 shows the spectrum of the beat note obtained by heterodyning two laser beams derived from the beams stabilized to the two independent cavities. A nearly uniform relative cavity drift of ≈1 Hz/s is suppressed by mixing the beat note with a swept frequency synthesizer. A common-mode drift of similar size is also observed. The width of the spectrum at the half-power point is about 0.9 Hz (32 s averaging time). Assuming that the length fluctuations of the cavities are similar in frequency distribution and amplitude but otherwise independent implies that the frequency fluctuations of neither laser contributes a frequency width greater than 0.6 Hz at 563 nm. This corresponds to a fractional linewidth of only 1×10^{-15}. Time-domain measurements described in Ref. [7] were used to determine the fractional frequency instabilities $\sigma_y(\tau)$ for the 30 ms to several second time scale relevant for our proposed optical standard. These results, and the Allan deviations reported for a number of other stable laser systems, are compared in Figure 2.

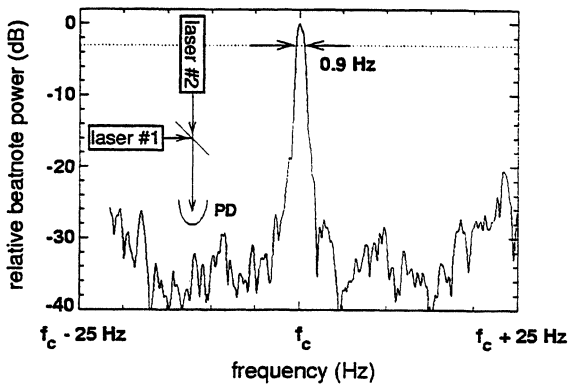


Figure 1: Power spectrum of the beat note between two laser beams stabilized to two independent cavities. The resolution bandwidth of the spectrum analyzer is 0.447 Hz, and the averaging time is 32 s.

3 SINGLE-ION REFERENCE

We work with singly charged ions of ^{199}Hg because they offer transitions suitable for both optical [4] and microwave standards [14, 15]. Figure 3 shows the $^2S_{1/2}$–$^2P_{1/2}$ electric-dipole transitions used for laser cooling, optical pumping, and state detection, and the $^2S_{1/2}$–$^2D_{5/2}$ electric-quadrupole transition that is the reference for the optical standard. A number of technical barriers must be surmounted in working with Hg^+. One is associated with the presence of residual elemental Hg in the trap vacuum vessel. At room temperature, the storage lifetime is reduced to a few minutes through recombination of the trapped ions with neutral partners from the background. We minimize the ambient Hg by producing it only when needed via thermal disassociation of isotopically enriched HgO. To obtain the desired ion storage times

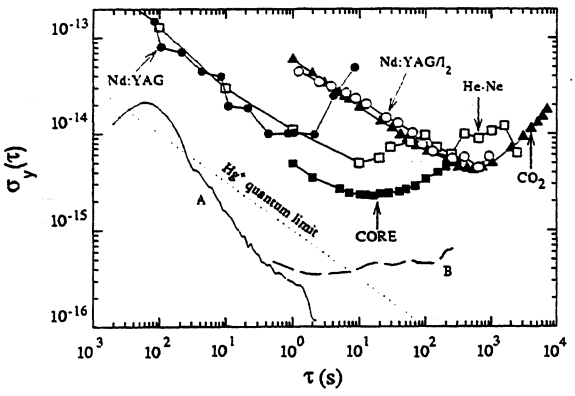


Figure 2: Allan deviation curves for stabilized lasers. We calculate $\sigma_y(\tau)$ for one of our sources from an analog-to-digital sample of the beat signal (curve A) and via a dual-mixer measurement system (curve B). The dotted line shows the quantum noise limit for a Hg^+ optical frequency standard with a single ion and $T_R = 30$ ms. Results for other stabilized lasers: Nd:YAG, Nd:YAG lasers locked to cavities [8]; Nd:YAG/I_2, iodine-stabilized Nd:YAG lasers [9]; He-Ne, methane-stabilized He-Ne lasers [10]; CO_2, CO_2 lasers locked to OsO_4 [11] (see comparable results in Ref. [12]); CORE, Nd:YAG lasers locked to cryogenic resonator oscillators [13].

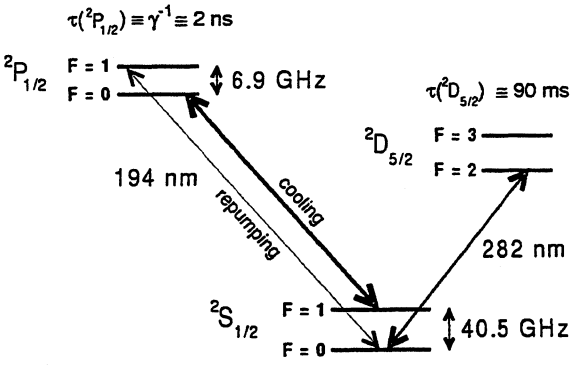


Figure 3: Simplified energy-level scheme for $^{199}Hg^+$. 194 nm radiation is used for Doppler cooling of the ion's motion; the long-lived optical clock transition is at 282 nm. The 40.5 GHz microwave clock transition is also shown.

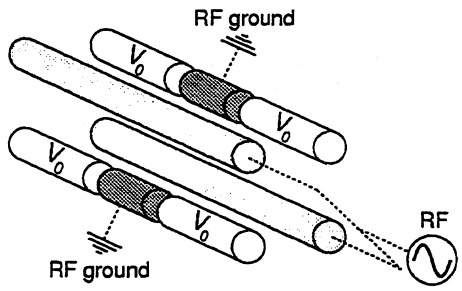


Figure 4: Basic geometry of the linear ion trap. Radial confinement is provided by a ponderomotive potential resulting from the RF voltage (several hundred volts at 10 MHz) applied to two opposing rods, with the complementary pair at RF ground. Trapping along the axis is provided by a DC potential V_0 (0-200 V) applied to the outer electrode segments.

(days), it is necessary to further refrigerate the trap enclosure to liquid He temperatures, so that any remaining free Hg is sequestered by cryopumping on the vessel walls. Cryogenic operation, however, introduces a host of other challenges, most prominently the accumulation and "freezing out" of patches of charge on the trap electrodes. The resulting stray electric fields can add enough additional bias to prevent the trapping of ions. Therefore, heaters must be incorporated into the trap structure so that its temperature may be elevated enough to permit the dissipation of any charge accumulated during the time the Hg oven and ionizing electron beam are activated. After loading, the trap is returned to cryogenic operation without further evidence of fluctuating bias fields.

We are experimenting with a number of heated trap structures, all variations of the linear geometry shown schematically in Figure 4. The operating principles of the linear Paul trap and laser cooling are described in detail elsewhere [16], but a few comments relevant to the present discussion are in order. Ions stored in radiofrequency traps exhibit a driven micromotion at the trap drive frequency if they are not confined to nodal points of the radiofrequency trapping field. Even small amounts of micromotion can lead to parametric heating of the secular motion of the trapped ions. For a single ion, it is possible to minimize this micromotion with the application of static bias fields which move the ion to the minimum of the ponderomotive potential. The amplitude of micromotion can be directly measured by correlating the phase of the ion scattered-light signal (which is modulated by the Doppler-shifted atomic absorption) with that of the trap drive [17]. By observing the correlation and applying static voltages to compensate for the existing bias, the driven motion can

be reduced to negligible levels.

The transition strength of the optical carrier is sensitively dependent on the vibration amplitude, and uncontrolled fluctuations in transition strength may result if the mean excitation number of the harmonic motion is large or changing. This effect can be overcome by strongly confining the ion so that its maximum excursions are much less than a wavelength of the 282 nm interrogating radiation, the Lamb-Dicke regime. The desire for rigid confinement requires that linear traps capable of supporting large electric potentials ($\approx$1-2 kV) with electrode dimensions and spacings $\lesssim$1 mm be realized in the laboratory. Presently, we have trapped and cooled ions to crystallization in one geometry which is capable of Lamb-Dicke confinement at secular frequencies of $1-2$ MHz during cryogenic operation, and have constructed a smaller, tighter trap which should yield much stronger confinement for identical applied voltages. In the future, we may move to micro-miniature lithographic traps similar to those used in our group at the National Institute of Standards and Technology in Boulder for quantum-state engineering investigations, modified for cryogenic operation.

Another difficulty specific to the Hg^+ experiments is the necessity of a reliable and economical source of narrow-band 194 nm radiation for laser Doppler cooling. Currently, the cooling light is generated using a single-mode Ar^+ laser oscillating at 515 nm that is frequency-doubled in β-barium borate (BBO) to 257 nm. The second-harmonic radiation is then sum-frequency mixed in BBO with the output of diode sources at 792 nm to produce light at 194 nm. We have constructed a Yb:YAG laser at 1.03 μm that is frequency-doubled in $KNbO_3$ to obtain over 400 mW of power at 515 nm, which is sufficient to replace the Ar^+ laser. Because the harmonic generation efficiency is limited by blue-light-induced infrared absorption at high powers in $KNbO_3$, we anticipate even better conversion efficiency by frequency-doubling with lithium triborate in the future. We have also constructed a Nd:FAP laser to replace the 563 nm dye laser and its Ar^+ pump. The inherent frequency stability and low maintenance of these all-solid-state systems make them very attractive for metrological applications and high-precision spectroscopy.

A generic impediment to the realization of any optical frequency standard is the delivery of the phase-stable radiation from the local oscillator to the atom. We transport the 563 nm light via optical fiber from the suspended high-finesse cavity table to the optical table supporting the trap cryostat. Doppler shifts due to the relative motion of the two tables, as well as vibration and insertion noise from the fiber itself, contaminate the phase of the interrogating radiation. This additive noise is sensed by heterodyning light on the cavity table with light which has made a round-trip to the trap table. Phase-locking this beat signal to a stable RF source actively reduces the frequency noise acquired during the transport of the light from the cavity to the ion trap [4, 18]; hence the phase purity of the cavity-stabilized laser is restored at the site of ion. The 563 nm light is frequency-doubled in deuterated ammonium dihydrogen phosphate to near resonance with the ion at 282 nm near the trap. Measurement or locking of the 282 nm light to the Hg^+ transition begins with laser cooling of the ion to near the Doppler limit of 1.7 mK with a pair of copropagating 194 nm beams: a strong (several hundred microWatts) cooling beam, and a weaker ($\approx$1 μW) repumping beam which prevents the ion from becoming trapped in the $^2S_{1/2}$ $F = 0$ level. Turning off the repumping radiation prepares the ion in the lower ($F = 0$) hyperfine level of the ground state. The clock transition is then probed with the time-domain Ramsey method in the absence of any 194 nm radiation. Finally, transitions to the $^2D_{5/2}$ level are detected using optical-optical double resonance ("electron shelving") [19], in which we detect light scattered from the strong cycling $^2S_{1/2}-^2P_{1/2}$ electric-dipole transition at 194 nm. If the atom is in the $^2D_{5/2}$ state, no scattering above background is observed. If the atom remains in the ground state, we detect strong scattering as the ion is driven up to the $^2P_{1/2}$ state and decays spontaneously back to the ground state. For a single ion starting in the ground state, a peak fluorescence count rate of tens of kiloHertz has been observed using a photomultiplier. This quantum amplification enables state detection with nearly unit efficiency. A digital integral servo loop will adjust the average frequency of the 282 nm radiation by means of an AOM to match the center of the detected atomic resonance. The values of the steered frequency will be recorded after each measurement cycle for evaluation [4, 20].

4 PERFORMANCE

Single laser-cooled ions stored in traps offer significant advantages over other atomic references in terms of relative immunity from systematic frequency shifts. For $^{199}Hg^+$ the fractional magnitude of the second order Doppler shift is approximately 2×10^{-18} at the Doppler cooling limit. We have chosen an isotope with nonzero nuclear spin which makes available transitions which are first-order independent of magnetic field strength at zero field. Hence the first-order Zeeman shift can be made very small. In our trap, collisional and blackbody radiation shifts are minimized by the cryogenic operation required for long storage times. Although light shifts from the strong cooling radiation can broaden

the narrow reference transition, we shutter the cooling beams during the probe cycle to prevent this from occurring. In probing the $^2S_{1/2}$–$^2D_{5/2}$ resonance, AC Stark effects are negligible as we rely on a single-photon quadrupole transition. Because we use an optical transition to a state with a nonvanishing electronic quadrupole moment, the limiting accuracy of our standard may arise from the uncertainty in the interaction of the $^2D_{5/2}$ atomic moment with the electric fields of the trap. While the quadrupole shift from the AC trapping fields can be precisely calibrated since they affect the motion of the ion in a known way, static fields or patches of charge on the trap electrodes may be difficult to control. Shifts larger than 1 Hz may be expected. We can significantly reduce the uncertainty in these shifts by measuring the quadrupole transition frequencies for three mutually orthogonal orientations of a quantizing magnetic field, in which case the quadrupole shift averages to zero.

From Eqn. (1) with $T_R = 0.03$ s, we estimate a projection noise-limited measurement stability for a single ion of approximately 1 part in 10^{15}, or 1 Hz, for averaging times of 1 s. We have demonstrated an optical local oscillator with a linewidth of less than 0.6 Hz for averaging times up to 32 s, which is sufficient for such an Hg^+ optical standard. We are hopeful that the system we are assembling, in combination with a frequency synthesis chain connecting the optical transition to microwave frequencies [21], might eventually provide a time standard with an accuracy near 10^{-18}, and stability surpassing the best present-day clocks.

The authors acknowledge the support of the U. S. Office of Naval Research and the National Institute of Standards and Technology.

References

[1] L. A. Rawley, J. H. Taylor, M. M. Davis, and D. W. Allan, Science **238**, 761 (1987).

[2] J. W. Armstrong, F. B. Estabrook, and H. D. Wahlquist, Astrophys. J. **318**, 536 (1987).

[3] *Special Issue on Time and Frequency*, Proc. IEEE **79** (1991).

[4] J. C. Bergquist, W. M. Itano, and D. J. Wineland, in *Frontiers in Laser Spectroscopy, Proceedings of the International School of Physics "Enrico Fermi" Course 120*, edited by T. W. Hänsch and M. Inguscio (North-Holland, Amsterdam, 1994), pp. 359-376.

[5] R. W. P. Drever, J. L. Hall, F. V. Kowalski, J. Hough, G. M. Ford, A. J. Munley, and H. Ward, Appl. Phys. B **31**, 97 (1983).

[6] Commercial products are identified in order to adequately specify the experimental procedure. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology.

[7] B. C. Young, F. C. Cruz, W. M. Itano, and J. C. Bergquist, Phys. Rev. Lett., in press.

[8] N. M. Sampas, E. K. Gustafson, and R. L. Byer, Opt. Lett. **18**, 947 (1993).

[9] J. L. Hall, L. -S. Ma, S. Swartz, P. Junger, and S. Waltman, in *1998 Conference on Precision Electromagnetic Measurements Digest*, edited by T. L. Nelson (IEEE, New York, 1998), pp. 151-152.

[10] S. N. Bagayev, A. K. Dmitriyev, P. V. Pokasov, and B. N. Skvortsov, in Ref. [2] pp. 289-296.

[11] O. Acef, Opt. Comm. **134**, 479 (1997).

[12] V. Bernard, C. Daussy, G. Nogues, L. Constantin, P. E. Durand, A. Amy-Klein, A. Van Lerberghe, and C. Chardonnet, IEEE J. Quant. Electr. **33**, 1282 (1997).

[13] S. Seel, R. Storz, G. Ruoso, J. Mlynek, and S. Schiller, Phys. Rev. Lett. **80**, 2089 (1998).

[14] D. J. Berkeland, J. D. Miller, J. C. Bergquist, W. M. Itano, and D. J. Wineland, Phys. Rev. Lett. **80**, 2089 (1998).

[15] L. L. Tjoelker, J. D. Prestage, and L. Maleki, in *Proceedings of the Fifth Symposium on Frequency Standards and Metrology*, edited by J. C. Bergquist (World Scientific, Singapore, 1996), pp. 33-38.

[16] J. C. Bergquist, W. M. Itano, and D. J. Wineland, Phys. Rev. A **36**, 428 (1987) and references therein.

[17] D. J. Berkeland, J. D. Miller, J. C. Bergquist, W. M. Itano, and D. J. Wineland, J. Appl. Phys. **83**, 5025 (1998).

[18] L. -S. Ma, P. Junger, J. Ye, and J. L. Hall, Opt. Lett. **19**, 1777 (1994).

[19] H. G. Dehmelt, IEEE Trans. Inst. Meas. **31**, 83 (1982).

[20] D. J. Wineland, W. M. Itano, J. C. Bergquist, J. J. Bollinger, S. L. Gilbert, and F. Diedrich, in *Frequency Standards and Metrology: Proceedings of the Fourth Symposium*, edited by A. De Marchi (Springer-Verlag, Berlin, 1989), pp. 71-77.

[21] B. Frech, J. S. Wells, C. W. Oates, J. Mitchell, Y-P. Lan, T. Kurosu, L. Hollberg, B. Young, and J. C. Bergquist, this volume.

TOWARDS AN INDIUM SINGLE-ION OPTICAL CLOCK

E. Peik, J. Abel, Th. Becker, M. Fries, J. von Zanthier, H. Walther

Max-Planck-Institut für Quantenoptik and Sektion Physik der Ludwig-Maximilians-Universität,
Hans-Kopfermann-Str. 1, 85748 Garching, Germany
e-mail: peik@mpq.mpg.de

ABSTRACT

We are investigating the $5s^2\,{}^1S_0 \rightarrow 5s5p\,{}^3P_0$ transition of a single trapped laser-cooled $^{115}In^+$ ion as a possible optical frequency standard. This line with a natural linewidth of only 1.1 Hz, corresponding to a Q factor of $1.2 \cdot 10^{15}$, is highly immune to systematic frequency shifts due to external perturbations. For sideband laser cooling and fluorescence detection of the indium ion the $5s^2\,{}^1S_0 \rightarrow 5s5p\,{}^3P_1$ transition at 230.6 nm is excited. Temperatures below 100 μK and a mean vibrational quantum number $\langle n \rangle < 1$ of the ion in the trap have been reached. For the clock transition a resolution of $3 \cdot 10^{-13}$ (linewidth 350 Hz) has been obtained so far, limited by the laser linewidth. The absolute frequency of the ${}^1S_0 \rightarrow {}^3P_0$ transition was measured by making a link to known reference frequencies of molecular iodine and methane using a frequency chain.

1. INTRODUCTION

A single laser-cooled ion in a radiofrequency trap represents a good approximation to the spectroscopic ideal of a motionless point-like absorber in a perturbation-free environment. It is consequently regarded as a perfect system for an optical atomic clock of very high accuracy [1]. With a frequency-stable laser locked to a narrow absorption resonance of a trapped ion a relative stability $\sigma_y(1\,\mathrm{s}) = 10^{-15}$ and an accuracy of a few times 10^{-18} seem possible. High-resolution spectroscopy of forbidden optical transitions has yielded sub-kHz linewidths in Hg^+ [2], Ba^+ [3] and, more recently, in Sr^+ [4], Yb^+ [5] and In^+. In the first four alkali-like ions, quadrupole transitions between S and D states are investigated. We are studying the earth-alkali-like spectrum of In^+, where the lowest two levels $5s^2\,{}^1S_0$ and $5s5p\,{}^3P_0$ are connected by a hyperfine-induced electric dipole transition at a wavelength of 236.5 nm which has a natural linewidth of only 1.1 Hz [6]. Figure 1 shows the relevant energy levels of In^+.

As a candidate for a primary optical frequency standard In^+ has three main advantages: (i) Since the electronic angular momenta in both states of the clock transition ${}^1S_0 \rightarrow {}^3P_0$ vanish, the shifts of the transition frequency due to external electromagnetic fields are very small. In fact, the ${}^3P_0$ state is not a pure

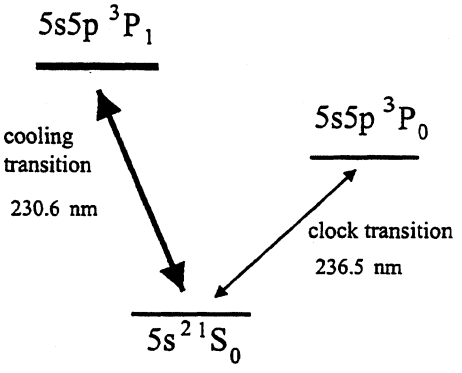


Figure 1: The lowest energy levels of the In^+ ion.

$J = 0$ state but contains small hyperfine admixtures of the $5s5p$ levels ${}^3P_1$ and ${}^1P_1$. These perturbations are responsible for the non-vanishing electric dipole moment between this state and the ground state [6]. For the same reason there is a small difference between the Landé g-factors of ${}^1S_0$ and ${}^3P_0$. This leads to a linear Zeeman shift of the clock transition which we measured to be 223 Hz/G for the $m_F = \pm 1/2 \rightarrow m_F = \pm 1/2$ Zeeman components. Shielding of magnetic fields to the 10 μG level will ensure the predicted 10^{-18} reproducibility of the frequency standard. Electric fields lead to a quadratic Stark shift of the transition frequency. In positive ions the static polarisabilities are generally smaller than in neutral atoms because the remaining valence electrons are more tightly bound. For the cadmium atom, which is isoelectronic to In^+, the ground state polarisability is $8 \cdot 10^{-36}$ J/(V/cm)^2. We estimate the quadratic Stark shift of the In^+ clock transition to be below 1 mHz/(V/cm)^2 (a shift of 1 mHz corresponds to a relative change of the transition frequency of $8 \cdot 10^{-19}$). The Stark shift induced by the trap-field is proportional to the mean quadratic distance of the ion from the trap center and consequently proportional to the temperature of the ion. The shift is smaller than 100 mHz/K for typical trap conditions and therefore negligible at the sub-millikelvin temperatures achievable with laser cooling. The averaged quadratic electric field strength of the black-body radiation emitted by the vacuum chamber is given by $\langle E^2 \rangle = 69.2\,(\mathrm{V/cm})^2\,(T/300\ \mathrm{K})^4$. At room temperature the spectrum of black-body radiation has its peak spectral density at a wavelength of 9 μm, whereas both,

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

the ground state and the metastable $^3P_0$ level of In^+ have strong electric-dipole-allowed transitions to the next-higher-lying states only at VUV wavelengths (159 nm and 194 nm, respectively). Consequently, the influence of the thermal radiation can be estimated by using the static polarisabilities. At a trap temperature of 300 K control of the temperature to ±1 K will be found sufficient to reduce the uncertainty in the blackbody shift below 1 mHz. The electric field gradient of the trap has no influence since both levels of the clock transition have vanishing quadrupole moment.

(ii) To excite the clock transition a reliable and frequency-stable laser source is required. The In^+ transition has the advantage that it coincides with the fourth harmonic of the 946 nm laser line of Nd:YAG. The availability of this diode-laser-pumped solid-state laser is of significance for long-term operation of the optical frequency standard. The measurements described here are still done with a conventional setup made from discrete optical elements [7,8], but in the future it is conceivable to use a monolithic ring-laser [9,10] offering high intrinsic stability and compact design. (iii) Laser cooling of In^+ is performed by using the narrow intercombination line $5s^2\ ^1S_0 \rightarrow 5s5p\ ^3P_1$. With a natural linewidth of 360 kHz, the photon scattering rate on this transition is sufficiently high to detect a single ion via the resonance fluorescence. At the same time the line is narrow enough to allow resolution of the ≈ 1 MHz vibrational frequencies of the ion in the trap. In this parameter regime, sideband cooling [1,11,12] allows the ion to be cooled to the motional ground state of the trap and to reach temperatures below 100 μK. This reduces the fractional second-order Doppler shift to values below 10^{-19}.

2. SIDEBAND COOLING OF THE INDIUM ION

The ion is trapped in a quadrupole radiofrequency trap that is a geometrical variant of the original Paul trap and basically consists of only a ring electrode (Paul-Straubel trap [13,14]). This type of trap is relatively simple to fabricate in miniature size (1 mm diameter of the ring), making it easy to confine the ion to a region in space that is smaller than the optical wavelength (Lamb-Dicke regime). It is also a geometrically quite open structure that allows good optical access to the trapped ion. The trap is driven with a RF field at 10 MHz; the oscillation frequencies in the time-averaged pseudopotential are 1.4 MHz in the axial and 0.9 MHz in the radial direction.

Sideband cooling is performed through laser excitation of the $5s^2\ ^1S_0 \rightarrow 5s5p\ ^3P_1$ intercombination line at a wavelength of 230.6 nm. In this transition the hyperfine component $F \rightarrow F+1$ (where $F = I = 9/2$) is excited with circularly polarised light in a vanishing magnetic field, so that optical pumping between the Zeeman sublevels results in a closed two-level system. The laser radiation is produced by a frequency-doubled stilbene-3 ring dye laser. The laser is frequency-stabilised to a reference cavity by using a high-bandwidth electronic servo system, resulting in a laser linewidth below 10 kHz.

In order to increase the initial cooling power for a highly vibrationally excited ion and be able to record high resolution spectra of the cooling transition we use a bichromatic cooling method [15]. The laser beam is passed through an electro-optical modulator with variable modulation frequency. Through phase modulation two weak sidebands containing about 0.1% of the total power each were created around the original laser frequency. The carrier of this spectrum was far detuned (by about −40 MHz) and efficiently cooled the ion in higher vibrational levels when excited, for example, by collisions with molecules of the residual gas. The high-frequency phase modulation sideband was used to cool the ion to the quantum ground state when tuned to the first low-frequency vibrational sideband and could also be scanned over the carrier frequency by changing the drive frequency of the electro-optical modulator. An experimental excitation spectrum obtained in this way is shown in Fig. 2. The spectrum is dominated by the carrier and the excitation of vibrational sidebands is very weak, as can be seen from the magnified deviations between the data and a fitted single Lorentzian curve.

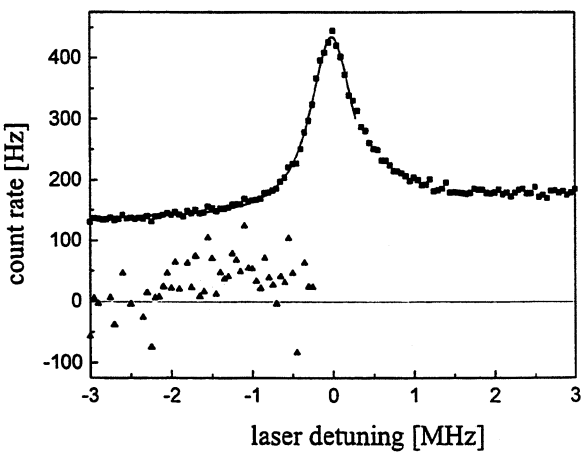


Figure 2: Excitation spectrum of the cooling transition of a single indium ion obtained with bichromatic sideband cooling (averaged over 32 scans). Below: deviations between the data points and a fitted Lorentzian curve (solid line) in tenfold magnification. From the height of the low-frequency vibrational sidebands at −0.9 MHz and −1.4 MHz detuning it is possible to determine the mean vibrational quantum number $\langle n \rangle = 0.7(3)$.

In the region of laser detunings corresponding to the low-frequency vibrational sidebands an increase of the

signal by about 2(1)% of the height of the Lorentzian is observable. In the Lamb-Dicke limit the height of the first red sideband relative to the carrier is given by $\langle n \rangle k^2 x_0^2$, where k is the wavenumber of the cooling laser light and $x_0 = \sqrt{\hbar/2m\omega}$ the extension of the ground state wavefunction (ω: vibrational frequency). We can therefore determine the mean vibrational quantum number $\langle n \rangle = 0.7(3)$, corresponding to a temperature of about 60 μK. This result indicates that the ion is in the vibrational ground state with a probability exceeding 50% for a basically unlimited time duration (several minutes in the present experiment).

3. HIGH-RESOLUTION SPECTROSCOPY OF THE CLOCK TRANSITION

Spectroscopy of the narrow $^1S_0 \rightarrow {}^3P_0$ line is performed in optical-optical double resonance using Dehmelt's idea of electron shelving [1]: Excitation of the metastable $^3P_0$ level leads to cessation of the single-ion fluorescence signal on the cooling transition until the level decays or the valence electron is brought back to the ground state by a stimulated process. This method leads to strong quantum amplification (absorption of one photon prevents subsequent scattering of some 10^5 to 10^6 photons) and allows detection of transitions to the metastable state with practically 100% efficiency.

The laser system used for excitation of the $^1S_0 \rightarrow {}^3P_0$ resonance is described in [8]. It consists of a diode-pumped Nd:YAG laser emitting at 946 nm. This laser contains all the necessary tuning elements and is frequency-stabilised to a passive resonator of high finesse. A second diode-pumped Nd:YAG laser is used for power amplification and efficient generation of the second harmonic at a wavelength of 473 nm. It is a ring laser, containing only a Nd:YAG crystal and a $KNbO_3$ crystal for frequency doubling. Infrared light from the stable master laser is coupled into this laser to transfer the frequency stability via injection locking. The blue light is coupled into an enhancement cavity to generate the UV radiation at 236.5 nm with a BBO crystal.

In order to obtain high-resolution spectra of the clock transition any light shift and line broadening by the cooling laser have to be avoided. Both laser beams are applied alternately and blocked by means of mechanical shutters. After a clock-laser pulse of 20 ms duration the cooling laser is turned on and the fluorescence photons are counted in a 40 ms time interval. If the count rate corresponds to the single-ion fluorescence level, the excitation attempt of the clock transition is regarded as unsuccessful and repeated either at the same or at a different frequency of the clock laser. If the ion is not fluorescing, the cooling laser is kept switched on to wait for the decay of the metastable state (lifetime 140 ms) and one excitation is registered. Typically, the frequency of the clock-laser

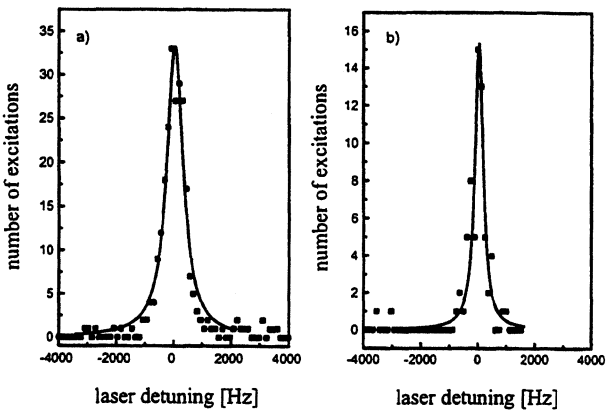


Figure 3: Excitation spectra of the $^1S_0 \rightarrow {}^3P_0$ resonance of a single indium ion obtained in optical-optical double resonance using electron shelving. a) Laser pulses of 20 ms duration and 1 μW power were used. The line is saturation-broadened to 700 Hz linewidth (FWHM). b) The laser power was reduced to 35 nW. The linewidth of 350 Hz is determined by the linewidth of the laser.

radiation was changed in steps of 8 Hz and four excitation attempts were performed at each frequency.

Two high-resolution spectra of the clock transition are shown in Fig. 3. At a laser power of 1 μW (spectrum 3a) the resonance is still saturation-broadened to 700 Hz linewidth (full width at half maximum on the ultraviolet frequency scale); at 35 nW laser power the linewidth is reduced to 350 Hz (spectrum 3b). This value corresponds to a spectroscopic resolution $\delta\nu/\nu$ of $3 \cdot 10^{-13}$. According to the above-mentioned measurement of the ion's temperature and our present experimental control of electromagnetic fields and vacuum conditions, no significant Doppler, Zeeman, Stark or collisional broadening is expected beyond the level of the natural linewidth of 1.1 Hz. The smallest linewidth that we observe experimentally is presently determined by the laser linewidth. The value of 350 Hz for this linewidth is in agreement with an independent determination of the frequency stability of the laser using a second stable optical reference cavity. It is determined mainly by residual vibrations of the high-finesse cavity that is used for the frequency stabilization. Improvements of the isolation of this cavity will probably allow us to resolve the natural linewidth of the ion, leading to a resolution of $8 \cdot 10^{-16}$.

Realisation of an optical clock requires first of all a narrow and stable resonance – as is the case for a trapped indium ion – and, secondly, precise frequency determination in comparison with the cesium clock. We measured the absolute frequency of the $^{115}In^+$ $5s^2\ {}^1S_0 \rightarrow 5s5p\ {}^3P_0$ clock transition with an accuracy of 3 parts in 10^{11} [16]. For this measurement a frequency synthesis chain was used which links the in-

dium transition to a methane stabilised He-Ne laser at 3.39 μm [17] and a Nd:YAG laser at 1064 nm whose second harmonic was locked to a hyperfine component of molecular iodine [18]. The frequency of the 946 nm Nd:YAG laser, whose fourth harmonic drives the In^+ transition, was set to the arithmetic mean of the frequencies of the 1064 nm Nd:YAG laser and a diode laser at 852 nm by means of an optical frequency interval divider [19]. The frequency of the diode laser was compared with another diode laser at 848 nm that was phase-locked to the fourth harmonic of the He-Ne laser at 3.39 μm. A frequency gap of 1.43 THz between the two diode lasers was bridged with the help of an optical frequency comb generator [20]. The frequency of the $^{115}In^+$ clock transition was determined to be 1 267 402 452 914 (41) kHz, where the accuracy is limited by the uncertainty of the iodine reference.

Parts of this frequency chain were already used in a previous measurement to determine the $1S \rightarrow 2S$ transition frequency in atomic hydrogen with an accuracy of 3 parts in 10^{13} [21]. An interesting prospect afforded by this work is the possibility of linking the two narrow optical resonances of atomic hydrogen and the indium ion and using the high precision of this frequency comparison in a search for temporal variations of the fine-structure constant (see Refs. [22,23]).

The combination of the very small systematic uncertainties in the transition frequency of the trapped indium ion (see introduction) and the availability of the technically convenient Nd:YAG laser to drive this transition makes this system a most promising candidate for a future optical clock.

ACKNOWLEDGMENTS

We gratefully acknowledge the collaboration with A. Yu. Nevsky, M. N. Skvortsov, S. N. Bagayev, R. Holzwarth, J. Reichert, Th. Udem, and Th. Hänsch in the absolute optical frequency measurement.

REFERENCES

[1] H. Dehmelt, IEEE Trans. Instrum. Meas. **31**, 83 (1982).

[2] J. C. Bergquist, W. M. Itano, F. Elsner, M. G. Raizen, and D. J. Wineland, in *Light Induced Kinetic Effects on Atoms, Ions and Molecules*, Eds.: L. Moi, S. Gozzini, C. Gabbanini, E. Arimondo, and F. Strumia (ETS Editrice, Pisa, 1991).

[3] N. Yu, X. Zhao, H. Dehmelt, and W. Nagourney, Phys. Rev. A **50**, 2738 (1994).

[4] J. E. Bernard, L. Marmet, and A. Madej, Opt. Commun. **150**, 170 (1998).

[5] C. Tamm and V. Bühner, Abstract QFD5, European Quantum Electronics Conference, EQEC '98, Glasgow, 1998.

[6] E. Peik, G. Hollemann, and H. Walther, Phys. Rev. A **49**, 402 (1994).

[7] G. Hollemann, E. Peik, and H. Walther, Opt. Lett. **19**, 192 (1994).

[8] G. Hollemann, E. Peik, A. Rusch, and H. Walther, Opt. Lett. **20**, 1871 (1995).

[9] Th. J. Kane and R. L. Byer, Opt. Lett. **10**, 65 (1985).

[10] I. Freitag, R. Henking, A. Tünnermann, and H. Welling, Opt. Lett. **20**, 2499 (1995).

[11] D. J. Wineland and W. M. Itano, Phys. Rev. A **20**, 1521 (1979).

[12] F. Diedrich, J. C. Bergquist, W. M. Itano, and D. J.Wineland, Phys. Rev. Lett. **62**, 403 (1989).

[13] N. Yu, W. Nagourney, and H. Dehmelt, J. Appl. Phys **69**, 3779 (1991).

[14] C. Schrama, E. Peik, W. W. Smith, and H. Walther, Opt. Commun. **101**, 32 (1993).

[15] E. Peik, J. Abel, Th. Becker, J. v. Zanthier, and H. Walther, Phys. Rev. A., in press.

[16] J. v. Zanthier, J. Abel, Th. Becker, M. Fries, E. Peik, H. Walther, R. Holzwarth, J. Reichert, Th. Udem, T. W. Hänsch, A. Yu. Nevsky, M. N. Skvortsov, and S. N. Bagayev, submitted to Opt. Commun.

[17] S. N. Bagayev, A. K. Dmitriyev, and P. V. Pokasov, Laser Phys. **7**, 989 (1997).

[18] P. A. Jungner, S. Swartz, M. Eickhoff, Jun Ye, J. L. Hall, and S. Waltman, IEEE Trans. Instrum. Meas. **44**, 151 (1995).

[19] H. Telle, D. Meschede, and T. W. Hänsch, Opt. Lett. **15**, 532 (1990).

[20] M. Kourogi, T. Enami, and M. Ohtsu, IEEE Photon. Techn. Lett. **6**, 214 (1994).

[21] T. Udem, A. Huber, B. Gross, J. Reichert, M. Prevedelli, M. Weitz, and T. W. Hänsch, Phys. Rev. Lett. **79**, 2646 (1997).

[22] J. D. Prestage, R. L. Tjoelker, L. Maleki, Phys. Rev. Lett. **74**, 3511 (1995).

[23] V. A. Dzuba, V. V. Flambaum, and J. K. Webb, Phys. Rev. Lett. **82**, 888 (1999).

DEVELOPMENT OF AN OPTICAL FREQUENCY STANDARD BASED UPON THE $^2S_{1/2}$-$^2D_{5/2}$ TRANSITION IN Sr^+

GP Barwood, P Gill, G Huang and HA Klein
Centre for Basic, Thermal and Length Metrology
National Physical Laboratory, Queens Road, Teddington, UK TW11 0LW

ABSTRACT

Recent work is presented on the development of an optical frequency standard, based on the $^2S_{1/2}$-$^2D_{5/2}$ transition at 674 nm in $^{88}Sr^+$. Work on characterising the probe laser is also presented and the observed linewidth of one of the $\Delta m=0$ Zeeman components of the 674 nm transition is 290 Hz. Two traps have been operated simultaneously and a null offset between them confirmed at the 3 parts in 10^{13} level. The feasibility of cooling and probing $^{87}Sr^+$ is also discussed.

1. INTRODUCTION

In recent years, there has been significant progress towards the use of narrow optical transitions in cold trapped ions as optical frequency standards. A number of candidate ions are being studied worldwide and, at the UK National Physical Laboratory (NPL), both ytterbium and strontium are being studied. The strontium ion is confined in an RF Paul trap, of radius ≈0.5 mm. A partial term scheme for $^{88}Sr^+$ is shown in figure 1. The ion is cooled using the transition at 422 nm, using a frequency doubled diode laser [1-3]. An additional laser at 1092 nm, is required to prevent optical pumping into the $^2D_{3/2}$ state. Until recently, this source was a Nd^{3+}-doped fibre laser, optically pumped by a diode laser at 825 nm. However, this has now been replaced by a specially-manufactured 50-mW distributed Bragg reflector (DBR) laser.

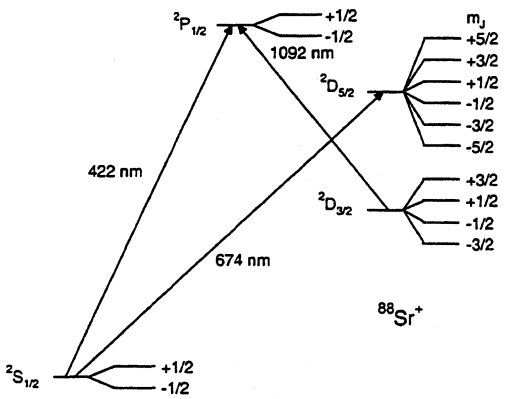


Figure 1: Partial term scheme for strontium

Beat frequency measurements between two similar 1092 nm DBR diodes indicate that the laser linewidth is ≈5 MHz full width at half maximum (FWHM). The laser is drive current modulated at 0.7 MHz over a frequency range of 0.7 GHz peak to peak. This frequency modulation allows the transition to be more easily located with a commercial wavemeter and relaxes the diode laser heatsink temperature stability required to maintain the ion fluorescence.

The optical frequency standard is provided by the quadrupole $^2S_{1/2}$-$^2D_{5/2}$ transition at 674 nm, driven by a AlGaInP diode laser. An overall schematic of the laser system is shown in figure 2. The transition frequency has been measured both at NPL [4] and at the National Research Council (NRC) in Canada [5-7]. The transition has been recommended by the Comité International des Poids et Mesures as a new radiation for the realisation of the metre [8]. It is the first ion-trap based optical frequency standard to be so included.

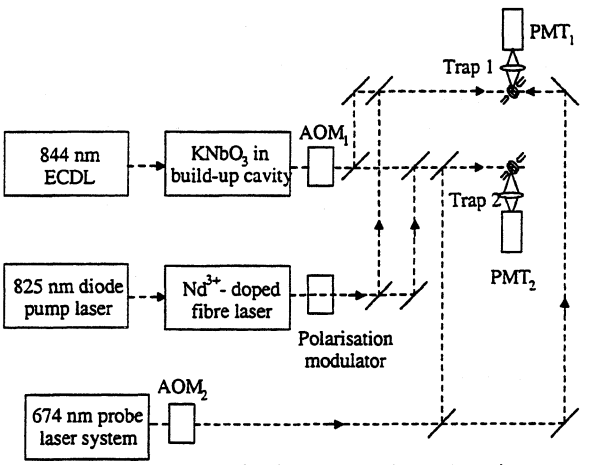


Figure 2: Cooling and probe laser schematic

The lifetime of the $^2D_{5/2}$ state has been measured to be (347 ± 33) ms [9-11], corresponding to a natural linewidth of 0.4 Hz. In $^{88}Sr^+$, which is the most abundant isotope, this transition splits into up to ten Zeeman components. The DC magnetic field is controlled through the use of three pairs of coils mounted around the trap. In order to determine the centre frequency of this transition, it is desirable to operate in a small magnetic field, for example less than 1 μT, so that the shifts of individual components are not too far away from line centre. However, in these low magnetic fields, the fluorescence falls to zero, since the 1092 nm laser is then being driven coherently from a manifold of unresolved Zeeman levels. In general, for any fixed polarisation state of the 1092 nm laser, there are two independent dark states into which the ion becomes optically pumped [12]. This problem

0-7803-5400-1/99/ $ 10.00

is resolved by rapid modulation of the 1092 nm laser polarisation [1, 13]. The ion fluorescence then provides a convenient diagnostic for minimisation of the magnetic field to below ≈2 μT, when the 1092 nm polarisation is static.

The ion is confined to the Lamb-Dicke region and interrogated using a computer-controlled cycle [2], where the ion is alternately cooled and then probed with the lasers being switched via acousto-optic modulators. The number of jumps to the $^2D_{5/2}$ state are then counted at each probe laser frequency for a particular number of interrogation cycles. Individual Zeeman components are broadened either through the effect of ambient AC magnetic field at the mains frequency (50 Hz) or are limited in linewidth by the 674 nm probe laser. In order to reduce the ambient AC magnetic field at the trap centre, the trap is housed inside a mu-metal shield. The contribution to the linewidth due to AC magnetic field of the $\Delta m = 0$ components has an upper limit of 140 Hz. This may be inferred from the maximum observed 1 kHz linewidth of the outermost $\Delta m = \pm 2$ components, since the magnetic field sensitivity of these components is 7 times that of the $\Delta m = 0$ lines. Provided the residual AC magnetic fields are less than a few nT, interrogation with a sufficiently narrow probe laser should produce spectrum with a dominant carrier and only small sidebands at the mains frequency.

2. PROBE LASER SYSTEM

The probe laser at 674 nm is an extended cavity diode laser (ECDL) which is sideband injection locked [14, 15] to a master laser. The master laser is an optically narrowed [16] AlGaInP laser locked to a highly-stable, but non-tunable high-finesse etalon.

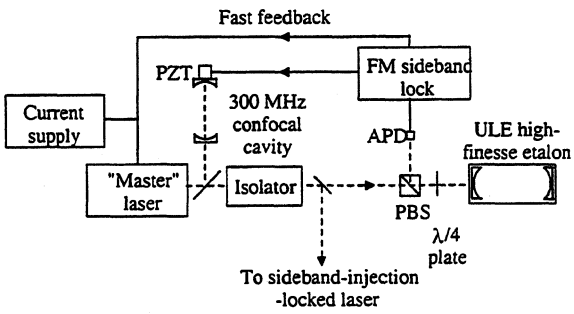


Figure 3: Master laser schematic

This is shown in figure 3 and has previously been described in a number of publications [eg 4]. The cavity has a free spectral range of ≈1500 MHz and a measured finesse of ≈2 x 10^5. The sideband injection locking of the ECDL provides precise offset locking from the master laser and also transfers the short-term frequency stability of the master laser to the slave system. The master laser linewidth is determined partly by the electronic lock to the cavity and also the degree of vibration isolation and short-term stability of the cavity itself. The master laser is shown in figure 3 and comprises a diode laser pre-stabilised through resonant optical feedback [16] and locked to a ULE cavity via a Pound-Drever lock. For this lock, the laser is drive-current modulated at 2.2 MHz. In order to determine the ability of the electronics to lock to the ULE cavity, two lasers were built and locked to adjacent longitudinal modes of the cavity. The beat frequency of 1500 MHz was then monitored on an RF spectrum analyser. As may be seen from the result in figure 4, the beat linewidth is encouragingly narrow. The fast phase noise of the diode laser is reduced through the use of resonant optical feedback and the remaining frequency noise, which is 1/f noise dominated [17]. This noise is therefore mainly at lower frequencies, and is reduced substantially through the electronic lock, which only requires a bandwidth of less than 100 kHz.

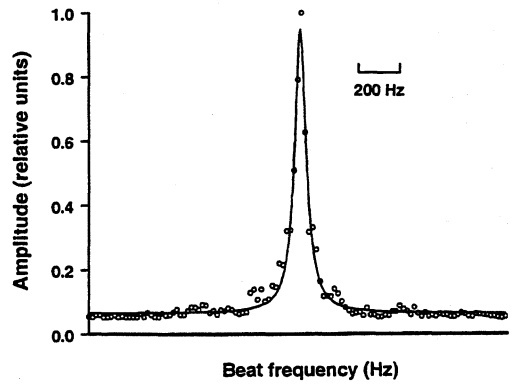


Figure 4: Beat frequency of two lasers locked to adjacent modes of the same cavity.

The data of figure 4 were observed using an analyser with an instrumental resolution of 30 Hz and fitted to a Lorenztian of FWHM 65 Hz. Assuming that both lasers contribute equally to this beat width, this implies that the electronic servo-control is currently able to produce laser linewidths of less than 30 Hz.

Whilst the combined optical and electronic lock produces encouraging results, the actual laser linewidth will also be determined by the cavity itself. A second cavity was assembled and housed in a temperature controlled and evacuated housing similar to that of the first cavity. Both cavities were suspended inside their respective housings. The beat frequency was down-converted in a double balanced mixer to around 50 kHz. This output was then sent to a frequency-to-voltage converter. The Fourier components were

determined with a low-frequency spectrum analyser and the vertical scale converted to a peak to peak frequency jitter. The result, for a laser linewidth of ≈1 kHz, is shown in figure 5. This shows that the noise spectrum predominantly consists of discrete frequencies at or below 100 Hz. These frequencies may also be observed on the output of an accelerometer, configured to monitor displacement and placed close to the cavity.

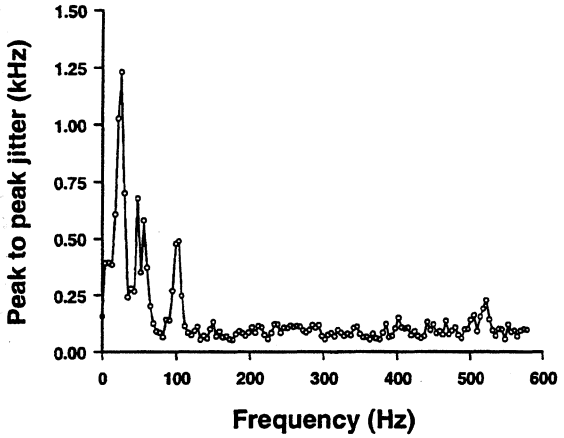


Figure 5 Noise spectrum of the beat frequency between two lasers locked to two independent cavities

Frequency variations of the probe laser over the course of a few minutes are more likely to originate from temperature variations of the cavity. The cavity is temperature controlled to a level of a few mK, but the effect of this on the frequency will depend on the thermal expansion coefficient, α. For the particular ULE spacer used, the thermal expansion coefficient is zero near 18°C although the cavity is generally held at 22.5°C, where $\alpha \approx -5.6$ MHz/K.

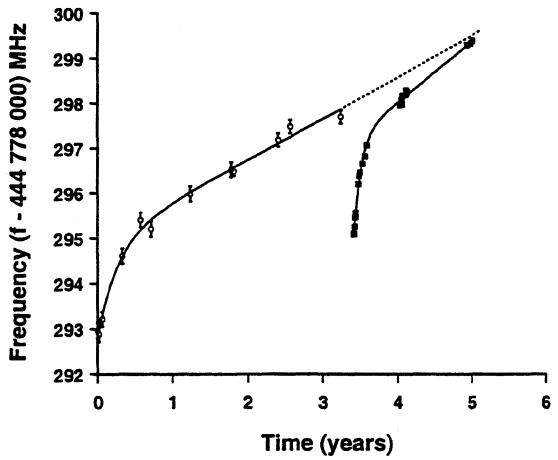


Figure 6: Isothermal drift of the ULE cavity over the course of 5 years. The discontinuity after 3.3 years is due to the cavity being exposed to air during re-mounting.

The observed stability of ±20 kHz over the course of a day therefore corresponds to a temperature variation of ≈±4 mK. Longer-term isothermal frequency drift of the cavity is also observed and the variation of frequency of one mode of this particular cavity is shown in figure 6. The long term drift is shown fitted to a formula of the form:

$$f = f_0 + at + b.\exp(-t/\tau)$$

There is a discontinuity after 3.3 years, when the cavity was exposed to air for a short period to install a suspension system inside the vacuum housing. It is interesting to note that the cavity length has now almost returned to that which might have been expected had the cavity not been exposed to air. This extrapolation is shown as a dotted line in figure 6, and represents an average drift of ≈ 2.5 kHz/day

3. COMPARISON BETWEEN TWO TRAPS

Two traps have been assembled and two ions separately trapped and cooled, in order to provide an initial determination of the reproducibility of this ion-trap based optical frequency standard. The two traps had a nominally identical design.

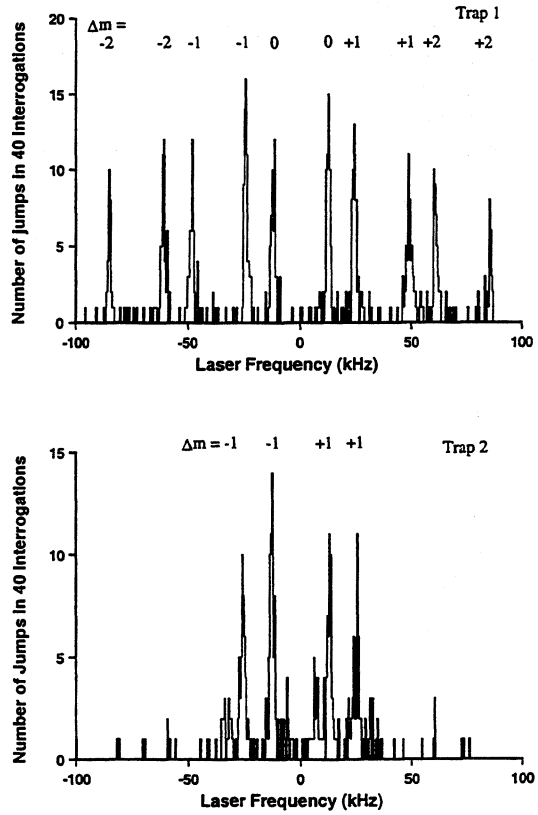


Figure 7: Scan over 674 nm transition with two traps

For this initial determination, the two cooling laser beams at 422 nm and 1092 nm were split for the two traps. The same linearly polarised source at 674 nm was also used and switched with a common acousto-optic modulator. The number of quantum jumps at each probe laser frequency was then recorded for the two traps. Pairs of Zeeman components were used to estimate the line centre for the two traps and a limit to the offset between them was then established. A typical scan using two traps is shown in figure 7 and a fuller description of the comparison is given in [18]. The scan of figure 7 is over 200 kHz, with a step resolution of 500 Hz. Prior to taking this scan, the magnetic field was nulled in the two traps. Initial reduction of the magnetic field, to less than 2 μT was made by using the fluorescence null observed when the 1092 nm laser had a static polarisation state. Successive 674 nm scans were then used to reduce the fields further in the three directions. For trap 1, the residual magnetic field was not parallel to the direction of the laser beams, nor parallel or orthogonal to the 674 nm polarisation allowing observation of all ten Zeeman components. For trap 2, the residual magnetic field was along the direction of the red laser propagation to enhance only the $\Delta m = \pm 1$ components. The average difference between the centre frequencies of the single ion 674 nm transition frequency measured in two traps for eight simultaneous pairs of scans was 120(90) Hz. This mean was calculated according to the estimated standard uncertainties from the individual scans. Thus, at this level of uncertainty, there is no statistically significant difference between the centre frequencies of the 674 nm $^{88}Sr^{+}$ $^{2}S_{1/2}$-$^{2}D_{5/2}$ transitions in the two traps. For the data of figure 7, the nominal conditions in the two traps were identical, except for the magnetic field.

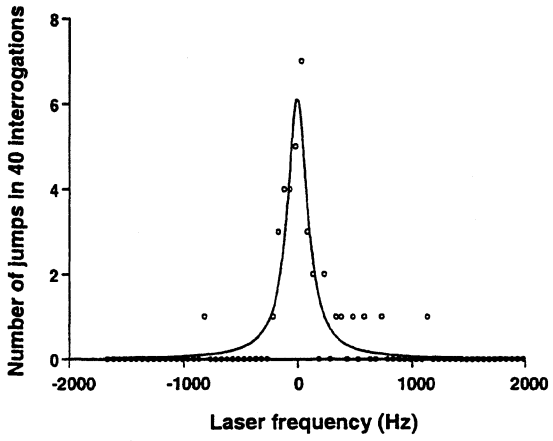


Figure 8: High resolution scan over the $\Delta m=0$ ($^{2}S_{1/2}$ (m_J=-½) $\rightarrow$$^{2}D_{5/2}$ (m_J=-½)) component, with a 50 Hz step resolution.

The micro-motion in the two traps was minimised in the direction along the lasers before the scans were taken, using RF-photon correlation techniques [19]. A further study is planned, with a more systematic variation in the trap operation parameters, including for example, the level of micro-motion, the magnitude of the RF trapping voltages and magnetic field.

Higher resolution scans over particular Zeeman components of the 674 nm transition were made in order to determine an upper limit on the laser linewidth. From 12 scans over the $^{2}S_{1/2}$ (m_J=-½) $\rightarrow$$^{2}D_{5/2}$ (m_J=-½) transition, with a 50 Hz step resolution, it was determined that the average observed linewidth was ≈290(40) Hz (FWHM), where a standard uncertainty is added in brackets. The observed width depended on whether the scan direction was in the same direction as the cavity drift, or in the opposite sense. A typical high resolution scan is shown in figure 8 and is shown fitted to a Lorenztian curve of FWHM 210 Hz. This improvement over the previously-reported probe laser linewidth at NPL from 1 kHz [1, 15] has been achieved mainly by suspending the optical breadboard mounting the cavity from 0.6 m above the breadboard.

For an ion transition with a linear Zeeman shift, magnetic fields produce the largest frequency shifts. The $\Delta m = 0$ transitions are the least susceptible to magnetic field, and the Zeeman shift for these transitions is calculated to be $\pm\frac{eB}{10\pi m_e}$ in a magnetic field B, where e is the electronic charge and m_e the electron mass. This frequency shift corresponds to $\pm 5.6\ \mathrm{Hz.nT^{-1}}$ so that nT magnetic field stability would be required in order to achieve a frequency stability approaching one part in 10^{14}. Therefore, one needs to interrogate both the $\Delta m = 0$ transitions which are symmetrically positioned about line centre to estimate the mean frequency, as was done in figure 7. In order for this estimation to be performed satisfactorily, the magnetic field must be sufficiently stable during the course of the scan. Slow magnetic field variations, for example over timescales of a minute, could ultimately limit the standard in the even isotope.

Other effects, principally Stark effects and second order Doppler shifts, may be expected to contribute to the non-reproducibility of an ion-trap based optical frequency standard. These effects have been calculated in detail for the case of a $^{199}Hg^{+}$ trap standard [20]. Calculations suggest that these effects will contribute to shifts at the few parts in 10^{14} level or below, which is around one order of magnitude less than the 3 parts in 10^{13} levels of comparison achieved here. Second order Doppler effects will arise if the micromotion is not properly reduced in all three directions. In the present

NPL strontium traps, RF-photon correlation is used to minimise the micromotion along the direction of the cooling laser only. Transverse motion may therefore be contributing a frequency shift. As discussed in [20], a Stark shift may arise, for example, if the oven coats the electrode surfaces unevenly with metal. With a narrower-linewidth probe laser and averaging using more scans, it should be easier to investigate and quantify systematic frequency shifts.

4. COOLING AND PROBING $^{87}Sr^+$

In $^{88}Sr^+$, all ten Zeeman components of the 674 nm transition shift linearly with magnetic field. The line-centre must be derived by determining the position of pairs of symmetrically placed Zeeman components. This problem may be solved by use of the odd isotope $^{87}Sr^+$, which has a nuclear spin of I = 9/2. A partial terms scheme for $^{87}Sr^+$ is shown in figure 9.

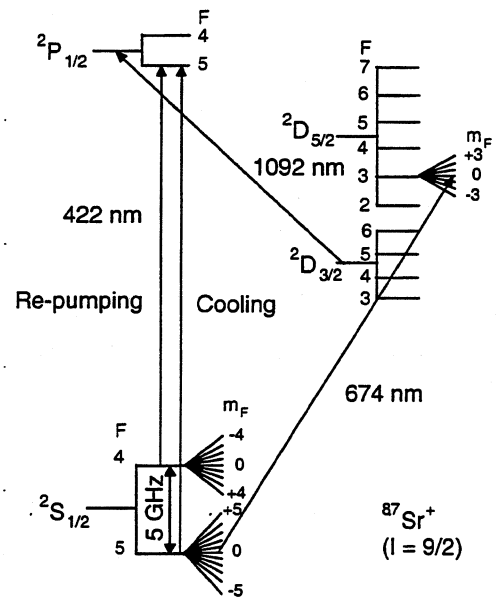


Figure 9: Partial term scheme for $^{87}Sr^+$ showing the cooling and probe lasers required

In order to cool this isotope, two lasers at 422 nm are required, separated by the ground state splitting of 5 GHz [21]. Whilst the hyperfine splitting of the $^2D_{3/2}$ state has not been measured to our knowledge, the transitions are unlikely to extend beyond the 0.7 GHz linewidth of the frequency modulated DBR laser in use for $^{88}Sr^+$. The principal advantage of using this ion is that the manifold of transitions at 674 nm contains some lines that have no linear Zeeman shift, for small magnetic fields. However, the high nuclear spin of $^{87}Sr^+$ results in optical pumping when attempting to cool and interrogate the ion. This may be solved by modulating the polarisation state of both the cooling laser, tuned to the $^2S_{1/2}$ (F=5) $\rightarrow$ $^2P_{1/2}$ (F=5) transition, and also the 1092 nm laser [1, 13]. Linear (π) polarisation of the cooling laser also allows selective optical pumping into the m_F=0 sub-levels of the $^2S_{1/2}$ state, required in order to enhance the observed line-strength of the 674-nm transitions which are independent of the first order Zeeman shift. A full description of the theory describing the cooling laser polarisation and probe sequence which will need to be used will be published in [22].

5. SUMMARY

This paper has described current work at the UK National Physical Laboratory (NPL) on the development of an ion-trap based optical frequency standard based on strontium. The standard uses the $^2S_{1/2}$-$^2D_{5/2}$ quadrupole transition at 674 nm, which has a natural linewidth of 0.4 Hz. In order to realise this linewidth, the probe laser should eventually have a sub Hz linewidth and current progress towards this aim has been reported. The currently-observed linewidth of the $^2S_{1/2}(m=-1/2)$-$^2D_{5/2}$ (m=-1/2) transition is 290 Hz (FWHM). Following the NRC and NPL measurements on the frequency of this transition, work has begun on the direct comparison between two traps. Early indications are that the reproducibility is at least at the 3 parts in 10^{13} level.

A theoretical study has been undertaken into using the most abundant odd strontium isotope, ^{87}Sr. This has a nuclear spin of I = 9/2 and has transitions which are free from the first order Zeeman effect.

6. REFERENCES

[1] GP Barwood, P Gill, G Huang and WRC Rowley, "Sub-kHz "clock" transition linewidths is a cold trapped $^{88}Sr^+$ ion in low magnetic fields using polarisation switching" Opt Commun., vol 151, No 1-3, 50-5 (1998)

[2] GP Barwood, CS Edwards, P Gill, HA Klein and WRC Rowley, "Observation of the $5s^2S_{1/2}$-$4d^2D_{5/2}$ transition in a single laser-cooled trapped Sr^+ ion by using an all-solid state system of lasers" Opt Lett, vol 18, No 9, 732-4 (1993)

[3] GP Barwood, CS Edwards, P Gill, G Huang, HA Klein, WRC Rowley, "Development of diode and fibre laser sources for a Sr^+ trapped ion optical frequency standard" Proc Soc Photo-Opt Instrum Eng., vol 1837, 271-7 (1992)

[4] GP Barwood, P Gill, HA Klein and WRC Rowley, "Clearly resolved secular sidebands on the $^2S_{1/2}$-$^2D_{5/2}$ 674-nm clock transition in a single trapped Sr^+ ion" IEEE Trans Instrum Meas., vol 46, 133-6 (1997)

[5] L Marmet, AA Madej, KJ Siemesen, JE Bernard and BG Whitford, "Precision frequency measurement of the $^2S_{1/2}$-$^2D_{5/2}$ transition of Sr^+ with a 674-nm diode laser locked to an ultrastable cavity" IEEE Trans Instrum Meas., vol 46, 169-73 (1997)

[6] AA Madej, L Marmet, KN Siemesen, JE Bernard and BG Whitford, "Sub-kHz linewidths and absolute frequency measurements for the single Sr^+ atom at 445 THz", Laser Spectroscopy: XIIIth International Conference, Hangzhou, China, 1997, World Scientific, p102-5

[7] JE Bernard, L Marmet, AA Madej, "A laser frequency lock referenced to a single trapped ion" Opt Commun vol 150, 170-4 (1998)

[8] The International System of Units, 7th edition, 1998, Paris: Bureau International des poids et mesures, p137.

[9] GP Barwood, CS Edwards, P Gill, HA Klein and WRC Rowley, "Laser cooling and probing of a trapped strontium ion using all solid state lasers" ELICOLS '93, AIP Conf Proc, vol 290, 35 (1994)

[10] Ch Gertz, Th Hilberath, G Werth, "Lifetime of the $4D_{3/2}$ and $4D_{5/2}$ metastable states in Sr II" Z Phys D, vol 5, 97 (1987)

[11] AA Madej and JD Sankey "Single trapped Sr+ atom: Laser cooling and quantum jumps by means of the $4d^2D_{5/2}$-$5s^2S_{1/2}$ transition", Opt Lett, vol 15, 634 (1990)

[12] MG Boshier, private communication

[13] GP Barwood, P Gill, HA Klein and WRC Rowley "Observation of fluorescence from cold trapped Sr^+ ions in a near zero magnetic field and 674 nm probe laser development", Proceedings of the 5th symp. Frequency Standards and Metrology, World Scientific, Singapore, p509 (1996)

[14] MJ Snadden, RBM Clarke, E Riis, "Injection locking technique for heterodyne optical phase locking of a diode laser" Opt Lett vol 12, 892-4 (1997)

[15] RBM Clarke, E Riis, GP Barwood, P Gill, G Huang, HA Klein, "A sideband injection locked extended cavity diode laser for interrogating cold trapped strontium ions" Opt Commun vol 158, 36-40 (1998)

[16] B Dahmani, L Hollberg and R Drullinger, "Frequency stabilization of a semiconductor laser by resonant optical feedback" Opt Lett vol 12, 876-8 (1987)

[17] P Laurent, A Clairon, C Breant "Frequency noise analysis of optically self-locked diode lasers" IEEE J Qu Electron. vol QE-25, 1131 (1989)

[18] GP Barwood, G Huang, HA Klein and P Gill, "Sub-kHz comparison of the single-ion optical-clock $^2S_{1/2}$-$^2D_{5/2}$ transition in two $^{88}Sr^+$ traps" Phys. Rev A (Rapid Communications), accepted 1999

[19] I Siemers, M Schubert, R Blatt, W Neuhauser, PE Toschek "The trapped state of a trapped ion- Line shifts and shape", Europhys. Lett vol 18, 139 (1992)

[20] D J Berkeland, J D Miller, J C Bergquist, W M Itano and D J Wineland "Minimisation of ion micromotion in a Paul trap", J Appl Physics, vol 83, 5025 (1998)

[21] Y Fukashiro, H Sunaoshi, S Hayashibe, T Shinozuka, M Fujioka, M Yagi, I Satoh, M Wada, S Matsuki "Nuclear laser spectroscopy using a laser-microwave double-resonance method with an ion trap", Hyperfine Interactions, vol 74, 41-6 (1992)

[22] GP Barwood, P Gill, G Huang, HA Klein, M Boshier "Polarisation dependent optical pumping for interrogation of a magnetic-field-free "clock" transition in laser cooled trapped $^{87}Sr^+$" (to be submitted)

NEAR-RECOIL-LIMITED TEMPERATURES OBTAINED BY LASER TRAPPING ON THE NARROW $^1S_0 - {}^3P_1$ INTERCOMBINATION TRANSITION OF NEUTRAL STRONTIUM

Kurt R. Vogel*, Timothy P. Dinneen, Alan Gallagher and John L. Hall
JILA, University of Colorado and National Instiute of Standards and Technology
Boulder, CO 80309-0440 USA

ABSTRACT

About 2×10^7 neutral strontium atoms have been loaded into a magneto-optical trap (MOT) that uses the narrow ($\Gamma = 2\pi\,7$ kHz) intercombination line at 689 nm as the cycling transition. The atoms are second-stage laser cooled to $\sim 10\,\mu$K from an initial 5 mK atomic sample ($\sim 5 \times 10^7$ atoms), generated in a vapor cell MOT that employs the strong ($\Gamma = 2\pi\,32$ MHz) cycling transition at 461 nm. To demonstrate the low temperatures of these atoms, optical Ramsey spectroscopy is performed upon the $^1S_0 - {}^3P_1$ transition using only *two* excitation pulses; the time dependence of the fringe contrast indicates an rms velocity of 3 cm/s for the 689 nm MOT-cooled sample.

1. INTRODUCTION

Laser-cooled alkaline earth atoms are prime candidates for constructing optical frequency standards because they can be trapped and cooled on their strong $^1S_0 - {}^1P_1$ cycling transition and interrogated using the narrow, spin-forbidden $^1S_0 - {}^3P_1$ intercombination line as the clock transition. To date, strontium [1, 2, 3], calcium [1, 4, 5], and magnesium [6] have been successfully collected into a magneto-optical trap (MOT). Calcium has been the recent focus of some impressive efforts [4, 5] to harness the 400 Hz intercombination line ($Q = \nu/\Delta\nu \sim 10^{12}$) at 657 nm as an optical frequency standard.

Because the absence of ground-state Zeeman structure inhibits sub-Doppler cooling in the even-isotope alkaline earth atoms, their broad $^1S_0 - {}^1P_1$ transitions (32 MHz for Sr and Ca; 80 MHz for Mg) only cool atoms to a Doppler limit of a few mK, and the Doppler width of the trapped sample remains much greater than the intercombination linewidth. These relatively large rms velocities (50–80 cm/s) are particularly troublesome for the high-resolution optical Ramsey spectroscopy of Ca and Mg. High optical powers are necessary for making the $\pi/2$ pulses short enough such that the entire velocity distribution is excited, and the limited interaction time due to ballistic expansion of the atoms outside the Ramsey beams places a constraint on the achievable signal-to-noise ratio.

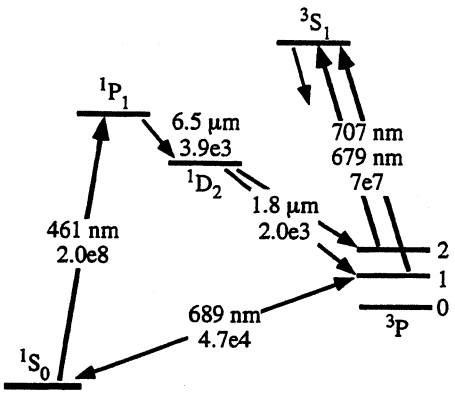


Figure 1: Level diagram for strontium showing relevant transition wavelengths and decay rates (s^{-1}).

One method of achieving lower temperatures is to Doppler cool on a more narrow cycling transition, such as the $^1S_0 - {}^3P_1$ intercombination lines of the alkaline earths. Unfortunately, the small scattering rates for these transitions in Ca and Mg result in cooling times that are prohibitively long given the initial velocity distributions of these MOT-cooled atoms. The 7 kHz Sr intercombination line at 689 nm, in contrast, has a scattering rate which is fast enough for second-stage Doppler cooling. In fact, the acceleration due to 689 nm photon recoils is enough to counteract gravity, making it feasible to laser trap using this narrow transition. In this regime where the recoil frequency step (ν_r=10 kHz) is comparable to the cooling transition linewidth, recoil-limited velocities are attainable [7] and have been independently observed by Katori et. al. [3]. The emphasis of this work is to probe the underlying physics of cooling on a transition where $E_r \sim \hbar\Gamma$, as well as evaluate the potential of using these ultracold atoms for future optical frequency standards, particularly with regard to the ultranarrow $5s^2\,{}^1S_0$–$5p\,{}^3P_2$ M2 transition at 671 nm.

2. EXPERIMENTAL SETUP

Figure 1 is a level diagram showing the pertinent cooling transitions for Sr. A frequency-doubled Ti:Sapphire laser provides the 461 nm light necessary for trapping on the $5s^2\,{}^1S_0 - 5p\,{}^1P_1$ cycling transi-

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

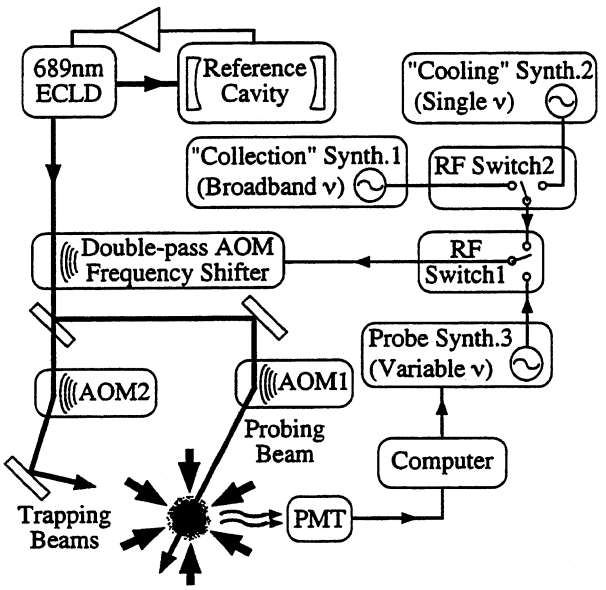


Figure 2: Experimental schematic.

tion and is offset-locked 56 MHz to the red of the ^{88}Sr resonance center in a Sr saturation spectrometer. The MOT is loaded from a background Sr vapor (n_{Sr}=10^8 cm^{-3}) and is composed of three orthogonal pairs of σ^+-σ^- beams and a 70 G/cm quadrupole field. Shelving losses to the long-lived $5p\,^3P_2$ state via the $4d\,^1D_2$ leak are eliminated by repumping to the $5p\,^3S_1$ state with diode lasers tuned to 679 nm and 707 nm [2]. Repumping allows as many as 10^8 atoms to be trapped at densities up to 10^{10}cm^{-3}.

The ~5 mK sample of Sr atoms from this MOT is further Doppler laser-cooled with 689 nm light that is red-detuned from the narrow $^1S_0 - {}^3P_1$ transition. This light is produced by a 689 nm extended-cavity laser diode (ECLD) that is locked very tightly to a high-finesse, Zerodur reference cavity. The ECLD frequency is servo-controlled to one of the reference cavity fringes by a Pound-Drever lock in a reflection topology [8], which reduces the optical linewidth to <1 Hz relative to the cavity. The Zerodur spacer of the reference cavity is suspended in a 10^{-7} torr vacuum housing whose temperature is servo-controlled to <1 mK. The resulting long-term frequency drift rate of the cavity is measured to be ~0.5 Hz/s. Radio-frequency (RF) precise tuning of the frequency-stable light is achieved by twice-diffracting the light through an acousto-optic modulator (AOM). The light is then split into six power-balanced beams (1 cm dia.) that are overlapped with the 461 nm trapping beams. A schematic of the experimental setup is shown in Fig. 2.

A primary concern of laser cooling on such a narrow (Γ=2π 7 kHz) transition is its limited velocity capture range: atoms with Doppler shifts $|kv| \gg \Gamma$ will not be resonant with the cooling light and therefore will not be radiatively damped. The capture velocity is $v_c \sim \Gamma/k$=0.5 cm/s, much less than the velocity distribution of the 461 nm MOT (v_{rms}~80 cm/s), and a very small percentage of the atoms will be cooled by single-frequency light. To increase the collection efficiency into the 689 nm MOT, the bandwidth of the cooling light is broadened by frequency modulation applied to synthesizer 1 in Fig. 2. A modulation depth of 2 MHz ensures that all atoms in our distribution are addressed by the light, while a modulation frequency of 10 kHz spaces the numerous frequency sidebands by 1 recoil frequency step (ν_r=10 kHz). The power in each cooling beam is such that its average optical power spectral density is $\sim 1 I_{sat}/\nu_r$ (I_{sat} = 3μW/cm^2 for $^1S_0 \rightarrow {}^3P_1$), and the photon scattering rate can be approximated as $\Gamma/4$ over the entire 2 MHz frequency distribution of the light. RF switch 2 allows the cooling light to be switched between broadband and single-frequency operation; broadband operation allows for maximizing the collection efficiency, while single-frequency operation can be tailored to achieve the lowest temperatures.

The >1000-fold reduction in the cycling transition linewidth severely impacts the allowed magnitude of the magnetic gradient field. If the atoms are to be stopped within the trap beam radius (R=0.5 cm), the Zeeman shift imposed by the gradient field should not exceed the optical power bandwidth, i.e., the gradient should satisfy the simple condition $g\mu_B(\partial B/\partial z)R \leq$ 2 MHz. Thus the magnetic field gradient must be reduced from 70 G/cm to $\leq$2 G/cm for the 689 nm MOT. Three orthogonal shim coils placed around the vapor cell cancel troublesome stray magnetic fields allowing good overlap of the centers of the two MOTs.

To accurately measure the velocity distribution of the MOT-cooled atoms, the same frequency-stable 689 nm light used for trapping and cooling also probes the first-order Doppler shift of the atoms. A cycling strategy facilitates non-perturbative probing of the velocity distribution. The 461 nm trap is first loaded in 40–60 msec with about 5 × 10^7 atoms, after which the 461 nm trapping light is turned off and the quadrupole B-field is reduced to 1–2 G/cm. The 689 nm trapping light is then switched on by AOM2 to further cool the atoms. About 30–60 ms of broadband cooling is followed by 10–20 ms of single frequency cooling. This trap light is then turned off and a 100 μs probe pulse (from AOM1) follows during which 689 nm fluorescence is collected with a photomultiplier tube (PMT). The probe beam (0.8 cm dia.) is linearly-polarized to address the magnetic-insensitive Δm=0 transition. The probe frequency is stepped to map the velocity distribution. Each data point is an accumulation of 100 measurement cycles.

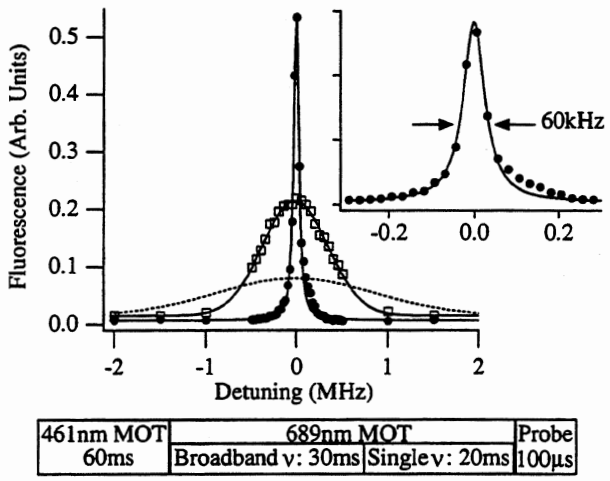


Figure 3: Narrow velocity distribution obtained by trapping on $^1S_0$–$^3P_1$ transition. Dashed line is the initial Gaussian velocity distribution of 461 mn MOT. Open squares are the Doppler-limited distribution after post-cooling on the 461 nm transition; solid line is a Gaussian fit. The 689 nm MOT distribution is shown as the solid circles; the solid line is a fit to a Lorentzian with 60 kHz FWHM. The timing sequence for a measurement cycle is shown below the graph.

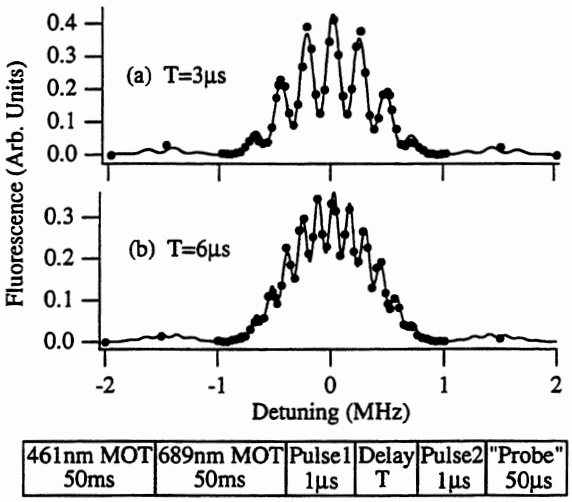


Figure 4: Two-pulse optical Ramsey spectroscopy of $^1S_0$–$^3P_1$ transition after atomic sample has been further cooled using the narrow 689 nm line. Timing sequence is shown below the graph. Circles indicate data points; lines are theoretical curves.

3. RESULTS

The velocity distributions before and after collection into the 689 nm MOT are shown for comparison in Fig. 3. The initial velocity distribution of the 461 nm MOT, shown as the dashed line, has an rms velocity of about 80 cm/s. To increase the collection efficiency into the 689 nm MOT, 7 ms of "post-cooling" occurs in which the 461 nm trap light is ramped down in intensity and its detuning decreased to -28 kHz. This post-cooling operates with unit efficiency and results in a Doppler-limited velocity distribution with $v_{\rm rms}$=27 cm/s (open squares). Next, 30 ms of spectrally-broadened 689 nm light collects about 40% of the atoms into a 200 kHz FWHM Gaussian velocity distribution. An additional 20 ms of cooling with single-frequency light that is red-detuned by 50 kHz results in a more narrow (60 kHz FWHM) yet Lorentzian distribution, shown as the solid circles in Fig. 3. A quarter of the atoms are lost while single-frequency cooling, and so the total collection efficiency into this narrow distribution is 30% of the 461 nm MOT-cooled atoms.

Analyzing the lineshape becomes problematic when the Doppler broadening of the atoms approaches the natural linewidth of the probe transition. Several distortion/broadening mechanisms contribute to the lineshape, thus masking the desired velocity information. In our experiment, saturation of the probe ($I_{\rm sat}$=4) contributes 16 kHz of Lorentzian broadening to the observed 60 kHz linewidth. The residual 44 kHz Lorentzian linewidth seems to be an artifact of the single-frequency cooling where velocity-space optical pumping is expected to produce non-Gaussian velocity distributions [7]. In addition, the asymmetric lineshape suggests that mechanical action by the probe light also plays a role, as evidenced by the additional heating on the blue side of the cooling peak. Further studies are necessary to fully determine the influence of the probe light on the experimental lineshape.

Because the residual velocity distribution appears non-Gaussian, we are reluctant to assign a temperature based solely on this lineshape. As an alternative way of probing the atomic velocities, we configured our experimental apparatus to perform optical 2-pulse Ramsey spectroscopy on the $^1S_0$–$^3P_1$ transition. Typically, optical Ramsey fringes are observed using a 4-pulse, traveling-wave (or 3-pulse standing-wave) methodology to cancel the spatial phase shifts that an atom accrues as it traverses the light field during the Ramsey interaction time [9]. Here we use a 2-pulse, traveling-wave excitation to probe these velocity-dependent phase shifts. For these experiments, the probe beam intensity is increased to provide a $\pi/2$ excitation in 1 μs. The number of atoms excited to the $^3P_1$ state is determined by collecting 689 nm fluorescence for 50 μs after the Ramsey excitation sequence.

Ramsey fringes are shown in Fig. 4 for two delay times. With 2-pulse excitation, fringes on the or-

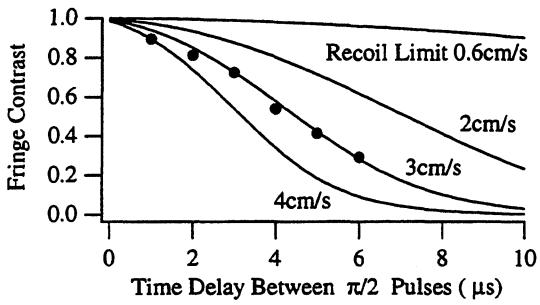


Figure 5: Dependence of Ramsey fringe contrast on delay time. Solid circles are the data, while lines are theoretical calculations for various rms velocities of the atomic sample.

der of the 7 kHz linewidth are unobtainable due to washout of the fringe contrast as the time delay is increased. The solid lines are fits to a modulated sinc function, from which the contrast ratios are extracted. These ratios are plot as a function of Ramsey delay time in Fig. 5 along with calculated dependencies for various rms velocities obtained from the theoretical treatment by Bordé et. al. [10] in which the individual atomic responses are integrated over a Gaussian velocity distribution. The data is consistent with an rms velocity of 3 cm/s, about 5 times the recoil-limited velocity of 0.66 cm/s. From this rms velocity we infer a $\sim 10\,\mu$K temperature for the narrow-line MOT.

In summary, we have demonstrated that trapping on the narrow 689 nm intercombination line results in a 25-fold velocity compression of the initial 461 nm trapped sample. Further studies are underway to determine the exact dependence of the velocity distribution on such trap parameters as magnetic field gradient and trap light powers/detunings, so that ultimately sub-μK temperatures are attained [3]. The realization of such low temperatures makes it possible to perform high-resolution spectroscopy of the doubly-forbidden $^1S_0$–$^3P_2$ transition at 671 nm. With a projected linewidth in the milli-Hertz range [11], the resonance Q will certainly be limited by the interaction time, which may approach tens of milliseconds for these free-falling, ultracold atoms. The resulting line Q could potentially reach $\sim 10^{13}$, offering the intriguing possibility of using this transition as a future optical frequency standard.

The authors thank Marla Dowell and Alex Olivas for their assistance in this experiment. We acknowledge Chris Oates for insightful discussions and keeping us abreast of the advances in calcium clock development. We also thank Carl Wieman and Eric Cornell for generous equipment loans. This work has been supported by the NSF and NIST.

*Present address: Time and Frequency Division, National Institute of Standards and Technology, MS 847.10, 325 Broadway, Boulder CO 80303

REFERENCES

[1] T. Kurosu and F. Shimizu, "Laser cooling and trapping of calcium and strontium," *Jpn. J. Appl. Phys.*, vol. 29, pp. L2127–L2129, 1990.

[2] K. R. Vogel, T. Dinneen, A. Gallagher, and J. Hall, "Experiments with strontium in a vapor cell magneto-optic trap," in *Methods for Ultrasensitive Detection*, Bryan L. Fearey, Ed., Bellingham, WA, 1998, vol. 3270, pp. 77–84, SPIE.

[3] H. Katori, T. Ido, Y. Isoya, and M. Kuwata-Gonokami, "Magneto-optical trapping of strontium atoms with photon recoil temperature," *Phys. Rev. Lett.*, vol. 82, no. 6, pp. 1116–1119, 1999.

[4] Th. Kisters, K. Zeiske, F. Riehle, and J. Helmcke, "High-resolution spectroscopy with laser-cooled and trapped calcium atoms," *Appl. Phys. B*, vol. 59, pp. 89–98, 1994.

[5] C. Oates, M. Stephens, and L. Hollberg, "An all-diode-laser optical frequency reference using laser-trapped atomic calcium," in *Proceedings of the 51st International Frequency Control Symposium*, 1997, vol. 13, pp. 219–224.

[6] K. Sengstock, U. Sterr, G. Hennig, D. Bettermann, J. Müller, and W. Ertmer, "Optical Ramsey interferences on laser cooled and trapped atoms, detected by electron shelving," *Opt. Commun.*, vol. 103, pp. 73–78, 1993.

[7] Y. Castin, H. Wallis, and J. Dalibard, "Limit of Doppler cooling," *J. Opt. Soc. Am. B*, vol. 6, pp. 2046–2057, 1989.

[8] R. W. P. Drever, J. L. Hall, F. V. Kowalski, J. Hough, G. M. Ford, A. J. Munley, and H. Ward, "Laser phase and frequency stabilization using an optical resonator," *Appl. Phys. B*, vol. 31, pp. 97–105, 1983.

[9] Ye. V. Baklanov, B. Ya. Dubetsky, and V. P. Chebotayev, "Non-linear Ramsey resonance in the optical region," *Appl. Phys.*, vol. 9, pp. 171–173, 1976.

[10] Ch. J. Bordé, Ch. Salomon, S. Avrillier, A. Van Lerberghe, and Ch. Bréant, "Optical Ramsey fringes with traveling waves," *Phys. Rev. A*, vol. 50, no. 4, pp. 1836–1848, 1984.

[11] R. H. Garstang, "Magnetic quadrupole line intensities," *The Astrophysical Journal*, vol. 148, pp. 579–584, 1967.

PROBING Ca^+ IONS IN A MINIATURE TRAP

M. Knoop, M.Herbane, M. Houssin, T. Pawletko, M. Vedel, and F. Vedel
Physique des Interactions Ioniques et Moléculaires (UMR 6633 CNRS - UAM1),
Université de Provence, Centre de St-Jérôme, Case C21, F - 13397 MARSEILLE CEDEX 20

ABSTRACT

The electric quadrupole transition of confined Ca^+ ions has been proposed as a frequency standard in the visible domain. To access metrologic conditions, we have built a cylindrical miniature trap, allowing to obtain high frequencies of motion and preparing the access to the Lamb-Dicke regime where the first order Doppler broadening is suppressed. Laser cooling and interrogation of the confined ions require lasers at 397nm and at 729nm. We propose the generation of these wavelengths using multimode laser diodes and a frequency-doubling set-up.

1. INTRODUCTION

Different earth-alkaline ions have been put forward for a frequency standard in the optical domain. Among these, Ca^+ takes a special place due to the fact that all the wavelengths necessary for cooling and probing of the ion can in principle be generated by solid-state lasers. This allows to imagine an ultimate experimental setup which would be simple and compact.

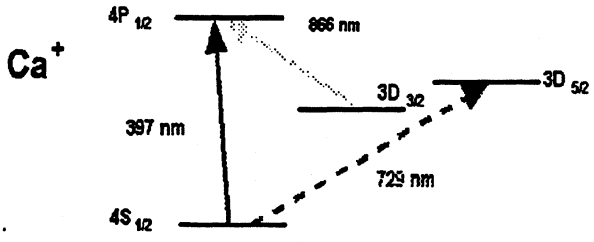


Figure 1: *Lowest energy levels of the $^{40}Ca^+$ ion*

The possible clock transition is the $4S_{1/2}$-$3D_{5/2}$ electric quadrupole transition at 729 nm in the red domain (fig. 1). The metastable $3D_{5/2}$ level has a lifetime of the order of one second [1], thus leading to a natural linewidth inferior to 200 mHz and giving rise to a Q-factor of the transition higher than 10^{15}.

In a metrological set-up $^{43}Ca^+$ will be used. This isotope has an odd nuclear spin (I=7/2) and may thus allow to work on transitions which are independent of residual magnetic fields to first order.

2. MINIATURE TRAP

For any application of the proposed transition as an atomic clock or a frequency standard, it is necessary to attain the resolution of its natural linewidth which requires extreme performances for the confined ions. One of the priorities is the elimination of the first order Doppler broadening which may cause linewidths in the GHz range for (uncooled) ions stored in a radiofrequency trap. Efficient reduction of the linewidth can be achieved by lasercooling to a few mK. The elimination of the residual Doppler broadening of some MHz can be completed in the Lamb-Dicke regime [2]. To get into this regime the ion's amplitude of motion must be inferior to a fraction of the emitted wavelength. The spectrum of the considered transition is then discrete showing a central carrier and sidebands which are separated by multiples of the ion's frequencies of motion [3, 4].

For technical reasons the access to the Lamb-Dicke regime is only possible for small trap dimensions. Moreover, a miniature trap geometry has the advantage that the total trapping volume is filled out by the cooling laser beam, thus avoiding residual hot ions in the trap. For our miniature trap design we chose a Paul-Straubel geometry [5] which offers good access for the laser beams and a large solid angle of observation.

2.1 The experimental setup

The trap consists of a cylindrical ring of molybdene with an inner ring diameter of 1.4mm and a wall thickness of 0.3mm, the total height of the cylinder ($2z_0$) equals 0.85mm (fig 2). On top and bottom of the ring compensation electrodes are placed at a distance of 5.5mm from the trap center. They have a diameter of 11mm and are made from molybdenum mesh with a transmission of 86%. For compensation of small potential defects in the trap a voltage can be applied

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

between these two mesh electrodes. Additional point electrodes in x and in y direction allow to correct field errors and to position a single ion with precision.

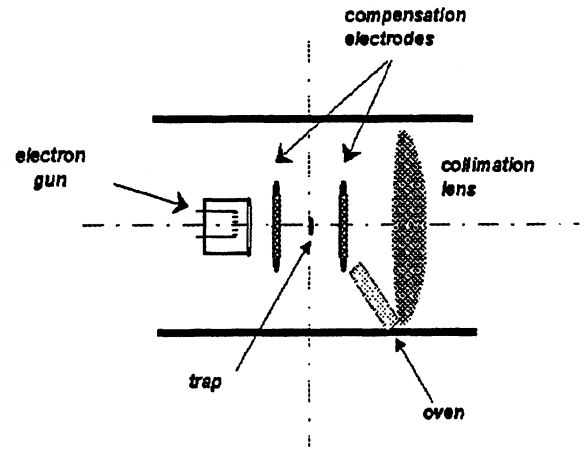


Figure 2: *Schematic design of the miniature trap; the compensation electrodes in x and y are not shown.*

The confinement frequency is $\Omega/2\pi$ = 11.6MHz with an AC amplitude V_{AC} up to $1500V_{rms}$.

Ions are created by electron bombardement of a slow atomic calcium beam produced by an oven. After loading of the trap, the oven and the electron gun are turned off to avoid collisions and a high stray light level. Collimation of the ion's fluorescence is made by an aspheric lens at about 16mm away from the trap center. This very open structure provides a solid angle of observation of almost 90°.

The base pressure in the vessel is below 5×10^{-10} mbar as measured by a Bayard-Alpert-gauge. The composition of the partial pressures of the residual gas can be determined by a mass-spectrometer with a resolution up to 1×10^{-9} mbar.

In the presented miniature trap, fluorescence signals can only be observed for laser-cooled ions. Actually, we can confine very small ion clouds (20 - 200 particles). In Figure 3 are plotted the signals of clouds containing approximately 50 to 150 ions. These signals have been obtained with the use of a light buffer gas (p(He)=1×10^{-7} mbar) for the pre-cooling of the clouds. Different cloud sizes can easily be produced by changing the heater current of the calcium oven. The temperature of the clouds can roughly be estimated from the envelopping Doppler profile. They lie between 400 and 700K, which represents an equilibrium value between the initial ion temperature

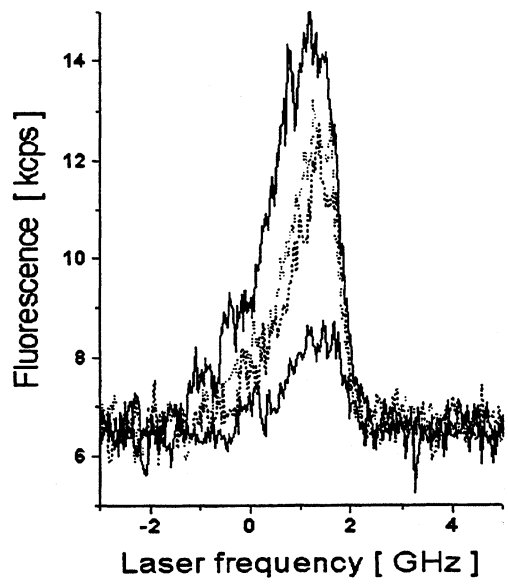


Figure 3: *Fluorescence of Ca^+ ion clouds of different sizes in the miniature trap. The different curves correspond to various values of the heater current of the calcium oven. These signals have been obtained with a moderate buffer gas pressure*

due to rf heating, the infinite heat reservoir of the buffer gas at 300K and the laser-cooling process.

Colder ions (of the order of 50K) have been trapped without buffer gas. After termination of the testing process of our trap, lower temperatures will be reached by reducing the size of the cloud in the trap.

2.2 Characterization of the miniature trap

To check the quality of the new trap we have explored the stability diagram along the q_z-axis and compared our measurements with the ideal case of the Paul trap. In Figure 4 is plotted the amplitude of the fluorescence signal of a confined cloud for various values of the amplitude of the trapping voltage V_{AC}. In the couse of the presented measurements, we did not apply a DC voltage ($a_z = 0$).

The plot shows a great number of points where the signal almost vanishes. These are most likely due to anharmonic perturbations in the trap as have been observed earlier [6]. This is indeed very probable as the simulation of the potential shows that it is quasi-harmonic (deviations less than 1%) only in a very small region (<5% of the total volume) around the center of the trap.Furthermore, the observed ion clouds are

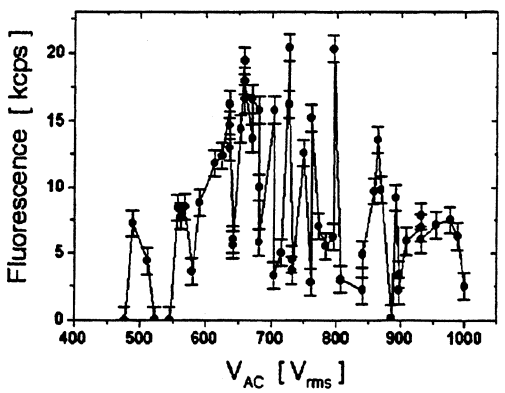


Figure 4: *Fluorescence of the confined ion cloud as a function of the trapping parameter V_{AC}. Each point corresponds to a new cloud with identical creation conditions.*

large with respect to the trap size, they do thus experience the influence of the non-ideal geometry of the trapping potential.

Different experimental and numerical methods [7, 8] have found that the maximum trapping efficiency for an ideal Paul trap is situated around $0.55 < q_z < 0.6$ with a_z close to zero. Furthermore, the limit of confinement for the first region of stability in the Mathieu diagram is q_z=0.908. Comparison of our data with these values yields a loss factor of $L = 6.8 \pm 0.5$. This factor is defined to be the ratio of the potential well depth for the ideal Paul trap with respect to the well depth in the presented trap at a given value of the trapping voltage V_{AC} [5].

3. PREPARATION OF LASER DIODES

In principle all the wavelengths involved in the Ca^+ experiment can be generated by diode lasers. However, besides the 866nm which can directly be obtained from a commercial singlemode laser diode, our laser set-ups require frequency-doubling and the use of multimode laser diodes.

3.1 397 nm

Laser radiation at 397nm from a diode laser can be generated by frequency doubling. We have built an enhancement cavity with a LBO crystal which has been tested using a titanium:sapphire laser [9].

To generate sufficient power for frequency doubling from a laser diode at the fundamental wavelength (794nm) we use an injection locking scheme. The master laser is a singlemode laser diode with a nominal power of 30mW which is placed into an external cavity to reduce its linewidth and to be able to tune its wavelength. The resulting spectral linewidth is inferior to 10.5MHz as measured by a scanning Fabry-Perot cavity (resolution limit $\sim$ 10MHz). Up to 5mW of singlemode power can be used for injection into the slave laser.

The slave laser is a 500mW broad-area laser with a free-running linewidth of about 2nm and a spatial profile that shows a couple of lobes depending on the ouput power. The master laser beam is injected into the slave via the ouput polarizer of an optical isolator. Careful matching of the size of the master laser focus on the lasing junction of the slave diode is necessary for an optimized injection. This is made by a set of two cylindrical lenses mounted on micrometer driven linear stages. Further parameters which we varied to achieve an optimized slave output are the injection angle of the master laser beam, the slave power, and the temperature of the slave laser [10].

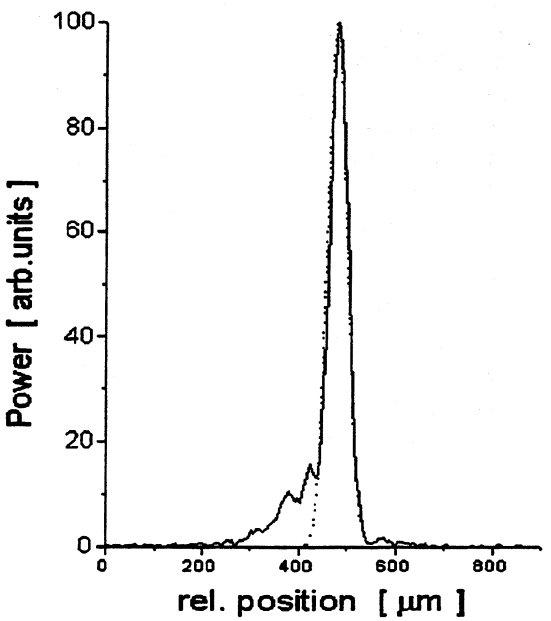


Figure 5: *Injection locking of a broad-area laser diode. The figure shows the spatial emission spectrum in the far-field. Almost 90% of the total output power is contained in the fitted Gaussian mode (dotted).*

An output with very good spectral properties has been obtained. In fact, the slave laser follows the frequency fluctuations of the master laser as closely as 1.5kHz. This could been de-

duced from a beat-note measurement between both lasers. Moreover, the injection locking has been optimized to obtain an almost Gaussian spatial output profile of the slave laser. Figure 5 shows a typical far-field emission pattern where 90% of the total output power is contained in a single Gaussian peak. The ratio of the power in the single spatial lobe with respect to the total output power decreases with increasing slave power, as the mode competition due to the free-running modes of the slave becomes more important. However, singlemode output power up to 120mW in a single Gaussian output lobe could be obtained for the described system.

3.2 729 nm

The 729nm laser will be used to probe the $4S_{1/2}$ - $3D_{5/2}$ electric quadrupole transition, its performances regarding linewidth and stability have to be comparable to the potentialities of the clock transition. In the wavelength range between 700 and 750nm, very few diodes are commercialized; we have used a multimode laser diode with a room-temperature wavelength of 729nm. This laser diode has been placed into an external cavity (Littrow configuration) to render its output singlemode. The free-running linewidth of about 2nm is then reduced to a few tens of MHz. We plan to stabilize this laser by an ULE cavity of high finesse (15000). The major difficulty lies in the poor spatial behavior of the external cavity laser. To facilitate the injection into the ULE cavity, the 729nm laser diode has been pre-stabilized onto a Fabry-Perot reference cavity (finesse=200, FSR=300MHz). The pre-stabilization set-up yields a laser linewidth of about one MHz.

4. Outlook

We have set up a cylindrical miniature trap and observed very small clouds of Ca^+ ions laser-cooled to temperatures of the order of 50K. Experiments to reduce the cloud to a single particle are currently undertaken.

Multimode laser diodes at 729nm and 794nm have been set-up using two different techniques to achieve singlemode output. While the 794nm source will be used for frequency-doubling in an external enhancement cavity, the 729nm laser is going to be stabilized on a high-finesse ULE cavity to meet the specifications necessary for a laser probing the clock transition.

REFERENCES

[1] M. Knoop, M.Vedel, F.Vedel, Phys.Rev.A **52**, 3763-3769 (1995), for a review see E. Biemont and C.J. Zeippen, Comments At. Mol. Phys. **33**, 29 (1996)

[2] R.H. Dicke, Phys.Rev. **89**, 472-473 (1953)

[3] F.G. Major and J.L. Duchene, J.Phys. (France) **A 36**, 953-959 (1975)

[4] D.J. Wineland, W.M. Itano, J.C. Bergquist, and R.G. Hulet, Phys. Rev. **A 36**, 2220-2232 (1987)

[5] C.A. Schrama, E. Peik, W.W. Smith, H. Walther, Opt.Comm.**101**, 32 (1993)

[6] M. Vedel, J. Rocher, M. Knoop, and F. Vedel, Appl. Phys. **B 66**, 191-196 (1998) and references therein

[7] R. Ifflaender and G. Werth, Metrologia **13**, 167-170 (1977)

[8] F.Vedel, J.André, Phys.Rev.**A 29**, 2098 (1984)

[9] M.Knoop, M.Vedel, M.Houssin, T.Schweizer, T.Pawletko, and F.Vedel, Proceedings of the Conference on Trapped Charged Particles and Fundamental Physics, CP 457, eds. Daniel H. Dubin and Dieter Schneider, The American Institute of Physics 1999, p. 365-368.

[10] T. Pawletko, M. Houssin, M. Knoop, M. Vedel, and F.Vedel, submitted for publication

THE OPTICAL Ca FREQUENCY STANDARD

F. Riehle, H. Schnatz, B. Lipphardt, G. Zinner, T. Trebst, T. Binnewies, G. Wilpers, J. Helmcke
Physikalisch-Technische Bundesanstalt
Bundesallee 100
D-38116 Braunschweig, Germany
email: fritz.riehle@ptb.de

ABSTRACT

PTB operates an optical wavelength / frequency standard by stabilizing the frequency of a laser to the intercombination transition $^3P_1 - {}^1S_0$ of laser cooled ballistic atomic ^{40}Ca at $\lambda \cong 657$ nm. The uncertainty of the standard was evaluated by the use of two independent systems. After comparison of the optical frequency with the Cs atomic clock the frequency is determined to be 455 986 240 494.13 (12) kHz.

1. INTRODUCTION

The progress in laser manipulation of atomic absorbers and measurement of optical frequencies paved the road to optical frequency standards to be used e. g. for the determination of fundamental constant or as optical clocks. We report on the recent achievements at PTB using the Ca–stabilized laser recommended for the realization of the meter [1] leading to one of the standards in the visible with the lowest uncertainty.

2. EXPERIMENTAL SET–UP

We have set up two different systems (see Fig. 1) which are described in more detail elsewhere [2]. They allow us to compare the transition frequencies of the intercombination line measured at two independent ensembles of Ca atoms and to check the reproducibility of the frequency delivered by the Ca standard. In order to unveil possible systematic frequency shifts we have chosen significantly different setups for the two traps. Each one comprises a magneto-optical trap (MOT) to cool and trap the Ca atoms by applying the methods of laser cooling on the $^1P_1$ - $^1S_0$ transition at $\lambda \cong 423$ nm. One trap, which will be labeled as MOT 1 in the following, is filled with atoms from a laser cooled and deflected Ca beam. To generate this beam, thermal atoms with a mean velocity of v $\leq$ 1000 m/s effusing from an oven are slowed down by a counterpropagating resonant laser beam at $\lambda \cong 423$ nm. A Zeeman slower was employed to keep the atoms in resonance with the laser during the deceleration. The decelerated atomic beam is deflected in a tilted one-dimensional optical molasses to separate the slow atoms from the laser beam. The deflected atomic beam propagates with a velocity of v $\cong$ 30 m/s towards the magneto-optical trap. The atoms are just slow enough to be captured by MOT 1. In MOT 1 the six trapping laser beams are generated by sending one beam three times

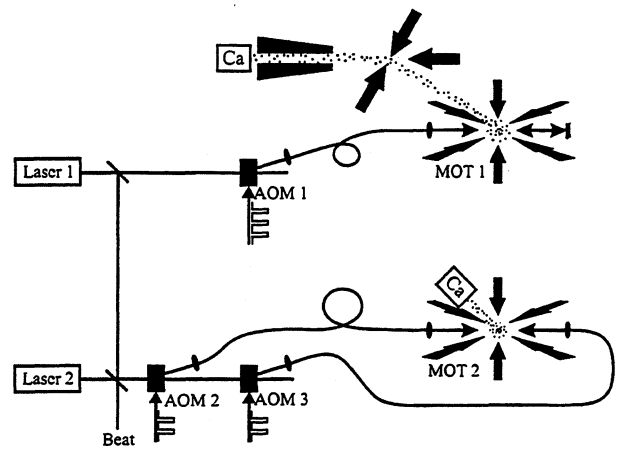


Figure 1: Experimental setup with two independent optical Ca standards

through the trap center before it is retroreflected to form the three beams in the opposite direction. This method makes the MOT's setup particularly simple since only a single beam has to be supplied. However, at high densities the MOT is not well balanced due to the unequal power in the counterpropagating beams.

In the other trap apparatus (MOT 2) we omitted the Zeeman slower and directed an effusive thermal beam towards the trap center (Fig. 1). As a consequence, only a low velocity fraction of the atomic flux of the beam can be captured by the trap i. e. those atoms from the Boltzmann velocity distribution ($v_{prob} \cong 600$ m/s, $T_{oven} \cong 900$ K) which are slower than the trap's maximum capture velocity of about 30 m/s. To obtain an acceptable loading rate the oven was placed close to the trap center ($d = 15$ cm). Furthermore, we increased the capture velocity by using two laser frequencies for the horizontal trapping beams. The additional laser frequency obtained by an acousto–optic modulator allows faster atoms to get in resonance with the trapping beams of the additional frequency. The number of atoms could be increased by roughly a factor of seven to about $5 \cdot 10^6$ atoms stored in MOT 2. This value is comparable to the result obtained with MOT 1 of about 10^7 atoms.

The radiation to excite the Ca intercombination line is generated by a high-resolution laser spectrometer comprising a dye laser or a diode-laser system

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

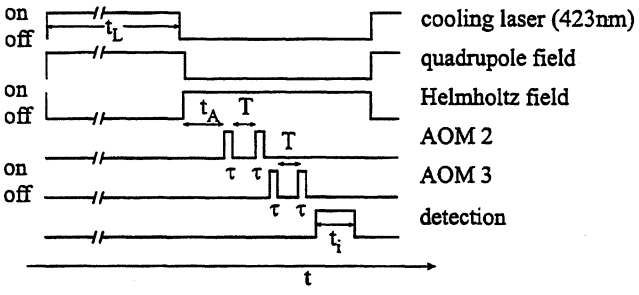


Figure 2: Timing diagram of the pulsed trapping, interrogation and fluorescence detection using a magneto–optical trap (see text and Fig. 1)

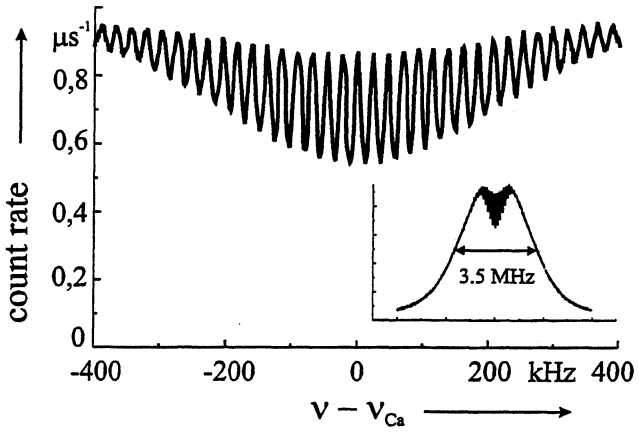


Figure 3: Optical Ramsey resonances excited in an expanding cloud of Ca atoms released from a magneto–optical trap

prestabilized to a suitable resonance frequency of a Fabry-Perot resonator. The observed line widths of both lasers are below 0.5 kHz and the drift of the prestabilized dye-laser system is less than 0.4 Hz/s whereas the drift of the diode laser system is higher [3], predominantly due to temperature fluctuations.

The laser beams used to interrogate the intercombination transition of the cold Ca atoms are derived from acousto–optic modulators (AOMs) (Fig. 1) used as "light switches"for the pulsed excitation (see Fig. 2). The radiation at $\lambda \cong 423$ nm used to cool, deflect, and trap the Ca atoms in MOT 1 was produced by a dye-laser pumped by an argon ion uv–laser. In MOT 2 a solid state laser system was employed consisting of a high power diode laser (200 mW) and a frequency doubling $KNbO_3$ crystal in a buildup ring resonator. In both experiments atoms are loaded for about fifteen to twenty milliseconds. After this time the number of atoms detected via the fluorescence at 423 nm saturates. To avoid systematic frequency shifts due to the Zeeman effect of the magnetic field of the MOT and light shifts due to the 423 nm radiation, both the magnetic field and the light of the trapping laser have to be shut off before interrogating the Ca atoms (see Fig. 2). The atomic cloud expands according to the root–mean–square (rms) velocity of the ballistic atoms. The rms velocity can be determined from the Doppler broadening of the intercombination line at $\lambda \cong 657$ nm and it is in the range of $v_{rms} \cong 1$ m/s for a Doppler broadening of approximately 3 MHz corresponding to a temperature of a few millikelvin. The velocity distribution for MOT 1 does not differ significantly from that measured in MOT 2. The low velocity of the ballistic atoms allows to utilize the full natural linewidth of the Ca intercombination transition of about 0.4 kHz corresponding to an excited state life time of 0.4 ms. During this time the atoms of a velocity of about 1 m/s move only by about 0.4 mm.

3. OPTICAL RAMSEY EXCITATION

A total interaction time that allows to resolve the natural line width of about 0.4 kHz would excite only a very narrow velocity group of the Doppler broadened distribution with a considerably reduced signal. In order to achieve high spectral resolution combined with a good signal–to–noise ratio (S/N), we apply the method of separated field excitation. In contrast to the Ramsey excitation originally devised for the interrogation of atomic beams by microwaves we apply the optical Ramsey excitation now in the time domain. Short pulses of $\tau = 1$ μs duration are used to excite a significant part of the cold ensemble of atoms. The necessary high spectral resolution is then achieved by a sufficiently large time separation T between two consecutive pulses. Subsequently, the red fluorescence of the excited atoms is detected. The corresponding time sequence (Fig. 2) consists of atom trapping ($t_L \cong$ 15 ms), turning off the trapping fields ($t_A \cong 0.5$ ms), separated–field excitation ($T \cong 0.1$ ms - 0.5 ms), and detection ($t_i \cong 0.5$ ms).

Time domain separated–field excitation was performed either by three pulses of a standing laser field or by two pulses of counterpropagating traveling laser fields, respectively. When the frequency of the interrogating laser is tuned close to the resonance, the fluorescence intensity contains a contribution (Fig. 3) which varies with the cosine of the laser detuning. Provided that the lengths of the pulses are small compared to their separation, the width of the interference fringes $\delta\nu = 1/(4T)$ is inversely proportional to T. The narrowest fringe width obtained with trapped atoms was close to the natural linewidth of about 400 Hz. In fact, the signals shown in Fig. 3 represent a superposition of two interference signals caused by the two recoil components which are separated by 23.1 kHz. Consequently, the time T between the pulses is adjusted in such a way that the interference structures resulting from both recoil components add to give a maximum visibility of the signal. Increasing the time T reduces the period of the interference structure and consequently the linewidth of the fringes.

4. FREQUENCY STABILIZATION

Any one of these fringes can be used to stabilize the frequency of the laser. The error signal for the stabilization is generated from the interference signal by modulating the laser frequency and simultaneously measuring the fluorescence intensity. In the most straightforward way, the frequency is square-wave modulated between two discrete values with the mean frequency tuned close to the center of the central fringe. The difference of the corresponding fluorescence intensities can be used as error signal. This method corresponds to a first harmonic detection (1f-method) of a servo–control system using analog electronics and harmonic modulation. The maximum slope is obtained for a total modulation width of $\delta\nu_{mod} = 1/4T$, i.e. if the frequency alternates between the points of maximum slope of the interference signal. The 1f–method can be applied to signals of which the background is symmetric to the center of the absorption line. However, slight asymmetries of the background may occur since the interference structure is superimposed to an incoherent background (see Fig. 3). It is caused by the saturation dip and the Doppler broadened spectrum since the Doppler profile is blue shifted by half of the recoil splitting ($\cong$ 11.5 kHz). In servo control–systems using analog electronics, the influence of such asymmetries is reduced by detecting the clock transition at the third harmonic of the modulation frequency (3f–method). We have approximated this method by measuring the fluorescence signal S(ν) at the following four frequencies $\nu = \nu_L \pm \delta\nu_{res}/4$ and at $\nu = \nu_L \pm 3/4\delta\nu_{res}$ where ν_L is the mean value of the modulated laser frequency and $\delta\nu_{res} = 2\delta\nu = 1/2T$ corresponds to the period of the interference structure. The error signal $E(\nu_L)$ was then calculated as follows:

$$E(\nu_L) = 3 \cdot [S(\nu_L + \delta\nu_{res}/4) - S(\nu_L - \delta\nu_{res}/4)] - [S(\nu_L + 3\delta\nu_{res}/4) - S(\nu_L - 3\delta\nu_{res}/4)].$$

In this equation, $E(\nu_L)$ is not sensitive to a linear and a quadratic frequency dependence of the background. We have checked the suppression of the background by stabilizing the laser frequency alternatingly with the 1f- and the 3f–method. The results obtained at a fringe period of 23.1 kHz show that the shift of the central fringe is in the range of 7 Hz for the 1f-method. The residual shift from the 3f-method was below the detection capability of our setup. A calculation from the measured spectrum resulted in a 50–fold suppression compared to the 1f-method [2].

After the detection, the error signal is used to step the frequency of the laser spectrometer which corresponds to a digital integrating servo control. The frequencies of both, the dye laser spectrometer and the diode laser spectrometer [3] are prestabilized to highly stable reference Fabry–Perot resonators. The linear drift of the eigen frequency of the reference resonator can be determined by the servo control and compensated for by adding a corresponding feed–forward signal to the signal controlling the laser frequency thus obtaining a doubly integrating feed–back loop.

The spectral resolution $\delta\nu = \nu/Q$ and the S/N of the interference structure (depending on the averaging time τ) are important parameters of a frequency standard. The instability of the frequency $\sigma(2,\tau)$ is determined ultimately by $\sigma(2,\tau) \cong 1/(S/N(\tau) \cdot Q)$ in the case of white noise. The observed values of S/N as taken from spectra like the ones presented in Fig. 3 decrease with increasing T due to the residual phase noise of the laser and to the natural linewidth of the clock transition. With our present setup, we obtain the maximum value of $S/N \cdot Q = 4.5 \cdot 10^{12}$ at about T = 350 μs. In most cases, we have performed the stabilization at a fringe width of approximately 1.2 kHz, i.e. at a resolution which is slightly larger than the optimum value for maximum stability.

From the periodicity of the fringes it might seem ambiguously to find the proper frequency of the intercombination transition. There are, however, several means to determine the proper fringe. One method relies on the fact that for variations of the fringe period i. e. for different resolutions only the frequencies of the minima at either recoil component are constant. Another possibility would make use of the beautiful method of applying a series of pulses in either direction to obtain Fabry–Perot–like high–finesse interference structures [4].

5. UNCERTAINTY OF THE Ca STANDARD

To convert a stable and reproducible oscillator to a frequency standard it is necessary to reference the frequency of the oscillator to the frequency of the primary standard of time and frequency, the Cs atomic clock. Hence, three independent sources contribute to the uncertainty of an optical standard. First, there is the uncertainty to realize the line center of the unperturbed Ca atom which mostly depends on the ability to reduce and correct perturbations affecting the frequency of the atomic transition. A second contribution results from the uncertainty of the primary Cs standard. Third, the comparison of the two frequencies of the Cs and Ca standards differing by five orders of magnitude may additionally contribute to the measurement uncertainty. These contributions have been investigated experimentally and theoretically [2] and they are listed in Table 1.

We first discuss the different effects affecting the uncertainty to realize the center of the clock transition of the Ca standard. An important contribution is caused by the residual first–order Doppler effect which will be looked at into in more detail at the end of this section. Due to the low velocities the second–order Doppler effect caused by the time dilation contributes as little as a few millihertz and can be neglected at present. We have investigated the influences of magnetic fields by the Zeeman effect, of electric fields by the Stark effect, of gravitational and other accelerations. The ac Stark shift caused by scattered light

Table 1: Contributions to the standard uncertainty in the frequency of the Ca optical frequency standard during a recent frequency measurement and attainable uncertainty.

effect	uncertainty achieved	uncertainty attainable
1^{st} order Doppler effect	6 Hz	1 Hz
2^{nd} order Doppler effect	3 mHz	1 mHz
magnetic fields	3 Hz	0.2 Hz
ac-Stark effect	5 Hz	0.1 Hz
quadratic Stark effect	0.06 Hz	0.06 Hz
gravitation	2 Hz	0.2 Hz
black–body radiation	3.5 Hz	0.5 Hz
second recoil component	1 Hz	–
collision of cold atoms	10 Hz	1 Hz
stabilization scheme	1 Hz	0.4 Hz
reproducibility	50 Hz	–
sum in quadrature	53 Hz	1.6 Hz
counting errors (chain)	100 Hz	0.1 Hz
statistical uncertainty	4 Hz	0.5 Hz
H-maser	5 Hz	0.5 Hz
Cs-clock	7 Hz	0.5 Hz
total uncertainty $\delta\nu$	113 Hz	1.8 Hz
relative uncertainty $\delta\nu/\nu$	$2.5 \cdot 10^{-13}$	$4 \cdot 10^{-15}$

of the cooling laser results from uncomplete blocking of the AOMs. We estimate its influence at present to 5 Hz and it can be largely eliminated by the use of fast mechanical shutters. Furthermore, there is a contribution of the field of temperature radiation (so–called black–body shift) which is of particular importance in our MOT 2. There, the oven is very close to the cloud of atoms. Even though the atoms are trapped behind a circular beam stop they are exposed to the radiation from the oven during their ballistic flight or due to radiation reflected from the walls. With the help of a shutter we expect to reduce this influence and the remaining uncertainty (0.5 Hz) will mainly result from the calculation of the correction. A particular contribution is due to the influence of the superposition of the interference patterns due to the recoil effect. Since we have chosen such a period that both interference structures add constructively, there is an uncertainty associated with this procedure which we estimate at present to be below 7 Hz. It can be completely removed if one recoil component is suppressed by one of the methods developed earlier (see e.g. [5]). The influence of the collisions of the cold Ca atoms has not been investigated yet, but it is expected to be below 10 Hz. A small contribution results from the stabilization. In the digital servo system we do not expect integrator offsets and we estimate the influence of the stabilization to 0.1 % of the linewidth.

The estimated uncertainties were checked by comparisons between two independent systems where the clouds of Ca atoms were released from the two independent magneto–optical traps depicted in Fig. 1. The comparisons have been performed during the past years with the laser frequency stabilized to the atomic clouds from the two traps (Fig. 4). The frequency of the laser stabilized to the atoms released from the MOT 1 filled with the laser cooled beam seems to be consistently higher than the frequency derived from the second apparatus (MOT 2) with a mean fractional frequency difference of about $7.5 \cdot 10^{-14}$. Since the origin of this difference was not known previously we took it into account in Table 1 as a contribution termed "reproducibility". Latest investigations indicate that this difference is mainly due to a residual first–order Doppler shift in the laser beams with non ideal wavefronts. Measurements of the Doppler profile showed that when the trapping fields were switched off the center of mass of the atomic cloud started to move. By acceleration of the cloud it was possible to achieve frequency offsets of up to about 80 Hz for different diameters of the counterpropagating clock laser beams and purposely increased acceleration. During all of the measurements used for the determination of the frequency this effect is supposed to be smaller, however, we assume that the influence of the residual first–order Doppler effect has been underestimated. This effect might therefore be responsible for the contribution termed "reproducibility" in Table 1. In subsequent investigations with identical microscope lenses the frequency differences could be reduced to a few hertz. We therefore conclude that at present the uncertainty due to this effect can be restricted to 6 Hz. We expect to further reduce it to about 1 Hz if the intensity of the six independent trapping beams will be carefully balanced. Besides the curvature of the wave fronts, acceleration due to gravity results in a frequency shift if the laser beams are not aligned horizontally (Fig. 5). The corresponding frequency shift (Fig. 6) depends linearly on $T + T'$ where T is the time between the two pulses of either direction and T' is the time between

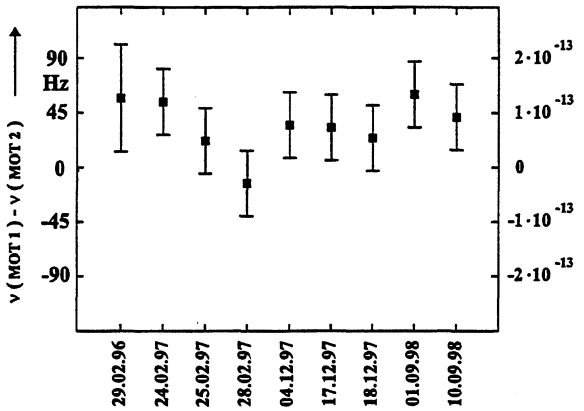


Figure 4: Difference between the transition frequencies realized by two independent ensembles of cold Ca atoms.

the second and the third pulse. The frequency shift $\delta\nu$ measured by reversing the time order of the pulses allows to identify and to minimize this contribution.

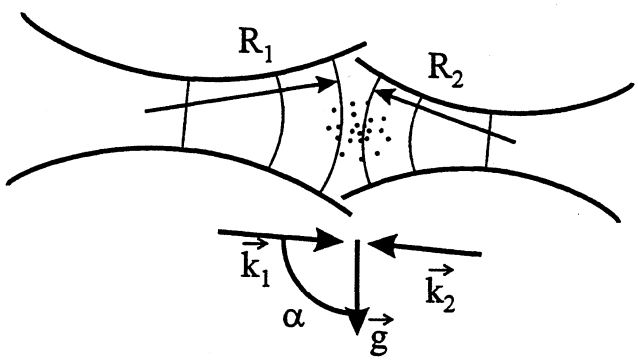


Figure 5: Excitation of an atomic cloud by tilted laser beams with curved wave fronts

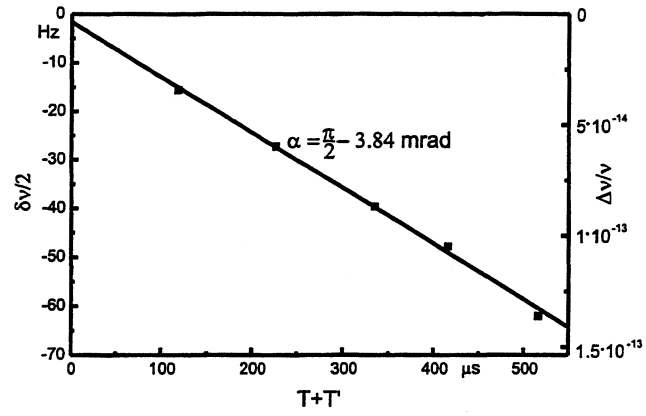


Figure 6: Measured frequency shift due to tilted laser beams and gravitation

From the contributions to the uncertainty to the Ca stabilized laser added in quadrature we derived 53 Hz for the previous measurements [2] and 14 Hz for the status with our present setup. Estimations based on these results (third row of Table 1) show that a relative uncertainty of a few parts in 10^{15} can be expected [2].

6. MEASUREMENT OF THE FREQUENCY

The (optical) frequency of the Ca standard was measured by comparison with PTB's primary Cs atomic clock, using a phase-coherent frequency chain. This chain is described in more detail elsewhere [6] and it suffices to recall that it basically applies the method of harmonic mixing thereby connecting the Ca standard via various intermediate oscillators (diode laser, colour center laser, CO_2 lasers, methanol laser, backward wave oscillator, Gunn oscillator, and H-maser) to the Cs clock. We point out that the described frequency chain allows a phase-coherent optical frequency measurement which, in principle, does not contribute

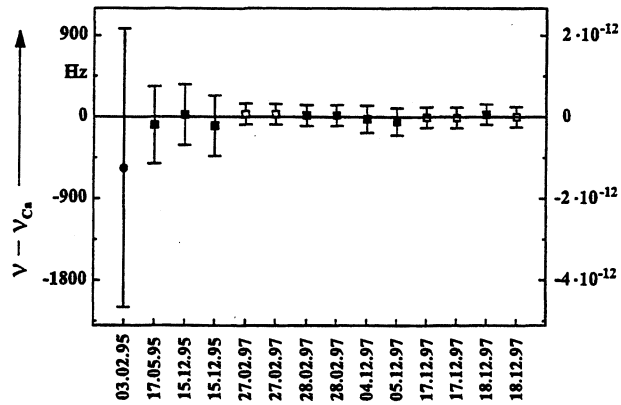


Figure 7: Measured frequency values of the Ca optical frequency standard. The measurements were taken with an effusive beam (circle), MOT 1 filled by a laser cooled beam (filled squares) and MOT 2 filled by a thermal beam (empty squares). Different values of the same day correspond to different pulse separations.

to the uncertainty of the measured frequency value. However, cycle slips can occur in each multiplication step. The rate of cycle slips critically depends on the S/N ratio of the beat signal. We estimated this rate in the most crucial stages of the chain and concluded that they contributed less than 100 Hz to the uncertainty (Table 1).

Several frequency measurements during more than two years (Fig. 7) have been performed using both MOTs, different resolutions, and different stabilization schemes. The weighted mean of all frequency measurements up to now is $\nu_{Ca} = 455\ 986\ 240\ 494.13\ (12)$ kHz. We emphasize that this uncertainty makes the Ca stabilized laser one of the most accurate optical frequency standards in the visible, today. Consequently, this standard has been recommended recently by the CIPM [1] for the realization of the meter with the lowest uncertainty. Current investigations look for the origin of the difference between the frequencies of both setups. Provided, that the cause for this discrepancy has been identified we expect from Table 1 that the uncertainty can be reduced by more than an order of magnitude.

As a result of the frequency measurements the relative frequency instability of the Ca standard was shown to decrease with $1/\sqrt{\tau}$ for integration times between $\tau = 1$ s and $\tau = 1000$ s. It reaches its flicker floor at approximately 10^{-14}. It has been shown how the stability can be improved by applying a modified detection scheme [7] allowing to reach the 10^{-14} level within an integration time of about $\tau = 1$ s.

7. DISSEMINATING THE Ca FREQUENCY

Accurate optical reference frequencies have various applications e. g. in precision spectroscopy, determination of fundamental constants, and length / wave-

length metrology. As a particular application we recall the measurement of the frequencies of the hyperfine structure lines R(180) 0-16 and R(42) 0-17 in molecular iodine at 815 nm wavelength [8]. The authors made use of the fact that these lines nearly coincide with the known frequency difference between the methane stabilized He–Ne laser and the Ca standard [1]. In the same way the sum frequency of these two standards allows one to set up frequency references in the green spectral range. Similarly, the combination of optical frequency standards already existing or under development might be used to establish a grid of reference frequencies covering the spectral range from the ultraviolet to the infrared range. Together with the conventional and novel techniques of division and measurement of optical frequencies discussed in this volume, virtually any optical frequency can be measured or synthesized. For the time being, however, as long as there are no small sized, cheap and reliable optical frequency synthesizers of ultimate accuracy available, the dissemination of optical frequencies will have to rely on the development of transportable optical frequency standards. Furthermore, these standards will allow to compare different concepts of frequency chains as well as the primary standards. For dissemination of the frequency of the Ca standard, a transportable optical frequency standard based on an effusive Ca beam has been developed at PTB [9] with a fractional uncertainty of $1.2 \cdot 10^{-12}$. This represents an improvement of more than an order of magnitude with respect to the widely used iodine stabilized He–Ne laser operating at $\lambda \cong 633$ nm wavelength. The use of existing small and efficient light sources will allow to generate the cooling radiation and the development of a transportable standard based on laser cooled Ca atoms.

ACKNOWLEDGEMENTS

This work was supported in part by the Deutsche Forschungsgemeinschaft (DFG) under SFB 407.

REFERENCES

[1] T. Quinn, "Mise en Pratique of the Definition of the Metre (1992)", Metrologia, Vol. 30, pp. 523-541, 1994 and
Report of the 86th meeting of the Comité International des Poids and Mesures (CIPM). Bureau International des Poids et Mesures, Sevres, Paris Cedex, 1997

[2] G. Zinner, "Ein optisches Frequenznormal auf der Basis lasergekühlter Calciumatome", PTB–Bericht, PTB-Opt-58, Braunschweig 1998

[3] V. Vassiliev, V. Velichansky, P. Kersten, T. Trebst, F. Riehle, "Subkilohertz enhanced–power diode–laser spectrometer in the visible", Opt. Lett. Vol. 23, pp. 1229-1231, 1998

[4] H. Hinderthür, F. Ruschewitz, H.–J. Lohe, S. Lechte, K. Sengstock, W. Ertmer, "Time–domain high–finesse atom interferometry", Phys. Rev. A Vol. 59, pp. 2216-2219, 1999

[5] F. Riehle, A. Witte, Th. Kisters, J. Helmcke, "Interferometry with Ca atoms", Appl. Phys. B Vol. 54, pp. 333-340, 1992

[6] H. Schnatz, B. Lipphardt, J. Helmcke, F. Riehle, G. Zinner, "First Phase-Coherent Frequency Measurement of Visible Radiation", Phys. Rev. Lett. Vol. 76, pp. 18-21, 1996

[7] T. Kurosu, G. Zinner, T. Trebst, F. Riehle, "Method for quantum–limited detection of narrow–linewidth transitions in cold atomic ensembles", Phys. Rev. A Vol. 58, pp. R4275-R4278, 1998

[8] B. Bodermann, M. Klug, H. Knöckel, E. Tiemann, T. Trebst, H. R. Telle, "Frequency measurement of I_2 lines in the NIR using Ca and CH_4 optical frequency standards", Appl. Phys. B Vol. 67, pp. 95-99, 1998

[9] P. Kersten, F. Mensing, U. Sterr, F. Riehle, "A transportable optical calcium frequency standard", Appl. Phys. B Vol. 68, pp. 27-38, 1999

1999 Joint Meeting EFTF - IEEE IFCS

ALL-DIODE-LASER OPTICAL FREQUENCY STANDARD BASED ON LASER-TRAPPED Ca ATOMS

C. W. Oates, F. Bondu, and L. Hollberg

Time and Frequency Division, 847.10
National Institute of Standards and Technology, 325 Broadway, Boulder, CO 80303
Tel: 303-497-7654 Fax: 303-497-7845 e-mail: oates@boulder.nist.gov

ABSTRACT

We describe a high performance optical frequency standard at 657 nm based on laser-trapped Ca. Using two semiconductor laser systems (423 nm for trapping and 657 nm for spectroscopy), we have been able to obtain sub-kilohertz optical linewidths and a relative Allan variance of 1.3×10^{-14} at one second, with prospects for considerable improvement. This Ca standard will be used to make a phase-coherent frequency measurement of the 282 nm $^{199}Hg^{+}$ optical standard.

1. INTRODUCTION

We have developed a table-top apparatus which is beginning to realize the considerable potential of the $^1S_0$ (m=0) $\rightarrow$ $^3P_1$ (m=0) intercombination line at 657 nm in neutral ^{40}Ca for use as a frequency standard. This transition is an attractive optical frequency standard due to its narrow linewidth (400 Hz), insensitivity to external perturbations, and convenient wavelengths for trapping and spectroscopy (see Figure 1).

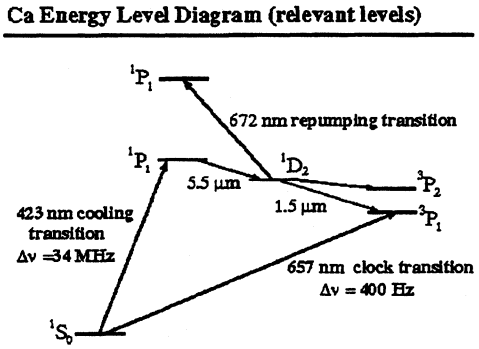


Figure 1. Cooling and clock transitions in Ca.

Since the pioneering efforts of Barger et al. in 1979, this line has been the subject of many experimental investigations[1]. The group of Helmcke and Riehle at Physikalisch Technische Bundesanstalt (PTB) has achieved numerous milestones with this transition, including a recent measurement of its absolute frequency with only 120 Hz uncertainty[2]. This makes it the most precisely know optical frequency of those recommended for the realization of the meter[3].

As an optical frequency standard, this transition offers remarkable potential for both stability and accuracy. For practical experimental parameters, we estimate an attainable fractional-frequency instability of $\leq 3\times10^{-16}\tau^{-1/2}$, nearly two orders of magnitude smaller than the projected instability of Cs fountain standards. Moreover, it has been estimated that a fractional frequency uncertainty of 10^{-15} or less is possible[4], although critical systematic effects such as collision shifts still need to be evaluated.

In order to approach this performance with a relatively straightforward apparatus, we had to address several critical issues. The first was a simplified design for a Ca magneto-optic trap (MOT), which included the development of a frequency-doubled semiconductor laser system to generate 50 mW of trap light at 423 nm. Next, we had to construct a 657 nm diode-laser system with the frequency stability required to resolve sub-kilohertz optical linewidths. Finally, to approach atom shot-noise-limited performance, we needed to implement a "shelving" detection technique, which enables us to achieve a high signal-to-noise ratio even at sub-kilohertz resolution.

2. APPARATUS

This apparatus has been described in detail elsewhere, so we will just outline the essential aspects here[5,6]. Our Ca MOT uses 50 mW of trapping light generated by frequency doubling 200 mW of 846 nm light from a master-oscillator power-amplifier (MOPA) semiconductor laser system. We load the trap directly with slow atoms from a thermal beam emerging from a 600°C oven. To increase the atom flux through the trap region, we locate the trap only 13 cm from the Ca oven. We also add a laser beam counter-propagating to the atomic beam to slow the atoms and increase the loading rate (by an order of magnitude), but we do not employ Zeeman slowing magnets as used in other schemes. This slowing/trapping combination can load as many as 10^7 atoms into our trap in 20 ms. The trapped sample has a temperature of 2 mK and a lifetime of ~20 ms, which is limited by optical pumping to the $^3P_2$ state through the $^1D_2$ state.

We perform the high resolution spectroscopy with a frequency-stabilized 657 nm laser diode in an extended-

cavity configuration. The frequency of the laser is locked to a high finesse cavity (finesse = 60 000) by feeding back primarily to the diode laser injection current. With a servo bandwidth of more than 3 MHz, we can achieve a fast laser linewidth of <20 Hz. Our reference cavity is environmentally isolated but still contributes measurable noise to the laser frequency. Recent improvements in acoustic and vibration isolation have reduced the cavity noise to <100 Hz in 1 s. We are able to tune the laser frequency relative to this stable reference cavity with a microwave synthesizer which feeds a double-passed acousto-optic modulator (AOM).

Since we require high power for time-domain optical Ramsey spectroscopy, we send the output from the frequency-stabilized master oscillator to a power amplifier. After spatially filtering the power amplifier output with an optical fiber, we have ~40 mW of useful probe power. We then chop this output beam with AOM's to generate the desired probe pulses.

3. OPTICAL RAMSEY SPECTROSCOPY WITH SHELVING DETECTION

After turning off the trapping fields, we excite the narrow 657 nm transition with a sequence of red pulses in accordance with the technique of time-domain optical Ramsey spectroscopy[4,7,8]. This four-pulse sequence enables us to maintain a high signal-to-noise ratio when going to high resolution. To improve the signal-to-noise ratio further, we have implemented a shelving detection technique first developed for probing trapped ions[9,10]. In our version we first probe the atoms with a nearly resonant 423 nm probe pulse. The induced fluorescence level gives us a measure of the number of ground-state atoms present at the beginning of the measurement cycle. We then excite a fraction of the atoms with the four-pulse Ramsey sequence. Before the excited atoms can spontaneously decay back to the ground state, we read out the modified ground state population with a second blue probe pulse. The ratio of fluorescence induced by the second blue pulse to that induced by the first then gives the depletion of the ground state due to red excitation. Since we can cycle many blue photons per atom per measurement, we get a much cleaner measure of the excitation and can approach 100 % detection probability per atom. Moreover, since we measure the *fraction* of atoms excited, we are effectively normalizing the signal against trap number fluctuations, which can be as large as 10 %.

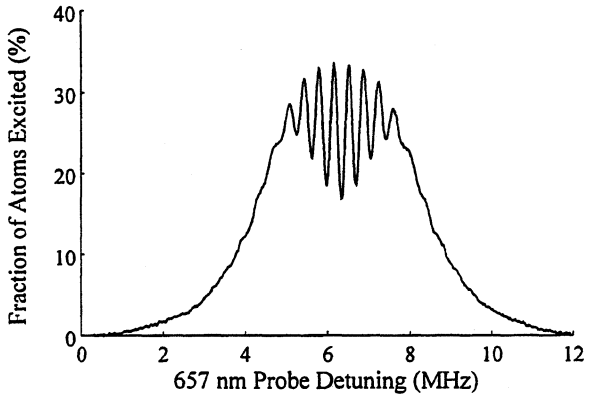


Figure 2. Low resolution optical Ramsey fringes. Total measurement time was 15 s.

An example of low resolution optical Ramsey spectroscopy taken with this system can be seen in Figure 2. The Ramsey fringes are superimposed on a Doppler background that corresponds to a temperature of 2.4 mK. By increasing the time between red pulses we can generate fringes with a much shorter period. Figure 3 shows the central fringes taken at a full-width half-maximum (FWHM) resolution of 960 Hz (resulting from a Ramsey time of 260 μs for each pulse pair).

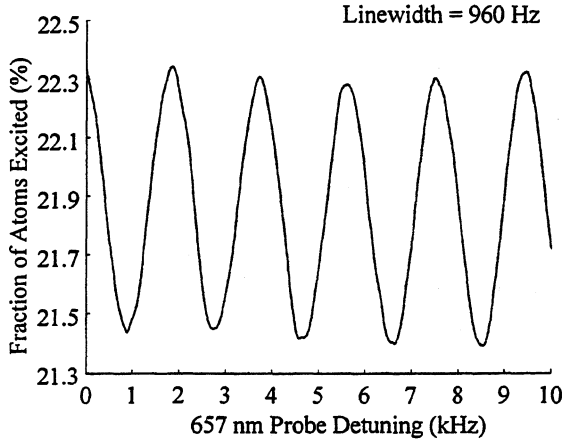


Figure 3. Sub-kilohertz optical Ramsey fringes with shelving detection. Total measurement time was 30 s.

From this figure we estimate a fractional frequency instability of close to $10^{-14}\tau^{-1/2}$. With our system we have actually been able to resolve linewidths as narrow as 400 Hz, equal to the natural linewidth of this transition.

The choice of optimal Ramsey resolution for the frequency standard is a compromise between narrow linewidth and high signal-to-noise ratio, which for our present system yields an optimal operating linewidth of ~1.6 kHz. As we improve the system and reduce excess low-frequency noise, we expect to operate nearer to the theoretical optimal resolution of 630 Hz (FWHM). To lock the red laser to a Ramsey fringe, we modulate the probe laser frequency with a 100 Hz square wave, then demodulate the resulting excitation, and feed back to the laser frequency. To evaluate the fractional frequency instability properly, we should either build a second system for comparison or compare our system to a more stable one. Lacking either of these at present,

we can still make an estimate of the Allan variance by measuring the fluctuations of the Ca signal relative to the reference cavity (after subtraction of the cavity drift). This "relative" Allan variance actually includes the cavity noise, so we think that our Ca system may actually be better than this measurement, although any relative linear drift of the Ca system is suppressed. Nonetheless, we see in Figure 4 promising performance with a relative instability of 1.3×10^{-14} at one second.

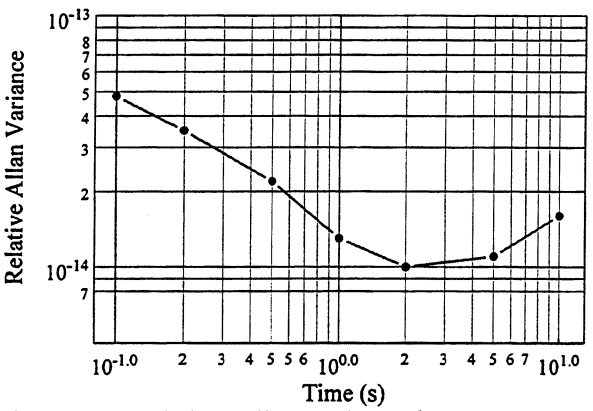


Figure 4. Relative Allan variance between the Ca transition and the optical reference cavity. A Ramsey fringe linewidth of 1.9 kHz was used for this measurement.

We have not yet reached the shot-noise capability of our system, because there is still some residual amplitude noise due to imperfect normalization.

4. FUTURE WORK

As a first application of our frequency standard, we are currently constructing a frequency chain to transfer knowledge of the Ca frequency (based on the PTB value) to the 282 nm transition in the Hg^+ ion frequency standard under development at NIST[11]. This phase-coherent chain is discussed in another paper at this conference[12]. The ultimate goal of this optical frequency measurement work will be to connect both the Ca and Hg^+ optical transitions to the Cs microwave standard with a (nearly) all-optical, phase-coherent frequency chain.

An essential part of this work involves sending Ca-stabilized light through a 100 m fiber to another room for frequency comparison. An initial measurement of the frequency noise written on to the light by the low frequency fluctuations caused by the fiber is shown in Figure 5. Here we have sent light through the fiber and then shifted its frequency with a double-passed AOM before sending it back. Beating this return light against the incident light yields twice the single-pass fiber noise contribution. It was important to perform this measurement with the laser locked to the optical cavity to prevent laser frequency noise from contributing to the noise spectrum. The ~1.5 kHz single-pass broadening seen in Figure 5 can be removed with a fiber-noise-cancellation technique first demonstrated by Ma et al.[13].

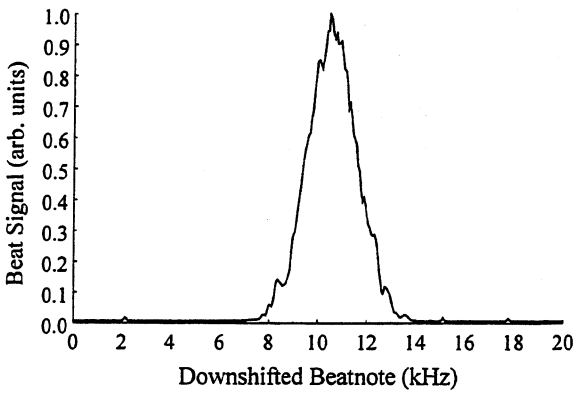


Figure 5. Beatnote between light before and after fiber with an averaging time of 5 s. Note that this shows the noise accumulated due to *two* passes through the fiber.

We acknowledge valuable discussions with J.L. Hall, J. Bergquist, K. Vogel, B. Young, R. Fox, J. Helmcke, and F. Riehle.

5. REFERENCES

[1] R.L. Barger, J.C. Bergquist, T.C. English, and D.J. Glaze, "Resoluton of photon-recoil structure of the 6573- calcium line in an atomic beam with optical Ramsey fringes," Appl. Phys. Lett., vol. 34, pp. 12-14, June 1979.

[2] H. Schnatz, B. Lipphardt, J. Helmcke, F. Riehle, and G. Zinner, "First phase-coherent frequency measurement of coherent visible radiation," Phys. Rev. Lett., vol. 76, pp. 18-21, Jan. 1996.

[3] T. Quinn, "Mise en Practique of the Definition of the Metre (1992)," Metrologia, vol. 30, pp. 523-541, Jan. 1994.

[4] Th. Kisters, K. Zeiske, F. Riehle, and J. Helmcke, "High-resolution spectroscopy with laser-cooled and trapped calcium atoms," Appl. Phys. B, vol. 59, pp. 89-98, Aug. 1994.

[5] C.W. Oates, M. Stephens, and L. Hollberg, "An all-diode-laser optical frequency reference using laser-trapped calcium," in Proceedings of the 51st Frequency Control Symposium, 1996, pp. 219-224.

[6] C.W. Oates, F. Bondu, R.W. Fox, and L. Hollberg, "A diode-laser optical frequency standard based on laser-cooled Ca atoms: sub-kilohertz spectrsocopy by optical shelving detection," accepted for publication in Eur. J. Phys. D.

[7] Ye. V. Baklanov, B. Ya Dubetsky, and V.P. Chebotayev, "Non-linear Ramsey resonance in the optical region," Appl. Phys., vol. 9, pp. 171-173, Feb. 1976.

[8] J.C. Bergquist, S.A. Lee, and J. L. Hall, "Saturated absorption with spatially separated laser fields: observation of optical Ramsey fringes", Phys. Rev. Lett., vol. 38, pp. 159-162, Jan. 1977.

[9] D.J. Wineland, J.C. Bergquist, W.M. Itano, and R.E. Drullinger, "Double-resonance and optical-pumping experiments on electromagnetically confined, laser-cooled ions," Opt. Lett., vol. 5, pp. 245-247, June 1980.

[10] W. Nagourney, J. Sandberg, and H. Dehmelt, "Shelved Optical Electron Amplifier – Observation of Quantum Jumps," Phys. Rev. Lett., vol. 56, pp. 2797-2799, June 1986.

[11] B.C. Young, R.J. Rafac, F.C. Cruz, J.C. Bergquist, W.M. Itano, and D.J. Wineland, "$^{199}Hg^{+}$ Optical Frequency Standard" in these proceedings.

[12] B. Frech, J.S. Wells, C.W. Oates, J. Mitchell, Y.P. Lan, T. Kurosu, L. Hollberg, B.C. Young, and J.C. Bergquist, "Sub-systems for optical frequency measurements, application at the 563 nm Hg^{+} transition and the 657 nm Ca line", in these proceedings.

[13] L.S. Ma, P. Jungner, J. Ye, and J.L. Hall, "Delivering the same optical frequency at 2 places – accurate cancellation of phase noise introduced by an optical fiber or other time-varying path," Opt. Lett, vol. 19, pp. 1777-1779, Nov. 1994.

PRESENT PERFORMANCE OF THE TRANSPORTABLE HE-NE/CH_4 OPTICAL FREQUENCY STANDARDS

M.Gubin, A.Shelkovnikov, E.Kovalchuk, D.Krylova, E.Petrukhin, D.Tyurikov

P.N.Lebedev Physical Institute, Leninsky pr.53, Moscow 117924, Russia

ABSTRACT

The paper presents frequency stability, reproducibility and performance limitations of the transportable He-Ne/CH_4 optical frequency standards. These type of devices can be used as precise (~10^{-13}) mid-Infrared references in optical metrology and frequency synthesis.

1. INTRODUCTION

New generation of transportable He-Ne/CH_4 (λ=3.39 μm) optical frequency standards (TOFS) [1,2] with resolved Magnetic Hyperfine Structure of $F_2^{(2)}$ methane line and improved signal-to-noise ratio can play an important role in the field of optical frequency metrology and precise spectroscopy.

It seems quite real that frequency repeatability for these devices could reach the level of 10^{-13}-10^{-14} while relative stability may exceed the parameters of H-maser at averaging times (10^{-3}-10^{4}) s. Now such kind of TOFSs after calibrating at the radio-optical frequency chain with respect to Cs primary standard can disseminate the absolute value of optical frequency ($\omega_{CH4} \cong$ 88 THz) with the uncertainty of ~10^{-13} and can be used in laboratories interested into build-up the accurate IR-visible frequency grid. This way to avoid the problem of a cumbersome low frequency part (10^{10} Hz to 10^{14} Hz) of the phase-coherent radio-optical frequency chains seems to be attractive particularly now with appearance of the cw optical parametric oscillators (OPO) in the mid-IR range (3.0-4.0) μm. Powerful (~400 mW) OPO phase locked to the TOFS frequency can overcome the problem of relatively low output intensity (~1 mW) of the TOFS and make much easier the non-linear mixing and conversion of the 88 THz radiation.

The present paper describes a recent progress of a new version of the transportable He-Ne/CH_4 OFS: the improvement of a middle-term stability and investigation of the physical reasons influencing on the device reproducibility and repeatability.

2. OPTICAL SCHEME OF THE TOFS

A physical principle of TOFS are based on the sub-Doppler resonances of saturated absorption (SA) and saturated dispersion (SD) of $F_2^{(2)}$ methane line [3].

To narrow the line width down to $\gamma \cong$ (3-5) kHz (HWHM) and to resolve completely the Magnetic Hyperfine Structure (MHFS) the telescopic expander is combined with a low pressure (30-50) μTorr CH_4 cell.

The device is built on the standard scheme of precise spectroscope and, in addition to the laser with a telescopic expander, it contains a Reference laser and a Heterodyne laser (see Fig.1 in ref.[4]). The Reference laser is permanently stabilised by SA / SD resonances with a half width $\gamma \cong$ 150 kHz. It provides radiation with narrow spectrum which is transferred to the Telescopic laser via Heterodyne laser and offset phase-lock servos.

The servo system based on personal computer and phase-lock loops with programmable synthesizers, records super-narrow resonances at the Telescopic laser. It determines a top of the central MHFS component and with interval (10-50) s introduces an error correction signal for the Reference laser to eliminate its slow drifts. Finally the output signal of the TOFS is the radiation of the Heterodyne laser, with the average frequency determined by the central component of $F_2^{(2)}$ methane line and a fixed offset (-600 kHz), determined by one of the synthesizers.

3. FREQUENCY STABILITY

Since 1996 several TOFSs with beam diameter D ≅ 60 mm in telescopic part were built in the Lebedev Institute. Absolute frequency measurements (AFM) were done at the radio-optical frequency chain (ROFC) of PTB and also by comparison with a stationary OsO_4 OFS (λ = 10.6 μm) and the renewed ROFC of BNM-LPTF [5, 4].

The relative stability (square root of Allan variance), determined for TOFS-2 during the AFM at PTB in 1997 is shown on Fig.1 (curve 1).

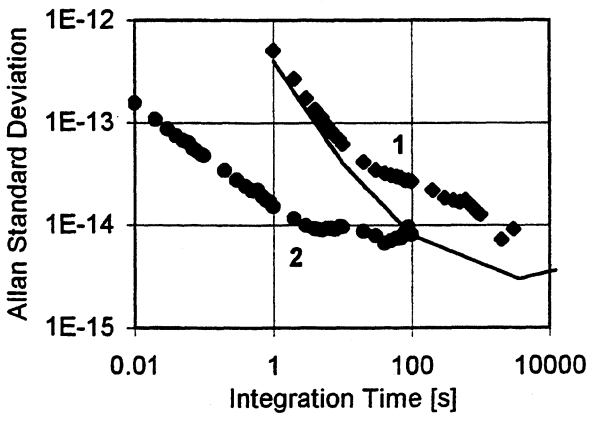


Fig.1. Allan standard deviation measured for output frequencies of: (1) TOFS-2 relatively H-maser (PTB); (2) improved Reference laser relatively stationary one. Solid line - H-maser accuracy limit.

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

For sample times $\tau > 10^2$ s the relative stability exceeds 2×10^{-14}; for sample times $\tau < 10$ s the stability is better than the stability of the H-maser which served as a reference for PTB chain. At the intermediate range of τ = (10 - 100) s TOFS's stability was limited by the Reference laser flicker noise.

At present the performance of the Reference laser in TOFS is improved. The curve 2 at Fig.1 shows its stability measured at Lebedev Institute relatively stationary laser. The range of sample times where the Reference laser stability exceeds the H-maser performance is enlarged up to τ = 100 s.

4. FREQUENCY REPEATABILITY AND REPRODUCIBILITY

The AFM at PTB with different TOFSs have shown that each device has repeatability $(2 - 3)\times10^{-13}$ on the long period of the time (months - year). But the frequency offsets between independently assembled devices are in the range of 200 Hz (2×10^{-12}) [4].

Fig.2 shows the results of the simultaneous AFM carried out with TOFS-2 and TOFS-3. The initial frequency difference between devices was $\Delta \cong 80$ Hz. The manipulations by some parameters (the CH_4 pressure, the field intensity, the amplitude of modulation, the mirrors adjustment) during AFM didn't bring shifts exceeding 20 - 30 Hz.

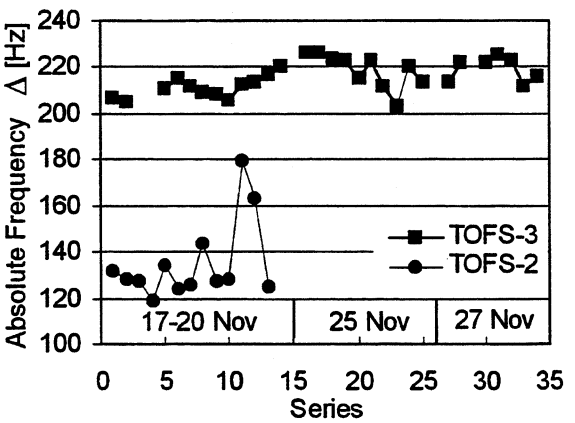


Fig.2. Frequency of TOFS-2 and TOFS-3 measured at PTB in 1997. f_{CH4} = 88 376 181 600 000 + Δ [Hz] Series 11 and 12 were obtained under laser heating.

Analysing the reasons of such frequency difference between TOFSs we finally came to conclusion that the only origin of such big shifts is connected with different optical adjustment of their telescopes.

Measurements made in laboratory have shown the noticeable dependence of output frequency versus defocusing of the telescopic cavity within a stability range of the resonator (~800 μm). Fig.3 shows that the full scale frequency shift versus longitudinal telescopic defocusing can reach 100 Hz for ±100 μm longitudinal lens shift.

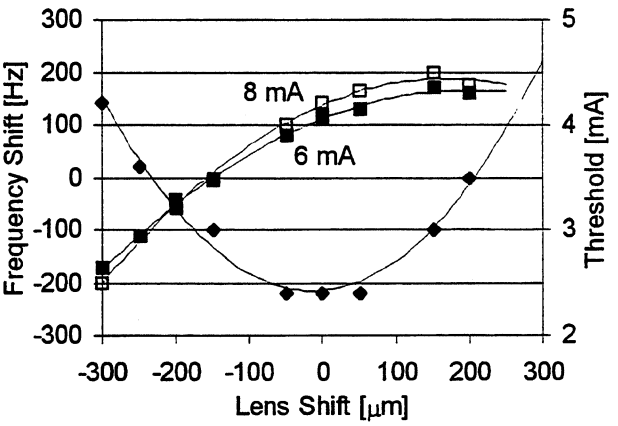


Fig.3. TOFS's frequency shift vs longitudinal telescopic lens detuning from the center of cavity stability range for two discharge currents. Lower curve - threshold current.

Thus it seems reasonable that the shift between different TOFSs could be caused by uncertainty in their telescopes adjustments. Even temperature difference in the room during the process of assembling could influence on the adjustment due to thermal expansion of a stainless steel jacket of the telescope tube. For example, heating of the TOFS-2 jacket on 10°C (this temperature change corresponds to ~ 100 μm detuning from the center of the resonator stability range) leads to the frequency difference between devices of 20 Hz instead of 80 Hz (Fig.2, series 11 and 12). That is in agreement with the measurements presented on Fig.3, where the slope of the frequency shift in the center of the resonator stability range can be evaluated as ~ 60 Hz / 100 μm.

5. THEORETICAL ESTIMATIONS OF THE FREQUENCY SHIFTS

We described analytically the influence of defocusing of the telescope's lenses on the form of the modes of empty resonator. The longitudinal defocusing of telescope doesn't introduce extra wave front curvature (WFC) to the Gaussian beam inside of the telescopic part. The plane mirror fastens the flatness of the plane beams in the absorbing cell situated there (Fig.4).

It worth mentioning [6], that for such optical configuration, like a present one, in the domain of $\gamma\tau \leq 1$ (γ and $1/\tau$ are correspondingly the homogeneous and transit time line-widths), where these devices usually operate, the saturation resonances originated from the conical parts of the beams (due to the large WFC) have much smaller intensity and much wider width, than the saturation resonances from the almost plane parts of the beams. So for description of SA/SD resonances in such kind of standards we can describe only the resonance in the field of two counterpropagating plane beams in the telescopic part of the resonator.

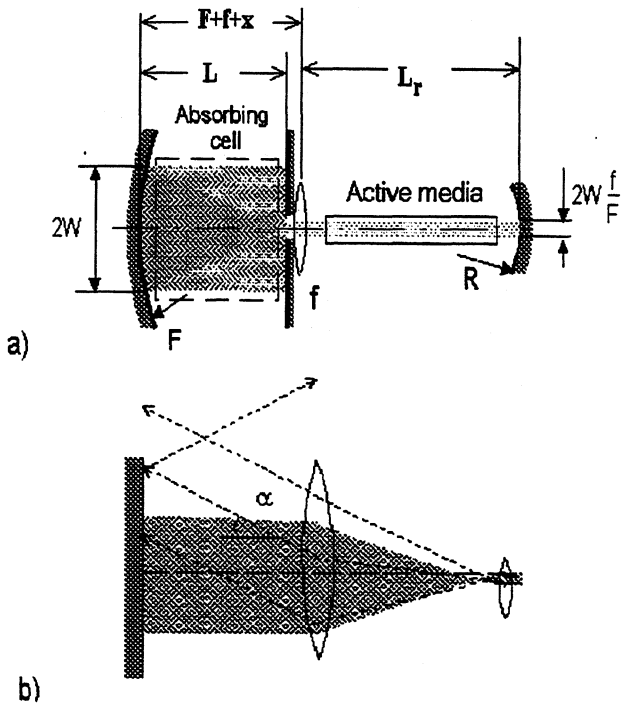


Fig.4. The optical scheme (*a*) and equivalent (unwrapped) (*b*) scheme of the TOFS.
The angle α between the main and parasitic plane beams in the telescopic part is essentially enlarged for good layout.

The influence of longitudinal defocusing of the telescope on the field inside the telescopic part consists in the appearance of the dependence of the beam width W in the absorbing cell and the dependence of the frequency distance (Δ) between nearest transverse modes of empty resonator from the longitudinal distance between focuses of the telescope's lenses x:

$$W^2 = \frac{F^2}{f^2}\frac{c\lambda}{\pi\Delta},\ \Delta = 2c\left(\frac{1}{L_r}\left(\frac{1}{R}+\frac{x}{2f^2}\right)\right)^{1/2}, \quad (1)$$

where λ=3.39μm, L_r =1m is the length of the resonator outside the telescope, R=5m is the radius of curvature of the outside mirror, F and f are correspondingly the focal distances of the large and small telescope's lenses. Formulas (1) were written in approximation of close-to-plane-mirror-resonator [5,7]. It does not correspond exactly to the working domain of present standard, but it permits us to make estimations of the influence of some transverse effects.

Longitudinal defocusing of the telescope does not lead directly to any frequency shift. It leads only to the mediate ones, i.e. to the changes of the order of magnitude of the already existing frequency shifts of the present standard, connected with the transverse structure of the field.

We estimated the influence of the two such effects, like the inhomogeneity of active media and the influence of the diaphragms, and found the following:

1) in the domain of $\gamma\tau \sim 1$, in which the present standard is operating ($\gamma \sim$ 2-3 kHz and $1/\tau \sim$ 3 kHz), the influence of the circular diaphragms is notably reduced due to transit time effects (on a factor of 50-100 as compared with the pure homogeneously broadened case), and leads to the frequency shifts of the center of the resonance, which do not exceed the order $10^{-3}\gamma$ even when the diameter of the diaphragm is comparable with the beam's transverse size;

2) The inhomogeneity of the active media leads to appearance of additional not conjugated WFC (contrary to custom conjugated WFC fastened by the mirrors curvature or by the telescope defocusing), which can be taken into account, like in [7], by introducing the new Gaussian beams with the complex width W, in which Δ is equal to

$$\Delta = 2c\left(\frac{1}{L_r}\left(\frac{1}{R}+\frac{x}{2f^2}-i\frac{g_0}{kR_0^2}\right)\right)^{1/2},$$

where $g_0 \sim 1$ is the gain coefficient per total pass throw the resonator, $R_0 \sim$ 0.5 cm is the diameter of the tube with the active media. Our numerical calculations for such field in the domain of $\gamma\tau \sim 1$ gave us the estimations of the total shift of the SD resonance due to such effect of the order of 60 Hz, and it's changes due to longitudinal defocusing of the telescope on the ± 0.5 mm within the range of resonator stability ≤ 40 Hz (and tends to zero with increasing of x, i.e. going away from the close-to-plane-mirror-resonator case).

Other possible reasons of the resonance shift are the *"parasitic" waves* caused by reflection (back to the telescopic part) from the several optical surfaces inside the resonator. The presence of such surfaces leads to the parasitic beam under angle α to the plane beam in the absorbing cell. This angle α between the counterpropagating main and parasitic plain waves in the telescopic part in the case of $\gamma\tau \le 1$ leads to emphasising of the role of molecules with some non-zero velocity in the formation of saturation resonance for the main beam in the presence of parasitic one and finally to the shift of SD resonance

$$\frac{\Delta\omega}{\gamma} = -2r_p\alpha^2\frac{L}{\lambda}\eta(\gamma\tau),$$

where r_p is the coefficient of parasitic reflection from the optical surface, L - the length of the absorbing cell, and $\eta(\gamma\tau)$ is some function of parameter $\gamma\tau$, which at $\gamma\tau \sim 1$ is of the order $\eta(1) = 0.5$. For values of $r_p \approx 0.01$, $L \approx F = 0.5$ m, and $\alpha \approx 5\times10^{-3}$ the uncertainty of the TOFS frequency is on the level of $|\Delta\omega| \sim 4\times10^{-2}\gamma \sim 160$ Hz.

CONCLUSIONS

At present the TOFS systems demonstrate frequency repeatability (for one device during several months, year) of 2×10^{-13}, and frequency stability up to 1×10^{-14} (for averaging time in the range of 1-100 s). Frequency reproducibility (for different devices) is of the order of 2×10^{-12}.

It was found that the frequency differences between devices are caused mainly by initial adjustment of the telescopes. The telescope defocusing changes the beam parameters and results in frequency shifts through transverse inhomogeneity of the active media. Theoretical estimations partially explain this difference.

More precise tuning of the telescopes, decreasing of the cavity losses, elimination of the back reflections give us the hope to decrease several times this kind of uncertainty in future.

Acknowledgments. This work was financially supported by the US Civilian Research and Development Foundation (CRDF Grant No. RE1-193), by the Russian Foundation for Basic Research (Grant No. 98-02-16724) and the Russian Federal Scientific-Technical Program 'Fundamental Metrology'.

REFERENCES

[1] M.Gubin, D.Tyurikov, A.Shelkovnikov, E.Kovalchuk, G.Kramer and B.Lipphardt, «Transportable He-Ne/CH_4 optical frequency standard and absolute measurement of its frequency», IEEE J. of Quant. Electr., v. 31, No. 12, 1995, p. 2177.

[2] S.Bagayev, A.Dmitriev, P.Pokasov, «Laser Physics», v. 7, No. 4, 1997, p. 989.

[3] See reviews: V.Alekseev, M.Gubin, E.Protsenko, «High precision optical frequency standards», Laser Physics, v. 1, No. 3, 1991, p. 221; V.Chebotayev, S.Bagayev, A.Titov, «Saturated absorption lineshape under transit-time conditions», Laser Physics, v. 4, No. 2, 1994, p. 224; and references therein.

[4] O.Acef, A.Clairon, G.D.Rovera et al , «Absolute frequency measurements with a set of transportable methane optical frequency standards», (in this issue).

[5] O.Acef, A.Clairon, L.Hilico, D.G.Rovera, G.Kramer, B.Lipphardt, A.Shelkovnikov, E.Kovalchuk, E.Petrukhin, D.Tyurikov, M.Petrovsky, and M.Gubin, "Absolute frequency measurements and intercomparisons of He-Ne/CH_4 (λ = 3.39 μm), CO_2/OsO_4 (λ = 10.6 μm) frequency stabilized lasers and Cs primary standard", in CPEM'1998 Digest, p.258, Washington D.C. (USA) July 6-10, 1998.

[6] V.Alekseev, M.Gubin, D.Krylova, D.Tyurikov, A.Shelkovnikov, «Methane frequency standard (λ = 3.39 μm) with the bandwidth of the reference line of 800 Hz», Laser Physics, v. 4, No. 2, 1994, p. 356.

[7] V.Alekseev, D.Krylova, «An effect of gain inhomogeneity on the laser standard frequency shift», Sov.J.Quantum Electron., 20, 1990, p. 789.

1999 Joint Meeting EFTF - IEEE IFCS

PERFORMANCES OF OsO_4 STABILIZED CO_2 LASERS AS OPTICAL FREQUENCY STANDARDS NEAR 29 THz

F. Ducos, , G.D. Rovera, C. Daussy and. O. Acef

BNM-LPTF, Observatoire de Paris
61, Avenue de l'Observatoire 75014 Paris - France
e-mail : ouali.acef@obspm.fr

ABSTRACT

In this paper, we report on the metrological capabilities of CO_2/OsO_4 optical frequency standards operating around 29 THz. Those frequency standards are currently involved in various fields such as frequency metrology, high resolution spectroscopy and Rydberg constant measurements. The most impressive features of the standards, lies in the 10^{-15} level frequency stability, allied to a long-term reproducibility (1 year) of 1.3×10^{-13}.

1. INTRODUCTION

During the last decade, the role of CO_2 laser frequency stabilized onto narrowed saturated absorption of OsO_4 molecular line, have been emphasised by the needs of absolute frequency measurements in the mid IR and visible domains (HeNe/I2 at 633 nm, DL/Rb at 778 nm) [1,2], of accurate spectroscopy of hydrogen atom, and concurrently of the Rydberg constant measurements [3]. Nowadays, they are currently used as starting point to calibrate these frequencies with respect to cesium primary standard, using optical frequency chains (OFC).

In this way, they allow to bridge in two steps the large frequency gap from the microwave range to the visible domain. On the other hand, since 1997 a OsO_4 frequency grid, known with uncertainties ranging from 7×10^{-13} to 2×10^{-12} [4], is recommend by the CIPM for the mise en pratique of the definition of the meter. This frequency grid corresponds to a strong absorption band which covers partially (from 28.1 to 29.3 THz) the wide emission spectrum of CO_2 lasers from 26 to 33 THz range

2. EXPERIMENTAL SET-UP

Two similar CO_2/OsO_4 frequency standards have been developed and compared to evaluate their performance. These systems have been already described in more details elsewhere [5]. Each device (Fig. 1) comprises a sealed CO_2 laser which is coupled to a 1.5 m long and high finesse (F = 170) Fabry-Perot (FP) cavity for saturated absorption purpose.

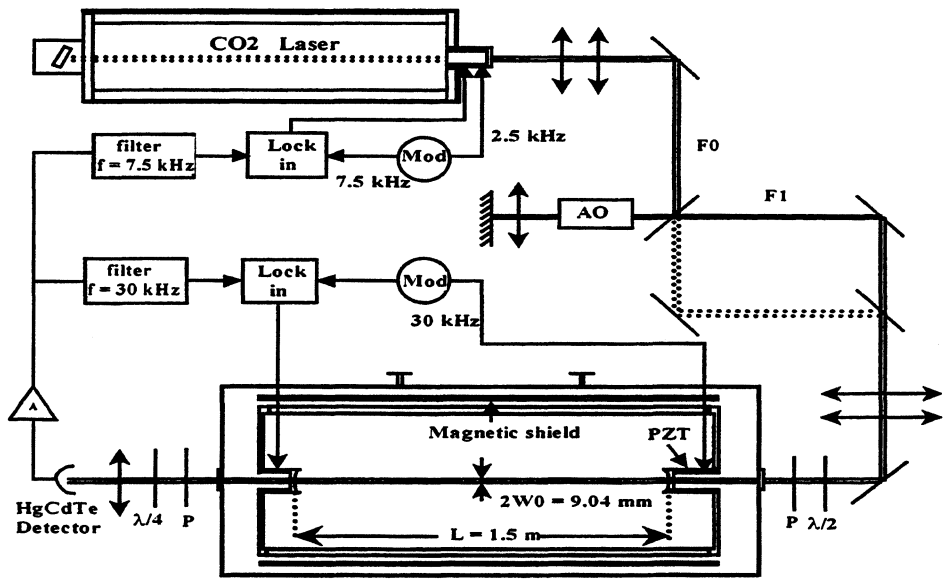


Fig. 1. CO_2/OsO_4 Experimental set-up

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

The FP cavity consists of two spherical ZnSe mirrors, with a 50 m radius of curvature. It is surrounded by a magnetic shield and the total apparatus is placed in a stainless-steel tube, sealed at the ends by two ZnSe windows. This tube is evacuated each day to 10^{-4} Pa range, before filling with 0.133 Pa of OsO_4 pressure. In this range the OsO_4 linewidth is typically about 20 kHz.

The 1 meter long sealed CO_2 laser is designed for low frequency noise and shielded against acoustic noise. The beat note between two independent free running CO_2 lasers shows a linewidth smaller than 1 kHz.

Thanks to the high finesse FP cavity, the required power to saturate the OsO_4 transitions is less then 1μW at the input of the FP cavity. Therefore, the CO_2 output power is attenuated by a factor of about 10^6. In this way, the laser is strongly isolated from optical feedback.

For the frequency stabilization of the overall system, we use the optical transmitted signal detected with a nitrogen cooled HgCdTe detector. The FP is weakly modulated at 30 kHz and the laser frequency at 2.2 kHz. The first harmonic of the FP transmission fringe is devoted to lock the FP , while the third harmonic of the OsO4 line is used to stabilize the laser frequency. The two independent CO_2/OsO_4 systems remain currently frequency locked more than 15 hours. This limitation is due to the warming-up of the nitrogen cooled detector [5].

3. MAIN FEATURES

The OsO_4 natural mixture comprises 7 isotopic species of Os: 192 (40%), 190 (26%), 189 (16%), 188 (13%), 187 (1.6%),....[6] The best results have been performed using the saturated absorption transitions of the two most abundant isotopes (192 and 190). Their relative abundance, associated to the absence of an hyperfine structure (for even isotope species), allows achievement of strong and narrow resonances.

Nevertheless, corresponding probably to hot bands many OsO_4 transitions in the vicinity of the CO_2 laser lines are unidentified. This is the case of the one located at – 15.254 MHz from the $10R(10)CO_2$ laser center line and used in this work. This OsO_4 frequency line has been linked to a primary cesium clock using the BNM-LPTF optical frequency chain (OFC), first in 1984 [7] and more recently several times from April 1998 to March 1999. The use of this particular OsO_4 line is justified by its perfect coincidence with the microwave to IR frequency chain [8].

Fig. 2. reports on the results of the absolute frequency measurements of one of the two CO_2/OsO_4 standards, against cesium clock, over 1 year. This set of data has been obtained by the use of very restrictive experimental conditions which are 0.133 Pa for the OsO_4 pressure and 100 μWatt for the intra-cavity power. These two parameters are measured within 1 and 2% of accuracy respectively, yielding to 6.2% in term of the OsO_4 peak saturation contrast. Its achievement in the range of 6 to 6.4 % determines in part the ± 5 Hz shot-to-shot measurement observed on this Fig. 2.

The larger frequency dispersion (about 25 Hz) observed in March 99, is due to this OsO_4 peak contrast variations (of about 20 %) consecutive to an incomplete cavity out gassing following a PZT change inside the FP cavity at the beginning of March'99.

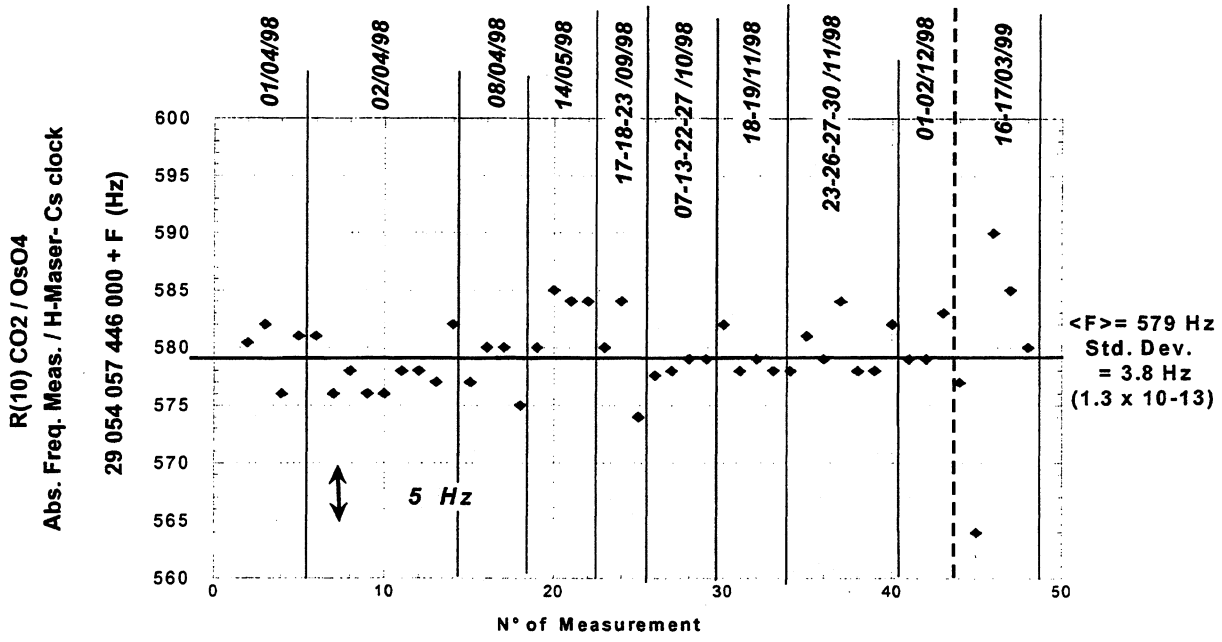


Fig. 2. Absolute frequency measurement of $R(10)CO_2/OsO_4$ laser against cesium clock over 1 year.

However, the mean value of this frequency standard, when the absorption corresponds to the standard conditions, is not affected at the level of 1×10^{-13}. The mean frequency of this CO_2/OsO_4 standard is:

F(1) = 29 054 057 446 579 (4) Hz,

with a relative uncertainty of 1.3×10^{-13} (1 σ), taking into account the total data.

The second similar CO_2/OsO_4 system, exhibit a slightly different frequency (12 Hz lower), with a relative uncertainty twice compared to the former one. This observed discrepancy at the beginning of our measurements in April'99 is due to a pollution of the second FP cavity. After several weeks of out gassing, a noticeable agreement with the first laser, (within 5 Hz) has been obtained.

Several parameters yield frequency shifts of the OsO_4 center line relatively to the unperturbed molecular transition. The major biases which affect the CO_2/OsO_4 reference have been reported previously in ref. [5], with a preliminary evaluation of the total uncertainty of about 3.5×10^{-13}. Improvement of the accuracy capability to 10^{-14} level is realistic, but requires an optical selection of very slow OsO_4 molecules associated to heterodyne detection [9].

Fig 3 shows the square root of the Allan standard deviation associated to the frequency difference between the two CO_2/OsO_4 frequency standards stabilized independently onto the OsO_4 transition in coincidence with the R(10) CO_2 line.

This figure 3 exhibits a $7 \times 10^{-14}/\sqrt{\tau}$ slope up to τ = 100 s, indicating a white frequency noise and a good correction of the frequency fluctuations. Consequently, the Flicker plateau is reached for 100 s<τ< 500 s.

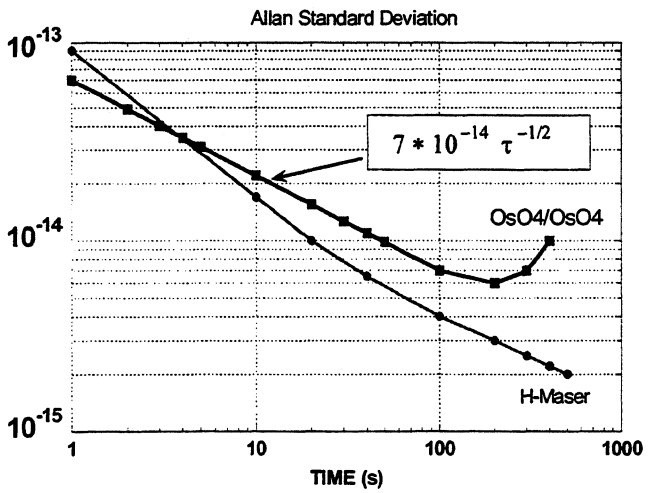


Fig. 3. Allan Standard Deviation of the beat note between two similar 10R(10) CO_2/OsO_4 systems (squares).

The long term stability is limited by the existence of a much more smaller absorption line on the hand side of the principal transition used for the stabilization purpose. This line shape has been reported elsewhere [9]. In spite of very small transition moment, the long term variations of the experimental parameters (pressure and power) have surely some influence on the long term stability.

This result brings us to check the metrological capabilities of the R(42)A1(3) fundamental transition of $^{192}OsO_4$.distant by 232.5 MHz from the center of the 10R(12) laser line. We have been able to achieve high saturation OsO_4 peak contrast (up to 36 %) using this transition.

Therefore, the short-term stability of the lasers has been enhanced in considerable way, reaching the $4.5\ 10^{-15}$ level from 10 to 40 s of integration time, as shown on Fig. 4.

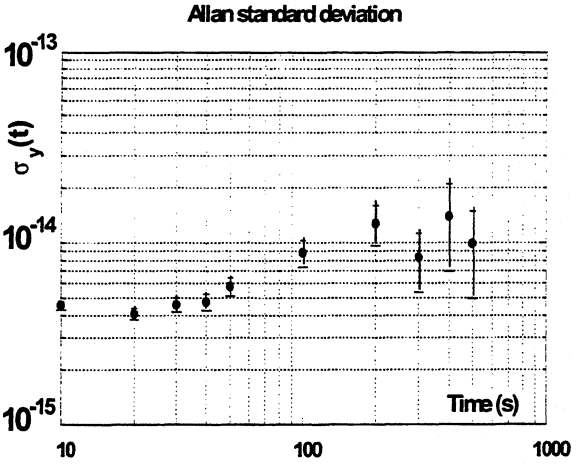


Fig. 4 Allan standard deviation of the CO_2/OsO_4 frequency standards, using the R(42) A1(3) $^{192}OsO_4$ absorption line.

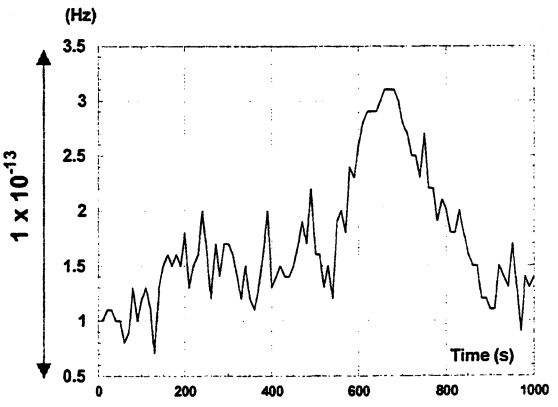


Fig. 5 Frequency drift of the frequency difference between two CO_2/OsO_4 standards independently locked using the R(42) A1(3) $^{192}OsO_4$ absorption line.

The long-term stability has been altered by the OsO_4 pressure variations mentioned above in this paper, and which induce frequency changes of the stabilized lasers as shown on Fig. 5. We expect more stable frequency evolution in the near future as observed before [5], after more complete out gassing of the FP cavity.

4. CONCLUSION

The CO_2/OsO_4 exhibit very attractive metrological characteristics as optical frequency standards, with short-term and long term stabilities in the 10^{-15} range and a reproducibility at the level of 1×10^{-13} . These results emphasize their interest in various fields, such as accurate spectroscopy and extension of frequency measurements to the visible domain. The recent and preliminary use of the fundamental and strong absorption line in coincidence with the 10 R(12) line at 29.096 THz, opens the way for progress in the accuracy capability to the 10^{-14} level.

5. ACKNOWLEDGEMENTS

This work is a part of the scientific developments of the Bureau National de Métrologie (BNM-France). We are indebted to the electronic staff for the renewal of the electronics stabillization devices of the lasers. Particular thanks are due to A. H. Gérard for the help in the course of this work.

6. REFERENCES

[1] O. Acef, J. J. Zondy, M. Abed, D. G. Rovera, A. H. Gérard, A. Clairon, Ph. Laurent, Y. Millerioux and P. Juncar, "A CO2 to visible optical frequency synthesis chain : accurate measurement of the 473 THz He-Ne/I2 laser", Optics Comm., Vol; 97, p. 29 (1993).

[2] D. Touahri, O. Acef, A. Clairon, J. J. Zondy, R. Felder, L. Hilico, B. de Beauvoir, F. Nez and F. Biraben, "Frequency measurement of the 5S1/2-5D5/2 two-photon transition in rubidium", Optics Comm., Vol. 133, p. 471 (1997).

[3] B. de Beauvoir, F. Nez, L. Julien, F. Biraben, D. Touahri, L. Hilico, O. Acef, A. Clairon, and J. J. Zondy, "Absolute frequency measurement of the 2S-8S/D transitions in hydrogen and deuterium: new determination of the Rydberg constant", Phys. Rev. Lett. , Vol. 78, p. 440 (1997).

[4] O.Acef,"Accurate frequency measurements in the infrared using stabilized CO_2/OsO_4 lasers", presented at the CPEM'98 conference, Washington D.C., USA, July 6-10, 1998.

[5] O. Acef, « *Metrological properties of CO2/OsO4 optical frequency standard* », Optics Comm. **134**, p. 479, (1997).

[6] Ch. Chardonnet and Ch. J. Bordé, "*Hyperfine interactions in the v3 band of Osmiom tetroxyde: Accurate determination of the spin-rotation constant by crossover resonance spectroscopy*", J. of Mol. Spect. Vol. 167 p. 71-98 (1994).

[7] A. Clairon, B. Dahmani, O. Acef, M. Granveaud, Yu. S. Domnin, S. B. Pouchkine, V. M. Tatarenkov, and R. Felder, «Recent experiments leading to the characterisation of the performance of portable He-Ne/CH4 lasers », Metrologia **25**, p. 9 (1988).

[8] G. D. Rovera, O. Acef, and A. Clairon, "Absolute frequency measurement in the infrared", Joint Meeting of the 13th European frequency and time forum and 1999 IEEE international frequency control symposium, Besançon, France 13-16 April, 1999. this issue.

[9] A. Clairon, O. Acef, Ch. Chardonnet and Ch. J. Bordé, "State-of-the-art for high accuracy frequency standards in the 28 THz range, using saturated absorption resonances of OsO4 and CO2", in Proceedings of *Frequency Standards and Metrology*, A. De Marchi (Ed.), Berlin, Springer Verlag (1989) p. 212.

1999 Joint Meeting EFTF - IEEE IFCS

SINGLE ION SPECTROSCOPY OF YTTERBIUM

M Roberts, P Taylor and P Gill

National Physical Laboratory, Teddington, Middlesex TW11 0LW, UK

ABSTRACT

Recent work is presented on the search for the weak electric octupole transition in a single cold ion of the odd-isotope $^{171}Yb^+$, and includes new $^{171}Yb^+$ spectroscopy on two other reference transitions in the optical and infra-red regions, which are used to locate the octupole transition. A successful observation of this $^{171}Yb^+$ $^2S_{1/2}$ - $^2F_{7/2}$ reference transition is presented.

INTRODUCTION

A forbidden transition in an ytterbium ion, is one of a handful of reference systems that are being considered as highly reproducible optical frequency standards. A single laser cooled ytterbium ion held in an electrodynamic trap is a near ideal atomic sample for a frequency standard. The quantum jump technique is used to observe an extremely forbidden transition in the ion. Tight confinement of the ion in the Lamb-Dicke regime, by the trap, removes the first order Doppler effect and recoil shift. Laser cooling and careful control of the ion's motion minimises the second order Doppler and Stark shift. Operation under ultra-high-vacuum conditions reduces pressure shifts to a negligible level.

Past work has concentrated on observing the extremely weak $^2S_{1/2}$ - $^2F_{7/2}$ reference transition at 467 nm. The upper level of this electric octupole transition has a measured lifetime of 10 years, making it potentially the narrowest known resonance in the optical spectrum. The difficulty in observing such a weak transition, led to its first observation in the technically straightforward $^{172}Yb^+$ isotope [1].

The 172-isotope is, however, not ideal as a frequency standard as the reference frequency is susceptible to the linear Zeeman effect. This has necessitated a switch to the $^{171}Yb^+$ isotope, which has $m_F = 0$ - $m_F = 0$ transitions that are free from the linear Zeeman effect. This reduces the effect of perturbing magnetic fields to an almost negligible level. Due to the presence of hyperfine structure and the difficulty in performing accurate calculations of isotope shifts, it has been necessary to undertake high resolution spectroscopy in this $^{171}Yb^+$ isotope so that the extremely weak 467 nm reference transition can be located.

LASER COOLING OF $^{171}Yb^+$

A single ion of $^{171}Yb^+$ was confined in a miniature rf Paul trap comprising a 1 mm ring and two end-cap electrodes by means of an AC drive voltage of 350 V peak - peak. An additional DC voltage applied to the ring equalised the axial and radial secular frequencies at 1.1 MHz. The rf-photon correlation technique was used to minimise the ion's micromotion in the direction of the laser cooling beam. This arrangement was used to perform the spectroscopy on the 411 nm $^2S_{1/2}$ - $^2D_{5/2}$ quadrupole transition and the 3.4 μm $^2F_{7/2}$ - $^2D_{5/2}$ infrared transition. More recently a new trap with end-cap electrode geometry has been developed. This trap has a similar rf drive frequency of $2\pi \times 12.8$ MHz. Most importantly, the trapped ion can be accessed in 3 directions by the cooling laser light, so that the rf-photon correlation technique can be used to reduce the ion micromotion in 3 dimensions, thereby minimising second order Doppler and Stark shifts. This new trap geometry was used for the search for the $^{171}Yb^+$ octupole transition. Both traps had background pressures of less that 10^{-10} mbar, and typical ion storage times of weeks to months are routinely achieved.

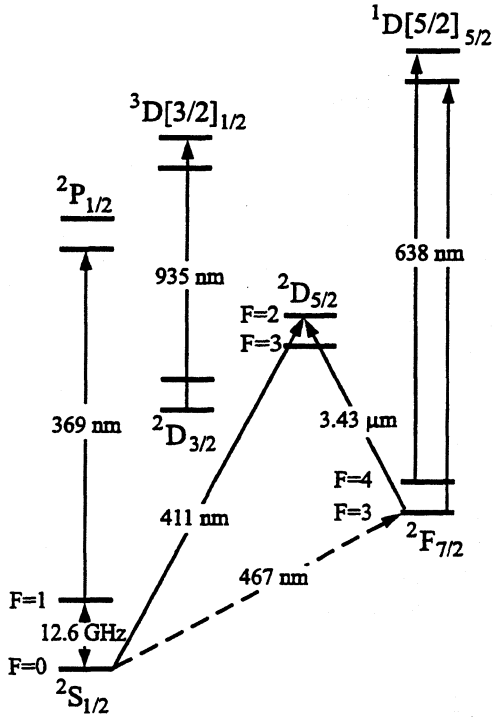


Figure 1: $^{171}Yb^+$ Term Scheme

A partial term scheme for $^{171}Yb^+$ is shown in figure 1. The 171-isotope of Yb^+ has spin ½, which causes hyperfine doublets. The single ion is laser cooled by repeatedly driving the F=1 - F=0 component of the $^2S_{1/2}$ - $^2P_{1/2}$ transition with laser radiation at 369 nm. From the $^2P_{1/2}$ (F=0) state there is a small possibility for decay into the metastable $^2D_{3/2}$ (F=1) state. To maintain the cooling cycle, the $^2D_{3/2}$ level is rapidly depopulated by a laser at 935 nm, returning the ion to the F=1 ground state via the $^3D_{[3/2]1/2}$ (F=0) level. This cooling scheme

0-7803-5400-1/99/ $ 10.00

forms a closed loop, but non-resonant excitation of the $^2P_{1/2}(F=1)$ state by the cooling laser leads to optical pumping of the F=0 ground state [2]. In order to maintain effective cooling, the F=0 ground state must be coupled to the cooling cycle. This is achieved by driving the F=0 - F=1 ground state transition with 1 watt of microwave radiation at 12.6 GHz. A maximum fluorescence count rate of 12 kHz was obtained.

For the $^{171}Yb^+$ reference transitions, it is desirable to work at low magnetic field in order to minimise the second order magnetic field contribution to the $m_F=0$ - $m_F=0$ transitions. However, in zero magnetic field no fluorescence is observed from the cooling transition. This is due predominantly to coherent population trapping in the $m_F=\pm1$ sub-levels of the F=1 ground state. A significant field of 1mT is necessary to fully restore the fluorescence, about two orders of magnitude larger than that needed in the even isotope $^{172}Yb^+$ case, where the population trapping only occurs in the $^2D_{3/2}$ level. As a result, the large 1 mT magnetic field is switched off in conjunction with the cooling radiation during the period when the reference transition is probed, in order to minimise AC Stark and Zeeman perturbations. A small μT field remains during this probe period.

411nm $^2S_{1/2}$ (F=0, m_F=0) - $^2D_{5/2}$ (F=2, m_F=0) REFERENCE TRANSITION

The $^2S_{1/2}$ - $^2D_{5/2}$ 411 nm quadrupole transition profile is observed by monitoring the quantum jumps in the fluorescence as a function of 411 nm probe laser frequency. The 411 nm probe light is generated using an 822 nm extended cavity diode laser FM-stabilised to an evacuated ULE reference cavity, and frequency doubled in a lithium triborate crystal within a resonant enhancement cavity. Frequency scanning through the reference transition is achieved by a double passed acousto-optic modulator between laser and ULE cavity. Due to the decay from the $^2D_{5/2}$ level to the highly metastable $^2F_{7/2}$ level, it is necessary to drive additional transitions at 638 nm from this $^2F_{7/2}$ level in order to recover the fluorescence and observe repeated jumps. A cavity-stabilised 638 nm extended cavity diode laser is used for this, and is switched in frequency by a few GHz in order to depopulate both F=3 and F=4 hyperfine levels of the F-state [3].

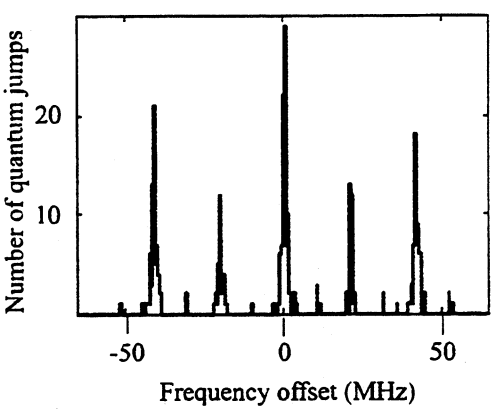


Figure 2: $^{171}Yb^+$ $^2S_{1/2}(F=0)$ - $^2D_{5/2}(F=2)$ Transition

The quantum jump line profiles of the full Zeeman structure of all three possible hyperfine transitions F=0 - F=2, and F=1 - F=2,3 at 411 nm were recorded at an applied magnetic field of 1 mT. The simplest transition is the $^2S_{1/2}(F=0)$ - $^2D_{5/2}(F=1)$, which has five well resolved Zeeman components shown in figure 2. The central component is the linear-Zeeman-effect free $m_F=0$ - $m_F=0$ transition, and is therefore suitable as a reference. A higher resolution scan of this $m_F=0$ - $m_F=0$ transition is shown in figure 3.

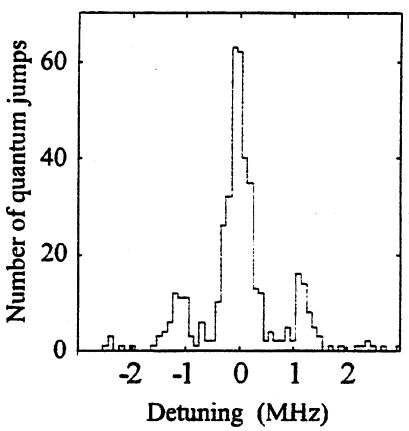


Figure 3: $^{171}Yb^+$ 411 nm (F=0, m_F=0) - (F=2, m_F=0) Transition

The transition is still susceptible to the second order Zeeman effect, and for best reproducibility, this shift needs to be taken account of. The shift was measured by plotting the transition jump profile line-centre as a function of magnetic field up to a few mT. The expected linear frequency shift with square of the magnetic field is obtained, with a measured value of 0.38 (8) Hz / $(\mu T)^2$.

The absolute frequency of the $m_F=0$ - $m_F=0$ Zeeman component was obtained by measuring the frequency of the ULE fringe to which the 822 nm laser was stabilised, coupled with a measurement of the AOM frequency used to tune the 411 nm doubled light across the quantum jump profile. The ULE fringe frequency was measured by means of the NPL 1 m evacuated Fabry-Perot comparator, with an accuracy of comparison of 2 in 10^{10}. The value of the transition was 729 487 779 566 (153) kHz.

The hyperfine splitting of the $^2D_{5/2}$ level was measured by scanning the 411 nm radiation between the F=1 - F=2 and F=1 - F=3 transitions, while the 822 nm fundamental was locked to the same ULE fringe. A hyperfine splitting of 191(2) MHz was obtained. In addition, the $^{171}Yb^+$ - $^{172}Yb^+$ $^2S_{1/2}$ - $^2D_{5/2}$ isotope shift of +1317.1(1.3) MHz was determined by combining the weighted splittings of the $^2S_{1/2}$ and $^2D_{5/2}$ levels with the absolute value of the F=0 - F=2 transition, and then comparing this value with the $^2S_{1/2}$ - $^2D_{5/2}$ value previously measured in $^{172}Yb^+$ [4].

3.43 μm $^2F_{7/2}$ (F=3) - $^2D_{5/2}$ (F=2) IR REFERENCE TRANSITION

The 3.43 μm $^2F_{7/2}$ - $^2D_{5/2}$ $^{171}Yb^+$ transition has particular significance as a mid infra-red frequency standard due to its close proximity (<1 THz) to the 3.39 μm methane-stabilised He-Ne laser frequency. With a better Q of $4x10^{12}$ and a low level of systematic shifts, it should prove a better choice as an intermediate standard within microwave-optical frequency chains [5].

The lower level of the 3.43 μm transition is the long-lived $^2F_{7/2}$ metastable level and thus the ion needs to be prepared in this level prior to driving the transition. The ideal state preparation technique in the future could be to prepare the ion in the $^2F_{7/2}$ (F=3, m_F=0) state by directly driving the 467nm $^2S_{1/2}$ - $^2F_{7/2}$ octupole transition. In our current situation, however, an absolute frequency measurement of the 3.43 μm transition is needed to supplement the 411 nm absolute frequency in order to determine the location of this weak octupole transition. As a result F-state preparation is made by driving the 411 nm $^2S_{1/2}$ - $^2D_{5/2}$ transition, followed by spontaneous decay from the $^2D_{5/2}$ level. This has a lifetime of 7 ms, and a branching ratio to the $^2F_{7/2}$ level of 83% [4]. Thus, if the fluorescence has not returned after some 70 ms after driving with 411 nm radiation, the ion is deduced to be in the $^2F_{7/2}$ level with high probability. One disadvantage of this state preparation technique is that there is no control over which $^2F_{7/2}$ (F=3, m_F) sub-level is populated, and thus one cannot selectively drive the m_F=0 - m_F=0 3.43 μm transition. However, by making use of a relatively broad 3.43 μm laser linewidth (~300 kHz), and recognising that we only need to measure the 3.43 μm transition at the part in 10^{10} level in order to locate the octupole transition, this is not a limitation if the magnetic field is small enough to allow all the Δm_F=0 components to contribute. A field of 14 μT was used during the 3.43 μm interrogation cycle. The 3.43 μm line profile is built from the number of quantum ON jumps at each frequency step. When no ON jump is recorded after ~70 msec, the ion is recovered to the $^2S_{1/2}$ ground state by driving the 638 nm transitions.

The 3.43 μm radiation is generated by difference mixing of 1064 nm and 812 nm radiations in a crystal of $AgGaS_2$ within a 1064 nm resonant enhancement cavity [5]. About 500 mW of single frequency 1064 nm light and 30 mW of single frequency 812 nm light from an extended cavity laser gave rise to 5 μW at 3.43 μm in a few hundred kHz bandwidth, dominated by the diode linewidth. The experimental layout is shown in figure 4.

The absolute frequency of the $^2F_{7/2}$ (F=3) - $^2D_{5/2}$ (F=2) hyperfine transition was measured via individual measurements of the 812 nm and 1064 nm laser frequencies. The 812 nm frequency was measured by interferometric comparison with an iodine-stabilised

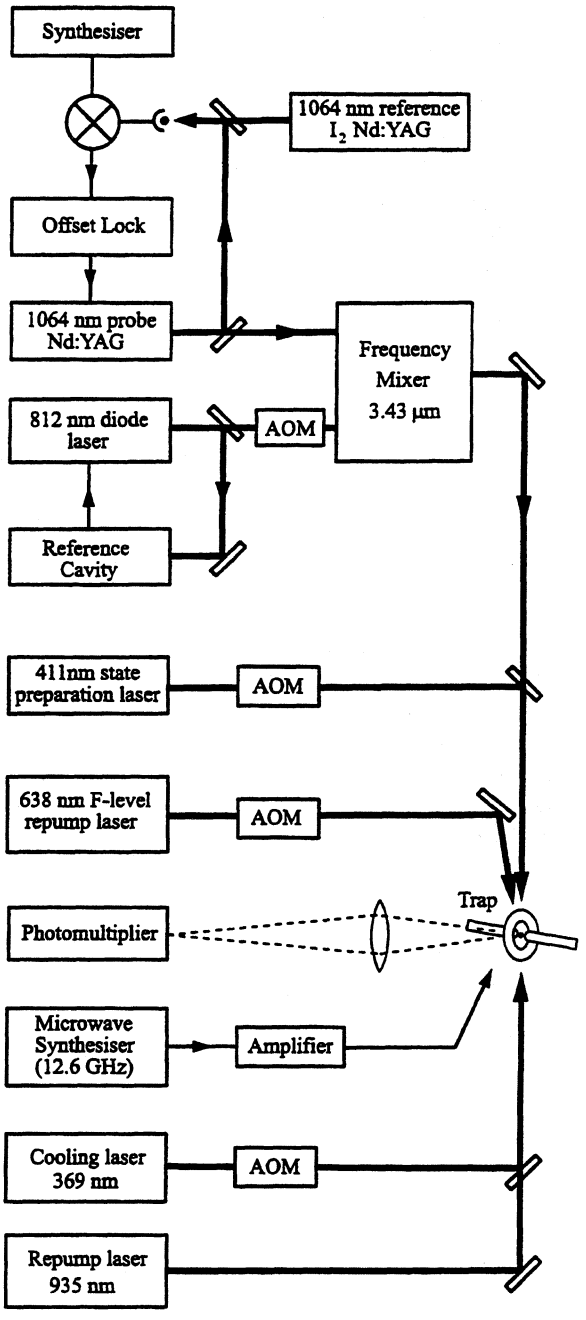


Figure 4: 3.43 μm Experimental Layout

He-Ne reference, using the NPL comparator. The 812 nm laser was stabilised to an evacuated ULE reference cavity. The 1064 nm laser frequency was measured by beating against the fundamental of a frequency-doubled Nd:YAG reference laser, where the 532 nm output was stabilised to the a_1 component of the 32 - 0 P(54) reference transition in $^{127}I_2$. The accuracy of the 1064 nm frequency was thus ~2 in 10^{10}. This frequency was offset-locked against the fundamental Nd:YAG reference under control from a scanned synthesiser frequency. The $^{171}Yb^+$ 3.43 μm $^2F_{7/2}$ (F=3) - $^2D_{5/2}$ (F=2) transition comprises five unresolved Δm_F=0 components, with a total unresolved splitting of 90 kHz in the 14 μT field. The transition frequency was measured as 87 366 282 258 (88) kHz [6].

An additional measurement was made of the F=4 - F=3 hyperfine component. When combined with the $^2D_{5/2}$ hyperfine interval measurement, this gives a hyperfine $^2F_{7/2}$ splitting of 3620 (2) MHz. The $^2F_{7/2}$ and $^2D_{5/2}$ hyperfine splittings can then be used to determine the $^{171}Yb^+$ 3.43 μm centroid frequency, and by comparison with the previous measurement of the $^{172}Yb^+$ 3.43 μm transition, a 3.43 μm $^{171}Yb^+$ - $^{172}Yb^+$ isotope shift of +4048 (4) MHz is deduced.

$^{171}Yb^+$ $^2S_{1/2}$ (F=0) - $^2F_{7/2}$ (F=3) 467 nm OCTUPOLE TRANSITION

The difference in measured values for the $^{171}Yb^+$ 411 nm and 3.43 μm transitions corresponds to the deduced location of the $^2S_{1/2}$ - $^2F_{7/2}$ octupole transition. This difference is 642 121 497 308 (176) kHz. The uncertainty of this location is about a factor x50 better than the original uncertainty derived for the $^{172}Yb^+$ octupole transition [1].

A search for the $^{171}Yb^+$ $^2S_{1/2}$ - $^2F_{7/2}$ octupole transition has been carried out using a new design end-cap trap. The probe laser was a frequency-doubled 934 nm Ti-Sapphire laser, the doubled blue light being acousto-optically tuned from a reference tellurium vapour transition positioned some 502 MHz from the projected octupole position.

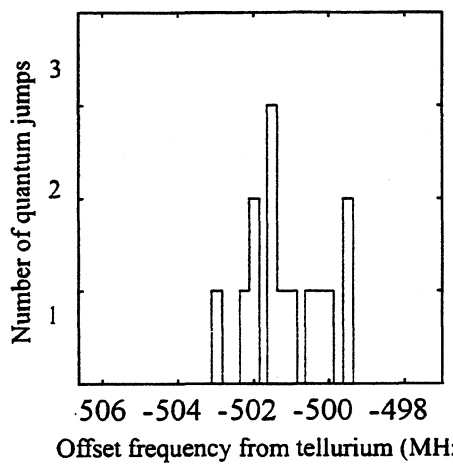


Figure 5: $^{171}Yb^+$ $^2S_{1/2}$ - $^2F_{7/2}$ (m_F=0 - m_F=0) Octupole Transition

A quantum jump profile of the octupole transition has now been observed by scanning a 10 MHz search range in the first instance, and 14 OFF jumps were observed over a period of 9 hours, which included 7680 s of interrogation. After each jump the ion was recovered to the ground state and fluorescence re-established by driving the 638 nm transitions. The data is shown in figure 5. The relatively few events observed so far, and the 1 MHz blue laser linewidth, contribute to a transition with unresolved secular sidebands. However the centre frequency is determined to be 642 121 498.1 (1.2) MHz, in excellent agreement with the deduced location value.

CONCLUSION

Absolute frequency measurements of the $^{171}Yb^+$ $^2S_{1/2}$ - $^2D_{5/2}$ 411 nm and $^2F_{7/2}$ - $^2D_{5/2}$ 3.43 μm reference transitions have been obtained. These transitions have theoretical Qs of $3x10^{13}$ and $4x10^{12}$ respectively. Their frequency values have been used to locate the $^{171}Yb^+$ $^2S_{1/2}$ - $^2F_{7/2}$ 467 nm octupole transition. A search at this location has resulted in successful observation of this $^{171}Yb^+$ octupole transition. Future work on this transition will concentrate on reducing the probe laser linewidth, and hence the transition drive rate. This should set the basis for a $^{171}Yb^+$ single ion octupole optical frequency standard.

REFERENCES

[1] M Roberts, P Taylor, GP Barwood, P Gill, HA Klein and WRC Rowley, "Observation of an electric octupole transition in a single ion", Phys. Rev. Lett., vol 78, pp 1876-9, March 1997.

[2] D Engelke and C Tamm, "Dark times in the resonance fluorescence of trapped $^{171}Yb^+$ ions caused by spontaneous quantum jumps to the $^2D_{3/2}$ (F=2) state", Europhys. Lett., vol 33, p 347, 1996

[3] M Roberts, P Taylor, SV Gateva-Kostova, RBM Clarke, WRC Rowley and P Gill, "Measurement of the $^2S_{1/2}$ - $^2D_{5/2}$ clock transition in a single $^{171}Yb^+$ ion", Phys. Rev A, 1999 (Submitted)

[4] P Taylor, M Roberts, SV Gateva-Kostova, RBM Clarke, GP Barwood, WRC Rowley and P Gill, "Investigation of the $^2S_{1/2}$ - $^2D_{5/2}$ clock transition in a single ytterbium ion", Phys. Rev. A, vol 56, pp 2699-2704, October 1997

[5] P Taylor, M Roberts, GP Barwood and P Gill, "Combined optical-infrared single-ion frequency standard", Opt. Lett., vol 23, pp 298-300, February 1998

[6] P Taylor, M Roberts, GM MacFarlane, GP Barwood, WRC Rowley and P Gill, "Measurement of the infrared $^2F_{7/2}$ - $^2D_{5/2}$ transition in a single $^{171}Yb^+$ ion", Phys. Rev. A, 1999 (Submitted)

Cs-REFERENCED OPTICAL FREQUENCY MEASUREMENT OF THE SINGLE, TRAPPED Sr^+ ION STANDARD AT 445 THz

J.E. Bernard, A.A. Madej, L. Marmet, K.J. Siemsen, and B.G. Whitford
National Research Council of Canada

Institute for National Measurement Standards
National Research Council, Ottawa, Ontario, Canada K1A 0R6

ABSTRACT

The frequency of the electric quadrupole allowed 5s $^2S_{1/2}$ – 4d $^2D_{5/2}$ transition at 445 THz in a single, trapped and laser cooled $^{88}Sr^+$ ion has been measured with an optical frequency chain referenced to a Cs atomic clock. A transition frequency of f_{S-D} = 444 779 044 095.4 kHz with an estimated standard uncertainty of 0.2 kHz has been determined. The intrinsic offsets of the measured ion transition frequency due to environmental perturbations have been calculated to be at the 10^{-15} level.

1. INTRODUCTION

Atomic frequency standards in the visible are important for time and length metrology, fiber-optic communications, the measurement of fundamental constants and tests of basic atomic theory. In recent years, advances in precision laser spectroscopy has led to the development of optical frequency standards with accuracies approaching those obtainable with the best microwave standards [1]. However, systematic frequency shifts due to collisions, distributions in velocity, and the confining electromagnetic fields currently limit the accuracy of these standards.

The ultimate frequency standard, in terms of accuracy, is probably one based on a narrow electronic transition in a single atom or ion which is isolated and at rest. Using the methods of single ion trapping and laser cooling, it is now possible to localize a single atomic particle to a submicron volume under an ultra-high vacuum almost indefinitely and reduce its kinetic energy to the milli-Kelvin level. The effects of collisions, interrogation time limits, first and second order Doppler shifts, and external fields and field gradients can be reduced to practically negligible levels. In addition, through the method of quantum jumps, it is possible measure the resulting narrow transition spectrum with almost 100% efficiency.

For a number of years, our group at NRC has studied the narrow electric quadrupole allowed 5s $^2S_{1/2}$ – 4d $^2D_{5/2}$ transition at 445 THz (674 nm) in the single trapped $^{88}Sr^+$ ion. This transition has a natural linewidth of approximately 0.4 Hz and was recently selected by the Comité International des Poids et Mesures (CIPM) as a recommended optical frequency for the realization of the metre [2]. Its currently recommended value, based on previous experiments [3,4], is 444 779 044.04 MHz with a standard uncertainty of ±0.06 MHz. The object of the work reported here was to remeasure this transition frequency with the highest available accuracy by means of a Cs-based frequency chain. This is the first reported Cs-based measurement of a visible transition frequency in a single ion [5] and only the second reported Cs-based chain measurement to reach the visible part of the spectrum [6].

2. EXPERIMENT

Figure 1 shows the energy level diagram for the lowest electronic states of the Sr^+ ion. In our experiment, a frequency-doubled diode laser at 422 nm is used for laser cooling and quantum-jump detection on the 5s $^2S_{1/2}$ – 5p $^2P_{1/2}$ resonance transition. An auxiliary diode-pumped, Nd-fiber laser at 1092 nm is used to pump the ion out of the 4d $^2D_{3/2}$ state since the 5p $^2P_{1/2}$ state has a 1:13 branching ratio to this level. The 5s $^2S_{1/2}$ – 4d $^2D_{5/2}$ "clock" transition is split into 10 Zeeman components [3] and is probed by a portion of the output power of an ultra-stable diode laser [7] which is shifted by a double-passed, acousto-optic modulator (AOM) in order to allow scanning of the S-D spectrum. The single Sr^+ ion is held in an rf Paul trap described elsewhere [7,8]. Stray electric fields are partially compensated by applying a differential bias to the trap electrodes. A single-layer magnetic shield reduces the effects of slowly varying dc magnetic fields and ac broadening on the positions and widths of the Zeeman components. A small magnetic field of approximately 14 μT is applied in the region of the trap by pairs of coils driven by a low-noise current source resulting in a splitting of the two inner Zeeman components of approximately 156 kHz. The current source is highly stable and the splitting has been found to remain constant to better than 50 Hz over a period of 2000 s [9].

0-7803-5400-1/99/ $ 10.00

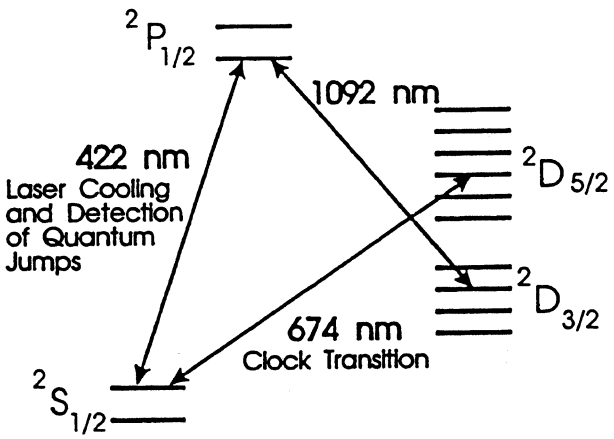


FIG. 1. Energy level of the Sr^+ ion.

In previous work [9] employing an optimized diode laser source, linewidths on the level of 250 Hz were obtained for a particular Zeeman component. In the present studies using a newer replacement diode laser, scans of a single component over periods of the order of a minute showed a linewidth of 500 to 1000 Hz, almost entirely due to the probe laser. A locking technique based on measurements of the quantum jump rate at four frequencies, one on each side of each of the two inner Zeeman components was used to control the AOM frequency [9]. Lock cycle periods were typically 20 s and drifts in the ultra-stable laser of up to 5 Hz/s could easily be followed by the lock. A correction term due to the temporal lag in the lock was calculated [9] from the measured drift rate and added to the measured offset of the ultrastable laser from the centre of the S-D transition manifold. This correction never exceeded 100 Hz and was usually only a few tens of Hz.

The frequency chain shown in Fig. 2 was used to measure the frequency of the ultra-stable laser and, from simultaneous measurements of the offset of the laser from the centre of the S-D spectrum, to determine the Sr^+ transition frequency. These measurements are reported in a forthcoming publication [5]. Parts of this chain have been described in previous publications. The NRC phase-locked infrared chain [10] was used to phase-lock the frequencies of four CO_2 lasers (A, B, C, and D) and two microwave oscillators (X and V) to a 5 MHz signal from a hydrogen maser which was referenced to an ensemble of three Cs standards. A Tm:YAG laser at 148 THz (laser F) [11] was phase-locked to the ultra-stable laser through an optical divide-by-3 system [12]. A counter monitored this lock. The frequency of the Tm:YAG laser was measured by mixing it with an auxiliary CO_2 laser (laser E) on a tungsten-nickel, point-contact (MIM) diode and counting the frequency of a tracking oscillator which was phase-locked to the resulting beat, $F-5E \simeq 1525$ MHz. Finally, the frequency of laser E was measured through the beat $E + A - 2C - 3V - X \simeq 2760$ MHz, also obtained with a MIM diode. Attempts to phase-lock laser E through this beat, which would result in a phase-coherent chain, were only partly successful due to excess noise in this laser. Therefore, most of the measurements were obtained by counting the frequency of a tracking oscillator which was phase-locked to the beat at 2760 MHz. The two counters were synchronized to within 0.1 s through software triggering. Since the drift rates were only a few Hz/s better synchronization of the counters was unnecessary for the present level of accuracy.

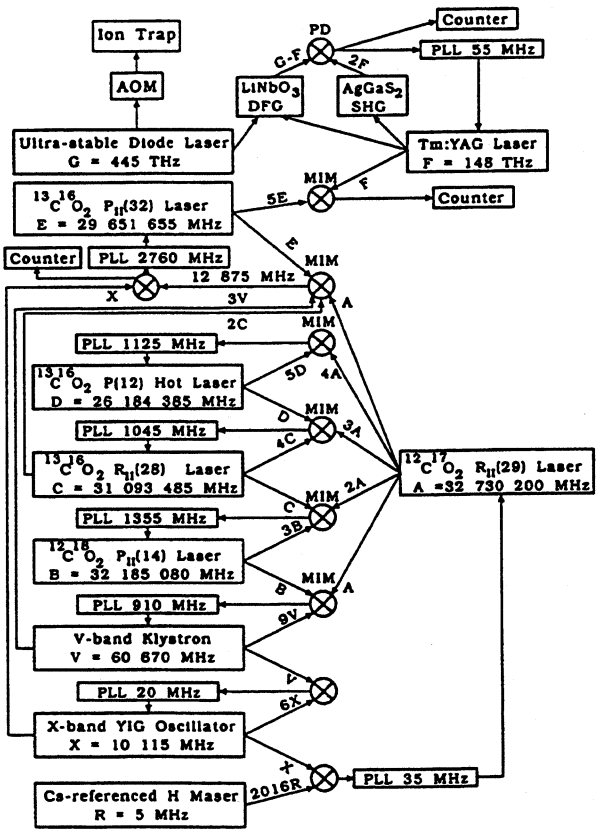


FIG. 2. The frequency chain used in the measurement of the frequency of the S-D transition in the $^{88}Sr^+$ ion.

Chain measurements were made on four separate days ranging over almost 15 months. Only the data of 11-Aug-97 was obtained with the entire chain phase-locked, for a total period of 21 s. On 1-Sep-98, 3-Sep-98, and 11-Nov-98 laser E was frequency locked to the IR chain and the beat frequency counted for integrated measurement periods of 3260 s, 6880 s and 1100 s, respectively. The duration of some continuous experimental runs exceeded

10 minutes. Gradual improvements in the S/N of the various beats resulted in tighter locking of the infrared chain and smaller fluctuations in the measured laser frequency. The observed 1-s Allan deviation of the chain-measured probe laser frequency ranged from 1×10^{-11} on 1-Sep-98 to 3×10^{-12} on 11-Nov-98.

The chain measurements were averaged in bins corresponding to the 20-s measurement cycles of the lock to the ion transition. Each bin resulted in one transition frequency measurement sample and a weighted mean (a weighting factor was applied which was inversely proportional to the square of the standard deviation) of these samples was determined for each experimental run. On the best runs, the standard deviation of the mean transition frequency was less than 60 Hz. The distribution of samples for each particular day was observed to be Gaussian having HWHM's of 1300 Hz for 1-Sep-98, 680 Hz for 3-Sep-98 and 450 Hz for the 11-Nov-98 data. A less tight chain lock and the small number of measurements on 11-Aug-97 resulted in an uncertainty for that day of 3000 Hz. Examination of the Allan deviation showed that the fluctuations decreased as $\tau^{-1/2}$ for averaging times, τ up to 100 s and stability floors of 1.4×10^{-12}, 6×10^{-13}, and 4.5×10^{-13} were observed for 1-Sep-98, 3-Sep-98, and 11-Nov-98, respectively. These limits were likely the result of either drifts in the lock of the AOM-shifted probe laser to the ion, possibly due to changes in the probe laser spectrum, or long term drifts in the chain, possibly due to slow drifts in the amplified and distributed reference signal. From previous experiments, cycle slips in the phase-locks are expected to be insignificant at the current level of precision.

The average transition frequency for each day was determined from the weighted mean of the 20-s samples and the uncertainty in the mean value was determined from the respective stability floors. Sample values which fell into a region having a probability of less than half an event in the fitted Gaussian distribution (Chauvenet Criterion) were not included in the calculation of the mean for a day's set of readings. These outliers were small in number (2–4% of the total set) and their exclusion shifted the mean value by only 10 to 20 Hz. The mean ion frequency for each day's run is plotted in Fig. 3. The S-D transition frequency determined from a weighted average of these 4 values is, f_{S-D} = 444 779 044 095.4±0.2 kHz (1σ). Also shown in Fig. 3 is the frequency determined in a previous measurement at NRC which employed a chain using difference frequency generation (DFG) between transfer standards at 474 THz (HeNe/I_2 laser) and 29 THz (CO_2/OsO_4 laser) [4].

3. FREQUENCY SHIFTS DUE TO THE ION

Any systematic shifts in the ion transition frequency due to its environment are much smaller than the measurement uncertainty of ±200 Hz. These shifts are summarized in Table 1.

If the ion is confined to a region of space with linear dimensions much smaller than the wavelength of the probing radiation (Lamb-Dicke regime), the absorption spectrum shows a carrier frequency which

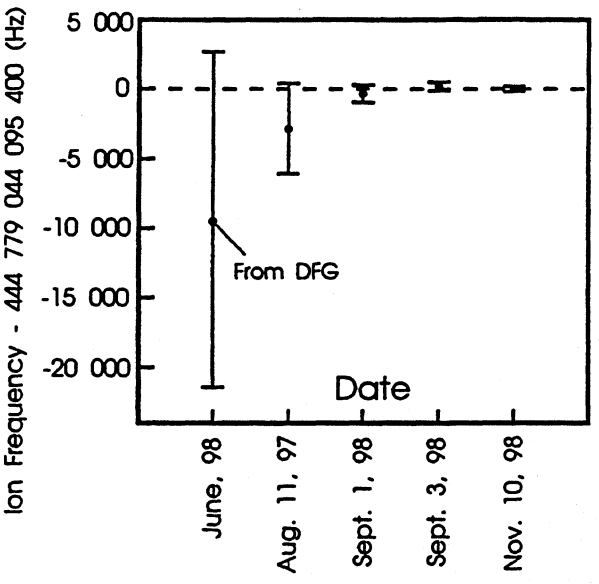


FIG. 3. The frequency of the Sr^+ ion S-D transition frequency as measured on each of the experimental days. The dashed line represents the average value of the Cs-based chain measurements (f_{S-D} = 444 779 044 095.4 kHz). Also shown is the average value from the difference frequency generation measurement.

is free of first order Doppler broadening and sidebands whose intensities yield information on the amplitude of the secular and driven oscillations of the ion in the trap potential well. Measurements of the

TABLE I. Ion related sources of systematic uncertainty and their calculated values in the present experiment.

Source	Line Center Shift	Magnitude
2nd Order Doppler Effect	0.13 Hz	3×10^{-16}
Quadratic Stark Shift	0.2 Hz	5×10^{-16}
Electric Quadrupole Shift of 4d $^2D_{5/2}$ level	< 0.5 Hz	$<1\times10^{-15}$
Blackbody ac Stark Shift	0.16 Hz	4×10^{-16}
ac Magnetic Fields	< 0.2 Hz	$<5\times10^{-16}$
Quadratic Zeeman Shift (static field)	15 mHz	3×10^{-17}
Collisions	< 10 mHz	$<2\times10^{-17}$

sideband intensities in our experiment indicate that the Sr^+ ion had a kinetic temperature of (15±5) mK and exhibited no first order Doppler shifts. From sideband to carrier power measurements, the oscillation velocity and amplitude away from the zero-field node at the trap centre can also be calculated and used to determine the second-order Doppler (time-dilation shift) and quadratic Stark shifts. For our ion trap, incomplete compensation of stray, static electric fields forced the ion slightly out of the node and the rf-driven micromotion dominated.

Since the 4d $^2D_{5/2}$ level possesses a quadrupole moment which can interact with gradients in the trap fields, a quadrupole shift can occur. Conservative estimates of displacement of the Sr^+ ion away from the zero-field node at the trap centre, indicate that this shift is less than 0.5 Hz.

The room temperature thermal radiation which bathes the ion can also cause a Stark shift of the clock transition levels. Since the dipole-allowed transition frequencies which couple to the S and D states in Sr^+ are well above the peak of the room-temperature black-body curve, the shifts are due primarily to the time averaged electric field intensity. For Sr^+ at T=300 K, the calculated black-body shift is only 0.16 Hz.

The S-D transition spectrum of Sr^+ exibits a first order Zeeman shift and therefore it is possible for stray ac magnetic fields to produce an asymmetric broadening of the spectrum, leading to a shift. From measurements of the broadening of Zeeman components with different shift sensitivies, we estimate that any shift due to ac magnetic fields is below 0.2 Hz.

The two remaining perturbations listed in Table 1, quadratic Zeeman shifts and collisions, are calculated to be very small.

4. CONCLUSIONS

In summary, we have performed the first Cs-based frequency measurement of a visible transition in a single trapped and laser-cooled ion standard. The frequency of the 5s $^2S_{1/2}$ – 4d $^2D_{5/2}$ transition in the single trapped $^{88}Sr^+$ ion was determined to an accuracy of 0.2 kHz ($\delta\nu/\nu = 4.5 \times 10^{-13}$) which represents a factor of 100 improvement in the knowledge of the single Sr^+ ion reference frequency and makes this transition one of the best known in the visible region of the electromagnetic spectrum. Systematic shifts associated with the ion standard itself are at the 10^{-15} level and can be reduced by further reduction of ion micromotion, reduction of stray fields in the trap structure and further shielding.

We would like to acknowledge helpful discussions with G.R. Hanes, J.S. Boulanger, and S. Cundy together with the excellent technical support provided by R. Pelletier, B. Hoger, W. Cazemier, W. Boland, and T. Cassidy.

[1] F. Riehle, H. Schnatz, B. Lipphardt, G. Zinner, T. Trebst, and J. Helmke, "The Optical Calcium Frequency Standard," pp. 299-300, In: Nelson T.L. (ed) 1998 Conference on Precision Electromagnetic Measurements at Washington USA, July 6-10, 1998.

[2] BIPM Proc. Verb. Com. Int. Poids et Mesures, vol. 65, pp. 63-71, 1998.

[3] G.P. Barwood, P. Gill, H.A. Klein, and W.R.C. Rowley, "Clearly Resolved Secular Sidebands on the $^2S_{1/2}$ –^{2} $D_{5/2}$ 674-nm Clock Transition is a Single Trapped Sr^+ Ion," IEEE Trans. Instrum. Meas., vol. IM-46, pp. 133-136 (1997).

[4] A.A. Madej, K.J. Siemsen, L. Marmet, J.E. Bernard, and O. Acef, "Linking the 474 THz HeNe/ I_2 Standard to the 445-THz Single Sr^+ Trapped Ion Standard: Heterodyne Frequency Measurements Using an OsO_4 Stabilized 29 THz Laser System," IEEE Trans. Instrum. Meas. (to be published April, 1999).

[5] J.E. Bernard, A.A. Madej, L. Marmet, B.G. Whitford, K.J. Siemsen, and S. Cundy, "Cs-based Frequency Measurement of a Single, Trapped Ion Transition in the Visible Region of the Spectrum," Phys. Rev. Lett., vol. 82, pp. 3228-3231 (1999).

[6] H. Schnatz, B. Lipphardt, J. Helmcke, F. Riehle, and G. Zinner, "First Phase-Coherent Frequency Measurement of Visible Radiation," Phys. Rev. Lett., vol. 76, pp. 18-21 (1996).

[7] L. Marmet, A.A. Madej, K.J. Siemsen, J.E. Bernard, and B.G. Whitford, Precision Frequency Measurement of the $^2S_{1/2}$ - $^2D_{5/2}$ Transition of Sr^+ with a Diode Laser Locked to an Ultrastable Cavity," IEEE Trans. Instrum. Meas., vol. 46, pp. 169-173 (1997).

[8] A.A. Madej. and K.J. Siemsen, "Absolute Heterodyne Frequency Measurement of the $^{88}Sr^+$ 445-THz S-D Single Ion Transition," Opt. Lett., vol. 21, pp. 824-826 (1996).

[9] J.E. Bernard, L. Marmet, and A.A. Madej, "A Laser Frequency Lock Referenced to a Single Trapped Ion," Opt. Comm., vol. 150, pp. 170-174 (1998).

[10] B.G. Whitford, "Uncertainty in Frequency Measurements at 88 THz made with the NRC Frequency Chain: Frequency of the NRC HeNe/CH_4 Laser," Metrologia, vol. 30, pp. 145-154 (1993).

[11] J.E. Bernard, B.G. Whitford, and A.A. Madej, "A Tm:YAG Laser for Optical Frequency Measurements: Mixing 148 THz Light with CO_2 Laser Radiation," Opt. Comm., vol. 140, pp. 45-48 (1997).

[12] J.E. Bernard, B.G. Whitford, and L. Marmet, "Phase-Locked Optical Divide-by-3 System for Visible Radiation," Opt. Lett., vol. 24, pp. 98-100 (1999).

PROGRESS IN THE DEVELOPMENT OF AN OPTICAL FREQUENCY STANDARD AT 192.6 THZ BASED ON A TWO-PHOTON TRANSITION OF RUBIDIUM ATOMS AT LAVAL UNIVERSITY

C. Latrasse, M. Poulin, D. Touahri, M. Allard and M. Têtu.
Centre d'Optique, Photonique et Laser
Département de Génie Electrique et de Génie Informatique
Université LAVAL
Ste-Foy, PQ, Canada, G1K 7P4
Tel: (418) 656-2146
Fax: (418) 656-3159

ABSTRACT

We have developed frequency standards at 192.6 THz (1556.2 nm) based on the two-photon transitions in rubidium at 385.2 THz (778.1 nm). These standards use a high power DFB laser at 1556.2 nm and second harmonic generation (SHG) in a periodically poled lithium niobate (PPLN) cristal. The linewidth of the DFB is reduced to the kHz level using optical feedback from a confocal cavity. The SH light is used to injection-lock a 778.1 nm laser diode which allows to observe and lock the 1556.2 nm laser to the ($5S_{1/2}$, F_g=2 - $5D_{5/2}$, F_e=4) two-photon transition in ^{87}Rb. Allan variance measurements between two identical standards show a beat stability of $2.5 \times 10^{-13}/\tau^{1/2}$ for observation times between 100 ms and 10 s and a level of 5.8×10^{-14} for 100 s is obtained. These standards have been moved recently to the National Research Council, Ottawa, Canada in order to measure their absolute frequency.

1. PURPOSE OF THIS WORK

The motivation for developing these frequency standards is two-fold: first, it is becoming very important for the telecommunication industry to standardize the channels allocation in dense WDM (wavelength-division-multiplexing) systems [1]. Some very precise and stable reference standards are now required in the 1.5 μm band to establish worldwide standardization of the communication systems. The Rb two-photon transitions offer the best metrological properties to realize a high performance practical frequency standard [2,3]. They have a theoretical linewidth of only ~ 165 kHz (at 1.55 μm) and an absolute frequency known with an uncertainty of 2 kHz [4]. The second objective of our work is to make a measurement of the absolute frequency of the ($5S_{1/2}$, F_g=2 - $5D_{5/2}$, F_e=4) transition in ^{87}Rb in collaboration with the Institute for National Measurement Standards at the National Research Council in Ottawa. For this, our colleagues from INMS/NRC are developing a frequency chain between the 192.6 THz radiation of our standards and their Cs standard.

2. DESCRIPTION OF THE FREQUENCY STANDARD

At Laval University, we have developed two identical frequency standards at 1556.2 nm based on the use of the ($5S_{1/2}$, F_g=2 - $5D_{5/2}$, F_e=4) two-photon transition in ^{87}Rb [2], as shown on figure 1.

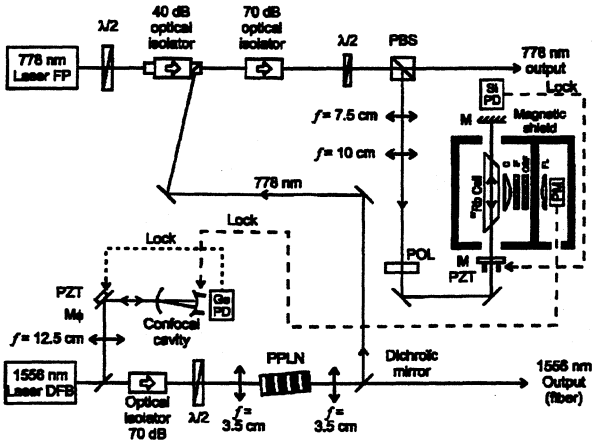


Figure 1: Schematic diagram of the 1556.2 nm frequency standard.

As the two-photon lines exhibit a very narrow linewidth, we first reduce the natural linewidth (a few MHz) of the DFB laser using weak optical feedback from a confocal cavity. The second harmonic of the laser generated in a periodically poled lithium niobate cristal is used to injection-lock a powerful 778 nm laser. The light from this laser is coupled into a plano-concave cavity containing a rubidium cell. A photomultiplier on the side of the cell collects the fluorescence emitted by the atoms at resonance and

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

allows to control the length of the confocal cavity, therefore locking the DFB laser frequency to that of the two-photon transition.

3. LINEWIDTH REDUCTION

We achieve linewidth reduction of the 1556.2 nm DFB laser using weak optical feedback from an off-axis confocal cavity [5]. In order to maintain optimum phase of the feedback light, we control the position of a PZT-mounted mirror positionned between the DFB laser and the confocal cavity. This is performed by modulating the laser frequency at 70 kHz (through its injection current) and by demodulating the transmission signal of the confocal cavity detected by a germanium photodetector. This provides an error signal which is fed back to the PZT.

By recording the beat note between the DFB lasers of our two similar systems, we found that the linewidth was reduced from 1 MHz (nominal linewidth) to about 1 kHz with optical feedback. We also measured the beat between the two 778.1 nm slave lasers and found that their linewidth was a few kHz which is narrow enough to observe the two-photon transitions properly.

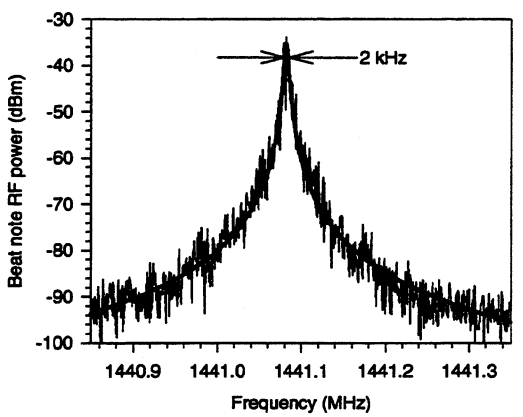


Figure 2: Optical beatnote between the 1556.2 nm DFB with optical feedback from a confocal cavity.

4. SECOND HARMONIC GENERATION IN A PPLN CRISTAL

In order to link the DFB laser frequency at 192.6 THz to that of the two-photon transition in rubidium at 385.2 THz, we use second harmonic generation in a PPLN cristal [6]. This cristal was developed in collaboration with INO, Québec and the Defense Research Establishment at Valcartier [7]. We realized a sample having a length of 9.6 mm with a 19.2 μm grating period. The temperature and wavelength bandwidths of this sample are respectively 10°C and 1.3 nm.

For our frequency standard, the sample was operated at 17°C to achieve the best second harmonic power at 1556.2 nm. In these conditions, we obtained a SH power of 7 μW with 53 mW at 1556.2 nm incident on the cristal.

A weak green light could also be observed at the output of the PPLN. This radiation corresponds to the sum frequency of the fundamental and second harmonic radiations in the PPLN cristal [8]. The theoretical grating period required to sum these frequencies is about 6.9 μm which is on the order of 1/3 of the 19.2 μm period grating we use. Therefore, the PPLN can generate the sum frequency using a third order quasi-phase matching process. However, the output power is weak because of the mismatch of the grating periods and also because the second harmonic light is in the μW range. A more efficient solution to have a useful green output at 577.8 THz (518.8 nm) would be to use another PPLN crystal with a 6.9 μm period and sum the 1556.2 nm and 778.1 nm outputs of our standard.

5. INJECTION LOCKING OF A 778 nm FABRY-PEROT LASER AND RESONANT CAVITY

After the PPLN cristal, the 7 μW SH beam is separated from the fundamental with a dichroïc beamsplitter. The transmitted beam at 1556.2 nm is the main optical output of our standard. The reflected second harmonic signal is focused on the front facet of a 40 mW Fabry-Perot laser (SDL 5402-H1) through a single-stage optical isolator equiped with a side port. This allows to injection-lock the 778 nm laser and control its frequency. The locking bandwidth of the slave laser was measured using a resonant cavity and is on the order of 1.5 GHz.

The 778 nm slave laser beam is then sent into a plano-concave resonant cavity having a free spectral range of 500 MHz. This resonant cavity has two functions: the first one is to enhance the power used to observe the two-photon transition; the second one is to ensure that the two counter-propagating beams probing the rubidium atoms are perfectly overlapping. This is necessary to avoid first-order Doppler effect. The optical power incident on the cavity can be controlled with a λ/2 plate and a polarizing beam splitter. The 778 nm beam transmitted by the polarizing beamsplitter is available to the user and constitutes the second output of our frequency standard. The finesse of the cavity is about 150 without the rubidium cell and decreases to 80 when the cell is inserted. The beam waist in the cavity is about 420 μm.

In order to maintain the laser and the resonant cavity frequencies in coïncidence, we demodulate the transmitted beam at 778.1 nm detected by a silicium photodetector and control the position of the input mirror of the cavity.

6. TWO-PHOTON RUBIDIUM SPECTROSCOPY

The rubidium (isotopic ^{87}Rb) atoms are contained in an evacuated fused silica cell which is placed in a copper oven and heated to about 90°C. The cold finger is maintained at about 82°C. When the laser is in resonance with the transition, the excited atoms return to the ground state through the radiative cascade 5D-6P-5S and emit a blue fluorescence at 420 nm. We use an aspheric condenser and a Fresnel lens placed on the side of the cell to collect this fluorescence which is detected with a Hamamatsu R928 photomultiplier (PM). We supply it with a voltage of 800 V and its load resistor is 20 kΩ. An interference filter centered at 420 nm and a colored glass filter are placed in front of the PM to reduce the noise due to the background light and the 778 nm stray light. The Rb cell, the fluorescence collection optics and the photomultiplier are placed in a specially designed μ-metal magnetic shield. On figure 3, we present the fluorescence profile corresponding to the transitions from $F_g = 2$ to $F_e = 4$, 3, 2, 1.

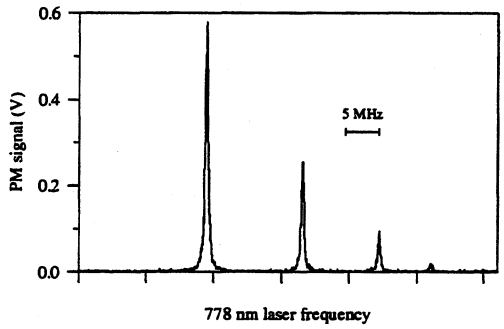


Figure 3: Fluorescence spectrum for the Fg=2 - Fe=4, 3, 2, 1 transition in ^{87}Rb.

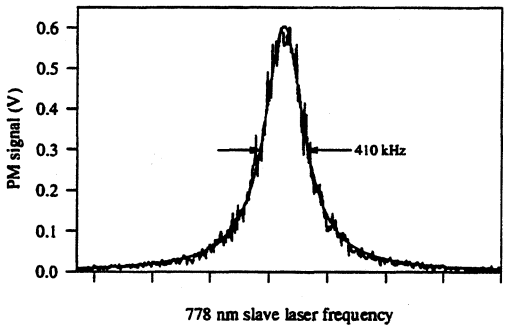


Figure 4: Fluorescence spectrum for the Fg=2 - Fe=4 transition in ^{87}Rb with a lorentzian profile fit.

We use the strongest transition $F_g = 2$ - $F_e = 4$ to frequency-lock our 1556 nm DFB laser. Figure 4 shows the corresponding profile with a 778 nm power of 0.5 mW incident on the cavity. A lorentzian profile was used to fit the fluorescence profile and we obtain a full width at half maximum of 410 kHz (at 778 nm). The main contributions to this linewidth are the natural linewidth (330 kHz) and the transit time broadening (70 kHz) [9].

The photomultiplier signal is demodulated at 70 kHz using a third lock-in amplifier, this allows to control the length of the confocal cavity, therefore locking the DFB laser frequency to that of the two-photon transition.

7. FREQUENCY STABILITY MEASUREMENTS

The frequency stability of our standards was measured by recording the beat between two independant 1556.2 nm DFB lasers. Both were locked to the same transition and one of them was shifted by 100 MHz using an acousto-optic modulator. Figure 5 presents the standard deviation of the beat note relative to the optical frequency.

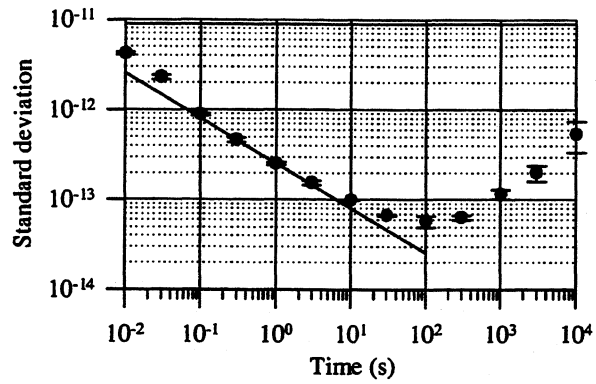


Figure 5: Frequency stability of the 1556.2 nm frequency standards.

We obtain a slope of about $2.5\times10^{-13}/\tau^{1/2}$ for averaging times between 0.1 s and 10 s. The best stability is 5.8×10^{-14} for 100 s. For greater averaging times, we observe a degradation of the stability. The main reason for this is probably the variation of the optical power probing the rubidium transition causing frequency fluctuations due to the change in the light shift.

8. FREQUENCY CHAIN FOR AN ABSOLUTE MEASUREMENT OF THE $5S_{1/2}$, F_g=2 – $5D_{5/2}$, F_e=4 TWO-PHOTON TRANSITION IN ^{87}Rb

We are presently working on the preparation of a frequency measurement of the ($5S_{1/2}$, F_g=2 - $5D_{5/2}$, F_e=4) two-photon transition in ^{87}Rb against a frequency chain developed at the NRC/INMS. This will allow us to confirm the value previously measured by F. Nez *et al* [10]. The two 1556.2 nm standards have been moved to the INMS/NRC in Ottawa, Canada. A specially designed frequency chain (figure 6) has been developed by A.A. Madej, J.E. Bernard, K.J. Siemsen and L. Marmet. It is based on the use of a CO laser (6.756 μm) and a Tm:YAG (2.022 μm) laser whose

sum frequency is very close to that of our 1556.2 nm standard (10.5 GHz). The Tm:YAG frequency is referenced to a trapped Sr^+ ion standard [11] through a divide-by-three step while the CO laser frequency is measured by comparison with CO_2 lasers and a klystron referenced to the Cs standard.

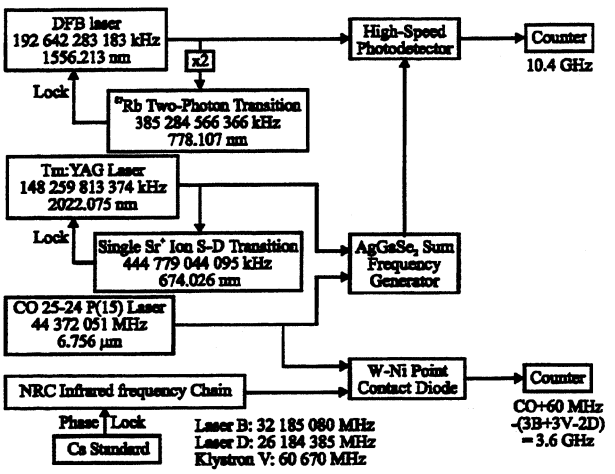


Figure 6: Frequency chain developed at the INMS/NRC for the absolute measurement of the ($5S_{1/2}$, F_g=2 - $5D_{5/2}$, F_e=4) transition in ^{87}Rb (courtesy of NRC).

We are presently evaluating systematic effects that shift the frequency of our standards (light shift, amplitude of the frequency modulation, lock-in amplifiers offsets, etc...). Absolute frequency measurements are under progress.

CONCLUSION

Two independant frequency standards at 192.6 THz (1556.2 nm) based on the two-photon transition of Rb have been realized at Laval University. They present a frequency stability of $2.5\times10^{-13}.\tau^{-1/2}$ for averaging times between 0.1 s and 10 s. Absolute frequency measure ments are under progress at the INMS/NRC in Ottawa, Canada. These standards presently provide two stabilized frequencies at 192.6 THz (1556.2 nm) and 385.2 THz (778.1 nm), and possibly at 577.8 THz (518.8 nm).

ACKNOWLEDGEMENT

We acknowledge B. Villeneuve and M. Svilans, Nortel Technology, for providing the high power DFB lasers. We are also grateful to Carl Paquet, INO, for his participation in the development of the PPLN crystal. We thank EXFO Electro-Optical Engineering and NSERC for their financial support through a NSERC/NRC grant.

REFERENCES

[1] M. Têtu, *OFC'98*, Tutorial Sessions SC145, "Absolute Frequency Control of semiconductor lasers for WDM applications", San José, CA, USA, February 1998.

[2] M. Poulin, C. Latrasse, D. Touahri, M. Têtu, P. Tremblay, "Compact frequency standard at 192.6 THz (1556 nm) based on the 5S - 5D two-photon transition in rubidium at 778 nm", paper WE2A-3, CPEM'98 Meeting, Washington, USA, July 10-16, 1998.

[3] M. Zhu and R. W. Standridge, "Optical frequency standard for optical fiber communication based on the Rb 5s → 5d two-photon transition", Opt. Lett., Vol. 22, No. 10, pp. 730-732, May 1997.

[4] D. Touahri, O. Acef, A. Clairon, J. J. Zondy, R. Felder, L. Hillico, B. de Beauvoir, F. Biraben and F. Nez, "Frequency measurement of the $5S_{1/2}$(F=3) - $5D_{5/2}$(F=5) two-photon transition in rubidium", Optics Commun., Vol. 133, pp. 471-478, 1 January 1997.

[5] B. Dahmani, L. Hollberg and R. Drullinger, "Frequency stabilization of semiconductor lasers by resonant optical feedback", Opt. Lett., Vol. 12, No. 11, pp. 876-878, November 1987.

[6] M. M. Fejer, G. A. Magel, D. H. Jundt, R. L. Byer, "Quasi-phase-matched second harmonic generation tuning and tolerances", IEEE J. Quantum Electron., Vol. 28, No. 11, pp .2631-2654, Nov. 1992.

[7] M. Poulin, C. Latrasse, M. Têtu, C. Paquet, C. Tremblay, P. Mathieu, "Second-harmonic generation at 1556 nm in periodically poled $LiNbO_3$", DAMP/DOP'97 Meeting, Ottawa, Canada, October 1997.

[8] O. Pfister, J.S. Wells, L. Hollberg, L. Zink, D.A. Van Baak, M. D. Levenson, W. R. Bosenberg, "Continuous-wave frequency tripling and quadrupling by simultaneous three-wave mixings in periodiclly poled crystals: application to a two-step 1.19-10.71 µm frequency bridge", Opt. Lett., Vol. 22, No. 16, pp. 1211-1213, August 1997.

[9] F. Biraben, M. Bassini, and B. Cagnac, "Lineshapes in Doppler-free two-photon spectroscopy. The effect of finite transit time", Journal de Physique, vol. 40, pp. 445-455, May 1979.

[10] F. Nez, F. Biraben, R. Felder, and Y. Millerioux, "Optical frequency determination of the hyperfine components of the $5S_{1/2}$-$5D_{3/2}$ two-photon transitions in rubidium", Opt. Commun., vol. 102, nos. 5-6, pp. 432-438, Oct. 1993.

[11] J. Bernard, A. A. Madej, L. Marmet,. B.G. Whitford, K.J. Siemsen and S. Cundy, "Cs-based frequency measurement of a single, trapped ion transition in the visible region of the spectrum", Phys. Rev. Lett., Vol. 82, pp. 3228-3231, 1999.

1999 Joint Meeting EFTF - IEEE IFCS

EXPERIMENTAL CHARACTERIZATION OF FM SPECTROSCOPY OF $^{127}I_2$ AT 532 nm USING A FREQUENCY-DOUBLED Nd:YAG LASER

G. Galzerano[(1)], C. Svelto[(1)], P. Ceriani[(1)], L. Re[(1)], F. Bertinetto[(2)], and E. Bava[(1)]

[(1)]Dipartimento di Elettronica e Informazione del Politecnico di Milano,
P.zza Leonardo da Vinci 32, 20133 Milano, Italy
E-mail: galzeran@elet.polimi.it Phone: +39 02 239936-09/01 Fax: +39 02 23993413

[(2)]Istituto di Metrologia Gustavo Colonnetti, Strada delle Cacce 73, 10135 Torino, Italy
Phone: +39 011 3977454 Fax: +39 011 3977459

ABSTRACT

This work deals with the experimental characterization of the Frequency Modulation Spectroscopy method by means of a frequency-doubled Nd:YAG laser and saturated absorptions of $^{127}I_2$ at the wavelength of 532 nm. A power of 8.7 mW at 532 nm with a continuos frequency tuning range of 80 GHz has been obtained by doubling 80 mW of the infrared Nd:YAG radiation incident on a self made external cavity doubler. In order to obtain a high degree of stability and reproducibility of the 532 nm radiation locked to a saturated absorption of $^{127}I_2$, linewidth, sensitivity, and signal to noise ratio of the demodulated signals have been characterized for different phase modulation conditions, iodine vapor pressures and for different optical power levels.

1. INTRODUCTION

Sub-Doppler optical Frequency Modulation Spectroscopy FMS [1] represents a very sensitive technique to realize practical optical frequency standards for a variety of scientific applications. Frequency-doubled monolithic diode-pumped Nd:YAG lasers due to their intrinsic characteristics of short term frequency stability and in conjunction with the FMS applied to saturated absorption of $^{127}I_2$ at 532 nm, are very attractive sources for applications in space communications, high resolution spectroscopy [2], fundamental physics [3] and metrology [4].

In the FMS technique, an unmodulated pump beam saturates the molecular transition whereas a counter-propagating probe beam, phase-modulated at a frequency higher than the resonance linewidth, is used to detect the saturated absorption which is, to the first order, Doppler-free and therefore homogeneously broadened. Because of the absorption and dispersion profiles of the saturated resonance, the originally phase modulation is converted into a corresponding amplitude modulation that can be detected by a photodiode. The in-phase component of the synchronously demodulated photocurrent (related to the dispersion profile) is an odd function of the detuning between the laser frequency and the resonance frequency (discriminating signal) and can therefore be used as an error signal to lock the laser frequency to the resonance center [1].

2. EXPERIMENTAL SET-UP

As a first step, the efficient generation of the 532 nm has been obtained by doubling the fundamental frequency of a diode-pumped Non-Planar Ring-Oscillator Nd:YAG laser (Ligthwave mod. 122). Fig. 1 shows the frequency-doubling setup adopted. The doubler consists of an external linear resonator (near concentric) where a non-linear crystal of $LiNbO_3$, 7% doped with MgO, is inserted. To keep the resonance condition between the fundamental laser frequency and the center of the cavity line, which is necessary to achieve the maximum green power, a frequency modulation or Pound-Drever technique [5] was used. By modulating the laser frequency at a rate of 196 kHz with a frequency deviation of 50 kHz by means of the laser piezo transducer (PZT), detecting either the cavity reflection or transmission, and using a proportional / integral servo closed to the cavity PZT, the locking condition has been maintained for several hours. With an incident infrared power of 80 mW and at the non-critical phase-matching temperature of 51.5 °C a maximum green power of 8.7 mW (with a rms stability of 1 % in 1 h) has been obtained.

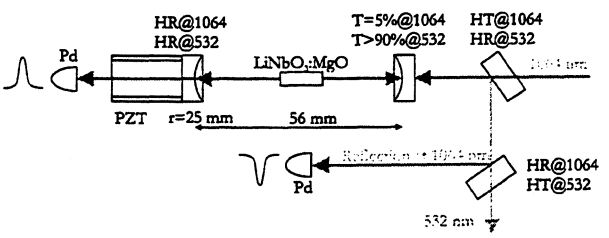


Figure 1. The adopted linear doubler.

By changing the NPRO temperature it is possible to span the frequency of the 532 nm radiation by ≈80 GHz

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

with a tuning coefficient of ≈-2 GHz/ °C. In this frequency range eight rovibronic iodine transitions can be addressed, namely from 1106 to 1113. One of these lines, the 1110 or R(56)32-0, is of great metrological interest because has been recently recommended by the CIPM for the realization of the meter [6].

The observed iodine transitions have been identified by comparing the hyperfine structures of these lines, obtained in the saturated absorption measurements set-up, with those available in literature [2, 7].

In the Fig. 2 the optical and electronic arrangement adopted for the implementation of the FMS technique is shown. The probe and pump beams are counterpropagating inside a 50 cm long iodine cell with circular polarizations (using two quarter wave plate) and with collimated beam diameters of ~1.5 mm. The iodine vapor inside the cell is kept at a fixed pressure by controlling the cold finger temperature with a stability of 0.1 °C.

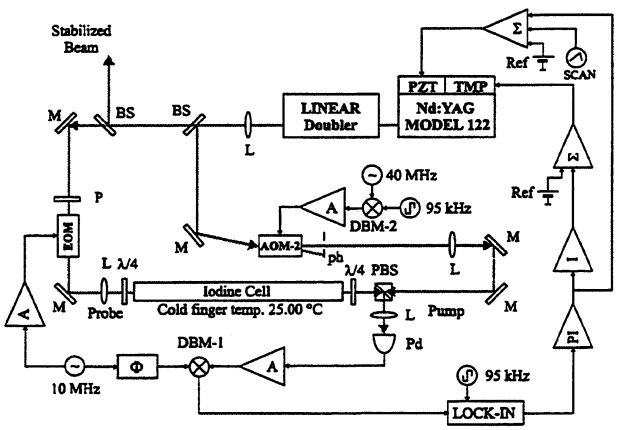


Figure 2. Experimental set-up for the FMS scheme.

Phase modulation of the probe beam is realized by an electrooptic modulator (EOM) driven with a RF signal at typical frequency of 10 MHz. By means of the acousto-optic modulator (AOM) the saturating beam is frequency shifted by 40 MHz and amplitude chopped at a rate of 80 kHz, avoiding in this way interference patterns due to optical feedback on the photodetector, and allowing, with a lock-in detection scheme, the suppression of the linear absorption (Doppler) background profile.

3. RESULTS

In order to understand how the probe and pump powers, modulation parameters (modulation frequency and phase index), and iodine vapor pressure can influence the frequency of the stabilized laser the demodulated signal lineshapes have been recorded by means of a Analog to Digital system, changing the parameters of interest.

As an example of recordings, in Figs. 3 (a) and 3 (b) are shown the dispersion profiles of the R(56)32-0 and P(53)32-0 transitions, which consist of 15 and 21 hyperfine structure components, respectively.

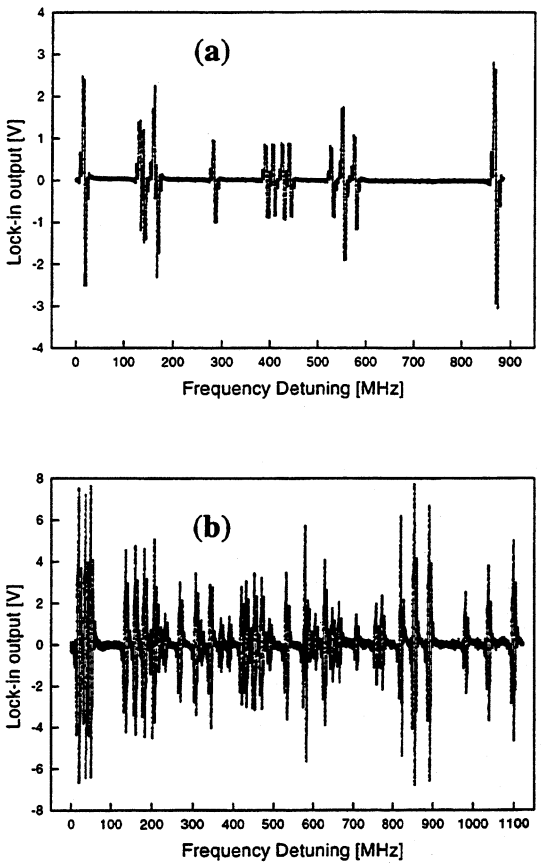


Figure 3. Dispersion profiles of **a)** R(56)32-0 and **b)** P(53)32-0 iodine line.

Particular attention has been paid to the recommended component a_{10} of the R(56)32-0 and, for this reason, the parametric dependence was carried out using this component. Usual working conditions of a modulation frequency of 10 MHz, a phase modulation index ≈1, pump and probe powers of 0.5 mW and 1.6 mW, respectively, and an iodine cold finger temperature of 22°C have been adopted. The obtained lineshapes for the dispersion and for the absorption profile of the a_{10} component are reported in Fig. 4. Using this type of recordings it was therefore possible to measure a slope at the center of the dispersion shape (frequency to voltage discriminating signal), a resonance linewidth, and a signal to noise ratio (expressed as the ratio between the slope and the statistical dispersion of the background line) of 1.2 MHz, 20 V/MHz, and 62 dB in a integration bandwidth of 150 Hz, respectively.

The following Figs. 5 (a) and 5 (b) show the linewidth of the a_{10} component for different cold finger iodine temperature and saturating power.

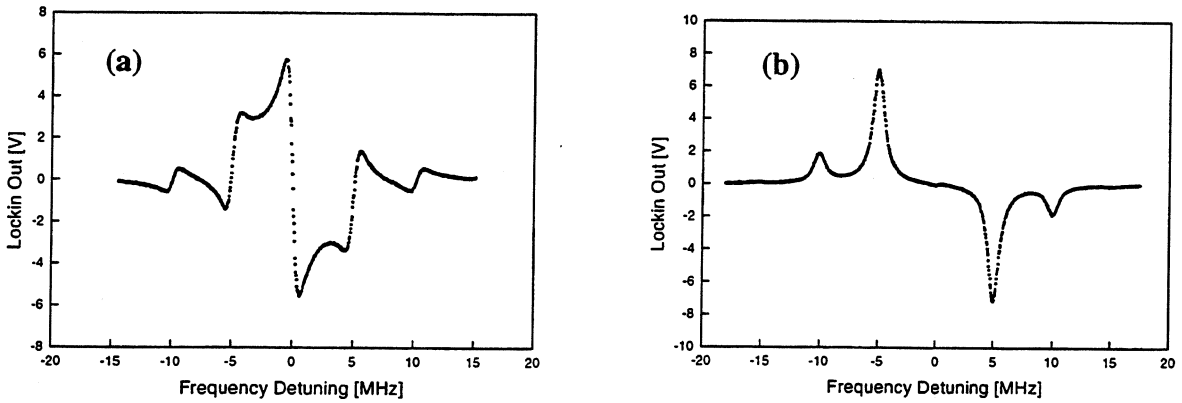


Figure 4. a) Dispersion and **b)** absorption profiles for the a_{10} component of the transition R(56)32-0.

From these diagrams a linear dependence of the homogeneous linewidth from the iodine pressure of 50(15) kHz/Pa and a saturating intensity of 0.45(4) mW/mm^2 have been estimated (by fitting). Furthermore, for iodine temperature higher than 5 °C, a constant linewidth was observed. This behavior shows that the iodine cell was filled with an iodine saturated-vapor pressure corresponding to ≈5 °C.

For stabilization purposes the realized slope (sensitivity) of the dispersion profile which acts as the error signal for locking the laser frequency to the iodine resonance is of great importance. Also the parametric dependence of this slope has therefore been investigated. As an example the Figs. 6 (a) and 6 (b) show the obtained slope as a function of the modulation index and frequency, respectively.

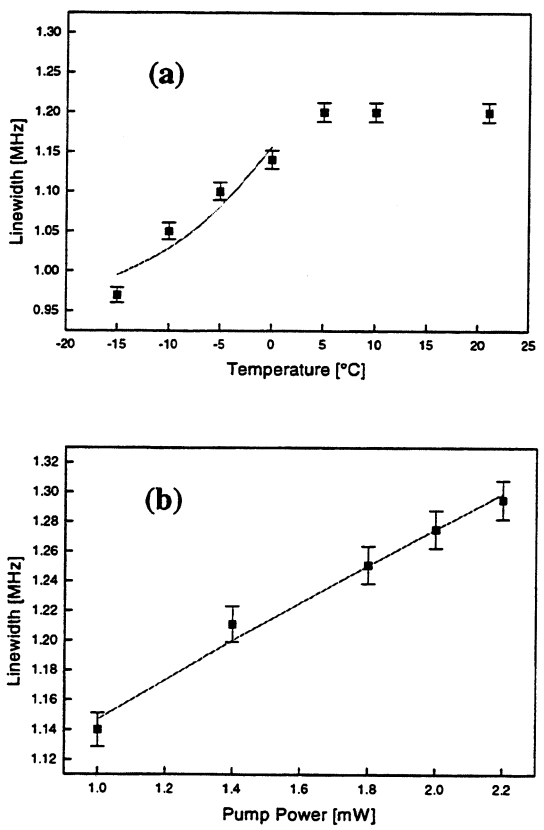


Figure 5. Linewidth at different **a)** cold finger iodine temperature and **b)** pump power level. Modulation frequency and index were 10 MHz and ≈1, respectively.

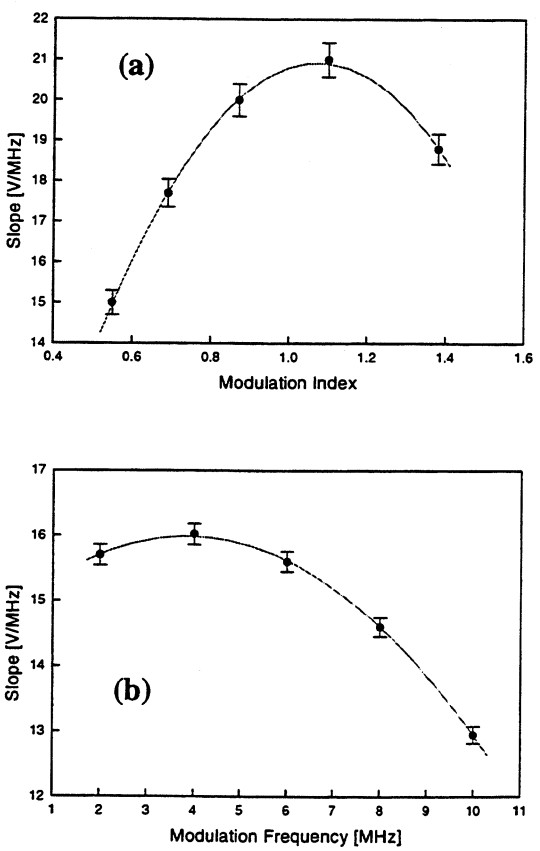


Figure 6. Sensitivity at different **a)** phase modulation index and **b)** phase modulation frequency.

From Fig. 6 (a), obtained at a fixed modulation frequency of 10 MHz, the maximum slope has been achieved for a phase modulation index of 1.1. On the other hand, keeping the modulation index at the constant value of 1 and changing the modulation frequency, the slope show an optimum value at 4 MHz, as it show in Fig. 6 (b). However, for this modulation frequency, the signal to noise ratio turns out to be lower than that realized at a modulation frequency of 10 MHz.

4. CONCLUSIONS

Experimental characterization of the Frequency Modulation Spectroscopy has been reported using frequency-doubled Nd:YAG laser and saturated absorption lines of $^{127}I_2$ at 532 nm. The Frequency Modulation Spectroscopy signals, both absorption and dispersion, have been characterized for different phase modulation conditions, iodine vapor pressures and for different optical power levels. In this way the optimum conditions to realize an optical frequency standard at 532 nm have been carried out.

Also the adopted scheme for the second-harmonic generation of the infrared Nd:YAG radiation has been reported.

From a metrological point of view, the 532 nm wavelength is of great interest since it has been recently recommended as an absolute reference in the CCL meeting in '97 by means of the frequency value of the a_{10} hyperfine component of the iodine transition R(56)-32-0, with an uncertainty of 40 kHz.

REFERENCES

[1] J. L. Hall, L. Hollberg, T. Baer and H. G. Robinson, Optical heterodyne saturation spectroscopy, *Appl. Phys. Lett.*, vol. 39, pp. 680, 1981.

[2] A. Arie and R. L. Byer, Laser heterodyne spectroscopy of $^{127}I_2$ hyperfine structure near 532 nm, *J. Opt. Soc. Am. B.*, vol. 11, pp. 1990-1997, 1993.

[3] A. Abramovici, W. E. Althouse, R. W. P. Drever, Y. Gursel, S. Kawamura, F. J. Raab, D. Shoemaker, L. Sievers, R. E. Spero, K. S. Thorne, R. E. Vogt, R. Weiss, S. E. Whitcomb, and M. E. Zucker, LIGO: The laser interferometer gravitational-wave observatory, *Science*, vol. 256, pp. 325-333, 1992.

[4] M. Eickhoff and J. L. Hall, Optical frequency standard at 532 nm, *IEEE Trans. IM*, vol. 44, pp. 155-158, 1995.

[5] T. Day, E. K. Gustafson, and R. L. Byer, Sub-hertz relative frequency stabilization of two-diode laser-pumped Nd:YAG lasers locked to a Fabry-Perot interferometer, *IEEE J. Quantum Electr.*, vol. 28, pp. 1106-1116, 1992.

[6] Quinn, T., Mise en pratique of the definition of the metre (1997), *Metrologia*, in preparation, see also Quinn, T., *Metrologia*, vol. 30, 523-541, 1994.

[7] Ye, J., Robertsson, L., Picard, S., Ma L.-S., and Hall, J. L., Absolute Frequency atlas of molecular I_2 lines at 532 nm, to be published in *IEEE Instrum. Meas.*

COMPACT, BROADLY TUNABLE MID-IR SOURCE FOR THE SPECTROSCOPIC INVESTIGATION OF MOLECULAR REFERENCE LINES IN THE 27 TO 33 THz RANGE

T. KAING AND J.-J. ZONDY
Laboratoire Primaire du Temps et des Fréquences (BNM-LPTF)
Bureau National de Métrologie/Observatoire de Paris
61, avenue de l'Observatoire, F-75 014 Paris (France)

A. YELISSEYEV, S. LOBANOV AND L. ISAENKO
Design & Technological Institute of Monocrystals (DTIM), SB RAS
43, Russkaya Str., 630058 Novosibirsk (Russia)

ABSTRACT

We report on the improvement of a tunable diode-laser-based difference-frequency spectrometer using $AgGaS_2$ nonlinear crystal. This spectrometer is being developed as an alternative to CO_2 laser spectrometers in the 9 µm to 11 µm range. We use a type-II cut crystal as a part of the improvement compared to a type-I cut, which was used in our preliminary set-up [1,2]. Using a Fabry-Perot cavity to enhance the 780-nm power (35 mW in front of the cavity) and 120 mW of the 843-nm radiation, 30-nW power of the tunable 10-µm radiation is detected. This power level is enough to investigate the linear absorption spectroscopy of SF_6. Doppler limited spectra over 2 GHz are recorded, showing the wide tunability of the spectrometer.

1. INTRODUCTION

Heavy spherical molecules such as O_sO_4, SF_6 or SiF_4 are potential accurate frequency references for an IR to visible frequency measurement chain. The BNM-LPTF secondary frequency standard at 29 THz is presently based on a CO_2 laser frequency-stabilized on a saturated absorption line of osmium tetroxyde (OsO_4) at 10.3 µm in coincidence with the R(10) CO_2 line [3,4,5]:

$$\nu_{IR} = 29\ 054\ 057\ 446\ 579 \pm 4\ \text{Hz}$$

In the 9 µm to 11 µm range, however, high resolution spectroscopy is often limited by the tunability of the CO_2 laser. As a matter of fact, the frequency tuning range of a single-mode low-pressure (10 mTorr) CO_2 laser of 0.8 m discharge length is 100 MHz per line , with a 30 GHz line-to-line separation. Besides, there is a spectral gap between the two transition bands centered around 9.4 µm and 10.6 µm. Spectroscopy with such a laser is then subjected to accidental coincidences between one laser line and the absorbing species of metrological interest (OsO_4, SF_6, SiF_4, O_3...). Although the CO_2 laser tunability can be extended with the use of acousto-optic or electro-optic modulators [6], their potential as easily and widely tunable sources is somewhat limited compared to that of a cw all-diode difference-frequency-generation (DFG) source producing microwatt range radiation, which is the typical level of power required for external-cavity enhanced saturation spectroscopy [3]. This latter source, owing to its compactness and extended tuning capability, is a very attractive device for the investigation of further narrow lines of OsO_4 that cannot be probed by a CO_2 laser. Even though the dye or Ti:Sapphire lasers allow wider continuous tuning range, the use of room-temperature, all-diode-laser-based DFG spectrometer is advantageous due to the portability of the device [7]. DFG spectrometers with nanowatt or microwatt-level output using Ti:Sapphire, diode pumped Nd:YAG or dye lasers, sometimes combined with diode lasers, have already been demonstrated for high resolution spectroscopy and trace gas detection [8,9,10,11]. In order to have a tunable, compact and portable DFG spectrometer, we have decided to use two low-power diode lasers. A microwatt power level is targeted by using a dual-arm cavity to enhance both pump lasers.

2. HIGH RESOLUTION EXPERIMENTAL SET-UP

The DFG experimental set-up is sketched in figure 1. Two free-running diode lasers with power P_3=50 mW and P_2= 200 mW are injection-locked by two master extended cavity diode lasers (ECDL) operating around λ_3= 780 nm (pump) and λ_2=843 nm (signal). For high resolution purpose, the masters are stabilized to the same high finesse cavity. Due to the common-mode residual frequency fluctuations, sub-kilohertz IR spectral resolution can be achieved with this set-up by difference frequency mixing [2,12]. The 843-nm master radiation is double-passed in a 200-MHz acousto-optic modulator before the injection locking stage. This allows to fine tune the frequency of the master over 150 MHz while the slave is still injection-locked. Owing to the 5 GHz injection locking range of the 780-nm slave

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

laser, broad continuous tuning range of the spectrometer over ~2.5 GHz can be also achieved by applying a voltage ramp to the grating PZT of the ECDL master, without the need to readjust the phase-matching angle. Such a broad tuning range enables to scan several Doppler limited absorption resonances of the SF_6 gas.

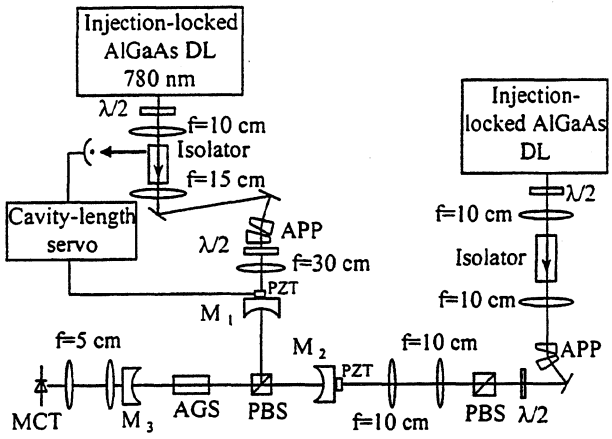


Fig. 1: Difference-frequency generation experimental set-up. APP: anamorphic prism pair, PBS: polarizing beam splitter.

In our preliminary set-up described in ref. [1], a 5 mm×5 mm×15 mm $AgGaS_2$ or AGS crystal cut for type-I (o-o→e) phase-matching, was used to generate 40 nW of 10.2-μm in a dual-cavity configuration. The results of this preliminary set-up were limited by the high round-trip loss of the dual-arm cavity (16% for each arm). This was due to poor broadband high-reflection (BBHR) coating in the near-IR of the ZnSe output coupler M_3 (90% reflection in the near IR and 90% transmission at 10 μm) and to the lossy intra-cavity elements (2.5% round-trip loss for the coupling Foster prism and 3% for the crystal). The high loss led to a poor coupling efficiency of the incident pump radiation into each arm of the dual-cavity. In the present experiment, these elements have been replaced by higher quality ones. The new BBHR ZnSe output coupler has a ≥99% reflection in the near-IR and 90% transmission at 10 μm. A broadband antireflection (BBAR) coated polarizing beam splitter with less than 1% round-trip loss is used instead of the Foster prism. In ref. [2], we have demonstrated that the use of type-II (e-o→e) cut $AgGaS_2$ crystal (θ=46°, φ=0) resulted in two times higher nonlinear conversion efficiency with respect to type-I for the same crystal length. In spite of good BBAR coatings (R=0.6%) in the near-IR region, the measured transmissivity of the 5 mm×5 mm×15 mm type-II $AgGaS_2$ used in ref. [2] was only about T=86% (at 778 nm and 842 nm). This was due to its high residual absorption in this near-IR region ($\alpha \approx 10\%$ cm^{-1}). Since then, a new higher quality 5 mm×5 mm×15 mm type-II $AgGaS_2$ crystal, grown at the Design & Technological Institute of Monocrystals, is used in our set-up. This crystal is broadband antireflection coated in the near IR with 98% transmission (absorption=0.25% cm^{-1}) for the ordinary 843-nm wave and 97% (absorption=1.33% cm^{-1}) for the extraordinary 780-nm wave. The dual-arm cavity allows to independently lock each arm length to the respective laser frequency. The use of low power diodes calls for a special care on input power coupling issue. To couple the maximum pump energy in each arm, slave diode beam reshaping with the use of anamorphic prism pair is necessary.

3. SINGLE PASS AND SINGLY-RESONANT CONVERSION EFFICIENCY

With these optical components, we first measure the conversion efficiency without any cavity (single-pass DFG) and then close the 780-nm arm (singly-resonant DFG). Without crystal, the finesse and the mode-matching rate are respectively 46 and 67% for a 12.9 cm optical cavity length (waist w_0=75 μm). These values dropped to 28 and 45% with the crystal. The reflection fringe displays no significant transverse cavity modes, hence this poor coupling efficiency is due to impedance mismatch of the input coupling mirror transmission (T=4.3%) to the total intra-cavity losses (6.6%). In the final configuration, this input coupler will be replaced by a new one with ~6% transmission in order to have a better mode-matching rate. With a finesse of 28 and a pump coupling efficiency of 45%, the calculated and measured enhancement of the 780-nm arm is only 2.5 compared to the single-pass case. The experimental conversion efficiencies are shown in figure 2. They are obtained with 35 mW of the 780-nm power in front of the cavity and by varying the power of the 843-nm radiation.

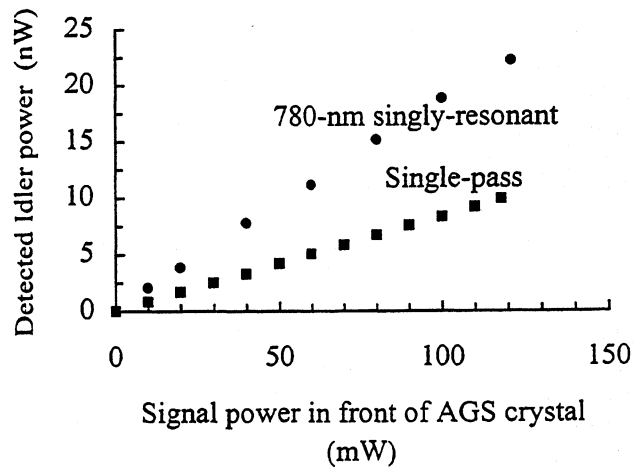


Fig. 2: DFG conversion efficiency using type-II phase-matching. The measured enhancement factor of the 780-nm cavity is only 2.5.

The spatial intensity distributions in the walkoff plane in the case of a single-pass DFG and singly-resonant DFG are compared in figure 3. In the resonant configuration, the profile is closer to a Gaussian mode than the other configuration. This is due to the TEM_{00} mode pump filtering effect of the resonator. The single-pass transverse profile of the DFG radiation displays an elliptical cross-section, with the minor axe of the ellipse located in the walkoff plane. This beam distortion is induced by the walkoff (ρ=1.2°) and focusing effects [13]. A nearly Gaussian beam section is important in our future saturated spectroscopy experiments, which should be performed within a Fabry-Perot resonator.

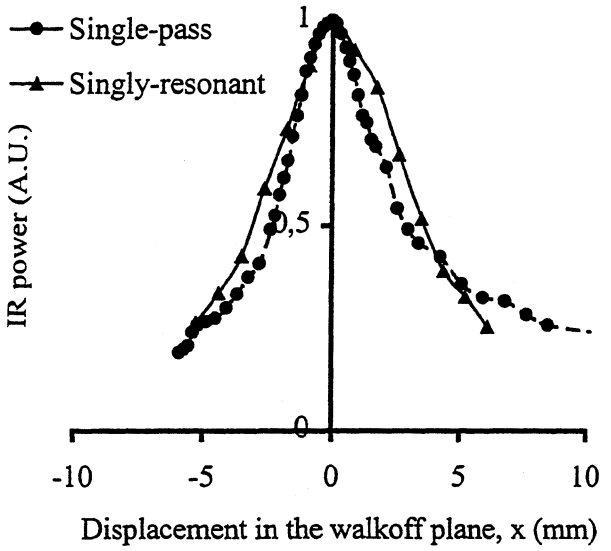


Fig. 3: Comparison between the transverse power distribution of the IR output in the walkoff plane in a single-pass DFG and pump-resonant DFG.

4. DOPPLER LIMITED SPECTROSCOPY OF SF_6

Using the singly-resonant configuration, we performed preliminary linear absorption spectrum of SF_6 over 2 GHz to demonstrate the wide tunability of our set-up. The 25-cm long Brewster cell is filled with less than 0.05 mTorr SF_6 pressure. In this experiment, the master lasers were free-running, with typical short term linewidth of 100 kHz. Their wavelengths are calibrated using a wavemeter with 500 MHz reading accuracy. They are set to 780.87 nm and 843.28 nm, in order to probe the transition band in coincidence with the P(16) line of the CO_2 laser. This band is rich in saturated transitions (55) spaced by ~10 MHz [14]. The 30-nW generated 10.55-µm radiation (idler) is then tuned in the vicinity of the Q48 (F2 5) saturated absorption line of SF_6. The IR radiation is focused into the cell with a ~1.5-mm waist. When the pressure of SF_6 in the cell is about 0.12 mTorr, almost all of the IR radiation is absorbed. The frequency of the 780-nm ECDL master laser is tuned over 2.2 GHz at a fixed diode temperature and injection current by applying a triangular voltage ramp to its grating PZT transducer (f=0.1 Hz). During the tuning process, the injection locking and the 780-nm cavity length stabilization to the laser frequency must be maintained. The 780-nm cavity is stabilized by dithering its length at 20 kHz and demodulating the reflected light intensity by a lock-in amplifier. This technique provides a lock at the maximum fringe transmission but, owing to the low modulation frequency, the bandwidth of the active length servo is limited to 200 Hz. The residual 20 kHz amplitude modulation could not be filtered out from the locking error signal. Furthermore, acoustical perturbation at frequencies above 500 Hz are not well corrected, which results in a poor intensity stability (~ 5% rms) of the probe DFG output. Preliminary Doppler spectra are shown in figure 4. The large intensity noise caused by the limited servo bandwidth must be cancelled prior to an attempt of saturation spectroscopy experiment. A side-of-fringe locking servo would provide a much more stable IR output intensity. The whole targeted 9-11 µm spectral range can be explored using a single crystal [1].

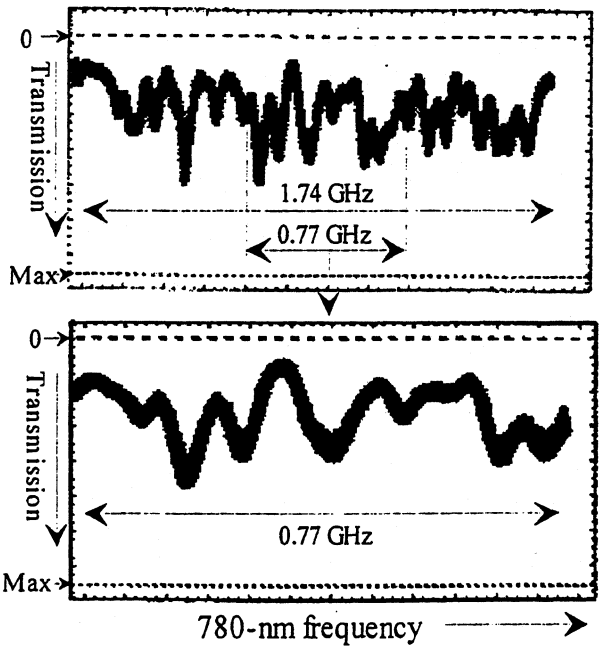


Fig. 4: Linear absorption spectrum of the SF_6 in cell while the 780-nm master frequency is tuned over 1.74 GHz (up) and 0.77 GHz (low). The dashed lines indicate the base line and the maximum IR power in the absence of the gas. The two spectra are taken separately: the lower spectrum is not a zoom of the upper one. The spectra show many overlapping Doppler-broadened transitions.

5. CONCLUSION

We have shown the potential wide tuning capabilities of our all-diode-laser-based spectrometer by scanning a portion of SF_6 spectrum. The 30-nW level IR power is enough for linear spectroscopy on the strongly absorbing SF_6 molecule. For saturated high resolution spectroscopy, microwatt level is necessary [4]. The next step is to improve the output power of the 10-μm radiation by closing the 843-nm cavity arm and by optimizing the impedance-matching efficiency. Further progress in reducing the residual amplitude noise is also under development. This spectrometer may work as an accurate 30-THz difference-frequency standard in the near-IR region, extending thus the reference grid of our optical frequency chain [15]. This will also allow long distance dissemination of the LPTF IR reference standards by using optical fiber links to convey the two near-IR frequency markers.

6. REFERENCES

[1] D. Lee, T. Kaing, J.-J. Zondy, "An all-diode-laser-based, dual-cavity $AgGaS_2$ cw difference-frequency source for the 9-11 μm range", Appl. Phys. B **67**, pp. 363-367, 1998

[2] T. Kaing, J.-J. Zondy, A. Yelisseyev, S. Lobanov, and L. Isaenko, "Improving the power and spectral performance of a 27 to 33 THz $AgGaS_2$ difference-frequency spectrometer", to be published in IEEE Trans. Instr. Meas.

[3] O. Acef, "Metrological properties of CO_2/OsO_4 optical frequency standard", Opt. Commun. **134**, pp. 479-486, Jan. 1997

[4] F. Ducos, G. D. Rovera, C. Daussy and O. Acef, "Performances of OsO4 stabilized CO2 lasers as optical frequency standards near 29 THz", LPTF'99 issue

[5] A. Clairon, O.Acef, C. Chardonnet, and C. J. Bordé, "State-of-the-art for high accuracy frequency standards in the 28 THz range using saturated absorption resonances of OsO_4 and CO_2", in Frequency Standards and Metrology, A. de Marchi, ed. (Springer-Verlag, Berlin, 1989), pp. 212-221

[6] F. Herlemont, M. Khelkhal, J. Legrand, and G. Pierre, "Doppler-free two-photon spectrum of SF_6 for metrological purposes", Opt. Lett., vol. 23, pp. 957-959, June 1998

[7] S. Sanders, R. J. Lang, L. E. Myers, M. M. Fejer and R. L. Byer, "Broadly tunable mid-IR radiation source based on difference frequency mixing of high power wavelength-tunable laser diodes in bulk periodically poled $LiNbO_3$", Electron. Lett. **32**, pp. 218-219, Feb. 1996

[8] W. Chen, G. Mouret, D. Boucher, "Difference-frequency laser spectroscopy detection of acetylene trace constituant", Appl. Phys. B **67**, 375-378, 1998

[9] B. Sumpf, T. Kelz, M. Nägele, H.-D. Kronfeldt, "A cw $AgGaS_2$ difference frequency spectrometer with diode lasers as pump sources", Appl. Phys. B **64**, pp. 521-524, 1997

[10] D. G. Lancaster, D. Richter, R. F. Curl, F. K. Tittel, "Real-time measurements of trace gases using a compact difference-frequency-based sensor operating at 3.5 μm", Appl. Phys. B **67**, pp. 339-345, 1998

[11] M. Seiter, D. Keller, M. W. Sigrist, "Broadly tunable difference-frequency spectrometer for trace gas detection with noncollinear critical phase-matching in $LiNbO_3$", Appl. Phys. B **67**, pp. 351-356, 1998

[12] D. Touahri, F. Nez, M. Abed, J.-J. Zondy, O. Acef, L. Hilico, A. Clairon, Y. Millerioux, F. Biraben, L. Julien, and R. Felder, "LPTF Frequency Synthesis Chain : Results and Improvement for the Near Future," IEEE Trans. Instrum. Meas. **44**, pp. 170-172, 1995

[13] J.- J. Zondy, "The effects of focusing in type-I and type-II difference-frequency generations", Opt. Commun. **149**, pp. 181-206, 1998

[14] Ch. Bréant, "Development of tunable infrared lasers with high spectral purity, application to hyperfine spectroscopy of the HF and SF6 molecules", PhD dissertation (Université de Paris-Nord, Laboratoire de Physique des lasers), 1985

[15] T. Kaing, D. Lee, J.-J. Zondy : « Diode laser pumped, 9-11 μm range doubly-resonant difference frequency spectrometer using $AgGaS_2$ crystal », in OSA Trends in Optics and Photonics, vol. 19 Advanced Solide Statd Lasers, Walter R. Bosenberg and Martin M. Fejer eds. (Optical Society of America, Washington DC 1998), p. 285-290.

1999 Joint Meeting EFTF - IEEE IFCS

TOWARD THE REALIZATION OF A FREQUENCY STANDARD AT 1.5 μm BASED ON NARROW-LINEWIDTH ERBIUM LASERS AND SATURATED ACETYLENE LINES

A. Onae[1], C. Svelto[2], K. Nakagawa[3], G. Galzerano[2], S. Taccheo[4], P. Laporta[4], and E. Bava[2]

[1]National Research Laboratory of Metrology,
1-1-4 Umezono, Tsukuba, Iabaraki 305, Japan
E-mail: Onae@NRLM.go.JP Phone: + 81 298 54 4045 Fax: + 81 298 54 4135

[1]INFM-Dipartimento di Elettronica e Informazione del Politecnico di Milano and CNR-CSTS,
Piazza Leonardo da Vinci 32, 20133 Milano, Italy
E-mail: Cesare.Svelto@PoliMi.IT Phone: +39 02 2399 3610 Fax: +39 02 2399 3413

[2]Institute for Laser Science, University of Electro-communications,
1-5-1 Chofugaoka, Chofu, Tokio 182-8585, Japan

[4]INFM-Dipartimento di Fisica del Politecnico di Milano and CNR-CEQSE,
Piazza Leonardo da Vinci 32, 20133 Milano, Italy
E-mail: Paolo.Laporta@PoliMi.IT Phone: +39 02 2399 6151 Fax: +39 02 2399 6128

ABSTRACT

This work is aimed at the study and development of a high-accuracy frequency standard at 1.5 μm, by combining the intrinsic frequency stability of a solid-state laser to the accuracy and reliability of Doppler-free molecular lines. A diode-pumped Er-Yb microlaser will be frequency stabilized to the center of $^{13}C_2H_2$ saturated absorption lines by means of the frequency modulation technique. A resonant Fabry-Perot cavity containing acetylene at low pressure is used to enhance the field intensity in order to saturate the 1.5 μm overtone transitions of the molecule. Two stabilization loops will be used to lock the Fabry-Perot cavity to the laser frequency and this last to the peak of the saturation dip. Preliminary experimental results on the Er-Yb laser frequency noise are reported.

1. INTRODUCTION

The demand for 1.5 μm wavelength/frequency standards is becoming a challenging task for several research groups working in the fields of optical metrology [1–3] and optical fiber communications [4, 5]. To achieve the required absolute frequency stability, different solutions have been proposed or performed using as a reference several transitions from a variety of molecular (NH_3, H_2O, HCN, C_2H_2, $^{13}C_2H_2$) and atomic (Kr, Rb, K) species. After choosing the proper molecular/atomic frequency reference, a suitable laser source has to be selected in order to realize a reliable frequency standard with good metrological properties. Typical high-performance laser sources in the 1.5-μm spectral region are the external-cavity semiconductor laser diodes (EC-SLD). These devices [6], in fact, allow for wide wavelength tunability (a few tens of nm) and narrow oscillating linewidth (~100 kHz) but exhibit a limited output power (a few mW) with rather poor spatial quality (elliptic and astigmatic beam). An interesting alternative source for the highest-demanding applications is found in diode-pumped Er-Yb microlasers [7]. In fact, these relatively novel laser devices, yet undergoing some technological improvements, allow for several tens of mW output power in single-frequency operation (1–ms linewidth below 50 kHz) with wide wavelength tunability (~40 nm) and excellent beam quality (circular beam with M^2 factor of 1.1).

Due to the abundance of well resolved and almost equispaced absorption lines, the acetylene molecule is one of the best suited reference in the 1.5-μm spectral region. From a chemical and technological point of view, this molecule is stable, non-toxic and non-corrosive, and allows for an easy and inexpensive preparation in glass cells. Regarding the metrological properties, the acetylene molecule is rather insensitive to external electric fields, having no permanent electric dipole moment, and exhibits a single saturated absorption within each Doppler-broadened line (no hyperfine structure). Unfortunately, quite intense optical fields are required to achieve sub-Doppler spectroscopy and frequency locking because the saturation intensity of the acetylene molecule is rather high, in the order of 5 W/mm^2. Presently, the compact and reliable external-cavity semiconductor laser diodes at 1.5-μm can provide for a few milliwatt output powers. Hence, the output field of these devices needs to be strongly enhanced, by means of Fabry-Perot build-up cavities [3] or by the commercially available but yet expensive

0-7803-5400-1/99/ $ 10.00

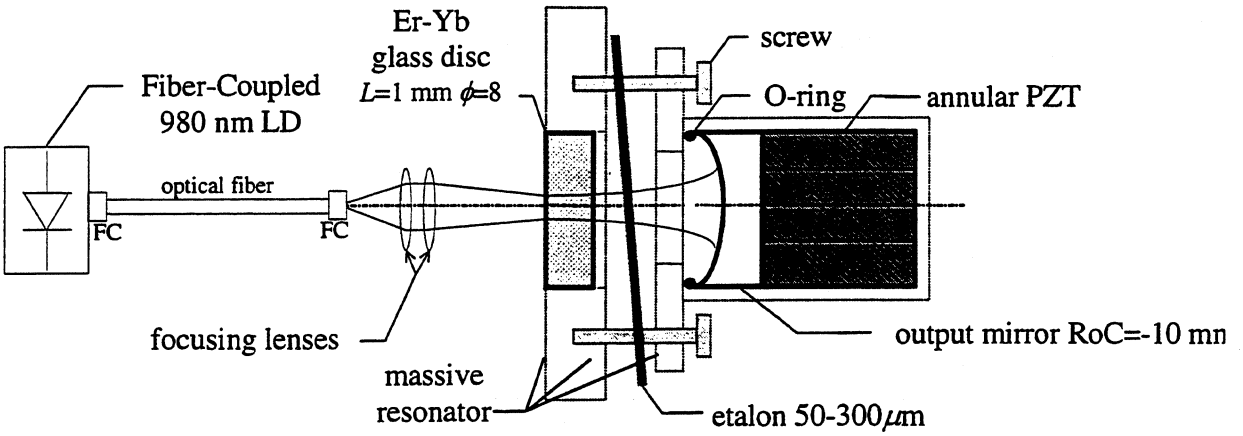


Figure 1. The Er-Yb tunable microlaser.

optical amplifyers [8], to observe sub-Doppler lines in this spectral region.

In this paper preliminary experiments on the frequency stabilization of Er–Yb laser against saturated absorption lines of $^{13}C_2H_2$ at 1.5 μm, are reported. A resonant Fabry-Perot cavity containing acetylene at low pressure is used to enhance the field intensity in order to saturate the 1.5-μm overtone transitions of the molecule. In particular the power spectral density of the laser frequency noise has been measured by means of the Pound-Drever-Hall scheme [9] and the Fabry-Perot resonator.

2. EBIUM LASER AND $^{13}C_2H_2$ FABRY-PEROT CELL

The laser source used in the experiment is schematically shown in Fig. 1. This single frequency Er–Yb:glass microlaser can be continuously tuned in the 1532-1536 nm wavelength range [10] and gives a diffraction limited output beam with an output power of several milliwatt. The laser cavity is based on a 2.5-mm long Er-Yb:glass disc (Kigre, QE-7 phosphate glass), longitudinally pumped at 978 nm wavelength by an InGaAs laser (Spectra Diode, SDL-6460). The output coupler consists of a concave mirror with 10-mm radius of curvature and 1% transmission and, for fine frequency tuning of the laser output, the mirror is glued to an annular piezoelectric transducer (PZT). To allow for single frequency operation and coarse wavelength selection, an uncoated 200 μm thick, BK7 etalon is inserted into the resonator. In the picture reported in Fig. 2 an example of practical realization of the Er-Yb:glass microlaser is shown.

To obtain the saturation of the acetylene lines, a Fabry-Perot resonator containing a sealed-off Brewster cell, filled with $^{13}C_2H_2$ at a gas pressure of 4 Pa, was used. As it is shown in Fig. 3, the Fabry-Perot consists of a plano-spherical cavity, with a nominal finesse of ≈200 and a free-spectral-range of 600 MHz. Moreover, in order to change the resonance frequency, the flat mirror was mounted on a piezo-electric transducer (PZT).

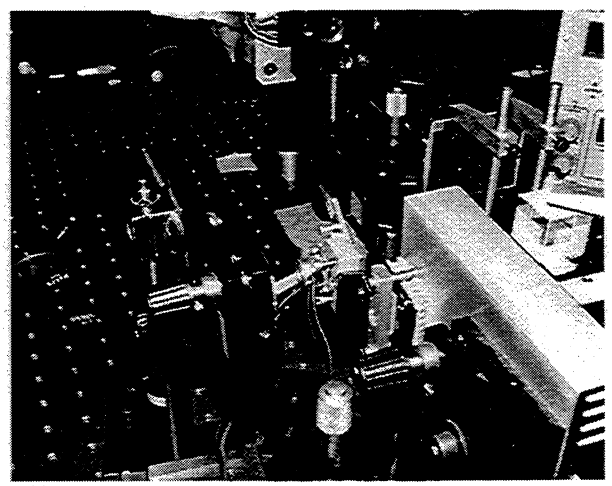


Figure 2. The Er-Yb tunable microlaser. A Peltier TEC is mounted on the laser cavity to allow for long-term frequency stability.

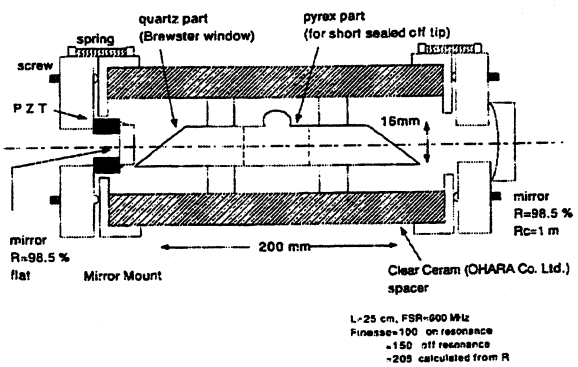


Figure 3. Plano-spherical Fabry-Perot resonator containing a sealed-off Brewster cell filled with $^{13}C_2H_2$ at 4 Pa pressure.

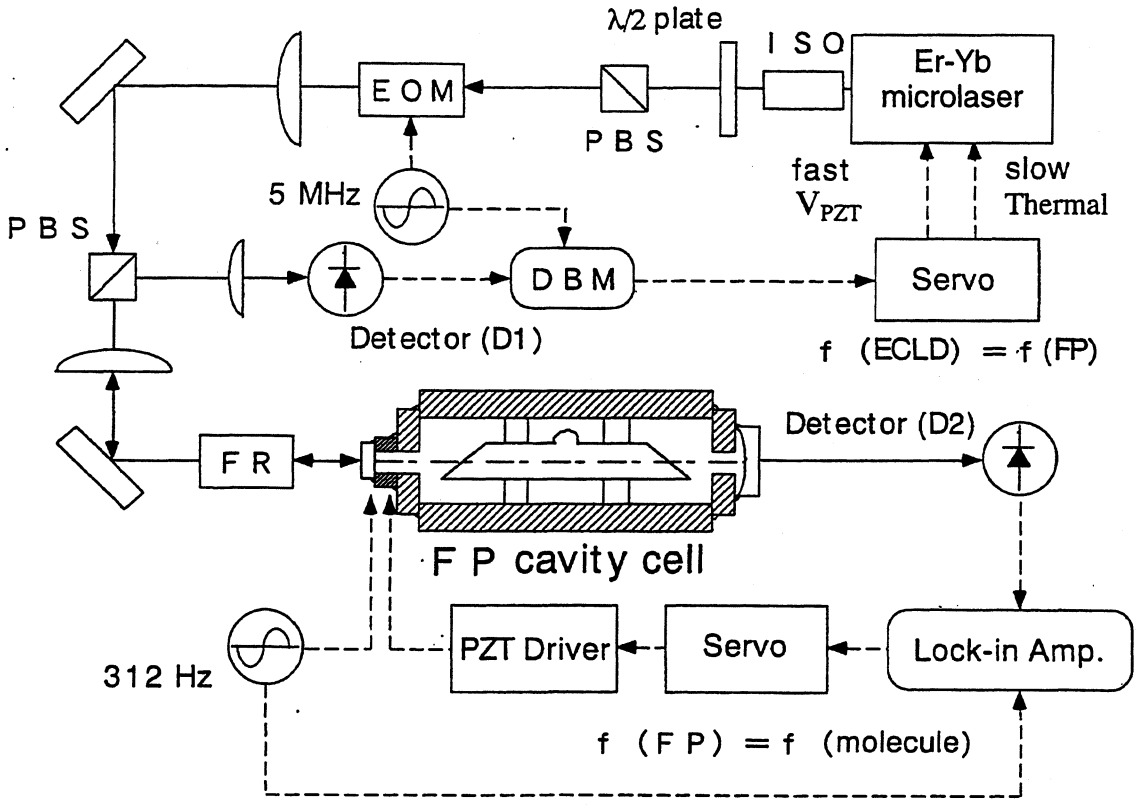


Figure 4. Experimental set-up to lock the laser to the FP cell (Pound-Drever-Hall technique with a modulation frequency of 5 MHz) and the FP cell to the $^{13}C_2H_2$ saturation dip (first harmonic method at 312 Hz)

3. EXPERIMENTAL SET-UP

To lock the laser frequency against Doppler-free acetylene lines, two stabilization loops [9] have to be used. According to the scheme shown in Fig. 4, one loop keeps the laser frequency locked to the Fabry-Perot resonance, whereas the second loop locks the resonance of the cavity to the saturated absorption line. The first opto-electronic loop uses a frequency modulation of 5 MHz, applied to an electro-optic modulator external to the laser cavity. This loop exploit as a discriminator signal the cavity reflection (Pound-Drever-Hall technique); the second loop uses the first-harmonic detection scheme (first derivative) by means of a lower modulation frequency (dithering of the cavity PZT) of 312 Hz.

4. PRELIMINARY RESULTS

In these preliminary experiments, as shown in Fig. 5, about 65 % of the 1.5 mW incident power was coupled into the Fabry-Perot resonator. By measuring the Full Width at Half Maximum of the reflected lines, a finesse of ≈150 has been estimated. This value is mainly limited by residual cavity losses due to the acetylene cell. When the laser frequency is in coincidence with a $^{13}C_2H_2$ absorption lines, usually line P(6), additional cavity losses further reduce the Fabry-Perot finesse to ≈100.

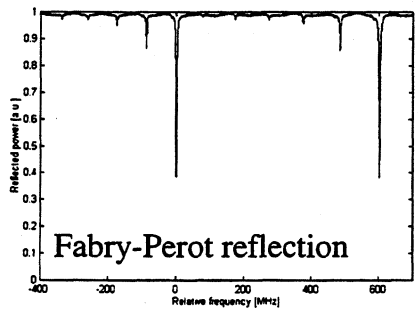


Figure 5. Reflection from the Fabry-Perot cavity.

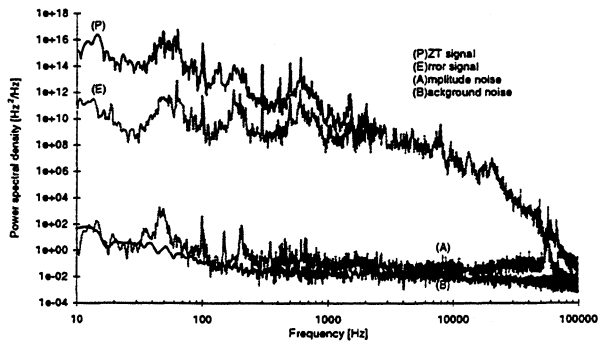


Figure 6. Power spectral density of the Er-Yb laser frequency noise.

By analyzing both the error and control signals of the first loop at 5 MHz, the laser frequency noise measurement has been realized. The laser was locked to the Fabry-Perot with a control bandwidth of about 2 kHz, essentially limited by laser PZT resonances. The obtained power spectral density laser frequency noise is reported in Fig. 6. The sensitivity of this measurement is fixed by the electronic noise at 10^{-2} Hz^2/Hz level for Fourier frequency higher than 1 kHz.

5. CONCLUSIONS

Diode-pumped Er-Yb microlaser with narrow linewidth can be a suitable source, locked to acetylene absorptions, to build frequency references in the 1.5 μm region of interest for High Density Wavelength Multiplexed and for coherent optical communication.

Efforts have been started to obtained cw linearly polarized single line microlaser which can provide enough power to saturate acetylene rotovibrational transitions inside a Fabry-Perot resonator.

REFERENCES

[1] M. Breton, P. Tremblay, C. Julien, N. Cyr, M. Tetu, and C. Latrasse, Optically pumped rubidium as a frequency standard at 196 THz, *IEEE Trans. Instrum. Meas.*, Vol. 44, pp. 162-165, 1995.

[2] A. Brunel, V. Mahal, I. Kiryushev, A. Arie, M. A. Arbore, and M. M. Fejer, Appl., ,"Frequency stability at the kilohertz level of a rubidium-locked diode laser at 192.114 THz, *Appl. Opt.*, Vol. 37, No. 27, pp. 6410-6414, 1998.

[3] A. Onae, K. Okumura, J. Yoda, K. Nakagawa, A. Yamaguchi, M. Kourogi, and K. Imai, Towards an accurate frequency standard at 1.5 μm based on the acetylene overtone band transition , *in Proc. of the Conference on Precision Electromagnetic Measurements*, pp. 400-401, Washington, July 1998.

[4] C. Svelto, E. Bava, S. Taccheo, and P. Laporta, Absolute frequency stabilization of two diode-pumped Er-Yb:glass lasers to the acetylene P(15) line at 1534 nm, *Appl. Phys. Lett.*, Vol. 73, No. 13, pp. 1778-1780, 1998.

[5] ITUT-T - *Draft recommendation G.692* (May 1997) "Optical interfaces for Multichannel Systems with Optical Amplifiers"

[6] New Focus catalog. The citation of a company's name is for sole purpose of technical communication.

[7] S. Taccheo, P. Laporta, S. Longhi, O. Svelto, and C. Svelto, Diode-pumped bulk erbium-ytterbium lasers, *Appl. Phys. B*, Vol. 63, No. 5, pp. 425-436, 1996.

[8] A. Onae, K. Okumura, Y. Miki, T. Kurosawa, E. Sakuma, J. Yoda, and K. Nakagawa, Saturation spectroscopy of an acetylene molecule in the 1550 nm region using an erbium doped fiber amplifier, *Opt. Commun.*, Vol. 142, pp. 41-44, 1997.

[9] R. W. P. Drever, J. L. Hall, F. V. Kowalski, J. Hough, G. M. Ford, A. J. Munley, and H. Ward, Laser phase and frequency stabilization using optical resonator, *Appl. Phys. B*, Vol. 31, pp. 97-105, 1983.

[10] S. Taccheo, P. Laporta, and C. Svelto, "Widely tunable single-frequency erbium-ytterbium phosphate glass laser", Appl. Phys. Lett., Vol. 68, No. 19, pp. 2621-2623, May 1996.

ABSOLUTE FREQUENCY MEASUREMENTS WITH A SET OF TRANSPORTABLE METHANE OPTICAL FREQUENCY STANDARDS.

O.Acef[1], A.Clairon[1], G.D.Rovera[1], F.Ducos,[1] L Hilico[1], G. Kramer[2], B.Lipphardt[2]
A. Shelkovnikov[3], E.Kovalchuk[3], E.Petrukhin[3], D.Tyurikov[3], M.Petrovskiy[3], M.Gubin[3], R.Felder[4], P.Gill[5], S.Lea[5]

[1] *BNM-LPTF, Observatoire de Paris, F-75014, Paris, France*
[2] *Physikalisch-Technische Bundesanstalt, Bundesallee 100, 38116 Braunschweig, Germany*
[3] *P.N.Lebedev Physical Institute, Moscow 117924, Russia*
[4] *Bureau International des Poides et Mesures, Pavillon de Breteuil F-92312, Sevres Cedex, France*
[5] *National Physical Laboratory, Teddington, Middlesex TW11 OLW, Great Britain*

ABSTRACT

For the first time absolute frequency measurements (AFM) referenced to Cs primary standards with a set of 3 double-mode He-Ne/CH_4 transportable optical frequency standards (TOFS, λ = 3.39 μm , ν = 88 THz) were carried out. The radio-optical frequency chain (ROFC) and the 29-THz CO_2/OsO_4 secondary frequency standard of BNM-LPTF (France) and the ROFC of PTB were involved in the experiment. The level of accuracy achieved for these cross-measurements is better than $5*10^{-13}$.

1. PROGRAM OF THE ABSOLUTE FREQUENCY MEASUREMENTS

The program of 1998 was a part of the joint research of several national laboratories and international organization in the field of precise spectroscopy and optical frequency metrology.

From 1991 to 1998 several sessions of AFM were carried out with the participation of Physikalisch-Technische Budesanstalt (Braunschweig, Germany). During this period a few different versions of transportable methane stabilized OFS developed in the Lebedev Institute (Moscow, Russia) were tested and calibrated with respect to Cs primary time standard at the ROFC of PTB.

Since 1995 the program of cooperation between Laboratoire primaire du temps et des fréquences (BNM-LPTF) and Lebedev Institute was started resulting in the first frequency comparisons of the stationary CO_2/OsO_4 OFS and He-Ne/CH_4 Transportable OFS in Nov.1996.

In 1997–1998 a few other laboratories expressed their interest to the stable methane frequency reference:

-National Physical Laboratory (NPL), Teddington, GB,
-Joint Institute for Laboratory Astrophysics (JILA) of University of Colorado and National Institute of Standards and Technology, Boulder, USA,
-Bureau International des Poids et Measures (BIPM), Sèvres, France.

This type of device having frequency repeatability near 10^{-13} could be used in conjunction with frequency synthesis to produce an absolute frequency measurements in the IR, visible, and UV ranges.

The goals of the present measurements were as follows:

1) comparison the renewed ROFC of BNM-LPTF and the ROFC of PTB;
2) independent calibration of the parameters of 3 new TOFS developed at Lebedev Institute in 1997-1998 and used now as frequency references at NPL, BIPM and Lebedev Institute.

The program of measurements realized in 1997-1998 included the following stages:

Dec.97 - TOFS(NPL) was calibrated on the PTB ROFC.

July.98 - TOFS(BIPM) was calibrated on the PTB ROFC.

Sept.98 -TOFS(BIPM), TOFS(Leb) were delivered to JILA and frequency calibration of a visible He-Ne/I_2 OFS (633nm) were made by mixing the radiation of Rb OFS (778nm) and 3.39 μm radiation [1].

Oct.98 - TOFS(BIPM), TOFS(NPL), and TOFS(Leb) were calibrated on the new ROFC of BNM-LPTF.

Nov.98 – TOFS(BIPM), TOFS(NPL), TOFS(Leb) were transported from Paris to Braunschweig and calibrated on the ROFC of PTB. Then back transportation the TOFS (LEB) from PTB to BNM-LPTF and recalibration on the BNM-LPTF ROFC have been done.

2. MAIN FEATURES OF THE APPARATUS INVOLVED

All transportable devices (Fig1.) were built on the same optical scheme of a precise optical spectrometer [2] and consisted of 3 lasers. Reference laser was frequency locked to the methane resonance by double-channel feedback loop (absorption + dispersion frequency discriminators were applied)[2].

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

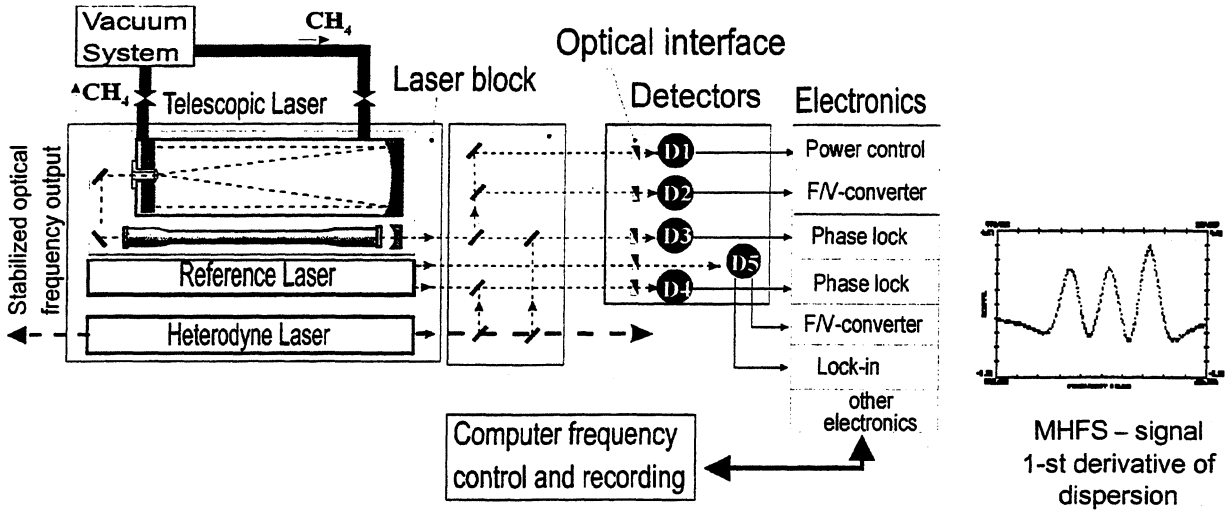


Fig.1 Scheme of the transportable methane frequency standard based on the double-mode He-Ne $\lambda = 3.39$ μm laser.

Heterodyne laser was offset phase locked to the Reference laser, and 3-rd laser with telescopic expander (beam waist in the CH_4 cell is 60mm) was offset frequency tuned via phase lock servo relatively to the Heterodyne laser frequency. The central component of triplet representing the magnetic hyperfine structure (MHFS) of the $F_2^{(2)}$–line of CH_4 (line-width about 4.0 kHz, see Fig.1 insert) was used for the long term frequency stabilization of the Heterodyne laser. Unmodulated Heterodyne emission served as output signal of the TOFS with constant offset relatively to the top of the central component of MHFS.

The phase-coherent ROFC of PTB [3,2] has been routinely used during many years for AFM up to the visible range. The performance of the chain estimated in the previous AFM with two TOFS (Leb) is practically ideal. It means that frequency errors are dependent of H-maser instability only amounting 10^{-14} -10^{-15} in the range of 100s-1000s averaging time.

The radio-optical frequency chain of BNM-LPTF has been completely refurbished in 1997. The most important improvement made since 1984 is the use of H-maser output as starting point of the AFM[4].

From April 1998 the chain was used to measure the absolute frequency of the stationary $CO2/OsO_4$ frequency standard of the BNM-LPTF [5].

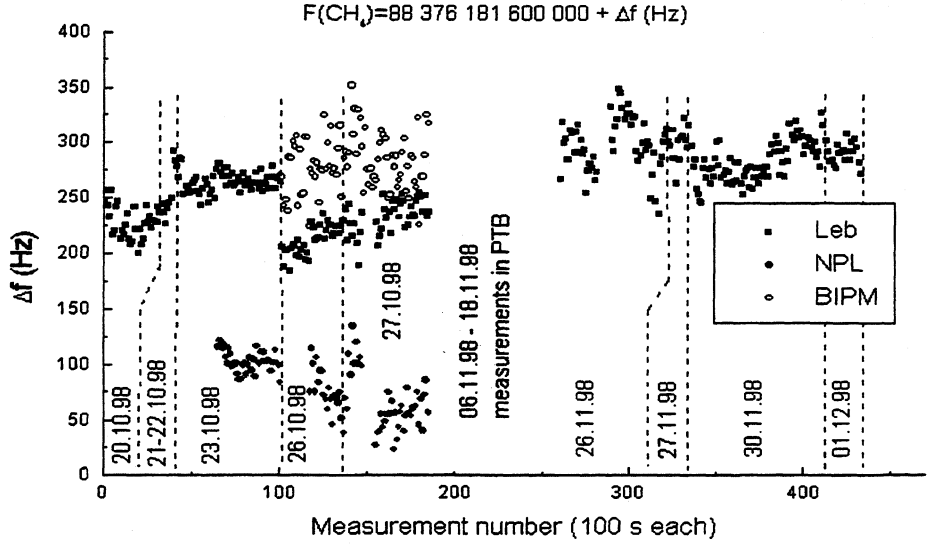


Fig.2 Results of AFM obtained in BNM-LPTF.

3. RESULTS OF AFM

Fig. 2 presents the series of **AFM in BNM-LPTF** in Oct. and Nov.-Dec. 98. with an interruption for the visit to PTB. The whole set of the data on Fig. 2 can be divided on 3 intervals.

Data on Oct.20-Oct.22. (Fig.3), represent preliminary measurements with TOFS(Leb). The device after transportation on the route Boulder-Paris as a luggage required some restoring and its parameters were adjusted several times during these 3 days.

Data on Oct.23 (Fig.2) were obtained with 2 systems only: TOFS(Leb), and TOFS(NPL). (The last was delivered from Teddington by car). AFM were made only for the BNM-LPTF CO_2/OsO_4 frequency standard. The frequency of He-Ne lasers were measured by counting the beat frequencies between OsO_4 standard and methane standards. The lasers were mounted on different optical tables. In this case the most reliable results were obtained. Estimation of frequency uncertainty for the data of Oct. 23 gives standard deviation σ of 100 s measurement as σ=8 Hz for TOFS(Leb) and σ=9 Hz for TOFS(NPL). This corresponds to the uncertainty of previous measurements (σ=6 Hz) obtained for TOFS(Leb) in the BNM-LPTF in 1996 [6].

Data on Oct.26-27 show the frequencies of all 3 methane standards measured simultaneously. The larger uncertainty increased up to 30 Hz due to more noisy conditions of the experiment. Mutual influence of these devices resulted from perturbation of the phase-lock servos by noise from mechanical pumps. and probably parasitic back reflections of the light from beats detection setup.

Some frequency shift between 23.10 and 26.10 of the TOFS(Leb) (about 50 Hz) was caused in part by temperature variation of about 5°C in the laboratory during weekend when work was stopped.

Second period of AFM at BNM-LPTF (26.11-01.12 Fig.2) was done with the HeNe/CH4-TOFS(Leb) and the CO_2/OsO_4 OFS (BNM-LPTF) connected to the ROFC. The standard deviation for the 4 days amounts 20 Hz for 100 s time of measurement.

The uncertainties of the AFM can not be associated only with methane standards frequency shifts. It was found that standard deviation of the frequency of the OsO4 –standard amounted to 2.9 Hz what efficiently gives near 9 Hz in the 3.39 μm range. Taking into consideration that many measurements in LPTF were not made synchronously with AFM of the OsO_4-standard, it should also contribute to the frequency scatter.

Fig.3 presents **results of AFM in PTB.** With better environmental conditions all 3 lasers demonstrated comparable frequency scatter with standard deviation near 10 Hz for 100 s sample time.

Several times during the measurements two devices – TOFS(NPL) and TOFS(Leb) – were realigned for the maximal output power, whereas the TOFS(BIPM) was aligned only once before the experiment. Some degradation of the MHFS signal resulted in some frequency drift of the TOFS(BIPM) during several days of AFM (see Fig.3).

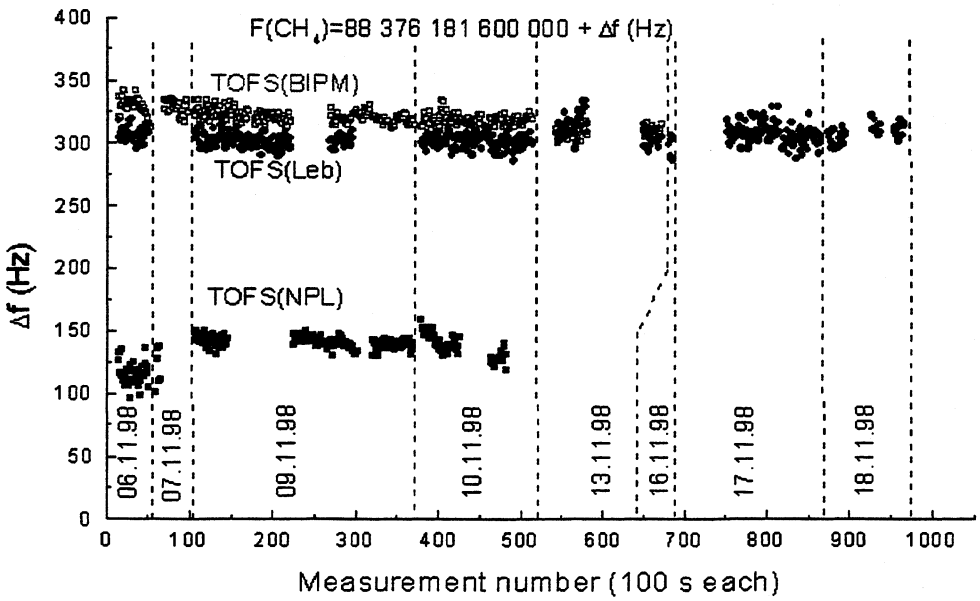


Fig.3 Results of AFM obtained in PTB.

In a separate experiment in PTB the frequency shifts of all methane standards were measured as a function of different parameters of the setup. There was an noticeable effect found which is caused by variations of the Heterodyne laser offset. The stabilized frequency exhibits a shift up to 50 Hz in the range of Heterodyne frequency offset variations from 550 kHz to 700 kHz. This shift was well reproducible for each device, so the effect is negligible if the Heterodyne frequency offset is constant. The origin of this shift is probably an effect of changes of frequency modulation index in the phase-lock servo.

The summarized results for average frequencies presented in the Table 1. All numbers are given for "standard" Het.- Ref .offset of (-600,0 kHz). Only during measurements in LPTF the value of Heterodyne offset for TOFS(NPL) and TOFS(BIPM) was set to –700 kHz and –550 kHz respectively. The Table 1 contains also data obtained in the previous sessions of AFM performed in Dec.97 with TOFS(NPL) and in Jul.98 with TOFS(BIPM).

Possible reasons of the averaged frequency difference between devices are discussed in Ref.[7].

Table 1. Results of AFM obtained in BNM-LPTF and PTB with methane standards of BIPM, NPL, and Lebedev Inst.

Chain:		LPTF	PTB
TOFS:	Date:	Δf = 3 Least of the TOFS	Significant Digits Frequency (Hz)
NPL	Oct.-Nov.98 Nov.97	71 σ=34	136 σ=11 126 σ=6
BIPM	Oct.-Nov.98 Jul.98	281 σ=25	320 σ=7 323 σ=9
Lebedev	Oct.-Nov.98 Nov. Dec. 98	239 σ=25 287 σ=32	273 σ=10

CONCLUSIONS

The AFM carried out in collaboration of several laboratories have shown that mid-Infrared transportable OFS based on double-mode He-Ne lasers stabilised over resolved MHFS of $F_2^{(2)}$ CH_4 line demonstrate frequency reproducibility (for different devices) of $1*10^{-12}$, and frequency repeatability (for one device during several months, year) of $2 * 10^{-13}$.

The research has proved that the transportable CH_4 OFS can disseminate the 88 THz reference frequency with an accuracy $\sim 10^{-13}$. Thus it opens new possibilities for laboratories interested in precise optical metrology without necessity of setting up the full radio-optical frequency chains.

Comparisons of the two radio-optical frequency chains fulfilled with the TOFS devices demonstrated the coincidence of the frequency records for the chains within interval of 40 Hz ($5*10^{-13}$). The prospects for decreasing the uncertainty of the future chain comparisons are clear now.

Acknowledgements. Financial support of the current investigations was partially provided by the U.S. Civilian Research and Development Foundation (CRDF Grant No.RE1-193), by the Russian Foundation for Basic Research (Grant No.98-02-16724) and the Russian Federal Scientific-Technical Program "Fundamental Metrology".

References

[1] J.Hall, L.S.Ma, M.Taubman, B.Tiemann, F.L.Hong, O.Pfister, and J.Ye, "Stabilization and frequency measurement of the I_2-stabilized Nd:YAG laser" in CPEM'98 Digest p. 151, Washington D.C. (USA) July 6-10, 1998.

[2] M.Gubin, D.Tyurikov, A.Shelkovnikov, E.Kovalchuk, G.Kramer and B.Lipphardt, "Transportable He-Ne/CH_4 Optical frequency Standard and Absolute Measurements...", IEEE, J. of Quant.Electr.,v.31, No.12, p.2177 (1995).

[3] G.Kramer, B.Lipphardt and C.O.Weiss, "Coherent Frequency Synthesis in the infrared", in Proc. IEEE Frequency Control Symp., May 27-29, 1992, pp.39-43.

[4] D.G.Rovera and O.Acef, in Proc. IEEE, Special issue of CPEM'98, Washington D.C. (USA) July 6-10, 1998, to be published.

[5] O.Acef, D.G.Rovera, F.Ducos, and L.Hilico, "Performance of OsO_4-stabilized CO_2 Lasers as Optical Frequency Standards near 29 THz", in Proc. Joint Meeting of the European Frequency and Time Forum and 1999 IEEE International Frequency Control Symposium, 13-16 April, 1999, Besancon, France (this issue).

[6] O.Acef, A.Clairon, L. Hilico, D.G.Rovera, G.Kramer, B.Lipphardt, A.Shelkovnikov, E.Kovalchuk, E Petrukhin, D.Tyurikov , M.Petrovsky, and M.Gubin, "Absolute Frequency Measurements and intercomparisons of CH_4 and OsO_4 stabilized...", in CPEM'98 Digest, p. 258, Washington D.C. (USA) July 6-10, 1998.

[7] M.Gubin, A.Shelkovnikov, E.Kovalchuk, D.Krylova, E.Petrukhin, and D.Tyurikov , "Present Performance of the Transportable He-Ne/CH_4 Optical Frequency Standard", in Proc. Joint Meeting of the European Frequency and Time Forum and 1999 IEEE International Frequency Control Symposium, 13-16 April, 1999, Besancon, France (this issue).

CONTRIBUTION TO THE DETERMINATION OF HIGH-ORDER ELASTIC STIFFNESSES FROM THE MEASUREMENT OF BT-CUT QUARTZ RESONATORS

Jaroslav Nosek, Lenka Kretschmerová, Petr Kretschmer
Technical University of Liberec, Hálkova 6, CZ-461 17 LIBEREC, Czech Republic
E-mail: jaroslav.nosek@vslib.cz

ABSTRACT

The paper deals with the determination of the high-order constants possibility in the quartz using the measurement of the intermodulation products of the BT- and other cuts quartz resonators. The method is based on the Tierstens work [6], which are provided on the AT-cut resonators vibrated in thickness - shear mode. The measurements performed on BT– cut quartz and other Y-cuts resonators given access to obtain of the effective non-linear constants of higher order. The experimental results and the comparison with the resonant method results are discussed.

1. INTRODUCTION

The properties of the acoustic waves of the resonator or of other piezoelectric selective devices are usually influenced by internal or external parameters on account of the non-linear properties of the crystal. These nonlinearities are at the origin of many non-linear effects, as frequency-amplitude effects, intermodulation, and the sensitising to various physical quantities, like electric field, temperature, mechanical forces, acceleration or pressure as it was shown by Gagnepain [2]. It is obvious that a part of this non-linear effects is non desired, and, at the opposite, nonlinearities can be used for sensing functions - Ref. [4], making the selective BAW or SAW resonators.

The description of the non-linear behaviour of the resonator requires the knowledge of not only the regular second order fundamental material constants, but also of the higher order constants of the third, and in some cases of the fourth order. The paper will be related to the phenomena resulting mainly in elastic nonlinearities of the quartz.

Since the piezoelectric coupling is small in quartz and the linear oscillatory behaviour of quartz is essentially elastic, it seems reasonable to assume that the nonlinear interaction in quartz is purely elastic and that the electroelastic interaction is due to linear piezoelectricity. This assumption is made in this treatment. Moreover, since in essentially thickeness-shear modes in trapped energy resonators, the wavenumber or decay number along the plate is much smaller than the thickness wave-number, only the thickness-shear deformation is included in the nonlinear interaction terms.

An analytical investigation of nonlinear interaction in AT-cut quartz plates vibrating essentially thickness-shear modes has been explicitly evaluated by Tiersten [6]. To determine some elastic coefficients, two nonlinear effects in the quartz resonators are studied:

- nonlinear forced vibrations (frequency-amplitude effect) [5], [12],
- intermodulation phenomena [11], [12].

In our investigations, the different Y-cuts of quartz resonators under intermodulation test are used for determination of elastic stiffnesses of the 4th order.

All resutls are obtained for quartz resonators vibrating on 1st harmonics. The resonators were made by Krystaly a.s. Hradec Králové, and automatically measured in Laboratory of Resonant Methods at Technical University of Liberec, Faculty of Mechatronics and Interdisciplinary studies.

2. INTERMODULATION OF PIEZOELECTRIC RESONATOR

The non-linear behaviour of the quartz Y-cuts resonators vibrating by thickness shear modes of deformation is the consequence of non-linear elastic properties of quartz. We kept the electric and electro-elastic terms linear.

The description of the intermodulation phenomena is made using the non-linear elastic coefficients of high order.

If two (or more) harmonic signal levels V_1 and V_2 with frequencies ω_1 and ω_2 positioned uniformly with respect to resonant frequency acts simultaneously in electronic circuit, the intermodulation signal of the voltage $\hat{V}_\Omega$ and angular frequency $\Omega = 2\omega_1 - \omega_2$ is created due to the cubic nonlinearities. This frequency Ω will be located also within the bandwidth, and therefore will not be filtered.

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

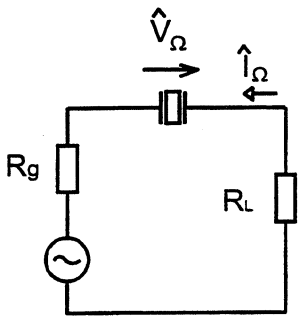


Fig. 1: Circuit with intermodulation current

The complex value of the intermodulation voltage for the circuit in Fig. 1 can be derived from Ref. [6]. Now, if we calculate the current $\hat{I}_\Omega$, after simple arrangement we get to an interesting relation for :

$$\hat{I}_\Omega = -(\hat{Y}_\Omega \cdot \hat{V}_\Omega + jA_{1\Omega} \cdot A_4 \frac{V_1^2 V_2}{\Delta\Omega}), \quad (1)$$

where $\hat{Y}_\Omega$ is the resonator admittance at angular frequency Ω. Significance of symbols for admittance $\hat{Y}_\Omega$, $A_{1\Omega}$, A_4 and $\Delta\Omega$ is shown in Ref. [11].
It follows from Eqn. (1), that the intermodulation current consist of a component dependent on the admittance $\hat{Y}_\Omega$ and a component affected by control signals of levels V_1, V_2, and effective elastic stiffness γ, which include the linear and non-linear elastic stiffness of 2nd, 3rd and 4th order.

In our case, the Eqn. (1) was used for the determination of nonlinear elastic stiffness γ of the Y-cuts quartz resonators introducing the measured values V_Ω, voltages V_1 and V_2, angular frequencies ω_1, ω_2, ω_N, Ω, and values of impedance $|Z|$ and φ for all frequencies. The values of S_e – surface of electrodes, 2h – thickness of the resonator, and material constants as nonlinear elastic stiffness, permitivity, and coefficient of electromechanical coupling for different cuts are introduced.

It is important so say that

$$\hat{V}_\Omega = |Z|\hat{I}_\Omega \quad (2)$$

and

$$Z \neq \frac{1}{\hat{Y}_\Omega} \quad . \quad (3)$$

3. DETERMINATION OF HIGH ORDER ELASTIC CONSTANTS

Tiersten [6] derived the relation, from which it is possible determine the nonlinear elastic stiffness $\tilde{c}_{6666}$ of AT-cut quartz resonator. Based on the same principle, we are trying to specify an effective elastic stiffness $\tilde{c}_{6666}$ of the 4th order for the BT-cut quartz resonators or some other cuts of quartz resonators.
Read the relation for a nonlinear stiffness

$$\gamma = \frac{1}{2}\tilde{c}^E_{2^{22}} + \tilde{c}^E_{3^{266}} + \frac{1}{6}\tilde{c}^E_{4^{6666}}, \quad (4)$$

where the symbols $\tilde{c}_{22}$, $\tilde{c}_{266}$, $\tilde{c}_{6666}$ denotes the effective elastic stiffnesses of the 2nd, 3rd, or 4th order. We express

$$\tilde{c}^E_{4^{6666}} = 6\gamma - 3\tilde{c}^E_{2^{22}} - 6\tilde{c}^E_{3^{266}}, \quad (5)$$

from Ref. [12].
The basic elastic stiffnesses are found out from tabular values of basic elastic stiffnesses of the 2nd and the 3rd orders for different cuts. Now, a computer program to carry the computations of effective elastic stiffnesses was made. It calculates effective elastic stiffness of the 2nd and the 3rd order for any cutting angle.
We found out that the stiffness $\tilde{c}_{6666}$ consists of five basic stiffnesses by means by the application of this program on the elastic stiffness of the 4th order:

$$\tilde{c}^E_{4^{6666}} = A_1 c^E_{4^{5555}} + A_2 c^E_{4^{5556}} + A_3 c^E_{4^{5566}} + A_4 c^E_{4^{5666}} + A_5 c^E_{4^{6666}} \quad (6)$$

The coefficients A_1, A_2, A_3, A_4 and A_5 in the relation (6) depend on the rotation angle of the resonator. We measure $\tilde{c}_{6666}$ for five different cuts of piezoelectric resonators and specify the constants A_1, A_2, A_3, A_4, A_5 for each of these cuts. We get a set of five equations. By the solution of this equations set, we obtain the values of basic elastic stiffnesses of the 4th order c_{5555}, c_{5556}, c_{5566}, c_{5666}, and c_{6666}.
Nevertheless, it is obvious that the values being obtained in this way neither allow to set the linkage with other elastic stiffnesses of the 4th order or the symmetry of this matrix. But, the elastic module $\tilde{c}_{6666}$ for any angle of rotated cuts of a quartz resonator can be determined by substitution in the relation (6).
The following processing was provided by the program for computing of the coefficients $A_1 - A_5$ in the equation (6), or in (7) generally:

$$\tilde{c}_{ijklmnop} = \sum_{abcdefgh} c_{abcdefgh}\alpha_{ia}\alpha_{jb}\alpha_{kc}\alpha_{ld}\alpha_{me}\alpha_{nf}\alpha_{og}\alpha_{ph} \quad (7)$$

In the Eqn. (7), α_{ia} - α_{ph} are direction cosines of rotation angles, which means that this is equation for uncompounded components of elastic stiffnesses tensor of 8th order. As seen in Fig. 2 of Ref. [12] , when indexes *i, j, k, l, m, n, o, p* and rotation angle α are entered, the program in the right window will write out the influence of stiffnesses $c_{abcdefgh}$ of basic matrix of elastic stiffnesses tensor of 8th order. In this description, the symbol α denote the angle of the rotation around X-axis. If the right side of the Eqn. (7) is converted back to compounded indexes with assuming symmetry, we get to Eqn. (6) in our case. The calculation of components of elastic stiffnesses tensor of 8^{th} order is shown in Ref. [12].

4. EXPERIMENT

The automated measurement set of the intermodulation products (IM), as signal of the V_Ω and angular frequency Ω is shown in the Fig. 2.

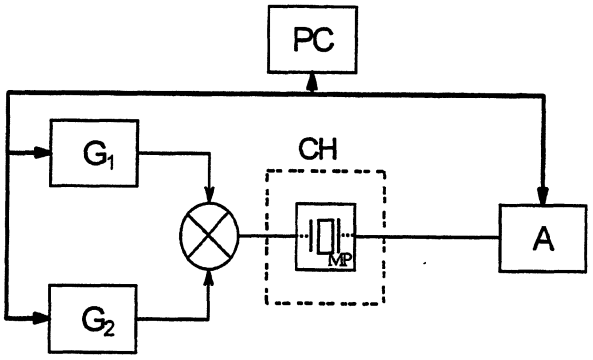


Fig. 2 Experimental set for IM products measurement

In the Fig. 2, the G1 and G2 are the generators of intermodulation test signals with angular frequencies ω_1, and ω_2 (given by HP 33120 A in the frequency range from 0,1 to - 13 MHz with the excitation level to 23,95 dBm for harmonics signal), CH is the test temperature chamber 4220 A Saunders with the 2 – port measurement device designed as MP,
A is Network/Spectrum analyser HP 4195 A , which is useful to 500 MHz in the frequency range, and from - 90 to +20 dBm of level of measured signal.

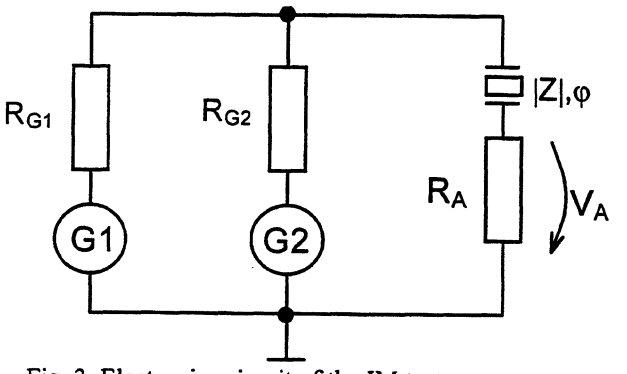


Fig. 3 Electronics circuit of the IM test

In the Fig. 3, the electric circuit used for direct measurement of intermodulation voltage V_Ω and angular intermodulation frequency Ω is shown.
The resonant frequencies ω_N , the impedance values $|Z|$ and phases φ for angular frequencies ω_1, ω_2 and Ω are measured by analyser HP 4192 A.

The 5 resonators for each type of Y-cut quartz resonator and 2 measurements for each resonator are realised.

Table 1: Calculated values of $|\gamma|$ [Nm^{-2}] of Y-cuts quartz resonators 10 MHz

Angle of Y-cut	-36,25	-35,2	-22,5	0	49,05	52,5
Resonator No.1	$1,13.10^{10}$	$1,84.10^{11}$	$7,11.10^{10}$	$9,16.10^{10}$	$2,06.10^{11}$	$5,50.10^{11}$
Resonator No.2	$6,02.10^{9}$	$2,74.10^{11}$	$6,77.10^{10}$		$1,76.10^{11}$	$5,83.10^{11}$
Resonator No.3	$5,14.10^{9}$	$3,28.10^{11}$	$7,25.10^{10}$	$5,98.10^{10}$	$1,70.10^{11}$	$4,97.10^{11}$
Resonator No.4		$4,26.10^{11}$	$6,23.10^{10}$	$1,62.10^{11}$	$1,56.10^{11}$	$5,45.10^{11}$
Resonator No.5		$2,19.10^{11}$	$6,14.10^{10}$	$1,18.10^{11}$		$5,03.10^{11}$
Dr [mm]	13,95	9	9	9	13,95	13,95
De [mm]	7	4	4	4	7	4

In the Table 1 the symbols Dr and De denote the diameter of the resonator and diameter of the electrode respectively.

The calculated results shows the small difference between the theory and experimental results. For the AT-cut resonators, the better results are given using the resonance method for determination of the nonlinear elastic stiffness γ .

5. CONCLUSION

This review shows that the intermodulation products of the quartz Y-cuts resonators depends not only on the 2nd and 3rd order fundamental elastic constants, but also on the 4th order ones. This is a serious difficulty for evaluating this effects, from such models, because the values of the 4th order fundamental elastic stiffness are almost completely unknowns, and even those of quartz crystal.
The measurement performed on quartz Y-cuts resonators give access to obtain of the effective non-linear constants. A computer determination of the basic non-linear constants of higher order is possible.

ACKNOWLEDGEMENT

This work was supported in part by the Grant of Ministry of Education of the Czech Republic under Contract No. VS 96006.
The authors wish to thank J. Zelenka from Technical University of Liberec for valuables recommendations.

REFERENCES

[1] J.Zelenka, Piezoelectric Resonators and their Applications. Prague: Academia, Amsterdam: Elsevier Science Publishers, 1986.

[2] J.J. Gagnepain, "Non-linear constants and their significance," in Proceedings of the 41st Annual Frequency Control Symposium, 1987, pp. 266-276.

[3] J.Nosek, "Some Second-order Effects in the Electromechanical Systems with Bulk- and Surface Acoustic Waves", in Proc. 11th European Frequency and Time Forum, 1997, pp. 243-248.

[4] J.Nosek , J.Zelenka, "Quartz Strip Resonator as a Temperature Sensor for Mechatronics", in Proc. of the 3rd International Conference Mechatronics 96 and M2VIP96, 1996, pp. 233-237

[5] J.Nosek, "Drive Level Dependence of the Resonant Frequency in BAW Quartz Resonators and his Modeling", IEEE Transactions on Ultrasonics, Ferroelectric, and Frequency Control (to be published in the Jul 1999).

[6] H.F. Tiersten, "Analysis of intermodulation in thickness-shear and trapped energy resonators", J.Acoust.Soc.Am., vol. 57, pp. 667- 681, March 1983.

[7] H.F.Tiersten, D.S.Stevens, "An analysis of nonlinear resonance in contoured-quartz crystal resonators", J. Acoust. Soc. Am., vol. 80, pp. 1122-1132, 1986.

[8] Ch.K.Hruska , IEEE Trans.Son.Ultrason. SU-28, pp. 108, 1981.

[9] J.Nosek,, J.Zelenka, "Recent Activities of Crystal Properties at the Technical University of Liberec", Annales de Chimie. Science des Matériaux. Masson Paris, Milan, Barcelone, vol. 22, pp. 661-668, 1997.

[10] J. Nosek, "New Experimental Possibilities of Investigation of Basic Properties of Crystals and Smart Materials", Annales de Chimie. Science des Matériaux. Masson Paris, Milan, Barcelone, vol. 22, pp. 691- 693, 1997.

[11] J. Nosek , "Intermodulation products of quartz resonators and their measurement", in Proc. of 9th European Frequency and Time Forum EFTF 95, 1995, pp. 447-449.

[12] J.Nosek, L.Kretschmerova, P.Kretschmer , „Determination of high – order constants using the resonant and laser interferometry methods", in Proceedings of the 12th European Frequency and Time Forum, 1998, pp. 72-75.

Calculations of Parameters of BAW Microwave Composite Resonators

B. Dulmet, G. D. Mansfeld

Laboratoire de Chronométrie Electronique et Piézolectricité
Ecole Nationale Supérieure de Mécanique et des Microtechniques,
26 Chemin de l'Epitaphe F 25000 Besançon
Bernard.Dulmet@ens2m.fr

Institute of Radioeengineering and Electronics, Moscow
Russian Academy of Sciences, 11 Mokhovaia st, 103907, Moscow, Russia.
mans@mail.cplire.ru

Abstract

This paper uses available models to calculate the main operating parameters of BAW microwave composite resonators. Although such resonators can be used as devices for characterization of the underlying substrate, they can be expected to reach comparatively good performances if the substrate is made from very low loss material, such as YAG driven into vibration by a thin layer of ZnO or AlN. Operating on very high overtone is then quite feasible and solves many technological issues related to the achievement of actually thin membranes. Also, earlier investigations have shown a surprisingly efficient energy trapping on such resonators. The paper presents first investigations about microwave composite resonators operating around 3GHz, using existing analytical models and shortly discusses issues attached to a two-dimensional FEM. The authors conclude that this technique should evolve to fully develop its expected benefits and must still be substitutted by analytical calculations to allow for accurate calculation of parameters in actual devices. Use of existing model for energy trapping analysis seems possible even for very high overtone numbers.

I. Studied Configuration

We have studied a BAW composite resonator in the ZnO on YAG configuration shown on Fig.1. The resonator is designed to operate on very high overtone, around 300. In this circumstance, the thickness of ZnO layer is very small in compare with thickness of substrate (0.70 *vs.* $430\mu m$). A large mass electrode covers the whole upper surface of substrate, and a small electrode is deposited on top of the ZnO layer. Both electrodes are made from Al $80nm$ thick. The c-axis of ZnO is normal to the substrate surface, so that thickness extensional modes are driven into vibration. The purpose of this work was to validate motional elements predicted by unidimensionnal model in compare with results obtained from 3D model. Checking effectiveness of energy trapping of the studied configuration of Fig.1 was especially wished from the point of view of Q-factor.

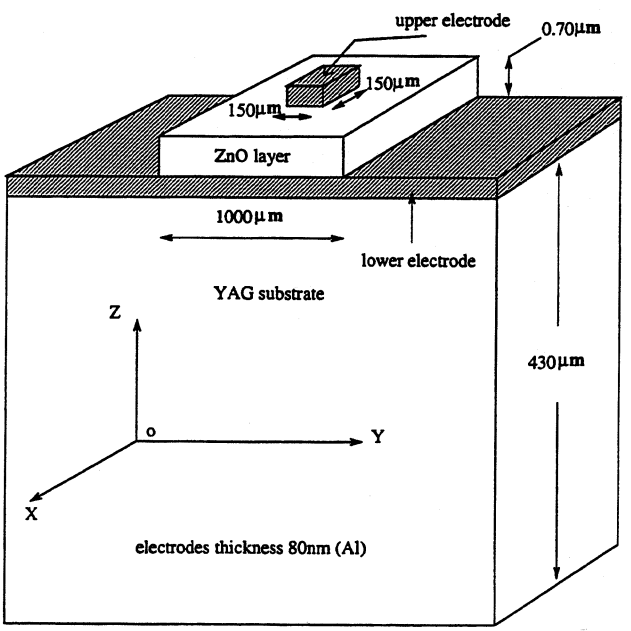


Figure 1: Studied configuration of YAG-ZnO composite resonator

II. Basic modelling

Since the thickness of electrodes is not so small with respect to thickness of ZnO layer (only one order of magnitude), the impedance method analysis takes into account the electrodes as a medium for wave propagation, as shown on Fig.2. Then the impedance of composite structure is given by

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

$$Z_e = \frac{1}{\omega C_0}\left[1 + \frac{K_t^2}{qd}\frac{i\left(Z_1^{in}+Z_2^{in}\right)\sin qd - 2Z_t^2(1-\cos qd)}{\left(Z_t^2+Z_1^{in}Z_2^{in}\right)\sin qd - i\left(Z_1^{in}+Z_2^{in}\right)Z_t\cos qd}\right] \quad (1)$$

where d is the thickness of the piezoelectric layer and q the complex wavenumber in it according to notations of Fig.2. After some manipulations, the equivalent circuit can be put in the form given on Fig.2, *i.e.* as a succession of series connected equivalent circuits, each one describing a given overtone

$$Z_e = \frac{1}{i\omega C_0'} + \sum_{n=1}^{\infty}\frac{i/\omega C_n}{\left(\frac{\omega_n}{\omega}\right)^2 - 1 + \frac{i}{\omega C_n R_n}} \quad (2)$$

and

$$\frac{C_0'}{C_0} = 1 - \frac{\sin q'd + (1-\cos q'd)\sin(q'd+2\phi_2)}{q'd/K_t^2} + \frac{\alpha}{2} \quad (3)$$

where q' is the real part of q, and α only depends on electromechanichal coupling factor of piezoelectric layer and properties of electrodes layers (See [1] for more details). Parallel resonances frequencies ω_n can be found by equating the real part of the acoustic path θ in the whole structure to $n\pi$, where the integer n is the overtone number, thereby fixing the values of all wavenumbers in structure. The motional parameters for n-th overtone can then be obtained by closed forms in terms of above mentionned quantities also presented in [1], and which only depend on n, θ, C_0, ω_n and α. The Q-factor is thus given by $Q \approx n\pi/2\Im\theta$. Data for losses in YAG are available (See [1] for a more general review on connected topics) so that the complex impedance can be simulated in terms of frequency. For the here-studied configuration, the following results were predicted at 3GHz :

f_a	3.034663 GHz
f_r	3.034524 GHz
Q	4.46 10^5

Two resonators with Lu-doped substrate were made and measured at IREAS. The impedance curve of Fig.3 was obtained on the best sample, which indicates a value of Q-factor

$$Q \approx 0.84 \quad 10^5 \quad (4)$$

which is extremely good in compare with known results for BAW resonators on quartz, LiNb03, or ZnO/Si substrates but is cleary below predicted value.

Although the model does include elastic effects in electrodes layers, it does not account for trapping effect, which may impact on the Q-factor, in case it is not sufficient. So, some effort was devoted to evaluating the transverse behavior of resonant modes in the studied configuration. We initially planned to investigate it by use of 2D FEM analysis, since a previously presented 3D program exists at LCEP

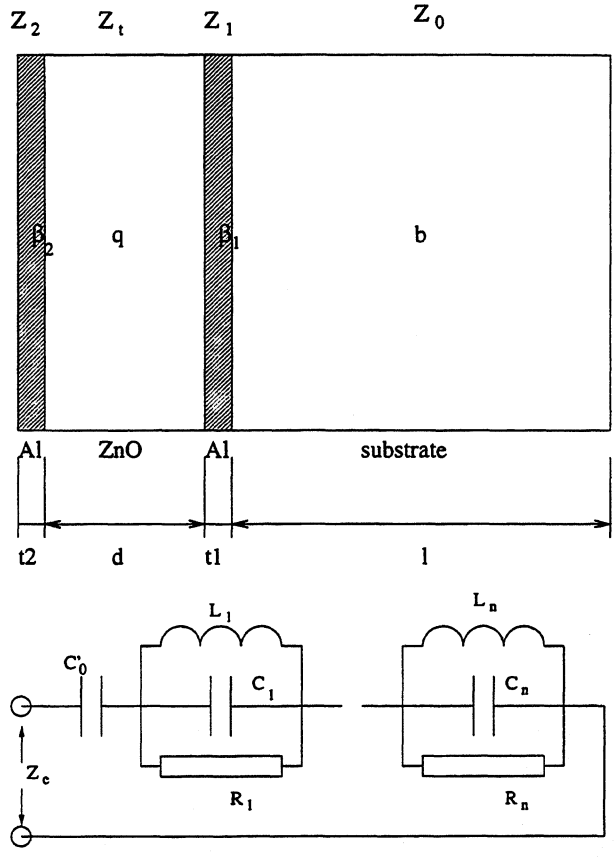


Figure 2: Impedance model for high overtone composite structure

for the modal analysis of piezoelectric resonators ([2-3]. Membrane elastic effects in thin electrode layers are taken into account. At the moment, only classical solving methods are considered for large eigenvalues problem resulting from the required dense mesh along the thickness. The program makes use of cubic interpolating functions along the thickness of plate, which weer found much more suitable for the modelling of sine waves than parabolic interpolations encountered in most 3D FEM programs. Unfortunately, it was found that, due to the extremely high overone number of studied configuration, the bandwith of obtained matrix is so large that even switching to 2D analysis results in intractable problem at the moment, unless special techniques could be implemented to drastically reduce the size of problem by restricting the FEM analysis to electrodes and piezoelectric layers, for instance. This was eventually left for further research and we put into effect calculations from the analytical model of Tiersten and Stevens [4] for composite structures.

The model consists of two parts : first, finding the dispersion equations for electroded and unelectroded regions, and second, matching the solutions at the interfaces betwen regions, after the dispersion equation is extended as a partial derivatives equations in terms of transverse coordinates (corresponding to in-plane directions in the treatment en-

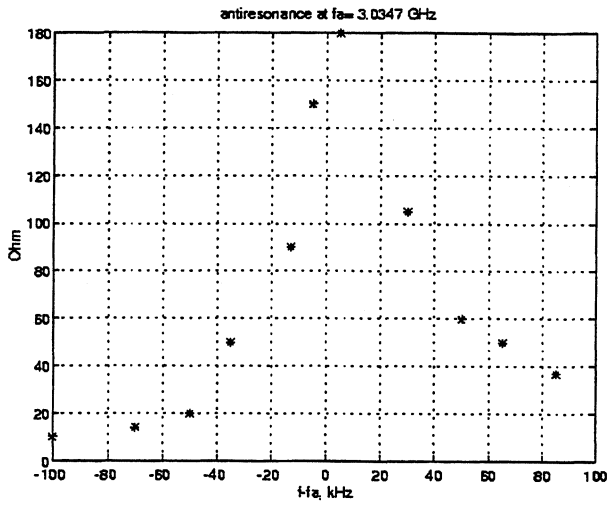
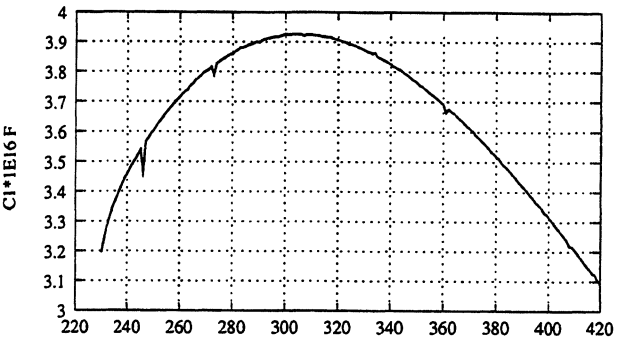


Figure 3: Measured impedance curve

p



Motional capacitance vs overtone number

Figure 4: Motional series capacitance in terms of overtone number

ergy trapped thin plates, see [5] for instance). In this model, only the mass loading effect of electrodes is taken into account,not the elastic effect. The steps of initial part can be summarized as follows :

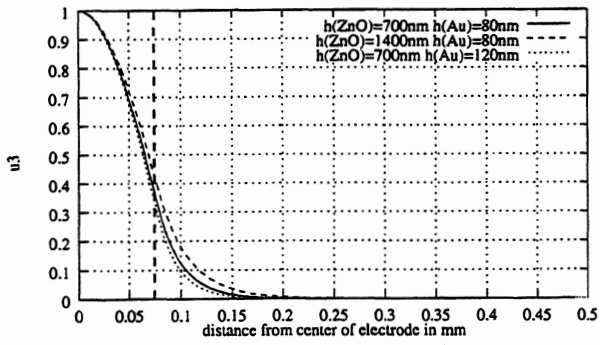
- finding unidimensional thickness solutions for *extensional modes*, with rigourous boundary conditions, both in electroded and unelectroded regions.
- asymptotic expansion of the solutions of balance equations around this limiting case when a small lateral wavenumber component is present. This should be made both in piezo layer and in substrate. Quasi–thickness extensional and shear modes appear already coupled at this level.
- combination of independent solutions of propagation equations in the bulk, obtained for dominant thickness extensional and thickness shear modes to satisfy boundary and continuity conditions at interfaces between layers.
- solving by Taylor's expansion the determinantal equation for boundary conditions.

The solutions in electroded and unelectroded regions differ only by their thickness–related terms. We carefully checked all formulas and followed the approach detailed at Ref [1]. Although the model yields predicted values of C_0, C_1, L_1, they are obtained in the framework of series equivalent circuit and thereby cannot be directly compared with data from the impedance model. Since the model gives closed form for the displacement amplitude along transverse directions, we used it to plot the results of Fig 5 which represents a semilog plot of the squared amplitude of thickness extensional component (normalized to 1 at the center of resonator) along transverse axis. Due to cubic and hexagonal symmetries of YAG and ZnO and symmetric shape of the upper electrode, the same predicted dependences are identical along both transverse directions. Since the ZnO layer is deposited as a strip $1000\mu m$ wide, we plot u_3^2 from the center of resonator up to $X_1 = 500\mu m$. In this model, energy trapping effect occurs only from the matching between solutions in electroded and unelectroded regions and the boundary conditions at the edges are not taken into account. Results of Fig 5 show that the actual configuration is well trapped since the predicted u_3^2 (practically proportional to the vibrating energy) remaining at the edges of resonator is below 10^{-10}, which is sufficiently small for the model to hold. In BAW resonator construction, it is admitted that a ratio 10^{-9} between predicted amplitudes at the edge and at the center is sufficient to get optimal values of Q-factor. Comparatively, Fig.6 shows that the residual amplitude of anharmonic modes at the edge is very high, so that such modes are practically killed since they are prone to dissipate a non negligible fraction of energy (a few 1/1000s) in the periphery of resonator. Interestingly enough, the predicted $C_1 = f(n)$ curve presented on Fig.4 shows an optimal value of the motional series capacitance near overtone number 300, which is in agreement with the fact that the amplitude distribution along the thickness of piezoelectric layer is extremely close to an exact half wavelength.

From this analysis, it was found that predicted energy trapping should be strong enough to allow for optimal value of Q-factor. Measurements of the roughness of surface performed by AFM microscopy during a stay of G.D. Mansfeld at LCEP indicated an averaged value of 14nm which is expected to lower the Q-factor by several units from the optimal value for perfectly smooth surface.

IV. Conclusion

The analysis shows that the energy trapping in the composite resonator is better achieved when the electrodes do not cover the whole ZnO area. Actual parameters of the experimental configuration were found to allow for an optimal value of C_1 parameter and a sufficiently strong trapping to



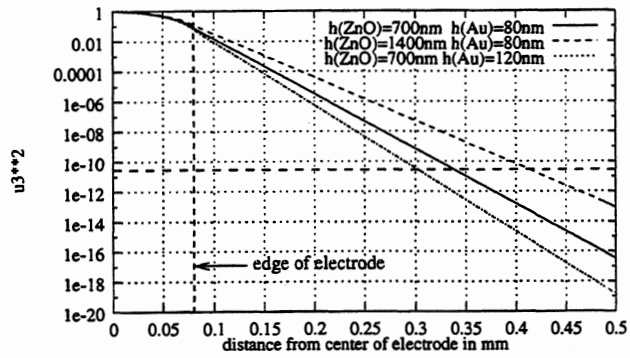
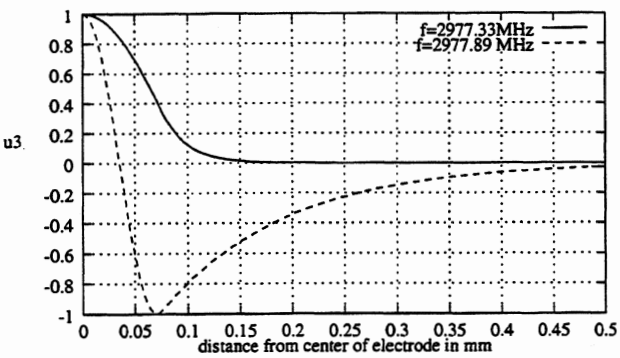


Figure 5: $|u_3|^2 = f(X_1)$ for overtone mode



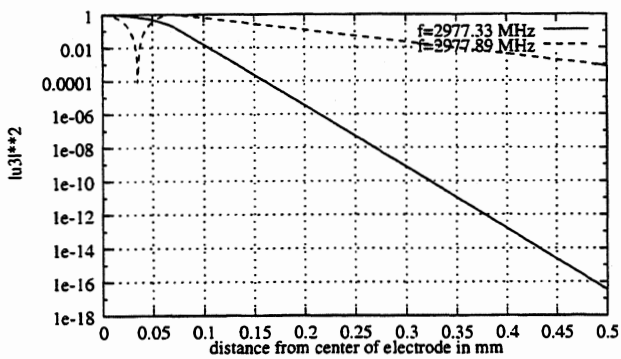


Figure 6: $|u_3|^2 = f(X_1)$ for anharmon,ic mode

ensure a good Q-factor. Although experimentally obtained Q factor is intrisically very good, it is below the expected value, and measurements of the roughness of surface tend to explain this behavior. FEM analysis has to be dramatically worked out to allow for the modelling of overtone numbers as high as 300, even if benefiting from simplification down to a 2D model. Further commitment is still necessary to gather modelling of elastic effects in electrode layers and transverse variations of mode in a single state-of-art analytical description.

References

[1] G.D. Mansfeld, *Modern State and Perspectives of Bulk Acoustic Wave Composite Resonators*, in the same proceedings as the present paper, 1999.

[2] F. Raoelijaona, B. Dulmet, *Finite element analysis using trigonometric interpolations for high overtones thickness modes resonators*, Proc. EFTF, pp 121–126, 1995.

[3] F. Raoelijaona, B. Dulmet, J.T. Stewart, D.S Stevens, *FEM Modeling of Spectrum and High Order Thermal Sensitivities of Flat Plate Resonators Operating in Fundamental Modes of essentially Thickness Shear*, Proc. E.F.T.F, pp 24–28, 1997.

[4] H.F. Tiersten, D.S Stevens, *An Analysis of Thickness-Extensional Trapped Energy Resonant Device Structures with Rectangular Electrodes in the Piezoelectric Thin Film on Silicon Configuration*, J. Appl. Phys. **54**, pp 5893 sq, 1983.

[5] H.F. Tiersten, *Analysis of Trapped Energy resonators Operating in Overtones of Coupled Thickness Shear and Thickness Twist*, J. Acoust. Soc.Am., **59**, pp879 sq, 1976.

EXTENSIONAL VIBRATIONS OF PIEZOELECTRIC CRYSTAL PLATES AND STRIPS

P. C. Y. Lee & Nicholas P. Edwards

Department of Civil Engineering & Operations Research
Princeton University, Princeton, New Jersey 08540 USA

ABSTRACT

A system of approximate second-order governing equations are extracted from an infinite system of 2-D equations for piezoelectric plates with general symmetry and charge-free faces which is deduced from the 3-D equations of linear piezoelectricity using a similar approach as that in Lee, Yu and Lin. These equations improve the accuracy of calculated dispersion curves as compared with those from the 3-D equations because a new truncation procedure has been employed, and they are applicable for frequencies up to and including the cut-off frequencies of the first symmetric thickness-stretch and the second symmetric thickness-shear modes. Furthermore, a system of 1-D second-order equations for the vibrations of piezoelectric strips with thin rectangular cross-section is deduced by averaging the 2-D equations over the narrow width of the plate. Closed form solutions of these 1-D equations are obtained for the second-order extensional vibrations of a barium titanate strip with finite length and narrow rectangular cross-section. Predicted frequency spectra are compared with the experimental data by Onoe and Pao and calculated results by Medick and Pao with good agreement. The present second-order theory for plates and strips is an improvement in that it contains no correction factors, includes the piezoelectric effect and is applicable for any crystal symmetry.

1. INTRODUCTION

In a previous paper[1], a new term representing displacements which vary linearly through the thickness (x_2 direction) was included with the infinite series expansion of trigonometric functions of the thickness direction. This term, induced by the gradients of the transverse displacement, improved the dispersion curves for antisymmetric vibrations of piezoelectric plates especially at lower frequencies. The first order equations including thickness shear, flexure, and face shear varying in the x_1 direction, and thickness twist and face shear varying in the x_3 direction predict frequency spectra that agrees closely with experimental results by both Koga and Fukuyo[2] and Nakazawa, Horiuchi, and Ito[3].

In the present paper, the same expansion is adopted for the mechanical displacements. This expansion has the advantage of having no correction factors and predicting the correct cut-off frequencies for any order theory. The electric potential is expanded in a series of cosine functions which easily satisfies charge-free boundary conditions. A set of approximate second-order equations is extracted from the infinite set of 2-D equations by using a special truncation procedure. The idea is to eliminate the third and fourth order displacements without setting them to zero. The relations for these higher order displacements are found by setting associated stresses to zero. This approach allows for the free development of motions which are coupled with the second-order equations. The truncation procedure has a direct effect on the accuracy of the dispersion curves calculated from the 2-D equations as compared to the 3-D solution. The most common improvement is in the first order thickness stretch branch and the second order thickness shear branch which results from the elimination of the fourth order thickness shear mode and the third order thickness stretch mode respectively.

2. SERIES EXPANSION

The components of mechanical displacement and electric potential are expanded in an infinite series of trigonometric functions of the thickness coordinate as follows

$$\begin{aligned} u_j &= -u_{2,j}^{(0)} x_2 + \sum_{n=0}^{\infty} u_j^{(n)} \cos\frac{n\pi}{2}(1-\psi_2), \\ \varphi &= \sum_{n=0}^{\infty} \varphi^{(n)} \cos\frac{n\pi}{2}(1-\psi_2), \end{aligned} \tag{1}$$

$$\text{with } u_j = u_j(x_1, x_2, x_3, t), \qquad \varphi = \varphi(x_1, x_2, x_3, t),$$
$$u_j^{(n)} = u_j^{(n)}(x_1, x_3, t),$$
$$\varphi^{(n)} = \varphi^{(n)}(x_1, x_3, t), \qquad \psi_2 = \frac{x_2}{b}.$$

The expansion $(1)_1$ has been shown to produce accurate dispersion curves and frequency spectra in a previous paper.[1] In $(1)_2$ the cosine series allows for the easy satisfaction of charge free faces.

3. 2-D EQUATIONS

The infinite set of 2-D equations are obtained in the same manner as in Lee *et al.*[1] and are the same as the previous work except for the following expressions:

$$\begin{aligned} E_i &= \sum_{n=0}^{\infty}\left[E_i^{(n)} \cos\frac{n\pi}{2}(1-\psi_2) + \bar{E}_i^{(n)} \sin\frac{n\pi}{2}(1-\psi_2)\right], \\ E_i^{(n)} &= -\varphi_{,i}^{(n)}, \qquad \bar{E}_i^{(n)} = -\delta_{2i}\frac{n\pi}{2b}\varphi^{(n)}, \\ D^{(n)} &= D_2(b) - (-1)^n D_2(-b), \\ D_{i,i}^{(n)} &- \frac{n\pi}{2b}\bar{D}_2^{(n)} + \frac{1}{b}D^{(n)} = 0, \qquad n \geq 0, \\ T_{ij}^d &= \frac{2}{3}c_{ijkl}S_{kl}^d + \sum_{m=0}^{\infty} c_m\left[c_{ijkl}S_{kl}^{(m)} - e_{kij}E_k^{(m)}\right] \\ &\quad + \sum_{m=0}^{\infty} s_m\left[c_{ijkl}\bar{S}_{kl}^{(m)} - e_{kij}\bar{E}_k^{(m)}\right], \end{aligned} \tag{2}$$

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

$$T_{ij}^{(n)} = c_n c_{ijkl} S_{kl}^{d} + (1+\delta_{n0})\left[c_{ijkl}S_{kl}^{(n)} - e_{kij}E_k^{(n)}\right] + \sum_{m=0}^{\infty} B_{mn}\left[c_{ijkl}\bar{S}_{kl}^{(m)} - e_{kij}\bar{E}_k^{(m)}\right],$$

$$\bar{T}_{ij}^{(n)} = s_n c_{ijkl} S_{kl}^{d} + (1-\delta_{n0})\left[c_{ijkl}\bar{S}_{kl}^{(n)} - e_{kij}\bar{E}_k^{(n)}\right] + \sum_{m=0}^{\infty} B_{nm}\left[c_{ijkl}S_{kl}^{(m)} - e_{kij}E_k^{(m)}\right],$$

$$D_i^{(n)} = c_n e_{ikl} S_{kl}^{d} + (1+\delta_{n0})\left[e_{ikl}S_{kl}^{(n)} + \varepsilon_{ik}E_k^{(n)}\right] + \sum_{m=0}^{\infty} B_{mn}\left[e_{ikl}\bar{S}_{kl}^{(m)} + \varepsilon_{ik}\bar{E}_k^{(m)}\right],$$

$$\bar{D}_i^{(n)} = s_n e_{ikl} S_{kl}^{d} + (1-\delta_{n0})\left[e_{ikl}\bar{S}_{kl}^{(n)} + \varepsilon_{ik}\bar{E}_k^{(n)}\right] + \sum_{m=0}^{\infty} B_{nm}\left[e_{ikl}S_{kl}^{(m)} + \varepsilon_{ik}E_k^{(m)}\right],$$

The face and edge conditions are also obtained in the same manner as Lee *et al.*, and they are not given here. The important difference in the 2-D equations is that the charge on the faces is easily specified when the cosine series expansion is used and the potential on the faces is easily specified when the sine series expansion is used.

4. SECOND ORDER EQUATIONS

From the infinite system of the preceding section, a finite set of approximate equations are extracted to form the second-order equations for piezoelectric crystal plates in which the zero, first, and second order displacements are kept and the higher order displacements are eliminated. The fifth and higher order displacements are set to zero:

$$u_j^{(n)} = 0, \qquad \text{for} \quad n \geq 5. \tag{3}$$

The third and fourth order displacements can be eliminated by setting

$$\begin{aligned} &T_{21}^{(2)} = 0, \quad T_{22}^{(2)} = 0, \quad T_{23}^{(2)} = 0, \\ &T_{21}^{(3)} = 0, \quad T_{22}^{(3)} = 0, \quad T_{23}^{(3)} = 0, \end{aligned} \tag{4}$$

and then solving for the third and fourth order displacements as functions of the zero, first, and second order displacements. However, we use the alternative relations defined by

$$\begin{aligned} &\bar{T}_{21}^{(3)} = 0, \quad \bar{T}_{22}^{(3)} = 0, \quad \bar{T}_{23}^{(3)} = 0, \\ &\bar{T}_{21}^{(4)} = 0, \quad \bar{T}_{22}^{(4)} = 0, \quad \bar{T}_{23}^{(4)} = 0, \end{aligned} \tag{5}$$

because they lead to expressions involving lower order displacements when compared to expressions from equation (4). The following simple expressions are obtained from equations (5):

$$\bar{S}_{21}^{(3)} = -\frac{2}{3\pi}\frac{c_{21kl}}{c_{2121}}S_{kl}^{(0)} - \frac{6}{5\pi}\frac{c_{21kl}}{c_{2121}}S_{kl}^{(2)} + \frac{2}{3\pi}\frac{e_{k21}}{c_{2121}}E_k^{(0)} + \frac{6}{5\pi}\frac{e_{k21}}{c_{2121}}E_k^{(2)},$$

$$\bar{S}_{22}^{(3)} = -\frac{4}{3\pi}\frac{c_{22kl}}{c_{2222}}S_{kl}^{(0)} - \frac{12}{5\pi}\frac{c_{22kl}}{c_{2222}}S_{kl}^{(2)} + \frac{4}{3\pi}\frac{e_{k22}}{c_{2222}}E_k^{(0)} + \frac{12}{5\pi}\frac{e_{k22}}{c_{2222}}E_k^{(2)},$$

$$\bar{S}_{23}^{(3)} = -\frac{2}{3\pi}\frac{c_{23kl}}{c_{2323}}S_{kl}^{(0)} - \frac{6}{5\pi}\frac{c_{23kl}}{c_{2323}}S_{kl}^{(2)} + \frac{2}{3\pi}\frac{e_{k23}}{c_{2323}}E_k^{(0)} + \frac{6}{5\pi}\frac{e_{k23}}{c_{2323}}E_k^{(2)}, \tag{6}$$

$$\bar{S}_{21}^{(4)} = -\frac{1}{2\pi}\frac{c_{21kl}}{c_{2121}}S_{kl}^{d} - \frac{8}{15\pi}\frac{c_{21kl}}{c_{2121}}S_{kl}^{(1)} + \frac{8}{15\pi}\frac{e_{k21}}{c_{2121}}E_k^{(1)},$$

$$\bar{S}_{22}^{(4)} = -\frac{1}{\pi}\frac{c_{22kl}}{c_{2222}}S_{kl}^{d} - \frac{16}{15\pi}\frac{c_{22kl}}{c_{2222}}S_{kl}^{(1)} + \frac{16}{15\pi}\frac{e_{k22}}{c_{2222}}E_k^{(1)},$$

$$\bar{S}_{23}^{(4)} = -\frac{1}{2\pi}\frac{c_{23kl}}{c_{2323}}S_{kl}^{d} - \frac{8}{15\pi}\frac{c_{23kl}}{c_{2323}}S_{kl}^{(1)} + \frac{8}{15\pi}\frac{e_{k23}}{c_{2323}}E_k^{(1)}.$$

In (6), we have neglected some smaller elastic constants which are zero in isotropic materials. By substituting equations (6) into $(2)_{6\text{-}10}$ and neglecting other third and fourth order strains, the modified constitutive relations are obtained:

$$T_{ij}^{(0)} = 2c_{ijkl}^{(0)}S_{kl}^{(0)} + \frac{4}{\pi}c_{ijkl}\bar{S}_{kl}^{(1)} - c_{ijkl}^{(3)}S_{kl}^{(2)} - 2e_{kij}^{(0)}E_k^{(0)} - \frac{4}{\pi}e_{kij}\bar{E}_k^{(1)} + e_{kij}^{(3)}E_k^{(2)},$$

$$T_{ij}^{(1)} = \frac{8}{\pi^2}c_{ijkl}^{(d)}S_{kl}^{d} + c_{ijkl}^{(1)}S_{kl}^{(1)} + \frac{8}{3\pi}c_{ijkl}\bar{S}_{kl}^{(2)} - e_{kij}^{(1)}E_k^{(1)} - \frac{8}{3\pi}e_{kij}\bar{E}_k^{(2)},$$

$$T_{ij}^{(2)} = -c_{ijkl}^{(3)}S_{kl}^{(0)} - \frac{4}{3\pi}c_{ijkl}\bar{S}_{kl}^{(1)} + c_{ijkl}^{(2)}S_{kl}^{(2)} + e_{kij}^{(3)}E_k^{(0)} + \frac{4}{3\pi}e_{kij}\bar{E}_k^{(1)} - e_{kij}^{(2)}E_k^{(2)},$$

$$\bar{T}_{ij}^{(1)} = \frac{4}{\pi}c_{ijkl}S_{kl}^{(0)} + c_{ijkl}\bar{S}_{kl}^{(1)} - \frac{4}{3\pi}c_{ijkl}S_{kl}^{(2)} - \frac{4}{\pi}e_{kij}E_k^{(0)} - e_{kij}\bar{E}_k^{(1)} + \frac{4}{3\pi}e_{kij}E_k^{(2)},$$

$$\bar{T}_{ij}^{(2)} = \frac{2}{\pi}c_{ijkl}S_{kl}^{d} + \frac{8}{3\pi}c_{ijkl}S_{kl}^{(1)} + 2c_{ijkl}\bar{S}_{kl}^{(2)} - \frac{8}{3\pi}e_{kij}E_k^{(1)} - e_{kij}\bar{E}_k^{(2)}, \tag{7}$$

$$D_i^{(0)} = 2e_{ikl}^{(0)}S_{kl}^{(0)} + \frac{4}{\pi}e_{ikl}\bar{S}_{kl}^{(1)} - e_{ikl}^{(3)}S_{kl}^{(2)} + 2\varepsilon_{ik}^{(0)}E_k^{(0)} + \frac{4}{\pi}\varepsilon_{ik}\bar{E}_k^{(1)} + \varepsilon_{ik}^{(3)}E_k^{(2)},$$

$$D_i^{(1)} = \frac{8}{\pi^2}e_{ikl}^{(d)}S_{kl}^{d} + e_{ikl}^{(1)}S_{kl}^{(1)} + \frac{8}{3\pi}e_{ikl}\bar{S}_{kl}^{(2)} + \varepsilon_{ik}^{(1)}E_k^{(1)} + \frac{8}{3\pi}\varepsilon_{ik}\bar{E}_k^{(2)},$$

$$D_i^{(2)} = -e_{ikl}^{(3)}S_{kl}^{(0)} - \frac{4}{3\pi}e_{ikl}\bar{S}_{kl}^{(1)} + e_{ikl}^{(2)}S_{kl}^{(2)} + \varepsilon_{ik}^{(3)}E_k^{(0)} - \frac{4}{3\pi}\varepsilon_{ik}\bar{E}_k^{(1)} + \varepsilon_{ik}^{(2)}E_k^{(2)},$$

$$\bar{D}_i^{(1)} = \frac{4}{\pi}e_{ikl}S_{kl}^{(0)} + e_{ikl}\bar{S}_{kl}^{(1)} - \frac{4}{3\pi}e_{ikl}S_{kl}^{(2)} + \frac{4}{\pi}\varepsilon_{ik}E_k^{(0)} + \varepsilon_{ik}\bar{E}_k^{(1)} - \frac{4}{3\pi}\varepsilon_{ik}E_k^{(2)},$$

$$\bar{D}_i^{(2)} = \frac{2}{\pi}e_{ikl}S_{kl}^{d} + \frac{8}{3\pi}e_{ikl}S_{kl}^{(1)} + 2e_{ikl}\bar{S}_{kl}^{(2)} + \frac{8}{3\pi}\varepsilon_{ik}E_k^{(1)} + \varepsilon_{ik}\bar{E}_k^{(2)},$$

where

$$c_{ijkl}^{(d)} = c_{ijkl} - \frac{\pi^2}{24} c_{ijkl}^{(3)}, \qquad e_{kij}^{(d)} = -\frac{\pi^2}{24} e_{kij}^{(3)},$$

$$c_{ijkl}^{(0)} = c_{ijkl} - \frac{5}{18} c_{ijkl}^{(3)}, \qquad e_{kij}^{(0)} = e_{kij} - \frac{5}{18} e_{kij}^{(3)},$$

$$c_{ijkl}^{(1)} = c_{ijkl} - \frac{16}{45} c_{ijkl}^{(3)}, \qquad e_{kij}^{(1)} = e_{kij} - \frac{16}{45} e_{kij}^{(3)},$$

$$c_{ijkl}^{(2)} = c_{ijkl} - \frac{9}{5} c_{ijkl}^{(3)}, \qquad e_{kij}^{(2)} = e_{kij} - \frac{9}{5} e_{kij}^{(3)},$$

$$c_{ijkl}^{(3)} = \frac{16}{5\pi^2}\left(\frac{c_{ij21}c_{21kl}}{c_{2121}} + \frac{c_{ij22}c_{22kl}}{c_{2222}} + \frac{c_{ij23}c_{23kl}}{c_{2323}}\right),$$

$$e_{kij}^{(3)} = \frac{16}{5\pi^2}\left(\frac{c_{ij21}e_{k21}}{c_{2121}} + \frac{c_{ij22}e_{k22}}{c_{2222}} + \frac{c_{ij23}e_{k23}}{c_{2323}}\right),$$

$$\varepsilon_{ik}^{(0)} = \varepsilon_{ik} + \frac{5}{18}\varepsilon_{ik}^{(3)}, \qquad \varepsilon_{ik}^{(1)} = \varepsilon_{ik} + \frac{16}{45}\varepsilon_{ik}^{(3)},$$

$$\varepsilon_{ik}^{(2)} = \varepsilon_{ik} + \frac{9}{5}\varepsilon_{ik}^{(3)}, \tag{8}$$

$$\varepsilon_{ik}^{(3)} = \frac{16}{5\pi^2}\left(\frac{e_{i21}e_{k21}}{c_{2121}} + \frac{e_{i22}e_{k22}}{c_{2222}} + \frac{e_{i23}e_{k23}}{c_{2323}}\right).$$

The stress equations of motion and charge equations of electrostatics are

$$T_{ij,i}^{(0)} + \frac{1}{b}F_j^{(0)} = 2\rho\ddot{u}_j^{(0)},$$

$$T_{ij,i}^{(1)} - \frac{\pi}{2b}\bar{T}_{2j}^{(1)} + \frac{1}{b}F_j^{(1)} = \rho\ddot{u}_j^{(1)} - \frac{8}{\pi^2}\rho b\ddot{u}_{2,j}^{(0)},$$

$$T_{ij,i}^{(2)} - \frac{\pi}{b}\bar{T}_{2j}^{(2)} + \frac{1}{b}F_j^{(2)} = \rho\ddot{u}_j^{(2)},$$

$$D_{i,i}^{(0)} + \frac{1}{b}D^{(0)} = 0, \tag{9}$$

$$D_{i,i}^{(1)} - \frac{\pi}{2b}\bar{D}_2^{(1)} + \frac{1}{b}D^{(1)} = 0,$$

$$D_{i,i}^{(2)} - \frac{\pi}{b}\bar{D}_2^{(2)} + \frac{1}{b}D^{(2)} = 0.$$

The strain-displacement relations are

$$S_{ij}^{d} = -bu_{2,ij}^{(0)},$$

$$S_{ij}^{(0)} = \frac{1}{2}\left(u_{i,j}^{(0)} + u_{j,i}^{(0)} - \delta_{2i}u_{2,j}^{(0)} - \delta_{2j}u_{2,i}^{(0)}\right),$$

$$S_{ij}^{(1)} = \frac{1}{2}\left(u_{i,j}^{(1)} + u_{j,i}^{(1)}\right),$$

$$S_{ij}^{(2)} = \frac{1}{2}\left(u_{i,j}^{(2)} + u_{j,i}^{(2)}\right), \tag{10}$$

$$\bar{S}_{ij}^{(1)} = \frac{\pi}{4b}\left(\delta_{2j}u_i^{(1)} + \delta_{2i}u_j^{(1)}\right),$$

$$\bar{S}_{ij}^{(2)} = \frac{\pi}{2b}\left(\delta_{2j}u_i^{(2)} + \delta_{2i}u_j^{(2)}\right).$$

The field-potential relations are

$$E_i^{(0)} = -\varphi_{,i}^{(0)}, \qquad E_i^{(1)} = -\varphi_{,i}^{(1)},$$

$$E_i^{(2)} = -\varphi_{,i}^{(2)}, \tag{11}$$

$$\bar{E}_2^{(1)} = -\frac{\pi}{2b}\varphi^{(1)}, \qquad \bar{E}_2^{(2)} = -\frac{\pi}{b}\varphi^{(2)}.$$

The second order equations for a sine series expansion of the electric potential has also been obtained and they differ slightly from the given equations.

5. AVERAGED EQUATIONS FOR A STRIP

The nth order components of displacement, stress and charge, and the nth order electric potential are averaged by integrating across the x_3 direction and dividing by 2c.

$$v_j^{(n)}(x_1,t) = \frac{1}{2c}\int_{-c}^{c} u_j^{(n)}(x_1,x_3,t)dx_3,$$

$$\sigma_{ij}^{(n)}(x_1,t) = \frac{1}{2c}\int_{-c}^{c} T_{ij}^{(n)}(x_1,x_3,t)dx_3, \tag{12}$$

$$\bar{\sigma}_{ij}^{(n)}(x_1,t) = \frac{1}{2c}\int_{-c}^{c} \bar{T}_{ij}^{(n)}(x_1,x_3,t)dx_3,$$

$$d_k^{(n)}(x_1,t) = \frac{1}{2c}\int_{-c}^{c} D_k^{(n)}(x_1,x_3,t)dx_3,$$

$$\bar{d}_k^{(n)}(x_1,t) = \frac{1}{2c}\int_{-c}^{c} \bar{D}_k^{(n)}(x_1,x_3,t)dx_3,$$

$$\gamma^{(n)}(x_1,t) = \frac{1}{2c}\int_{-c}^{c} \varphi^{(n)}(x_1,x_3,t)dx_3.$$

The stress equations of motion and charge equations of electrostatics (9) are also integrated through the x_3 dimension and divided by 2c:

$$\sigma_{1j,1}^{(0)} + H_j^{(0)} = 2\rho\ddot{v}_j^{(0)},$$

$$\sigma_{1j,1}^{(1)} - \frac{\pi}{2b}\bar{\sigma}_{2j}^{(1)} + H_j^{(1)} = \rho\ddot{v}_j^{(1)},$$

$$\sigma_{1j,1}^{(2)} - \frac{\pi}{b}\bar{\sigma}_{2j}^{(2)} + H_j^{(2)} = \rho\ddot{v}_j^{(2)}, \tag{13}$$

$$d_{1,1}^{(0)} + d^{(0)} = 0,$$

$$d_{1,1}^{(1)} - \frac{\pi}{2b}\bar{d}_2^{(1)} + d^{(1)} = 0,$$

$$d_{1,1}^{(2)} - \frac{\pi}{b}\bar{d}_2^{(2)} + d^{(2)} = 0,$$

where

$$H_j^{(n)} = \frac{1}{2c}\left(T_{3j}^{(n)}[c] - T_{3j}^{(n)}[-c]\right),$$

$$d^{(n)} = \frac{1}{2c}\left(D_3^{(n)}[c] - D_3^{(n)}[-c]\right).$$

Considering the case where the x_3 dimension of the plate, 2c, is much smaller than both of the other dimensions, $c \ll \mathrm{Min}(a,b)$, and for traction free x_3 faces, we assume

$$T_{33} = 0. \tag{14}$$

For the remainder of the paper, the focus is on the in-plane symmetric motions of a thin plate which has the material properties of a hexagonal crystal, class 6mm. Using equation (14) in $(9)_2$, we have

$$T_{33}^{(0)} = 0, \qquad T_{33}^{(2)} = 0. \tag{15}$$

By approximately satisfying (15), and for shorted electrodes on the x_3 faces,

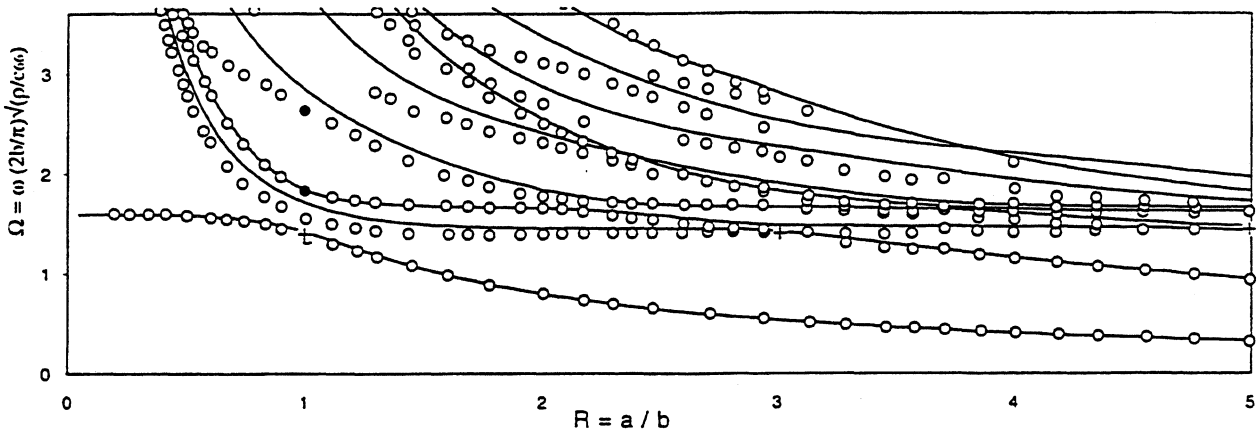


Figure 1: Frequency Spectrum Calculated using Present Theory and Measured by Onoe and Pao for a Thin Plate.

$$u_{3.3}^{(0)} = -\frac{c_{31}^{(0)}}{c_{33}^{(0)}} u_{1.1}^{(0)} - \frac{1}{b}\frac{c_{32}}{c_{33}^{(0)}} u_2^{(1)} + \frac{1}{2}\frac{c_{31}^{(3)}}{c_{33}^{(0)}} u_{1.1}^{(2)},$$

$$u_{3.3}^{(2)} = \frac{c_{31}^{(3)}}{c_{33}^{(2)}} u_{1.1}^{(0)} + \frac{2}{3b}\frac{c_{32}}{c_{33}^{(2)}} u_2^{(1)} - \frac{c_{31}^{(2)}}{c_{33}^{(2)}} u_{1.1}^{(2)}. \qquad (16)$$

Substituting (16) into (7) and the result into (12) gives the modified constitutive relations which will be used in the stress equations of motion:

$$\sigma_1^{(0)} = 2\tilde{c}_{11}^{0} v_{1.1}^{(0)} + \frac{2}{b}\tilde{c}_{12}^{4} v_2^{(1)} - \tilde{c}_{11}^{3} v_{1.1}^{(2)},$$

$$\sigma_6^{(1)} = c_{66}^{(1)} v_{2.1}^{(1)} + \frac{8}{3b} c_{66} v_1^{(2)},$$

$$\bar{\sigma}_2^{(1)} = \frac{4}{\pi}\tilde{c}_{12}^{4} v_{1.1}^{(0)} + \frac{\pi}{2b}\tilde{c}_{22}^{1} v_2^{(1)} - \frac{4}{3\pi}\tilde{c}_{21}^{5} v_{1.1}^{(2)}, \qquad (17)$$

$$\sigma_1^{(2)} = -\tilde{c}_{11}^{3} v_{1.1}^{(0)} - \frac{2}{3b}\tilde{c}_{21}^{5} v_2^{(1)} + \tilde{c}_{11}^{2} v_{1.1}^{(2)},$$

$$\bar{\sigma}_6^{(2)} = \frac{8}{3\pi} c_{66} v_{2.1}^{(1)} + \frac{\pi}{b} c_{66} v_1^{(2)},$$

where the modified elastic constants are

$$\tilde{c}_{11}^{0} = c_{11}^{(0)} - \frac{c_{13}^{(0)} c_{31}^{(0)}}{c_{33}^{(0)}} - \frac{1}{2}\frac{c_{13}^{(3)} c_{31}^{(3)}}{c_{33}^{(2)}},$$

$$\tilde{c}_{22}^{1} = c_{22} - \frac{8}{\pi^2}\frac{c_{23} c_{32}}{c_{33}^{(0)}} - \frac{16}{9\pi^2}\frac{c_{23} c_{32}}{c_{33}^{(2)}},$$

$$\tilde{c}_{11}^{2} = c_{11}^{(2)} - \frac{1}{2}\frac{c_{13}^{(3)} c_{31}^{(3)}}{c_{33}^{(0)}} - \frac{c_{13}^{(2)} c_{31}^{(2)}}{c_{33}^{(2)}}, \qquad (18)$$

$$\tilde{c}_{11}^{3} = c_{11}^{(3)} - \frac{c_{13}^{(0)} c_{31}^{(3)}}{c_{33}^{(0)}} - \frac{c_{13}^{(3)} c_{31}^{(2)}}{c_{33}^{(2)}},$$

$$\tilde{c}_{12}^{4} = c_{12} - \frac{c_{13}^{(0)} c_{32}}{c_{33}^{(0)}} - \frac{c_{32} c_{13}^{(3)}}{c_{33}^{(2)}},$$

$$\tilde{c}_{21}^{5} = c_{21} - \frac{3}{2}\frac{c_{23} c_{31}^{(3)}}{c_{33}^{(0)}} - \frac{c_{23} c_{31}^{(2)}}{c_{33}^{(2)}}.$$

By substituting the modified constitutive relations (17) into the stress equations of motion (13) we obtain the equations of motion for symmetric in-plane average displacements:

$$\tilde{c}_{11}^{0} v_{1.11}^{(0)} + \frac{1}{b}\tilde{c}_{12}^{4} v_{2.1}^{(1)} - \frac{1}{2}\tilde{c}_{11}^{3} v_{1.11}^{(2)} = \rho \ddot{v}_1^{(0)},$$

$$-\frac{2}{b}\tilde{c}_{12}^{4} v_{1.1}^{(0)} + c_{66}^{(1)} v_{2.11}^{(1)} - \frac{\pi^2}{4b^2}\tilde{c}_{22}^{1} v_2^{(1)} + \frac{2}{3b}\left[4c_{66} + \tilde{c}_{21}^{5}\right] v_{1.1}^{(2)} = \rho \ddot{v}_2^{(1)}, \qquad (19)$$

$$-\tilde{c}_{11}^{3} v_{1.11}^{(0)} - \frac{2}{3b}\left[4c_{66} + \tilde{c}_{21}^{5}\right] v_{2.1}^{(1)} + \tilde{c}_{11}^{2} v_{1.11}^{(2)} - \frac{\pi^2}{b^2} c_{66} v_1^{(2)} = \rho \ddot{v}_1^{(2)}.$$

The dispersion curves are calculated from the equations of motion (19) in the usual manner. The boundary conditions which are satisfied at $x_1 = \pm a$ are

$$\sigma_1^{(0)} = 0, \quad \sigma_6^{(1)} = 0, \quad \sigma_1^{(2)} = 0, \qquad (20)$$

The dimensionless frequencies which allow the satisfaction of (20) are shown in figure 1 in solid lines and they compare closely with the experiment by Onoe and Pao[4] on a barium titanate plate (data points extracted from Medick and Pao[5]).

ACKNOWLEDGMENT

This work was supported by grant DAAH 04-95-1-0614 from the U.S. Army Research Office.

REFERENCES

[1] P.C.Y. Lee, J.D. Yu and W.S. Lin, "A New Two-Dimensional Theory for Vibrations of Piezoelectric Crystal Plates with Electroded Faces," J. Appl. Phys., vol. 83 (3) p. 1213. 1998

[2] I. Koga and H. Fukuyo, "Vibration of Thin Piezoelectric Quartz Plates," J. Inst. Elec. Comm. Engrs. of Japan, vol. 36, p. 59. 1953

[3] M. Nakazawa, K. Horiuchi, and H. Ito, "Frequency Mode Charts of Rectangular AT-cut Quartz Plates," in Proceedings of the IEEE Ultrasonics Symposium, New York, p. 547. 1990

[4] M. Onoe and Y.H. Pao, "Edge Mode of Thin Rectangular Plate of Barium Titanate," J. Acoust. Soc. Am., vol. 33, p. 1628. 1961

[5] M. Medick and Y.H. Pao, "Extensional Vibrations of Thin Rectangular Plates," J. Acoust. Soc. Am., vol. 37, p. 59. 1965

1999 Joint Meeting EFTF - IEEE IFCS

SYNTHETIC MODELING OF QUARTZ CRYSTAL OSCILLATOR

R. Brendel, N. Ratier, L. Couteleau, G. Marianneau, F. Lardet-Vieudrin, P. Guillemot*

Laboratoire de Physique et Métrologie des Oscillateurs du CNRS
associé à l'Université de Franche-Comté-Besançon - 32, avenue de l'Observatoire - 25044 Besançon Cedex – France
Brendel@lpmo.univ-fcomte.fr
*Centre National d'Etudes Spatiales - 18, avenue Edouard Belin - 31055 Toulouse Cedex - France

ABSTRACT

This paper presents a new modeling technique to describe the nonlinear behavior of complicated oscillator circuits. The simulation program being developed first removes the resonator from the oscillation loop and call on SPICE to calculate the large signal y-parameters of the amplifier circuit considered as a nonlinear two-port circuits. The oscillation condition, obtained by reinsering the resonator across the two-port circuit, is expressed under the form of a complex polynomial in the harmonic variable $j\omega$, the coefficients of this polynomial being nonlinear functions of the signal amplitude. Solving the real and imaginary parts of this characteristic polynomial by using nonlinear analysis algorithm, it is possible to accurately calculate both amplitude and frequency of the oscillation.

1. INTRODUCTION

For several years a dedicated software is being developed so as to accurately predict the behavior of quartz crystal oscillators [1, 2, 3]. The program is based on the nonlinear Barkausen criterion leading to a complex characteristic polynomial whose the solution is the amplitude and frequency of the oscillation. In this approach the nonlinear behavior of the active device usually a BJTs is described by a large signal y-parameter two-port representation [4]. The other passive components of the circuit are then progressively "encapsulated" in the nonlinear two-port and the corresponding transformations are stored and coded [5, 6]. At the end of the process, the resonator is inserted in the circuit and the oscillation condition is derived so as to obtain the aforementioned characteristic polynomial.

This procedure works quite well with circuits including only one active device but is not well suited for complex circuits including several nonlinear devices such as amplifier loop with two BJT or automatic gain control, electronic tuning circuit (varicap) and so on.

Because of the increasing computer speed, another nonlinear approach can be used in these more complicated configurations. Instead of describing only the active device as a nonlinear component, the amplifier loop as a whole is considered as a large signal two-port. To this end, a SPICE-like netlist describes the oscillating circuit, a parser identifies the position of the quartz resonator and removes it from the amplifier circuit, the SPICE subcircuit of the loop amplifier is then automatically generated so as to analyze its nonlinear behavior.

Reintroducing the resonator across the two-port circuit, the oscillation condition is expressed under the form of a complex polynomial in the harmonic variable $j\omega$, the coefficients of this polynomial being nonlinear functions in the signal amplitude.

The amplitude and oscillation frequency are then obtained by a method similar to the one used in the previous program.

2. LARGE SIGNAL TWO-PORT AMPLIFIER CIRCUIT DESCRIPTION

The oscillator circuit is initially described under the form of a SPICE netlist with additional informations such as the frequency-temperature coefficients and the amplitude-frequency coefficients of the quartz resonator. Then, the result is parsed by a topological analysis program (TAP) that identifies the resonator and removes it so as to isolate the amplifier circuit. The TAP then rebuilts two SPICE netlists. The first one is obtained by shorting the output of the amplifier with a large capacitor value and by putting a sine function generator at the input, the input and transfer nonlinear admittances are then obtained as function of the input amplitude by performing a set of SPICE transient analyses with larger and larger input signal amplitudes.

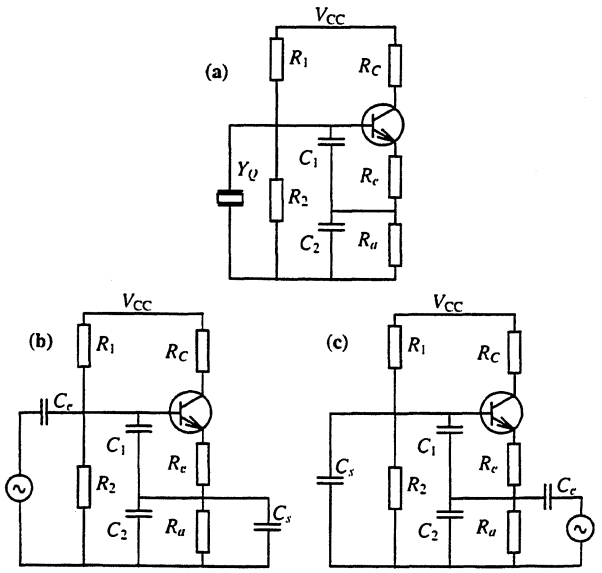


Fig. 1. Initial oscillator circuit and rebuilded amplifier test circuits

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

Likewise, the second circuit is obtained by shorting the input and putting the sine generator at the output, the output and reverse nonlinear admittances are then obtained as functions of the output amplitude. Fig. 1 shows an example of a Colpitts oscillator circuit and the two-amplifier test circuits rebuilded by the TAP.

Performing the suitable ratios of currents and voltages obtained from the SPICE transient simulations performed at the assumed oscillation frequency for every excitation level, the amplifier circuit can be reduced to the large signal two-port circuit shown in Fig. 2. In this representation, the y-parameters are nonlinear functions of either the input signal amplitude ($y_{ia}(u)$ and $y_{fa}(u)$) or the output signal amplitude ($y_{ra}(v)$ and $y_{oa}(v)$).

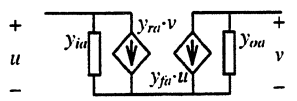
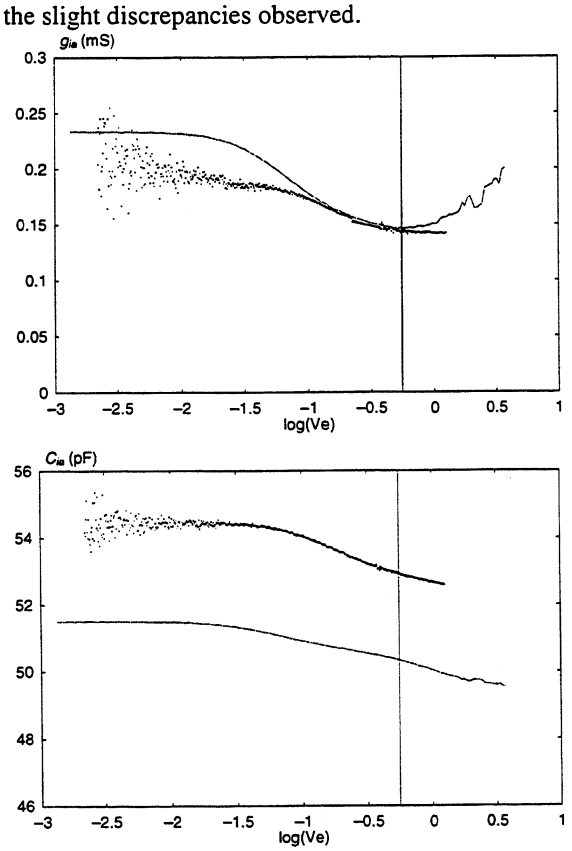


Fig. 2.Large signal two-port equivalent circuit of the amplifier

These complex y-parameters can be split into their real and imaginary parts under the form :

$$y_{ia}=g_{ia}+sC_{ia} \qquad y_{ra}=g_{ra}+sC_{ra}$$
$$y_{fa}=g_{fa}+sC_{fa} \qquad y_{oa}=g_{oa}+sC_{oa} \qquad (1)$$

Fig. 3 shows a selected number of y-parameters obtained for the amplifier circuit represented in figure 1 as well as experimental data obtained by measurements performed on the actual circuit. Parasitic capacitances or inaccurate transistor model parameters can explain the slight discrepancies observed.



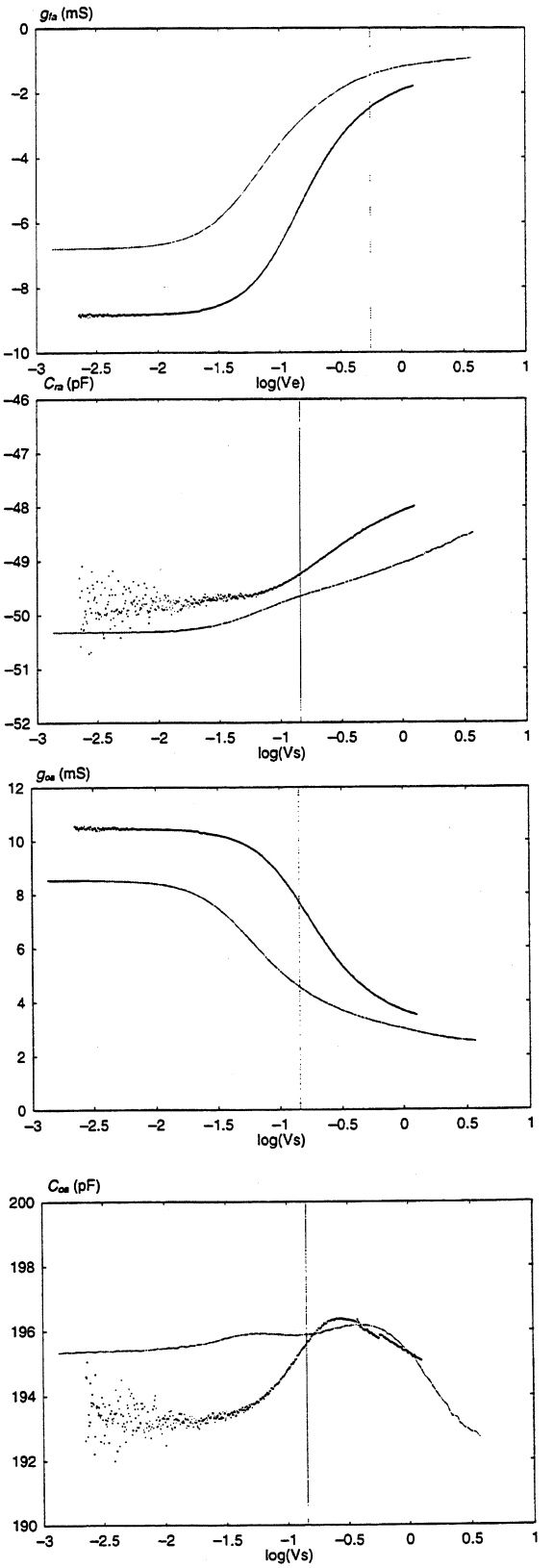


Fig. 3. Large signal y-parameters functions of the signal amplitude

3. OSCILLATION CONDITION

Whatever the quartz crystal oscillator circuit and whatever the complexity of the amplifier circuit may be, when the resonator is reintroduced across the amplifier, the oscillator circuit can always be reduced to one of the two forms shown in fig. 4.

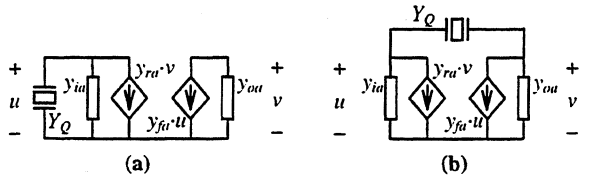


Fig. 4. Reduced forms of the oscillator circuit

In the first case the resonator is connected between input and ground while in the second case it is connected between the amplifier input and output.

In both cases (a) and (b) in fig. 4 the oscillation condition can be expressed as :

$$\Delta_y + y_a Y_Q = 0 \tag{2}$$

Where

$$\begin{aligned} y_a &= y_{oa} \text{ in the case of Fig. 4.a} \\ y_a &= y_{ia} + y_{oa} + y_{ra} + y_{fa} \text{ in the case of Fig. 4.b} \end{aligned} \tag{3}$$

In both cases :

$$\Delta_y = y_{ia}\, y_{oa} - y_{ra}\, y_{fa} \tag{4}$$

Y_Q designates the admittance of the series branch (C_q, L_q, R_q) of the resonator :

$$Y_Q = \frac{1}{L_q} \frac{s}{s^2 + \frac{R_q}{L_q} s + \omega_q^2}, \qquad \omega_q^2 = \frac{1}{L_q C_q} \tag{5}$$

(In this case, the parallel capacitance of the resonator should be part of the amplifier circuit).

It can be shown that developing the oscillation condition Eq. (2) using relations (1), (3), (4) and (5) leads to a characteristic polynomial of degree 4 in the Laplace's variable s :

$$\sum_{n=0}^{4} \alpha_n s^n = 0 \tag{6}$$

The polynomial coefficients α_n are functions of the resonator parameters and of the amplifier y-parameters, the latter depend on either the input or the output signal amplitude. These two variables are connected by expressing the amplifier gain that takes the form :

$$\begin{aligned} \frac{v}{u} &= -\frac{y_{fa}}{y_{oa}} \text{ in the case of Fig. 4.a} \\ \frac{v}{u} &= -\frac{y_{ia} + y_{fa}}{y_{oa} + y_{ra}} \text{ in the case of Fig. 4.b} \end{aligned} \tag{7}$$

The isochronism defect (amplitude-frequency effect) is taken into account by using the equation :

$$\omega_q = \omega_{q0} (1 + a_f P_q) \tag{8}$$

Where ω_{q0} is the low level resonant frequency of the crystal a_f is an anisochronism coefficient depending on the resonator used and P_q is the actual resonator drive level expressed as :

$$\begin{aligned} P_q &= \frac{1}{2} \Re(Y_Q)\, u^2 \quad \text{in the case of Fig. 4.a} \\ P_q &= \frac{1}{2} \Re(Y_Q)\, (u - v)^2 \quad \text{in the case of Fig. 4.b} \end{aligned} \tag{9}$$

4. STEADY STATE OSCILLATIONS ALGORITHM

The steady state oscillations amplitude and frequency can be obtained by solving the characteristic polynomial Eq. (6) in the frequency domain. Substituting the harmonic variable $j\omega$ for the Laplace's variable s splits the initial equation into two equations, one for the real part, the other one for the imaginary part

$$\alpha_4 \omega^4 - \alpha_2 \omega^2 + \alpha_0 = 0 \tag{10}$$

$$\alpha_3 \omega^2 - \alpha_1 = 0 \tag{11}$$

Taking into account Eqs. (7), (8) and (9), Eqs. (10) and (11) enables to obtain the steady state frequency ω_0 and the input and output amplitudes u_0 and v_0.

So as to settle a compromise between computing speed and accuracy, the steady state solution is calculated through a two pass algorithm. The DC supply voltage being known, it is possible to sweep the input and output generator voltages (Fig. 1 b and c) from small value up to DC supply voltage value amplitude to obtain the amplifier two-port y-parameter curves as shown in Fig. 3. Thus, for each value of the input voltage amplitude u, it is possible to calculate the corresponding output voltage v satisfying Eq. (7) as well as the value of the y-parameter and the resonator drive level at the actual calculation step. The oscillation frequency is calculated by using Eqs. (8) and (11). At this point Eq. (10) must also be satisfied, if not, the input voltage is increased and the process is repeated until the sign of Eq. (10) changes. If the sign of the polynomial Eq. (10) never changes over the input amplitude range the circuit cannot oscillate, in the opposite case, the solution is located in the vicinity of the last calculated point. The second pass consists in performing a new set of transient simulations so as to obtain a better accuracy of the y-parameters in the vicinity of the solution. Furthermore, because the voltage amplitude range involved is quite reduced, the y-parameters can be expanded as a linear function in the input or output voltage amplitudes as shown in Fig. 5. At this stage, it becomes possible to linearize Eqs. (10) and (11) to obtain an accurate value of oscillation frequency ω_0 and input and output voltages as well.

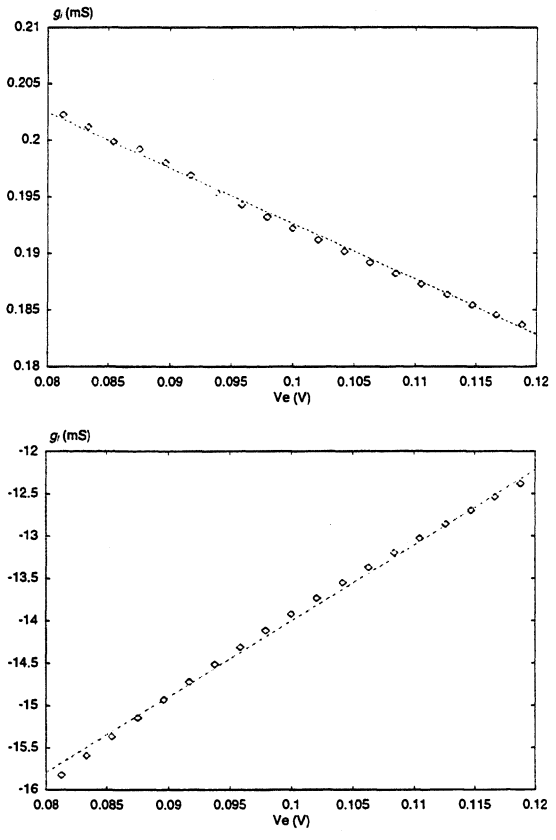


Fig. 5. Linear expansion of the y-parameter in the vicinity of the steady state amplitude

5. SIMULATION PROGRAM

A preliminary version of a simulation program has been developed upon the principles previously described. The input file looks like a SPICE circuit file except some statements specific to the oscillator structure. Starting from this input file, the program performs the various tasks and creates the auxiliary files necessary to the oscillation condition calculation. Eventually, results are edited in an output file and on the computer screen.

6. CONCLUSION AND FUTURE PROSPECTS

Higher and higher computer speed as well as increasing software availability has opened a new way for the development of a simulation program dedicated to quartz crystal oscillators. This program, grounded on a synthetic nonlinear description of the loop amplifier performed by a set of SPICE transient analyses, allows the user to accurately predict the frequency and amplitude of the oscillations whatever the topology and/or the complexity of the circuit may be.

The feasibility of this new approach is now well established. Nevertheless, many features of the program have to be developed or improved :

- introduction of useful functionalities such as effect of a circuit parameter change on the oscillation characteristics, sensitivity calculation, tolerance and worstcase analysis, ...
- oscillation frequency and amplitude transients calculation using slowly varying functions method developed in the previous program can also be [7, 8].implemented in this version of the program
- Likewise, perturbation method previously used to analyze the noise mechanisms and to calculate AM and PM noise spectra in oscillators [9, 10, 11] can also be used here to obtain similar results.

Moreover, some improvements in the user interface still have to be carried out so as to make this program an accurate, efficient and easy-to-use tool for the quartz oscillator design.

ACKNOWLEDGEMENTS :

This work has been supported by CNES and DGA under contracts :
#832/CNES/88/5374/00, #962/CNES/91/1476/00, #962/CNES/94/1230/00, #714/CNES/97/7000/00.

REFERENCES

[1] R. Brendel, G. Marianneau, E. Robert, "Nonlinear modeling techniques for quartz crystal oscillators", Proc. of the 5th European Frequency and Time Forum, Besançon, France., March 1991

[2] R. Brendel, F. Djian, E. Robert, "High precision nonlinear computer modeling technique for quartz crystal oscillators", Proc. of the 45th Annual Symposium on Frequency control, Los Angeles, May 1991

[3] R. Brendel, G. Marianneau, T. Blin, M. Brunet, "Computer aided design of quartz crystal oscillators", IEEE Trans. On Ultrasonics, Ferroelectrics and Frequency control, vol. 42, n°4, pp. 700-707, July 1995

[4] M.E. Frerking, Crystal oscillator design and temperature control, new York, Van Nostrand Reinhald Co, 1978

[5] N. Ratier, R. Brendel, P. Guillemot, "Quartz oscillators : deriving oscillation condition by symbolic calculus", Proc. of the 10th European Frequency and Time Forum, Brighton, UK, March 1996

[6] N. Ratier, L. Couteleau, . R. Brendel, P. Guillemot, "Automatic formal derivation of the oscillation condition", Proc. of the 51st International Symposium on frequency Control, Orlando, USA, May 1997

[7] R. Brendel, N. Ratier, L. Couteleau, G. Marianneau, P. Guillemot, "Transient simulation in quartz crystal oscillators", Proc. of the 11th European Frequency and Time Forum, Neuchâtel, Switzerland, March 1997

[8] R. Brendel, N. Ratier, L. Couteleau, G. Marianneau, P. Guillemot "Slowly varying function method applied to quartz crystal oscillator transient calculation", IEEE Trans. on Ultrasonics, Ferroelectrics and Frequency Control, vol. 45, n°2., pp. 520-527, March 1998

[9] R. Brendel, N. Ratier, L. Couteleau, G. Marianneau, P. Guillemot, "Nonlinear analysis of noise in quartz crystal oscillators", Proc. of the 12th EFTF, Varsaw, Poland, March 1998.

[10] N. Ratier, L. Couteleau, R. Brendel, P. Guillemot, "AM and PM noise analysis in quartz crystal oscillators : symbolic calculus approach", Proc. of the 52nd, ISFC, Los Angeles, USA, May 1998.

[11] R. Brendel, N. Ratier, L. Couteleau, G. Marianneau, P. Guillemot, "Analysis of noise in quartz oscillators by using slowly varying function method", IEEE Trans. on UFFC, Vol. 46, n°22, March 1999.

INVESTIGATION OF EXCITATION CIRCUIT INFLUENCE ON CRYSTAL OSCILLATOR F-T CURVE LOCAL DISTURBANCES

Alexander Lepetaev, Anatoly Kosykh
Omsk State Engineering University. 644050, Mira Avenue, 11, Omsk, Russia

ABSTRACT

This article is devoted to investigation and simulation of oscillator circuit influence on frequency jumps anf F-T curve irregularity. Till now the influence of parasitic resonator modes on oscillator frequency was considered only empirically and was not linked with properties of the resonator excitation circuit. However it is possible to show, that magnitudes of local perturbations and the fact of appearance of oscillator F-T curve hysteresis are depend on a type oscillator circuit and on values of its elements.

In the represented article the mathematical model of the given phenomenon is offered. The analysis of conditions that lead to stimulation of frequency jumps or shifts by the oscillator circuit is provided. The technique of computer simulation of F-T curve irregularity is offered.

The results of F-T curve irregularity simulation for various combinations of resonator and oscillator, recommendation influencing on choice of the oscillation circuit and outcomes of comparisons with experimental data are presented.

It is known, that near to the basic resonance frequency there can be parasitic resonances appropriate to other modes of oscillations [1]. For example, in the resonator working at shift oscillations of a C- mode, the harmonics of face shear oscillations can be observed. The activity of these parasitic modes is insignificant, but, nevertheless, in some cases they can result in occurrence of local peaks of frequency vs. temperature characteristic of crystal resonator. In temperature compensated oscillators it leads to sharp increase of a thermocompensation mistake of spasmodic or hysteresis type. If the quartz resonator is used as the temperature sensor, the presence of such parasitic modes can result in occurrence of local errors of temperature measurement.

Till now the influence of parasitic modes of resonator on the oscillator characteristics was considered only empirically and was not connected with properties of the quartz resonator excitation circuit. However it is possible to show, that the fact of presence of a hysteresis in F vs. T curve of the oscillator and the value of its local disturbances depend on a type and face values of elements of the oscillator circuit.

We shall consider the quartz resonator having except the basic mode an additional parasitic one, the frequency of which strongly depends on temperature. The equivalent electrical circuit of such resonator is represented in Fig. 1, where consecutive oscillatory circuit with parameters R, L, C corresponds to the basic mode of oscillations, and oscillatory circuit with parameters Ra, La, Ca coresponds to additional (parasitic) mode.

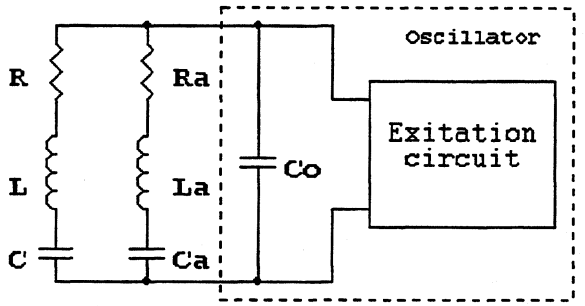


Fig. 1 Equivalent circuit of quartz oscillator

Static capacity of the quartz resonator and all other elements, simulating the properties of resonator electrodes, pins and package (see, for example, [1]), are referred to the oscillator circuit and are considered as its accessory. The resistance Z of high – Q consecutive oscillatory circuit is convenient for expressing through value of generalized detuning

$$Zq = R\cdot(1 + j\cdot\xi),$$

where R - motional resistance of a consecutive oscillatory circuit; $\xi = 2Q\cdot f/F_0$; Q - oscillatory circuit quality; F_0 - resonance frequency of a consecutive oscillatory circuit; $f = F - F_0$ (the difference between the current frequency and frequency of a consecutive resonance, or absolute detuning).

This formula is used in the frequency range, where $f/F_0 << 1$.

Let's enter the following designations:

F_0 - consecutive resonance frequency of the basic mode;

F_a - consecutive resonance frequency of a parasitic mode;

$\Delta = F_0 - F_a$ - the difference of consecutive resonance frequencies of the basic and parasitic modes;

$a = 2Q/F_0$ - factor of proportionality between the absolute detuning of frequency f and generalized detuning for the basic mode of oscillations ($f = F - F_0$);

$b = 2Q_a/F_a \cong 2Q_a/F_0$ - factor of proportionality between absolute detuning of frequency $f_a = F - F_a = f + \Delta$ and generalized detuning ξ_a for a parasitic mode of oscillations;

From the point of view of these designations the total resistance of the system with two oscillatory circuits can be expressed with the following formula:

$$Zq = R\{(1+j\cdot a\cdot f)^{-1} + P[1+j\cdot b\cdot(f+\Delta)]^{-1}\}^{-1} \qquad (1),$$

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

where $P = R/R_a$.

The point of a steady state oscillation can be determined from the equation: $Zq + Zg = 0$, where Zg - oscillator input resistance , which can be measured by means of technique given in [2]. It is more convenient to define the steady state regime point as a point of quartz resonator resistance Zq hodograph and negative oscillator input resistance Zg hodograph crossing. It is possible to consider that oscillator input resistance hodograph is defined only by volume of an input current, and doesn't depend on frequency, as it is possible to neglect the frequency properties of the oscillator in comparison with frequency properties of the resonator). Using the oscillation circuit given in [2] as an example (which is rather widespread),it is seen from the diagrams shown there that at $F_0 = 10^7$, the oscillator reactance isn't practically changed and equals approximately to 300 Ohm in the whole range of input currents, where the ratio $10 < Re(-Zg) < 100$ is carried out. In this case the hodograph of oscillator negative input resistance will look like a direct line parallel to the real axis.

The hodographs of impedanse Zq, calculated with (1) under frequency shift, are presented in Fig. 2 for different values of Δ.

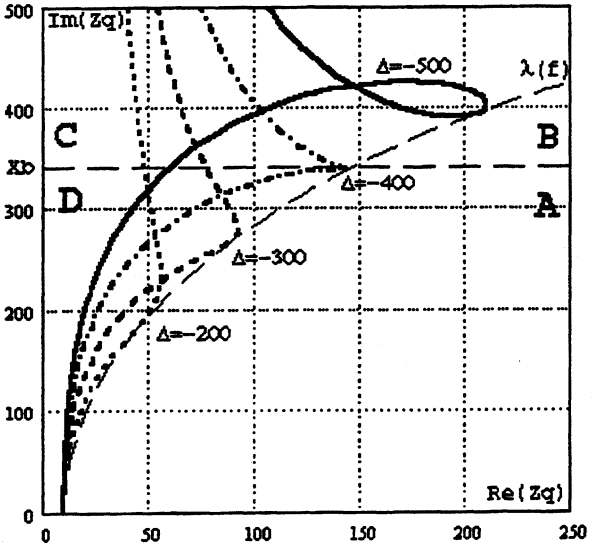


Fig.2. Hodographs of oscillator input impedance Zq

The following meanings of oscillation circuit parameters were used: $F0 = 10^7$; $Q = 0.5 \cdot 10^6$; Qa = 32150; R = 10; Ra = 1000. It is obvious from these diagrams, that the influence of a parasitic mode is displayed as changing of active components of crystal resonator resistance (real part of Zq). Fig. 2 shows that under some values of generalized detuning ξ, there is the "loop" of impedance curve. The presence of this loop results in the appearance of two additional points of crossing the hodograph of Zq and hodograph of oscillator negative input impedance -Zg. These three points of crossing correspond to three possible frequencies of oscillation, however the middle (by frequency) point is unstable, as the fixative ability of quartz is negative in this point. Two other points are steady state. The answer to the question: on which of two possible frequencies the oscillator will oscillate, depends on prehistory of oscillator work. During loop location change because of the parasitic mode frequency change oscillator frequency jumps can be caused by the fact that the current working point ceases to exist, and the oscillator is readjusted on the other working point. The disappearance and occurrence of working points takes place when hodograph of -Zg becomes the tangent line in relation to a loop. As the presence of hodograph loop of the resonator impedance inevitably results in hysteresis of oscillator F-T curve occurrence, it is necessary to define, at what values of generalized detunings ξ it can occur, and how the boundary condition of "loop" occurrence depends on ratios of both Q and motional resistance R of the basic mode and parasitic ones.

It is obvious from Fig. 2 that the loop on hodograph of impedance of two consecutive oscillatory circuits system moves and simultaneously changes in values when a frequencies difference Δ changes. At small Δ values the loop disappears (degenerates). The boundary value of Δ, when there is a loop disappearance, may be determined by zero equating of real and imaginary part of Zq derivatives (1) on basic mode f absolute detuning value. Thus the system of the equations turns out:

$$\begin{cases} \dfrac{d}{df}\left(\dfrac{1}{1+a^2f^2}+\dfrac{P}{1+b^2(f+\Delta)^2}\right)=0 \\ \dfrac{d}{df}\left(\dfrac{a\cdot f}{1+a^2f^2}+\dfrac{P\cdot b\cdot(f+\Delta)}{1+b^2(f+\Delta)^2}\right)=0 \end{cases} \qquad (2)$$

Having made differentiation in (2), we receive the following system of the equations:

$$\begin{cases} \dfrac{a^2f}{\left(1+a^2f^2\right)}+\dfrac{P\cdot b^2(f+\Delta)}{\left(1+b^2(f+\Delta)^2\right)^2}=0 \\ \dfrac{a\cdot\left(1-a^2f^2\right)}{1+a^2f^2}+\dfrac{P\cdot b\cdot\left(1-b^2(f+\Delta)^2\right)}{\left(1+b^2(f+\Delta)^2\right)^2}=0 \end{cases} \qquad (3)$$

Let's enter designations:

$$u = a\cdot f; \quad v = b\cdot(f+\Delta). \qquad (4)$$

Then after transformations we get

$$\begin{cases} a\cdot u\cdot\left(1+v^2\right)^2+P\cdot b\cdot v\cdot\left(1+u^2\right)^2=0 \\ a\cdot\left(1-u^2\right)\left(1+v^2\right)^2+P\cdot b\cdot\left(1-v^2\right)\left(1+u^2\right)^2=0 \end{cases} \qquad (5)$$

Further we make a variable replacement:

$$m = a\cdot(1+v^2)^2, \quad n = P\cdot b\cdot(1+u^2)^2. \qquad (6)$$

Then

$$\begin{cases} m \cdot u + n \cdot v = 0 \\ m \cdot (1 - u^2) + n \cdot (1 - v^2) = 0 \end{cases} \quad (7)$$

The decision of this system is:

$$\begin{cases} u = \pm\sqrt{\frac{n}{m}} \\ v = \mp\sqrt{\frac{m}{n}} \end{cases} \quad (8)$$

By substituting (6) in (8), we receive:

$$\begin{cases} u = \pm\sqrt{\frac{a}{P \cdot b}} \\ v = \mp\sqrt{\frac{P \cdot b}{a}} \end{cases} \quad (9)$$

By substituting these expressions in (4), we finally receive:

$$\begin{cases} f_b = \pm\frac{1}{\sqrt{P \cdot a \cdot b}} \\ \Delta_b = -f_b \cdot (P + 1) \end{cases} \quad (10)$$

From (10) it is possible to define the maximal value of generalized detuning ξ_b, which can be brought in by the oscillator without occurrence of frequency hysteresis effects:

$$|\xi_b| = \sqrt{\frac{Q}{Q_a} \cdot \frac{R_a}{R}} \quad (11)$$

In this point the reactance module of the resonator is equal to $Xb = R \cdot |\xi_b|$. In Fig. 2 the line corresponding to the equation $Im(Zq) = Xb$ divides the area of resonator into two parts: above this line there is a loop on resonator impedance hodograph, (so, there is a frequency hysteresis), lower this line there is no hysteresis (we consider area, where $\xi > 0$).

If the ratio (10), connecting together Δ_b and f_b values, is used in the expression (1), we 'll get the following function:

$$\lambda(f) = R \cdot \{(1 + j \cdot a \cdot f)^{-1} + P \cdot [1 - j \cdot P \cdot b \cdot f]^{-1}\}^{-1} \quad (12)$$

The diagram of this function also is given in Fig. 2. It is seen, that this function is bending relative to the excitations brought by the parasitic mode shift. The diagram of this function also divides the area of resistance into two parts. If the module of a real part of input oscillator resistance begins in the point located to the right of this line, the failure of fluctuations will not be observed; if it is to the left, there can be a failure of fluctuations.

Thus the area of resonator impedance (at positive generalized frequency shifts) can be divided into four zones designated in Fig. 2 by symbols A, B, C, and D accordingly. Depending on location of the start point of -Zg hodograph, four models of behavior of the generator in a range of temperatures turn out:

A - there is no failure of fluctuations and there is no hysteresis of frequency;
B - there is no failure of fluctuations, but there is a hysteresis of frequency;
C - there is a failure of fluctuations and there is a hysteresis of frequency;
D - there is a failure of fluctuations, but there is no hysteresis of frequency.

Proceeding from a generalized detuning to a relative one, the boundary relation of resistance and quality factor of parasitic mode may by written as:

$$\frac{R_a}{Q_a} = 4 \cdot R \cdot Q \cdot \left(\frac{\Delta F}{F_0}\right)^2 \quad (12)$$

Besides, using (1) it is possible to define a real part of a system of two oscillatory circuits resistance in a point of loop occurrence:

$$R_S = R \cdot \left(\frac{Q_a + Q}{Q_a + Q \cdot P}\right) \quad (13)$$

Let's consider the use of the received expressions on an example of thermocompensated oscillator based on AT-cut resonator, working in a temperature range of -60...+90 °C. For such range of temperatures the amplitude of the resonator F-T curve makes value about $40 \cdot 10^{-6}$. Hence, the oscillator, as a minimum, should provide the same value of relative frequency adjustment. Let's assume, that the resonator has the following parameters: $R_0 = 10$; $Q_0 = 0.5 \cdot 10^6$. Then from (12) we get, that for the absence of resonator F-T curve hysteresis anomalies it is necessary to carry out the ratio $Qa < 31.25 \cdot Ra$ for parasitic mode of oscillations. If we accept $R_a = 1000$, then it is necessary, that $Q_a < 31250$. On the border of this expression (when instead of an inequality there is an equality) from (13) we get, that the motional resistance of the quartz resonator can be increased by 14.7 times, that in some cases can result in failure of oscillations.

In Fig. 3 and 4 the diagrams of a deviation of resonator frequency (Fig. 3) and its motional resistance (Fig. 4) are represented depending on frequency shift Δ of a parasitic mode at different values of capacitor reactance of the oscillator Xg. The frequency change of a parasitic mode on these diagrams is equivalent to temperature change (determined by a F-T curve slope), therefore the meaning along the axes X in Fig. 3, should be considered as resonator temperature expressed in some conventional units.

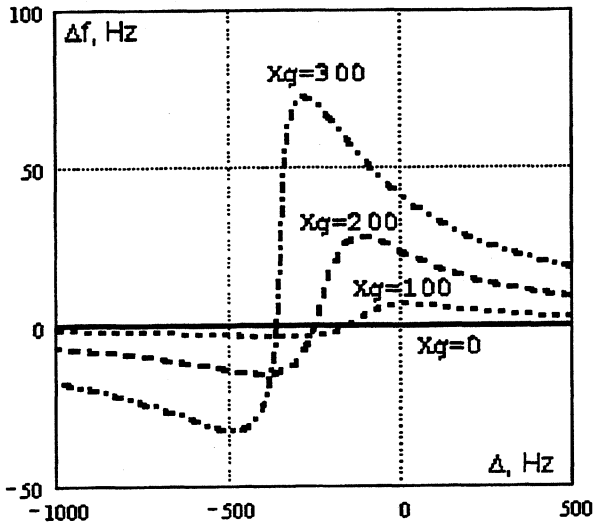


Fig. 3. Main mode frequency shift vs. parasitic mode frequency offset

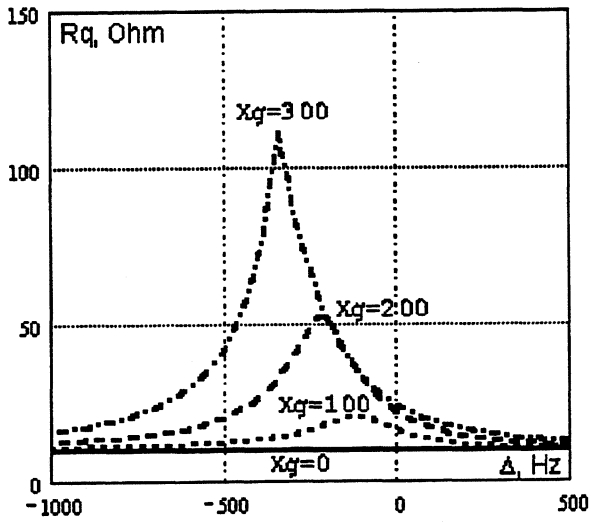


Fig. 4. Resonator motional resistance vs. parasitic mode frequency offset

From these figures it is obvious, that the increase of reactance of the oscillator can result in deterioration of temperature stability of the oscillator, creating local irregularity of frequency and motional resistance vs. temperature curves.

CONCLUSIONS:

- The increase of the frequency shift of basic oscillation mode in the crystal oscillator increases the parasitic modes influence on oscillator F-T curve and motional resistance. Therefore for reduction of parasitic modes influence in TCXO it is desirable to use the excitation circuits with negative initial frequency shift, that will allow to minimize the maximal value of the frequency shift module.
- In resonators, which are supposed to be used in TCXO, it is necessary to pay special attention to activity of parasitic modes. For control of resonator parameters it is necessary to take into account the dependence of activity dips and F-T curve of the resonator on inserted frequency shifts.
- The frequency shift inserted in the oscillator with the purpose of correction of frequency, (for example, because of aging) changes the value and location of resonator local F-T curve irregularity, that breaks thermocompensation accuracy. In this connection it is necessary to pay attention to alternative methods of reorganization of frequencies connected to influence on the resonator, for example, management of a constant electrical field, management of local heating piezoplate etc.

REFERENCES.

[1] Piezoelectric resonators: The handbook/ V.G. Androsova, E. G. Bronnikova and others. /Under P. E. Kandyba and P. G. Pozdnyakov edition. Moskow. Radio and Communication. 1992 – 392 p. (in Russian)

[2]. M. Toki, Y. Tsuzuki. Measuring Method Of Equivalent Series Capacitance And Negative Resistance Of Quartz Crystal Oscillator Circuits. Proc. 37th Annual Frequency Control Symposium. Philadelphia, Pa, 1-3 June 1983, pp. 300-305.

CHARACTERIZATION OF LOW-DISLOCATION SYNTHETIC QUARTZ GROWN ON HIGHLY DISTORTED SEED BY X-RAY TOPOGRAPHY

Armando H. Shinohara*, Marcos C. Iano*, Carlos K. Suzuki* and Yutaka Mikawa**
*UNICAMP-University of Campinas, Campinas-SP, Brazil
**Fine Crystal Co. Ltd, Muroran, Sapporo, 051, Japan

Dept. of Materials Engineering, School of Mechanical Engineering, University of Campinas, CP6122, CEP 13083-970, Campinas-SP, Brazil, Tel:+55-19-7883309, Fax: +55-19-2893722, E.mail: hideki@fem.unicamp.br

ABSTRACT

In the present study, the feasibility to grow a dislocation-free synthetic quartz with large Z-region from a seed containing high density of dislocation is reported. For such purpose, a seed with new geometric design was prepared. A seed long in Y-direction and containing V-shaped cuts with cutting angle of 90° made on Z-face {0001} was prepared and grown in the hydrothermal process.

As a result, new growth regions usually not found in the conventional Y- and Z-bars synthetic quartz crystals have been imaged by the X-ray topography. The new growth region grown from the V-shaped cut was composed with two sub-regions of distinct textures. Due to their high growth velocity perpendicularly to cut faces, they disappeared and were replaced by the so-called Z-region. However, these newly grown sub-regions played an important role to change the direction of dislocation propagation present in the seed trapping them and minimizing their propagation into the grown Z-region.

1. INTRODUCTION

Usually, synthetic crystals for quartz resonators are grown on the seed of Y-bar or Z-plate along Y-direction. The advent of cellular and satellite networks using high frequency, smaller, and cost effective resonator is driving force to improve uniformity in quartz material. It well known that seed quality is one the main factors which affects the final quality of synthetic quartz crystal grown in the hydrothermal growth conditions. From the technological point of view, defects such as dislocation affect significantly the quartz crystal performance. Increasing use of photolitography in quartz component manufacture, etch channel and dislocation densities play a critical role, as does etch pit formation. So far, several investigations have been conducted to improve the crystalline quality and purity of synthetic quartz crystal for such purposes [1-14].

In a previous study [9,14], synthetic quartz crystals were grown on S-, ξ-, intermediary-bar and cylindrical seeds and characterized by X-ray topography and chemical analysis of impurities. As a result, several new growth regions usually not present in the Y- and Z-bars synthetic quartz crystals are observed. Furthermore, even though using a low-quality seed, that investigation showed a possibility to grow a synthetic quartz crystal of high crystalline perfection with almost dislocation-free but with large Z-region. Therefore, motivated by those results, newly designed seed containing high density of dislocations was prepared and growth conducted to obtain a synthetic quartz crystal with large Z-region and low dislocation-density. Characterization study was conducted by the X-ray double-crystal topography.

2. MATERIALS AND METHOD

In the present study, Z-region of a Z-bar synthetic quartz crystal containing about 100 dislocation/cm^2 was chosen as seed material. The V-shaped cuts were made in Z-direction [0001] using a diamond saw, as schematically shown in Fig. 1.

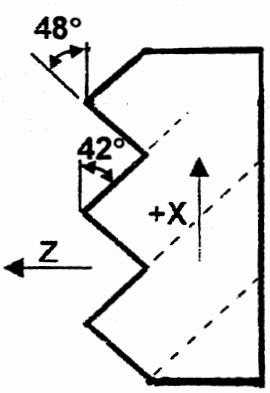


Figure 1. Schematic diagram of newly designed seed containing V-shaped cut of 90°.

After V-shape cutting process, the seed was polished using silicon carbide powders of #320 and #800 mesh. In order to eliminate the tension introduced in the cutting and polishing processes the seed was etched in a hydrofluoric acid solution (40%) for 30 min. The growth of synthetic quartz crystal was conducted in an experimental autoclave with the following growth condition listed in Table I.

The characterization study was conducted by a X-ray double-crystal topography of non-parallel (+,-)

0-7803-5400-1/99/ $ 10.00

topography of non-parallel (+,-) setting [15] using the Laue case (transmission geometry) and CuKα radiation generated in a sealed-off fine-focus X-ray tube of 1.5 kW maximum power.

Table I. Synthetic quartz crystal growth Condition.

Growth Rate	0.68mm/day
Growth Temperature	356°C
Pressure	1300 kgf/cm^2
Solution	3.8% NaOH
Growth Period	39 days
Autoclave Dimension	ϕ 200mm x 3.5m

X-ray radiation was generated by 3 kW stabilized transform X-ray generator DMAX2200 of Rigaku International and a high precision goniometer model KTG-11 of Kohzu Seiki for obtaining X-ray topographic images. The CuKα radiation was monochromated by a highly asymmetric Si(111) monochromator with asymmetric factor 1/b = 20. The distance from monochromator and sample was fixed to be 68 cm. This setting is slightly dispersive but it helps to obtain very sharp X-ray topographic images with high sensitivity respect to the lattice distortions. Sample for X-ray topographic imaging was extracted perpendicularly to Y-axis from the center of as-grown synthetic quartz crystal. Its thickness was adjusted to have μt = 3.0 to 3.5 (where μ is linear absorption coefficient and t was the sample thickness). In this case, X-ray topographic images are formed by the intermediary image of dynamical and kinematical diffraction effects giving a better contrast.

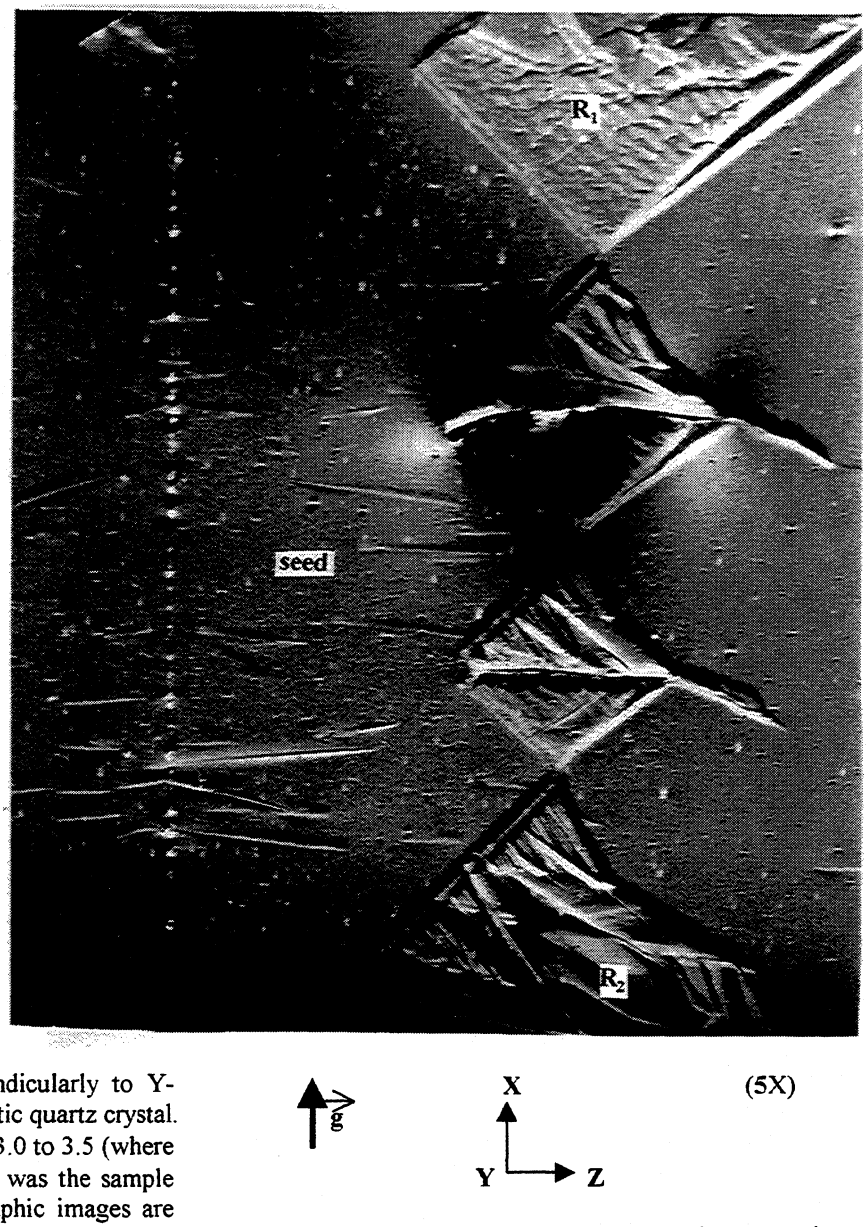


Figure 2. X-ray topographic image of Y-cut plate obtained using (20$\bar{2}$0) diffraction plane.

3. RESULTS AND DISCUSSION

Figure 2 shows the X-ray topographic image of synthetic quartz crystal grown from the seed with two V-shaped cuts and two large oblique faces to Z-direction obtained using asymmetric (20$\bar{2}$0) diffraction planes. From this image, a total of four new growth regions and two tails were observed. The region grown from V-shaped cut is basically composed with two sub-regions of distinct textures, which are separated by a plane of fluid inclusion grown from the vertex of V. The appearance of fluid inclusion plane is not well understood yet, but its size depended on the V-shaped cutting depth. The depth of V-shaped cut in the seed shown Fig.2 is 2.5 mm and formed a plane of fluid inclusion of about 5 mm. Concerning to the growth direction of the newly grown sub-regions, according to a previous study [9], they grow perpendicularly to the cut faces. The texture of each sub-region was similar to the texture of regions R_1 or R_2, depending on the growth direction. However, The sub-regions grew from the V-shaped cut faces and met each other at the plane of fluid inclusion forming square-like figure region. After at certain growth period, probably due to the high growth velocity, they are replaced by so-called Z-region. Furthermore, before a complete replacement by the Z-region, a "tail" was formed at the end of fluid inclusion plane and extended into Z-region for a determined length almost perpendicularly to one of V-cut face. Its origin can be attributed to the growth direction and presence of a plane of fluid inclusion. The size of the

"tail" seemed to be independent of the V-shaped cut depth.

Concerned to the present target, it was clearly observed from the X-ray topographic image (Fig.2) that, when a flat face is made on a seed perpendicularly to Z-direction dislocation present in it propagate into the grown Z-region, as well known. On the other hand, V-shaped cut practically stop the dislocation propagation into as-grown Z-region. This effect can also be explained based on our previous study on S- and ξ-bar synthetic quartz crystals growth and characterization by X-ray topography [9]. Dislocation tends to propagate almost perpendicularly to the growth face. Furthermore, the sub-regions grown in the V-shaped cut grow perpendicularly to the cut faces. The both sub-regions form a larger region containing a plane of fluid inclusion and it is replaced by Z-region. In this sense, V-shaped cut plays an important role to change the direction of dislocation propagation and trap them into the newly grown sub-regions.

4.SUMMARY

From the present study, the following conclusions could be obtained.

The size of fluid inclusion plane increased as function of increase of depth of V-shaped cut.

The present study results showed a method to minimize the propagation of dislocation and obtain a large Z-region even though using a seed of low quality making V-shaped cut in Z-direction.

5. ACKNOWLEDGMENT

The Authors thank FAPESP, Capes and CNPq for the financial support. We would like also to thank Dr. Y. Banno and Mr. M. Hatanaka of Fine Crystal Co., Ltd for continuous encouragements.

6. REFERENCES

[1] A. Zarka, L. Lin and M. Buison, M., "Influence de la Localisation Sectorielle du Germme sur la Qaulité Crystalline de Quartz de Synthese," J. Crystal Growth, vol.54, pp.398, 1981.

[2] A. Zarka, L. Lin and M. Buison, "Influence du Germe sur la Densité de Dislocations Produites lors de la Croissance de Quartz de Synthese," J. Crystal Growth, vol.57, pp.466, 1982.

[3] A.F. Armington and J.F. Balascio, "The Growth of High Purity, Low Dislocation Quartz," Proc. 38th Ann. Freq. Control Sympos., USA, 3-7, 1984.

[4] S. Taki, "Improvement of Growth Process and Characterization of Quartz Crystals," Prog. Crystal Growth and Charact., vol.23, pp.313-339, 1991.

[5] F. Iwasaki, "Line Defects and Etch Tunnels in Synthetic Quartz,", J. Crystal Growth, vol.39, pp.291-298, 1977.

[6] C.K. Suzuki, M.S. Tanaka and A.H. Shinohara, "Growth and Characterization of Optical Grade Synthetic Quartz", Proc. 1996 IEEE International Frequency Control Symposium, Honolulu, USA, 78-83, 1996.

[7] R.L. Barns, P.E. Freeland, E.D. Kolb, R.A. Laudise and J.R. Patel, "Dislocation-Free and Low-Dislocation Quartz Prepared by Hydrothermal Crystallization," J. Crystal Growth, vol. 43, 676-686, 1978.

[8] S. Ishigami, M. Sato, F. Uchiyama, K. Agatsuma and K. Tsukamoto, "Growth of High Quality Quartz Crystal and Its Application to Temperature Sensors,", Proc. 1994 IEEE International Freq. Control Symp., 99-106, 1994.

[9] A.H. Shinohara and C.K. Suzuki, "Study of S- and ξ-Bars Synthetic Quartz By X-Ray Topography," Proc. 1996 IEEE International Frequency Control Symposium, Honolulu, USA, 71-77, 1996.

[10] Y. Mikawa, M. Hatanaka and Y. Banno, to be published in this procceedings of Joint Meeting of EFTF and FCS, 1999.

[11] J.F. Balascio and T. Lind, "The Growth of Piezoelectric Alpha Quartz Crystals," Current Opinion in Solid State & Mater. Sci., vol. 2, 588-592, October 1997.

[12] M. Hosaka and T. Miyata, "Hydrothermal Growth of Alpha-Quartz Using High-Puirty Alpha-Cristobalite as Feed Material,", Mater. Res. Bul., vol.28, pp. 1201-1208, November 1993.

[13] C.K. Suzuki, C. Farias, S.M. Takiya and J. Kiss, "Seed Technology for Synthetic Quartz", Cerâmica, vol. 32, 303-306, 1986.

[14 F. Iwasaki, A.H. Shinohara, H. Iwasaki and C.K. Suzuki, "Effect of Impurity Segregation on Crystal Morphology of Y-Bar Synthetic Quartz," Jpn. J. Apply. Phys., vol.29, pp.1139-1142, June 1990.

[15] K. Kohra, H. Hashizume and J. Yoshimura, "X-Ray Diffraction Topography Utilizing Double-Crystal Arrangement of (+,+) or Non-Parallel (+,-) Setting,", Jpn. J. Appl. Phys., vol.9, pp.1029-1038, 1970.

1999 Joint Meeting EFTF - IEEE IFCS

HIGH VOLTAGE DIFFUSION ALONG Z- AND X-AXES ON QUARTZ CRYSTAL

M. SMAALI, J.J. BOY, J.B. BRIOT

ENSMM / LCEP – 26, ch. de l'Epitaphe - 25030 BESANÇON Cedex - FRANCE

ABSTRACT :

Quartz resonator quality is improved using sweeping treatment. It is well known that this operation is usually performed by applying a high electric field along the optical axis (Z) of the quartz lattice. This is due to the existence of channels formed by helicoidal structure of quartz crystal. This crystallographic structure exhibits other channels parallel to the electrical axis (X) which are larger than the previous ones.

In order to compare the influence of each kind of sweeping on the quality of quartz resonator, we have performed Z- and X-sweeping on natural and cultured quartz bars used to manufacture 10 MHz SC-cut resonators. As usually, the process of qualification has been achieved by comparisons between spectrometric measurements (infrared, laser absorption, ICP) before and after treatments.

INTRODUCTION :

The work presented here started after fruitfull discussions with Doctor Philippot, french researcher in crystallochemistry, who said that the X-channels (which means channels along A2) are structurally larger than these along Z axis. And so, first of all, the authors want to thank him. The study of previous references does not convince us of the uselessness of our work !

Indeed, since the first works of J.C. King [1], sweeping is usually achieved by applying the electric field along the Z-axis of quartz (and from seed side to the external one for synthetic crystal quartz).

So, quartz crystal sweeping is a widely used process to improve the quality of quartz crystal resonators especially for use in ultra stable oscillators. Some defects are associated with alkali ions (generally Na^+ or Li^+) that are always found inside both natural or synthetic quartz (due to the growth process, the Li "impurity" level is very low in high-Q synthetic quartz). Under a high electric field, these impurities migrate through the crystal which can be cleaned by a following lapping step (to prevent that impurities migrate into). Through IR spectroscopy and chemical analysis (ICP) [2], it has been shown previously that such defects form Al-M centers inside the crystal (where M means any alkali ion) and are replaced after sweeping process by either Al-OH centers or Al-h centers in case of a sweeping under vacuum (where h means "hole"). However there is no clear evidence for Z-axis to be the most efficient channel for ionic migration. Indeed, the channel diameters along X- and Z-axes are similar. Then it appears to us interesting to check the sweeping efficiency in applying electric field along X-axis.

Comparative studies between sweeping along Z- and X-axes will be reported in this paper. Systematic measurements have been done on both synthetic and natural quartz. Laser absorption, IR spectroscopy and chemical analyses have been correlated to evaluate the effect of sweeping on each kind of defects. Finally, some resonators have been cut in each case and electrical characteristics of these resonators will be presented ... in a next paper.

I. USED SAMPLES AND SWEEPING TECHNIQUES :

Two quartz samples are used in this work. The first one is synthetic, and four bars are cut inside. The second one is natural, and three bars are cut inside. Their dimensions are a few cm along Y-axis and 12 to 15 mm along the other directions. The table below includes all bars used and the type of treatment applied to each one.

Table I- Different quartz material used and its treatment received

Synthetic quartz	Natural quartz	Type of treatment
SB1	NB1	Untreated
SB2	NB2	Z-swept
SB3a and SB3b	NB3	X-swept

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

Sweeping experiment is achieved in a closed system under dry nitrogen atmosphere. Before each experiment, each bar has been dipped in hydrochloridric acid at temperature close to 70°C to be cleaned. The same treatment has been applied to the electrodes, which are in platinum.

After this step, sample is pressed between the electrodes, and introduced in the furnace. A low vacuum is then created in the closed system which is purged with dry nitrogen vapor. The heating process is started at rate of five degrees per minute from a room temperature to 500°C. The electric field is applied few hours after the desired temperature would be reached. To compare the effect of crystallographic orientation, one sample from each origin is swept along Z-axis, and the other one is swept along X-axis. To justify these choices and to attribute in which cases, sweeping efficiency is more important analytical measurements such as laser absorption, infrared at room and liquid nitrogen temperature are used before and after sweeping process. As for quantitative measurements, we have compared the results of untreated samples to those of swept ones.

II. SAMPLES ANALYSES :

II.1. The 632.8 nm laser absorption :

This technique consists in measuring the 632.8 nm laser radiation transmitted through the two parallel polished faces of the quartz sample [3]. The laser beam, emitted by a 2 mW He-Ne source, is 0.8 mm width. The transmitted intensity is measured by a silicon photodiode detector and stored. The sample is translated in front of the laser beam with two micrometric motion units controlled by computer. The sample is analysed plane by plane as indicated in [4], the step between two successive data points being 1 mm and between two planes 2 mm.

The absorption values are calculated from the following relationship :

$$\sigma_L = \frac{1}{d}.\log_{10}\frac{I_0}{I_n}$$

where :
d is the thickness of the sample, I_0 and I_n being the measured intensities outside the sample and at the point number n.

The mean value of σ_L for one plane is given by :

$$\overline{\sigma}_L = \frac{1}{d}.\frac{1}{N}.\sum_{n=1}^{N}\log_{10}\frac{I_0}{I_n}$$

The two following graphs concern the natural quartz just for a medium plane (P10). We show that the σ_L values obtain for a bar swept along Z-axis is very inhomogeneous and higher than the values obtained before sweeping. Furthermore the existence of streaks of brown coloration confirms the creation of a lot of Al-hole centers.
In contrary, when the sweeping is performed along X-axis, no coloration appears and the difference in the absorption of the laser beam between untreated and treated sample is very low, as indicated here and in the table below (with the normalized value α_{3500}).

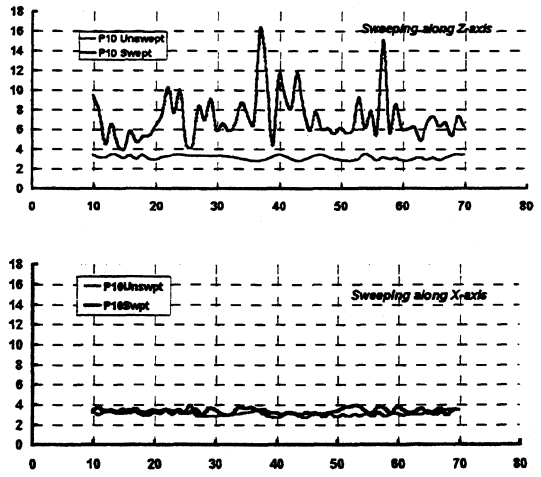


Fig. 1a, 1b : Laser absorption on natural quartz swept along Z- and X-axis [σ_L (x 10^{+2}) vs Point Nb]

Table 2 : $\overline{\sigma L}$ for natural quartz

	NB2		NB3	
cm^{-1}	Before sweep.	After X-sweep.	Before sweep.	After Z-sweep.
$\overline{\sigma L}$	0.0314	0.0341	0.0322	0.0724
α_{3500}	0.033	0.032	0.034	0.029

As the Table 3 (see below), it indicates that the change of the electric field sense along X-axis does not induced visible difference on the mean value of σ_L on the synthetic quartz samples (SB3a and SB3b).

Table 3 : $\overline{\sigma L}$ for synthetic quartz

	SB3a		SB3b	
cm^{-1}	Before sweep.	After X-sweep.	Before sweep.	After X-sweep.
$\overline{\sigma L}$	0.037	0.044	0.035	0.043
α_{3500}	0.025	0.023	–	–

II.2. InfraRed spectrometry :

In synthetic quartz :

All untreated samples present the same degree of defects, due to the same spectra obtained with each bar (Fig. 2). This shows the homogeneity of this quartz. The 3580 cm^{-1} band (also called s_4 band) characterizes untreated synthetic quartz. This absorption band is due to the as-grown OH vibrations. It is present in all spectra with the same

intensity. We can also observe the 3438 cm^{-1} band (s_3) and the 3396 cm^{-1} band (s_2).

The spectrum of the quartz which was swept along Z-axis shows an important decrease of the so called s_4 band. At the same time a new absorption band due to the sweeping treatment appears. This band is called e_2 and it is located at 3366 cm^{-1}. By comparison of these results with those obtained with the X-swept quartz, this last ones show a few modifications of its spectrum. Although the e_2 band is present, its intensity is lower than in the case of the Z-swept quartz. The decrease of s_4 band is very low too. The same result is obtained with other synthetic quartz swept along X-axis. This allows us to conclude that sweeping along X-axis is possible, but its efficiency is lower in comparison to sweeping along Z-axis, despite of the presence of a large channel along X-axis.

In natural quartz :

Infrared spectra of Z-swept quartz show the presence of the 3466 cm^{-1} band. In contrary, the 3473 cm^{-1} band, which appears in infrared spectrum of untreated natural quartz, disappears. As for e_1 and e_2 bands, the comparison is proved with no need because of their strong intensities in all infrared spectra.

As for X-swept quartz, few modifications have been observed between its infrared spectrum and those of the untreated sample (Fig. 3, 4). The 3473 cm^{-1} is well present and a small band at 3466 cm^{-1} appears. This let us tell that alkali impurities related to the 3473 cm^{-1} absorption band are not completly eliminated from the crystal. Then infrared analyses of both synthetic and natural quartz show that sweeping treatment along Z-axis is more efficient than this along X-axis.

α_{3500} measurements :

Other comparison of infrared data is induced in measuring the α_{3500} coefficient which determines the quality of quartz. These values presented in tables 2 and 3 well confirm, in each case and for each quartz quality, the efficiency of the sweeping process because of the decrease of this coefficient.

II.3. : ICP analyses :

As for quantitative analyses (see below table 4), their results obtained with Z-swept quartz are more significant. Nevertheless chemical analyses of the X-swept quartz - synthetic and natural - show that alkali impurities content has been reduced too, but not in the same ratios as those of Z-swept quartz.

Table 4 : Chemical analysis results of the principal impurities given in atomic ppm (atoms/10^6 Si)

Element	SB1	SB2 Z-swept	SB3 X-swept	NB3	NB3 X-swept	NB2	NB2 Z-swept
Fe	0.822	0.820	0.823	0.866	0.991	0.846	0.742
Al	2.635	2.709	2.775	18.532	19.690	19.961	19.520
Ca	0.860	0.797	0.734	0.377	0.402	0.461	0.361
Na	0.810	0.157	0.235	2.352	1.045	2.444	0.340
Li	0.087	0.000	0.087	8.741	3.029	9.261	0.000

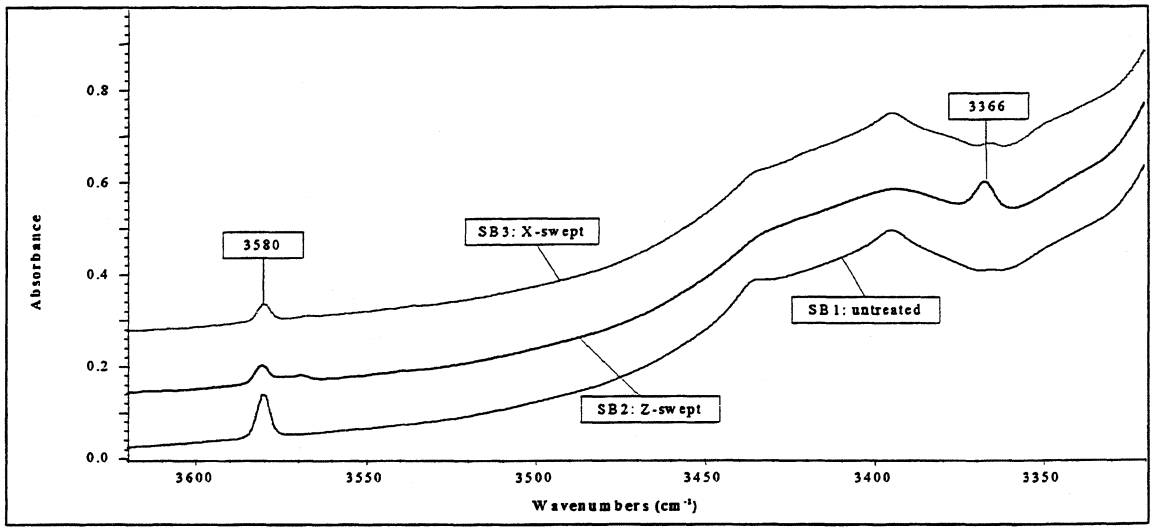


Fig. 2 : Infrared spectra of synthetic quartz

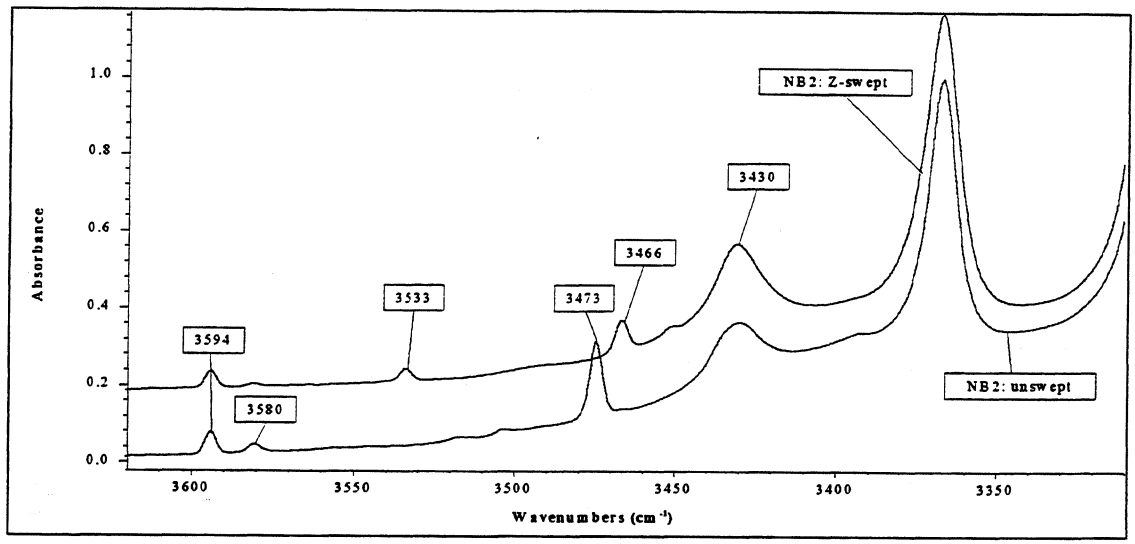


Fig. 3 : Infrared spectra of natural quartz before and after sweeping along Z-axis

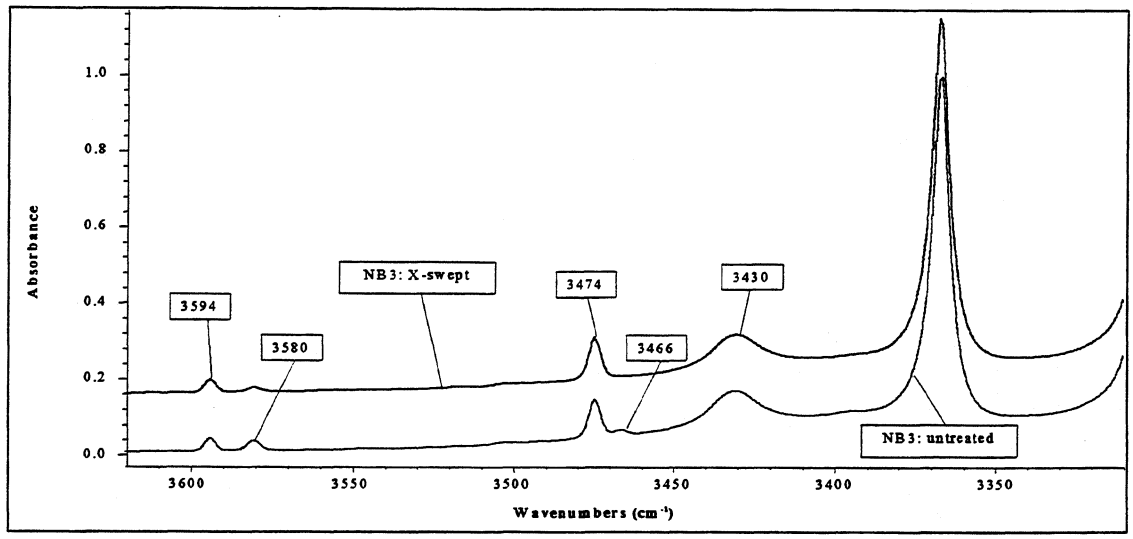


Fig. 4 : Infrared spectra of natural quartz before and after sweeping along X-axis

III. CONCLUSION :

Every experiment has confirmed that the sweeping along X-axis is not really efficient in comparison with the sweeping performed along the Z-axis.
Though synthetic quartz crystal exhibits a lot of dislocations almost parallel to the Z-axis, we do not think that they are responsible because, in natural quartz, they are rarely present...
We prefer to explain the small efficiency of the X-sweeping by observing the structure of the helicoidal channels parallel to X-axis. Due to the symmetry around Z-axis, we can found another symetric direction by rotating the crystal around Z-axis by 60°. By the way, this multicomponent of X-channels does not facilitate the interstitial impurities transport.

IV. BIBLIOGRAPHY :

1. **J.C. KING** : « Electrolysis of synthetic quartz : effect upon resonator performance », Proc. Of the Institution of Electrical Engineers, Vol. 109, Part B (1962), pp. 295, 301.
2. **H. BAHADUR** : « Electrodiffusion and Irradiation Effects in Crystalline Quartz », 5th European Frequency and Time Forum (1991), pp. 148, 156.
3. **P. ZECCHINI, H.G. LIPSON** : « Defect distribution mapping in quartz », 43rd Annual Symp. on Frequency Control (1989), pp. 490, 496.
4. **K. YAMNI, J.J. BOY, M.A. PASQUALI, P. ZECCHINI** : « Non-destructive study of Al-hole concentration and frequency change of quartz resonators », 8th European Frequency and Time Forum (1994), pp. 203, 214.

1999 Joint Meeting EFTF - IEEE IFCS

NEW TECHNIQUE TO DECREASE DISLOCATIONS IN SYNTHETIC QUARTZ CRYSTAL

Yutaka Mikawa, Motohide Hatanaka and Yasutaro Banno

Fine Crystal Co., Ltd.
9-1, Chatsu-machi, Muroran-shi, Hokkaido 051-0006, Japan
E-mail : mikawa@sanynet.ne.jp

ABSTRACT

Dislocations in synthetic quartz crystal are mainly propagated from the seed crystals. Therefore, low dislocation seed crystals are fundamental when growing low dislocation synthetic quartz. A new technique called "frame seed method" was developed to produce low dislocation seed crystal. In previous studies, it was confirmed that the technique cutting Z-cut seed from the +X region is effective to decrease dislocations. However, this technique has one problem; the Y-dimension becomes shorter than the original seed crystal as the result of the appearance of m-face. This new technique provided low dislocation seed without shortening of the Y-dimension by controlling the appearance of m-face. This technique can remove the other defects such as twins and cracks. Therefore applying this technique to the natural quartz containing twins or cracks, it is expected to remove these defects.

1. INTRODUCTION

These days, the quartz devices are getting higher frequency, higher stability and smaller size. The processes of devices are also developing, for example, the etching and photolithographic process are applied to tuning fork type resonators or high frequency filters. To satisfy these performances or to resist the processes, the quality of the quartz materials has been demanded to be high Q, low inclusions and low etch channels. Especially, the etch channel density directly effects the yield of devices using the heavy etching process. The low etch channel crystals are grown from only low dislocation seed crystals, because the dislocations are mainly propagated from the seed crystal. In previous studies, there are two confirmed techniques to get low dislocation seed crystals. One technique is to use high quality natural quartz as a seed crystal. But nowadays, these natural quartz crystals are very rare, expensive and sometimes it involves high risks. Another technique is to make Z-cut seed from the +X region of synthetic quartz. This technique is confirmed to be effective in many studies [1–8]. If the seeds were cut from the Z region, almost all the dislocations intersect the seed surface, therefore these dislocations are propagated to the next generation. On the other hand, in the +X region, the dislocations are nearly parallel to the surface of the Z-cut seed. Therefore few dislocations intersect the seed surface and few dislocations are propagated to the next generation. As a result, the low dislocation seed crystal is obtained. But one disadvantage of this technique is the shortening of the Y-dimension by appearance of m-face. For example, when the Z-cut seed is cut from the +X region: the X dimension is 70mm and the Y dimension is 250mm, the original seed crystal must be more than 492mm in the Y dimension as shown in figure 1. It is difficult to obtain such a long seed crystal.

In this study, we developed a new technique called "frame seed method" to overcome this disadvantage. Using this technique, almost all the area of the original seed was changed into the +X region without shortening of the Y dimension. Furthermore this technique can be applied for removal of the other defects such as twins or cracks.

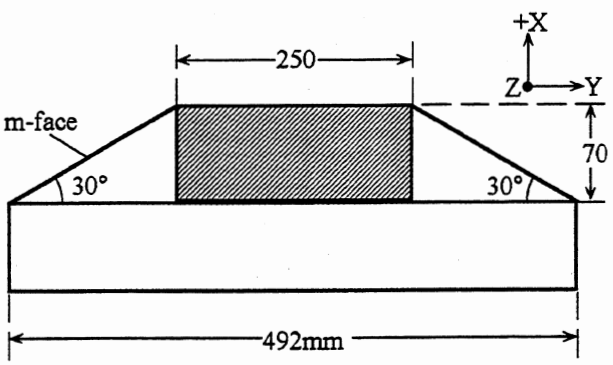


Figure 1 Z-cut seed from +X region

2. NEW TECHNIQUE

When the Z-cut seed is cut from the +X region, the Y dimension becomes shorter than the original seed as the result of the appearance of m-face. The appearance of m-face has been considered to be unavoidable on the growth of synthetic quartz crystal. If the appearance of m-face can be controlled, it is possible to make the Z-cut seed from +X region without shortening of the Y-dimension. To solve this problem, a new technique called "frame seed method" was developed. The process of this technique is described in figure 2. First, a Z-cut seed which has the required dimensions is prepared. Second, the seed is cut into "three sided frame shape" leaving the –X side and both the Y sides. The cutting direction of both the Y sides is parallel to the X-axis and cutting direction of the –X side is parallel to the Y-axis. Then, this seed is grown in the autoclave under normal growth condition. After growth, the removed area is filled with the +X region without appearance of m-face. If the X dimension of the +X region was not enough in one growth operation, more than once operation can be applied until it reaches required dimension.

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

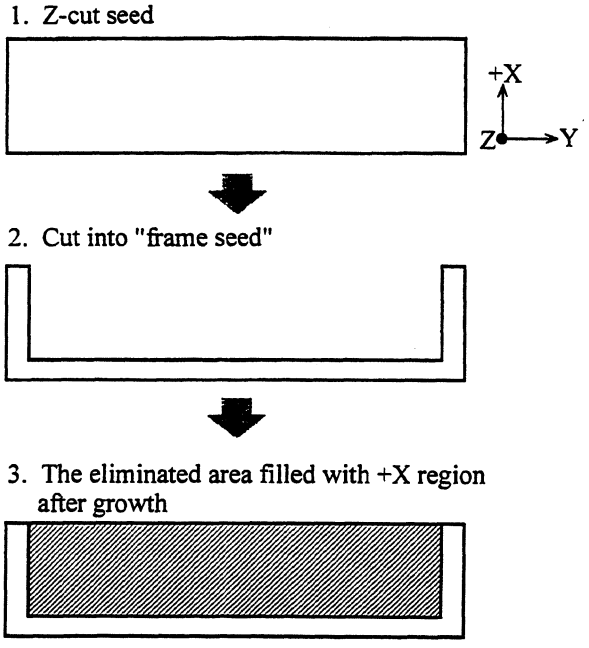


Figure 2 Process of "frame seed method"

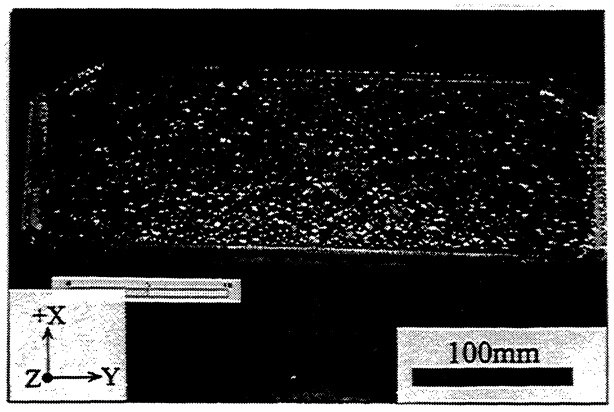


Figure 3 Z-plate as-grown crystal used as the starting material of the experiment

3. EXPERIMENTAL AND RESULT

This "frame seed method" was confirmed using commercial large autoclave. As the starting material, a large Z-plate quartz crystal was prepared as shown in figure 3. The dimensions of X, Y and Z are 100mm, 380mm and 25mm, respectively. The etch channel density of this crystal amounted to 150 per square centimeter. Then this crystal was sliced into Z-cut seeds using multi-blade-saw. The thickness of these seeds was 2mm. One of these seeds was cut into frame seed shape as shown in figure 4. The eliminated area is 310mm in Y dimension and 90mm in X dimension. This frame seed was grown in the autoclave under the conditions – autoclave size: 650mmID x 14m, pressure: 145MPa, temperature: 360°C and the solution: 4% NaOH. The total growth period amounted to 270 days (90 days x 3

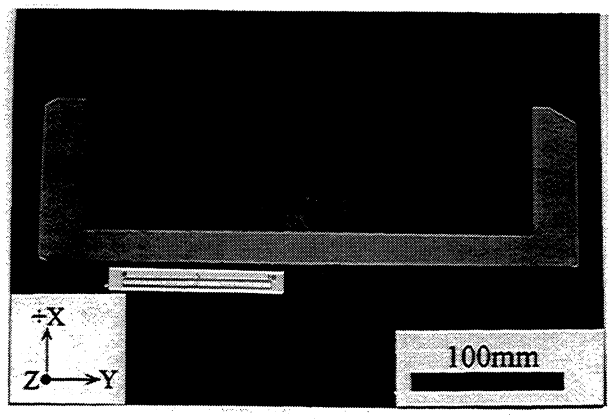


Figure 4 Frame seed (eliminated area: X90 x Y310 mm)

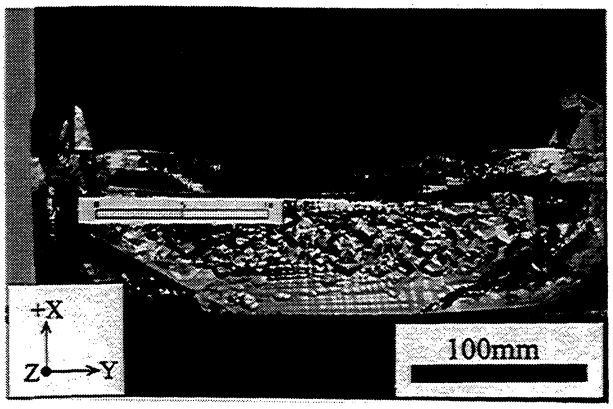


Figure 5 The grown crystal from the frame seed (observed from Z side)

runs). After growth, the eliminated area was filled with the +X region as shown in figure 5. The surface pattern of the filled area gives the information that this area is occupied by the +X region as shown in figure 6 and 7. From this crystal, Z-cut seeds ware sliced and then the seeds were grown in the autoclave to confirm the effect of decreasing dislocations. Figure 8 shows the as-grown crystal grown from the improved seed. There is an apparent difference in the cobble pattern on the Z face between the original crystal and the improved crystal. There are many cobbles on the original crystal caused by dislocations in the crystal (figure 3). On the contrary, the improved crystal shows that there are very few cobbles on the Z face; that suggests this crystal contains few dislocations. This crystal was etched by 40% HF to count the etch channel density and, as a result, it was less than 3 per square centimeter.

4. DISCUSSION

The "frame seed method" is expected to be useful for not only decreasing dislocations but also removing other defects such as twins or cracks. In the commercial growth, the seed crystals containing twins are not used. Therefore little attention is paid to how to control twins in usual commercial growth. However the initial seed

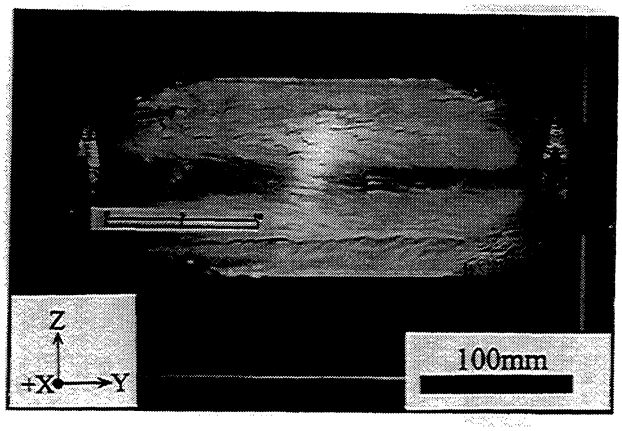


Figure 6 The grown crystal from the frame seed (observed from +X side)

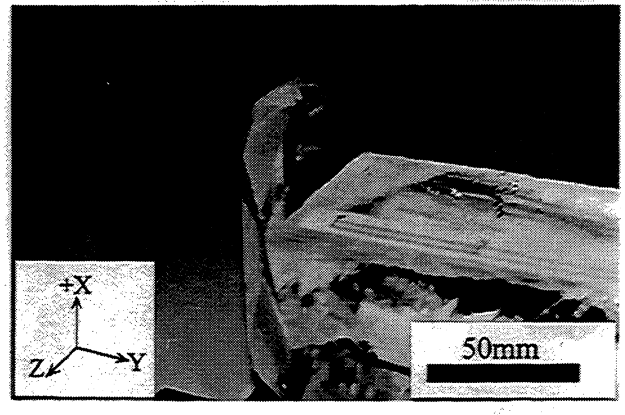


Figure 7 The grown crystal from the frame seed (near the edge of Y side)

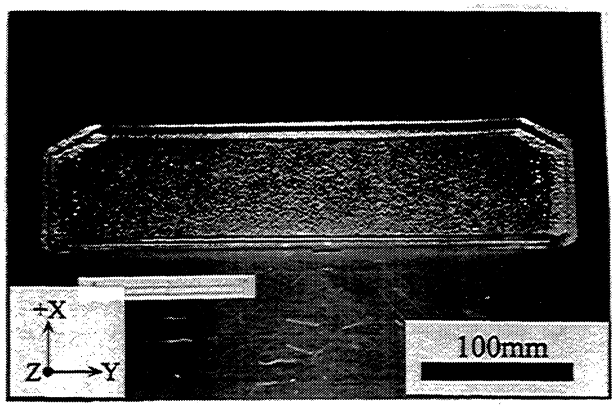


Figure 8 Z-plate as-grown crystal grown from the improved seed. (etch channel density: < 3per cm^2)

crystal must have been natural quartz crystal. The twins are frequently observed in natural quartz crystals, therefore the natural quartz crystals with both large dimensions and no (or small) twins or cracks are very rare and expensive. When the Z-cut seeds are cut from the natural quartz crystal containing twins, the dimensions become smaller than the initial crystal as the result of avoiding the twins. Figure 9 shows the process of removing the twins in natural quartz crystal

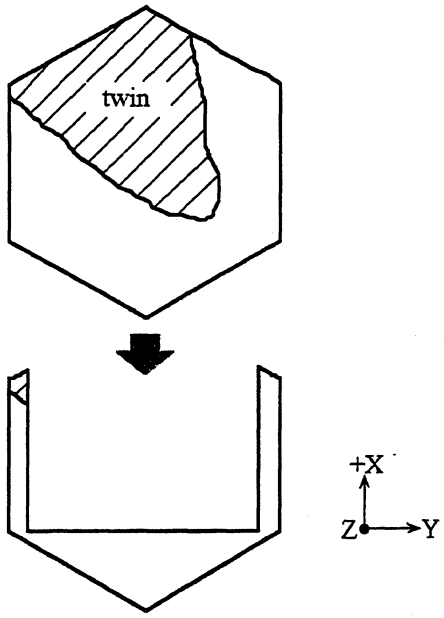


Figure 9 Removing the twin in the natural quartz

using the frame seed method. As quartz has three equivalent X-axes, a suitable X-axis can be selected for removing the twins most effectively. The eliminated area is filled with +X region after growth in the autoclave. Using this technique, the twins in the natural quartz crystal are removed still keeping the original size. Therefore, we can revive the natural quartz which were considered to be impossible to use as seed crystal because of their twins or cracks.

5. CONCLUSION

- The appearance of m-face was controllable using the frame seed method. As a result, it was possible to make Z-cut seed from +X region without shortening of the Y dimension.

- The etch channel density of improved crystal was less than 3 per cm^2.

- This technique can be used for removing the other defects such as twins or cracks. If this technique is applied to the natural quartz crystals with twins or cracks, these crystals which were considered to be impossible to use as seed crystal will be utilized.

6. REFERENCES

[1] A. Zarka, L. Liu and M. Buisson, "Influence de la Localisation Sectorielle du Germe sur la Qualite Crystalline de Quartz de Synthese", J. Cryst. Growth, vol 54, pp 394-398,1981

[2] A. F. Armington and J. J. Larkin, "The Growth of High Purity, Low Dislocation Quartz", J. Cryst. Growth, vol 71, pp799-802, 1985.

[3] A. F. Armington and J. F. Balasio, "The Growth of

High Purity, Low Dislocation Quartz", Proc. of the 38th Annual Freq. Cont. Symp., pp 3-7, 1984.
[4] A. F. Armington and J. F. Balasio, "The Growth of High Quarity Quartz in Commercial Autoclaves", Proc. of the 39th Annual Freq. Cont. Symp., pp 230-233, 1985.
[5] A. F. Armington, J. A. Horrigan, M. T. Harris and J. F. Balascio, "A Sutudy of Dislocations and Inclusions in Alfa Quartz", Proc. of the 41st Annual Freq. Cont. Symp., pp 213-215, 1987.
[6] J. F. Balasio and A. F. Armington, "Development Results for the Production of High Quarity Quartz", Proc. of the 40th Annual Freq. Cont. Symp., pp 70-75, 1986.
[7] G. R. Johnson and R. A. Irvine, "Etch Channels in Single Crystal Cultured Quartz", Proc. of the 41st Annual Freq. Cont. Symp., pp 175-182, 1987.
[8] R.A. Laudise, R.L. Barns, D.S. Stevens, E.E. Simpson, and H Brown , "High Performance Quartz", Proc. 42nd Ann. Freq. Contr. Symp., pp116-126, 1988.

HYDROGEN AND ITS RADIATION EFFECTS IN QUARTZ CRYSTALS

Harish Bahadur
National Physical Laboratory
Hillside Road, New Delhi-110012, INDIA
FAX:91-11-575-2678 & 91-11-576-4189, Phone(Work) 91-11-5787161 Extn: 2336

ABSRACT

Hydrogen trapped at numerous hydroxyl defects in quartz lattice serves to compensate the electron excess defects of trivalent aluminum during irradiation. We have measured near infrared (region 3100-3700 cm^{-1}) absorption in optical samples of quartz at 77 K in as-grown crystals and upon electron irradiation at 77 K before and after 300 K- irradiation.. The spectral changes have been utilized to measure the concentration of various Al-related defect centers in quartz.

1. INTRODUCTION

All crystalline quartz contains Al, H and Ge as ubiquitous impurities which get randomly incorporated in the growing crystal [1-7]. Their type and concentration limit the Q (quality factor) of a resonator. The Al^{3+} is the most pervasive impurity substitutionally replacing Si^{4+}in the quartz lattice. It is charge compensated by either alkalis or protons in natural quartz. In cultured crystal all the Al^{3+} is charge compensated with Li^{+} ions [4]. Quartz oscillator crystals upon exposure to ionizing radiations suffer both transient as well as steady-state frequency shifts and *Q* losses. These effects arise due to radiation induced modification of the impurity-related point defects. The ionizing radiation produces a large number of electrons and holes which move through the crystal until they recombine or get trapped by these impurities present in the quartz material. The electrons are trapped at Ge- sites while the holes are captured by Al atoms and form the Al-hole centers during the irradiation process which give their characteristic observable effects. We shall discuss here the point defects in quartz with special reference to hydrogen in crystals.We shall also present an extension of the Martin's work [8] to the cases of natural quartz crystals in quantitatively estimating the aluminum impurity and related point defects. As natural quartz is almost always used as the starting material in the hydrothermal synthesis of cultured quartz, these studies would help in appropriate selection of natural material for growing cultured quartz crystals of desired characteristics for frequency control purposes. The Al-related defects centers have been studied by using a variety of techniques [1-7]. The Al-OH^{-} centers can be detected by measuring the near infrared absorption and monitoring the absorption bands at 3306 and 3367 cm^{-1} [9]. Sweeping can exchange the aluminum charge compensators. First developed by King [10], the technique of sweeping is now commercially used to improve the radiation hardness of quartz. For studying the radiation effects, we have used good quality high- and low-H natural quartz crystals with nearly similar aluminum concentration (~50 ppm) and widely different hydrogen levels designated as' high-H' and 'low-H'.The cultured crystal used was a Sawyer Electronic Grade material. A detailed description of the sweeping, measurement and irradiation techniques is available elsewhere [3].

2. RESULTS AND DISCUSSION

Figures 1 and 2 respectively depict the representative spectra for as-grown conditions of cultured and natural quartz crystals and after their irradiation at 77 K and 300 K. It is evident from the series of spectra in these Figures that majority of the infrared bands initially present show a decrease in their strength at 77 K-irradiation. The 300 K-irradiation produces new bands at 3367 and 3306 cm^{-1} in the cultured sample related to the Al-OH^{-} centers [9]. In natural crystal, the intensity of these-as-grown Al-OH^{-} bands increases significantly after the 300 K-irradiation These results clearly show that the natural crystals have aluminum ions charge compensated with both alkalis as well as protons [3]. The 300 K-irradiation breaks the Al-M^{+} centers (M^{+} = Li^{+} or Na^{+}) in both types of crystals into a mixture of Al-hole and Al-OH^{-} centers. The hydroxyl defects in quartz can probably be best characterized by irradiating the crystal. At 77 K-irradiation, only protons move; they come to rest , mostly in shallow and some in deep traps. Warm up to room temperature releases most of them back to their original defect species; those not returning must have been captured by deep traps and need higher temperature for their release. With shallow and deep traps defined in this way, irradiation at 77 K and subsequent warm up can be used to estimate for the concentration of deep and shallow traps for protons in quartz crystals Irradiation at room temperature allows movement of both , protons as well as alkali ions. Spectra taken some time afterwards would then yield the combined concentration of deep trapping centers for both alkalis and protons. Subsequent effect of second irradiation at 77 K should then be compared with the effect of first irradiation at 77 K.

0-7803-5400-1/99/ $ 10.00

2.1 Measurement of Impurity-related Point Defects

(A) Natural Quartz:

We have used [11] the irradiation cycle for a quantitative measurement of deep and shallow H-filled, Li-filled deep traps, and Li- and Na-filled deep traps, $Al\text{-}Li^+$, $Al\text{-}OH^-$ and Al-hole centers. Figure 3 displays the results for both 'High-H' and 'Low-H' samples only for their Al-related centerss. As yet no direct method of measuring the $Al\text{-}Li^+$ centers exists as they do not exhibit any spectroscopic signal. On the other hand, $Al\text{-}OH^-$ and the Al-hole centers are independently observable by near-infrared and ESR absorption methods respectively. Also, the acoustic loss peaks at 23 K [12] and that around 600 K can also be monitored to measure the concentration of Al-hole and the $Al\text{-}OH^-$ centers respectively [13].

(B) Cultured Quartz:

With the above background of radiation-induced mobility of alkali ions and protons in quartz, the combined strength of 3306 and 3367 cm^{-1} bands for the unswept quartz obtained after irradiation at 300 K would directly give the strength of $Al\text{-}OH^-$ for a specific radiation dose. The difference between the concentration of H-swept $Al\text{-}OH^-$ and the unswept $Al\text{-}OH^-$ will straightaway give the measure of the $Al\text{-}Li^+$ centers.H-sweeping replaces all the Li^+ trapped at aluminum sites with protons and in the process forms the $Al\text{-}OH^-$ centers. Therefore, the combined strength of 3306 and 3367 cm^{-1} bands would represent the Al concentration and in retrospect, the concentration of $Al\text{-}Li^+$ centers as well. In the unswept or alkali swept quartz crystals, the difference in the combined strength of both the $Al\text{-}OH^-$ bands in H-swept and that obtained upon irradiating the unswept crystal at 300 K would be a measure of Al-hole concentration produced after a specific radiation dose. This operation can be expressed as

$$[(Al\text{-}OH^-)_{H\text{-}swept} - (Al\text{-}OH^-)_{unswept\ irrad.\ 300\ K} = Al\text{-hole}\ldots(1)$$

2.2 Estimation of Aluminum in Natural Quartz

Figures 4 shows a bar diagrams related only to $Al\text{-}OH^-$ point defects in two natural quartz samples designated as H_2 and L_2 respectively in their as-received conditions and after H-sweeping. To be certain that the sweeping is complete may be confirmed by the fact, as also displayed in these Figures, that the irradiation at 300 K in H-swept condition did not lead to any increase in the strength of the $Al\text{-}OH^-$ bands. This would be due to temporary migration from aluminum centers during irradiation followed by the back diffusion of protons to their precursors after the irradiation is over. During the course of irradiation, the $Al\text{-}OH^-$ centers would break to form Al-hole centers. The breaking of $Al\text{-}OH^-$ would be due to the dissociation of protons from aluminum centers when the energy of recombination of free electrons and holes is transferred to the $Al\text{-}OH^-$ centers in the case of this recombination occurring near the aluminum centers in

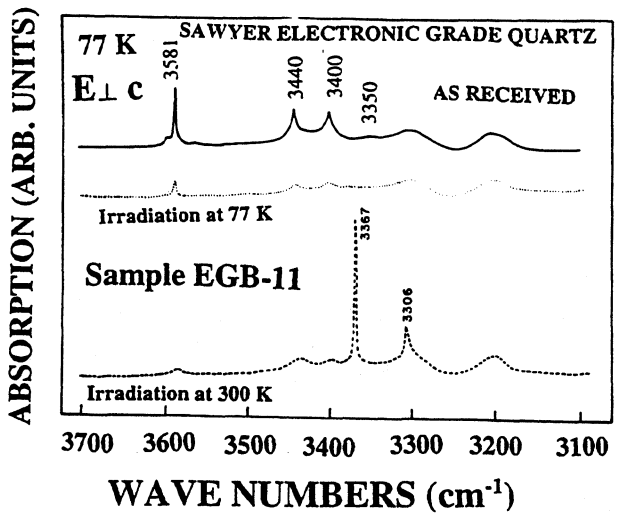


Fig.1 Near infrared absorption spectra of a cultured quartz crystal (Sawyer Electronic Grade, Sample EGB-11) measured at 77 K in its as-received, irradiated at 77 K and irradiation at 300 K conditions.

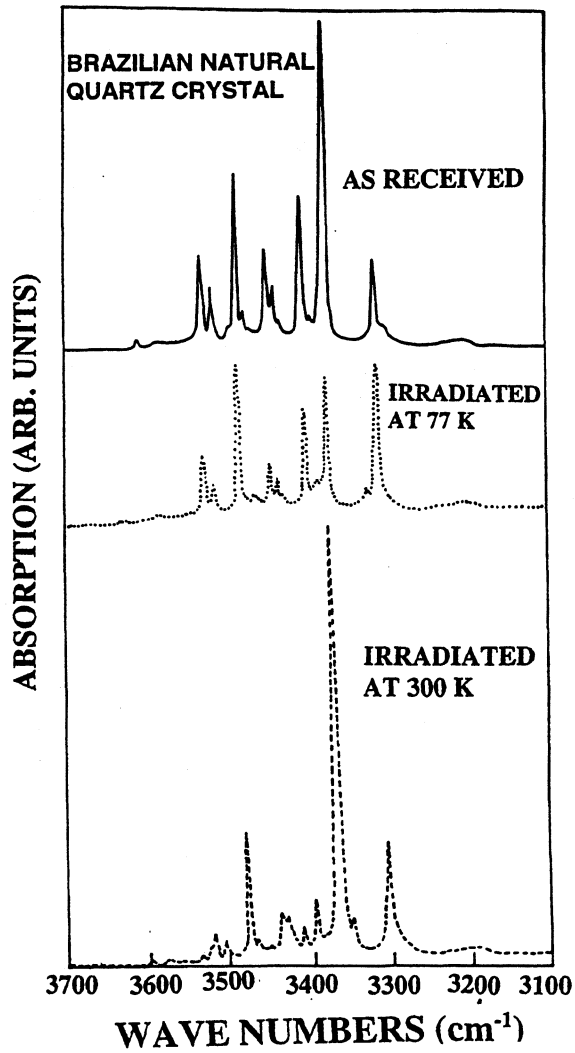


Fig.2 Near infrared absorption spectra of a natural Brazilian quartz crystal measured at 77 K in its as-received, irradiated at 77 K and irradiation at 300 K conditions.

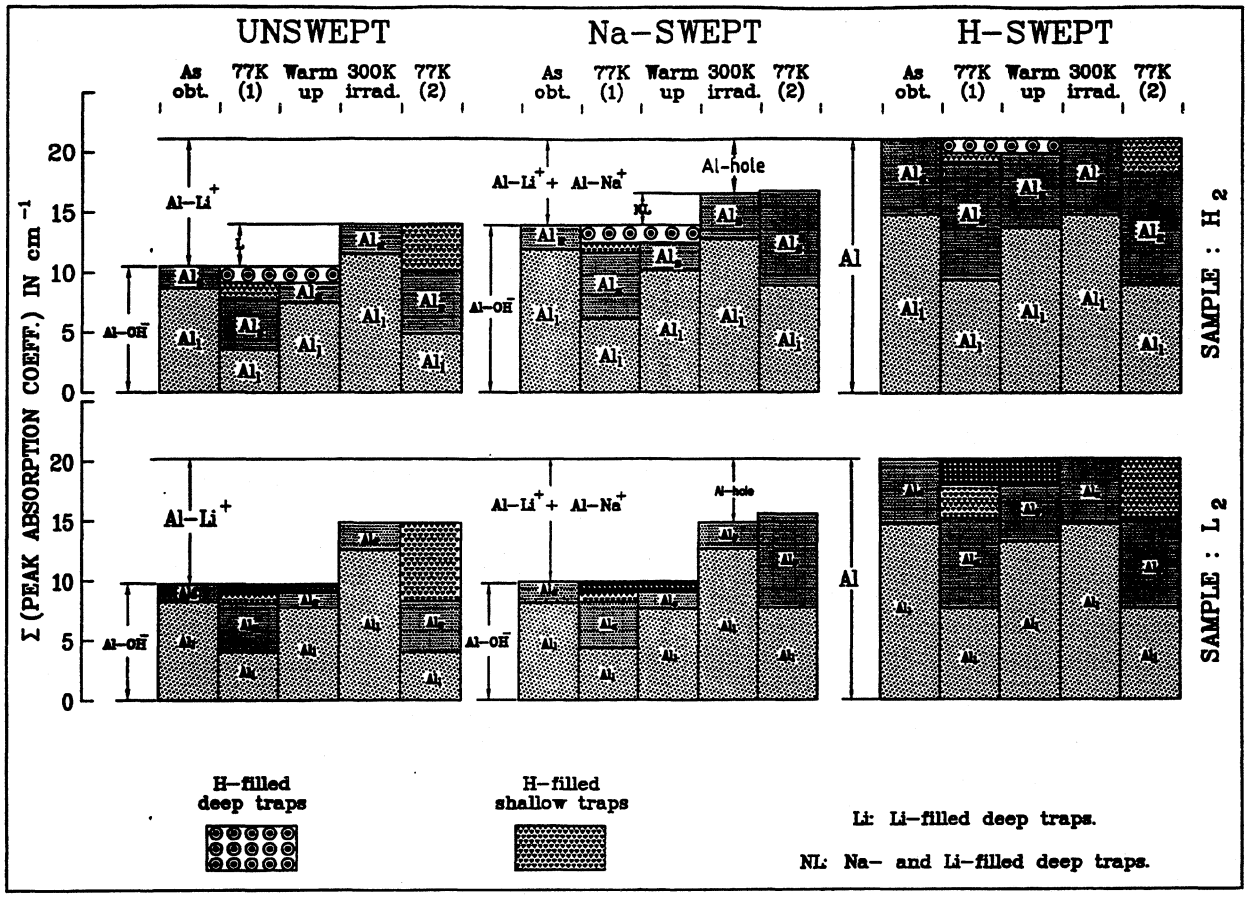


Figure 3: Bar diagrams showing the radiation effects on aluminum-related point defects in both 'High-H'(Sample designated as H_2) and 'Low-H"(Sample designated as L_2) samples of natural quartz crystals. 'As obt' represents as-received unswept or obtained after Na and H-sweeping; '77 K (1)' depicts the results after first irradiation at 77 K; '300 K irrad.' stands for after irradiating the sample at 300 K and '77 K(2)' is the stage of second irradiation at 77 K.

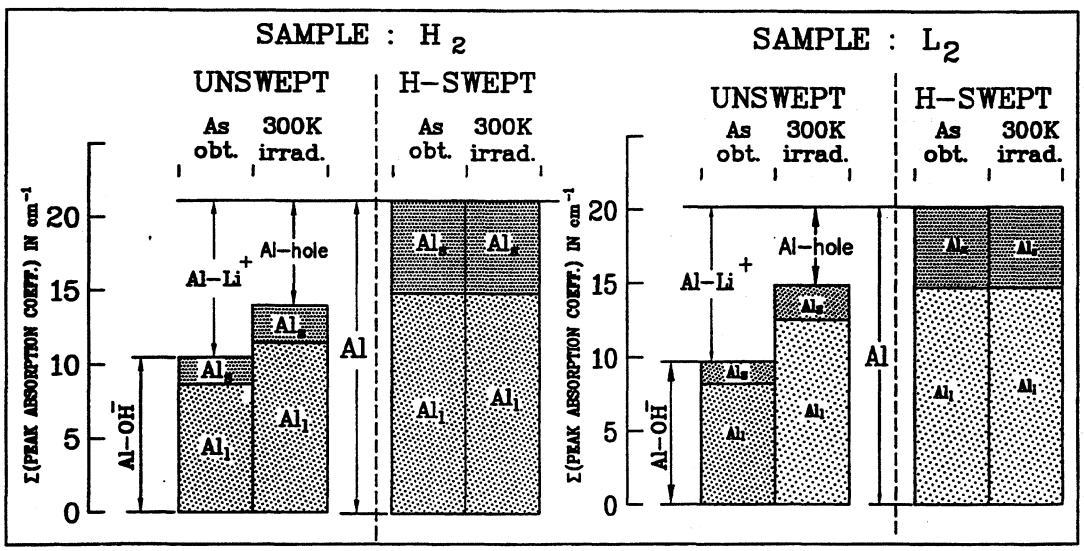


Figure 4. Bar diagrams showing the measurements of Al-OH^- and Al-hole centers and the quantitative estimation of aluminum in samples of High-H and Low-H unswept and H-swept natural quartz crystals. While the crystals have nearly the same aluminum concentration, the High-H crystal shows lesser Al-OH^- centers than the Low-H crystal and vice-versa Al-hole centers.

quartz lattice. As soon as the irradiation is over, the protons, due to their high affinity with aluminum, diffuse back to aluminum centers. Martin [8] has quantitatively estimated the aluminum concentration in ppm as given by

$$C_{Al}\ (\text{ppm}) = 3.3\ \alpha_{3367} \quad \text{.....(2)}$$

where α_{3367} is the absorption coefficient of the 3367 cm^{-1} Al-OH^- band in H-swept condition measured at 90 K. Using Martin's formulation, the aluminum concentration in the high-H, Brazilian natural quartz crystals was found to be 49.2 ppm and 48.5 ppm for low-H, Arkansas crystal. Thus, the two samples were found to have nearly same Al-concentration. The concentration of as-grown Al-OH^- centers in these crystals for high-H sample was 28.7 ppm and 27.4 ppm for low-H crystal. This shows that about 60% of total Al in natural quartz is charge compensated by protons and rest would be thus by alkalis; mostly Li^+.

Uniform irradiation at 300 K (dose = 2 Mrad) increased the concentration of Al-OH^- centers. This would be due to break away of Al-alkali centers and their subsequent conversion into a mixture of Al-OH^- and Al-hole centers. At such a high dose, Al-related defects will saturate. Thus, the difference in H-swept Al-OH^- strength and that of its unswept 300 K-irradiated state would exhibit the concentration of Al-hole centers and that in the unswept as-received state would correspond to the concentration of Al-Li^+ centers as in that case most of the aluminum will be charge compensated with Li+ [4].

It would be expected that high-H crystal upon irradiation would show up higher concentration of Al-OH^- due to abundance of hydrogen in it compared to the low-H quartz. Interestingly, in contrast, the high-H crystal showed a larger number of Al-hole centers (11.22 ppm) than the low-H one (7.11 ppm). Similar observations were noticed for other sets of samples also from these stones. Contemplated ESR measurements on these crystals will confirm a clear picture of our present results and calculations. At the present stage of work, we expect this due to jamming in the kinetics of large number of protons in high-H crystals for steric reasons which prevents them to reach Al-sites after irradiation. This view point is supported by the fact that hydrogen shows a much lower mobility than alkalis in the quartz lattice during sweeping [2,7]. On the other hand, irradiation - induced holes, due to their higher mobility, get trapped at Al-sites in the high-H quartz. In the case of low-H crystals the situation is normal due to nonsteric hindrances.

The present studies would find application in appropriate selection of natural quartz for precision electronic devices. The investigations suggest that a low-H and low Al natural quartz which would thus have a high mechanical Q should be the preferred material for many critical aerospace applications.

3. ACKNOWLEDGMENTS

The author thanks Professor J.J. Martin for supporting the work at Oklahoma State University, under contracts with Sandia National Laboratories and Rome Air Development Command, Hanscom, US Air Force. A sense of deep gratitude is expressed towards Dr. J.C. King of Sandia National Laboratory (now retired), Dr. F.L. Walls of NIST, USA and Professor R.J. Besson of ENSMM, France for their interest, appreciation and encouragement in the work.

4. REFERENCES

1. Harish Bahadur "Irradiation effects on Ge-H, Al-H and other H-related point defects in cultured quartz crystals", IEEE Trans. Nucl. Sci. Vol. 43, pp.2085-2095 (1996).

2 J.J. Martin, "Radiation-induced frequency offsets and acoustic loss in AT-cut quartz crystals", J. Appl. Phys. , Vol. 68, pp. 5095-5104 (1990).

3. H. Bahadur, "Infrared characterization of natural and cultured quartz: The effect of electrodiffusion and irradiation", J. Appl. Phys. Vol.66, pp. 4973-4982 (1989).

4. J.J. Martin, "Aluminum-related acoustic loss in AT-cut quartz crystals" J. Appl. Phys. Vol. 56, pp 2536-2540, 1984.

5. W.A. Sibley, J.J. Martin, M.C. Wintersgill and J.D. Brown, "The effect of radiation on the OH^- infrared absorption of quartz crystals" J. Appl. Phys.Vol. 50, pp. 5449-5452 (1979).

6. J.J. Martin, Ho B. Hwang, H. Bahadur and G.A. Berman, "Room temperature acoustic loss peaks in quartz", J. Appl. Phys. Vol. 65, pp. 4666-4671 (1989).

7. L.E. Halliburton, M.E. Markes and J.J. Martin,"Point defects in synthetic quartz: A survey of spectroscopic results", Proc. Proc. Annual Frequency Control Symp. Vol. 34, pp. 1-8, (1980).

8. J.J. Martin, "Estimation of aluminum and growth-defect content in cultured quartz by using infrared absorption", Proc. 1996 IEEE International Frequency Control Symposium, pp. 126-130 (1996).

9. A. Kats "Hydrogen in alpha quartz", Philips Res. Rept, Vol. 17, pp.133-195 (1962).

10. J.C. King, "Electrolysis of synthetic quartz:Effect upon resonator performance" Proc. IEE (London) Vol. , B109, Suppl. 22, pp. 295-301 (1962).

11. Harish Bahadur, "Low-temperature irradiation effects in natural crystalline quartz ", J. Appl. Phys. Vol.75, pp. 1420-1425. (1994).

12. J.J. Martin, Ho B. Hwang and H. Bahadur, "Radiation effects in the acoustic loss spectra of AT-cut quartz crystals", Proc. Annual Frequency Control Symposium, pp. 266-269 (1985).

13. J.J. Martin,"Acoustic loss in cultured quartz"Proc. 1996 IEEE International Frequency Control Symposium, pp.170-178 (1996).

MATHEMATICAL MODEL FOR COMPUTER SIMULATION OF QUARTZ HYDROTHERMAL GROWTH

M.A.Arkhipov

Scientific Production Company Goodwill, 1, Institutskaya St., Alexandrov, Vladimir Region, 601600 Russia

ABSTRACT

Mathematical model for heat-impulse-mass transfer in vessel is proposed. Theory of similarity is applied to transfer in vessel. Kinetic coefficient of stone growth have no significant dependence on geometry of growth zone but depends on solution physical properties and temperature difference between neighbor stone's faces. The phenomena of seed location influence on quartz growth rate are discussed. Optimum location of baffle plate in respect to isotherms is estimated.

1.MATHEMATICAL MODEL

Schematic drawing of vessel is represented on FIG.1. There are three zones of different thermal convection profile in vessel: zones A,B contains large scale whirlwind liquid movement (normally in opposite directions one in respect to other); zone C- zone of layered movement of liquid between stones along its surfaces. It is existed rising and falling flows along surfaces in one space between two stones.

As soon as quartz growth rate (rate of changing for internal geometry of vessel) is considerably smaller than rate of heat-impulse-mass transfer - the last may be considered as steady. It means that on different stages of hydrothermal cycle different distributions of concentration, velocity, temperature are performed but it depends only on steady boundary conditions (boundary conditions are constant for each stage of process).

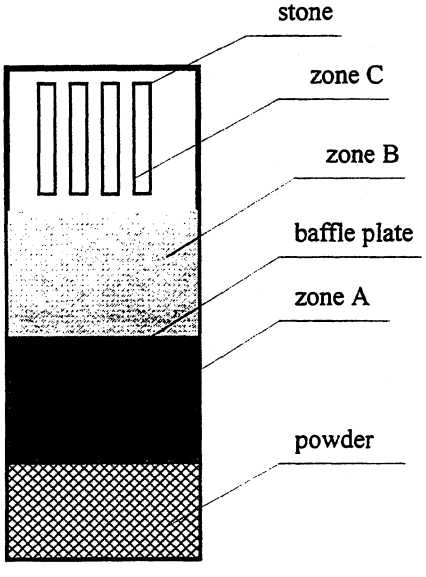


FIG.1. Thermal convection zones in vessel

Vessel for hydrothermal growth has cylindrical form and ordinary deposition of heaters on walls of autoclaves also has cylindrical symmetry. According to Curie principle: movement of solution inside vessel have to satisfy cylinder elements of symmetry. Here influence of seeds and holes location in baffle plate are neglected. Distributions of concentration, temperature and velocity are the functions only on height of considered point in vessel and its radius.

Under written above assumptions mathematical model of thermal convection may be written as follows:

$$\partial(\rho w_z)/\partial z + 1/r\ \partial(r\rho w_r)/\partial r = 0, \quad (1)$$

$$w_z \partial w_z/\partial z + w_r \partial w_z/\partial r = g\Delta\rho/\rho + \nu \nabla^2 w_z,$$

$$w_z \partial w_r/\partial z + w_r \partial w_r/\partial r = \nu \nabla^2 w_r, \quad (2)$$

$$w_z \partial T/\partial z + w_r \partial T/\partial r = a \nabla^2 T, \quad (3)$$

$$w_z \partial c/\partial z + w_r \partial c/\partial r = D \nabla^2 c, \quad (4)$$

where ρ - density, w- velocity , r- radius of considered point, z-height of considered point, g- gravitational acceleration, $\Delta\rho$ - difference between local density and average density in surrounding volume, ν - kinematic viscosity, T - temperature, a- thermal diffusivity, c- concentration of SiO_2 in solution, D- diffusion coefficient. Equation (1) is called continuity equation, equations (2) - Navier-Stokes equations for impulse transfer, equations (3) and (4) describe heat and mass transfer, respectively. We restrict ourselves here that only transport of SiO_2 exist. We suppose here that uniform distribution of other components concentrations is taken place.

Boundary conditions are as follows: distribution of temperature is given, components of velocity is equal to zero, gradient of concentration also is equal to zero on the internal walls of vessel and on the baffle plate. Concentration have to be equal to its equilibrium value on the surfaces of stones and powder.

Hydrothermal growth rate V of stones may be defined by following equation [1]:

$$V = \beta(c - c_s(P,T)), \quad (5)$$

β- kinetic coefficient of stone growth, $c_s(P,T)$ - equilibrium concentration under present pressure P and

0-7803-5400-1/99/ $ 10.00

temperature T. Practical question is about factors that have influence on β .

2.GRASHOF NUMBER AND KINETIC COEFFICIENT

Uniformity of stone's growth rate in different parts of vessels is strongly required for increasing production output of wafers from stone. It is clear that it is impossible to create the same growth conditions in different parts of vessel in the case of applying uniform location of seeds because of different thermal convection conditions. Kinetic coefficient is also very important factor for transfering of results obtained on small vessels into large mass production vessels.

According to Third Theorem of Similarity [2] steady thermal convection have to be defined only by numbers that contain parameters from boundary conditions and physical properties. In the case of free thermal convection that take place in vessel conditions - of similarity are defined by Grashof number:

$$Gr=g L^3 \alpha_V \Delta T/\nu^2, \quad (6)$$

where L- characteristic dimension of convection area, α_V- expansion coefficient of volume. Grashof number characterize relation of Archimedes' force in respect to viscosity force. In the case of increasing of Gr value intensity of thermal convection also have to be increased.

It is very interest to note that for different zones in vessel parameters L, Δ T have to be chosen differently. Thus, for zones A,B (FIG.1) this parameters are attached to the points in the upper and lower points of this zones because large distance between internal walls of vessel corresponds to well known task of free thermal convection along wall that have different temperature in respect to one of liquid. In zone C task is similar to problem of free thermal convection between two vertical walls that have different temperatures and corresponding points lie on the opposite walls of stones.

It goes without saying from equation (6) that it is impossible to receive the same condition of heat-mass-impulse transfer for small and large vessel because Gr increase as L^3 for large vessel. But larger Gr value mean more active transport of SiO_2 and that's why it create better conditions for growth.

As soon as equations (3) and (4) have similar form we may apply experimental results [2] obtained for Nu number (heat exchange) for receiving formula for β :

$$\beta L/D=K\, Gr^{0.3}\, (h/L)^{-0.1}, \quad (7)$$

where K- constant, h- height of area of thermal convection in zone C. From (6),(7) it is easy to receive that kinetic coefficient of growth not depends significantly on dimensions and may be changed very slightly (as $\Delta T^{0.3}$) by varying temperature gradient. It is necessary to note that here ΔT is temperature difference between neighbor stone's faces.

3.DISTRIBUTION OF TEMPERATURE AND TRANSPORT

Distribution of temperature, existed in vessel, activate free thermal convection in vessel. Thermal convection has influence on distribution of temperature. Experimental results [2] for isotherms from internal part of vessel are represented on FIG.2.

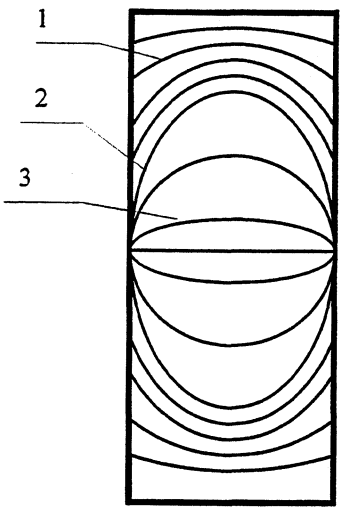


FIG.2. Isotherms in vessel

Origin of distribution may be understood, considering great difference of thermal diffusivity of metal walls and average thermal diffusivity for stones and solution. In the area of isotherm #1 and #3 there are no significant gradients of temperature in radial direction in the central area of vessel. Intermediary area (isotherm #2) are characterized by large temperature gradients not only near walls but also in the center.

As soon as free thermal convection in zone C is due to temperature gradients in radial direction it is clear that seeds have to be located in the area of isotherms similar to #2. In any case temperature gradients are varying in radial direction. It may cause different growth rates because of different Gr number values for each layer between stones in zone C. That's why it is necessary to find correspondence between seeds location and temperature distribution that may provide equal Gr number values for all layers between stones.

Transport of SiO_2 from dissolution zone to crystal growth zone is carried out by two processes: convection and diffusion. It is known that convection transport in vessel is considerably greater than ones for diffusion. That's why it is necessary to decrease length of diffusion transport as much as possible. Diffusion mechanism is

taken place on the borders between two whirlwinds. Thus it would be very helpful to increase intensities of whirlwinds, to enlarge its sizes and to decrease quantity of whirlwinds in zones A and B because each whirlwind's border lead to diffusion mechanism.

Position of baffle plate in respect to isotherms have strongly influence on the character of whirlwind liquid movement. Let's consider case when horizontal isotherms (zero temperature gradients) not corresponds to baffle plate - separate location (see FIG.3, where baffle plate is located higher in respect to isotherm). Direction of movement depends on temperature of wall (Tw) and temperature of liquid (Tl). In this case three whirlwinds are taken place if distance between baffle plate would be considerable. Decreasing of the distance may lead to disappearance of intermediary whirlwind because of stresses in liquid and the worst case may happened - large diffusion zone. When the distance would be small - the dimensions of diffusion zone would be decreased and one may receive only two closely located whirlwinds. Thus, the best case is when isotherm with zero radial temperature gradients would corresponds to baffle plate

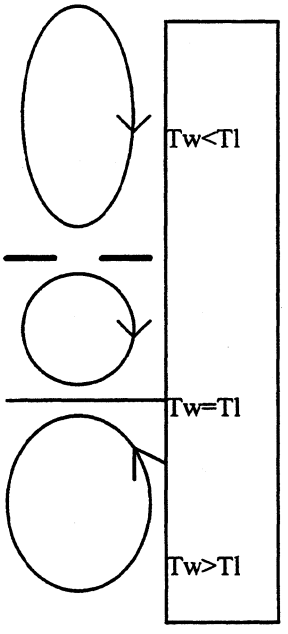


FIG.3. Whirlwinds in vessel near walls.

4.CONCLUSION

Kinetic coefficient of quartz growth rate is a function of Grashof number. There are no strongly dependence of kinetic coefficient on dimensions and location of stones in vessel and it may be changed slightly by varying temperature difference between neighbor stones. Grashof number define intensity of growth through its influence on free thermal convection transport.

For receiving uniform thickness of stones one have provide equal Grashof number values in vertical liquid layers between stones.

Positional non-correspondence in location of baffle plate and zero radial gradient isotherm lead to decreasing in transport of SiO_2 from dissolution zone to crystal growth zone.

5.REFERENCES

[1] A.A. Chernov and E.I. Givargizov, Crystal Growth, Moscow: Nauka, 1980, ch.7.5, pp.321-322.

[2] S.I. Isaev and I.A. Koginov, Theory of Heat and Mass Exchange, Moscow: Vishaia Schola, 1979, ch.VII.2, pp. 345-347.

ISOCHRONISM DEFECT FOR VARIOUS DOUBLY ROTATED CUT QUARTZ RESONATORS

N. Gufflet, R. Bourquin, J.J. Boy

ENSMM / LCEP – 26 Ch. de l'Epitaphe – 25030 BESANÇON Cedex – FRANCE
Fax : (33) 3 81 88 57 14 – E-mail : ngufflet@ens2m.fr

ABSTRACT

It has been shown in earlier work, that the amplitude-frequency effect (also called isochronism defect or anisochronism) could be a limitation factor on ultra-stable oscillators.

Theoretical studies based on the non-linear theory of the piezoelectricity have already been developed to explain the amplitude-frequency effect. So it is possible to estimate the dependence of the isochronism defect versus design parameters of resonators (radius of curvature, electrodes diameter, overtone rank,...). However, due to the unknown of the fourth order elastic coefficients, it is not possible to predict isochronism defect of any resonant frequency of a given energy trapped resonator.

To tentatively find the orientation of a plate which does not exhibit an isochronism defect, we have realised electroded resonators with different orientations and curvatures.

We present results which verify, particularly, $R^{-1/2}$ dependence of amplitude-frequency effect versus radius of curvature. Moreover, we show that the isochronism defect can be positive or negative and can vary from one orientation to other one of about two orders of magnitude.

1. INTRODUCTION

A previous paper [1] has shown that amplitude-frequency effect may have significant influence on the short-term stability of oscillators due to the phase noise caused by amplitude noise. Moreover, the amplitude frequency effect can also affect the long-term stability through the variation of the driving level on the resonator.

Several theoretical studies of the amplitude-frequency effect have been performed, showing that the isochronism defect is due to non-linearities, especially elastic one. But at our knowledge only two ones take the energy trapping into account.

In the first one [2], D. A. Nassour uses a perturbation method to determine the current through the resonator. A determination of the amplitude-frequency effect is obtained for singly and doubly rotated cuts and in both cases, non-linear elastic, piezoelectric and dielectric coefficients are included. In the second one [3], H. F. Tiersten uses an iterative approach, and determines the isochronism defect for a general doubly rotated cut. Just the non-linear elastic terms up to fourth order are included.

All of these studies lead to similar relation, which can be written :

$$\frac{\Delta F}{F} = \frac{\sqrt{2}\rho\alpha P}{\pi^2 h^{1/2} n (M_n P_n)^{1/4} \eta^{(1)} \bar{c}^{(1)3/2} R_c^{1/2}} \tag{1}$$

where n, h, ρ, R_c, $\eta^{(1)}$, $\bar{C}^{(1)}$ and P are respectively the overtone rank, the half-thickness of the resonator, the density, the radius of curvature, the effective viscoelastic coefficient of the first mode, the effective elastic coefficient of the first mode and the power applied to the resonator. M_n and P_n are the dispersion constants [4] which govern the energy trapping and α is a multiplication factor which differs slightly in the different theoretical investigations. Nevertheless α is a function of the elastic non-linear terms up to fourth order, but in the reference [2], piezoelectric and dielectric non-linear terms are included. However, these terms are not well known, moreover their influences on α seem to be negligible. In consequence α can be expressed as :

$$\alpha = \frac{9}{64}\gamma - \frac{\bar{c}^{(1)}}{4}\left(\beta_2^2 \nu_2 + \beta_3^2 \nu_3\right) - \frac{53}{512}\frac{\beta_1^2}{\bar{c}^{(1)}} + \frac{h\bar{c}^{(1)}}{n^2\pi^2}\left(\beta_2^2 \zeta_{n2} + \beta_3^2 \zeta_{n3}\right) \tag{2}$$

where :

$$\gamma = \frac{1}{2} C_{2\,2222} + Q_{1M} Q_{1R} C_{3\,2MR222} + \frac{1}{6} Q_{1M} Q_{1R} Q_{1N} Q_{1K} C_{4\,2MR2N2K2} \tag{3.1}$$

$$\beta_1 = \frac{3}{2} Q_{1M} C_{2\,2M22} + \frac{1}{2} Q_{1M} Q_{1K} Q_{1R} C_{3\,2MK2R2} \tag{3.2}$$

$$\beta_2 = \frac{1}{2} Q_{2M} \left(C_{2\,2M22} + \frac{1}{2} Q_{1K} Q_{1R} C_{3\,2MK2R2}\right) \tag{3.3}$$

$$\beta_3 = \frac{1}{2} Q_{3M} \left(C_{2\,2M22} + \frac{1}{2} Q_{1K} Q_{1R} C_{3\,2MK2R2}\right) \tag{3.4}$$

$$\nu_i = \frac{7\bar{c}^{(i)} - 6\bar{c}^{(1)}}{8\bar{c}^{(1)}\bar{c}^{(i)}\left(\bar{c}^{(i)} - \bar{c}^{(1)}\right)} \tag{3.5}$$

$$\zeta_{ni} = \frac{(-1)^{(n-1)/2}\eta_n}{8\bar{c}^{(1)}\left(\bar{c}^{(1)} - \bar{c}^{(i)}\right)} \times \left(\frac{\eta_n^{(i)+2}\cos\eta_n^{(i)+}h}{\left(\eta_n^2 - \eta_n^{(i)+2}\right)\cos\eta_{n0}^{(i)}h} + \frac{\eta_n^{(i)-2}\cos\eta_n^{(i)-}h}{\left(\eta_n^2 - \eta_n^{(i)-2}\right)\cos\eta_{n0}^{(i)}h}\right) \tag{3.6}$$

and $\eta_n = \frac{n\pi}{2h}$, $\eta_{n0}^{(i)} = 2\sqrt{\frac{\bar{c}^{(1)}}{\bar{c}^{(i)}}}\eta_n$, $\eta_n^{(i)\pm} = \eta_{n0}^{(i)} \pm \eta_n$, (3.7)

i=2,3

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

$\bar{c}^{(i)}$ is the elastic constant corresponding to the i[th] mode velocity, the C_2, C_3, C_4 are the elastic coefficients of second, third and fourth order in doubly rotated frame and Q is a 3x3 matrix whose elements are $Q_{ir} = q_r^{(i)}$, and $q_r^{(i)}$ is the r[th] component of the i[th] normalised eigenvector.

It should be noticed that in obtaining the relation (1) from the expression of the amplitude-frequency effect of the reference [3] (eq. 100), which is expressed as a function of current square, we have used the expression of the motional resistance given by the reference [2] :

$$R_m = \frac{1}{16} \frac{n^2 \pi h \eta_e \sqrt{\alpha_n \beta_n}}{e_{221}^2 erf\left(l\sqrt{\frac{\alpha_n}{2}}\right)^2 erf\left(l\sqrt{\frac{\beta_n}{2}}\right)^2} \quad (4)$$

where :

$e_{221} = Q_{1R} \hat{e}_{22R}$ and $\hat{e}_{22R}$ is the second order piezoelectric coefficient in the doubly rotated frame, l is the width of the square electrode, erf is error function and α_n and β_n come from reference [3] :

$$\alpha_n^2 = \frac{n^2 \pi^2 \bar{c}^{(1)}}{8 R_c h^3 M_n}$$

$$\text{and } \beta_n^2 = \frac{n^2 \pi^2 \bar{c}^{(1)}}{8 R_c h^3 P_n} \quad (5)$$

The value of the non-null second and third elastic coefficients can be found easily in the literature. To determine the null coefficients and the relations between the others, we write that all the crystal properties present the same symmetry than the geometric one. For the fourth order coefficients, computer calculations have been performed and are in accordance with reference [5]. We found 57 fourth order null coefficients, 28 independent and 41 relations.

Due to the unknowledge of the fourth order coefficients values, it is not possible to predict the isochronism defect. But we are able to estimate dependency of the amplitude-frequency effect versus design parameters as, for example, the radius of curvature. In the expression of α, all the coefficients, except γ, are calculable. So, we can characterise each cut by its γ coefficient (eq. 3.1).

2. MEASUREMENTS METHOD

It can be shown that the resonance frequency is very closed to the frequency for which the phase between the voltage and the current across the resonator is null. So we measure the frequency of zero phase at different driving level and we notice that this measurement does not present difficulty on thermally compensated modes.

We use the following bench, where the "PI" circuits are in accordance with IEC 444 norm :

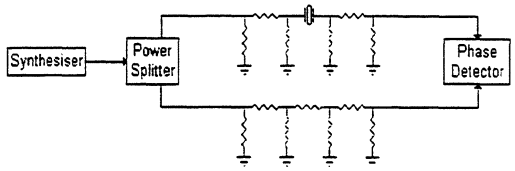


Figure 1 : Experimental set-up

However, the measurements of amplitude-frequency effect on non-thermally compensated mode are relatively difficult. Indeed, increasing the drive level will increase the temperature of the resonator by virtue of the heat dissipated by Joule effect moreover the thermal variations of the middle perturb the resonance frequency.

Measurements are possible by phase locking the synthesiser on the resonator close to zero and by recording the frequency at each step of driving level.

Here is an example of measurement of the B-mode (third overtone) of the SC-cut :

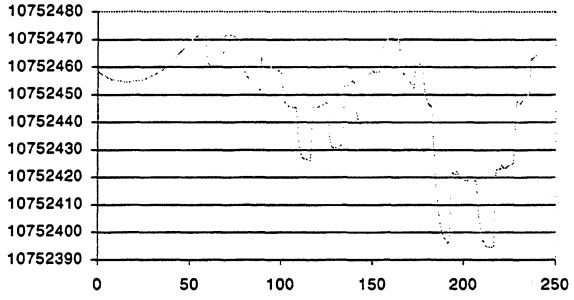


Figure 2 : Resonance frequency (Hz) vs. Time (s)

Performing an average on the steps, we can draw the amplitude-frequency effect :

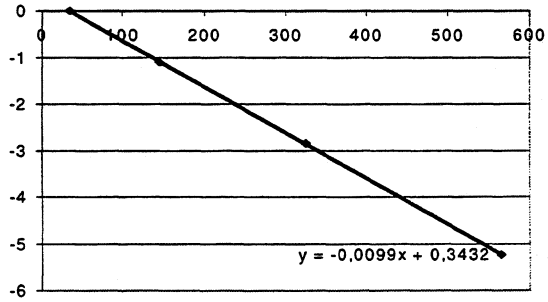


Figure 3 : Relative Frequency (10^{-6}) vs. Power (μW)

We find $\Delta F/F = -10.10^{-9}/\mu W$ and then $\gamma = -20.10^{11}$ N/m².

It should be noticed that the measurement is less precise than on thermally compensated cuts and it is not always possible to determine the isochronism defect on modes presenting high temperature dependence or (and) low amplitude frequency effect.

3. CHOICE OF RESONATORS

The interest of measuring amplitude-frequency effect on non-thermally compensated cuts is to reach the value of the fourth order coefficients. But, in this paper, we limit our measurements to thermally compensated cuts (with $\theta \cong 34°$).

It is admitted that the temperature-frequency behaviour can be fitted by the following third degree equation :

$$\frac{\Delta F}{F} = a(T-T_0)+b(T-T_0)^2+c(T-T_0)^3$$

where T_0 is the reference temperature (usually 25°C). A vibrating mode is thermally compensated when a=0 (or close to zero). On quartz resonators, A-mode can not be quasi-independent of the temperature, but B- and C-modes can be. And then, the locii a=0 can be calculated [6-7] as indicated in Figure 2.

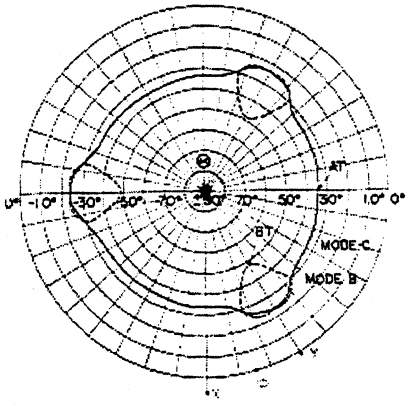


Figure 4 : Locii of null first order temperature coefficient of B- and C-modes (from [6])

Another important parameter is the choice of the radius of curvature. Indeed, to prevent edge effect we have to trap correctly the vibration in the centre of the resonator. Theory of energy trapping has already been performed [4], and the calculation can be summarised in the determination of M_n and P_n which govern with the radius of curvature the profile of the vibration. M_n and P_n depending on the rank overtone, the energy trapping can not be good for each fundamental, third, fifth... overtone rank. For example, the calculation of M_n and P_n on the C-mode of SC-cut leads :

Table 1 : M_n and P_n for the 1st, 3rd, 5th and 7th overtones of the C-mode of SC-cut

	n=1	n=3	n=5	n=7
M_n (10^9N/m)	96.51	52.24	70.34	56.60
P_n(10^9N/m)	56.41	63.08	78.20	110.46

Which means that to have an amplitude vibration on the periphery of the resonator of about 10^{-6} times of the amplitude vibration in the centre of the resonator (i.e. a good energy trapping) we have to use the following radius of curvature :

Table 2 : Radius of curvature for M_n and P_n

	n=1	n=3	n=5	n=7
R_c for M_n (mm)	21	350	710	1740
R_c for P_n (mm)	35	290	640	890

So, a 250mm radius of curvature, for example, can trap correctly third and fifth overtones but not the fundamental nor the seventh overtone (which becomes too much trapped).

4. EXPERIMENTAL RESULTS

4.1 Isochronism defect vs. radius of curvature

Measurements of the isochronism defect have been performed on several SC-cut quartz resonators having different radii of curvature. Table 3 shows the construction parameters of these resonators.

Table 3 : SC-cut resonators parameters

Res.	ϕ Quartz (mm)	ϕ Electrodes (mm)	R_c (mm)
1a	13	6	150
2a	13	6	170
3a	13	6	190
4a	13	6	230
5a	13	6	370

Resonance frequency, isochronism defect and resistance have been measured for the 3rd and the 5th overtone, moreover we have calculated the γ. Results are presented in Table 4.

Table 4 : Anisochronism for SC-cut resonators

Res.	n	F (Hz)	R (Ω)	$\Delta F/F$ ($10^{-9}/\mu W$)	γ (10^{11}N/m²)
1a	3	9852800	92	1.53	5.03
2a	3	9851800	86	1.42	4.99
3a	3	9852000	82	1.35	5.01
4a	3	9849600	74	1.22	4.99
5a	3	9853700	60	0.94	4.91
1a	5	16377200	395	1.18	5.97
2a	5	16378700	376	1.08	5.84
3a	5	16381200	340	1.01	5.79
4a	5	16381400	300	0.93	5.85
5a	5	16396200	240	0.72	5.76

Then we can draw the isochronism defect versus radius of curvature.

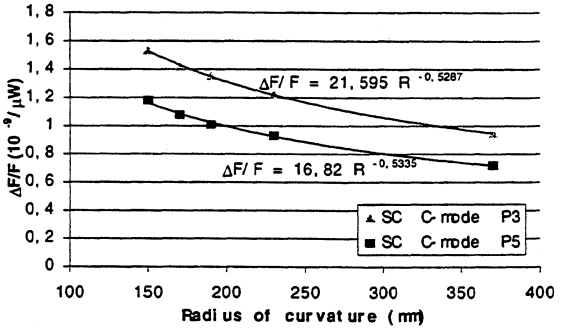


Figure 5 : Isochronism defect vs radius of curvature

4.2 Isochronism defect vs. phi

To tentatively find a cut exhibiting a low amplitude-frequency effect, we have manufactured several resonators, thermally compensated on C-mode, for various phi angles. The resonators parameters are the following :

Table 5 : Doubly rotated resonators parameters

Res.	φ (°)	θ (°)	ϕ Electrodes (mm)	R_c (mm)
1b	22.75	33.97	6	230
2b	24.3	33.8	7	230
3b	25.15	33.67	7	230
4b	26.28	33.55	7	230
5b	27.33	34.0	7	300
6b	28	34.1	7	300
7b	30	34.48	7	700

Here are the measurement's results :

Table 6 : Isochronism defect for doubly rotated cut

Res.	n	F (Hz)	a (10^{-6})	R (Ω)	$\Delta F/F$ ($10^{-9}/\mu W$)	γ (10^{11} N/m²)
1b	3	9849000	1.0	71	1.2	4.99
2b	3	9837500	1.3	89	0.97	4.39
3b	3	9865000	1.2	99	0.69	3.63
4b	3	9961300	1.7	140	0.35	2.66
5b	3	9972900	0.9	130	-0.11	1.28
6b	3	9974800	0.7	128	-0.32	0.55
7b	3	10338300	1.9	148	-0.82	-2.71

The amplitude-frequency effect and the γ coefficient can be plotted versus φ for thermally compensated cut :

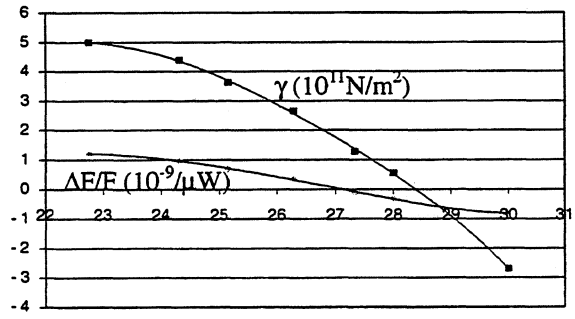


Figure 6 : $\Delta F/F$ and γ coefficient versus φ(°)

4.3 Isochronism defect vs. overtone rank

According to relation (1), the anisochronism effect exhibits a n^{-1} dependence, but it should be noticed that the M_n and P_n constants vary from one overtone to another one. Nevertheless, for cuts as SC-cut for which M_n and P_n are of the same order of magnitude for third and fifth overtones, the n^{-1} dependence can be observed. But we have observed that the isochronism defects of seventh and ninth overtones of an AT-cut change of sign in relation to third and fifth ones. This fact can not be explained by the theory of reference [2-3].

5. CONCLUSION

Our experiments confirm that the isochronism defect depend on radius of curvature by a factor $R^{-1/2}$.

The theoretical investigations give a relatively good description of the phenomenon for correctly trapped resonators, but seem unable to describe the behaviour of quartz resonators with vibrating energy too much confined.

Tested cuts exhibit an isochronism defect which has been decreased by a factor 10 in relation to SC-cut and we expect that adjusting the cut angle precisely can reduce the isochronism defect by a greater value. But we must point out that C-mode and B-mode are nearer than on SC-cut and problem in oscillators could appear due to the filtering process. So test in oscillators should be done.

ACKNOWLEDGEMENTS

Authors wish to thanks Eric Andrey, Séverine Michel, Daniel Thiebaud for resonators manufacturing and Gilles Robichon from BVA Industries for supplying quartz resonators.

REFERENCES

[1] R. J. Besson et al, "Phase noise limitation due to amplitude frequency-effects in state-of-the-art quartz resonators, Proc. of the IEEE AFCS, 1996, pp. 839-843

[2] R. Bourquin, D. Nassour, D. Hauden, "Amplitude Frequency effect of SC-cut quartz trapped energy resonator", Proc. of the 36th IEEE AFCS, Philadelphia (1981), pp. 200-207

[3] H. F. Tiersten, D. S. Stevens, "An analysis of non-linear resonance in contoured-quartz resonators", JASA., Vol. 80 (4), pp. 1122-1132, October 1986

[4] D. S. Stevens, H. F. Tiersten, "An analysis of doubly rotated quartz resonators utilizing essentially thickness modes with transverse variations", JASA, Vol. 79 (6), pp. 1811-1826, June 1986

[5] R. Brendel, "Independent fourth order elastic coefficients for all crystal classes", Acta Crystallographica, Vol. A-35, pp. 525-533, 1979

[6] R. Bechmann, A. D. Ballato, T. J. Lukaszek, "Higher-order Temperature Coefficients of the elastic stiffenesses and compliances in alpha-quartz", Proc. I.R.E., pp. 1812-1822, August 1962

[7] B. Dulmet, R. Bourquin, "Application of Lagrangian effective material constants to the study of the thermal behaviour of saw propagation in piezoelectric crystals", Proc. of the IEEE Ultrasonics Symposium, Cannes 1994, pp. 331-335

A Layerwise Plate Theory for the Vibrations of Electroded Crystal Plates

Ji Wang, Jiun-Der Yu, †*Yook-Kong Yong,* and ‡*Tsutomu Imai*

Epson Palo Alto Laboratory, 3145 Porter Drive, Suite 104, Palo Alto, CA 94304, USA. †Dept. Civil & Environmental Eng., Rutgers U., Piscataway, NJ 08855, USA. ‡EM Technology Lab., Seiko Epson Corp., 3-3-5 Owa, Suwa-shi, Nagano-ken 392, Japan

Abstract— **Electrodes on a crystal resonator has been traditionally considered as mass addition to the crystal plate, thus resulting the neglect of their stiffness. This assumption is considered reasonable if electrodes are thin in comparison with the crystal in terms of the mass ratio, the relative mass of electrodes. For thicker electrodes of high frequency resonators, this assumption has to be reexamined for better prediction on their effects on the resonance frequency and the frequency-temperature characteristics. In this study, the electrodes are considered as layers of plates with their own stiffness and mass properties. As a result, the deformation of electrodes also involve independent ones in addition to the crystal deformation. The theory employed also enables us to consider the deformation to a higher order of degree. The layerwise plate theory is derived and implemented for the finite element method. Numerical results are compared with the other computing schemes with and without electrodes considerations.**

I. Introduction

Mindlin plate theory has been widely used for the vibration analysis of quartz crystal plates from the thickness-shear vibration frequency [1], frequency-temperature relations [2,3], to the piezoelectrically forced vibrations [4]. Of course, in each of these aspects, modifications have been made to accommodate many special requirements arising from these specific problems, such as the temperature increment and the electrical boundary conditions on the electrodes. In case of the crystal resonators, the presence of the electrodes on the crystal faces is the primary consideration in the further modification of the plate theory.

From the beginning, it has been agreed and lately proved experimentally that for thin electrodes, only the mass effect of the electrodes needs to be take into account in the vibrations. This leads to the defining of the mass ratio, or the equivalent mass of the electrodes, in many studies started by Mindlin [5]. This concept has been adopted, refined, and expanded by many authors [6,7]. Lately, the consideration of the mass effect of the electrodes is generalized in a finite element implementation [8].

If the electrodes cannot be characterized as thin, we have difficulties in taking into consideration of the stiffness of the electrodes. In a study of thickness-shear and flexural vibrations of electroded crystal plates, Mindlin [5] considered the stiffness of these vibration modes by assuming that the displacements in the electrodes are the same with the crystal faces. This resulted in the equations which modified the elastic constants of the crystal plate in the similar manner for the mass terms. It should be regarded as an improvement in the consideration of the stiffness of the electrode stiffness, although there is a lack of experimental data to prove its effectiveness. Later, based on a similar assumption, Wang *et al.* [9] generalized this approach and implement it in a finite element procedure with the higher-order Mindlin plate theory.

As an attempt to consider the effect of the electrodes in a more precise manner, we start with assumption that there are deformations, or displacement components, in the electrodes which are independent of the crystal deformation. In other words, we consider the electrodes as layers of metal plate bonded on the faces of the crystal plate. In this case, although there are requirements for the displacements of electrodes to be continuous over the interfaces, we shall have the equation for the vibrations of the electrodes in addition to the crystal plate. By solving these equation together, we should be able to predict the vibration frequency and mode shapes in a more accurate manner, thus effectively improving the consideration of the electrodes for both the mass and stiffness properties. Treatment like this has been widely used in the study of composite plates [10], in which each layer has very close thickness but quite different properties. For the crystal resonator designs, in contrast, even in today, the electrodes are only about one or two percent of the thickness of the whole resonator structure. If such a technique will lead to the satisfactory solutions needed for resonator design is to be explored further.

II. Layerwise Plate Theory for Electrodes

The basic assumption in the layerwise theory for composite plate is that the displacements of each layer are independent but related to each other through the boundary conditions of the interfaces. For a resonator model showing in Fig. 1, there are only three layers, with the crystal plate with thickness $2b$ as the dominant one, to

0-7803-5400-1/99/ $ 10.00

be considered. It is natural for us to consider the displacement field as

$$u_j = \sum_{n=0} b^n u_j^{(n)} + \sum_{n=1} (x_2 - b)^n \bar{u}_j^{(n)}, \quad (1)$$

where $u_j^{(n)}$ and $\bar{u}_j^{(n)} (j = 1, 2, 3)$ are the displacement components in the crystal plate and electrodes, respectively.

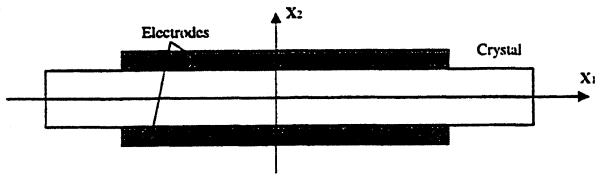


Fig. 1 An electroded quartz crystal plate.

It is apparent that the displacement is continuous at the interface of the electrode and crystal plate. This is important because there will be no special requirement in the treatment of the boundary conditions later on.

Starting with the displacement field given in (1) and the three-dimensional elasticity definitions, we have the strain field in the electrode as

$$S_{ij} = \sum_{n=0} b^n \bar{S}_{ij}^{(n)} + \sum_{n=1} (x_2 - b)^n \tilde{S}_{ij}^{(n)}, \quad (2)$$

where two components of the strain

$$\bar{S}_{ij}^{(n)} = \frac{1}{2}\left(u_{j,i}^{(n)} + u_{i,j}^{(n)}\right),$$

$$\tilde{S}_{ij}^{(n)} = \frac{1}{2}\left[\bar{u}_{j,i}^{(n)} + \bar{u}_{i,j}^{(n)} + (n+1)\left(\delta_{2i}\bar{u}_j^{(n+1)} + \delta_{2j}\bar{u}_i^{(n+1)}\right)\right], \quad (3)$$

are related to the displacements of the crystal plate and electrodes, respectively.

As a result, in the electrode the stress is

$$T_{ij} = \sum_{n=0} b^n \bar{c}_{ijkl}\bar{S}_{kl}^{(n)} + \sum_{n=0} (x_2 - b)^n \bar{c}_{ijkl}\tilde{S}_{kl}^{(n)}, \quad (4)$$

where $\bar{c}_{ijkl}$ are the elastic constants of the electrode. Again, there are two parts in the stress tensor.

For the electrode, the three-dimensional variational equations of motion are

$$\int_A \int_b^{b+2\bar{b}} (T_{ij,i} - \bar{\rho}\ddot{u}_j)\,\delta u_j dx_2 dA = 0, \quad (5)$$

or

$$\int_b^{b+2\bar{b}} \left\{T_{ij,i} - \bar{\rho}\left[\sum_{n=0} b^n \ddot{u}_j^{(n)} + \sum_{m=1} (x_2 - b)^m \ddot{\bar{u}}_j^{(m)}\right]\right\} \times \left[\sum_{n=0} b^n \delta u_j^{(n)} + \sum_{m=1} (x_2 - b)^m \delta\bar{u}_j^{(m)}\right] dx_2 = 0, \quad (6)$$

where A, $\bar{b}$, and $\bar{\rho}$ are the upper crystal face, half thickness of the electrode, and the density of the electrode material, respectively.

Utilizing the definition of the displacement field, now we can integrate the stress in (4) over the thickness of the electrode to define the two-dimensional stress components

$$\bar{T}_{ij}^{(n)} = \int_b^{b+2\bar{b}} T_{ij} b^n dx_2, n = 0, 1, 2, 3,$$

$$\tilde{T}_{ij}^{(n)} = \int_b^{b+2\bar{b}} T_{ij} (x_2 - b)^n dx_2, n = 1, 2, 3, \quad (7)$$

of the electrode.

Through integration by parts, we can easily combine the equations of motion of both electrode and crystal plate in variational form as

$$\sum_{n=0}\left[\int_A \bar{T}_{ij}^{(n)}\delta\bar{S}_{ij}^{(n)} dA + \int_A \bar{\rho}\left(\sum_{m=0} 2\bar{b}\bar{b}^{m+n}\ddot{u}_j^{(m)} + \sum_{m=1} \frac{2^{m+1}}{m+1} b^n \bar{b}^{m+1}\ddot{\bar{u}}_j^{(m)}\right) dA\right] = \sum_{n=0}\left[\int_C n_i \bar{T}_{ij}^{(n)} ds + \int_A \bar{F}_j^{(n)} dA\right]\delta u_j^{(n)},$$

$$\sum_{n=1}\left[\int_A \tilde{T}_{ij}^{(n)}\delta\tilde{S}_{ij}^{(n)} dA + \int_A \bar{\rho}\left(\sum_{m=0} \frac{2^{n+1}}{n+1} b^m \bar{b}^{n+1}\ddot{u}_j^{(m)} + \sum_{m=1} \frac{2^{n+m+1}}{n+m+1}\bar{b}^{n+m+1}\ddot{\bar{u}}_j^{(m)}\right) dA\right] = \sum_{n=1}\left[\int_C n_i \tilde{T}_{ij}^{(n)} ds + \int_A \tilde{F}_j^{(n)} dA\right]\delta\tilde{u}_j^{(n)},$$

$$\sum_{n=0}\int_A \left(T_{ij}^{(n)}\delta S_{ij}^{(n)} + \rho\sum_{m=0} B_{mn}\ddot{u}_j^{(m)}\right) dA = \sum_{n=0}\left(\int_C n_i T_{ij}^{(n)} ds + \int_A F_j^{(n)} dA\right)\delta u_j^{(n)}. \quad (8)$$

By combining the above equations, some of the face-traction terms will disappear. Of course, these equations will be the basis of the finite element implementation of the layerwise theory of the electrodes.

Obviously, the stresses in the electrode and crystal plate can be rewritten as

$$\bar{T}_{ij}^{(n)} = \sum_{m=0} \bar{C}_{ijkl}^{(nm)}\bar{S}_{kl}^{(m)} + \sum_{m=1} \tilde{C}_{ijkl}^{(nm)}\tilde{S}_{kl}^{(m)}, n = 0, 1, 2, 3,$$

$$\tilde{T}_{ij}^{(n)} = \sum_{m=0} \tilde{C}_{ijkl}^{(mn)}\bar{S}_{kl}^{(m)} + \sum_{m=1} \hat{C}_{ijkl}^{(nm)}\tilde{S}_{kl}^{(m)}, n = 1, 2, 3,$$

$$T_{ij}^{(n)} = \sum_{m=0} C_{ijkl}^{(nm)} S_{kl}^{(m)}, n = 0, 1, 2, 3, \quad (9)$$

where $\bar{C}_{ijkl}^{(nm)}$, $\tilde{C}_{ijkl}^{(nm)}$, $\hat{C}_{ijkl}^{(nm)}$, and $C_{ijkl}^{(nm)}$ are the generalized elastic constant matrices of the electrodes and crystal plate. The detailed structure and components of these matrices can be found from corresponding equations of the stress components.

III. Finite Element Implementation

For matrix implementation, we start with the definition of the displacement vector

$$\mathbf{u} = \{\mathbf{u}_p, \mathbf{u}_e\}, \tag{10}$$

where $\mathbf{u}_p$ and $\mathbf{u}_e$ are the displacement vectors of the crystal plate and electrode, respectively. It is obivouis that we can treat the displacement of the crystal plate as a subset of the displacement vector by defining

$$\mathbf{u}_p = \mathbf{V}\mathbf{u}, \tag{11}$$

where $\mathbf{V}$ is a transform matrix.

For discretization, as usual, we define

$$\mathbf{u} = \mathbf{N}\mathbf{U}, \tag{12}$$

where $\mathbf{N}$ and $\mathbf{U}$ are the shape function matrix and nodal displacement, respectively [9]. As a result, the discrete strain-displacement and constitutive relations can be written as

$$\begin{aligned}
&\mathbf{S} = \{\mathbf{S}_p, \mathbf{S}_e\}, \mathbf{S} = \mathbf{B}\mathbf{U}, \mathbf{S}_p = \mathbf{B}_p\mathbf{U}_p = \mathbf{B}_p\mathbf{V}\mathbf{U}, \\
&\mathbf{T} = \{\mathbf{T}_p, \mathbf{T}_e\}, \mathbf{T} = \mathbf{C}\mathbf{S}, \mathbf{T}_p = \mathbf{C}_p\mathbf{S}_p = \mathbf{C}_p\mathbf{V}\mathbf{U},
\end{aligned} \tag{13}$$

where subscript p denotes the quantities related to the crystal plate.

Accoordingly, the equations of motion (8) will be given in matrix form as

$$\begin{aligned}
&\left(\mathbf{V}^T\mathbf{K}_p\mathbf{V} + \mathbf{K}\right)\mathbf{U} + \left(\mathbf{V}^T\mathbf{M}_p\mathbf{V} + \mathbf{M}\right)\ddot{\mathbf{U}} \\
&= \mathbf{V}^T\mathbf{F}_C^p + \mathbf{F}_C + \mathbf{V}^T\mathbf{F}_A^p + \mathbf{F}_A,
\end{aligned} \tag{14}$$

where

$$\begin{aligned}
&\mathbf{K}_p = \int_A \mathbf{B}_p^T\mathbf{C}_p\mathbf{B}_p dA, \mathbf{K} = \int_A \mathbf{B}^T\mathbf{C}\mathbf{B} dA, \\
&\mathbf{M}_p = \int_A \mathbf{N}^T\mathbf{m}_p\mathbf{N} dA, \mathbf{M} = \int_A \mathbf{N}^T\mathbf{m}\mathbf{N} dA, \\
&\mathbf{F}_C^p = \int_C \mathbf{N}\mathbf{T}_p ds, \mathbf{F}_A^p = \int_A \mathbf{N}\mathbf{F}_p dA, \\
&\mathbf{F}_C = \int_C \mathbf{N}\mathbf{T} ds, \mathbf{F}_A = \int_A \mathbf{N}\mathbf{F} dA,
\end{aligned} \tag{15}$$

are stiffness matrix of the crystal plate, stiffness of the electrodes, mass matrix of the crystal plate, mass matrix of the electrodes, and element force vectors. Again, we skip the discussions of these matrices.

For the computation of the mass matrices $\mathbf{M}_p$ and $\mathbf{M}$, some of the elements are

$$\begin{aligned}
&\mathbf{m}_p^{(nm)} = B_{nm}\rho\mathbf{I}_3, n, m = 0, 1, 2, 3; \\
&\mathbf{m} = \begin{bmatrix} \bar{\mathbf{m}} & \tilde{\mathbf{m}} \\ \tilde{\mathbf{m}} & \hat{\mathbf{m}} \end{bmatrix}, \\
&\bar{\mathbf{m}}^{(nm)} = 2b^{n+m}\left[\bar{\rho}\bar{b} + (-1)^{n+m}\tilde{\rho}\tilde{b}\right]\mathbf{I}_3, n, m = 0, 1, 2, 3, \\
&\tilde{\mathbf{m}}^{(nm)} = \frac{2^{m+1}}{m+1}b^n\left[\bar{\rho}\bar{b}^{m+1} + (-1)^{n+m}\tilde{\rho}\tilde{b}^{m+1}\right]\mathbf{I}_3, \\
&\quad n = 0, 1, 2, 3; m = 1, 2, 3, \\
&\hat{\mathbf{m}}^{(nm)} = \frac{2^{n+m+1}}{n+m+1}\left[\bar{\rho}\bar{b}^{n+m+1} + (-1)^{n+m}\tilde{\rho}\tilde{b}^{n+m+1}\right]\mathbf{I}_3, \\
&\quad n, m = 1, 2, 3.
\end{aligned} \tag{16}$$

For the effects of the mass and stiffness, we can use the resonance frequencies of the free vibration for comparison. In this case, we have the free vibration problem

$$\left(\mathbf{V}^T\mathbf{K}_p\mathbf{V} + \mathbf{K}\right)\mathbf{U} - \omega_0{}^2\Omega^2\left(\mathbf{V}^T\mathbf{M}_p\mathbf{V} + \mathbf{M}\right)\mathbf{U} = 0, \tag{17}$$

where ω_0 and Ω are the fundamental thickness-shear and normalized frequencies, respectively.

IV. Numerical Examples

To demonstrate the layerwise theory for the electrodes, we use the finite element method to evaluate the frequency shift for an electroded crystal plate. The plate we use for the evaluation is essentially the sample measured by Koga [11], but we covered about a quarter of the plate in the central region with electrodes on both faces. Four-node elements with incompatible nodes are used for the computation.

In Fig. 2, we compare the frequency spectra of the pure crystal plate with the partially electroded crystal plate, but we used a model based on the assumption that the deformation in the electrodes are uniform. As a result, both the stiffness and mass have been considered in the computation [10]. It can be observed that there is a negligible It shows that for small length (width) to thickness ratio, the effect of the electrodes are actually limited.

Next in Fig. 3 we analyzed the same resonator with the first-order layerwise model discussed in this paper. We observe the general trend of the frequency shift, but it is hard to compare with the ones shown in Fig. 2 because there are much more frequency branches due to the presence of the electrodes. A close examination of the frequency spectrum reveals that many of the branches actually relate to the vibrations of the electrodes and there are virtually no vibrations inside the crystal plate.

If we turn to Fig. 4, the results from the third-order electrodes plate theory, it is clear that the elimination

of the extraneous modes is important. Without such a procedure, these results will be difficult to use in the design process.

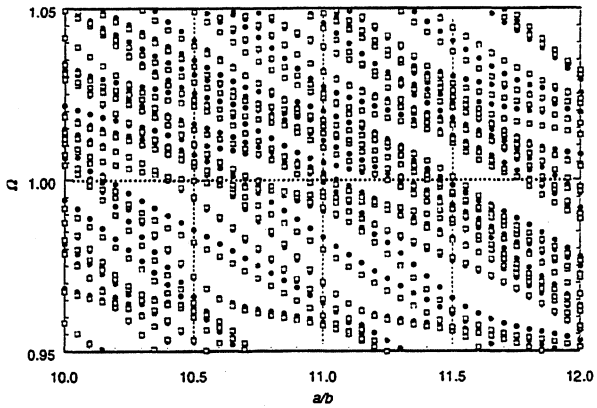


Fig. 2 Frequency spectra of a rectangular crystal plate (dot; c/b=16.3660) in comparision with a partially electroded crystal plate (square; $R = 4.803\%$).

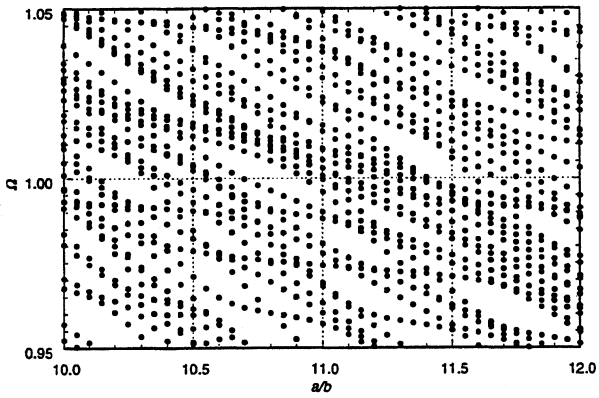


Fig. 3 Frequency spectrum of a partially electroded crystal plate ($c/b = 16.3660$ and $R = 4.803\%$) with the first-order electrode consideration.

V. Conclusions

The layerwise plate theory for the electrodes of a crystal resonator structure is derived. These equations are further formulated and implemented for the application of the finite element method to the analysis the effect of electrodes on crystal resonators.

The disadvantage we have at this moment is that the increase of the variables in the electrodes will inevitably increase the number of frequency solutions thus making it practically impossible to use the frequency spectrum if possible elimination procedure is not in place. This is indeed a challenge worth of further studying.

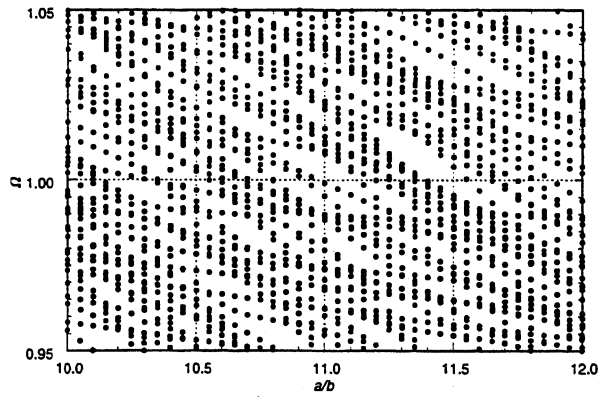


Fig. 4 Frequency spectrum of a partially electroded crystal plate ($c/b = 16.3660$ and $R = 4.803\%$) with the third-order electrode consideration.

References

[1] Y. -K. Yong, J. Wang, T. Imai, S. Kanna, and E. Momosaki. "A set of hierarchical finite elements for quartz plate resonators". In *Proceedings of the 1996 IEEE Ultrasonics Symposium*, pages 981–985, San Antonio, TX, November 1996. IEEE.

[2] Y. -K. Yong, J. Wang, and T. Imai. "On the accuracy of Mindlin plate predictions for the frequency-temperature behavior of resonant modes in AT- and SC-cu quartz plates". *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 46(1):1–13, 1999.

[3] J. Wang, Y. -K. Yong, and T. Imai. "Higher-order plate theory based finite element analysis of the frequency-temperature relations of quartz crystal resonators". In *Proceedings of the 1998 IEEE International Frequency Control Symposium*, pages 956–963, Pasadena, CA, May 1998. IEEE.

[4] J. Wang, J. D. Yu, Y. -K. Yong, and T. Imai. " A new theory for electroded piezoelectric plates and its finite element application for the forced vibrations analysis of quartz crystal resonators ". In *Proceedings of the 1998 IEEE International Ultrasonics Symposium*, pages 913–918, Sendai, Japan, October 1998. IEEE.

[5] R. D. Mindlin. "High frequency vibrations of plated, crystal plates". In *Progress in Applied Mechanics*, pages 73–84. Macmillan, New York, 1963.

[6] H. F. Tiersten. *Linear Piezoelectric Plate Vibrations*. Plenum Press, New York, 1969.

[7] P. C. Y. Lee, S. Syngellakis, and J. P. Hou. "A two-dimensional theory for high-frequency vibrations of piezoelectric crystal plates with or without electrodes". *J. Appl. Phys.*, 61(4):1249–1262, 1987.

[8] J. Wang, Y. -K. Yong, and T. Imai. "Finite element analysis of the piezoelectric vibrations of quartz plate resonators with higher-order plate theory". *Int. J. Solids Struct.*, 36(15):2303–2319, 1999.

[9] J. Wang, Y. -K. Yong, and T. Imai. "Higher-order plate theory based finite element analysis of the frequency-temperature relations of quartz crystal resonators". To appear.

[10] J. N. Reddy. *Mechanics of Laminated Composite Plates: Theory and Analysis*. CRC Press, Boca Raton, Florida, 1997.

[11] I. Koga. "Radio-frequency vibrations of rectangular AT-cut quartz plates". *J. Appl. Phys.*, 34(8):2357–2365, 1963.

RELATIONSHIP BETWEEN QUARTZ CRYSTAL ORIENTATION AND THE SURFACE QUALITY OBTAINED BY ULTRASONIC MACHINING

Pedro L. Guzzo, Alberto A. Raslan, José Daniel B. De Mello
Universidade Federal de Uberlândia,
Departamento de Ciências Físicas, Laboratório de Tribologia e Materiais
Campus Santa Mônica, Bloco 1R, 38400-902 Uberlândia - MG- BRASIL.

ABSTRACT

Ultrasonic machining (USM) of natural quartz is investigated in relation to crystal orientation. The rate of cutting was measured along directions perpendicular to X-, Y-, Z- and AT-cut planes. Surface profilometry was carried out on lateral surfaces achieved by USM. It is shown that cutting rate and surface roughness are dependent on crystal orientation. These dependencies are preliminary discussed in relation to indentation fracture mechanics and fracture toughness anisotropy of natural quartz. Scanning electron microscopy analysis shows that brittle microcracking is the main wear mechanism involved with material removal.

1. INTRODUCTION

Since the 1940s, α-quartz has been the material of choice in the majority of frequency control devices. Due to its piezoelectric and elastic properties, quartz crystals have been used in the production of electromechanical devices such as resonators, filters and sensors [1,2]. Virtually, all devices fabricated today use cultured quartz. These crystals exhibits greater uniformity than natural quartz and are frequently avaiable free from inclusions, twins and with low amounts of impurities. One of the remaining uses of natural electronic-grade material is in the manufacture of resonators used as pressure sensors in deep wells. Besides crystal quality, of the primary importance in frequency control devices is the manufacturing technology employed to fabricate these units. The metrological accuracy required for high-Q resonators depends on both surface finish and accuracy of plate geometry [1]. Because quartz is a brittle, hard and nonconductive material, ultrasonic machining (USM) appears as one of the few types of manufacturing processes showing the ability to machine electromechanical devices with complex shapes. For instance, the edge shape and the bridges of BVA resonators and the monolithic strucutre of pressure sensors have been achieved by USM [3].

From the tribological point of view, USM can be classified as a three-body abrasive wear mode [4]. USM is an impact grinding process where abrasive particles are hammered against the surface of the workpiece. A cutting tool attached to a transducer/horn combination is vibrated longitudinally at a resonant frequency. A slurry with abrasive grains is flushed through the gap between tool and the work surface. The tool progressively reproduces its shape into the workpiece, as long as machining takes place.

Similar as in grinding, lapping or polishing, research on USM presents certain difficulties due to the large number of operating conditions [5,6]. Material removal rate, surface finish and machining accuracy are influenced by various process parameters such as amplitude and frequency of ultrasonic oscillations, hardness and size of abrasive particles, static load and properties of the workpiece material. The influence of these parameters on surface finish and metrological accuracy of resonator's components was not systematically investigated. Until now, just the effect of abrasive grain size on the out-of-roundness of BVA components was reported [7]. Additionally, there is any investigation relating mechanical properties of quartz crystals and the surface finish obtained by USM.

The present work investigates the influence of crystallographic orientation on the quality of quartz surfaces obtained by USM. For this, X-, Y-, Z- and AT-cut samples of natural quartz were machined keeping the operating conditions as constant as possible. The wear mechanisms and the surface roughness were characterized by scanning electron microscopy and surface profilometry, respectively. It is observed that cutting rate and the roughness of machined surfaces are sensitive to specimen orientation. This dependence is preliminary discussed in relation to indentation fracture mechanics and fracture toughness anisotropy of natural quartz reported by other investigators.

2. EXPERIMENTAL PROCEDURE

A left-hand natural quartz block was oriented using polarized light, etching and X-Ray diffraction technique [8]. X-, Y-, Z- and AT-cut plates free from inclusions and cracks were cut up on a diamond saw to within ± 1°of the required crystallographic plane. Following the IRE standard [9], X-, Y- and Z- cuts are identified by $\{11\bar{2}0\}$, $\{10\bar{1}0\}$ and {0001}, respectively. The AT-cut is turned of 3.5° in relation to the minor rhombohedral plane $\{01\bar{1}1\}$. The plates were lapped with a coarse alumina (Al_2O_3) abrasive to eliminate surface flaws induced by sawing. Additionally, this lapping improved the parallelism between faces and produced a thickness of 3 mm.

USM experiments were carried out in a stationary Sonic-Mill® machine with a power output of 1790 W. A piezoelectric transducer converts the electric oscillations into mechanical vibrations in the frequency range of 20 KHz. A conical titanium horn with diameter ratio of 2.5 was used in all experiments. The amplitude of vibration of the tool was set at the minimum value

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

and the static load was kept constant at 81N by a pneumatic balance system. The slurry was made with boron carbide (B_4C) with mean grain size of 25 μm (400 mesh) and distilled water in the ratio of 1 : 2.5 by volume. The abrasive flow rate, set near to 3.3 l/min, was kept as uniform as possible. A cutting tool with hexagonal section (Fig. 1) was fabricated in stainless steel. The area of the hole is 95 mm^2.

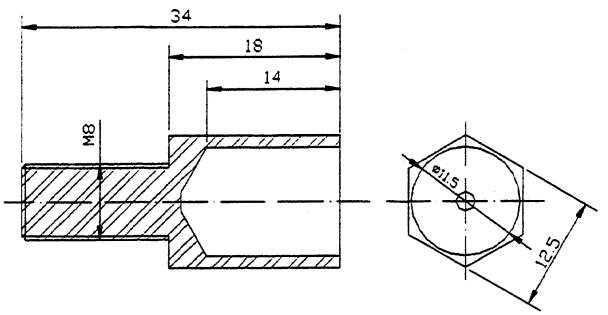


Fig. 1 – Drawing of the cutting tool.

The cutting rate was determined by measuring the time expended to machine 100 μm into depth. At least fifteen measurements were carried out during USM of each hole. Two holes rotated approximately by 90° in relation to machining direction were made on each plate. The roughness profiles were measured with a Rank Taylor Robson profilometer, type Surtronic 3+, using a cutt-off value of 0.8 mm. Seven profiles were measured perpendicularly to machining direction on lateral surfaces generated by USM. The centre-line average (R_a) was choosed to characterize the surface roughness. By using a Carl Zeiss DSM 94A microscope, scanning electron microscopy (SEM) was employed to characterize wear micromechanisms.

3. RESULTS AND DISCUSSION

Table 1 shows the cutting rate (CR) and the surface roughness (R_a) resultant from USM of natural quartz. During the experiments, several samples of plate glass were also machined in order to monitor severe variations in operating conditions. The CR of glass specimens showed small scattering near to 12 μm/s suggesting that fluctuations in operating conditions did not obscure the values of CR presented here.

The relationship between R_a and CR is shown in Fig. 2. The scattering of experimental points associated with natural quartz samples is large, becoming hard to find a clear relation between R_a and CR. This behavior can be explained by the fact that CR and R_a were not measured on the same crystallographic plane. For a given hole, CR was measured in relation to the crystallographic plane that was perpendicular to machining direction whereas R_a was measured on lateral surfaces generated by USM, i.e., crystallographic planes parallel to machining direction. The intrinsic dispersion related to abrasive processes of brittle solids could also be contributing to increase the scattering shown in Fig. 2. Further experiments are required to clarify the dependence between CR and R_a of quartz surfaces achieved by USM.

Table 1- Cutting rate (CR) and roughness (R_a) of machined surfaces of natural quartz and glass.

Machined Direction	CR (μm/s)	Machined Plane	R_a (μm)
X	10.9 (1.0)	Y	1.8 (0.3)
	10.8 (0.9)	Z	1.4 (0.1)
Y	10.4 (1.0)	X	1.4 (0.2)
	11.4 (0.6)	Z	1.0 (0.1)
Z	10.5 (0.8)	X	2.1 (0.4)
	9.8 (0.8)	Y	1.7 (0.2)
AT (Y')	9.6 (1.2)	X	1.3 (0.1)
	9.4 (1.3)	Z'	2.2 (0.3)
Glass	11.9 (0.9)		2.5 (0.6)
	12.0 (0.9)		2.6 (0.3)

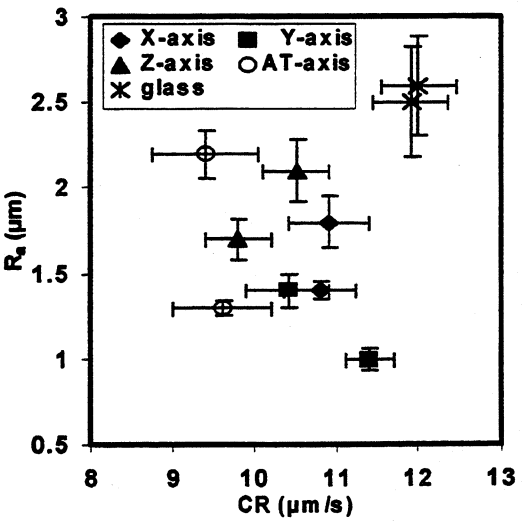
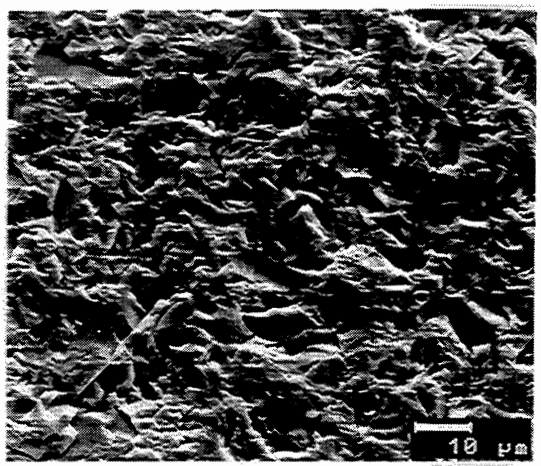
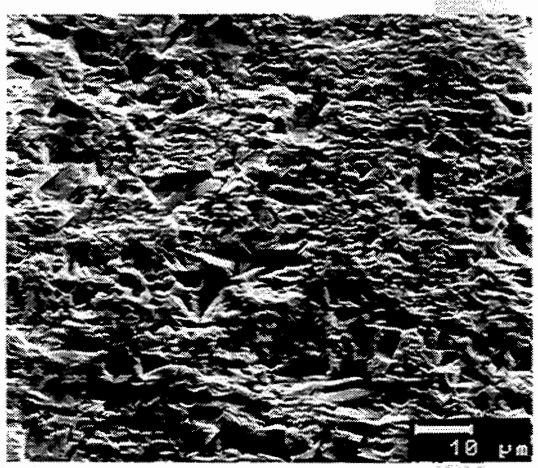


Fig.2 – Relationship between the cutting rate (CR) of machining directions and the roughness (R_a) of lateral surfaces generated by USM of natural quartz and glass.

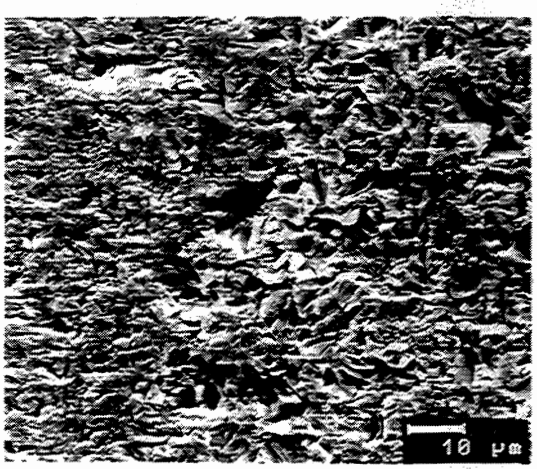
The micrographs shown in Fig. 3 illustrate the pattern of lateral surfaces obtained by USM. It is observed that the wear mechanism controlling the material removal is brittle microcracking. In this machining process, the hammering action of abrasive particles on the work surface [5,6] is responsible for crack formation and crack propagation into the workpiece material. Fig. 3 shows that cracking events appearing on lateral surfaces parallel to X- and Y-cut planes are larger than those observed on the surface parallel to Z-cut. SEM analysis also revealed the presence of channels-like patterns mainly on X-cut and glass surfaces. Previously, it was reported that these channels are probably shaped by the motion of cavitation streamers carring abrasive particles which tend to cut deep furrows on lateral surfaces [10].



(a) X-cut



(b) Y-cut



(c) Z-cut

Fig. 3 – SEM micrographs of lateral surfaces of natural quartz achieved by USM.

Indentation fracture analysis in brittle solids made clear that lateral crack extension can provide a basis for estimating material removal rates for several practical situations [11]. Extending this analysis to stationary USM [12], the volume of material (V) that might be removed per unit of time can be expressed as :

$$V \propto \frac{P^2}{K_{IC}^{3/2} H^{1/2}} \frac{f}{n} \qquad (1)$$

where P is the static load (N), f is the frequency of ultrasonic machining (Hz) and n is the effective number of abrasive particles in the gap between the tool and the workpiece. K_{IC} and H are the fracture toughness and the static hardness of the workpiece material, respectively. From Eq. (1) it is observed that material removal rate is inversely proportional to $K_{IC}^{3/2}$ and $H^{1/2}$.

The values of K_{IC} measured by Iwasa and Brad [13] for X-, Y-, Z- and AT-cut of natural quartz were 0.85, 0.97, 1.15 and 0.86 MPa $m^{1/2}$, respectively. It was not found in literature static hardness for natural quartz measured as a function of crystal orientation. The avaiable data were measured on X-, Y- and Z-cut samples of synthetic quartz [14]. As it was not detected influence of crystal orientation [14], H = 10 GPa was also adopted to AT-cut. In case of glass, it was found K_{IC} = 0.7 MPa $m^{1/2}$ and H = 5.5 GPa [11].

It is known that brittle microcraking depends on the resistance of crack propagation through the media [4]. From the reported values of K_{IC}, it is observed that Z-cut is tougher than X-, Y-, and AT-cut planes which suggests that Z-cut may be more resistant to microcracking. Thus, it is expected that the rate of material removal by the hammering effect of the abrasive particles on the Z-cut should be lower compared with other planes. This prediction is in agreement with Fig. 4 which correlates experimental values of CR with the product $K_{IC}^{3/2}H^{1/2}$. Neglecting

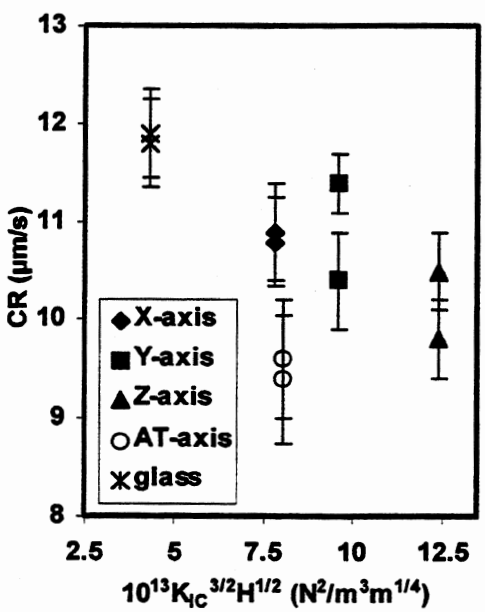


Fig. 4 – Relationship between the cutting rate (CR) of machining directions and the product $K_{IC}^{3/2}H^{1/2}$.

the points associated with AT-cut, it is observed that CR approximately decreases with increasing $K_{IC}^{3/2}H^{1/2}$. The dispersion shown in Fig. 4 suggests that further experimental work is required to determine static hardness of crystallographic planes of natural quartz.

The relationship between R_a and $K_{IC}^{3/2}H^{1/2}$ is plotted in Fig. 5. Although scattering of experimental points related to X-cut is large, it is observed that R_a is approximately linearly correlated with $K_{IC}^{3/2}H^{1/2}$. Machined surfaces parallel to Z-cut exhibit the lowest values of R_a which increases in the sequence, Z-cut < Y-cut < X-cut < glass. This behavior can also be explained by the intensity of brittle microcraking shown in Fig. 3. The occurrence of cracking events and the size of craters should increase with decreasing crack propagation resistance.

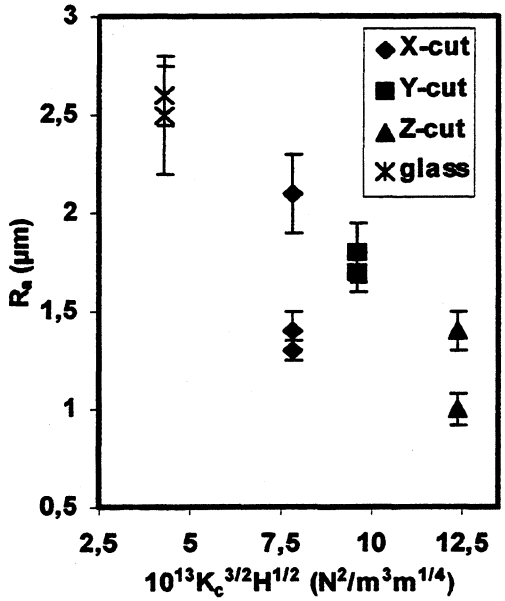


Fig. 5 – Relationships between the roughness (R_a) of lateral surfaces and the product $K_{IC}^{3/2}H^{1/2}$.

The relationships shown in Figs. 4 and 5 made clear that the rate of cutting and the surface roughness resultant from USM are connected each other by the product $K_{IC}^{3/2}H^{1/2}$. It is observed that crystal directions which have higher rates of material removal are associated with lateral surfaces with inferior surface finish.

4. CONCLUSION

The rate of material removal and the roughness of lateral surfaces achieved by USM seem to be dependent on crystallographic orientation of natural quartz. Compared with other low-indice planes, Z-cut exhibited the lowest values of cutting rate and surface roughness due to its higher brittle microcraking resistance. The product $K_{IC}^{3/2}H^{1/2}$ appears to be suggestive to predict the influence of crystal anisotropy on both material removal rate and surface roughness resultant from abrasive machining processes. Further investigations are required to elucidate the relationship between the cutting rate and the surface roughness associated with USM of quartz crystals.

ACKNOWLEDGEMENTS

This wok was financed by CNPq (P:620136/92-8) and FAPEMIG (TEC 263/92). One of the authors (PLG) is grateful for finantial support provided by CNPq and FAPEMIG. The undergraduate students Israel C. Nuñes and Jean C. Rezende assisted with preparation of specimens and with ultrasonic machining tests.

REFERENCES

[1] J.C. Brice, "Crystals for quartz resonators," Rev. Mod. Phys., vol. 57, pp. 105-146, 1985.

[2] A. Ballato, "Advances in high-Q piezoelectric resonator materials and devices," IEEE Trans. on Ultras. Ferroelec. and Freq. Control, vol. 41, pp. 834-844, 1994.

[3] R.J. Besson, "Innovative frequency standards : Piezoelectric resonators and others," Proc.of the *4th European Frequency Time Forum*, 1990, pp.99-105.

[4] K.H. Zum Gahr, Microstructure and wear of materials. Amsterdam : Elsevier Sc. Pub., 1987, 554p.

[5] M.C. Shaw, "Ultrasonic grinding", Int. Institution for Production Eng. Res., vol. 10, pp. 257-265, 1956.

[6] V. Soundararajan, V. Radhakrishnan, "An study on the mechanisms involved in ultrasonic machining", Int. J. on Machine Tool Design and Res., vol. 26, pp. 307-321, 1986.

[7] A.A Raslan, J.D.B. De Mello, "Usinagem do quartzo por processos mecânicos", Proc. of the *II Congresso Internacional de Metalurgia e Materiais*, 1997.

[8] R.A. Heising, Quartz crystals for electrical circuits. New York: Van Nostrands Co. Inc., 563p.

[9] I.R.E. Standard on piezoelectric crystals, Proc. of the Inst. of Radio Enginners, vol. 37, pp. 1378-1395, 1949.

[10] E.A. Neppiras, R. Foskett, "Ultrasonic machining I – Technique and Equipament", Philips Tech. Review, vol. 18, pp. 325-334, 1956/57.

[11] A.G. Evans, T.R. Wilshaw, 'Quasi-static solid particle damage in brittle solids I – Observations analysis and implications", Acta Metallurgica, vol. 24, pp. 939-956, 1976.

[12] M. Komaraiah, P.N. Reddy, "A study on the influence of workpiece properties in ultrasonic machining", Int. J. Mach. Tools. Manufact., vol. 33, pp. 495-505, 1993.

[13] M. Iwasa, R.C. Bradt, "Clevage of natural and synthetic single crystals." Mat. Res. Bulletin, vol. 22, pp. 1241-1248, 1987.

[14] N.E.W. Hartley, T.R. Wilshaw, "Deformation and fracture of synthetic α-quartz", J. Mat. Science, vol. 8, pp. 265-278, 1973.

TRAPPED-ENERGY VIBRATORY GYROSCOPES USING ROTATED *Y*-CUT $LiNbO_3$

Kiyoshi Nakamura and Masahiro Ohsaki

Department of Electrical and Communication Engineering,
Graduate School of Engineering, Tohoku University,
Aoba 05, Aramaki, Aoba-ku, Sendai 980-8579, Japan

ABSTRACT

In a rotated *Y*-cut plate of $LiNbO_3$, two orthogonal thickness-shear vibrations can be excited independently, by electric fields perpendicular and parallel to the plate surface. Using the trapped-energy thickness-shear vibrations, piezoelectric vibratory gyroscopes of easy mount with no need of trimming are obtained. In this paper, a trapped-energy gyroscope using a rotated *Y*-cut of $LiNbO_3$ is proposed. It is shown experimentally that by choosing appropriate cuts the two orthogonal thickness-shear vibrations have close resonance frequencies, and the energy of the thickness-shear vibrations is trapped around the electrodes. The basic characteristics of a vibratory gyroscope using a 163° rotated *Y*-cut $LiNbO_3$ plate is presented.

1. INTRODUCTION

Piezoelectric vibratory gyroscopes [1] have been put into practical use as angular velocity sensors. At present they are being used for applications such as VTR hand cameras and car navigation systems. Although some types of bending vibrators have been exclusively used in practical gyroscopes, such vibrators tend to be subjected to deviation and deterioration of their characteristics, due to bonding of the piezoelectric elements and wires, and mounting the vibrators. Furthermore, the device height is large, as the length direction of the vibrator has to coincide with the rotation axis to be detected.

As is well known, the energy of the thickness-shear vibrations in a piezoelectric plate can be trapped in the electroded portion due to the effects of the electrode mass loading and the piezoelectric unstiffening [2,3]. Hence, even if the vibrator is supported at the plate edge, the Q-factor of the trapped-energy resonances hardly decreases. Using such trapped-energy vibrations, it is possible to obtain thin vibratory gyroscopes of easy mount [4]. Recently Abe et al. [5] reported a trapped-energy vibratory gyroscope using piezoelectric ceramics.

In this paper, a trapped-energy piezoelectric gyroscope using a single crystal of lithium niobate ($LiNbO_3$) is proposed[6]. The basic characteristics of the trapped-energy vibratory gyroscope, fabricated using a 163° rotated *Y*-cut $LiNbO_3$ plate are presented.

2. PRINCIPLE AND STRUCTURE

2.1 Perpendicular and Parallel-Field Excitation

The vibratory gyroscope as an angular velocity sensor is based on the principle that when an angular velocity acts on the gyroscope vibrating in one of the two orthogonal vibration modes, another one is excited via the Coriolis force proportional to the angular velocity. The two orthogonal vibrations are required to be piezoelectrically excited and detected independently and it is desirable for the resonance frequencies of these two vibrations to be almost the same.

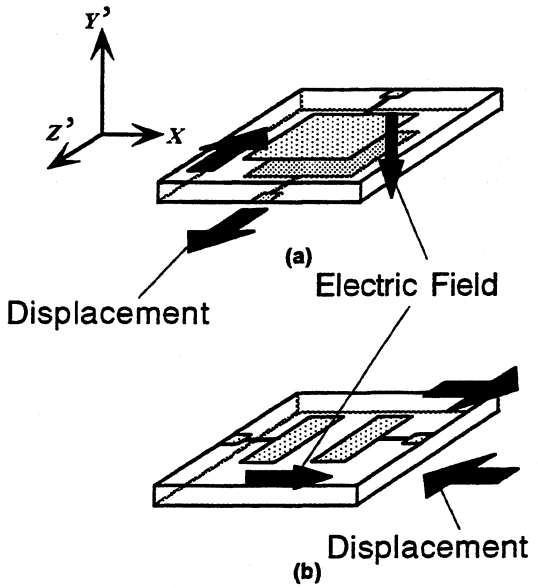


Fig.1. Excitation of the two orthogonal thickness-shear modes. (a) Perpendicular-field excitation. (b) Parallel-field excitation.

The vibratory gyroscope proposed in this paper utilizes the two orthogonal thickness-shear vibrations in rotated *Y*-cut plates of $LiNbO_3$, where the thickness-shear vibration modes can be independently excited and detected piezoelectrically, using the perpendicular and parallel-field excitations. In the perpendicular-field excitation shown in Fig.1 (a), the *Z'*-polarized thickness-shear mode can be excited by the electric field in the thickness direction, while the *X*-polarized thickness-shear mode cannot, because a voltage is applied between

0-7803-5400-1/99/ $ 10.00

the electrodes on the top and bottom surfaces. In the 163° rotated Y-cut $LiNbO_3$, the effective coupling factor of the Z'-polarized thickness-shear mode is as high as 62%, whereas that of the thickness-extensional mode is zero [7]. Therefore, only the Z'-polarized thickness-shear mode can be excited by the perpendicular-field.

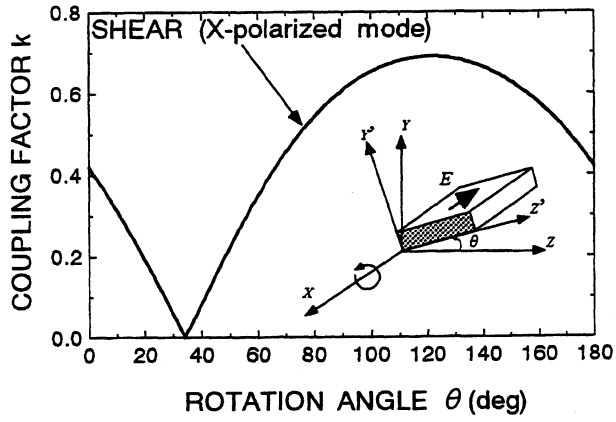


Fig.2. Rotation angle dependence of the coupling factor for parallel-field excitation of the X-polarized thickness-shear mode in rotated Y-cut $LiNbO_3$.

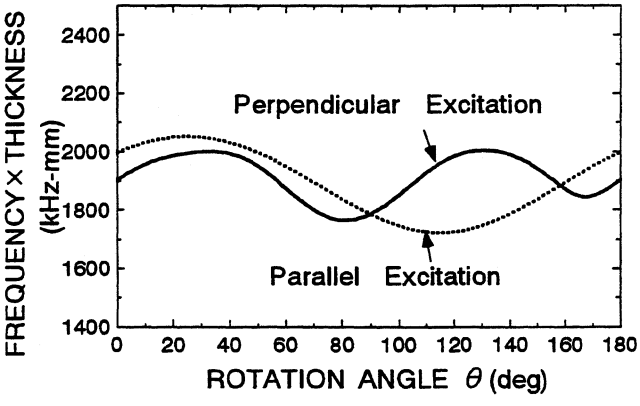


Fig.3. Frequency constants for the Z'-polarized thickness-shear mode vs. rotation angle θ.

On the other hand, in the parallel-field excitation shown in Fig.1(b), the X-polarized thickness-shear mode can be strongly excited by an electric field parallel to the X-axis, while the thickness-extensional mode and the Z'-polarized thickness-shear mode cannot be excited. The effective coupling factor of the X-polarized thickness-shear mode for parallel-field excitation was calculated using Yamada and Niizeki's method [8]. The results are shown in Fig.2. In the 163° rotated Y-cut $LiNbO_3$, the effective coupling factor of the thickness-shear mode is as high as 56%. Figure 3 shows the calculated frequency constants (resonance frequencies × plate thickness) for the two thickness-shear modes [7]. The resonance frequencies of the two thickness-shear modes are close to each other at θ=163°. Considering the above discussion, it may be concluded that the 163° rotated Y-cut plate is suitable for piezoelectric gyroscopes.

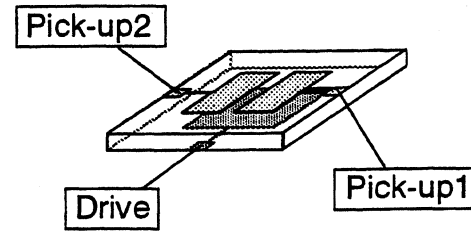


Fig.4. Trapped-energy piezoelectric gyroscope using roted Y-cut $LiNbO_3$.

2.2 Structure of Gyroscopes

The structure of the gyroscope using perpendicular and parallel-field excitations are shown in Fig. 4. In the gyroscope, the Z'-polarized thickness-shear mode is driven by applying a voltage between the electrodes on the top and bottom surfaces (perpendicular-field excitation). The parallel-field due to the X-polarized thickness-shear mode generated by the Coriolis force is detected by measuring the difference of currents from the pick-up 1 and pick-up 2 electrodes, which is proportional to the angular velocity.

3. EXPERIMENTS

Some experiments were carried out on the gyroscope shown in Fig.4. Figure 5 shows the dimensions of the gyroscope, fabricated using a 163° rotated Y-cut $LiNbO_3$ plate, 2mm thick. The admittance characteristics for the perpendicular-field excitation were measured using the drive electrodes and the short-circuited pick-up 1 and pick-up 2 electrodes. The admittance characteristics for the parallel-field excitation were measured using the pick-up 1 and pick-up 2 electrodes, with floating the drive electrode. The resonance frequencies of the two orthogonal thickness-shear modes were plotted against the electrode gap g in Fig. 6. The resonance frequencies of the two thickness-shear modes tend to approach each other with increasing the electrode gap, but cannot be brought into coincidence with each other by adjusting the electrode gap.

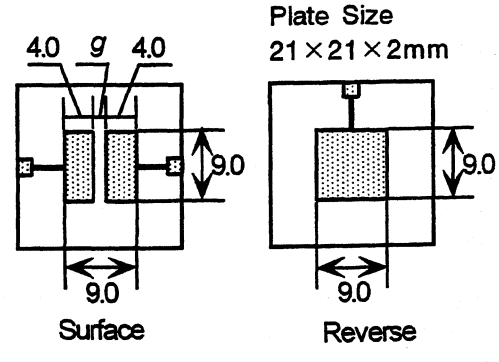


Fig.5. Dimensions of a gyroscope.

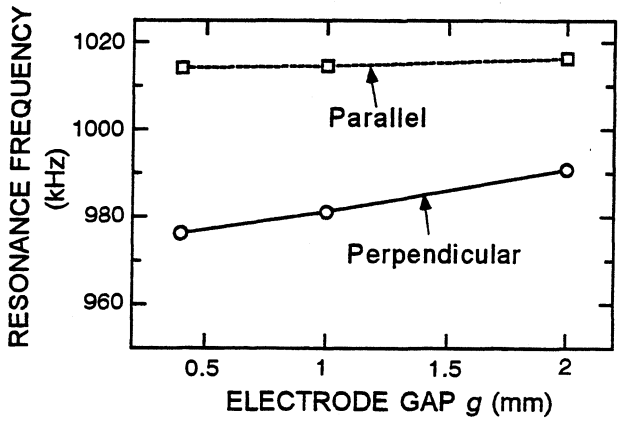
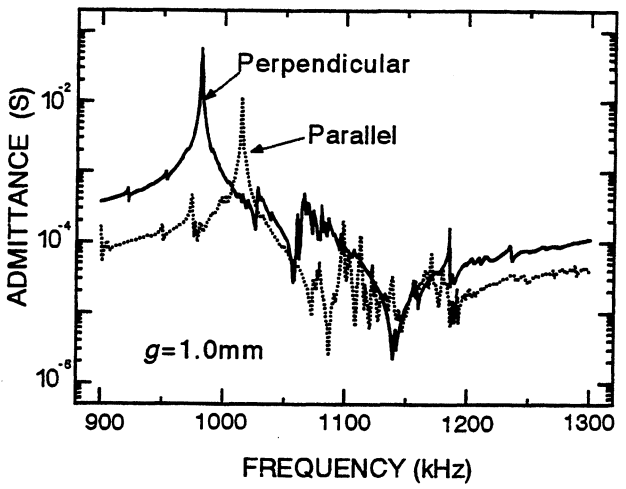


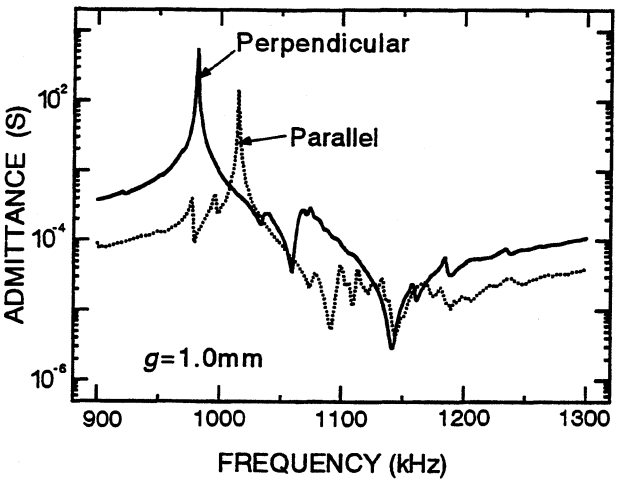
Fig.6. Resonance frequencies as functions of electrode gap g.

The admittance characteristics of the gyroscope with and without the acoustical absorber (silicone rubber) on the plate edge are shown in Figs.7(a) and 7(b), respectively. The absorber has little effects on the Q-factor of the main resonance, while some spurious responses decreased remarkably. This implies that the energy of the two vibration modes might be trapped in the central electroded portion. Some spurious responses still remain in the frequency region higher than the main resonance. These spurious resonances might be trapped-energy inharmonic overtones.

Figure 8 shows the schematic diagram of the driving and detection circuits, which are composed of an oscillation circuit, two current detection circuits, a differential amplifier, and an AC-DC converter. The characteristics of the gyroscope were measured by rotating the gyroscope by an electric motor. The driving voltage and frequency of the vibratory gyroscope were 5Vp-p and 981 kHz, respectively. In Fig.9 the output voltage of the gyroscope is plotted against the angular velocity. The output voltage is seen to increase with increasing angular velocity, although the linearity is not so good.



(a)



(b)

Fig.7. Admittance characteristics of the vibratory gyroscope (a) with and (b) without acoustical absorber at the plate edge.

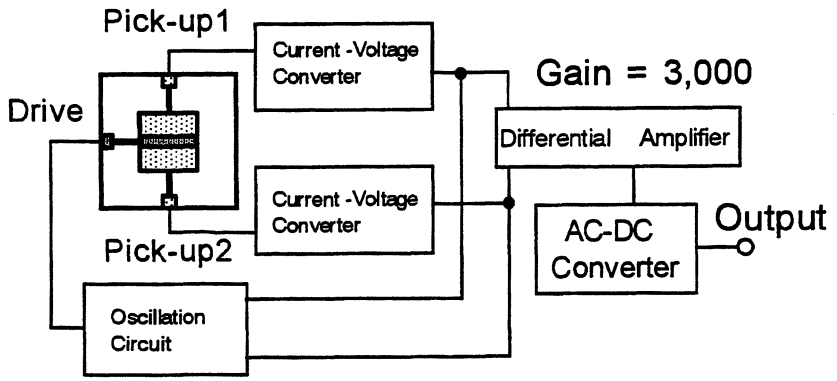


Fig.8. Schematic diagram of the driving and detection circuits of a vibratory gyroscope.

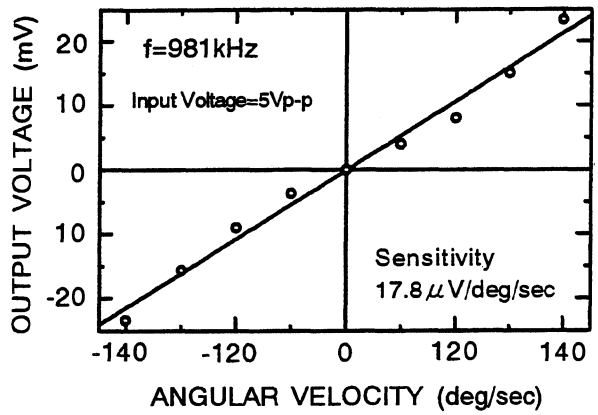


Fig.9. Output voltage vs. angular velocity.

4. CONCLUSION

In this paper, the trapped-energy piezoelectric gyroscope using a rotated *Y*-cut $LiNbO_3$ plate was proposed. This gyroscope is characterized by the easy mount with no need of trimming, the thin structure, and the facility of an easy manufacturing process. It was shown that by using the 163° rotated *Y*-cut of $LiNbO_3$ the resonance frequencies of the two orthogonal thickness-shear modes are close with each other. The basic characteristics of the piezoelectric gyroscope were confirmed, although the sensitivity was not high. By choosing more appropriate rotated *Y*-cuts between 155° −160°, the resonance frequencies of the two orthogonal thickness-shear modes may be brought into coincidence with each other, thereby hopefully yielding a higher sensitivity.

REFERENCES

[1] M. Konno: Jpn. J. Appl. Phys., **27**, p.7 (1987).
[2] W. Shockley, D. R. Curran and D. J. Konval: Proc.17th Freq. Control Symp., p.88 (1963).
[3] M. Onoe and H. Jumonji: J. Inst. Elec. Com. Eng. Jpn. **48**, p.1574 (1965) [in Japanese].
[4] G. M. Reese, E. L. Marek and D. W. Lobitz: Proc. IEEE Ultrasonics Symp., Montreal, Canada, p.419 (1989).
[5] H. Abe, et al.: Jpn. J. Appl. Phys. **37**, p.5345 (1998).
[6] K.Nakamura and M.Ohsaki : Jpn. J. Appl. Phys., **37**, p.2864 (1998).
[7] A. W. Warner, M. Onoe and G. A. Coquin: J. Acoust. Soc. Am., **42**, p.1223 (1966).
[8] T. Yamada and N. Niizeki: J. Appl. Phys., **41**, p.3604 (1970).

1999 Joint Meeting EFTF - IEEE IFCS

INTERNAL STRESS IN SPUTTERED GOLD ELECTRODES AND ITS SIGNIFICANCE FOR QUARTZ RESONATORS

Greger Thornell, Fredric Ericson and Guy Portnoff*

The Ångström Laboratory, Uppsala University, Box 534,

SE-751 21 Uppsala, Sweden

*) Quartz Pro AB, Box 6062, SE-175 06 Järfälla, Sweden

ABSTRACT

Internal stress in magnetron sputtered gold films, as estimated with the cantilever beam deflection method, is correlated to deposition parameters such as pressure, temperature and deposition rate, as well as morphology and the amount of gas entrapped in the film. The compressive stress was found to increase with higher deposition rate, lower temperature or lower pressure which is in accordance with the elsewhere proposed atomic peening model.

1. INTRODUCTION

Quartz is in many aspects a material brought to all desirable perfection and one of the main factors limiting the long term stability of plated resonators is today perhaps the chemical, thermal and mechanical stability of the electrodes. Electrode material is often deposited at non-equilibrium conditions implying a tendency to change with time. By using inert materials like gold the problem is reduced to controlling the internal stresses which, although being easily compensated for initially, acts like a driving force for diffusion, grain growth and dislocation movement. Of course this relaxation changes the balancing stress in the crystal and hence its frequency. Means of minimizing such internal stress are easily obtained by utilising a PVD-method like sputtering with a wide range of adjustable process parameters, compared to evaporation.

2. RESIDUAL STRESS

Residual stress can be divided into thermally induced stress, which is the result of mismatch of coefficients of thermal expansion between film and substrate material together with a deposition temperature differing from the component's normal operating temperature, and intrinsic stress closely related to the microstructure of the film. Hence, intrinsic stress is affected by defects and impurities incorporated in the film during growth, and is primarily governed by the thermal and kinetic energy delivered during deposition. At low temperature and in the absence of particle bombardment, the mobility of the adatoms is rather reduced and the film becomes porous and ends up in a tensile state. If the substrate is heated or bombarded during deposition the resulting microstructure is denser and the evolving stress usually compressive.

There exist quite a few methods to measure thin film stress. In most of them, like X-ray diffraction Raman Spectroscopy and neutron diffraction, the lattice parameters are measured and so it is actually the strain that is registered. Less expensive and faster are the methods where the balancing stress in the substrate, a whole wafer or a thin beam, causes it to bend whereupon the curvature can be recorded by optical microscopy, interferometry or even stylus profilometry. Here, the reason for employing the cantilever beam deflection method was at least fourfold. Primarily, there was a wish to separate the growth related stress from that of the mounting clamps (see Fig. 1), secondly the observation of frequency shifts of real resonators left us expecting a rather low level of stress, thirdly, preliminary experiments with X-ray diffraction and Raman spectroscopy resulted in a very low signal to noise ratio, and finally the many measurements intended over a long time (several months) made this method very attractive.

Assuming a biaxial symmetry, the stress of the film, σ, is related to the coated beam's curvature, R, by Stoney's formula, Eq. 1 [1], where the subscripts s and f refer to substrate and film respectively, and E is the Young's modulus, ν the Poisson's ratio, and t the thickness. Given a small deflection compared to the beam length, the curvature can be substituted by the length of the beam squared, ℓ^2, divided by twice the deflection, 2δ, as shown in the right part of the equation.

$$\sigma_{res} = \frac{E_s}{(1-\nu_s)} \frac{t_s^2}{6t_f R} = \frac{E_s}{(1-\nu_s)} \frac{\delta t_s^2}{3t_f \ell^2} \qquad (1)$$

3. MATERIALS AND METHODS

Beams were fabricated from single crystalline, 130 µm thick, Z-cut quartz wafers by standard lithography followed by wet etching [2]. For easy removal of the samples, each containing a set of four 3.8 mm long cantilever beams, from the wafer, and mounting in a standard blank holder, Fig. 1, a circular trench bridged by two stalks was etched out around each set.

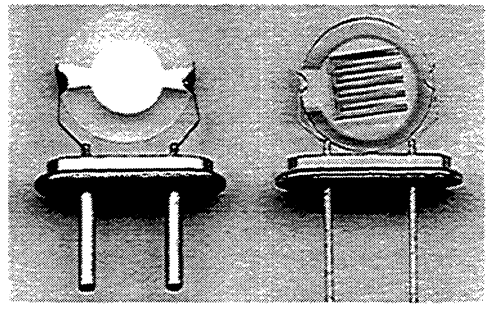


Fig. 1. A mounted Ø 7.62 mm AT-cut 13.2 MHz resonator together with a sample for stress measurements.

The vacuum deposition system used is designed for large scale production of high precision resonators. It consists of four vacuum chambers connected in line, capable of coating 1000 blanks in a cycle after thorough "dry cleaning" of the crystals' both sides The sputter source is a DC planar magnetron type specially designed to prevent electrons from reaching the crystal from the plasma and cause heating.

The following deposition parameters were varied in this study: pressure ($1.0 \cdot 10^{-3}$ and $3.0 \cdot 10^{-3}$ mbar),

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

deposition rate (2, 10 and 50 Å/s), substrate temperature (room temperature and 80°C) and deposition time (to achieve films of different thickness). The parameter set for each cantilever beam sample as well as surface condition are given in Table 1.

Table 1. The surface and deposition conditions for the beam samples.

sample	condition	pressure (mbar)	deposition rate (Å/s)	voltage (V)	current (mA)
I	polished	$3.0 \cdot 10^{-3}$	2	284	10
II	"	"	10	366	60
III	"	"	"	366	60
IV	lapped	"	"	366	60
V	"	"	"	366	60
VI	"	"	"	365	60
VII	polished	"	50	471	210
VIII	"	"	"	471	210
IX	"	$1.0 \cdot 10^{-3}$	2	333	10
X	"	"	"	333	10
XI	"	"	10	383	50
XII	"	"	"	383	50
XIII	"	"	"	383	50
XIV	lapped	"	"	383	50
XV	"	"	"	383	50
XVI	polished	"	50	531	210
XVII	"	"	"	531	210
XVIII*	"	"	"	531	210

*) like sample XVII, but at 80°C

For each deposition at least one cantilever beam sample was used, and for a few key samples (I, VII, IX, XVII, XVIII) some ordinary blanks were coated for Rutherford backscattering spectrometry (RBS) measurements of entrapped argon, as well as for transmission electron microscopy (TEM).

The deflection measurements were done by means of a light optical microscope equipped with a linear displacement transducer. At the highest magnification the depth of focus is below one micrometer.

TEM samples were prepared by immersing coated blanks into concentrated hydrofluoric acid (48% HF) for about 10 s, then moving them to deionised water and scooping up the off floated film on a gold grid. This relatively quick and simple sample preparation allowed a large area to be examined which ensured a representative view of the films.

Ageing was performed first at room temperature (RT) for four weeks and then at 85°C in an ordinary hot oven for 13 weeks. The TEM samples however were only annealed for 20 hours.

3. RESULTS

From the RBS analysis the argon content of samples I, VII, IX, XVII and XVIII as well as the film thickness of all samples could be established, Table 2 and 3 respectively. The accuracy of the amount of argon is ±0.2 atomic percent and the detection limit is 0.2 atomic percent.

The cantilever beam thickness was measured with an accuracy better than 0.1 µm with the help of a scanning electron microscope, Table 3.

The deflection data of each sample's four beams was averaged and used for stress calculation according to Eq. 1 with E_S=78 GPa, and ν_S=0.55 [3]. In most cases the deflection scatter within each set of four beams was about equal to the resolution of the measurement, i.e. 1 µm, which of course implies a relative difference of say 100% at the smallest recorded deflections, and a little more than 1% at the largest. In the chart of Fig. 2 the stress vs. time for sample I, VII, IX, XVII and XVIII is given. This figure serves the dual purpose of showing the typical rate of relaxation, common for all samples, and the number and frequency of the deflection measurements conducted. All samples are characterised by four stress parameters, as deposited stress, σ_{AS}, stress at end of RT ageing, σ_{RT}, the stress after 20 hours of annealing, σ_{20} , and stress at end of annealing, σ_F in Table 3. Following the convention, negative stress values imply compressive stress and positive values tensile stress.

Table. 2. The argon contents of "as deposited" and annealed samples I, VII, IX, XVII and XVIII.

sample	I	VII	IX	XVII	XVIII
as deposited argon-atomic-percentage	<0.2	0.3	<0.2	0.5	<0.2
aged argon-atomic-percentage	0.4	0.5	<0.2	0.3	-

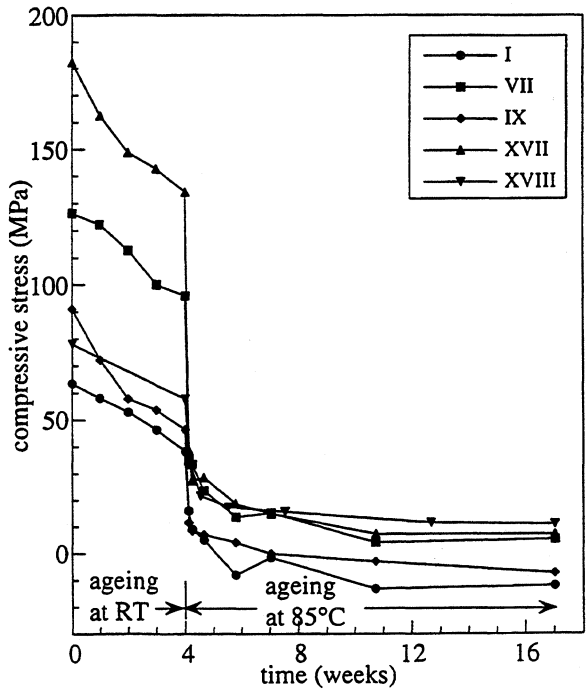


Fig. 2. Stress vs. time for samples I, VII, IX, XVII and XVIII.

In contrast to the RBS samples that were annealed for 13 weeks, the ageing of the TEM samples was interrupted after four weeks of RT ageing and just 20 hours of annealing. This decision was based on the rather drastic stress decrease encountered with the beam samples at that time. The results are shown in Fig. 3, where samples I, VII, IX, XVII and XVIII are shown at equal magnification. Even though the exact grain size couldn't be readily determined, the relative difference could be verified easily from the pictures.

4. DISCUSSION AND CONCLUSIONS

Focusing mainly on the initial stress, σ_{AS}, and evaluating the outcome of Table 3, in terms of compressive

Table 3. The beam and film thickness as measured by means of SEM and RBS respectively, together with the stress at the beginning and end of RT ageing, and after 20 h and 17 weeks of annealing at 85°C, σ_{AS}, σ_{RT}, σ_{20}, and σ_F respectively, for all samples.

sample	I	II	III	IV	V	VI	VII	VIII	IX
beam thickness (μm)	9.5	9.5	11.4	9.4	9.4	9.4	11.4	11.4	11.4
film thickness (Å)	365	520	845	440	605	840	500	485	480
σ_{AS} (MPa)	-63	-71	-74	-91	-69	-70	-126	-123	-91
σ_{RT} (MPa)	-38	-50	-64	-69	-51	-59	-96	-84	-46
σ_{20} (MPa)	-16	-24	-26	-21	-20	-22	-35	-29	-12
σ_F (MPa)	12	-3	-11	6	-4	-7	-6	-7	7
sample	X	XI	XII	XIII	XIV	XV	XVI	XVII	XVIII
beam thickness (μm)	11.4	11.3	11.3	11.3	9.4	9.4	11.3	11.3	5.5
film thickness (Å)	465	420	425	805	410	815	565	550	550
σ_{AS} (MPa)	-85	-97	-100	-139	-113	-169	-151	-182	-78
σ_{RT} (MPa)	-45	-63	-55	-103	-67	-122	-114	-134	-58
σ_{20} (MPa)	-16	-16	-16	-36	-26	-39	-40	-38	-35
σ_F (MPa)	4	3	8	-7	2	-8	-8	-7	-11

stress, the following could be concluded: A lower pressure gives, irrespective of deposition rate, a higher stress, Fig. 4. This is in agreement with the atomic peening model, since a higher voltage is needed to sustain the plasma which also accelerates the charged particles. In addition the longer mean free path implies less energy loss for the ions and neutrals on their way to the substrate. For a given pressure the stress increases with deposition rate, Fig. 4, which intuitively could be interpreted as either a higher accommodation of gas or less time for migration and grain growth during high rate deposition. From samples differing only in deposition time the stress increases with film thickness at least for the lower of the two pressures investigated. This signals stress gradients. Also clearly pronounced with the lower pressure is the higher stress in films deposited on lapped substrates compared to polished.

Finally the large reduction of stress with a higher substrate temperature, sample XVII vs. XVIII, suggests either a more ordered microstructure or less incorporation of impurities with heating.

Comparing these observations with the results from RBS (Table 2), a trend towards higher Ar-contents with higher deposition rate and higher pressure is clearly established. Equally evident is the much lesser presence of Ar, if any, in sample XVIII compared to XVII.

From a vertical comparison of the TEM pictures of Fig. 3, a remarkable grain growth is seen after just 20 hours of annealing for all but the XVIII sample, which agrees well with the fact that the stress of XVIII relaxed relatively little compared to the others - only around 55%, when I, VII, IX and XVII relaxed more than 70% after the same time of annealing as the TEM samples. From a horizontal comparison of the very same TEM photographs, one can see a much coarser microstructure for the slowly grown films. The structural change attributed to higher deposition rate (25 times higher) by far exceeds that for a lower pressure (one third).

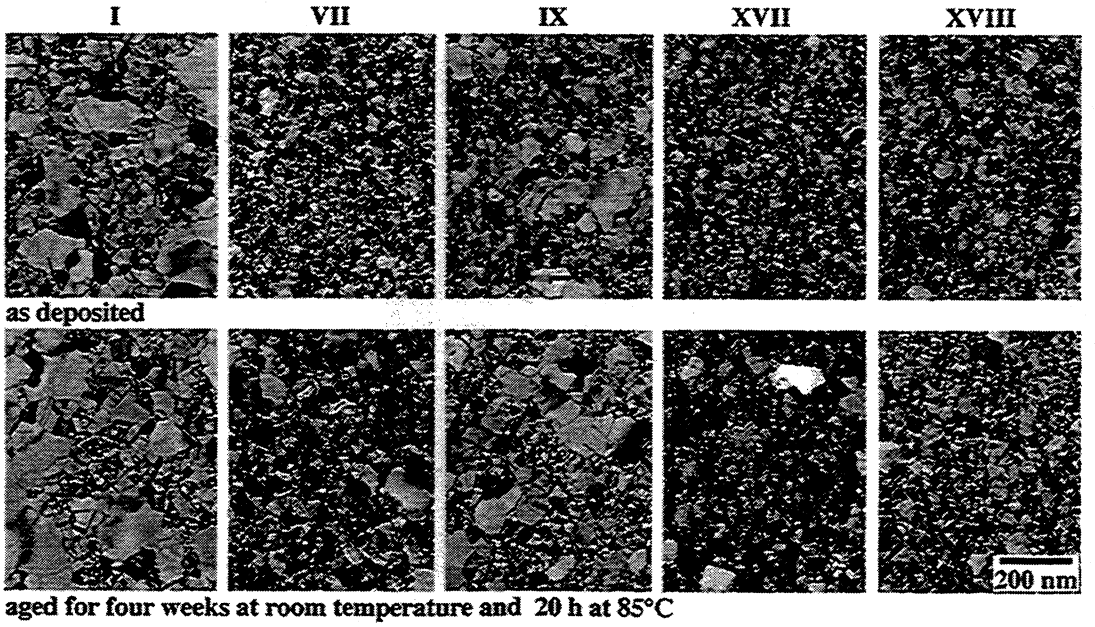


Fig. 3. A matrix of bright field TEM pictures, at the same magnification, showing the grains of as deposited (upper row) and aged (lower row) samples; I, VII, IX, XVII and XVIII.

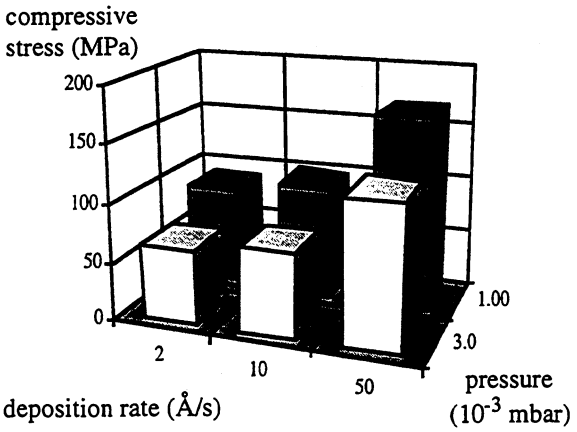


Fig. 4. Compressive stress as a function of pressure and deposition rate for polished, thin coated samples.

With the overall low contents of argon and the significant diversity in microstructure, we are prone to characterise the stress as morphology rather than impurity governed. Also, a decrease in pressure should either lower the compressive strain because of less gas incorporation, or increase the stress because of higher bombardment energy. Judging for instance from Fig. 4 the former alternative ought to be ruled out.

A legitimate question is whether there is a thermally induced component in the residual stress measured here. In contrast to the calculation of the total stress, this estimation requires some knowledge of the film properties namely the coefficient of thermal expansion, α, and the Young's modulus. In literature and handbooks the value for gold ranges from 14 to 16.2 ppm/°C [4,5], while that of quartz along the y-axis is approximately 13.7 ppm/°C [3] at room temperature. So, forming the product of the biaxial Young's modulus for gold, 76 GPa [4], the difference in thermal expansion and temperature difference, the tensile stress ought to be at worst around 16 MPa for a temperature difference of 85 degrees. But, actually the value may equally well be zero.

From Table 3 a tendency towards faster room temperature relaxation with higher initial stress could be perceived. In the view of the much finer grains structure accompanying the higher stress, the relief mechanism is believed to be stress promoted grain boundary diffusion as discussed by for instance Hieber et al. [6]. Given the many obstacles provided by the abundant grain boundaries, dislocation movements are not likely to contribute much to the relaxation however. Migrating point defects on the other hand, are often responsible for stress relaxation.

Returning to the fundamental question of this work, the effect of stress on the resonator frequency shift, $\Delta f/f$, should be calculated. In a work of E. P. Eernisse [7] this question has been addressed and Eq. (2) suggested;

$$\frac{\Delta f}{f} = K \frac{2\sigma_f t_f}{t_q} \qquad (2)$$

where K is a stress coefficient calculated and given in the same paper, and the quotient the average stress bias for blanks coated with equally stressed and thick films on both sides. For simplicity taken as uniform here, σ_f is the film stress, and t_f and t_q is the thickness of the film and quartz respectively. Taking the stress data from sample XVII (Table 3), representing the most commonly deposited electrode, and applying them to a 100 µm thick AT-cut quartz plate with 400 Å thick electrodes, the initial stress of -182 MPa would induce a tensile stress in the resonator resulting in a downward frequency shift of the order of 4 ppm (taking K to be $2.75 \cdot 10^{-11}$ Pa^{-1} [7]) as compared to a resonator with unstressed electrodes. After four weeks at RT this will decrease to approx. 3 ppm. Both the calculated initial frequency shift and its time dependence are in good agreement with our experience from measurements conducted on real resonators.

6. SUMMARY

In conclusion sputter deposition parameters have been altered in this study to change the residual stress in gold electrodes on quartz. In all cases the stress was found to be compressive and a transition to tensile stress was experienced only with annealing. From the observation that the stress increased, in absolute terms, with lower pressure, and that very little argon was found in the films, we are prone to blame the stress on the elsewhere proposed atomic peening mechanism (see e.g. [8]) and not on incorporation of sputter gas. On raising the deposition rate both an increasing compressive stress and a finer grain structure could be detected. With an exception for the sample prepared at elevated temperature, the amount of trapped argon increased with higher deposition rate. When translated to frequency shifts of ordinary AT-cut resonators, great agreement between stress relaxation and normally observed initial frequency drift could be verified.

7. ACKNOWLEDGEMENTS

We wish to express our sincere gratitude to Mr Johan Öhrmalm, Dr Christer Hedlund and Prof. Jan-Åke Schweitz for experimental assistance and/or stimulating discussions, and the Swedish Board for Technical Development (NUTEK) for its finacial support.

8. REFERENCES

[1] G. G. Stoney, "The Tension of Metallic Films Deposited by Electrolysis", Proc. Royal Soc., pp. 172-75(1909).

[2] C. Hedlund, U. Lindberg, U. Bucht, and J. Söderkvist, "Anisotropic Etching of Z-cut Quartz", J. Micromech. Microeng., vol. 3pp. 65-73 (1993).

[3] J. C. Brice, "Crystals for Quartz Resonators", Rev. of Mod. Phys., 57, pp. 105-146(1985).

[4] M. Zafrany and F. Meyer, "Stability of Cr/Au Layer on Quartz Substrates", Proc. of European Freq. and Time Forum, 1993.

[5] J. A. Thornton and D. W. Hoffman, "Stress-Related Effects in Thin Films", Thin Solid Films, 171, pp. 531(1989).

[6] H. Hieber and K. Pape, "Umkristallisation dünner Goldschichten", Zeitschrift für Metallkunde, 70, pp. 459-466(1979).

[7] E. P. Eernisse, "Quartz Resonator Frequency Shifts Arising from Electrode Stress", Proc. of The Ann. Freq. Control Symp., pp. 1-5(1975).

[8] H. Windischmann, "Intrinsic Stress in Sputter-Deposited Thin Films", Critical Reviews in Solid State and Materials Science, 17(6), pp. 547-96(1992).

1999 Joint Meeting EFTF - IEEE IFCS

DIRECT BONDING FOR TRUE ALL QUARTZ PACKAGE AND NEW RESONATOR DESIGNS

Örjan Vallin* and Pelle Rangsten
Uppsala University

The Ångström Laboratory, P.O. Box 534, SE-751 21 Uppsala, Sweden

ABSTRACT

We present the direct bonding method of mono-crystalline quartz. Demonstrator devices were designed to show the process compatibility of etching, patterning of electrodes and direct bonding for a true All Quartz Package. The ability to unite different crystallographic directions makes interesting laminates.

1. INTRODUCTION

When joining two surfaces to each other, any adhesive, solder or other intermediate layer will introduce uncertainty in the behaviour of the united structure. If one can obtain hermetically sealing without the need of any intermediate layer, many unwanted effects will vanish. The method is called direct bonding. The direct bonding was originally developed in microelectronics for so-called silicon-on-insulator wafers [1], or as a replacement for epitaxial silicon growth [2]. Today the technique is much used also in micro system technology both as an encapsulation technique and as an adding technique. The direct bonding has been shown to work for a variety of materials [3-5]. We have shown that the method is applicable also to mono-crystalline quartz [6]. Montress and Parker have done extensive work on All Quartz Package (AQP) for Surface Acoustic Wave (SAW) filters using glass frit as sealing material [7, 8]. Ando et al have used a direct bonding method to make a glass package [9]. Our aim is to use direct bonding of the quartz itself to make a true AQP, and eventually to integrate several components in an All Quartz System (AQS).

In the case of packaging of quartz resonators and filters the instability of the glue will affect the long-term frequency stability. The thermal mismatch of the layer and the quartz will lead to temperature dependence. The exclusion of adhesion layer eliminates effects as ageing, creep, out-gassing and temperature instability of the intermediate layer. Variations in layer thickness and thermal mismatch between the layer and substrate are prevented. The direct bonding method is also attractive for its simplicity as no externally applied forces or electrical field are needed. Wet chemical cleaning of the mirror-polished surfaces and a subsequent annealing after bringing the surfaces together is all that is needed to unite the surfaces.

The bonding method should be compatible with preceding and subsequent process steps, i.e. metallisation, etching, bonding, and forming of electrodes. Direct bonding is indispensable for a true all quartz package.

Wafer bonding is just the first step to an all quartz package, namely the basic sealing technique. The second step is the combination of a structuring technique together with the sealing technique. The third step is the need for electrical connections into the sealed cavity.

We fulfil the first and second step by combining wet chemical etching with direct bonding. Hermetically sealed cavities are presented in previous work, where the hermeticism of the cavities was confirmed in a temperature cycling test [6].

In this derivative work, laminates of different cuts are manufactured and demonstrator devices are designed to show the process compatibility of etching, patterning of metal electrodes and direct bonding for a true AQP.

2. BONDING THEORY

Direct bonding is alternative known as fusion bonding, thermal bonding or wafer bonding. The requirements for direct bonding are smooth, planar surfaces without particle contamination. If this is fulfilled two mirror-polished surfaces adhere to each other when brought together at room temperature without the use of any adhesive or outside applied force or electric field. When the surfaces are put together, the contact area will expand at a certain rate, called the contact wave velocity. If the wave front proceeds across the wafer without any help, like applying a pressure, the bonding is referred as spontaneous. The bonding starts spontaneous or can be initialised by applying a slight pressure. The bonded area then spreads over the wafers in a few seconds. A proceeding anneal at elevated temperature strengthen the bond over the interface.

The quartz surface is hydrophilic due to adhered hydroxyl (OH) groups. To achieve a higher degree of hydrophilicity, the surface can be wet chemical treated in for instance ammonia or nitric acid. Most probably the bonding mechanisms of quartz are similar to the case of hydrophilic silicon or silicon dioxide [10].

When brought together, the surfaces are attracted to each other by hydrogen bonds, Fig. 1a. During annealing stronger covalent bonds, over-bridging the interface, are formed, Fig 1b.

The bond strength can be measured by the crack opening method [11].

The phase transition at 573°C limits the temperature span for processing of quartz, however Tong *et al* claims that longer annealing times can compensate for lower annealing temperatures [10].

* Corresponding author. E-mail: Orjan.Vallin@Angstrom.uu.se. Tel. +46 18 4717254, Fax: +46 18 555095

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

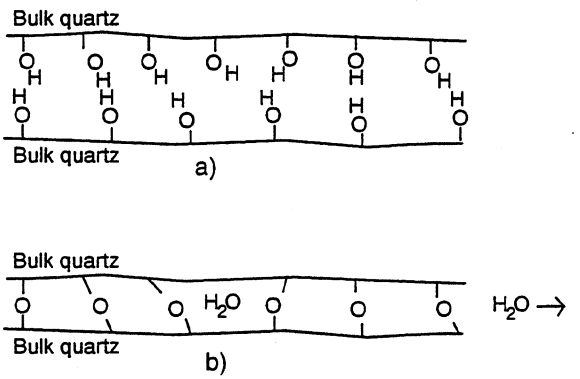


Fig 1. a) Room temperature bonded quartz surfaces attracted to each other by hydrogen bonds. b) Strong covalent bonds at the interface after annealing.

3. MATERIALS AND EQUIPMENT

We used 1.5'' × 1.5'' polished square wafers of α-quartz (MicroCrystal, Grenchen, Switzerland). The wafers were of different batches, purchased at different times. They were of various cuts and thickness: Z (200 µm), X+1°50' (130 µm) and AT (130 µm). All wafers were mechanically polished by the manufacturer to an optical grade and some wafers were coated with a Cr-Au film.

For one experiment we used circular quartz blanks in two sizes: 8.89 mm diameter, 160 µm thick, and 7.62 µm in diameter, 110 µm thick (Quartz Pro AB, Järfälla, Sweden).

The experiments were carried out in a clean-room equipped for semiconductor processes with controlled particle concentration, temperature and air humidity. We used standard laboratory equipment such as simple holders and tweezers when handling the wafers, and apparatus such as photoresist spinner, contact mask aligner, temperature controlled wet etching equipment, anneal furnace, chemical mechanical and a polishing machine for processing.

4. EXPERIMENTS

Following experiments were carried out:

1. Plain *as received* wafers of the same cut was bonded to each other and annealed at 200, 300, 400 and 500°C.
2. Plain *as received* wafers of different cut, as well as wafers of the same cut, but rotated in respect to each other, was bonded and annealed at 500°C.
3. Demonstrator structures for AQP were made. One SAW filter demonstrator made out of two quartz wafers and one web-moated AT-cut resonator demonstrator in a three wafer configuration. See Fig. 2 for schematic drawings and the work of H. Rapp and K. Hjort (this conference) for the web-moated resonator design.
4. Circular quartz blanks of different thickness were bonded to each other.

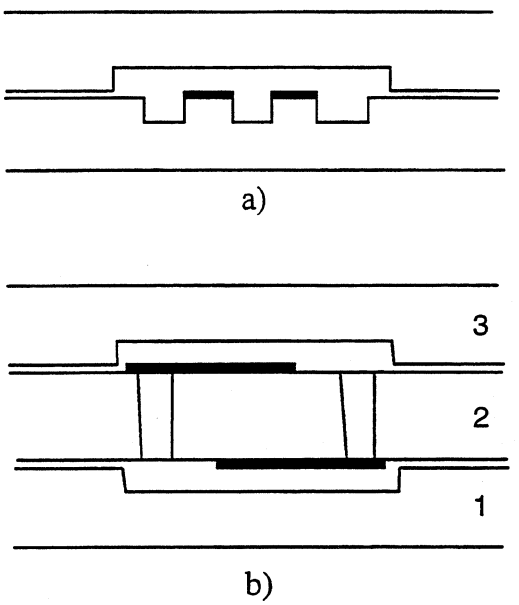


Fig. 2. Schematic drawings of the two true All Quartz Package demonstrators. a) the two-wafer SAW filter demonstrator, b) the three-wafer web-moated resonator demonstrator.

The experiments involved these process steps:

Masks for demonstrator structures were designed. The wafers were covered by photo resist (Shipley 1813) by spinning, and they were exposed in a contact mask aligner (Karl Süss MBJ21). The gold was etched in a KI solution, and the chromium in a commercial Cr-etch (Merck). The resist was removed in acetone and isopropanole, and for some cases in H_2SO_4:H_2O_2 (1:1), followed by a rinse in de-ionised (DI) water. The quartz was etched in HF:NH_4F (2:3) at 80 °C. The electrodes were patterned in a second lithography step.

In the hydrophilisation step the wafers were submersed in boiling HNO_3 (69%) for ten minutes and rinsed in hot DI water. They were blown dry one by one in nitrogen and immediately contacted in air at room temperature. Bonded areas are easy to distinguish from unbonded areas, i.e. voids, with a simple visual inspection.

All evaluation of the bond strength were performed by using the crack opening method and a 50 µm thin blade.

The bonded wafers were annealed in a vertical annealing furnace. The temperature was ramped up to anneal temperature at a rate of 10°C/min, annealed for one hour and ramped down to room temperature at 10°C/min.

5. RESULTS AND DISCUSSION

Pairs of AT-cut and pairs of X+1°50'-cut bonded spontaneously. They showed an increasingly bond strength with higher annealing temperatures, see Fig. 3.

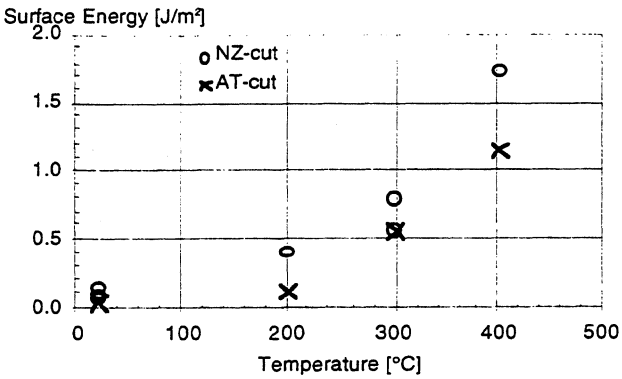


Fig. 3. Measured surface energies vs. annealing temperature for AT-cut and X+1°50'-cut (denoted NZ). Annealing time 1 hour. Reprinted from [6].

Annealing at 500°C gave too strong bonds to be evaluated by crack opening. The wafers shattered when inserting blade.

Pairs of as-received wafers, 130 µm thick, bonded spontaneously, independent of the cut or rotation of each wafer. This suggests that the bonding can be achieved for arbitrary cuts and directions. We have manufactured laminates of different cuts and of the same cut but with one wafer rotated or flipped. Oscillators could thus be formed, with different crystallographic orientation in different parts [12].

One batch of wafers, Z-cut of 200 µm thickness, did not bond as received. After chemical–mechanical polishing, the wafers bonded. We believe that the ability to bond is a combination of the surface roughness and the stiffness of the wafers. A similar behaviour was experienced in the three-wafer AQP demonstrator, where wafer 1 and 2 bonded spontaneously, but wafer 3 bonded only after pressing every area with the tweezers. In the same way, only the thinnest of the circular blanks bonded. A possible solution for better three-wafer packages is to thin down the wafers and hence decrease the stiffness.

The room temperature bonding of the SAW demonstrator was successful, but during the annealing the wafers separated except in a few spots. This might be due to remaining water in the cavities, evaporating during the annealing.

Although the direct bonding is fully compatible with metallization and etching, great care must be taken not to ruin the bondability of the surfaces. The metal covered wafers are preferable as the surfaces are protected during processing. When stripping the bond areas one must make sure that no residual metal is left, as it will result in unbonded areas.

The demonstrator devices were designed to show the process compatibility of etching, patterning of electrodes and direct bonding. The next problem to be solved for a true AQP is how to connect the electrodes from the outside. Suggested solutions are solder bumps in via holes [9] or ion implantation to obtain buried conductors [13].

6. SUMMARY

Direct bonding is a simple and feasible technique for encapsulation of oscillators and filters. However the electrical connection remains to be solved.

As long as the surfaces are plane, sufficiently polished and free of particles and other contamination, one is free to combine any crystallographic direction. New resonator designs could be made with different crystallographic orientation in different parts of the oscillators.

7. ACKNOWLEDGEMENTS

Håkan Rapp is acknowledged for lending us the design of the web-moated AT-resonator.

The Joint Meeting organisation is acknowledged for travelling funds.

8. REFERENCES

[1] J. B. Lasky, "Waver bonding for silicon-on-insulator technologies," *Applied Physics Letters*, vol. 48, pp. 78-80, 1986.

[2] M. Shimbo, K. Furukawa, K. Fukuda, and K. Tanzawa, "Silicon-to-silicon direct bonding method," *Journal of Applied Physics*, vol. 60, pp. 2987-9, 1986.

[3] A. Namba, M. Sugimoto, T. Ogura, Y. Tomita, and K. Eda, "Direct bonding of piezoelectric crystal onto silicon," *Appl. Phys. Lett.*, vol. 67, pp. 3275-6, 1995.

[4] R. W. Bower, M. S. Ismail, and B. E. Roberds, "Low temperature Si_3N_4 direct bonding," *Applied Physics Letters*, vol. 62, pp. 3485-7, 1993.

[5] K. Hjort, G. Thornell, J. A. Schweitz, and R. Spohr, "Quartz micromachining by lithographic control of ion track etching," *Applied Physics Letters*, vol. 69, pp. 3435-6, 1996.

[6] P. Rangsten, Ö. Vallin, K. Hermansson, and Y. Bäcklund, "Quartz-to-quartz direct bonding," *Journal of the Electrochemical Society*, vol. 146, pp. 1104-5, 1999.

[7] T. E. Parker, J. Callerame, and G. K. Montress, "A new all quartz package for SAW devices," in Proc. of the 39th Ann. Frequency Control Symp., 1985.

[8] G. K. Montress, T. E. Parker, and D. Andres, "Review of SAW oscillator performance," presented at 1994 IEEE Ultrasonics Symposium Proceedings 1994.

[9] D. Ando, K. Oishi, T. Nakamura, and S. Umeda, "Glass direct bonding technology for hermetic seal package," in Proc. of The Tenth Ann. Int. Workshop on Micro Electro Mechanical Systems An Investigation of Micro Structures, Sensors, Actuators, Machines and Robots 1997.

[10] Q. Y. Tong and U. Gosele, "A model of low-temperature wafer bonding and its applications," *J. of The Electrochem. Soc.*, vol. 143, pp. 1773-9, 1996.

[11] W. P. Maszara, G. Goetz, A. Caviglia, and J. B. McKitterick, "Bonding of silicon wafers for silicon-on-insulator," *Journal of Applied Physics*, vol. 64, pp. 4943-50, 1988.

[12] T. Ichinose, Y. Ishiai, M. Hatanaka, J. Yukawa, S. Kawasaki, and T. Uemura, "Quartz Angular Rate Sensor for Automotive Application," SAE Int. Cong. and Exp., Detroit, Michigan, 1999.

[13] J. S. Danel, P. Martin, M. Dufour, A. Ermolieff, S. Marthon, F. Pierre, and M. Dupuy, "Enhancement of quartz electrical conductivity by ion implantation," in Proc. of the 1993 IEEE Int. Frequency Control Symp., 1993.

BATCH FABRICATION OF AT-CUT CRYSTAL RESONATORS UP TO 200 MHZ.

Christian Wüthrich, Silvio Dalla Piazza*, Urs Rüedi*, and Bruno Studer*

Asulab, Rue des Sors 3, CH-2074 Marin, Switzerland
*Micro Crystal, CH-2540 Grenchen, Switzerland

ABSTRACT

AT-cut High Frequency Fundamental quartz resonators have been fabricated up to 200MHz. Specific problems relating to this range of frequency like the fabrication process or the design of the electrodes are presented. Finished devices have a resistance lower than 20 Ohms, exhibit only few weak spurious modes and are stable during an accelerated ageing measurement.

1. INTRODUCTION

AT-cut High Frequency Fundamental quartz resonators (AT-HFF) are attractive devices for high speed and mobile communications. The interest lies in their good thermal stability and the possibility to achieve easily high frequency resonators with high quality factor. The interest in AT-HFF is easily understood if one remembers that this kind of resonator has less inharmonic spurious modes and a much higher pullability due to a higher motional capacitance than their overtone counterparts.

Several groups are working in the field of AT-HFF [1,2] and all of them have a similar approach. A slice of quartz, with both faces polished at a high degree of parallelism, is etched in its central area until the desired thickness is reached. The frequency of the resonator is inversely proportional to the thickness of the membrane. Two different techniques are used for thinning the slice. Wet chemical etching in a buffered HF solution is a well adapted method for batch production while ion milling offers a high precision in the control of the frequency.

2. THEORETICAL OVERVIEW.

2.1 Confinement of the wave within the plane.

The thickness shear distortion of an AT resonator can be considered as the sum of a thickness shear wave and a thickness twist wave [3]. In the case of a purely thickness shear wave, a resonance frequency f_r can be determined. The thickness twist can propagate only for frequencies higher than f_r.

The resonance frequency of the thickness shear mode is lowered by the electrodes, mostly due to mass loading. One can get two different cut-off frequencies for the thickness twist wave: f_e and f_s which are for the electroded and the surrounding area respectively. f_e being smaller than f_s, it is possible to choose a frequency f_r in such a way that: $f_e < f_r < f_s$. In this case the wave with a frequency f_r will propagate within the electroded area but will not be able to go into the surrounding area.

By solving the wave equation, a propagation vector can be determined for the thickness twist wave propagating within the electroded area, whose amplitude is:

$$\zeta_e(f) = \frac{\pi}{d}\left(\frac{c'_{66}}{c'_{55}}\right)^{0.5}\left[\left(\frac{f}{f_e}\right)^2 - p^2\right]^{0.5} \qquad (1)$$

Where d is the thickness of the quartz resonator, c'_{66} and c'_{55} are the elastic constants for the AT-cut quartz and p is the order of the thickness shear mode.

While in the surrounding area only an evanescent wave is present with an extinction coefficient.

$$\gamma_s(f) = \frac{\pi}{d}\left(\frac{c'_{66}}{c'_{55}}\right)^{0.5}\left[p^2 - \left(\frac{f}{f_e}\right)^2\right]^{0.5} \qquad (2)$$

In the case of a square electrode, one can separate the solution along the two orthogonal axis and for each axis four boundary conditions are given by the continuity of displacement and continuity of shear stress at the interface. Standing waves satisfying the boundary conditions are given by the equation:

$$\tan(a\zeta_e(f)) = \frac{\gamma_s(f)}{\zeta_e(f)} \qquad (3)$$

Where a is the lateral dimension of the electrode.

It can be calculated that if the lowering of the cut-off frequency in the electroded area is excessive (i.e. thick electrode) many frequencies will satisfy the standing wave equation.

2.2 Design of the electrodes.

Usually, in an AT-quartz resonator, we want a single mode resonance with a high quality factor and no spurious modes. A high quality factor is easily achieved using a thick electrode leading to a good energy trapping under the electrode as well as low resistance of the electrode. However, the thickness of the electrode is limited if only one standing wave mode has to be obtained. The optimum thickness is achieved when the second standing-wave mode is close to be confined.

Excessive confinement can be avoided by using a bottom electrode much greater than the top electrode. The excited area is given by the overlap between the two electrodes and the thickness twist wave is only confined by the step high of one electrode.

2.3 Figure of merit of the electrodes metal.

The conductivity of the material of the electrodes must be as high as possible. On the other hand, its specific mass must be low, allowing the deposition of a relatively thick film in order to get a low sheet

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

resistance without a large decrease of the cut-off frequency due to mass loading. It is thus possible to define a figure of merit of the electrode which is the ratio of the conductivity over the specific mass. This figure of merit is given in table 1 for different metals.

Metal	Conductivity 10^7S/m	Density kg/m^3	Fig. of merit 10^3S*m^2/kg
Aluminium	3.65	2700	13.50
Silver	6.21	10490	5.92
Copper	5.88	8930	6.59
Gold	4.55	19300	2.36
Berylium	3.08	1850	16.60

Table 1: Density and conductivity of various metals as well as their figure of merit as electrode material for high frequency shear mode oscillators.

Beryllium has the best figure of merit but cannot be used because of its toxicity and lack of stability in air. We determined that aluminium is the best material for electrodes, with a figure of merit twice as good as copper and five times better than gold.

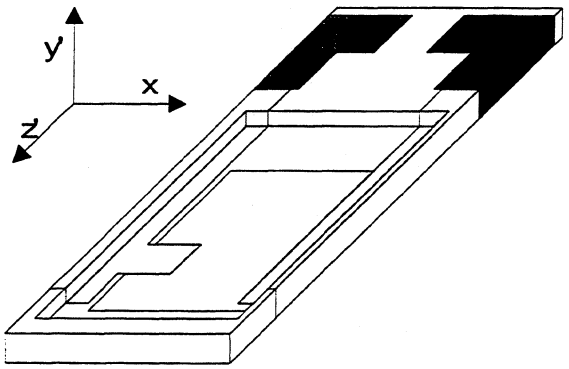


Figure 1: Schematic diagram of a 155 MHz AT-HFF crystal resonator with a length of 5.8mm, a width of 1.9mm and a thickness of the membrane of 11μm. The gold contacts are shown in dark grey and the aluminium electrodes are depicted in light grey.

3. DESIGN OF THE HFF RESONATOR

The inverted mesa structure of the HFF resonators is shown in fig. 1. The resonator consists of a slice of AT-cut quartz with a nominal thickness of 125 μm. The thickness is reduced in an area of about 4.5mm long and 1.7mm width down to 11μm for 155MHz resonators or 8μm for 200 MHz devices. The bottom electrode is plated on a large area while the top electrode has the shape of a strip overlapping the bottom electrode in a 500 μm by 500 μm square. The design adopted for the electrodes allows a high quality factor to be achieved because it minimises the series resistance of the electrodes and allows a good control over the confined mode.

4. FABRICATION OF THE RESONATORS

The process is derived from the fabrication of the 50MHz AT-HFF resonators [4]. The process flow chart is given in Fig. 2.

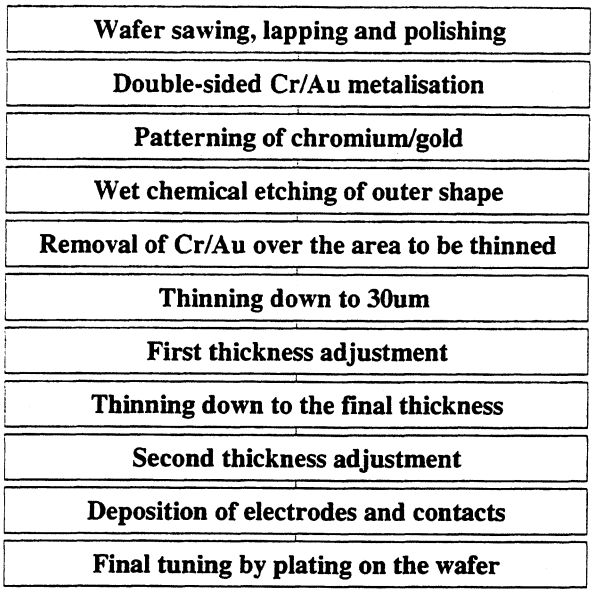


Figure 2: Block diagram of the wafer batch manufacturing process.

In a first step the AT-cut square wafers (38.1mm aside) are prepared by successive sawing lapping and polishing operations. The initial flatness and parallelism of the wafer are important because it will not be futher adjusted during manufacturing. Wafers are then coated with chromium and gold on both sides by vacuum evaporation. The Cr/Au layer is patterned by wet chemical etching and is used as a mask for the deep etching in HF/NH_4F at 85°C needed for micromachining the individual resonators on the wafer. At the end of this step, chromium and gold are removed, on each side, over the area where the thinning has to occur. Thinning is performed in a specially designed solution allowing to obtain a reasonable etching speed and a smooth surface at the end of the etch process. The etching is stopped when the blanks have a thickness of about 30μm.

The thickness of the blanks exhibits at that stage a discrepancy due to the waviness present at the end of the polishing process. The thickness spread is nearly constant, which means that the relative spread (in percents of the centre frequency) increases during the etch process. Etching in one step to the desired thickness would for this reason lead to an unacceptable frequency spread and even to measurement difficulties. The resonance frequency of the blanks is measured using a CNA 300 [5] in transmission mode[6].

The tuning of the blanks is made in three etching steps with different etching times. Once the blanks have been tuned, a second deep etching brings them close, but slightly under the frequency of the final devices. At

that stage a second cycle of tuning etch is used to diminish the residual dispersion in frequency.

Aluminium electrodes are deposited by vacuum evaporation through shadow masks made out of quartz. Chromium and gold electrical contacts are deposited by vacuum evaporation in order to obtain a reliable and stable contact with the package. Final tuning is performed in an automatic wafer plating equipment

5. CHARACTERISTICS OF THE RESONATORS

5.1 Electrical parameters.

The typical performances of 320 measured devices working at 155.52 MHz are summarised in table 2. The measured devices have 0.5 mm square electrodes of 50nm thick aluminium.

	Average	Std. error
Series resistance R_s [Ω]	15.1	4
Motional capacitance C_1 [fF]	4.19	0.5
Motional inductance L_1 [mH]	0.25	0.01
Shunt capacitance C_0 [pF]	2.3	0.1
Quality factor Q	16500	2000

Table 2: Typical performances of the 155.52 MHz resonators with standard errors.

The standard error on the parameters is only 10% of their average value, which shows the process reproducibility. The ratio C_0/C_1 is 550 which allows a good pullability.

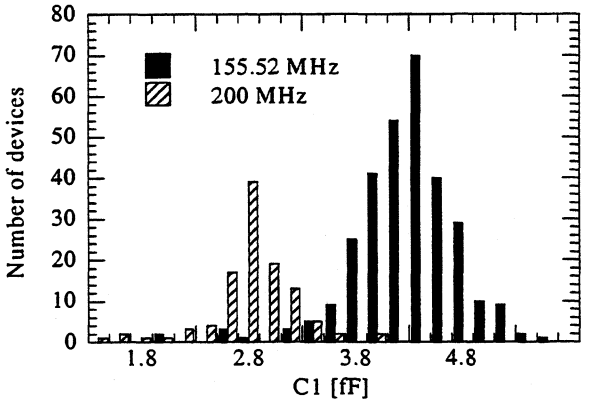


Figure 3: Histogram of the motional capacitance of 320 devices working at 155.52 MHz and of 128 devices working at 200MHz.

The distribution of the motional capacitance is shown in fig. 3. The values of the motional capacitance are well centred for both frequencies, showing that the mode of the resonators is well defined by the electrodes. No resonator with partial activity could be found.

The distribution of the frequencies is shown in fig. 4. before and after wafer tuning. The measurement of the blanks is affected by several parameters. For example the fact that the probe used to measure the blanks does not exactly match the area of the electrodes. This explains the frequency spread observed after the deposition of the electrodes. The standard error of the frequency can be reduced by a factor 40 after the wafer tuning. No significant change has been observed on the resonator parameters (C_1, Q, ...) after tuning.

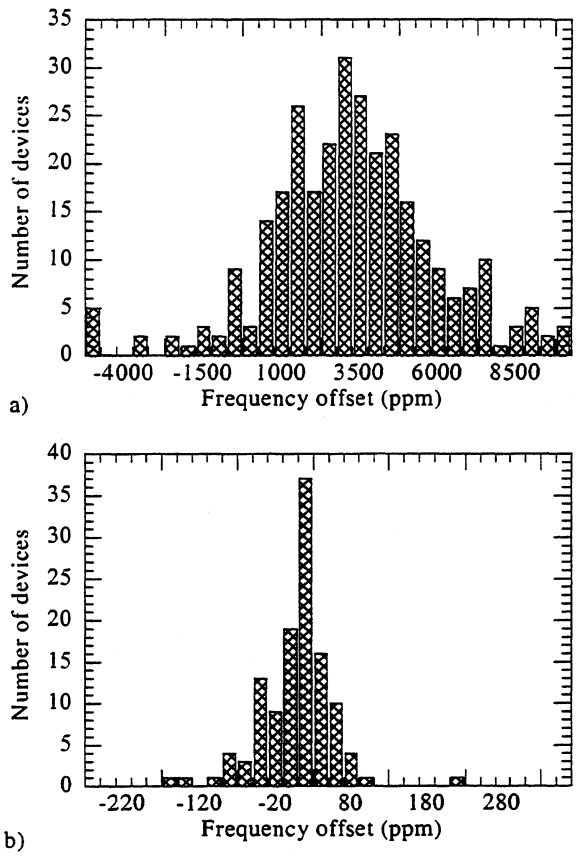


Figure 4: Dispersion of the frequency for devices working at 155.52 MHz. a) Histogram of the frequencies for the whole batch after the deposition of the electrodes. b) Histogram of the frequencies for two wafers after tuning.

5.2 Inharmonic modes.

For the dimension and thickness of the electrodes of the 155MHz resonators, the equation 3 allows to calculate that one inharmonic mode is present along the x axis and another one along the z' axis. Both modes are located 440 kHz above the fundamental mode. Statistical evaluation of the frequency of the strongest inharmonic modes respective to the fundamental at 155.52 MHz has been made. Spurious modes appear frequently at 380kHz and 800kHz above the fundamental. Considering the accuracy of the electrode geometry, the agreement with the calculated frequency offset is judged reasonable. The mode at 800kHz is probably due to an oscillation on the combination of the first inharmonic along the x and z' axis.

5.3 Thermal characteristics.

AT-cut shear mode quartz oscillators exhibit a frequency shift versus temperature which can be approximated by a linear and a cubic term as expressed by the following equation.

$$\frac{df}{f_0} = \alpha(T - T_i) + \gamma(T - T_i)^3 \qquad (4)$$

Where df/f_0 is the frequency shift expressed in ppm respectively to the frequency obtained at the temperature of the inflection point T_i. α is the linear coefficient , γ is the cubic coefficient and T is the actual temperature. The frequency of a good quality resonator will exhibit no activity dips when its temperature is changed.

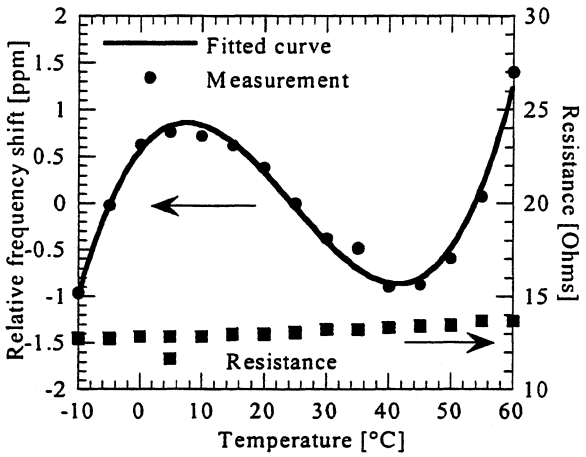


Figure 5: Shift of the resonator frequency versus the temperature for a typical device at 155.52 MHz. The resistance R_1 of the oscillator exhibits a slight increase when the temperature is raised.

The frequency versus temperature of a typical resonator is shown in fig. 5. The frequency is within a 3ppm band over the temperature range (-10°C to 60°C). The measured values can be fitted within a fraction of ppm using the following coefficients: α=-7.6*10^{-8}/°C, T_i=24.5°C , γ=8.85*10^{-11}/°C^3.

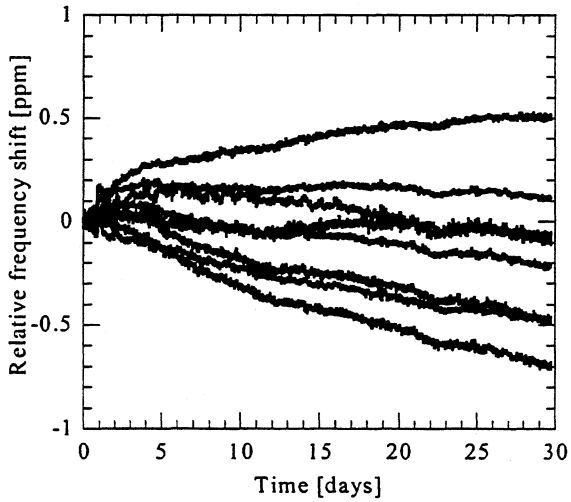


Figure 6: Drift of the frequency of 8 resonators working at 155.52 MHz with aluminium electrodes, placed in an oven at 85°C.

5.4 Accelerated ageing.

Eight 155.52MHz resonators with aluminium electrodes have been aged during 30 days at 85°C. The devices were continuously oscillating and measured at the ageing temperature. Figure 6 shows the relative frequency drift versus time, one day after temperature stabilisation.. The samples exhibit an increase of the frequency at the start of the experiment followed by a tendency to lower their frequency. All samples remain within 1.5ppm of the initial frequency.

6. CONCLUSION

AT-cut quartz resonators working at 155.52 MHz and 200 MHz have been fabricated. The dispersion of the parameters of the resonators shows that the process is well reproducible. The frequency shift measurements as well as the accelerated ageing test show that our resonators have characteristics which are as stable as devices working at much lower frequencies.

7. ACKNOWLEDGEMENTS

We would like to thank Joseph Ricchiuto for his technical support as well as Emil Zellweger for designing the electronic hardware used in the accelerated ageing measurement.

8. REFERENCES

[1] Masako Tanaka, Takayuki Ugajin, Nobushige Araki, and Yoshimasa Oomura, "The Resonating Properties of a Crystal Unit of Fundamental Frequency 150 MHz", Japanese Journal of Applied Physics, vol.36, pp.3022-3027, May 1997.

[2] Osamu Ishii, Hirokazu Iwata, Makoto Sugano, and Tsuyoshi Ohshima, "UHF AT-cut Crystal Resonators Operating in the Fundamental Mode", in Proceedings of the 1998 IEEE International Frequency Control Symposium, 1998, pp.975-980.

[3] W. Schockley, D. R. Curran, and D. J. Koneval, "Trapped-Energy Modes in Quartz Filter Crystals", Journal of the Acoustical Society of America, vol.41, pp.981-993, April 1967.

[4] B. Studer, W. Zingg, and S. Dalla Piazza, "Chacteristics of High Frequency Fundamental Rectangular Quartz Crystal Resonators", in Proceedings of the 10th European Frequency and Time Forum, 1996, pp. 14-20.

[5]Transat Corporation, 31000 Bainbridge Road, Solon, Ohio 44139

[6] Eduard A. Gerber, and Arthur Ballato, Precision Frequency Control, vol.2, Orlando: Academic Press, 1985, ch.7, pp. 1-44.

A Comprehensive Mapping of Surface Acoustic Wave Properties on Gallium Orthophosphate ($GaPO_4$)

E. Henry Briot, E. Bigler, W. Daniau, G. Marianneau and A. Pakfar,
LPMO/CNRS, associé à l'Université de Franche-Comté,
32 Avenue de l'Observatoire, 25044 BESANCON Cedex - FRANCE

Abstract

A comprehensive mapping of Rayleigh wave (SAW) properties (velocities, temperature effects, electro-mechanical coupling factors, power flow angles, reflection coefficients) has been carried out for Gallium Orthophosphate ($GaPO_4$), a piezoelectric material of the same crystalline 32-class as quartz. It is shown that several temperature-compensated cuts exist, and that maximum electro-mechanical coupling factors are 4 times higher than (ST,X) quartz. Experimental SAW velocities and temperature effects measured on devices built on singly-rotated plates are also reported.[1]

Introduction

Gallium Orthophosphate ($GaPO_4$), a piezoelectric material with the same symmetry (32) as quartz, has several advantages over quartz : a higher electro-mechanical coupling, a lower sensitivity of acoustic modes to temperature effects, the ability to operate at much higher temperatures. Progress in crystal growth and characterization of $GaPO_4$ [1, 9, 10] has renewed interest for surface acoustic wave properties on this crystal [2]. A comprehensive mapping of surface acoustic wave properties on singly and doubly rotated plates for all propagation directions is presented under the form of contour charts versus cut and propagation angles for the following SAW properties : phase velocities, electro-mechanical coupling factor, power flow angle, first and second order temperature coefficients, turnover temperature. A complete mapping is also given given for the first order reflection coefficient under a metal strip or groove. Calculations are based on the most recently published set of material constants for $GaPO_4$ [7].

It is shown that several interesting cuts satisfying at least two important criteria among listed properties exist. Velocities have been found in the range of 2300 to 2900 m/s, maximum electro-mechanical coupling factors up to 4 times higher than (ST, X) quartz, and reflection coefficients up to 1.5 h/λ (3 times higher than (ST,X) quartz). Among various cuts, a singly-rotated cut exist in the vicinity of the IEEE [3] cut angle $\theta = -15°$. This cut, close to the AT cut of bulk waves on $GaPO_4$ [9, 6], is the equivalent of the ST,X cut for quartz, is temperature-compensated, has a zero power flow angle, an electro-mechanical coupling factor $k^2/2$ of 1.4 10^{-3} (2.4 times (ST,X) quartz) and a reflection coefficient of 1.2 h/λ (2.4 times higher also).

Experimental velocities and temperature effects measured on singly-rotated $GaPO_4$ plates supplied by LPMS-CNRS (Montpellier, France) are presented and compared to various sets of already published material constants.

1 Mapping properties of SAW propagation on $GaPO_4$

SAW properties : velocities, electro-mechanical coupling coefficient $k^2/2 \simeq \Delta V/V$ (between free and metallized surface), power flow angle and temperature coefficients were computed according a classical method [4]. Reflection coefficients, an important parameter for the design of SAW resonators and filters has been computed according to the model proposed by Datta & Hunsinger [5]. Material constants were taken from the last data set published in 1998 by Reiter et al. [7]. Results are displayed under the form of contour charts as a function of IEEE [3] (YXwlt $\varphi/\theta/\psi$) angles θ (2nd IEEE plate cut angle) and propagation angle ψ (SAW propagation direction) for different values of the first plate cut angle φ. Examples of such charts results presented at the end of this paper for $\varphi = 0$ (singly-rotated plates), φ =20°, 40°, and -50°. General tendencies are as follows:

- velocities, in the range of 2300 to 2900 m/s, are significantly smaller than for quartz, taking into account the higher mass density of $GaPO_4$,
- electro-mechanical coupling coefficients can be 4 times higher than quartz ($\Delta V/V = 2.5 \times 10^{-3}$ versus $.58 \times 10^{-3}$ for quartz) ; the tendency for $\Delta V/V$ is to decrease for values of the propagation angles ψ departing from 0 to ±90°,

[1]This work is supported by CEC under contract Brite-Euram BRPR-CT97-0391, THOMSON MICROSONICS (France), CNRS and DGA-DRET (France)

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

- first-order temperature coefficients (TCF1) are mostly in the range of ±20 ppm/°C, which is smaller than quartz ; the consequence is a large number of predicted compensated SAW cuts,
- reflection coefficients on an aluminium strip can be found in the range between .5 to $1.5h/\lambda$ (to be compared to (ST,X quartz : $0.5h/\lambda$) ; reflection coefficients for grooves (not presented here) are typically 30% smaller per h/λ but very similar in their anisotropy dependence with respect to reflection coefficients under aluminium strips.

A systematic investigation of SAW cuts satisfying simultaneous criteria of zero TCF1, zero power flow angle and a coupling coefficient $\Delta V/V \geq 10^{-3}$ yields the following results for velocities, coupling and reflection coefficients (Table 1):

φ	ψ	θ	V(m/s)	$\Delta V/V$(%)	R_{alu} (h/λ)
0	0	-25	2360	.15	1.
10	0	-35	2390	.15	1.
20	10	-45	2440	.12	.9
30	20	±50	2490	.10	.8
40	10	50	2450	.15	.5
50	0	40	2400	.15	.7
-10	0	-40	2400	.15	.7
-50	0	40	2400	.15	.7

Table 1: Zero power flow angle temperature-compensated SAW cuts of $GaPO_4$

Table 1 shows that the most interesting cuts are very probably found for singly-rotated plates or for small values of the cut angle φ, to get maximum values of coupling coefficients and reflection coefficients. For this reason as well as for practical reasons (singly-rotated plates being easier to fabricate) experimental work has been performed on singly-rotated plates ($\varphi = 0$).

2 Experimental results

Test devices (SAW delay lines) were fabricated on Y-cuts, and (Y-20) plates with propagation direction along X, i.e. cut angles $\varphi = 0$, $\psi = 0$, $\theta = 0$ and $-20°$. Fig. 1 shows a comparison between experimental values of Rayleigh wave velocities and theoretical predictions according to 4 sets of published material constants for $GaPO_4$ [6, 7, 8, 9]. A discrepancy exists between the more recent set of constants published by C. Reiter et al [7] and earlier sets by the AVL team [6, 8] or Palmier [9]. The agreement is correct with ref. [8] and [9], the latter being satisfactory, since elastic constants in [9] were measured on crystals grown with the same method used to fabricate the samples presented in this work. Differences between free surface and metallized surface velocities are expected to be in the range of 2-4 m/s and cannot explain the differences between theory and experiments. Fig.2 shows a comparison between experimental values for the first order temperature coefficient at 25°C and predicted values as a function of cut angle theta, for the same set of published material constants of $GaPO_4$ [6, 7, 8, 9]. Major discrepancies exist with earlier sets of constants [8] and the best agreement is found with refs. [6] and [9].

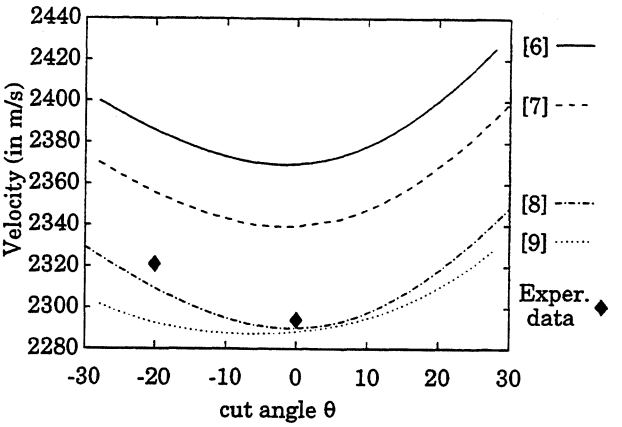


Figure 1: Comparison of measured SAW velocities with simulations for singly-rotated plates, propagation along X.

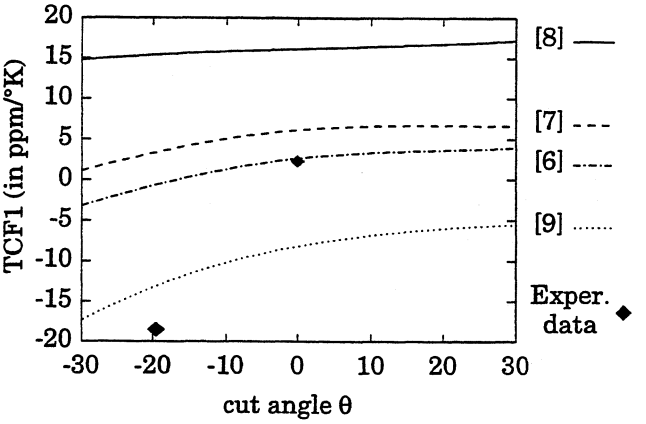


Figure 2: First-order temperature effects for SAW propagation on singly-rotated plates, propagation along X : comparison between simulations and measurements.

Ref. [7] as mentioned in Table 1 predicts a temperature-compensated cut at an angle $\theta = -25°$, whereas ref. [6] is in better agreement with experimental data showing evidence of a temperature-compensated cut in the θ-range ($-15°$ to $-5°$). Looking for a zero TCF1 at 25°C would suggest $\theta = -5°$ as the optimum cut, where an extrapolation based on the turnover temperature would favor $\theta = -15°$.

The temperature sensitivity measured on the Y-cut of $GaPO_4$, with a second order temperature coefficient of $-24\times10^{-9}/°C^2$ is already smaller than the value of the (ST, X) cut of quartz ($-34 \times 10^{-9}/°C^2$), and changes rapidly *vs.* the cut angle θ, since a cubic-like f-T curve is observed for $\theta = -20°$.

Conclusion

A comprehensive mapping of SAW properties on $GaPO_4$ has been obtained. Several temperature-compensated cuts with zero power flow angle are predicted theoretically. Singly rotated plates (Y-cut, (Y-20) cut) of $GaPO_4$ have been successfully tested for SAW propagation along the X-axis. Measured SAW velocities agree within 1% with theoretical predictions based previously published material constants. Discrepancies still exist between predicted values of the first order temperature coefficient and what was found experimentally. However, experimental evidence of a temperature compensated SAW cut ($\varphi = 0$, $\psi = 0$) in the range from $\theta = -15°$ to $-5°$ is found, together with a 2nd order temperature coefficient smaller than for the (ST, X) cut of quartz. A precise determination of what appears as the equivalent of the ST cut for $GaPO_4$ will be a logical consequence of future improvements in crystal growth and characterization.

Acknowledgements

The authors wish to thank Dr. E. Philippot and his team (LPMS-Montpellier) for providing $GaPO_4$ crystals. The authors are also very grateful to P. R. Bourquin (LCEP-Besançon) for sample orientation and cutting as well as Mrs. Jacquet and le Nain (University Paris XI) for high quality polishing.

References

[1] W. Krempl, F. Krispel, W. Wallnöfer, *Industrial development and prospects of $GaPO_4$*, proc. 2nd Workshop on Piezoelectric Materials, Edited by E. Philippot and A. Goiffon, Ann. Chim. Sci. Mat. vol 22, no8, pp. 623-626 (1997).

[2] M.P. Da Cunha and S.A. Fagundes, *Investigation on recent quartz-like materials for SAW applications*, IEEE Ultrasonics Symp., Sendai (Japan) october 5-8 1998.

[3] IEEE Standard on Piezoelectricity 176-1978.

[4] J.J. Campbell and W.R. Jones, *A method for estimating optimal crystal cuts and propagation direction for excitation of piezoelectric surface waves*, IEEE Trans. Sonics Ultrason., vol. SU-15, pp. 209-217 (1968).

[5] S. Datta, B. Hunsinger, *First order reflexion coefficient of surface acoustic waves from thin strip overlays*, J. Appl. Phys., vol. 50, n°9, pp. 5661-5665 (1979).

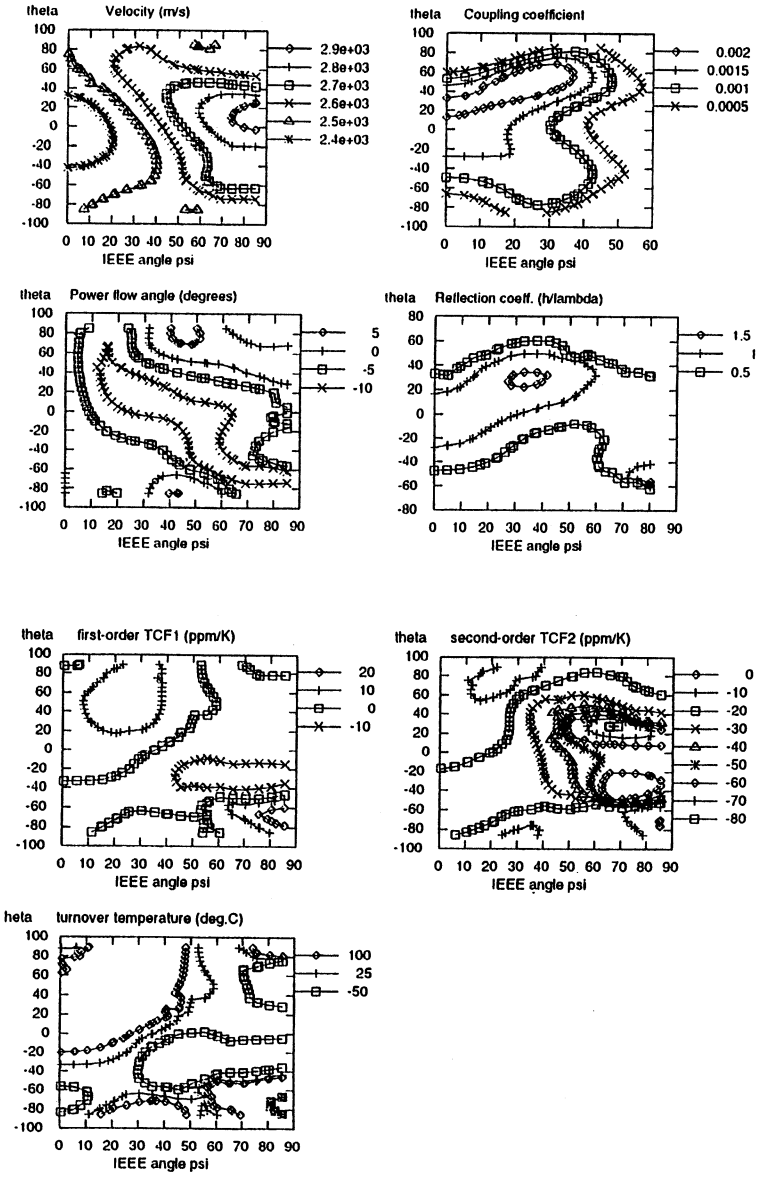
[6] C. Reiter, H. Thanner, W. Wallnöfer, P.W. Krempl, *Properties of $GaPO_4$ thickness-shear resonators*, Proc. 2nd Workshop on piezoelectric materials, Edited by E. Philippot and A. Goiffon, Ann. Chim. Sci. Mat., Vol 22, n°8 pp. 633-636 (1997).

[7] C. Reiter, H. Thanner, W. Wallnöfer, P.W. Krempl, *Temperature dependence and compensation of $GaPO_4$ thickness resonators*, Proc. European Frequency and Time Forum, Warszawa (Poland) pp 447-450 (1998).

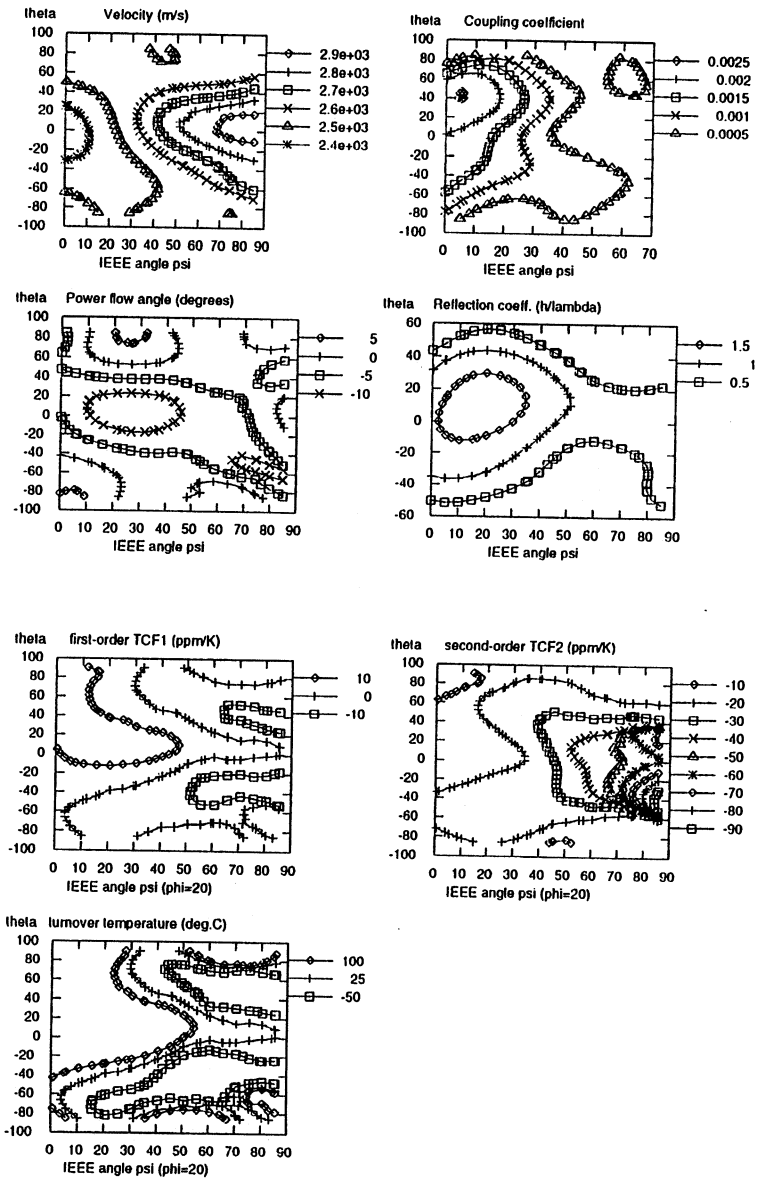
[8] W.Wallnöfer, P.W. Krempl, A. Asenbaum, *Determination of the elastic and photoelastic constants of quartz-type $GaPO_4$ by Brillouin Scattering*, Phys. Rev. B, vol. 49, pp. 10075-10080 (1994).

[9] D. Palmier, *Optimisation de la cristallogenèse et de la caractérisation des propriétés piézoélectriques du phosphate de gallium ($GaPO_4$), généralisation des relations "structures-propriétés" pour les matériaux de type quartz*, Thèse, Université Montpellier II, nov. 1996.

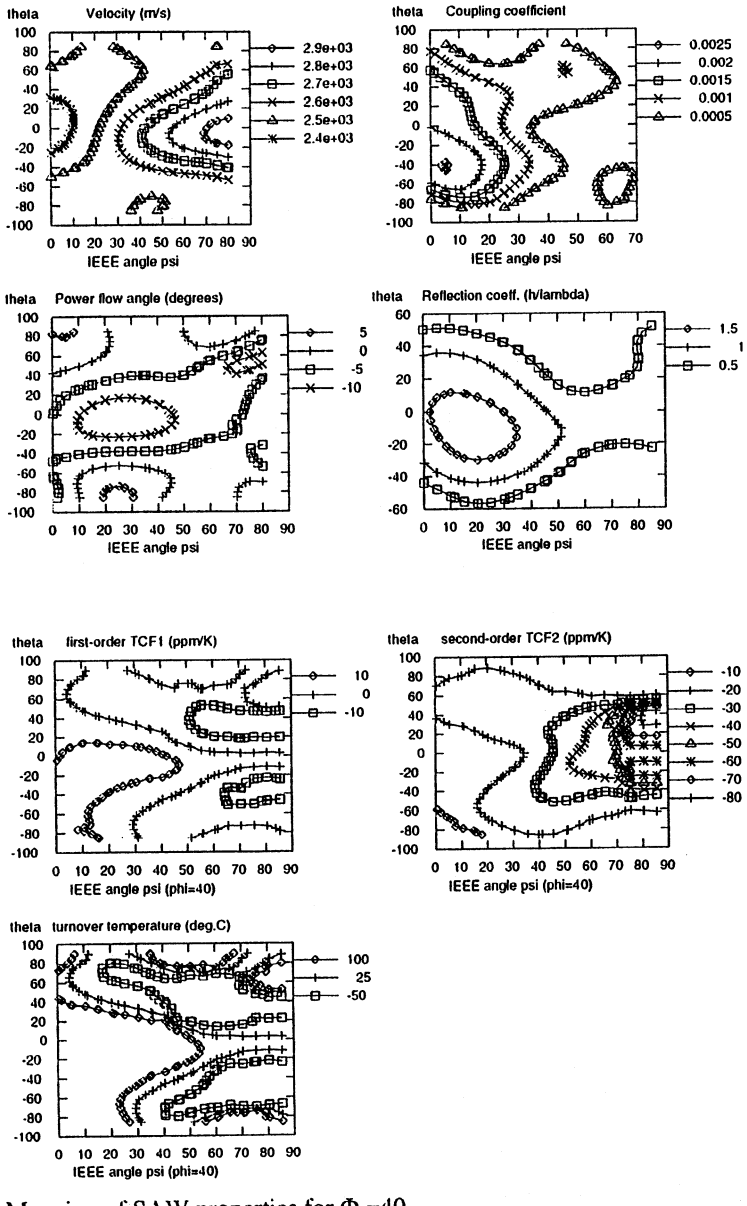
[10] J. Détaint, J. Schwartzel, A. Zarka, B. Capelle, D. Cochet-Muchy, E. Philippot, *Properties of the plane and plano-convex resonators using berlinite, gallium orthophosphate and langasite*, Proc. 1994 Ultrasonics Symp., IEEE cat.# 94CH3468-6, vol. 2, pp. 1051-1056.



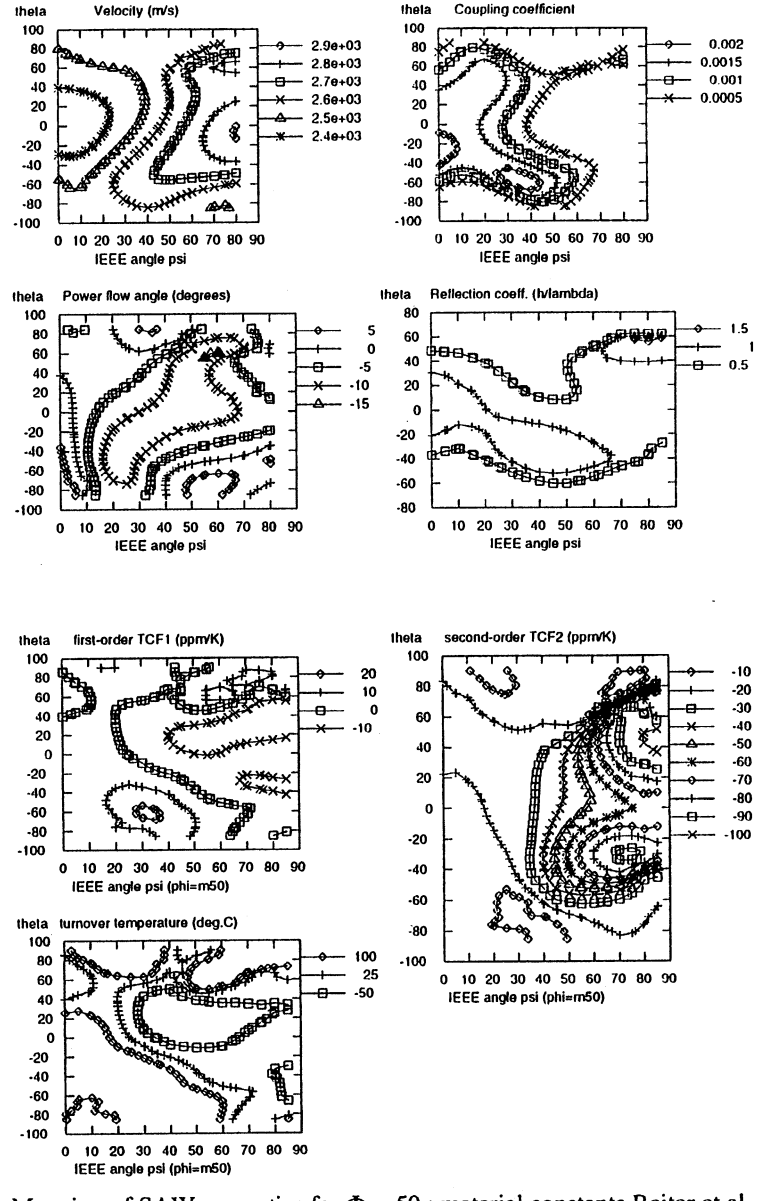
SAW Properties mapping for Φ=0 ; material constants Reiter et al. EFTF 98



SAW Properties mapping for Φ=20 ; material constants Reiter et al. EFTF 98



Mapping of SAW properties for Φ =40



Mapping of SAW properties for Φ =-50 ; material constants Reiter et al. EFTF 98

LANGASITE, LANGANITE, AND LANGATATE RESONATORS: RECENT RESULTS[1]

R.C. Smythe, R.C. Helmbold, G.E. Hague, & K.A. Snow
Piezo Technology, Inc., 2525 Shader Road, Orlando, Florida 32804, U.S.A.

ABSTRACT

Materials in the langasite family are of current interest for both bulk-wave and surface-wave devices. Plano-convex Y-cut bulk-wave resonators have been built and tested on overtones 1-9 using LGS (langasite, $La_3Ga_5SiO_{14}$), LGN (langanite, $La_3Ga_{5.5}Nb_{0.5}O_{14}$), and LGT (langatate, $La_3Ga_{5.5}Ta_{0.5}O_{14}$). Frequencies and motional inductances are compared with calculated values, with good agreement except for the motional inductance of LGT. For all 3 materials, frequency variation is an essentially parabolic function of temperature. For LGN and LGT, reported values of the Q-frequency product are significantly above the classical limit for AT-cut quartz.

1. INTRODUCTION

Materials in the langasite family are of current interest for both bulk-wave and surface-wave devices. In this paper, we report results for plano-convex, Y-cut bulk-wave resonators using LGS (langasite, $La_3Ga_5SiO_{14}$), LGN (langanite, $La_3Ga_{5.5}Nb_{0.5}O_{14}$), and LGT (langatate, $La_3Ga_{5.5}Ta_{0.5}O_{14}$). This is a continuation of work reported at this meeting last year. [1] These materials were selected from among the many members of the langasite family because of their potentially low acoustic loss and consequent high resonator Q-frequency product, and were grown by Crystal Photonics, Inc. [2, 3, 4]

2. RESONATOR DESIGN

The resonators are Y-cut, polished, plano-convex with a radius of curvature of 26.5 cm (2 diopters), a blank diameter of 14 mm, an electrode diameter of 6.35 mm, and a nominal 5th overtone frequency of 10 MHz using the slow shear (C) mode. They are mounted in evacuated HC-47 coldweld enclosures. This design allows adequate confinement of acoustic energy away from the mounting points and blank edge for the 5th and higher overtones. According to calculation, the 3rd overtone and, especially, the fundamental modes are not sufficiently confined to permit maximum Q; this is confirmed below by measurement.

3. RESONATOR PARAMETERS

The resonators were measured in a programmable temperature chamber at 5°C intervals from +5° to +125°C. At each temperature, all equivalent circuit parameters were measured using a high-precision reflection coefficient method, whose principles have been described earlier. [5]

3.1 Langasite (LGS)

Twenty-nine Y-cut resonators made from a single boule were tested. In Table 1, the median measured frequency and motional inductance at 25°C are compared with values calculated using Stevens-Tiersten theory and material constants given by Sakharov, *et al.* [6,7] Here and in the sequel, the calculated 5th overtone frequency has been adjusted to be equal to the measured value, since the blank thickness has not been measured with the accuracy required to use it as a basis for calculation.

Table 1. Equivalent Circuit Parameters, LGS 2 diop. Plano-Convex Resonators, 25°C

OT	Fs (kHz)		L1 (mH)	
	Calculated	Measured	Calculated	Measured
1	2012	2012	39	43
3	5983	5982	110	119
5	9974	9974	126	138
7	13953	13955	165	173
9	17947	17953	165	200

The agreement between measured and calculated frequency is excellent; motional inductance agrees within 10%, except for 21% for the 9th overtone. There is no measurable 11th overtone mode. The static capacitance is 10.0 to 10.2 pF, depending on overtone.

3.2 Langanite (LGN)

Twenty-eight resonators made from a single boule were tested. The median measured frequencies and motional inductances at 25°C are shown in Table 2 and compared with calculated values based upon the material constants of Pisarevsky, *et al.* [8] The static capacitance is 10.9 to 11.2 pF. The agreement between measured and calculated motional frequency and motional inductance is excellent for all overtones. Unlike LGS and LGT, the 11th overtone for LGN

[1] Work supported under U.S. Army Contract N66001-97-C-8634

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

appears to be trapped, in agreement with prediction, but has not yet been measured.

Table 2. Equivalent Circuit Parameters, LGN 2 diop. Plano-Convex Resonators, 25°C

OT	Fs (kHz)		L1 (mH)	
	Calculated	Measured	Calculated	Measured
1	2052	2051	36	35
3	6105	6102	105	106
5	10176	10176	121	120
7	14234	14237	167	159
9	18303	18308	174	175

3.3 Langatate (LGT)

Twenty-four resonators made from a single boule were tested over temperature for overtones 1-9. As for LGS, the 11th overtone of LGT is not measurable. The median measured frequency and motional inductance at 25°C are shown in Table 3 and compared with calculated values based upon the material constants given by Pisarevsky, *et al.* [9] For frequency, the agreement is good; however, the measured motional inductance is 33% to 40% higher than calculated. The static capacitance is 10.7 to 10.9 pF.

Table 3. Equivalent Circuit Parameters, LGT 2 diop. Plano-Convex Resonators, 25°C

OT	Fs (kHz)		L1 (mH)	
	Calculated	Measured	Calculated	Measured
1	2010	2021	24	32
3	6024	6023	72	99
5	10046	10046	81	114
7	14055	14058	111	148
9	18074	18080	116	163

3.4 Resonator Q and Q-frequency product

High Q is a desirable resonator property for both filter and oscillator applications. A convenient figure of merit is the product of Q and frequency. An investigation by Warner, published in 1960, found that for AT-cut natural quartz resonators measured at or near 25°C, the upper bound on Q·f was 15×10^6, where f is in MHz. [10] This limit was found by optimizing the design of the resonator at each frequency, using plates from 15 to 90 mm in diameter; the limit was essentially constant from 200 kHz to 100 MHz.

For the three materials, we have measured Q·f, both for unplated, unmounted contoured blanks and for finished resonators. The blank measurements are made at ambient temperature in vacuum, at a pressure of 10^{-4} Torr, using non-contacting electrodes, 6.35 mm in diameter spaced 1.3 mm apart, and a capacitively balanced hybrid network to cancel feedthrough due to inter-electrode capacitance. The blanks are supported at their edges. The measured data are corrected for loading due to the measurement circuit, so that the Q reported is the unloaded Q. Table 4 gives the maximum and median values of Q·f for the 5th, 7th, and 9th overtones of a representative sample.

Table 4. Q-frequency product of unplated blanks, measured in vacuum

OT	Max. (median) Blank Q·f/10^6		
	LGS	LGN	LGT
5	7.9 (7.2)	19.3 (13.2)	28.6 (25.4)
7	9.7 (8.7)	21.5 (14.4)	29.2 (21.9)
9	12.9 (9.0)		20.1 (11.4)

The finished resonators are measured in a temperature chamber, as described in section 3. Table 5 reports the maximum and median Q·f values at 25°C.

Table 5. Q-frequency product of finished resonators

OT	Max. (median) Resonator Q·f/10^6		
	LGS	LGN	LGT
5	6.2 (4.8)	14.6 (12.4)	21.2 (18.4)
7	8.8 (5.8)	17.5 (13.9)	25.6 (23.1)
9	6.0 (4.2)	17.0 (11.0)	22.6 (15.9)

The differences in Q·f among the 3 materials are quite striking, with LGT being easily the best at this stage of material development, followed by LGN. Both of these have exhibited maximum values in excess of Warner's value for AT-cut quartz, and for LGT, the median value exceeds the AT cut maximum. Values for LGS are much lower; Détaint, *et al.* have reported similar values. [11]

In all cases, the Q of the finished resonator is significantly below that of the blank. While we have not investigated this exhaustively, measurements show that the reduction is due in part, at least, to ohmic losses in the electrodes, especially for the 5th and 7th overtones of LGT.

3.5 Unwanted Modes

Unwanted mode plots have been made with the resonators in a capacitively-balanced hybrid network. Figures 1 and 2 show the first few unwanted modes for one of the LGT resonators.

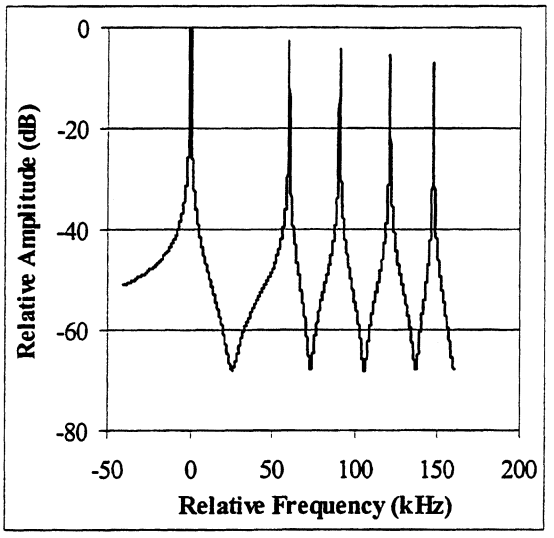


Figure 1. Unwanted Modes for the 5th Overtone of LGT Plano-Convex Resonator, Unit 22

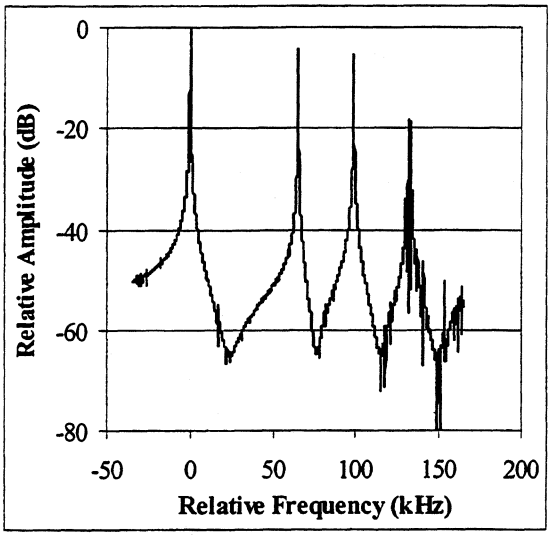


Figure 2. Unwanted Modes for the 7th Overtone of LGT Plano-convex Resonator, Unit 22

A very interesting feature in figure 2 is that, in addition to the usual anharmonic modes, we see a large number of other small modes. We have not attempted to identify them; however, such modes are not electrically excited in the case of AT- or SC-cut quartz resonators. They extend below the frequency of the main mode as well as above, and are also present in the vicinity of the 9th overtone, though less prominently. They have also been observed for the 7th overtone of LGN, but have not been seen for LGS.

Equivalent circuit parameters have been measured for the (n,0,2) and (n,2,0) modes. Table 6 presents the relative frequency, the mode resistance, and the resistance ratio for the 5th, 7th, and 9th overtones of the resonator plotted in figures 1 and 2. The agreement of the measured and calculated frequencies is good; the resistance ratios are approximately the same as the inductance ratios, except for the (7,0,2) and (9,0,2) modes, where the resistances are much higher.

Table 6. Unwanted mode relative frequency, resistance, and resistance ratio for LGT resonator unit 22

Mode (n,m,p)	Rel. Frequency [kHz]		Res. [ohms]	Res. Ratio
	Calc.	Meas.		
5,0,0	0.0		3.6	1.0
5,0,2	66.3	59.9	7.6	1.9
5,2,0	94.7	90.4	7.5	2.1
7,0,0	0.0		7.4	1.0
7,0,2	66.4	65.3	32.8	4.4
7,2,0	99.8	98.8	12.7	1.7
9,0,0	0.0		18.3	1.0
9,0,2	66.4	61.7	256.6	14.0
9,2,0	151.5	141.6	46.7	2.6

3.6 Frequency vs. Temperature

The resonators were measured at 5°C intervals from +5° to +125°C and for overtones 1-9. For all three materials, the fractional frequency variation with temperature closely approximates a parabolic function of the form

$$\Delta f/fo = (f_T - f_{To})/f_{To} = b\cdot(T - To)^2$$

where To is the turnover temperature, f_T is the frequency at temperature T, and b is the parabolic constant. The turnover temperatures and parabolic constants for overtones 1-9 of each material are given in Tables 7 and 8.

Table 7. Turnover temperature vs. overtone for Y-cut LGS, LGN, and LGT resonators

OT	Turnover Temp. [°C]		
	LGS	LGN	LGT
1	25.0	61.8	71.4
3	71.9	64.4	69.4
5	81.4	66.2	69.9
7	77.9	64.5	67.8
9	65.3	63.6	66.9

Table 8. Parabolic constants vs. overtone for Y-cut LGS, LGN, and LGT resonators

OT	Parabolic Constant, b [ppm/°C^2]		
	LGS	LGN	LGT
1	-0.0589	-0.0687	-0.0694
3	-0.0571	-0.0666	-0.0666
5	-0.0567	-0.0663	-0.0673
7	-0.0562	-0.0660	-0.0669
9	-0.0531	-0.0657	-0.0667

Note that for LGN and LGT, neither the turnover temperature nor the parabolic constant vary significantly with overtone – roughly 62° to 66° for LGN and 67° to 71° for LGT. For LGS, however, the turnover temperature is 25°C for the fundamental mode, and varies from 65° to 81° for overtones 3 through 9. This is in general agreement with measurements by Sakharov, *et al.* [6]

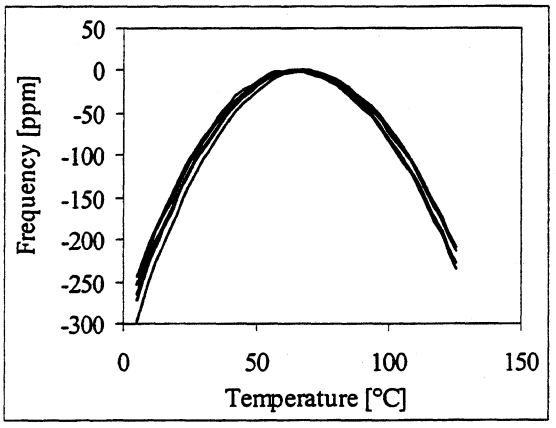


Figure 3. Relative Frequency vs. Temperature for an LGT 2 diop. Plano-convex Resonator (typical of 24 units)

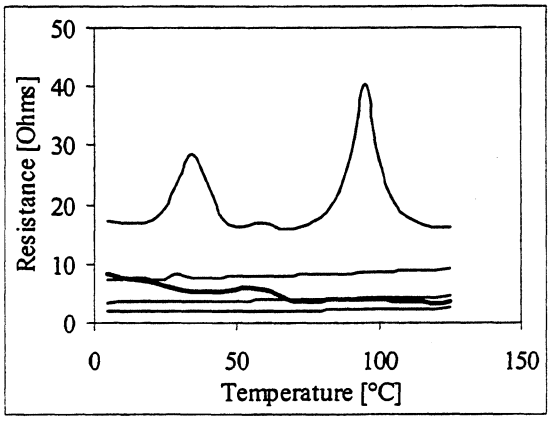


Figure 4. Relative Frequency vs. Temperature for an LGT 2 diop. Plano-convex Resonator (typical of 24 units)

Typical curves of LGT frequency vs. temperature are shown in figure 3. It can be seen that the five f-T curves nearly overlay each other. Figure 4 shows resistance vs. temperature for the same resonator. The resistances of overtones 3-9 are progressively higher; the resistance of the fundamental mode, shown by the heavy line, deviates from this progression due to its very inadequate confinement away from the edge of the blank. Strong coupled modes can be seen for the 9th overtone; this is typical of all 3 materials.

4. CONCLUSIONS

For each of the three materials being investigated the Y-cut has an essentially parabolic frequency-temperature characteristic. While not shown here, the turnover temperature can be adjusted to a desired value by small changes in the Y-Y' angle, θ. The parabolic constants are similar for the three materials. At this stage of development, langatate, LGT, exhibits the highest Q-frequency product, but langanite, LGN, is a strong contender. With further improvements in material and optimization of resonator orientation and design, it is not unreasonable to expect even higher values of the Q-frequency product.

REFERENCES

[1] Smythe, R. C., "Material and Resonator Properties of Langasite and Langatate: a Progress Report," *Proc. IEEE International Frequency Control Symp.*, pp.761-765, 1998.

[2] Chai, B., J.L. LeFaucheur, Y.Y. Ji, & H. Qiu, "Growth and Evaluation of Large Size LGS ($La_3Ga_5SiO_{14}$), LGN ($La_3Ga_{5.5}Nb_{0.5}O_{14}$) and LGT ($La_3Ga_{5.5}Ta_{0.5}O_{14}$) Single Crystals", *Proc. IEEE International Frequency Control Symp.*, pp.748-760, 1998.

[3] Chai, Bruce H.T., H. Qiu, Y.Y. Ji, & J-L Lefaucheur, "Growth of High Quality Single Domain Crystals of Langasite Family Compounds," These Proceedings.

[4] The materials were grown by Crystal Photonics, Inc., Sanford, Florida.

[5] Smythe, R.C., "An Automated Resonator Measurement System Using a Reflection Coefficient Bridge, *Proc. 35th Annual Frequency Control Symp.*, 1981, pp.280-285.

[6] Stevens, D.S., and H.F. Tiersten, "An analysis of doubly rotated quartz resonators utilizing essentially thickness modes with transverse variation," *J. Acoust. Soc. Am.*, v.79, no.6, June, 1986, pp. 1811-1826.

[7] Sakharov, P. Senushencov, A. Medvedev, Yu. Pisarevsky, "New Data on Temperature Stability and Acoustical Losses of Langasite Crystals", *Proc. IEEE International Frequency Control Symp.*, pp.647-652, 1995.

[8] Pisarevsky, Yu.V., P.A. Senushencov, P.A. Popov, B.V. Mill, "New Strong Piezoelectric $La_3Ga_{5.5}Nb_{0.5}O_{14}$ with Temperature Compensation Cuts," *Proc. IEEE International Frequency Control Symposium*, pp. 653-656; 1995.

[9] Pisarevsky, Yu. V., P.A. Senyushenkov, B.V. Mill, N.A. Moiseeva, "Elastic, Piezoelectric, Dielectric Properties of $La_3Ga_{5.5}Ta_{0.5}O_{14}$ Single Crystals," *Proc. 1998 IEEE International Freq. Control Symp.*, 1998, pp. 742-747.

[10] Warner, A.W., "Design and Performance of Ultraprecise 2.5-mc Quartz Crystal Units," *Bell Sys. Tech. J.*, Sept., 1960, pp. 1193-1215.

[11] Détaint J., B. Capelle, A. Zarka, & D. Cochet-Muchy, "Langasite Plano-Convex Resonators Using Lateral or Perpendicular Field Excitation," *Proc. 9th European Forum on Time and Frequency*, 1995, pp. 289-296.

GROWTH OF HIGH QUALITY SINGLE DOMAIN CRYSTALS OF LANGASITE FAMILY COMPOUNDS

B. Chai, H. Qiu, Y. Y. Ji and J. L. Lefaucheur
Crystal Photonics, Inc., 2729 N. Financial Court, Sanford, FL 32773, USA

Abstract:

Langasite family compounds have attracted a lot of attention because of the combination of a number of nice material properties such as high piezoelectric coupling, temperature compensation and low acoustic loss. Another very important property is that the compound melts congruently so that large single crystals of langasite can be produced by the conventional Czochralski melt pulling technique similar to that of $LiNbO_3$ and $LiTaO_3$. This makes the commercial scale production feasible.

Nevertheless, despite the promises, the production of large commercial size langasite single crystal has always been erratic. We reported a year ago that langasite has both twinning and domain structure problems which were not recognized before. We developed the etching technique to reveal these crystal defects.

In order to produce single domain crystals, a systematic process is established to screen the seed crystals so that the initial seed is twin-free, domain-free and low in dislocations. We also made special effort to reduce the temperature gradient within the growth chamber, since we have observed evidence of strain induced defects which are due to the growth anisotropy of the crystal. With all the changes in the growth process and growth environment, we are able to produce very high quality single domain crystals of the langasite family compounds including LGS, LGN and LGT. A parallel paper will report for the first time a Qf value of 29 million (where f is frequency in unit of MHz) from one of the LGT crystals produced at our place. This value greatly exceeds the bench mark value of 17 million that quartz crystal has maintained in the past half a century.

At present time, the production of high quality single crystals of langasite family compounds is very reproducible. We are still in the process to improve the crystal quality and to establish the commercial process of producing 3" diameter single crystals of langasite and its isomorphic compounds.

Instroduction:

Langasite family compounds have received a lot of attention recently because of the potential application of the new digital mobile communication system. They have several quite desired properties superior to the current material – the ST-cut quartz. Two of the desired properties are higher electromechanical coupling which will allow wider bandwidth which is 3 times that of ST-quartz, and lower acoustic loss or higher Q value which is about twice that of quartz. In addition, langasite has the same temperature compensation property as quartz so that within a reasonable temperature range near room temperature, the frequency is stable without the need of an oven. Given all these nice properties, langasite should be readily adapted for device fabrications.

Unfortunately, this is not quite the case. Langasite family contains more than 100 compounds. Both the composition and the structure of langasite are far more complex than any of the current commercial piezoelectric crystals. This add more difficulty to control the composition and to grow high quality crystals. Moreover, the majority of these compositions are not so useful. Most of them melt incongruently and thus difficult to produce. A large number of the langasite family compounds contain Ge which evaporates readily in both oxide (GeO_2) and suboxide (GeO) forms. High vapor pressure prevents them to be grown to high quality single crystals under controlled conditions. As a result, only a small number of the compounds that are growable. They are primarily gallium containing compounds which are also congruent melting. The three most commonly cited compounds are:

LGS	langasite	($La_3Ga_5SiO_{14}$)
LGT	langatate	($La_3Ga_{5.5}Ta_{0.5}O_{14}$)
LGN	langanite	($La_3Ga_{5.5}Nb_{0.5}O_{14}$)

At present time, only langasite (LGS) wafer is available commercially and is produced in a number of places. Both LGN and LGT are available only in small quantities for property evaluation. One of the very important reason

0-7803-5400-1/99/ $ 10.00

that langasite has the commercial potential is that it melts congruently. Therefore, it is totally feasible to grow the crystal in large sizes at reasonable growth rate by the conventional Czochralski melt pulling technique similar to $LiNbO_3$, $LiTaO_3$ and not like the slow flux growth technique similar to that of $KNbO_3$. Even though the basic growth process seems straightforward, the actual growth is much more complicated. There is evaporation during growth so that the melt has to be compensated. The distribution ratios of all the growth ingredients are not all unity. Depending on the specific crystal, a specific melt composition has to be used in order to maintain high quality growth and with high material yield. Finally, the crystal can form a number of defects such as twinning and domain structure. It requires special effort to optimize the growth condition in order to achieve the high quality single crystal growth. At present time, there is still a lot of problems to supply the material in production quantity with consistent crystal quality.

To be able to launch the product to a commercial success, one must be able to address three basic questions, namely:

(1) can we supply the crystal with consistent material property,
(2) can our production capacity meet the demand for device fabrication usage,
(3) can the crystal production price be low enough to be acceptable for device use as compared with current existing crystal wafers such as ST-cut quartz or $LiNbO_3$.

Growth of Langasite Family Compounds:

Langasite has the Ca3Ga2Ge4O14 structure. This family of compound was discovered in late 70's by Professor B.B. Mill and his associates [1] of the Russian Academy. The crystal structure is trigonal and has the space group of P321. It is similar but not the same of quartz which has the space group of $P3_121$ or $P3_221$. Langasite has higher symmetry than quartz without the handness. On the other hand, since both crystal is trigonal in symmetry, all the calculations developed for quartz can be used directly to langasite.

The growth langasite family compounds were first established in Russia. They found that a number of the compounds melt congruently. Single crystal were produced by pulling directly from the melt. Nevertheless, the exact phase diagram was never established. We proposed a schematic phase diagram by modify slightly the Japanese result [2]. It is clear that langasite is the only ternary compound within the phase boundary of the La_2O_3 – Ga_2O_3 – SiO_2 system.

The melting temperature of most of the gallium containing langasite compounds is below 1500°C with LGS at 1470°C. Therefore, it is possible to use either platinum or iridium crucible with either resistant or induction heating for the growth. In our place, we use Iridium crucible only in combination with induction heating. Because of the oxidation of the iridium crucible, all the growth is under nitrogen environment. It is know that the color of langasite crystal is sensitive to the oxidation state during growth. Under ambient oxygen condition, the crystal shows a deep reddish orange color. This is the typical color of langasite crystals available in the commercial market. Our crystals are grown under nitrogen environment, they were much lighter in color. Typically, they vary from colorless to light yellow. Although we all claim that the color of the crystal has no effect on the piezoelectric performance. However, there is no controlled test to prove one way or the other.

To start the growth, the chemicals are carefully weighed and mixed before loading. The compositions of the chemicals have been adjusted to accommodate the evaporation and the distribution of these elements. Charge was heated by RF induction to the melting temperature and let to melt homogeneously. A crystal seed with a pre-selected orientation is then lowered to touch to the melt surface and let to stabilize before starting the pull program. We use a two-loop weight feed back cascade control system to control the diameter of the crystal. In the early time, we are using the traditional PID (Proportional-integral-differential) control system for growth. More recently, we convert all the system to our latest two-loop adaptive controlled system. We are able to see dramatic change in the growth result. We will elaborate this issue in the discussion section.

Single crystals were pulled directly from the melt. Typical pull rate is about 1 mm per hour and rotation varies from 10 to 20 rpm. The primary seed orientation is c-axis or [0001]. We also explore the growth along the Y-axis [10-10]

direction as well as other orientations. We mentioned before that there are many factors that can change the composition of the melt. Based on careful observation of the growth result, we are able to adjust the melt composition to accommodate most of the losses. Currently, we are able to convert at least 80% of the melt to high quality crystal. In some composition such as LGT, the melt conversion is exceeding 90%. Typical growth cycle is 7 to 9 days depending on size of the charge. We currently use two sizes of crucibles to produce two sizes of crystals. For experimental work, the typical size of crystal is 45 – 50 mm in diameter and 150 mm in length. For production work, the crystal is 70 – 75 mm in diameter and up to 250 mm in length.

In addition to the large contrast in color between our crystals with those of commercial ones, our crystals also show distinctive roundness of the overall shape in sharp contrast to all the commercial crystals with strong facet development. At present time, we attribute the difference due to the temperature gradient within the furnace, even though our furnace is designed specifically with low heat loss and low temperature gradient. Another possibility of the difference is due to the melt composition. Since no one publish the exact melt composition that they are using for the growth, it is practically impossible to make comparison. In our own experience, we find that there are distinct differences in the tendency for facet formation among the three compositions that we have investigated. In our case, given the same growth environment, LGS has the smallest tendency of facet formation whereas LGN has the strongest tendency. Fig. 1 shows a picture of a typical LGS c-axis boule. Clearly there is little facet development.

Defects in Langasite Crystals:

To grow small size langasite crystals for research purpose is not a difficult work, since the crystal yield is not of concern. The difficulty increases drastically with increasing the crystal diameter. The reason is that the material intake by the crystal is much faster for larger boules and diffusion rate at the crystal-melt interface becomes more critical. In other words, unless the melt composition is correct, it is very easy to form defects. In our last year report [2], we describe in detail what we found as defects in langasite crystals. We found that there are three basic types of defects:

(1) Opaque white inclusions,
(2) Twinning -- only electric twins
(3) Domain structure -- low angle grain boundaries.

The opaque white inclusion is always associated with poor melt composition control. Even with proper adjustment of the melt composition, in some cases, we are only able to grow up to 80% of the melt into high quality crystals. The bottom portion of the crystal always turn into opaque white loaded with inclusions. Unfortunately, there is also limitation how much we can adjust the melt without causing growth difficulty at the very beginning of the growth cycle. At present time, we are satisfy with >80% yield. To us the white opaque inclusion is no longer a problem.

Twinning is expected in langasite crystal because of the symmetry. Since there is no handness, we do not have the optical twins, but we do have the electrical twins. Fig. 2 shows a typical electrical twin in a LGN crystal. The surface of the crystal has been etched. By shining with an oblique light, the twins are revealed by the reflections of the small facet of the etch pits. The distribution of electric twin looks nearly identical to that of the quartz crystals further confirm the close crystal symmetry relationships between the two crystals. On the other hand, we only see this type of twins in LGN and LGT crystals and very rarely seen in LGS crystal. At present time we do not have a physical model to explain the difference. One possibility is that the twinning might be related to the ordering of the ochahedral site of the structure. In this case, both LGN and LGT have disorder in this particular site due to 50/50 occupation by Nb^{5+} (or Ta^{5+}) and Ga^{3+} ions. Partial ordering of the two ions may be the cause of twinning. One other hand, in LGS, the octahedral site is occupied by Ga^{3+} ion only. This subtle difference may be the reason for the difference in twin occurrence.

The third type of defect is the low angle domain structure. This is the most serious defect and our primary concern at the present time. When we start a run, if the growth is stable from the beginning, we would expect to complete the growth with a perfect boule (Fig. 3). Unfortunately, this is not always the case.

In many occasions, the crystal starts with a smooth growth and then all of a sudden a small domain is formed at the rim of the crystal in contact with the melt. Once the domain is formed, it will grow along with the crystal quite persistently until the end of the run (Fig. 4). This domain structure is quite thin (a few millimeters up to a centimeter) along the skin of the crystal boule. Unfortunately, because its orientation is different from the original crystal, the differential thermal expansion will force the region to crack into small pieces as seen in Fig. 4 since the thickness of the domain region is much less than the bulk of the crystal. The fine cracking is always on the domain side. In some cases, if the domain region is very thin, the whole domain region will fracture out and the bulk of the crystal is saved (Fig. 5). However, if the domain region is thicker, then a few of the fractures will also penetrate into the main crystal and the crystal is totally ruined.

The formation of the domain structure is quite randomly and it can form at any stage during the growth period, although it most frequently formed during the growth of the cone section. Since the domain region has totally different crystallographic orientation, we believe that it is a nucleation and growth phenomenon. Once a stable nucleus is formed and attached at the side of the main crystal, it will grow continuously until the end of the run. The reason the domain structure forms more frequently during the cone growth period is that this is the period we are pushing the crystal to increase diameter to the final size. As a consequence, the degree of supercool is increased. It is possible that with a small fluctuation in the growth rate, a sudden correction of the heating power will cause the formation of a stable nucleus. Once the crystal reached its final diameter, the degree of supercool is stabilized. The chances of domain formation are much reduced, even though we did observed once during the growth of a 3" diameter LGT crystal. The domain structure is formed right at the flat facet side of a constant diameter crystal near the end of the growth. Of course the domain structure also caused massive fracture of the main crystal and it is totally ruined.

The formation of domain structure is quite unique for langasite. Over the years, we have grown quite a large number of crystals. Even though every crystal has its own unique characteristics, we have not observed the problem of this type of domain formation. The easiness of formation of stable nuclei indicates that the melt can tolerate very little supersaturation. We believe that this is cause of the complex ternary melt composition. We have tried a number of ways to overcome this problem such as reduce temperature gradient, reduce the cone angle, soften the gain in the control system. All these approaches help but can not totally eliminate the problem. Just recently, we are in the process of expansion and also upgrade our control system. We have developed a two loop cascade adaptive control program and begin to replace the existing conventional two loop PID control system. By using the new control system, we find that the problem of domain formation is totally eliminated. The new system has a much faster response in anticipate the changes ahead. By doing so, the chance of small fluctuation of growth rate is eliminated so that the melt will not be supercooled to the extend to cause stable nuclei formation. A typical LGT crystal grown by this adaptive control system is illustrated in Fig. 6.

Summary and Conclusions:

Despite of the problems of crystal defects, we are able to produce all three langasite family compounds, LGS, LGN and LGT, with superior properties. Our LGS crystal has a Qf value near one order of magnitude better than the commercial crystals. But more important of all, both our LGN and LGT crystals show a Qf value exceeding that of the quartz crystal which was the bench mark low loss piezoelectric crystal for the past half centuries [3, 4]. We believe that the superior values attribute both the intrinsic material properties as well as our growth process. At present time, the LGT crystal seems to have the best overall properties and also consistency in crystal growth. In the past, we have more problem of twin formation in LGN crystals. However, with our new adaptive control growth system, the quality of LGN crystal improves dramatically and the twin problem is more or less under controlled.

We have stated at the beginning that even though langasite family compounds have three different types of defects, the most troublesome one is the domain formation. After extensive investigation, we recognize that it is a nucleation and growth problem. Apparently, because of

the complex melt composition, the tolerance to the degree of supersaturation of the melt is very low. Small power fluctuation is sufficient to cause the formation of nuclei. To eliminate the domain problem, we will need very stable growth control system to minimize the power fluctuation. We are able to achieve such growth control stability by using our new two-loop cascade adaptive control system. We are pleased to say that since we install with this new control system, all the crystals produce since have no domain formation problem. Current test is with 2" diameter crystals. We are able to grow langasite crystals at any orientation without difficulty. We will soon extend the test for the growth of 3" diameter crystals.

In conclusion, langasite crystals have demonstrated excellent material properties for both bulk and SAW device applications. In order to achieve commercial success, it is necessary to have a production process which can produce the crystal with high yield and consistent crystal quality. Current commercial supply is limited and the crystal quality is not quite consistent enough for mass production. We have investigated in detail the growth defects in these crystals. Proper remedies to eliminate these defects are proposed and they have demonstrated to be effective. Among the three crystals that we have investigated, we find that both LGN and LGT have superior properties over LGS. Between the LGN and LGT, we find that the later one has less problem of twinning and more consistent in crystal yield. We are now concentrating our effort into only LGN and LGT crystals. We believe that they will be the materials of choice for the eventual device application. Our most difficult task is trying to increase the crystal yield and to reduce the crystal cost so that they can be more or less in line with that of quartz and $LiNbO_3$.

Acknowledgement:

The authors want to express their thanks to Dr. John Vig, Mr. Robert Smythe and Prof. Donald Maloca for their contribution in the material evaluation and also discussion to improve the crystal quality. This work is supported by the Navy contract No: N66001-97-C-8634 and is subcontracted from the General Technical Service contract No: GTS-97-02.

References:

[1] B.V. Mill, A.V. Buashin, G.G. Khodzhabagyan, E.L. Belokoneva, N.V. Belov, Sov. Phys. Dokl., v. 27, p.434, 1982.

[2] B. Chai, J.L. Lefaucheur, Y.Y. Ji and H. Qiu, "Growth and evaluation of large size LGS ($La_3Ga_5SiO_{14}$), LGN ($La_3Ga_{5.5}Nb_{0.5}O_{14}$) and LGT ($La_3Ga_{5.5}Ta_{0.5}O_{14}$) single crystals", Proceedings of the 1998 IEEE International Frequency Control Symposium, p. 748-760, 1998.

[3] R.C. Smythe, "Material, and resonator properties of langasite and langatate: A progress report", Proceedings of the 1998 IEEE International Frequency Control Symposium, p. 761-765, 1998.

[4] R.C. Smythe, R.C. Helmbold, G.E. Hague and K.A. Snow, "Langasite, langanite and langatate resonators: recent results", This Proceeding, 1999.

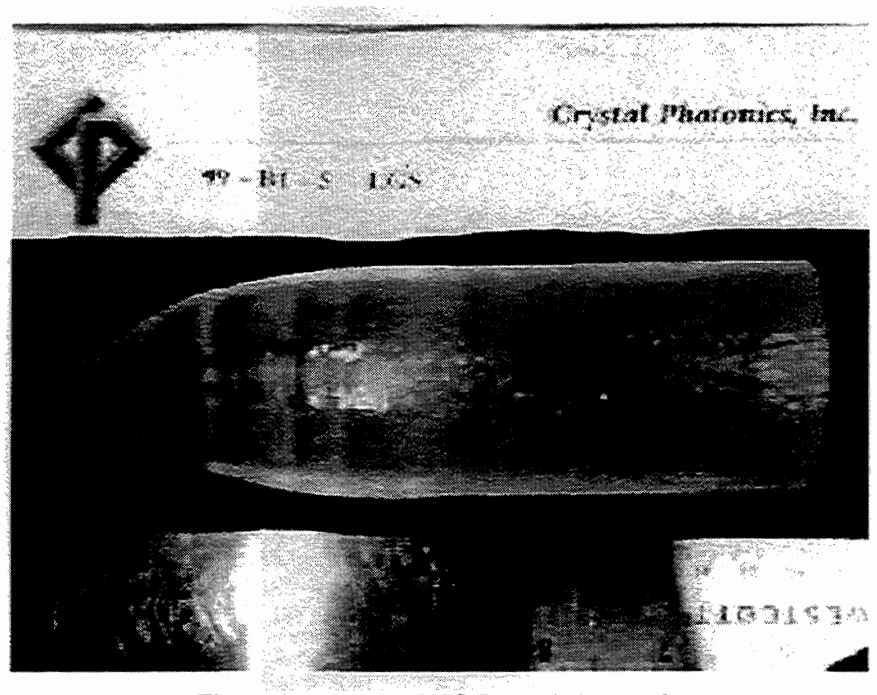


Figure 1: As grown LGS (langasite) crystal.

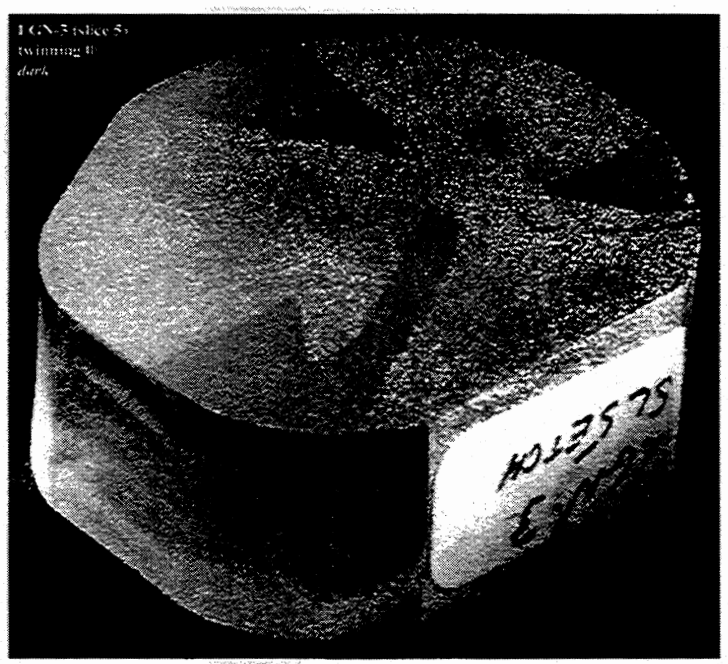


Figure 2: Electric twins in LGN (langanite) crystal.

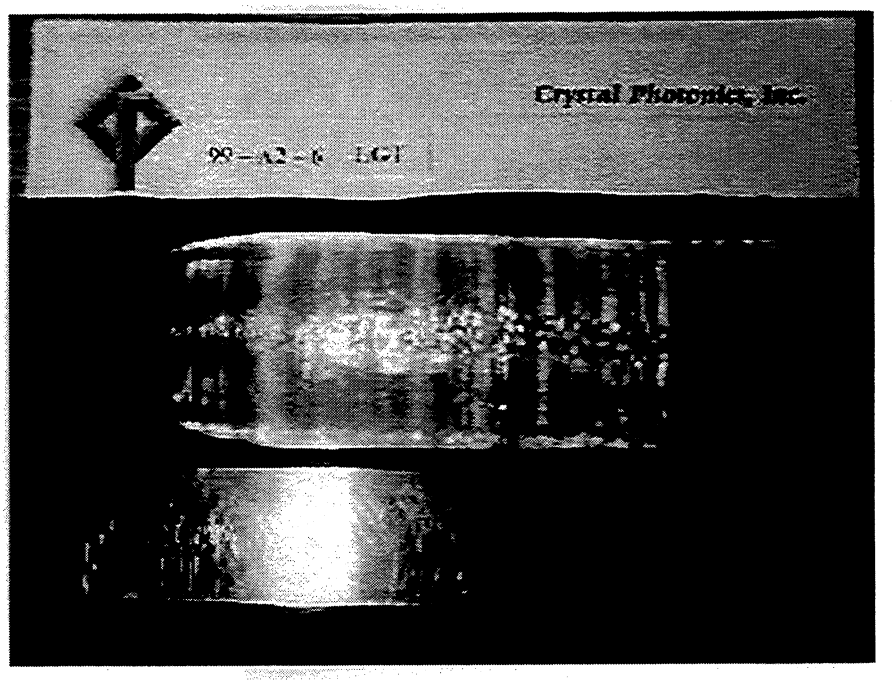


Figure 3: As grown LGT (langatate) crystal.

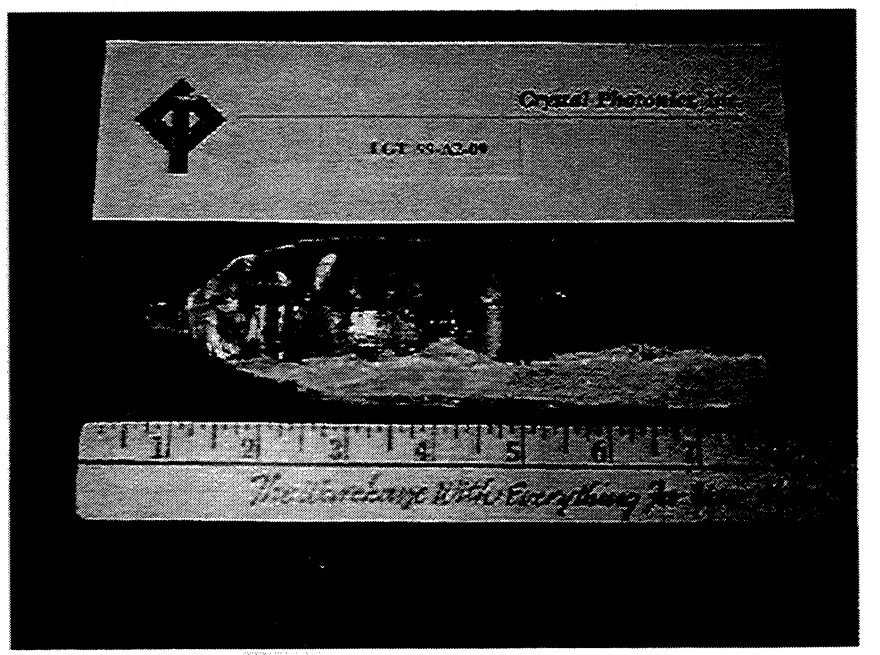


Figure 4: LGT crystal showing the domain structure.

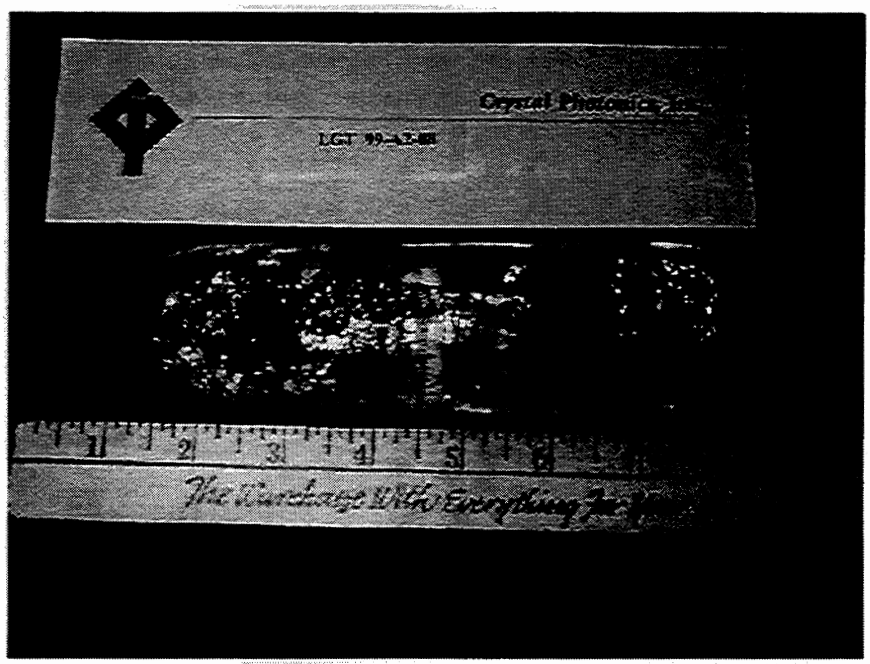


Figure 5: LGT crysal with the domain structure region removed by itself.

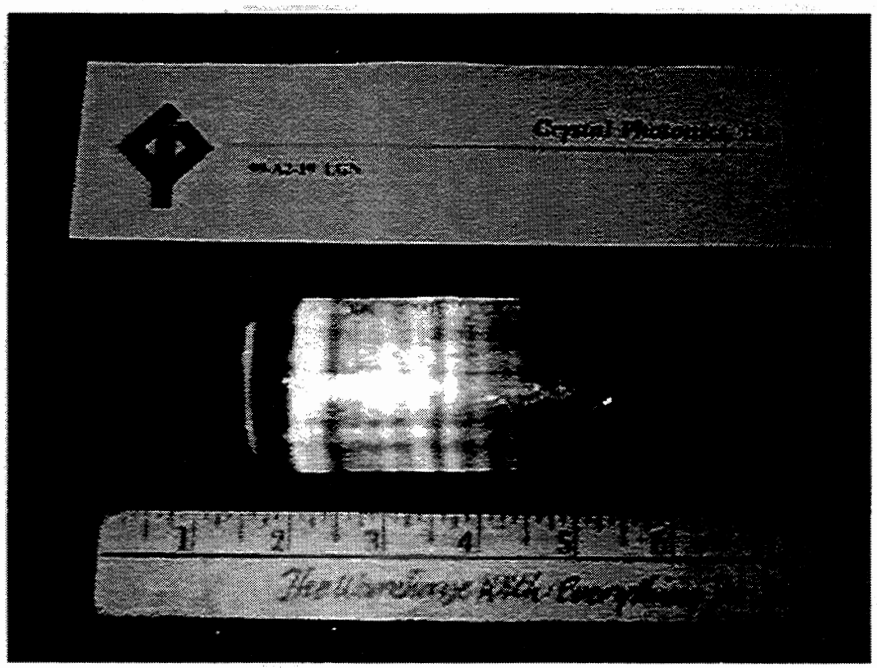


Figure 6: LGN crystal grown with the new computer control system.

1999 Joint Meeting EFTF - IEEE IFCS

SYNTHESIS, GROWTH AND SOME PROPERTIES OF SINGLE CRYSTALS WITH THE $Ca_3Ga_2Ge_4O_{14}$ STRUCTURE

B.V.Mill, Yu.V.Pisarevsky*, E.L.Belokoneva

Moscow State University, 119899 Moscow, Russia
*Institute of Crystallography RAS, Leninsky prospekt 59, 117333 Moscow, Russia

ABSTRACT

Brief review on synthesis, thermal stability and systematics of compounds with the $Ca_3Ga_2Ge_4O_{14}$ structure are reported. The crystal structure, isomorphism and order-disorder problems are discussed. Main growth parameters, including congruent melt composition and field of homogeneity, influence of ambient compositions on phase relations in these multicomponent systems, thermal stability of melt and crystal defects are under consideration. Structure-property correlation and some criteria for prediction of piezoelectric properties are formulated.

1. INTRODUCTION

Compounds with the $Ca_3Ga_2Ge_4O_{14}$ structure, discovered and studied in MSU and IC in 80's, are attractive piezoelectric materials now. From about 100 compositions 9, including $La_3Ga_5SiO_{14}$ (LSG or langasite), are grown as single crystals by the Czochralsky method. About 40 new quadruple compositions were synthesized by authors recently additionally, among them aluminates. Some systematics on compounds will be discussed in the paper.

Although 3 inch diameter single crystals of LSG are grown now in several places, there are some problems in receiving high quality material. The homogeneity field for all crystals under consideration is unknown and rather narrow one. Single crystal growth is complicated due to evaporation of some melt components (Ga_2O_3, GeO_2) and their dissociation. Influence of melt and ambient composition and other crystal growth parameters will be discussed.

Internal stresses, inclusions, striations and other defects of single crystals were studied by various optical and X-ray methods.

Some criteria for prediction of piezoelectric activity are formulated on the base of analysis of crystal structure and piezoelectric properties.

2. COMPOUNDS WITH THE $Ca_3Ga_2Ge_4O_{14}$ STRUCTURE

It is known about 100 compounds with this structure. Their compositions may be represented by the next formulas [1, 2]:

1. $A^{2+}{}_3B^{3+}{}_2Ge_4O_{14}$ (A = Ca, Sr, Ba, Pb; B = Ga, Al, Fe,Cr, In) [3];
2. $Na_2A^{2+}Ge_6O_{14}$ (A = Ca, Sr);$Na_{1.8}Ca_{1.1}Si_6O_{14}$ (high-pressure phase [4]);
3. $NaA^{2+}{}_2M^{3+}Ge_5O_{14}$ (A = Ca, Sr, Pb; M = Ga, Al, Fe, Sc, In);
4. $Sr_3B^{2+}Ge_5O_{14}$ (B = Mg, Zn, Ni, Co, Fe, Mn) [3];
5. $Ln_3M^{4+}Ga_5O_{14}$ (Ln = La, Pr, Nd; M = Si, Ge, Ti, Sn, Zr, Hf);
6. $Ln_3M^{5+}{}_{0.5}Ga_{5.5}O_{14}$ (Ln = La, Pr, Nd; M = Nb, Ta, Sb);
7. $Ln_3Al_{3+x}Si_{3-x}O_{12+x}N_{2-x}$ (Ln = La, Ce, Nd, Sm, Gd, Dy, Y; $0 \le x \le 1$) [5].

Recently synthesized compositions [6-8]:

8. $La_3SbZn_3X^{4+}{}_2O_{14}$ (X = Si, Ge);
9. $Sm_3Ga_{5-x}Al_xSiO_{14}$; ($3 < x < 5$);
10. $A^{2+}Ln_2SbGa_5O_{14}$ (A = Ba, Sr; Ln = La, Pr, Nd); $A_{1-2x}La_{2+2x}Sb_{1-x}Al_{5+x}O_{14}$ (A = Ca, Sr, Ba; $x > 0$);
11. $La_2SrGa_4Si_2O_{14}$;
12. $A^{2+}{}_3X^{5+}Y^{3+}{}_3Z^{4+}{}_2O_{14}$ (A = Ca, Sr, Ba, Pb; X = Sb, Nb, Ta; Y = Ga, Al, Fe, In; Z = Si, Ge).

The synthesis of oxynitrides (group 7) and high-pressure silicate (gr. 2) reveals additional possibilities for search of new compositions. $A_{1-2x}La_{2+2x}Sb_{1-x}Al_{5+x}O_{14}$ with more then five Al per formula is the first example of properly aluminates in this structure type (previous attempts to prepare aluminates failed [2]). The idea to search compositions in four-component oxide systems accordingly to a number of cation sites in the structure resulted in discovery of numerous new compositions (gr. 8, 12) [6, 8].

Much more reliable mixed oxide compositions do not produce phases with the $Ca_3Ga_2Ge_4O_{14}$ structure. Garnet structure phase $A_3B_2X_3O_{12}$, which is close in composition to our subject, forms in many cases. Another more stable then $Ca_3Ga_2Ge_4O_{14}$ structures are perovskite-type, monoclinic $Pb_3MgGe_5O_{14}$-type [3] and so on. No $Ca_3Ga_2Ge_4O_{14}$-like compounds occur among phosphates, arsenates and vanadates.

It is not too simple to systemize compounds correctly due to the lack of crystal structure data for the most of compounds. All they, except of gr. 7 (oxynitrides), are mixed oxides. Only compounds of gr. 2 and some compounds of gr. 3, 4 are pure germanates and silicate with all tetrahedral sites occupied by Ge^{4+} or Si^{4+} ions. Compounds of gr. 1, 12 and some of 2, 3, 4 may be described as mixed germanates and silicates (e. g. $Ca_3Ga_2Ge_4O_{14}$ is (calcium) gallo-germanate, $Sr_3NbGa_3Si_2O_{14}$ silico-gallate or gallo-silicate). Compounds of gr. 5 (M= Ti, Sn, Zr, Hf), 6 and 10 are gallates. Compounds of gr.

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

Table 1. Cell parameters (Å) and site occupation in some crystals.

Composition	a	c	Thomson cube 3e	Octahedron 1a	Tetrahedron 3f	Tetrahedron 2d	Refere-nces
$Na_2CaGe_6O_{14}$	8.183	4.836	Na,Ca	Ge	Ge	Ge	9
$Na_{1.8}Ca_{1.1}Si_6O_{14}$	7.903	4.595	Na,Ca	Si	Si	Si	5
$Ca_3Ga_2Ge_4O_{14}$	8.076	4.974	Ca	Ge	Ga,Ge	Ge	9
$Sr_3Ga_2Ge_4O_{14}$	8.270	5.040	Sr	Ga,Ge	Ge,Ga	Ge	9
$Pb_3Ga_2Ge_4O_{14}$	8.417	5.014	Pb	Ga,Ge	Ga,Ge	Ge	10
$Ba_3Fe_2Ge_4O_{14}$	8.541	5.214	Ba	Fe	Fe,Ge	Ge	9
$Ca_3NbGa_3Si_2O_{14}$	8.212	5.000	Ca	Nb	Ga	Si	8
La3Ga5SiO14	8.168	5.095	La	Ga	Ga	Ga,Si	9
La3Ga5GeO14	8.204	5.109	La	Ga	Ga	Ga,Ge	9
$Ln_3Nb_{0.5}Ga_{5.5}O_{14}$	8.235	5.129	La	Ga,Nb	Ga	Ga	9
$Ln_3Ta_{0.5}Ga_{5.5}O_{14}$	8.228	5.124	La	Ga,Ta	Ga	Ga	11
$Nd_3Al_{3.5}Si_{2.5}O_{12.5}N_{1.5}$	7.974	4.875	Nd	Al	Al,Si	Si	5

5 (M= Si, Ge) are mixed gallates. ($La_3Ga_5SiO_{14}$, so called langasite or LGS, is silico-gallate and more correctly to name it "lansigate" or LSG).

Investigation of thermal stability reveals that the most of compounds melt incongruently or decompose in solid phase and therefore are not suitable for melt single crystal growth by the Czochralski method. The congruent and nearly congruent melting compositions are found in gr. 1, 2, 5, 6, 12.

3. CRYSTAL CHEMISTRY

The $Ca_3Ga_2Ge_4O_{14}$ structure-type (space group P321) may be described as a mixed framework consisting of two types of tetrahedra, 2d (symmetry 3) and 3f (symmetry 2), and 1a octahedra (symmetry 32). The holes of framework (Thomson cubes) are occupied by large cations in positions 3e (symmetry 2). The tetrahedra form layers, which alternate along the c axis with the layers of octahedra and Thomson cubes. From two types of tetrahedra 2d are smaller ones because they share the half of edges (as well as octahedra 1a) with large polyhedra, while tetrahedra 3f do not share edges with any polyhedra. Smaller cations prefer 2d sites and larger cations - 3f sites in accordance with ion size. 3f tetrahedra can accommodate so large ions as Zn^{2+} and In^{3+}.

There are about 15 X-ray single crystal structure determinations now. The cation site occupation for some crystals is given in the Table 1. When structure data are absent, cation distribution is tentative.

Analysis of cation distribution (Table 1) indicates that the most of the crystals possess disordered structure because contain two kinds of atoms at least in one of cation sites. Only $Ca_3NbGa_3Si_2O_{14}$ has ordered structure with filling each polyhedron by single sort of atoms. All silicates of gr.12 evidently also have ordered structure. Their Ge-analogues may reveal structure disorder due to possible Y and Ge atoms redistribution over 2d and 3f tetrahedra. Other suspected compounds with ordered structure are $La_3SbZn_3X^{4+}{}_2O_{14}$ (X = Si, Ge) and $Ln_3Ga_5M^{4+}O_{14}$ with M = Ti, Sn, Zr, Hf. Compounds $Sr_3B^{2+}Ge_5O_{14}$ (B = Mg, Zn, Ni, Co) were described as ordered ones [1]. Synthesis of compounds of gr. 8, where all Zn atoms occupy 3f tetrahedra, let to suggest disorder in the structure of $Sr_3B^{2+}Ge_5O_{14}$. Hence, the most compounds with the $Ca_3Ga_2Ge_4O_{14}$ structure are disordered phases.

4. SINGLE CRYSTAL GROWTH

The first experiments on Czochralski growth of $Ca_3Ga_2Ge_4O_{14}$ single crystal were performed at the beginning 1979 and of $La_3Ga_5SiO_{14}$ - two years later. YAG <111> seed was used for LSG crystallization. The boule consisted of several crystallites, and it was possible to cut seeds. Single crystals were received by the seed selection. Single crystal growth of germanates originated from $Ca_3Ga_2Ge_4O_{14}$ seed and gallates - from $La_3Ga_5SiO_{14}$ seed. Single crystals grown in MSU in 80's are designated in the Table 2. Surprisingly that no new single crystal compositions were reported from that time, except of LSG modified by Al and Ti [12].

4.1 Homogeneity field

The Czochralski growth, despite of outer simplicity, is a very sophisticated process, especially in its physical-chemical aspects. It is necessary to know homogeneity field and congruent melting composition for high quality single crystal growth. Unfortunately these data are absent for compounds with the $Ca_3Ga_2Ge_4O_{14}$ structure. The situation is complicated due to evaporation of some melt components (GeO_2, Ga_2O_3).

By [13], the homogeneity field of $La_3Ga_5SiO_{14}$, if exists, is very small. We consider that it is narrow and certainly exists, as indicates the study of growth striations in crystals. The congruent melting composition apparently slightly deviates from the stoichiometric one. Discuss some possible mechanisms of nonstoichiometry in this crystal. It looks possible the vacancies (□) existence found in $Na_{1.8}Ca_{1.1}Si_6O_{14}$ in large holes (sites 3e). Vacancies formation in LSG requires corresponding substitution Ga by Si ($□_xLa_{3-x}Ga_{5-3x}Si_{1+3x}O_{14}$). Further, interstitial sites in tetrahedral layer between

Table 2. Single crystals grown in MSU by the Czochralski method*

Composition	T_m, °C	Diameter, mm	Main body length, mm	Weight, g	Crucible diameter	Ambient
$Ca_3Ga_2Ge_4O_{14}$	1370	25	80	200	42	O_2 (2-2.5 at)
$Sr_3Ga_2Ge_4O_{14}$	1370	25	80	250	42	O_2 (0.5-2.5 at)
$Na_2CaGe_6O_{14}$	1100	20	40	60	42	O_2 (1 at)
$La_3Ga_5SiO_{14}$	1470	45	100	1000	67	N_2:O_2=97:3
$Nd_3Ga_5SiO_{14}$	1470	23	70	200	42	N_2:O_2=97:3
$Pr_3Ga_5SiO_{14}$	1470	8	15	4	42	N_2:O_2=97:3
$La_3Ga_5GeO_{14}$	1470	23	70	200	42	N_2:O_2=97:3
$La_3Nb_{0.5}Ga_{5.5}O_{14}$	1450	30	80	300	42	N_2:O_2=90:10
$La_3Ta_{0.5}Ga_{5.5}O_{14}$	1470	27	80	250	42	N_2:O_2=97:3

* Also single crystals of wide range solid solutions & doped by Nd, Pr, Er, Ho, Cr, Fe, Co, Ni, Mn, Cu. Auto growth machine "Malvern MSR-2" (Metals Research), Pt crucibles, Pt afterheater, pull speed 2.5-5 mm/h (undoped), 1.5-2.5 mm/h (doped and solid solutions), rotation speed 25-50 rpm., interface shape transition, growth duration (including heating and cooling time) 30-48 h.

two Thomson cubes may be in use for La(3e) small disordering without compositional change or for partial occupation by extra La with simultaneous substitution Si by Ga ($La_{3+x}Ga_{5+3x}Si_{1-3x}O_{14}$). If such compositional shifts occur, they are beyond the X-ray structure analysis resolution. Reinvestigation of $La_3Ga_5SiO_{14}$ crystal structure did not show residual extra site electron density [10]. We investigated crystal growth from nonctoichiometric (by 1 mol. % of each component) melts around $La_3Ga_5SiO_{14}$ and did not receive positive results. This means that the deviation from stoichiometry is smaller, may be in order. The same situation exists apparently for $La_3Nb_{0.5}Ga_{5.5}O_{14}$ and $La_3Ta_{0.5}Ga_{5.5}O_{14}$ growth. Difficult to say something about the homogeneity field of $Ca_3Ga_2Ge_4O_{14}$ and $Sr_3Ga_2Ge_4O_{14}$ through the strong GeO_2 evaporation, but in above conditions (Table 2) we received from stoichiometric or with slight (0.1-0.2 mol. %) GeO_2 excess melts high quality single crystals with yield ~80 %. In the case of four-component compounds the homogeneity field may be enough large. For example, the compound $Sr_3NbGa_3Si_2O_{14}$ may be represented as 1:1 solid solution of $Sr_3Ga_2Si_4O_{14}$ and $Sr_3Nb_2Ga_4O_{14}$. Although the last compounds do not exist, $Sr_3NbGa_3Si_2O_{14}$ may have extended solid solution region with both. This discussion on stoichiometry does not concern possible oxygen nonstoichiometry.

4.2 Ambient influence on crystallization process

The ambient choice is very significant in the Czochralski growth of. compounds with the $Ca_3Ga_2Ge_4O_{14}$ structure, because some melt components (GeO_2, Ga_2O_3) dissociate and evaporate during the crystal growth. The ambient composition, found for developed in our group fast growth technique, changes drastically for different crystals (Table 2). Impossible to receive germanate single crystals at low oxygen pressure due to dissociation $GeO_2 \leftrightarrow GeO + O$. It is necessary to increase partial oxygen pressure up to 1 at or even more to suppress GeO_2 dissociation. Enlarged oxygen pressure does not influence appreciably on GeO_2 evaporation rate, but permits to receive high quality crystals and without Pt inclusions. The reason of such difference in crystallization quality, on our opinion, is that the system becomes more complicated (four-component instead of three-component) at melt component dissociation. The congruent composition shifts from the initial one and an admixture removal from crystallization interface need more time. If this time is insufficient (too high pull speed), the crystallization is nonequilibrium and defect one. Growth of large diameter single crystals of germanates looks problematic due to the strong GeO_2 evaporation, despite the lower T_m. The attempt to compensate GeO_2 loss by adding GeO_2 excess results in faster GeO_2 evaporaton, because the excess of volatile is weak connected with the melt. Consolation is that $Sr_3Ga_2Ge_4O_{14}$, possessing the highest from all crystals electromechanical characteristics, has no temperature compensation [14].

The same problems, but connected with the dissociation of Ga_2O_3, appear at single crystal growth of $La_3Ga_5SiO_{14}$, $La_3Nb_{0.5}Ga_{5.5}O_{14}$ and $La_3Ta_{0.5}Ga_{5.5}O_{14}$ (Nb and Ta also can change valence). At lower oxygen pressure (Table 2) poor crystallization occurs. Dissociation of $Ga_2O_3 \leftrightarrow Ga_2O + 2O$ results in complication of phase relation, Pt or Ir appearance in the melt and crystal, and requires pull rate decrease to give the melt possibility

to relax to the crystal composition. At oxygen deficiency Nb and Ta take away oxygen from Ga, but in rich oxygen ambient Nb and Ta containing melts dissolve extra oxygen and evolve it during crystallization with formation of gas voids in crystal. High quality single crystals with yield ~85 % are received at optimal oxygen pressure. Gallate crystals, grown in recommended ambient, are of pale yellow to brownish color.

The single crystal growth from Ir crucibles differs somewhat from growth in Pt crucibles, because Ir at high temperature has poor resistivity against oxygen. The more remarkable are the results received at crystal growth of $La_3Ga_5SiO_{14}$, $La_3Nb_{0.5}Ga_{5.5}O_{14}$ and $La_3Ta_{0.5}Ga_{5.5}O_{14}$ from Ir crucibles in pure nitrogen atmosphere [13]. Large, perfect and nearly colorless single crystals were grown at pull rate 1 mm/h and rotation rate 10 rpm, which are much slower then used by us. The crystal quality decreases as pull rate increases above 1 mm/h. These results cause to suggest that nitrogen, at oxygen absence, plays oxygen part and enter crystal structure (remember oxynitrides). So the crystal composition will be different from initial (e. g. $La_3Ga_{5-x}Si_{1+x}O_{14-x}N_x$). This requires crystallization rate decrease. The growth from Ir crucibles will be different in N_2 and Ar ambient, because Ar cannot enter crystal. The first small colorless $La_3Ta_{0.5}Ga_{5.5}O_{14}$ single crystal was received in Japan in Ar [15]. We consider that nitrogen "contamination" of crystal is possible also in oxygen containing ambient.

4.3 Thermal stability of melt and temperature gradients

Compound may melt congruently and do not grow as single crystal by the Czochralsky method if thermal stability of the melt, related in structure to the solid, is low. There are large temperature gradients in growth chamber, so the melt in whole is overheated against T_m and may lose structure similarity with the crystal. It is possible to crystallize such melt at increased supercooling, but difficult in a single crystal form. We observed such behavior for some compositions. Some compounds or solid solutions have reduced thermal stability. The best way is to grow them at low temperature gradients (10-30 deg·cm^{-1}) and improved temperature control. Low gradient growth reduces evaporation from the melt (the mean melt temperature and convection decrease) and formation of defects (twins, strains, inclusions).

5. DEFECTS IN CRYSTALS

Internal stresses, inclusions, striations, twins and other typical defects of single crystals were observed. Inclusions, both platinum and gas bubbles, usually less than 3 mkm, occur with concentration depending on growth conditions. Some crystals have very low laser beam scattering level. Usually concentration of inclusions increases at bottom part of a boule. Internal stresses were measured by optical and x-ray methods. Best crystals have low residual stresses. High level stresses, observed in twinned crystals, causes crystal cracking during the growth and cooling.

Periodical changing of lattice parameters and refraction index along the crystal growth direction was discovered by optical and X-ray methods. Laser beam after LSG sample along perpendicular to c-axis direction divides on a set of diffraction spots. Angular distances between spots match to volume diffraction grating with period 40-70 mkm. Two complanar reflexes with very close incidence angles were used for X-ray. For each reflex we used separate x-ray counter. Beam dimension was less than 0.1 mm on crystal surface. Angle difference between reflexes is connected with lattice parameters. Period of changing was 40-300 mkm and $\Delta c/c \sim 10^{-4}$. Fifteen $La_3Ga_5SiO_{14}$ crystals, grown from stoichiometic and nonstoichiometric melts, have the similar striation picture.

Twinning in langasite-type crystals, described in [13], usually does not have regular character. Misorientation of crystallites in the boule changes from several minutes through several to tens degrees. Because our crystals, grown at low temperature gradient with flat interface ({0001} facet), are strongly faceted, the twin boundaries, if present, are clearly seen by eye. Sometimes due to low melt stability, melt overheating or power instability (e. g. at interface shape transition) multiple crystallization take place on the periphery of boule. New crystallites can grow along boule or decline.

Chemical etching (etcnant on the base of orthophosphoric acid) reveals dislocations. Dislocation density is large in the central part of crystal (too fast cone growth plus interface shape transition) and low on the crystal periphery.

6. STRUCTURE – PROPERTY CORRELATION

The full set of elastic, piezoelectric and dielectric constants have been determined for $La_3Ga_5SiO_{14}$, $La_3Nb_{0.5}Ga_{5.5}O_{14}$, $La_3Ta_{0.5}Ga_{5.5}O_{14}$, $Ca_3Ga_2Ge_4O_{14}$ and $Sr_3Ga_2Ge_4O_{14}$ crystals [16]. Some data are published for $Na_2CaGe_6O_{14}$ crystal [17]. Between different parameters, the primary interest for materials application attract electromechanical coupling coefficient k, acoustic loss and temperature stability. The values of k_{26} and k_{12} are displayed in the Table 3, including our unpublished results for $Nd_3Ga_5SiO_{14}$, $La_3Ga_5GeO_{14}$ and $Na_2CaGe_6O_{14}$.

Let us try to find the correlation between structural parameters and properties. It is obvious that the ionic radii of the elements in sites 3e and k obey the rule: the more is the ionic radii of the element, the higher is the k (compare Ca^{2+} - Sr^{2+}, Nd^{3+} - La^{3+}). We may expect higher k for compounds with Ba^{2+} (1.42 Å) and Pb^{2+} (1.29 Å), but

Table 3. Electromechanical coupling coefficients k, ionic radii [18] and interatomic distances A-O (Å)

Crystal	3e cation radii	Average framework radii	Average A-O in Thomson cubes: found	calculated	k_{26}, Nm^{-2}	k_{12}, Nm^{-2}
$Ca_3Ga_2Ge_4O_{14}$	1.12	0.45	2.549	2.495.	0.125	0.145
$Sr_3Ga_2Ge_4O_{14}$	1.26	0.45	2.654	2.625	0.260	0.266
$Na_2CaGe_6O_{14}$	1.16	0.42	2.579	2.525	0.109	
$La_3Ga_5SiO_{14}$	1.16	0.46	2.586	2.555	0.134	0.161
$La_3Ga_5GeO_{14}$	1.16	0.48	2.592	2.555	0.111	0.160
$La_3Nb_{0.5}Ga_{5.5}O_{14}$	1.16	0.50	2.598	2.555	0.156	0.153
$La_3Ta_{0.5}Ga_{5.5}O_{14}$	1.16	0.50	2.599	2.555	0.187	0.168
$Nd_3Ga_5SiO_{14}$	1.11	0.46		2.485	0.075	0.106

unfortunately single crystals of $Ba_3Ga_2Ge_4O_{14}$ and $Pb_3Ga_2Ge_4O_{14}$ cannot be grown by the Czochralsky method.

Classical approach for explanation of structure-properties relationship in piezoelectric crystal was suggested by Abrahams [19]. The absolute polarization was related directly to the structural distortions caused in the structure under applied stress on the basis of change of charge distribution. In some examples is was displacement of the cation along the polar direction of the structure, in other - increase of tetrahedra distortion. Both effects - 3e cation displacement along two-fold axis and framework distortion - may be pronounced in the case of structure under discussion.

In the framework built from Ga,Ge-tetrahedra and octahedra the size of the hole with 3e site depends on the size of large cation what is accompanied with the change of unit cell dimensions. The comparison of the average interatomic distances Ca-O and Sr-O with the sum of ionic radii (Table 3) shows less difference for Sr-compound, what may be interpreted as the hole is a bit more large for Ca atoms. In such case Sr has less freedom for the displacement under stress, but Sr compound demonstrates better properties. It seems that the use of ionic radii only does not give simple explanation. Also the Thomson cube is asymmetrical along two-fold axis. We may suppose that the displacement of Sr atom with the higher atomic number (heavy atom), contributes more to the perturbation in the structure under applied stress. The structural defect connected with the disordering of large cations in $Na_2CaGe_6O_{14}$ may be responsible of less k.

The substitution in 3d tetrahedra Si by larger Ge ($La_3Ga_5SiO_{14}$ – $La_3Ga_5GeO_{14}$) decreases tetrahedra distortion and is accompanied by the less piezoeffect. Crystals with Ta and Nb atoms in the 1a sites possess better then LSG characteristics. The polarizability effect known for Nb^{5+}, Ta^{5+} and also Ti^{4+} may play a role. Contrary to piezoeffect, polarizability of Nb^{5+} is higher then of Ta^{5+}; the explanation may be again in a "heavy" Ta. It is impossible to grow $La_3Ga_5TiO_{14}$ by the Czochralsky method due to incongruent melting, but Ti-substituted $La_3Ga_5SiO_{14}$ single crystals have been grown [12]. "Light" Ti^{4+} may be contributes less in k then Nb^{5+} or Ta^{5+}. It seems also that the higher valence of cations in a large holes and octahedra favors the piezoeffect increase.

Investigation of crystal structure under applied pressure (direct piezoeffect) or under electric field (converse piezoeffect) with making clear change of structural peculiarities (framework distortion and atom displacement) may give the answer on the nature of structure-properties relationship for this structural type. Nevetherless, it is clear now, that the improvement of properties may be reached via growing of the compounds having the enlarged ionic radii cations in 3e sites, small ionic radii cations in a framework, high valence and heavy ions in 3e and 1a sites and polarizable cations in 1a site.

Disorder of crystal structure is a kind of defect influencing many physical properties including elastic and acoustic, in particular losses. It ought to expect an improvement of acoustic and piezoelectric characteristics in ordered materials.

It is quite unclear how to control temperature stability of characteristics. Three main La-crystals have temperature compensation, it is absent in $Sr_3Ga_2Ge_4O_{14}$ and may be in other germanate crystals.

7. CONCLUSION

About 100 multicomponent oxide compounds form the $Ca_3Ga_2Ge_4O_{14}$ family. They include a single silicate (high pressure phase), several germanates and gallates, 3 aluminates with all tetrahedral sites occupied by single atom kind (Si, Ge, Ga, or Al). The most of compounds are mixed germanates and silicates with ordered or disordered crystal structure. A special case are oxynitrides.

Only nine individual compounds are grown as single crystals by the Czochralski method and 3 of them, $La_3Ga_5SiO_{14}$, $La_3Nb_{0.5}Ga_{5.5}O_{14}$,

$La_3Ta_{0.5}Ga_{5.5}O_{14}$, have good chance to become 21 century piezoelectric material for digital mobile communication systems and other acoustics applications. To realize 3-4 inch diameter single crystal production some difficulties must be overcame. The main among them are melt and ambient composition control.

Some structure-property correlation is established. To improve piezoelectric properties of materials with the $Ca_3Ga_2Ge_4O_{14}$ structure we recommend to increase cation size in Thomson cubes, introduce into the framework highly polarizable ions and small ions, use high valence and heavy ions in 3e and 1a sites. The materials with ordered crystal structure will have better characteristics.

8. ACKNOWLEDGMENT

This work is supported by the Russian Fund for Fundamental Investigation Grant No 99-2-17242.

9. REFERENCES

1. B.V. Mill, A.V. Butashin, A.M. Ellern, "Germanates with the $Ca_3Ga_2Ge_4O_{14}$ structure," Izv. Akad. Nauk SSSR, Neorgan. Mater., vol. 19, pp.1715-1717, 1983.

2. B.V. Mill, A.V. Butashin, G.G. Khodzhabagyan, E.L. Belokoneva, N.V. Belov, "Modified rare-earth gallates with a $Ca_3Ga_2Ge_4O_{14}$ structure," Dokl. Akad. Nauk SSSR, vol. 264, pp. 1385-1389, 1982.

3. W. Eysel, U. Lambert, B.E. Mayer, C. Renkenberger, "Crystal structures and crystal chemistry of compounds $M_{5-p}T_{4+p}O_{14}$," Z. Kristallogr., vol. 201, pp. 235-251, 1992.

4. T. Gasparik, J.B. Parise, B.A. Eiben, J.A. Hriljac, "Stability and structure of a new high-pressure silicate, $Na_{1.8}Ca_{1.1}Si_6O_{14}$," Amer. Mineral., vol. 80, pp. 1269-1276, 1995.

5. P.-O. Kall, J. Grins, M. Nygren, "Structure of the Nd U-phase, $Nd_3Al_{3.5}Si_{2.5}O_{12.5}N_{1.5}$," Acta Cryst., vol. C47, pp. 2015-2019, 1991.

6. B.V. Mill, T. Fukuda, "New compounds with the $Ca_3Ga_2Ge_4O_{14}$ structure: $La_3SbZn_3Si_2O_{14}$ and $La_3SbZn_3Ge_2O_{14}$,"Rus. J. Inorg. Chem., vol. 43, pp.470-471, 1998.

7. B.V. Mill, E.L. Belokoneva, T. Fukuda,"Novel gallates and aluminates of the $Ca_3Ga_2Ge_4O_{14}$ structure," Rus. J. Inorg. Chem., vol. 43, pp.1032-1037, 1998.

8. B.V. Mill, E.L. Belokoneva, T. Fukuda, "New compounds with a $Ca_3Ga_2Ge_4O_{14}$ -type structure $A_3XY_3Z_2O_{14}$," Rus. J. Inorg. Chem., vol. 43, pp.1168-1175, 1998.

9. E.L. Belokoneva, B.V. Mill, Crystal chemical systematics of minerals. Moscow: Mosc. St. Univ. Ed., 1985, ch. II-3, pp. 140-156.

10. E.L. Belokoneva, B.V. Mill, S.Yu. Stefanovich," Structure peculiarities and some properties of langasite $La_3Ga_5SiO_{14}$ and $Ba_3Ga_2Ge_4O_{14}$.," Rus. J. Inorg. Chem., 1999, in press.

11. H. Takeda, K. Sugiyama, K. Inaba, K. Shimamura, T. Fukuda, "Crystal growth and structural characterization of new piezoelectric material $La_3Ta_{0.5}Ga_{5.5}O_{14}$," Jpn. J. Appl. Phys. P. 2, vol. 36, pp. L919-L921, July 1997.

12. M.F. Dubovik, V.N. Baumer, T.I. Korshikova, T.S. Teplitskaya, "Structural peculiarities of destoichiometrized langasite $La_3Ga_5SiO_{14}$ single crystals modified by aluminium and titanium ions," Functional Mater., vol. 4, pp. 310-312, 1997.

13. B. Chai, J.L. Lefaucheur, Y.Y. Ji, H. Qiu, "Growth and evaluation of large size LGS ($La_3Ga_5SiO_{14}$), LGN ($La_3Ga_{5.5}Nb_{0.5}O_{14}$) & LGT ($La_3Ga_{5.5}Ta_{0.5}O_{14}$) single crystals," in Proc. IEEE International Frequency Control Symp., 1998, pp. 748-760.

14. Yu.V. Pisarevsky, N.A. Moiseeva, B.V. Mill, P.A. Senyushenkov, O.Y. Silvestrova, "Comparision of elastic, piezoelectric, dielectric properties of some materials from $La_3Ga_5SiO_{14}$, $La_3Ga_{5.5}Nb_{0.5}O_{14}$, and $La_3Ga_{5.5}Ta_{0.5}O_{14}$ subgroups," in This Conference Abstract, 1999.

15. H. Kawanaka, H. Takeda, K. Shimamura, T. Fukuda, "Growth and characterization of $La_3Ta_{0.5}Ga_{5.5}O_{14}$ single crystals," J. Cryst. Growth, vol. 183, pp. 274-277, 1998.

16. Yu.V. Pisarevsky, P.A. Senyushenkov, B.V. Mill, N.A. Moiseeva, "Elastic, piezoelectric, dielectric properties of $La_3Ga_{5.5}Ta_{0.5}O_{14}$ single crystals," in Proc. IEEE International Frequency Control Symp., 1998, pp. 742-747.

17. H. Takeda, R. Uecker, M. Kumatoriya, K. Shimamura, P. Reiche, T. Fukuda, "Growth and characterisation of $Na_2CaGe_6O_{14}$ single crystals," Cryst. Res. Technol., vol. 32, pp. 939-945, 1997.

18. R.D. Shannon, "Revised effective ionic radii", Acta Cryst., vol. A32, pp. 751-767, 1976.

19. S.C. Abrahams, "Structure relationship to dielectric, elastic and chiral properties," Acta Cryst., vol. A50, pp. 658-685, 1994.

1999 Joint Meeting EFTF - IEEE IFCS

LANGASITE: WHAT TEMPERATURE COEFFICIENTS OF MATERIAL CONSTANTS ARE CORRECT?

Rinat M. Taziev

Institute of Semiconductor Physics of the Russian Academy of Sciences
630090 Novosibirsk, Russia

Abstract: the paper presents an evaluation of the temperature coefficients of elastic constants of langasite on both BAW and SAW delay-temperature dependence data. Least squares method is used to fit the temperature coefficients of the elastic constants of langasite to the measured BAW and SAW temperature coefficients of delay of the first order (TCD1). For sufficient decrease of computation time, the perturbation procedure is employed. That allows us to fit optimally the constants of crystals in order to get the coincidence with experimental BAW and SAW delay-temperature data using only 1 iteration step in a computer.

2. Introduction

Last time Langasite ($La_3Ga_5SiO_{14}$, LGS) single crystal has attracted an attention as a new piezoelectric material. LGS is a trigonal crystal and belongs to the point group 32 symmetry as α-quartz. LGS has a lower SAW velocity and a higher electromechanical coupling coefficient (K^2) in comparison with quartz. These properties seem useful for designing intermediate-frequency SAW filters with small size and a wide pass-band.

Many numerical investigations have been reported on the SAW propagation characteristics such as phase velocity, K^2, and temperature coefficient of delay on LGS [1,2,3], but only a few reports are found on the experimental studies [4,5]. The experimental results of SAW velocity and K2 are in good agreement with the numerical investigations of the SAW characteristics in Z,Y,X and Z^{+140}- cuts of LGS. However, the experimental values of the SAW temperature coefficient of delay are strongly deviated from numerical ones in the mentioned above substrate cuts and propagation directions of SAW in langasite substrates [4,5].

In the paper we will attempt to find the reason for that large deviation between experimental and numerical values of SAWs TCD. Then we will correct temperature coefficients of elastic stiffness of LGS using the least squares method for fitting numerical values of bulk wave TCD with the measured BAW delay-temperature coefficients of the first order. For sufficient decrease of computation time, the perturbation procedure is employed. That allows us to fit optimally the temperature coefficients of constants of crystal in order to get the coincidence with experimental BAW and SAW delay-temperature data in final step.

3. Sensitivity of SAWs TCD to different temperature coefficients of material constants of langasite

As a start position we will use only recently published "fresh" material constants and their temperature coefficients of langasite [6,7], which are presented in Table 1 and Table 3 (we inversed the signs of C_{14} and e_{14} because they were determined for inverse direction of the X-axes of langasite in [7]).

In our calculations of SAW parameters we use the data of work [7] because they provided an excellent accuracy between experimental and numerical values of bulk wave velocities (deviation din't exceed the value of 3-5 m/s). Also we preliminary computed SAWs TCD in langasite using both material constants [6,7]. As a result we can see from Fig.1 that it is preferable to use data from [7] because they provide better agreement with experiments [4,5].

Using the following thermal expansion presentation of fundamental material constants in series with respect to $\theta=T-To$, $C_{ij}(T)=C_{oij}(1+TC_{ij}\,\theta)$, we may estimate the sensitivity of SAW TCD to variation of the temperature coefficients of all material constants of langasite utilizing the following formulae for $TCD=\alpha_{11}-TCV$, where

$TCV=V'_T/V_o=\sum(a_{ij}TC_{ij}+b_{ij}Te_{ij}+d_{ij}T\varepsilon_{ij})-T\rho/2$,

here, TC_{ij}, Te_{ij} and $T\varepsilon_{ij}$ are the temperature coefficients of the elastic stiffness, piezoelectric, and dielectric permittivity constants, respectively; $T\rho$ denotes the mass density temperature coefficient; α_{ij} are the thermal expansion constants, the coefficients a_{ij}, b_{ij} and d_{ij} may be easy exactly computed using

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

perturbation formulae for relative change of wave velocity V.

For langasite which have the symmetry as α-quartz, TCD depends only on 10 temperature coefficients of material constants: TC_{11}, TC_{13}, TC_{14}, TC_{33}, TC_{44}, TC_{66}, Te_{11}, Te_{14}, $T\varepsilon_{11}$ and $T\varepsilon_{33}$. From comprehensive computations we conclude that the SAWs TCD in langasite is highly sensitive only to 3 temperature coefficients: TC_{66}, TC_{44} and TC_{14}. The coefficients TC_{11}, TC_{33} give contribution to TCD, but not such large as 3 above mentioned one. The contribution of the coefficients TC_{13}, Te_{11}, Te_{14}, $T\varepsilon_{11}$ and $T\varepsilon_{33}$ is small and for roughly estimation of TCD they may be neglected. From that we conclude that discrepancy between numerical and experimental values of TCD mainly is connected only with the accuracy of evaluation of 3 above mentioned temperature coefficients of material constants LGS from bulk wave measurements.

Table 1. The elastic stiffnesses (in 10^{-11} N/m^2) piezoelectric e_{ij} and dielectric constants ε_{ij} (in ε_o).

	[6]	[7]
C_{11}	1.893	1.888
C_{13}	0.9528	0.959
C_{14}	0.1493	0.141
C_{33}	2.624	2.614
C_{44}	0.5384	0.535
C_{66}	0.4216	0.42
e_{11}	-0.431	-0.44
e_{14}	0.108	0.08
ε_{11}	18.97	18.92
ε_{33}	52.	50.7
ρ, Kg/m^3	5743	5743
α_{11},ppm/ °C	5.07	5.11
α_{33}, ppm/ °C	3.6	3.6

Table 2. Experimental values of velocities and their TCV for bulk waves in langasite [8].

Propagation direction	Wave type	Velocity, m/s	TCV, ppm/°C
X	L	5748.7	-21.3
X	FS	3311.5	-43.7
X	SS	2379.6	53.5
Y	QL	5755.3	-29.9
Y	QS	3009.9	-14.1
Y	S	2738.0	13.6
Z	L	6746.7	-40.3
Z	S	3052.2	-27.8

4. Evaluation of the temperature coefficients of elastic stiffness of langasite

Another step is an analysis of all bulk wave measurements, presented in [8], to find inaccuracy of evaluation of the temperature coefficients. For that we used the experimental data from [8], which presented in Table 2. We especially selected only pure mode directions because, probably, they must provide minimal values of errors in the experiments. From given 8 bulk waves temperature coefficients of the wave velocity values we may evaluate 5 temperature coefficients of elastic stiffnesses: TC_{11}, TC_{33}, TC_{14}, TC_{66} and TC_{44}.

For evaluation of the temperature coefficients of the material constants $C_{ij}(T)$ of langasite, it is necessary to minimize the following quadratic residuals by varying the constants TC_{ij},

$$\Sigma[TCD_{cal}{}^{(n)}(TC_{ij})-TCD_{exp}{}^{(n)}]^2=\min$$

here, $TCD_{cal}{}^{(n)}$ and $TCD_{exp}{}^{(n)}$ are the calculated and experimental values of TCD for BAW or SAW on langasite crystal for n-th propagation direction of wave. The temperature coefficients TC_{13}, Te_{11}, Te_{14}, $T\varepsilon_{11}$ and $T\varepsilon_{33}$ are not varied and were taken the same as in [7]. To obtain the temperature coefficients of the material constants of crystal it needs to do only one iteration step in a computer.

New elastic temperature coefficients for langasite are shown in Table 3. These coefficients were evaluated from 8 experimentally measured values of TCD1 for BAW on langasite. Comparison shows that some temperature coefficients obtained are deviated from the data [7].

5. Numerical results

Fig.1 shows a comparison of TCD calculated both with new temperature coefficients of the elastic stiffness and data [1,2] with experimental data on X, Y, Z, Z^{+140}-cuts of langasite [4,5]. One can see that the numerical values of TCD with new temperature coefficients are in much better agreement with the experiments then those with the data [6,7]. Some experimental TCD points are deviated from numerical one (for example, Z^{+140}-cut, $\Psi=16°$, which may be explained by complex behavior of various acoustic modes near this point, (see Fig.3), where the surface wave, which has the real and lowest velocity is transformed to the exceptional bulk wave.

We also in Fig.2 presented contour map results for TCD, power flow angle and coupling coefficient for SAW, where the values of TCD varies from -4ppm to 4 ppm and power flow angle varies from -5° to 5°.

Table 3. First-order temperature coefficients of the elastic stiffness of langasite, 10^{-4}/°C.

	New,BAW	[7]	[6]
TC_{11}	-0.663	-0.675	-0.53
TC_{13}	-0.75	-0.75	-0.88
TC_{14}	-3.348	-3.175	-2.05
TC_{33}	-0.941	-1.025	-1.04
TC_{44}	-0.628	-0.675	-0.62
TC_{66}	-0.1346	0.079	-0.047
Te_{11}	5.87	5.87	4.56
Te_{14}	6.25	6.25	-6.28
$T\varepsilon_{11}$	1.5	1.5	1.37
$T\varepsilon_{33}$	-7.6	-7.6	-7.95

They were computed using new values of temperature coefficients of material constants of langasite. From that we can see that set of useful cuts and SAW propagation directions are located around the experimentally verified Z^{+140}-cut of langasite.

In final step, we also minimized quadratic residuals for BAW TCD together with SAW TCD for more decrease the deviation between experimental and numerical values of TCD. As result we obtain the temperature coefficients which gives slightly better agreement between them. (see Fig.4) These new temperature coefficient data are the following; TC_{11}=-0.641,TC_{13}=-0.75,TC_{14}=-3.509, TC_{33}=-0.91,TC_{44}=-0.587,TC_{66}=-0.1369.

Conclusion

New temperature coefficients of the first order of elastic stiffness of langasite were evaluated from bulk and SAWs TCD measurements on langasite substrates. Numerical values of TCD with new constants are in better agreement with experiments than those with data of [6,7].

Acknowledgements

The author thanks P.P.Turchin for the valuable discussions.

REFERENCES

1. N.F.Naumenko, "SAW and leaky waves in a new piezoelectric crystal of langasite", presented at the ISSWAS, pp.245-251, 1994.
2. I.B.Yakovkin, R.M.Taziev and A.S.Kozlov, "Numerical and experimental investigation SAW in langasite", in Proc. IEEE Inter. Ultrason. Symp.,1995, pp. 389-392.
3. T.Sato, M.Murota, and Y.Shimisu, "Characteristics of Rayleigh and leaky surface acoustic wave propagating on a $La_3Ga_5SiO_{14}$ substrate", Jpn. J. Appl. Phys., Part 1, Vol.37B, pp.2914-2917, May 1998.
4. E. H. Briot, E. Bigler, M. Solal, S.Ballandras, G.Marianneau and G.Martin, "Comparison between theoretical and experimental properties of SAW on Z-cut of langasite", presented at the EFTF, 1997.
5. K.Inou and K.Sato, "Propagation characteristics of surface acoustic waves on langasite", Jpn. J. Appl. Phys., Part 1, Vol.37B, pp.2909-2913, May 1998.
6. S.Sakharov, P.Senushenkov, A.Medvedev, and Yu. Pisarevsky, "New data on temperature stability and acoustical losses of langasite", in Proc. IEEE Intern. Freq. Contr. Symp. 1995, pp.647-651.
7. B.P.Sorokin, P.P.Turchin, S.I. Burkov, D.A.Glushkov, and K.S.Aleksandrov, "Influence of static electric field , mechanical pressure and temperature on the propagation of acoustic waves in $La_3Ga_5SiO_{14}$ piezoelectric single crystals", in Proc. IEEE Int. Freq. Contr. Symp., 1996, pp.161-169.
8. P.P.Turchin, "Nonlinear electromechanical properties and effects of external static loading on propagation of acoustic waves in $La_3Ga_5SiO_{14}$ ", Ph.D. thesis, Krasnoyarsk, 1997.

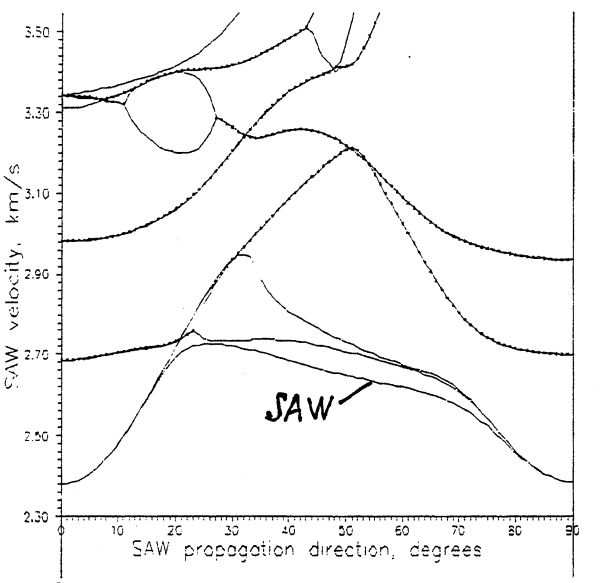


Fig. 2. Behavior of some branches of acoustic waves which satisfy both wave equation and boundary conditions on the free surface of Z^{+140}-cut of langasite as a function of propagation direction.
-*-* waves with real and imaginary parts of velocity
____ waves with only real value of velocity.

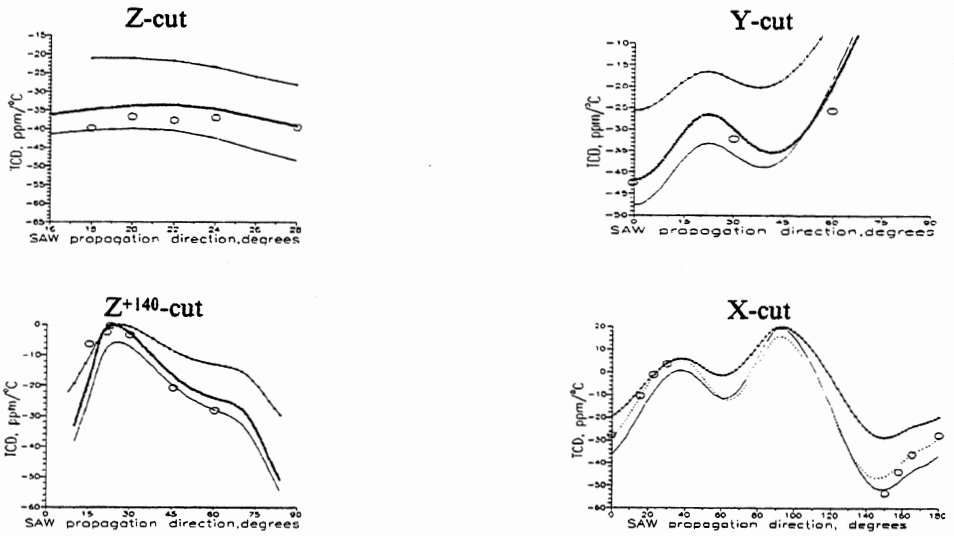


Fig.1. Numerical and experimental TCD values for SAW on different cuts of langasite.
-*-*- TCD with data [6], _____ TCD with data [7], ------ TCD with new data evaluated only from BAW TCD experimental data.

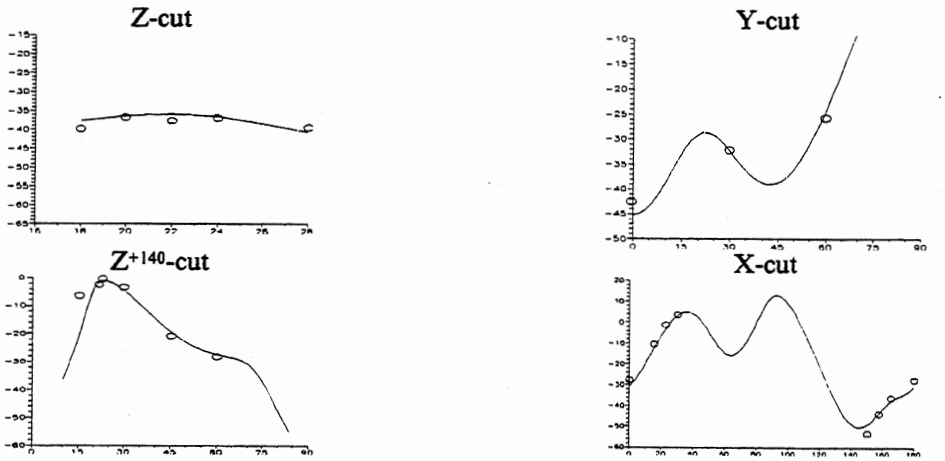


Fig.4. Numerical and experimental TCD values in langasite. Numerical TCD is computed with new temperature coefficient evaluated from the BAW +SAW experimental data.

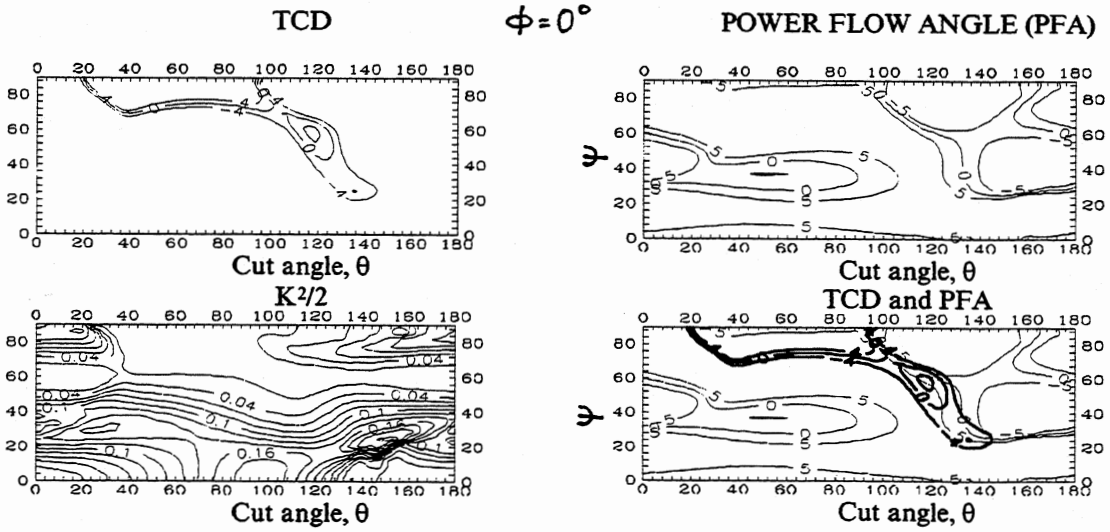


Fig.3. Contour maps for TCD,power flow angle and coupling coefficients ($K^2/2$) for SAW in langasite.

INVESTIGATION OF PHASE FORMING AT SYNTHESIS OF LANGASITE – PIEZOELECTRIC

M. F. Dubovik, T.I. Korshikova, O. M. Proskurnya*
Institute far single crystals of NAS of Ukraine,
310001, Kharkiv, Lenin Av., 60, Ukraine ; e-mail < dubovik @ isc. kharkov. ua >,
* State Politechnical University,
310002, Kharkiv, Frunze St., 21, Ukraine; e-mail < oles @ kpi. kharkov. ua >.

ABSTRACT

General thermodynamic characte-ristics of langasite S^0_{298} , ΔH^0_{298} , ΔG^0_{298} , $\Delta G(T)$, $C_p(T)$ and melting points of the eutectics were calcu-lated. It was established that optimal temperature of langasite synthesis T≡1723 K; the ions gallium and silicon are most mobile reagents.

1. INTRODUCTION

The oxides of lanthanum La_2O_3, gallium Ga_2O_3 and silicon SiO_2 are initial components of gallium silicate of lanthanum $La_3Ga_5SiO_{14}$ (langacite, LGS) , which single crystals have effective piezoelectric properties and serve laser medium [1, 2, 3].

Determining requirement of high-quality single crystals generation is the synthesis of initial charge, which is identical on composition and symmetry to a crystal. Especially it is important in case of low-symmetric crys-tals of multicomponent compound cultivation, such as LGS. It is caused by presence of free components in charge that leads to additional reactions in a melt that are accompanied by high volatility, interaction with a material of a crucible and, as the consequence, distur-bance of phase homogeneity of the crystal. The purpose of the present study is : carrying out a thermodynamic estimation of course of lanthanum-gallium silicate for-mation solid-phase reaction; finding requirements for LGS monophase charge synthesis; study of interaction interaction features of lanthanum, gallium and silicon oxides.

2. THEORY AND EXPERIMENTS

We have evaluated possible eutectic points of binary pseudo-systems, which are formed by LGS and other binary compounds included in ternary system La_2O_3 - Ga_2O_3 - SiO_2, such as $La_2Si_2O_7$ (T_{fu}=2010K), $La_{14}Si_9O_{39}$ (T_{fu}=2253K), $LaGaO_3$ (T_{fu}=1463K) and Ga_2O_3 (T_{fu}=1998K). The attempt to estimate direction and course of solid-phase reaction has been made by calculation the high-temperature heat capacity equation and dependence of the Gibbs energy on temperature. The given calculations allow defining probability of solid-phase reaction course and temperature band for technological process of generation LGS material with predetermined properties. The estimation of eutectic temperatures in binary and triple systems was carried out on the equation Abstain-Howlend and Le Shatelje [4] which canonical form is:

$T_i = T_{fu}(i) / 1 - \ln x_i N_i^{-1}$ (1),

where:

T_i - current liquid's temperatures, K;

$T_{fu}(i)$ – me-lting point of the i builde , K ;

x_i - molecular ratio of the i builder;

N_i - atoms number in formula unity.

Solving the set of equation (1) by metod [5] has made the estimation of eutectic compositions and tempera-tures in multicomponent mixtures.

The results of the calculations made are given in table.

Table. The estimation of eutectic compositions and temperatures in multicomponent mixtures.

N/N	LGS, mol. %	$LaGaO_3$, mol. %	$La_{14}Si_9O_{14}$, mol. %	$La_2Si_2O_7$, mol%	Ga_2O_3, mol%	T_{fu}, K
1	99,91	-	0,09	-	-	1743
2	1,20	98,90	-	-	-	1460
3	1,10	98,80	0,20	-	-	1460
4	99,95	-	0,05	-	-	1743
5	83,20	-	-	16,80	-	1729
6	83,20	-	0,40	16,40	-	1729
7	58,00	-	-	-	42,00	1703
8	83,20	-	-	16,80	-	1729
9	47,70	-	-	12,30	40,00	1689
10	58,00	-	-	-	42,00	1703
11	1,10	98,90	-	-	-	1460

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

As it is visible from the submitted results, eutectic temperatures of LGS and others binary compounds of system La_2O_3 - Ga_2O_3 - SiO_2 lay in a temperature interval 1460-1729K and along with LGS some binary compounds can be formed. Therefore during LGS synthesis it is necessary to adhere strictly to calculated composition of starting builders in charge and to their chemical cleanliness.In the absence of data to reference entropy $La_3Ga_5SiO_{14}$, enthalpy of formation, the Gibbs energy and high-temperature heat capacity equation, we made evaluating calculation of the LGS basic thermodynamic performances. LGS generation occurs on reaction:

$1.5La_2O_3 + 2.5Ga_2O_3 + SiO_2 = La_3Ga_5SiO_{14}$ (2)

The ΔH^0_{298} defining was carried out on the following formula:

$\Delta H^0_{298(LGS)} = \Sigma\Delta H^0_{298(oxide)} + 0.05\Sigma\Delta H^0_{298(oxides)}$ (3)

After substitution of values which correspond to starting oxides, we have obtained numerical value of enthalpy: $\Delta H^0_{298(LGS)}$ = -6679.57 kJ/mole. For obtaining the more precise performance of required value the calculation S^0_{298} was made on four methods: of K.B. Yacimirski [6], of I.D. Wood [7], of N.A. Landia [8] and on method of the total of oxides. S^0_{298} = 463,29 J/moleK

As LGS is the composite oxygen compound, which doesn't have any composite polymorph transmutations, consists of oxides and melts congruently, calculation the high-temperature heat capacity equation for $La_3Ga_5SiO_{14}$ was ordered according to the V rule of thermodynamics [9] :

C_p =495.18+0.090416286T– 6533292.26T^2 J/moleK (4)

The temperature dependence of the Gibbs energy was calculated up to the formula [9]:

$\Delta G_T^0 = \Delta H^0 - \Delta aTlnT - 0.5\Delta bT^2 - 0.5cT^2 + yT$ (5), where :

$\Delta H^0 = \Delta h_{298}^0 - \Delta a_{298}ln298 - 0.5\Delta b\,(298)^2 + c\,(298)^{-1}$ (6)

and y was found from expression (7) :

$\Delta G_{298}^0 = \Delta H^0 - \Delta a_{298}ln298 - 0.5\Delta b(298)^2 + \Delta c(298)^{-1} + 298y$.

The $\Delta G = \Delta f(T)$ dependence equation coefficients were obtained on the basis of calculations and reference data and the equation has a form:

ΔG_T^0 = –321444.03+18.97TlnT–1.93·$10^{-3}T^2$–13.18T^{-1} – -19248.74T J/mole (8)

The graph of dependence of the Gibbs energy on temperature is given in a fig.1.

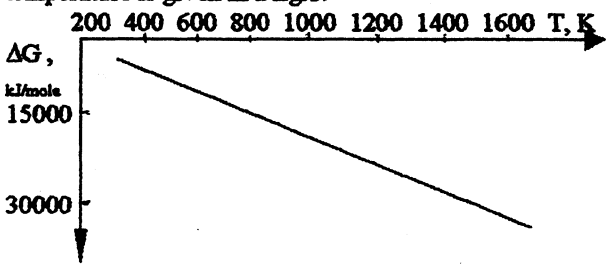


Figure 1. Dependence of the Gibbs energy on temperature

On the basis of physicochemical examinations including differential thermal analysis (DTA), dilatometric analysis (DLA) and X-ray phase analysis (XRPA) the temperature band of initial components intensive interaction has been determined.

DTA was carried out on derivatograph of a OD-103 type of firm "MOM" with rise in temperature about 7.5K/min in a band 293-1773K. The etalon was α-aluminium-oxide. XRPA was executed on X-ray diffractometer of a type DRON. Dilatometric measurements were carried out on differential dilatometer Netzsch 403 ED (Germany) on air with rate of heating 5 K/min.

For examinations were used samples as tablets with 15 mm diameter and 2-4 mm height, which were produced by a molding at pressure 100 mPa/cm. As binder substance was used three-percentage water solution of vinyl alcohol.

3. RESULTS

At DTA examination stochiometric mixture of oxides La_2O_3, Ga_2O_3, SiO_2 in the area of low temperatures two endothermic effects were fixed (at 653K and 823K, fig. 2, curve 2); the same effects were observed on DTA-curve of oxide of lanthanum (fig. 2, curve 1). In the area of high temperatures (up to XRPA data) the endothermic effect at 1663K corresponded to formation solid-phase crystallization (fig. 2, curve 2).

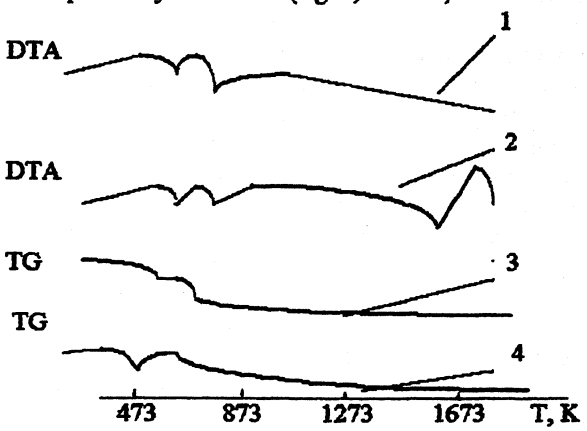


Figure 2. DTA curves; 1.3 – La_2O_3; 2,4 - $La_3Ga_5SiO_{14}$

The dilatometric data analysis for samples of the same composition showed that the contraction began at 1223K, was intensively passing with rise of temperature and $\Delta l / l$ achieved 25 % at 1673K (fig. 3), that indicates a very high relative affinity of builders mixture. These data specify that the solid-phase langasite synthesis should be carried out at temperature higher than 1683K with slow rise of temperature and particular time delay in contraction interval. The maximal content of LGS phase in charge, which had been made as a powder, was 92%. Tabletting of previously burnt mixture of starting oxides (T = 823K) with the subsequent high-temperature tablet annealing

(T = 1723K) allowed to receive monophase fusion mixture LGS.

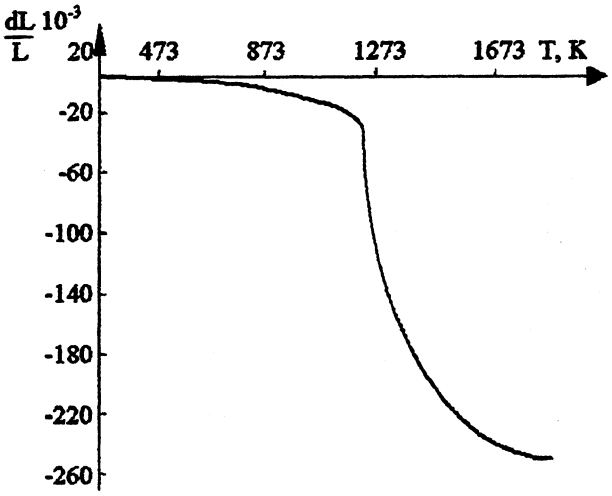


Figure 3. Curves of LGS caking

Mass transfer in LGS generation was studied in double systems La_2O_3 - Ga_2O_3 and $LaGaO_3$ - SiO_2 , as since according to the equilibrium diagram of system La_2O_3 - Ga_2O_3 - SiO_2 [2] gallates of silicon are not formed, and LGS can be obtained from gallate of lanthanum (GL) and dioxide of silicon. As experimental samples were used tablet surfaces, which had been press-fitted from previously synthesized gallate of lanthanum and oxide of gallium and when burnt for exception of contraction and deforming. The oxides of lanthanum and silicon were utilized as powders. After grinding and weighing tablets platinum layer were superimposed on their surfaces as an inert score. Tablets GL were covered by the crushed oxide of silicon and tablets oxide of gallium – by oxide of lanthanum and they were burnt during 5 hours at temperature of 1723K and 1573K correspondingly. Tablets GL and Ga_2O_3 were burnt simultaneously for taking into account mass loss as a result of builder fugitiveness. After burning mass of tablets GL in covering increased and Ga_2O_3 – decreased. The X-ray diffraction defining of compound, which had formed as powdery deposit on tablet surfaces, showed that the layer LGS was generated on tablets GL and GL was generated on tablets Ga_2O_3. These experiments specify preferred diffusion ions of silicon and gallium in solid-phase interaction process.

4. DISCUSSION

The analysis of the obtained results showed : theoretically beginning of reaction in system La_2O_3 - Ga_2O_3 - SiO_2 is possible from temperature of 500K and the reaction goes up to the end;dependence of Gibbs energy on temperature in a wide temperature interval indicate high probability of reaction course; monophase charge maybe received by synthesis at T ≅1723K.

With the Gibbs energy diminution the reaction rate decreases, that character of reaction rate changes. It is possible to assume, that this fact is explained by that LGS is composite crystalline substance with random structure has narrow congruente area with the possible doubles [10], which transition is not adequate to ΔG change in the course of time. In this case with ΔG diminution the reaction rate decreases, i.e. minimally inconvertible compound corresponds to maximal rate value, and maximum inconvertible one – to minimum rate value. Apparently, the transition less inconvertible modification to more inconvertible occurs with change of the Gibbs energy up to the V.Osvald steps rule [9].
Determination of the basic thermodynamic performances of LGS generation and examination of features of the solid-phase interaction mechanism of oxides of lanthanum, gallium and silicon allowed to develop synthesis monophase charge procedure with a diversion of the basic builders content from designed one no more than 0.5mass% for lanthanum and gallium and 1.5% for silicon with the content of checked impurities Al, Fe, Mg, Mn, Cu, V, Ni no more than 10^{-4} mol.% everyone. Use of such charge enables reproducibly grows high-quality single crystals LGS up to 60-mm diameter [11].

Nonstohimetricity of LGS single crystals can be well seen in the spectrum of their optic transmittance (fig. 4). For single crystals withe diviation of 0,5% for gallium and of 1,9% for silicon the self-absorption edge moves to the field of long waves for Δλ≈0,5mkm and goes dour the level of transmittance in the interval of the measured wavelengths (curve 2) in comparison with LGS single crystals of stohimetric composition (curve 1).

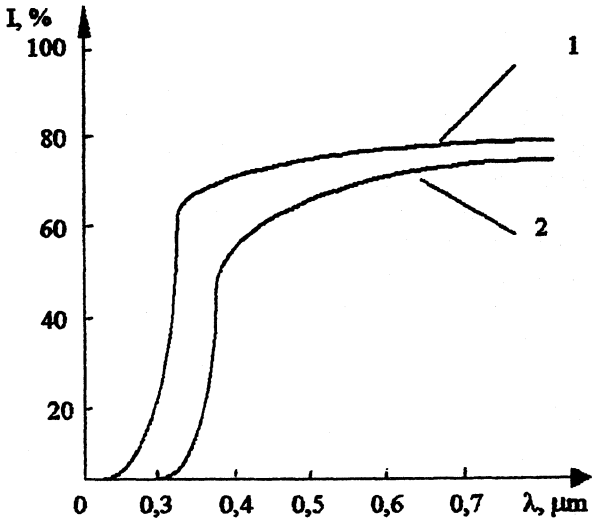


Figure 4. Spectrums of optic transmittance of the langa-.site single crystals at 300 K :1- $La_3Ga_5SiO_{14}$ (stohiometric) ; 2 - $La_3Ga_{4,975}Si_{1,01875}O_{14}$; $t_{(0001)}$=2mm

5. CONCLUSIONS

1. The estimation of eutectic temperatures in binary subsystems, which is produced by LGS with four compounds in the limits of ternary system La_2O_3 - Ga_2O_3 - SiO_2, has been made. A thermodynamic estimation of course solid-phase reaction of formation LGS has been made for the first time. Its basic thermodynamic performances have been designed: the enthalpy, the entropy, the Gibbs energy and also the high-temperature equation of heat capacity has been deduced.
2. Has been detected that the most mobile reagents in a mixture of lanthanum, gallium and silicon oxides at its high-temperature processing are the gallium and silicon ions.
3. The conditions of LGS synthesis have been ascertained.

REFERENCES

[1] А. А. Каминский , С. Э. Саркисов, Б. В. Миль, Г. Г. Ходжабагян , Новый неорганический материал с высокой концентрацией ионов Nd^{+3} для получения стимулированного излучения на переходах $^4F_{3/2} \rightarrow {}^4I_{11/2}$ и $^4F_{3/2} \rightarrow {}^4I_{11/2}$. Неорганиче-ские материалы , 1982, т. 18 , № 8 , с.1396 – 1397 .

[2] A. A. Kaminskii, B. V. Mill et al., Investigation of trigonal $(La_{1-x}$ Ndx$)_3Ga_5SiO_{14}$ crystals.Phys Stat . sol (a) , 1983, vol. 80 , N 1, pp. 387 – 396.

[3] И. А. Андреев, М. Ф. Дубовик , Письма в ЖТФ, 1984, т.10 , № 8 , с. 487 – 492 .

[4] L. F. Epsein, W. H. Houland, J. Amer. Ceram. Soc., 1953, vol. 36, N 10, p. 101 .

[5] Я. Н Питак, Г. Н. Шабанова , В. В. Тараненкова, тез. докл. Международной конференции «Компьютер: наука, техника, технологии, здоровье » , Харьков,1993, с. 142-143.

[6] А. С. Бережной , Многокомпонентные системы оксидов, Наукова думка, Киев ,1970, с. 544.

[7] Д. Вуд, Б. Фрейзер. Термодинамика для ионов , Москва, «Мир», 1981, с. 184 .

[8] Н. А. Ландия . Расчет высокотемпературных теплоемкостей неорганических веществ по стандартной энтропии , Тбилисси, АН.Груз.ССР 1962, с. 222 .

[9] В. И. Бабушкин, Г. М. Матвеев, О. П. Мчедлов-Петросян, Термодинамика силикатов, Стройиздат, 1986, с. 408.

[10] B. Chai, J. L. Lefancheur et al. Proccedings of the Conference 1998 IEEE Internet. Freq. Contr. Symposium, USA, 1998, p. 748-760 .

[11] М. Ф. Дубовик, Неорганические материалы, 1992, т. 28, № 9, с. 1999-2006.

MODERN STATE AND PERSPECTIVES OF BULK ACOUSTIC WAVE COMPOSITE RESONATORS

G.D.Mansfeld

Institute of Radioengineering and Electronics RAS. 11 Mokhovaya st., 103907 Moscow, Russia

ABSTRACT

The review of the modern state and the perspectives of bulk acoustic wave composite resonators is presented. Especial attention is paid to the following topics: what composite resonators are, modern composite resonators including multyfrequency resonators, the thin-films resonators with and without acoustic isolation from the substrate. A particular review of the microwave bulk acoustic spectroscopy of the materials - the measurements of attenuation and elastic coefficients is presented.

1. INTRODUCTION

First information about composite acoustic structures appeared in the 30 - 40th [1]. The interest to these structures was caused by the necessity to decrease the operating frequencies of receiving and radiating ultrasound elements and their acoustic impedance and impedance transformation [2]. Nowadays the interest to the composite acoustic structures (resonators) is connected with the possibility to develop some microwave devices based on layered- and multilayered acoustic structures [3]. Three classes of the BAW microwave resonators have been developed. The first one is high overtone BAW composite resonators schematically shown in Fig.1a. The second one is membrane resonators. The third one is so-called solidly mounted resonators (thin-film piezoelectric resonators acoustically isolated from the substrate by a set of quarter-wavelength layers). The special interest to the BAW composite resonators is caused by the possibility of their application to the acoustic spectroscopy of thin films and layers and the materials with very high attenuation [4]. A short review of all the problems listed is presented.

2. HIGH OVERTONE BAW COMPOSITE RESONATORS

The tendency to increase the operating frequencies of the piezoelectric resonators requires principally new technical solutions of this problem. As on microwaves the piezoelectric plates must be very thin the resonators made of such thin plates are very fragile and the required parameters are very difficult to reproduce. In order to keep high values of the quality factor Q it is necessary to use much more acoustically transparent materials than quartz. As a result Qf product in these materials is higher. YAG ([110], $2.0\ 10^{14}$), spinel ([111], $4.5\ 10^{14}$), Al_2O_3 ($3.3\ 10^{13}$), $LiNbO_3$ [111], $1.6\ 10^{14}$ are more preferable than traditional quartz (AT, Qf = $1.4\ 10^{13}$). The composite resonators made of a relatively thick plate and thin piezoelectric layer (transducer) is in fact a very wide band multyfrequency resonator. The transducer loaded by the plate has a wide bandwidth (up to a few GHz). The resonances in the structure occur each time when the total phase gain θ for the acoustic wave traveling through the structure is equal to $n\pi$, where n is integer. So one can observe a set of tens or hundreds of resonant peaks. The distant between them of a few MHz or tens MHz depending on the plate thickness. To obtain the equivalent circuit diagram of the high overtone BAW resonators we use a well known expression (1.4.23) [5] for the input electric impedance of the piezoelectric loaded by two acoustic loads. The first one is the input impedance of the electrode placed on the plate of the material with low acoustic losses covered by a layer with acoustic impedance Z_L (this layer can be used for frequency adjusting or this is the layer under investigation). The second one is the input impedance of the outer electrode. After some algebraic manipulations it may be represented in the form:

$$Z_e = \frac{1}{i\omega C_0'} + \sum_{n=1}^{\infty} \frac{i/\omega C_n}{\left[\left(\frac{\omega_n}{\omega}\right)^2 - 1\right] + \frac{i}{\omega C_n r_n}}$$

where

$$C_0' = C_0\left\{1 + \alpha/2 - \left(K_t^2/q'd\right)\left[\sin q'd + \right.\right.$$
$$\left.\left. + \left(1 - \cos q'd\right)\sin\left(q'd + 2\phi_2\right)\right]\right\}^{-1}$$

$$r_n = \frac{\alpha}{2y\omega_n C_0},\quad L_n = \frac{\alpha}{\omega_n^2 C_0 \pi n},\quad C_n = \frac{C_0}{\alpha}\pi n$$

$$\alpha = \frac{8K_t^2}{\pi n}\sin^2\frac{q'd}{2}\sin^2\left(\frac{q'd}{2} + \varphi_2\right)$$

The coefficient α only takes into account the electromechanical coupling and properties of the piezoelectric layer and the outer electrode. This formula is very convenient for analysis. It corresponds to the equivalent circuit diagram shown in Fig.1b. All the components under the sum correspond to the parallel

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

resonances which occur at the frequencies $\omega = \omega_n$. These resonant frequencies are found from the condition $\mathrm{Re}\,\theta = n\pi$ where

$$\theta = qd + \phi_1 + \phi_2$$

$$\phi_1 = \arctan\left[\frac{Z_1}{Z_t}\tan(\varphi_1 + w_1)\right]$$

$$\varphi_1 = \arctan\left[\frac{Z_0}{Z_1}\tan(\varphi_0 + bl)\right]$$

$$\varphi_0 = \arctan\frac{Z_L}{Z_0}\tan\delta L,\ \phi_2 = \arctan\frac{Z_2}{Z_t}\tan w_2$$

C_0 is the static capacitance of the transducer, K_t is the stiffened electromechanical constant, q, b, β_1, β_2, δ are the complex vectors (taking into account the losses in the layers e.g. $q = q' - iq''$) for the piezoelectric layer, plates, electrodes, and additional layer, d, l, t_1, t_2, L are the corresponding thickness of the layers, $w_1 = \delta_1 t_1$, $w_2 = \delta_2 t_2$. The total losses in the structure (y in the expression for r_n) are calculated as $y = \mathrm{Im}\,\theta$. The quality factor of the resonator on frequency ω_n is equal to $Q \approx \pi n/2y$, so it is evident that in order to obtain high Q-values the total losses in the structure must be as low as possible. Usually the losses in the transducer material are much higher than that in a specially chosen crystalline plate. To minimize the influence of the transducer acoustic losses it is necessary to select the thickness of the plate much thicker than that of the transducer. It is the usual way to obtain a high Q-value. The results obtained by different authors [6-9] are: for spinel [111] Q≈2 10^4, f=1.28 GHz, for Al_2O_3 Q=6564, f=7.66 GHz. The increase in Q is expected at low temperatures [10]. Some results were confirmed in [11] where for YAG crystals at T = 30K, the quality factor was experimentally found Q =310^5. For highly acoustically transparent materials the losses due to the scattering on the surface roughness may exceed the lattice losses. In accordance with [12,13] they grow as the square of frequency. The quality of the surface preparation should be perfect.

The resonant frequencies of the peaks may be controlled within some limits. The frequencies of the series resonances may be tuned in small intervals by series capacitance included into the circuit. Another original way to change the frequency is to substitute the layer L by an electromechanical transducer. The electric load connected to this transducer effects the reflection coefficient and causes the frequency change [14]. The wide band frequency tuning may be performed using magnetic materials in one of the layers. Frequency control is realized through the change in the acoustic wave velocity caused by the applied magnetic fields. Experimentally the properties of this effect were investigated in [9] . The theory giving a detailed description of the phenomenon has been developed in [15] .

If to make the layer L in BAW composite resonator of the material which can absorb the molecules of the gases the resonator will serve as a gas sensitive element - the resonant frequencies may be changed due to gas absorption. The resonant properties of this layer may play a dominant role in the value of sensitivity. The example showing the dependence of the frequency change of the gas sensors from the thickness of the layer for the same ambient gas concentration is presented in Fig.2. It is seen that the proper chose of the sensitive layer thickness makes it possible to increase the response many times [16].

3. THIN-FILM RESONATORS

In recent years many publications have been devoted to membrane resonators - composite structures which consist of a thin semiconductor layer (Si, GaAs) prepared by the method of selective etching and covered with thin piezoelectric (ZnO, AlN) and electrode films. The main resonances in such structures are usually within the 0.4 - 2.5 GHz region. Though because of high losses in the materials in use these structures cannot provide for high values of Q-factor (Qf < 10^{13}) such resonators can be used in sensors.

Another way to create the thin-film piezoelectric resonators operating on fundamental frequencies on microwaves is to use the effect of the acoustic isolation of the resonators from the substrate by a set of λ/4 layers (solidly mounted BAW thin-film resonators) [17]. The transforming properties of one layer with the impedance z_a placed on the substrate with impedance z_s for λ/4 thickness are $z_{in}^{(1)} = z_a^2/z_s$. For the two layers (the second quarter-wavelength layer with impedance z_b placed between the substrate and the layer with impedance z_a) impedance is $z_{in}^{(2)} = (z_b/z_a)^2 z_s$. For n pairs of the layers input impedance is $z_{in}^{(2n)} = (z_b/z_a)^{2n} z_s$. So, if one wants to decrease input impedance using an even number of the pairs one should take the ratio (z_b/z_a) as small as possible. Transforming properties of the layers strongly depend on the acoustic impedance ratio of the layers. To support the value of Q ≈10^4 in case of $LiNbO_3$-$LiTaO_3$ (z_b/z_a)=0.7 the number of the layers must be N>18. The detailed theory of the resonators acoustically isolated from the substrate is developed in [18]. The first experimental realization such type of the resonators

and filters based on such resonators on microwaves was demonstrated in [17] (up to 2.5 GHz). Some resonators with AlN-SiO_2 pairs were investigated in [19] (f=2 - 2.3 GHz, Q≈10^3). The Q-factor in such experiments was restricted in part by the losses in amorphous or texture films used. It is evident that in order to improve Q-factor it is necessary to develop the technology of the monocrystalline films [20]. The potential advantages of such type devices are: they can be fabricated using the thin-film technology, the requirements to the substrate material and the accuracy of λ/4 layer thickness are not very rigorous, they have a very low weight and extremely small dimensions typical for all the composite resonators on microwave.

4. BAW RESONATOR SPECTROSCOPY

The information about sound velocities and acoustic wave attenuation coefficient is contained in Θ where all the wave vectors are complex). Using the expressions for ReΘ and ImΘ and the results of the experimental measurements of the resonant frequencies and bandwidths of the resonances it is possible to investigate the acoustic properties of the layers and the films composing HBAR resonator structures. The method of acoustic characterization of the materials using BAW resonator spectroscopy is presented in [21]. In fact this type of spectroscopy transfers the ideology of the acoustic losses and sound velocities measurements from time domain («echo-pulse» method) to frequency domain. The advantages of the method suggested are: it makes it possible to characterize very thin plates and even thin films, to measure the frequency dependences of the attenuation coefficients in a wide band using the same sample (due to wide band excitation of the multiresonances), to measure very high values of the acoustic losses in the materials. This method was successfully used for acoustic characterization of LGS [22], $GaPO_4$ [23], quartz [24], high losses in polycrystalline ferrites [25] etc. As an example the frequency dependence of the losses in Lengmuir-Blodgett film on microwaves is shown in Fig.3. It demonstrates the possibility of the measurement of very high values of losses in thin films. Another example of BAW spectroscopy application is the investigation of carbon nanotube films (tubelens) deposited onto the surface of YAG.

The recent results of carbon nanotudes investigations are supported by the RFBR grant 99-02-17602.

REFERENCES

1. L.Bergmann, Der ultraschall und seine anvendung in wissenschaft und technik, Zurich, 1954.
2. L.M. Brekhovskikh, Waves in Layered Media. (Academic press, New-York, 1960)
3. Gagnepain J.J.. Recent Advances in Subminiature/Thin Membrane Resonances// Ultrason. Symp.:Proc. -1985.-P.291-301.
4. G.D.Mansfeld, A.D.Freik, B.N.Krutov. «High overtone BAW resonator as a tool for acoustic characterization of thin films and layers», in Proceeding of the Forum EFTF 93, Neuchatel, Switzerland, March 16 - 18, 1993, pp.261-266.
5. G.S.Kino, Acoustic waves: devices, imaging, and analog signal processing. Prentice-Hall Inc., Englewood Cliffs, NJ.
6. Moore R.A., Haynes J.T., McAvoy B.R. High Overtone Bulk Resonator Stabilized Microwave Sources // Ultrason.Symp.: Proc.- 1981.- P.414-424.
7. Bailey D.S., Driscoll M.M., Jelen R.A., McAvoy B.R. Frequency Stability of High Overtone Bulk Acoustic Resonators// 1990, Ultrason.Symp.:Proc. - 1990. - P.509-512.
8. Driscoll M.M., Jelen R.A., Matthews N. Extremely Low Phase Noise UHF Oscillators Using High Overtone, Bulk Acoustic Resonators// Ultrason.Symp.:Proc.- 1990. - P.513-518.
9. Salvo H.L., Jr., Moore R.A., Adam J.D., McAvoy B.R.. Properties of Tunable YIG HBAR'S// Ultrason.Symp.:Proc. - 1987. - P.337-340.
10. S.N.Ivanov, V.V.Medved, «Unharmonicity constant and attenuation of longitudinal acoustic waves in solid solutions of $Y_{3c}Lu_{(3-3c)}Al_5O_{12}$. Fiz.Tv.Tela, (Sov), vol.25, pp.2907-2910, October, 1983.
11. McAvoy B.R., Salvo H.L., Jr. Cryogenic Experiments with Microwave Acoustic Resonators// Ultrason.Symp.:Proc. - 1986. - P.343-347.
12. V.V.Kosachev, Yu.N.Lohov, V.N.Chukov, BAW scattering on statistically rough free boundaries of solids, Sol. State Phys. (Sov.) v.31,p.105-113. 1989.
13. S.N.Ivanov, E.N.Khazanov, Scattering of BAW by rough surface. Radiotechnics and Electronics (Sov.) v.2, p.402-408, 1981.
14.E.S.Kucheryavaya, G.D.Mansfeld, B.V.Sveshnikov, A.D.Freik, «Electrically tuned BAW composite resonators», Akusticheskii Zhurnal (Sov), vol.pp.346-348, Febryary, 1995.
15. G.D.Mansfeld, N.I.Polzikova, O.N.Kudryavtseva, On the spectra of BAW resonator with ferrite layer, Proc. International Symposium on acoustoelectronics, frequency control and signal generation, 17 - 19 Sept., 1996, Moscow, pp.52 - 56
16. S.G.Alekseev, S.V.Boritko, G.D. Mansfeld, Tech.Phys.Letters, v.25, October, 1999, to be published.
17. K.M.Lakin at all, Solidly mounted resonators and filters, Proc. IEEE Ultrason. Symp., 1995, pp. 905-908.
18. G.D.Mansfeld, S.G.Alekseev, I.M.Kotelyanskii, Bulk acoustic wave microwave resonator and filters with acoustic isolation of resonating layers, Proc. IEEE Ultrason. Symp., 1998.

19. M.A.Dubois, P.Muralt et al. «Solidly mounted BAW resonators based on aluminum nitride thin films», Proc. IEEE Ultrason. Symp., 1998.
20. G.D. Mansfeld, Tech. Phys. Letters, v.23 (10), October, 1997, p.750-752.
21. G.D.Mansfeld, «BAW composite resonator spectroscopy», Proc. of IEEE UFFC Simposium, 1994 pp.655-658, Cannes, France.
22. G.D.Mansfeld, «Measurements of acoustic wave attnuation in $La_3Ga_5SiO_{14}$ using HBAR technique», Proc.of IEEE Frequency Control Symposium, 1994, pp.35-39, Boston, USA.
23. G.D.Mansfeld, I.M.Kotelyanskii, V.S.Veretin, O.V.Zvereva, «Acoustic losses in $GaPO_4$», Proc. International Symposium on acoustoelectronics, frequency control and signal generation, 17 - 19 Sept., 1996, Moscow, pp.102-105.
24. G.D.Mansfeld, A.D.Freik, M..Moorey, «Characterization of losses in quartz resonators using electromagnetic waves reflection coefficient measurements», in Actes. Proc. EFTF-95, pp.20-23.
25. A.F.Kabychenkov, G.D.Mansfeld, V.N.Nagirniak, A.A.Rubtsov, «Peculiarities of attenuation of acoustic waves in polycrystalline YIG». Fizika Tverdogo Tela, 1991, v.33, N.3, pp. 675-680.

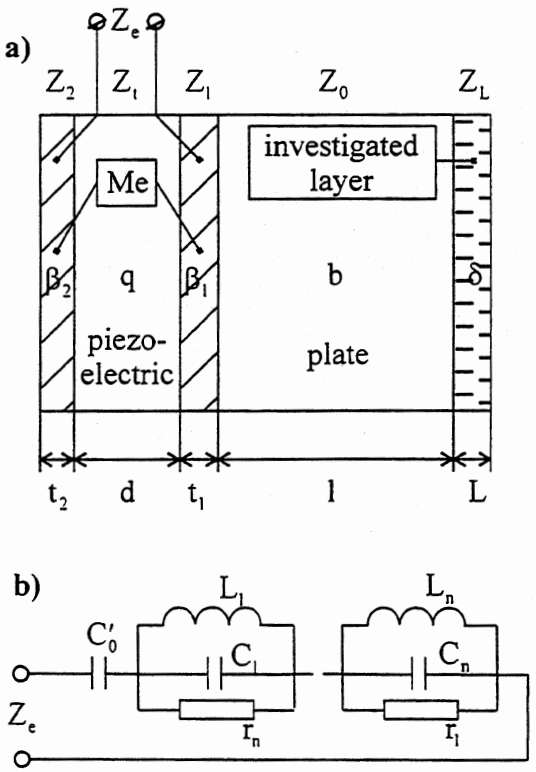


Fig.1. The schematic of the HBAR (a), the equivalent circuit (b)

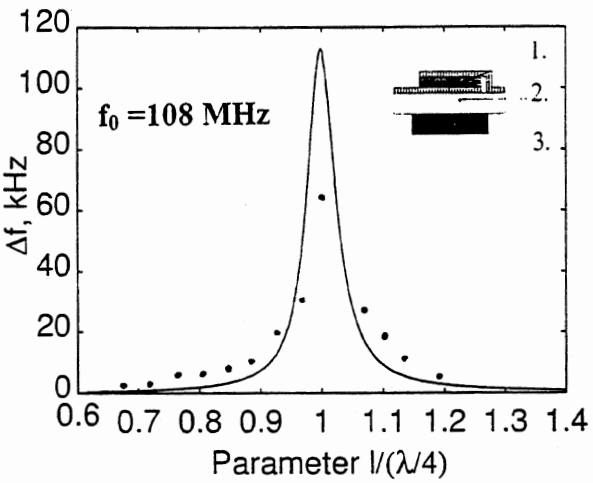


Fig.2. Resonance gas sensor sensitivity as a function of the thickness of the absorbing layer (1 - piezoelectric transducer, 2 - z-cut quartz plate, 3 - gas sensitive layer - polymer film $(NH_2)_2CHN\ HNO_3$).

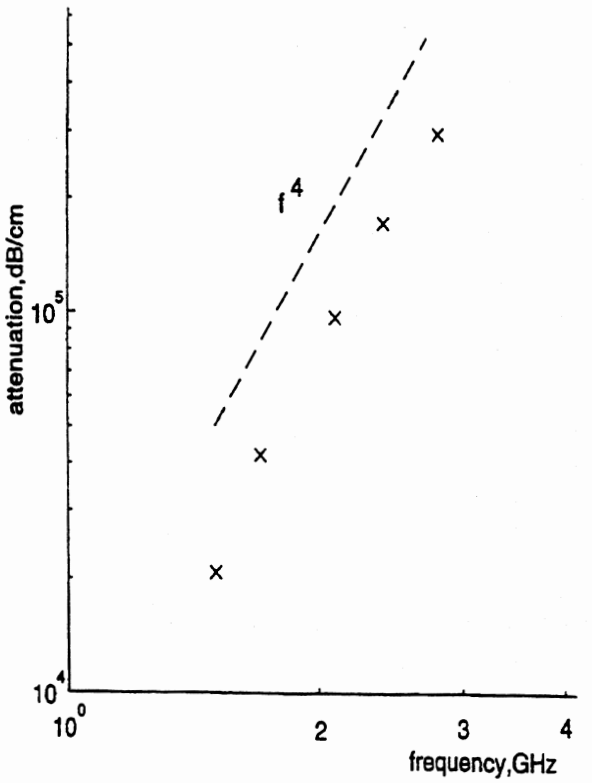


Fig.3. Acoustic losses in the Lenghmuir-Blodgett film $C_{17}H_{35}COOH$

1999 Joint Meeting EFTF - IEEE IFCS

SOME APPLICATIONS OF A STEREOGRAPHIC ANALYSIS OF ETCHING SHAPES FOR STRUCTURES MICROMACHINED IN SINGLY AND DOUBLY ROTATED QUARTZ PLATES

C. R. Tellier and T.G. Leblois
Laboratoire de Chronométrie, Electronique et Piézoélectricité

Ecole Nationale Supérieure de Mécanique et des Microtechniques
26 Chemin de l'Epitaphe – 25030 BESANCON - FRANCE

ABSTRACT

A stereographic analysis of 3D etching shapes is used to recognize rhombohedral faces limiting micromachined structures. Adjustment of the database is performed to generate for these faces extrema in the dissolution slowness. Theoretical shapes for microstructures derived from this new database are found to be in agreement with experiments.

1. INTRODUCTION

In the last two decades interest in the wet chemical etching of quartz crystals was revived. Most of recent works were devoted to the micromachining of resonant structures in Z-cut plates [1] or of tuning forks in Y-cut plates [2] which serve as sensing elements for accelerometers. However for thermal sensor applications a great number of non-standard cuts possess interesting metrological properties. When we are concerned with a lot of different cuts it is advantageous to use the kinematic and tensorial model for the anisotropic dissolution of crystals proposed by C.R. Tellier [3] which offers the possiblity to predict theoretical shapes for micromachined structures [4, 5]. The complete agreement between theoretical and experimental etching shapes depends on the adjustment of the database composed of dissolution constants which must be determined from experiments. Unfortunately for the quartz crystal etch rates depend markedly on the origin of the quartz blank. As a result up to now the proposed data base gives theoretical 3D shapes which are not fully satisfactory. This paper constitutes a tentative to provide a final adjustment of the database for structures micromachined in a concentrated $NH_4F.HF$ solution (60g per 100 ml of water).

2. ANALYSIS OF EXPERIMENTS

2.1. Characterization of the anisotropy

First of all it is essential to outline the specificity of the anisotropy of the chemical attack of the quartz crystal in a concentrated $NH_4F.HF$ solution. For this purpose we concentrate our attention on simple structures such as membranes and mesa micromachined in various singly rotated ($\varphi = 0°$, $\theta = \pm \theta_0$) and doubly rotated ($\varphi = 30°$, $\theta = \pm \theta_0$) plates designated as $X \pm \theta_0$ and $Y \pm \theta_0$ cuts respectively. To obtain these structures square and circular « masks » were patterned in evaporated Cr-Au films and plates are etched at a temperature of 65C.
The SEM images of final structures given in Figs. 1 and 2 illustrate conveniently the main features of the anisotropic chemical attack which can be summarized as follows :
(i) « Circular » membranes (Figs. 1, 2) are bordered by several planar facets or slightly curved regions which intersect to form cross-sectional dissolution profiles with obtuse or acute angle showing the complexity of the anisotropy.
(ii) Mesa produced with a starting circular or square Cr-Au mask (Figs 1, 2a) do not exhibit large convex corner undercuttings as observed for other crystals such as silicon [5].
(iii) Facets which bound circular membranes contribute also to bordering sides of mesa. This feature is typical of the quartz crystal.
(iv) The formation (Fig. 3) of an intermediate facet at the square corner of a membrane is less frequent than for other crystals.

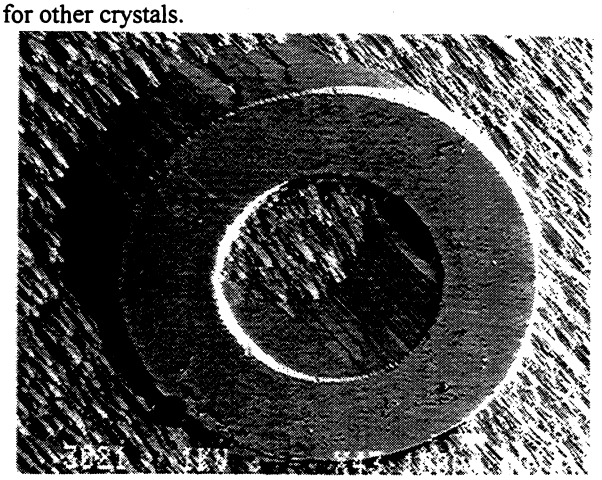


Figure 1 : Mesa-membrane structure micromachined in a Y + 30 plate

Among these features observations (ii) and (iii) are of prime importance. Effectively let us recall that kinematic models [6] for the anisotropic chemical etching predict that final shapes are governed by extrema in the dissolution slowness. In terms of dissolution criteria, a starting concave (convex) 3D structure is finally limited by facets associated with maxima (minima) in the dissolution slowness. As a result a membrane (i.e. a totally concave 3D shape) is bordered by facets associated with maxima in L. In contrast as for a mesa the top contour of the mask looks

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

convex when all minima in **L** are lower than the minor maximum we are sure that the etching results in a marked convex corner undercutting [5]. The absence for the quartz crystal of convex corner undercutting can be understood as follows : some minima in **L** are larger than some maxima or are of same order of magnitude. The existence of several minima and maxima with close magnitude together with the fact that the anisotropy ratio concerns extrema of converse nature (and not two maxima as in silicon [5]) provide a specificity to the anisotropic chemical etching of quartz which can be qualified of type 2 [7].

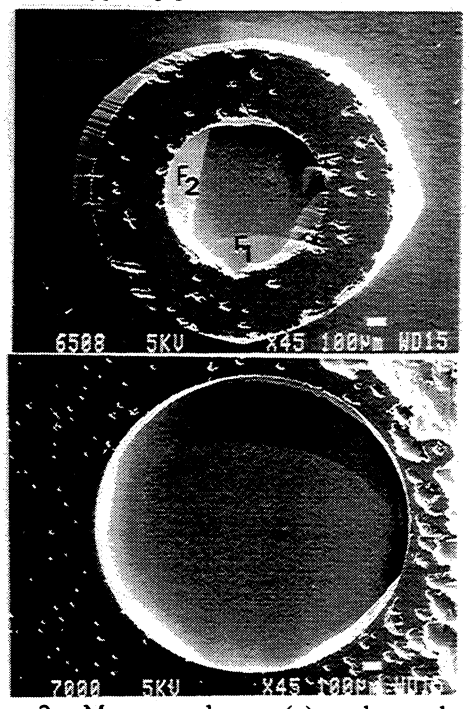


Figure 2 : Mesa-membrane (a) and membrane (b) micromachined in X+65 and X-70 plates respectively

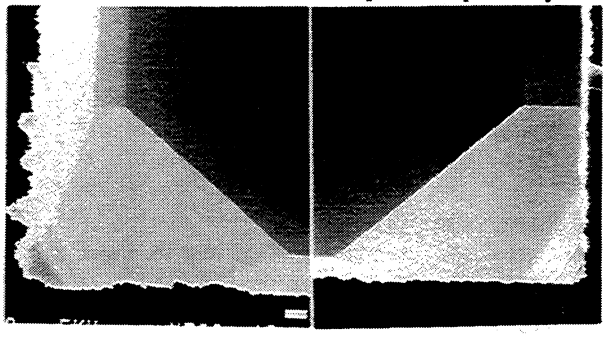


Figure 3 : Limiting facets at square corners of a X-70 membrane

2.2 The stereographic representation

Let be $\mathbf{n}$ (φ, θ) the outward unit normal to a doubly rotated plate with angles of cut (φ, θ) and $\mathbf{n_0}$ (φ_0, θ_0) the outward normal to a particular cut (φ_0, θ_0) which serves as projection plane $\boldsymbol{P}_p$ in a stereographic representation and suppose that the plate (φ_F, θ_F) corresponds to a limiting facet. If $\boldsymbol{P}_p$ is the Z-cut (φ_0=0°, θ_0 = 90°) it becomes evident that the facet (φ_F, θ_F) is represented by a point P_F (Fig. 4) on the sphere and by a point H_F in the projection plane in such a way that :

i) angles α and β correspond to polar angles and $\alpha = \varphi_F + 90°$

ii) the angle γ is just equal to $\beta/2$

iii) the intersection of the facet with the projection plane is a straight line L_F which lies perpendicular to OH_F.

Feature (iii) is still valid when the projection plane corresponds to a doubly-rotated cut.

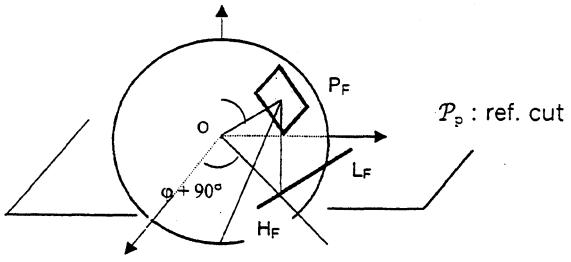


Figure 4 : The stereographic projection

Here we assume that the rhombohedral faces (r, R, S, X) of a quartz crystal (table 1) constitute the limiting facets which bound a micromachined membrane and are associated with maxima in **L**. So a stereographic projection of the rhombohedral faces on singly or doubly rotated reference cuts will give the successive intersecting lines L_F of these faces with the reference surface as viewed on SEM images (Figs 1 and 2). Inclinaisons δ" of faces, angles α" that segments OH_F perpendicular to intersecting lines L_F make with the rotated axis $x"_1$ and coordinates $(x"_1, x"_3)$ of representative points H_F were numerically evaluated for various projection planes corresponding to doubly (or to singly) rotated plates.

Faces	φ_F(°)	θ_F(°)
R	120	38,21
r	-60	38,21
S	30	24,44
X	8,94	8,048

Table 1 : The rhombohedral faces

Owing to the symmetry of the crystal class 32 representative points related to faces r, R and S exhibit some specific symmetries. Firstly for projection planes obtained by a rotation of θ_0 about the two-fold axis x_1 two faces R (or r) with angles of cut (120°, θ_F) and (240°, θ_F) have representative points symmetrical with respect to the rotated axis $x"_3$ (Fig. 5b). Secondly if we turn attention to a projection plane (Y+θ_o) it appears (Fig. 5a) that representative points H_F and H_R associated with a face r (φ_r, θ_r) and with a face R (φ_R, $\theta_R = \theta_r$) lie symmetrical with respect to the $x"_3$ axis provided the relation $\varphi_R + \varphi = 180°$ was verified.

Moreover for Y+θ_0 planes it is also possible to find a pair of faces S (Fig. 5a) whose representative points possess also this symmetry property.
As example some contour shapes for X+θ_0 and Y+θ_0 membranes composed of intersecting lines L_F are drawn in Fig. 5. For more clearness only representative points corresponding to faces which have been recognized without ambiguity on the SEM images (Figs.1,2) are indicated together with some supplementary points which illustrate the above-mentioned symmetry properties. Dotted lines correspond to faces with inward normal $\mathbf{n}_F$ which intersect the upper surface of membranes by forming sharp dissolution profiles.

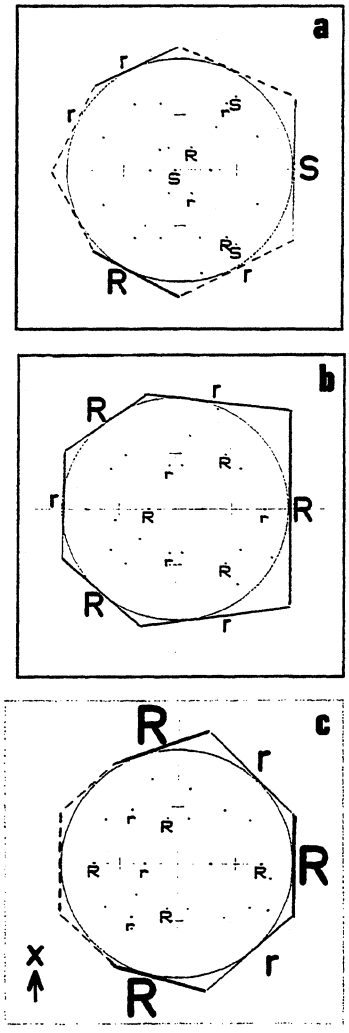


Figure 5 : Contours of membranes as given by a stereographic projection : (a), (b), (c) are for Y+30, X+65 and X-70 plates respectively

The stereographic analysis reveals that faces r and R participate to the sidewalls of membranes micromachined in X+θ_0 and Y+θ_0 plates. In addition faces S bound partly Y+θ_0 membranes. In particular the less-inclined facet ($\delta'' \approx 5°$) which is formed on a Y+30 membrane seems to be associated with a face S. Theoretical variations in the inclinaison of this face S with increasing values of θ_0 agree well with experimental variations [7]. However for several cuts some limiting facets cannot be connected with rhombohedral faces (for example facets F_3,F_4 on Figs 1 and 2). This behavior will be discussed in section 3.2.

3. APPLICATION TO MICROMACHINING

3.1. Adjustment of the database

In a first time we assume that faces r, R and S are associated with maxima in the dissolution slowness. Hence, dissolution constants in a first step were adjusted to generate maxima for these cuts (φ_F, θ_F). In a second step shapes of 3D micro-structures are compared with theoretical shapes as derived from the 3D numerical simulation "TENSOSIM" developped by C.R. Tellier [5]. Surprisingly when maxima in **L** occur for the S faces the adequation between theoretical and experimental Y+θ_0 membranes is deteriored. Consequently a second assumption is made and we suppose that S faces are in practice connected with accentuated minima in **L**. With this assumption it is easy to generate minima close to S face. Fig. 6 gives two polar plots of **L** lying the X and Y cuts where extrema associated with faces S, r and R are indicated.

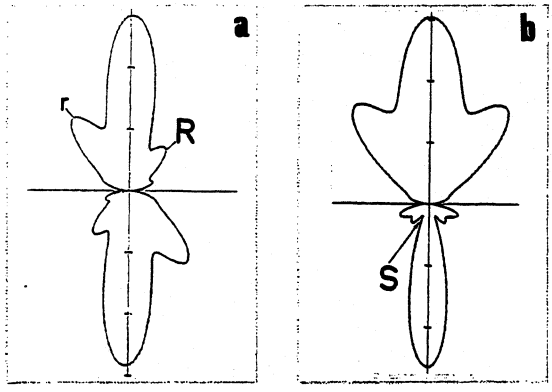


Figure 6 : Polar diagrams of **L** corresponding to X (a) and Y (b) sections

3.2. Theoretical shapes and discussion

Theoretical shapes (constant level contour diagrams) for holes micromachined in various plates are given in Fig. 7 as derived from the simulation tool TENSOSIM [5, 7] starting with the adjusted database. Turning our attention to the theoretical Y+30 membrane (Fig. 7a) we observe that as expected the presence of an accentuated minimum in the vicinity of the S face lying far from marked maxima in **L** causes effectively the formation of a sligthly curved facet (F_1) with a mean inclinaison of about 5°. Examination of the theoretical X + 65 membrane (Fig. 7b) reveals a qualitative agreement with the SEM image of Fig. 2a provided we assume a misalignment of about 20° for the rotated axis. In particular we recognize on the theoretical 3D shapes all

the facets ($F_i$and F_2) which bound the membrane as well as the curved region (CR). From a quantitative point of view we observe only small departures between theoretical and experimental inclinaisons of limiting facets. Comparison of Figs 2b and 7c indicates also a satisfactory accord between theoretical and experimental X-70 membranes if here again we accept a misalignment of about 15°. To verify the misalignment angle Ψ_A for the X-70 plate several simulations of the concave undercutting at square corners of hole are performed for various values of Ψ_A . Fig. 8 shows the theoretical corners for Ψ_A = 15°. We observe that the proposed data base generates effectively two intermediate facets at corners which are similar to facets depicted on SEM images (Fig. 3).

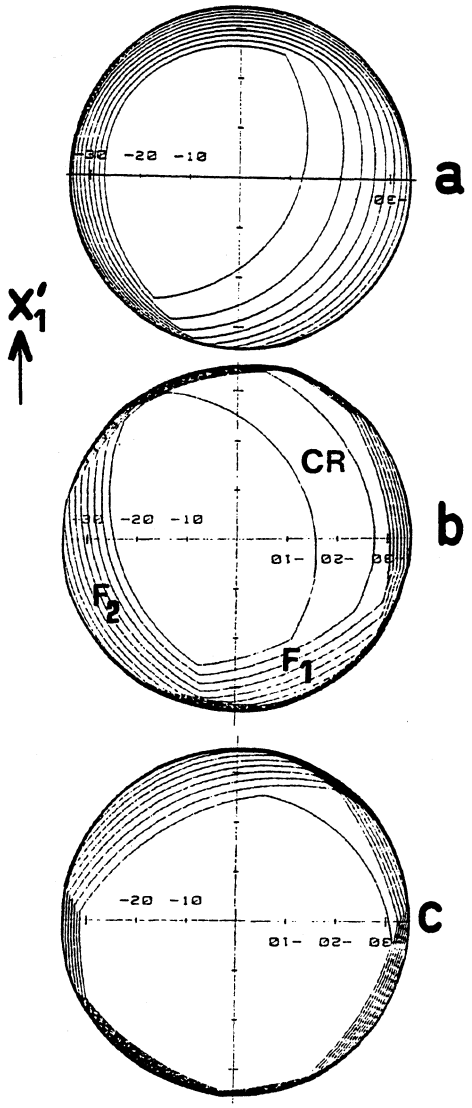


Figure 7 : Theoretical «circular » membranes in (a) Y+30, (b) X+65 and (c) X-70 plates.

CONCLUSION

In conclusion it is interesting to adopt a stereographic analysis for the identification of some limiting facets. The subsequent adjustment of the data base reveals that faces R and r are associated with maxima in the $(\varphi = 90°, \theta = 0°)$ polar diagram of the dissolution slowness whereas a minimum occurs in the vicinity of the face S. But other extrema are also operative. Finally with this new database we obtain a better agreement between theoretical and experimental 3D etching shapes even if theoretical shapes appear to be rotated with respect to stereographic projections.

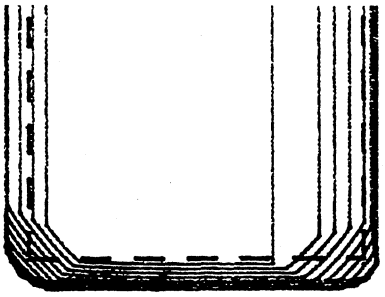


Figure 8 : Theoretical square corners for a X-70 membrane

Acknowledgements
The authors gratefully acknowledge Dr. D.Benmessaouda and Dr. T. Messaoudi for providing the etched structures

REFERENCES

[1] J. Soderkvist, « Micromachined gyroscopes » Sensors and Actuators A, vol. 43, pp 65-71, 1994.
[2] T. Ueda, F. Kohsaka, D. Yamasaki and T. Lino, « Quartz crystal micromechanical devices » in Proc. 3rd Int. Conf. Solid-State Sensors and Actuators, Philadelphia, PA, USA, 11-14 June 1985.
[3] C. R. Tellier and J.L. Vaterkowski, « A tensorial representation of the dissolution slowness. Application to etched singly rotated quartz plates », J. Mater Sci. vol. 24, pp 1077-1088, 1989
[4] C.R. Tellier and S. Durand, « Microfabricated silicon structures experimental and theoretical investigations of 3D etched shapes » in Proc. 9th EFTF, Besançon, France, 1995, pp 426-429
[5] C.R. Tellier and S. Durand « Micromachining of (hhl) silicon structures : experiments and 3D simulation of etched shapes », Sensors and Acuators A, vol. 60, pp 168-175, 1997.
[6] C.R. Tellier, J.Y. Amaudrut and A. Brahim-Bounab, « The dissolution slowness surfaces of cubic crystals : Part II : Applications to class 23 and to combined etching and lithography techniques », J. Mater. Sci. vol. 26, 1991, pp 5595-5607
[7] C.R. Tellier, T. Messaoudi and T. G. Leblois, « Contribution of a 3D numerical simulation to the design of micromachined quartz microstructures in Proc. of the 10th EFTF, Brighton, U.K., March 1996, pp 364-369

COMPLETE X-RAY ORIENTATION DETERMINATION OF QUARTZ BARS USING A SMALL SCANNING RANGE

H. Berger, H. Bradaczek, and G. Hildebrandt
EFG International Berlin, Research Center
Dueppelstr. 13, D-14163 Berlin, Germany

ABSTRACT

In order to determine the true *AT* cutting angle and the *XX'* miscutting at the *X* face perpendicular to the rotation axis of quartz bars, three X-ray diffraction peaks have to be considered. Reflection combinations have been found which can be measured with sufficiently small errors in the relevant angles using two beams incident in two slightly different directions and scanning over a limited angular range in the order of the reflection widths. So, an automated measuring and bar-adjusting process can be established. An example of a suited reflection combination is given, and a corresponding arrangement based on it is proposed.

1. INTRODUCTION

For the cutting process of *AT* blanks, the preceding X-ray orientation of the lumbered bars is a well-established procedure. It is usually performed by measuring one X-ray reflection by the so-called Θ scan (lattice-plane normal as well as incident and reflected beam coplanar). In this way, only one component of the cutting angle is measured. Orientation deviations of the bar surface (up to 30 arcmin.) are not revealed and lead to unknown cutting errors. Employing reflections at the (101) lattice plane, as are used in the Automated Bar-Adjustment Machine (produced by EFG International Berlin, Fig. 1), this so-called *XX'* miscutting error is moderate (<8 arcsec). Using other reflections, the miscutting error may be appreciably higher depending on the reflection geometry.

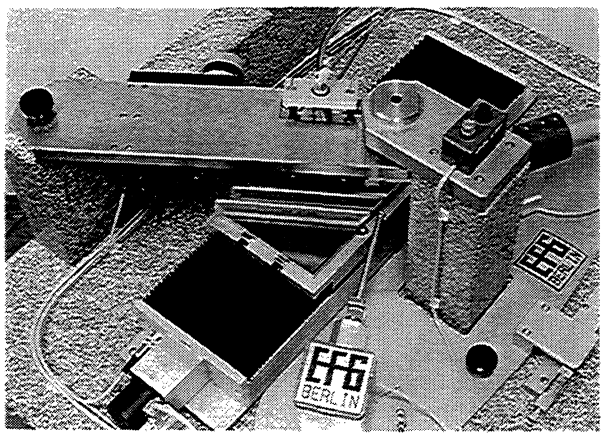


Fig. 1.
View of the EFG Automated Bar-Adjustment Machine (X-ray control by Θ scan).

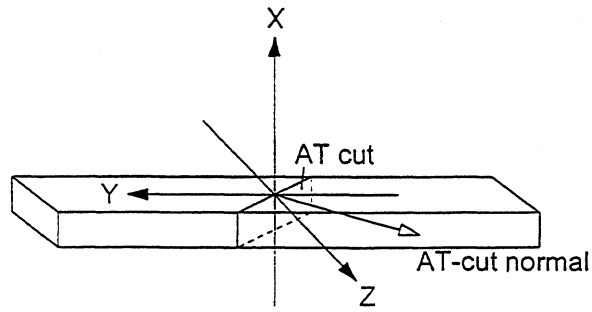


Fig. 2.
Lumbered quartz bar with coordinates and *AT* cut.

The bar adjustment using one X-ray reflection is feasible only at the surfaces parallel to the *X* axis of the bar (Fig. 2). So, it cannot be used for bars which are to be placed in two-dimensional arrays on the cutting support. For this purpose, a method has to be found which enables to measure at the *X* surface.

It is therefore necessary to develop a method of complete orientation determination of quartz bars applicable to the *X* surface. The complete orientation determination (resulting in two angular components) requires the measurement of at least three X-ray diffraction peaks. The method should need the rotation about only one axis, namely the *X* axis. Such a method is the Ω-scan method, developed several years ago for the complete orientation determination of *AT*-cut [1, 2] as well as *SC*-cut quartz blanks [2, 3]. A special variant with rotation about the *Y* axis of the bar has been applied for the X-ray controlled *SC* cutting [4]. With the Ω-scan method, an X-ray beam is two times reflected at two lattice planes (four peaks measured) using nearly a full turn of the specimen. However, for the *AT* bar adjustment the X-ray measurement should be performed by scanning over angular ranges as small as possible in order to change the angular positions of the pre-positioned bars only slightly. Therefore, in the ideal case the three reflections should occur simultaneously.

The simultaneous registration of three reflections means fulfilling the conditions of a so-called "four-beam case" of X-ray diffraction (one incident, three diffracted beams involved). Whereas there exists a large number of three-beam cases (two simultaneous reflections), true four-beam diffraction can be realised only under special conditions. The utilisation of the three-beam case for oriented cutting has been suggested earlier [3]. However, the conditions for this case must be fulfilled exactly, what can be implemented only by adjustment about at least two axes.

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

Therefore, a simple method has to be found which makes use of a pseudo four-beam case. The possibility of using two incident beams with an angular difference of a few degrees has to be taken into account. A general (skew) reflection geometry like that used in the Ω-scan method is to be applied.

In the following, the requirements for a complete X-ray orientation method for *AT* bars and possibilities of its realisation will be discussed. The draft of a bar-adjusting arrangement based on the proposed measuring method will be shortly described.

2. REQUIREMENTS FOR X-RAY CONTROLLED BAR ADJUSTMENT

For this X-ray method, three reflections have to be found which have the following properties and can be measured under conditions as follows:

i) The reflections should have sufficiently high and comparable intensities.
ii) For medium bar orientation, the three reflections should appear simultaneously.
iii) The condition ii) should be fulfilled for two X-ray beams having an angular difference small enough that it can be realised by one and the same X-ray tube.
iv) Each peak must be sharp and well separated from the second one arising at the same lattice plane.
v) It may be advantageous to observe a special geometric relation between the *Y* direction of the bar and the plane containing the incident beams (see below).
vi) The resulting errors of the orientation determination and of the adjustment must be sufficiently small. Therefore, the lattice planes corresponding to the reflections to be measured have to be related in a certain way one to each other as well as to the rotation axis.
vii) All the conditions mentioned above should be fulfilled over the whole orientation range of the lumbered bars.

It is clear that the requirements are partially contradictory. So, a high sensitivity for orientation changes will lead to small errors and, simultaneously, to larger shifts of the angular positions of the peaks, i.e. a larger scanning range has to be considered. Therefore, there is a necessity to make compromises.

3. PRINCIPLES OF THE X-RAY CONTROLLED BAR ADJUSTMENT

Let us suppose the principal geometry of the lattice (angular coordinates of the lattice planes with respect to the lattice) as well as of the diffraction (Bragg angles, angle between incident X-ray beam and rotation axis) are to be known. Then from the measured angular distances of the reflections at three different (non-coplanar) lattice planes, produced by rotation about an axis, the angles between lattice planes and this axis, and the angular coordinates of the rotation axis with respect to the lattice can be calculated.

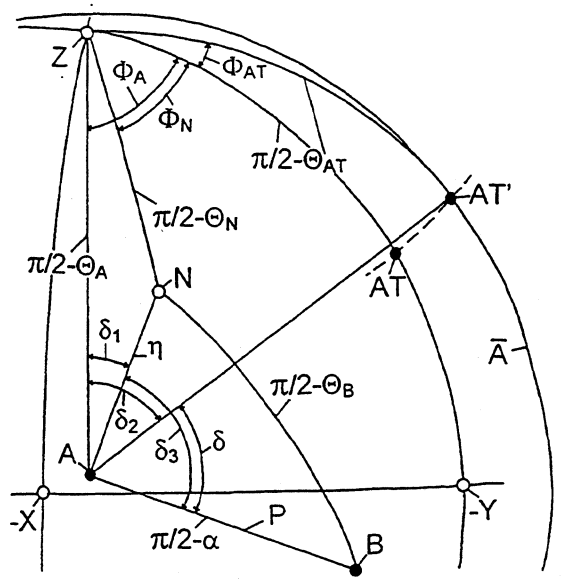


Fig. 3.
Projection of a section from a sphere surface showing the intersection lines and points, resp., of the relevant planes and directions as well as the principal angles for the X-ray measurement at the "*X* face" (projection plane) of a misaligned bar (only one lattice plane shown).
A, $\bar{A}$: rotation axis and plane perpendicular to it, resp.; *X*,*Y*,*Z*: coordinate axes of the quartz lattice; *AT*: normal of the *AT* cut; *AT'*: normal of the *AT* cut with *XX'* miscut (by the angle Φ_{AT}); *N*: lattice-plane normal; *B*: incident beam direction; *P*: plane containing the rotation axis and the incident beam direction (explanation of the angles see text).

After determining these values, the necessary setting angle for the bar adjustment and the *XX'* miscutting angle to be expected can be derived. The principal reflection geometry and the geometry of the bar to be adjusted are shown in Fig. 3. From the wanted cutting angle Θ_{AT} the adjusting angle δ (related to the plane *P*) and the *XX'* miscutting angle Φ_{AT} are obtained according to

$$\delta = \delta_1 - \delta_2 + \delta_3 \qquad (1)$$

with

$$\delta_1 = \sin^{-1}[\cos\Theta_N \sin(\Phi_A - \Phi_N) / \sin\eta], \qquad (1a)$$

$$\delta_2 = \cos^{-1}(\sin\Theta_{AT} / \cos\Theta_A), \qquad (1b)$$

$$\delta_3 = \cos^{-1}[(\sin\Theta_B - \sin\alpha\cos\eta) / (\cos\alpha\sin\eta)], \qquad (1c)$$

and

$$\Phi_{AT} = \Phi_A - \cos^{-1}(\tan\Theta_{AT}\tan\Theta_A) \qquad (2)$$

(Θ_A , Φ_A : angular coordinates of the rotation axis; Θ_N, Φ_N : angular coordinates of the lattice-plane normal; η: angle between lattice-plane normal and rotation axis; Θ_B: Bragg angle; α: incidence angle of the X-ray beam).

4. CHOICE OF OPTIMAL MEASURING CONDITIONS

There exists a very large number of combinations of three reflections in quartz. The choice was limited to reflections of higher intensity (structure factor >10). Regarding the geometrically possible reflections for CuKα radiation, some 10^5 of non-equivalent combinations have to be checked. This was made by means of a particular program which selected only combinations for which the angular difference of the incidence angles did not exceed a certain value, and the evaluation as described above led to small errors for the adjusting angle. The selected reflection combinations were further checked with respect to the remaining criteria (cf. chapter 2.). In order to determine the widths of the reflection profiles, these profiles have been calculated supposing realistic beam divergence [5].

A number of reflection combinations resulted which were worth analysing in more detail. Within these, the smallest errors to be expected for the adjusting angles are in the same order of magnitude as the measuring errors of the peak positions. The differences in the incidence angles of the X-ray beams are at least a few degrees. In all cases, the necessary scan range, due to the reflection widths and to the orientation range to be considered, must be in the order of 10°. For the reflection combination 043/143/31$\bar{2}$ representing the best compromise found up to now concerning all requirements, the most important data are given in Table 1.

Table 1.
Parameters for the X-ray measurement of the reflection combination 143/043/31$\bar{2}$, CuKα radiation.
Rotation axis (for medium bar orientation): X axis [X surface: (110)];
peak distance and peak position shift for bar misalignment of ±0.3° about Y and Z;
halfwidths for Kα_1 (* for Kα doublet);
errors relative to the common peak position error ≡1

reflection	1 4 3	0 4 3	3 1 $\bar{2}$
Bragg angle	68.9°	57.2°	45.4°
incidence angle	71.1°	79.2°	71.1°
halfwidth	0.63°	1.7°	2.1°*
maximum peak distance		2.7°	
maximum peak position shift		7.4°	
errors			
- adjusting angle		0.65	
- XX' miscutting		0.52	

The assumed uncertainty of the bar orientation of 0.3° for both angular coordinates is guaranteed by most producers.

5. DRAFT OF A MEASURING AND ADJUSTING ARRANGEMENT

Striking features of an arrangement realising the basic concept are: the skew geometry of the measuring system, three separate detectors to take the simultaneous reflections, and two incident X-ray beams with a certain angular difference. Considering only one X-ray tube, two beams can be principally implemented by beam deflexion using crystal monochromators or by utilising the angular convergence or divergence, resp., of emitted beams. Applying a long line focus, two converging beams can be used, connected with an appreciable loss of intensity. Besides that, the available focus length of at most 12 mm would lead to very small tube-bar distances. On the other hand, using two diverging beams emitted from a point focus and transmitted by two diaphragms, the impact points are correspondingly separated on the bar surface. In this case, the plane containing the beams should be nearly parallel to the long axis (Y) of the bar. In spite of this complication, this seems to be the best variant and will be considered in the following.

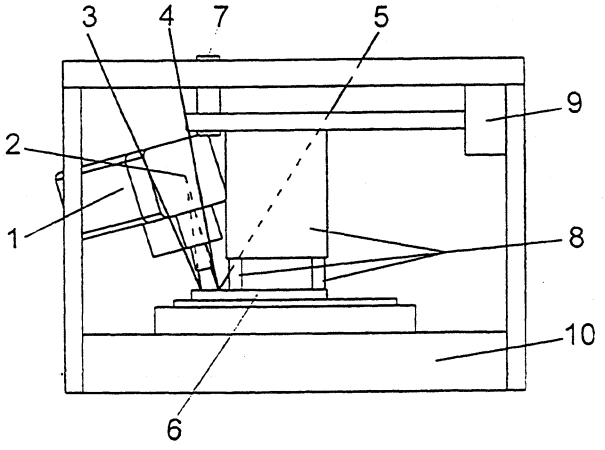


Fig. 4.
Sketch of an X-ray controlled bar-adjustment machine as proposed in the text (side view).
1: X-ray tube; *2*: tube focus (two diverging incident beams indicated by dashed lines); *3 - 5*: reflected beams; *6*: bar; *7*: rotation axis; *8*: bar holder enabling the bar rotation by drive *9*; *10*: translation support (*1*, *3 - 5*, and *10* inclined to the plane of the figure; see text).

An arrangement based on the reflection combination 043/143/31$\bar{2}$ and the parameters as given in Table 1 is sketched in Fig. 4. The Y direction of the bar and the bar holder are in one plane, the plane containing the incident beams is rotated of about 20° around the rotation axis, and the planes through rotation and detector axes are inclined by various amounts up to about 70°.

Using a distance tube focus - bar of 130 mm, the measuring points on the bar surface will be separated by roughly 20 mm. Diaphragm diameters of 0.3 mm give still total peak profile widths up to about 6°. Considering a sufficient tolerance range, a total range of at least 15° should be scanned during the measurement.

Only this rotation in one direction needs to be performed. After measuring and evaluating, the bar is brought into the calculated final position and fixed there. The initial angular positions of the bars have, therefore, to be changed by the necessary scan range taking care that each bar can be free adjusted (Fig. 5). In this way, an automated measuring and adjusting process can be established.

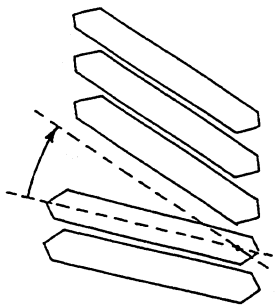


Fig. 5.
Scheme of the successive bar adjustment of a row. At the bottom: pre-positioned, at the top: finally positioned bars

Assuming the same peak-position error for all reflections, the deciding error of the adjusting angle is only somewhat smaller than the measuring error, which, therefore, should be as low as a few arcsec. An estimation and the comparison with existing measuring arrangements shows that this should be realistic if the time for measurement and adjustment of one bar is chosen to be in the order of one minute.

6. CONCLUSIONS

The bar adjustment controlled by an X-ray measurement at the X surface, based on the method discussed above, involves some problems. One of them may result from the two separated measuring points because orientation inhomogeneities, especially connected with the seed plate, cannot be excluded. The arrangement as well as the procedure are somewhat more complicated (three separate detecting systems; scanning range >15°, longer cycle time) than for the usual Θ scan. The continued search for further reflection combinations including also weaker reflections or the use of other wavelengths could eventually lead to more effective conditions. Also the calibration of the system will involve more effort. It will be necessary to correct the measured peak positions carefully with respect to shifts due to reflection profile asymmetry [6].

However, the described method fulfils the requirements as postulated above. In contrast to the Θ-scan method, the true AT-cutting angle is obtained also for misaligned bars, and the XX' miscutting angle is determined additionally with high precision. Bars having too large XX' miscutting angles could be marked or rejected before fixing.

It should also be possible to apply the principal procedure of complete orientation determination by taking three reflections at surfaces parallel to the X axis (performing an approximate "Θ scan"). The conditions for such a method have still to be found.

REFERENCES

[1] B. Nestler, H.-J. Kuhr, G. Hildebrandt, and H. Bradaczek, "Novel use of a commercial goniometer for sorting round quartz blanks", Meas. Sci. Technol., vol. 2, 1991, pp. 528-531.

[2] H. Bradaczek and G. Hildebrandt, "Applications of X-rays in the quartz-oscillator industry", Rigaku J., vol. 14, 1997, pp. 12 - 20.

[3] Berger, H. Bradaczek, H.-A. Bradaczek, and G. Hildebrandt, "Application of the Ω-scan to the sorting of doubly rotated quartz blanks", Proceedings of the International Frequency Control Symposium, 1996, pp.39-43.

[4] H. Berger, H. Bradaczek, and G. Hildebrandt, "X-ray control and adjusting arrangement for cutting SC-cut blanks using only two rotation axes", Proceedings of the 20th Piezoelectric Devices Conference, 1998, pp. 5/1-5/4.

[5] H. Berger, "Improvements of X-ray orientation determination methods applicable in quartz resonator industry", Proceedings of the 12th European Frequency and Time Forum, 1998, pp. 339-344.

[6] H. Berger, H. Bradaczek, and G. Hildebrandt, "Improvements of the Ω-scan method by eliminating systematic errors", Proceedings of the 19th Piezoelectric Devices Conference, 1997, pp. 15/1-15/15.

Theoretical Analysis of the Second Stop-band of Rayleigh Waves Propagation on Periodically Corrugated Anisotropic Substrates

Bernard Dulmet, Hervé Watchueng and Jean-Bernard Briot

Laboratoire de Chronométrie Electronique et Piézolectricité
Ecole Nationale Supérieure de Mécanique et des Microtechniques, F 25000 Besançon

Bernard.Dulmet@ens2m.fr Tel: 33 (0)3 81 40 28 25 Fax: 33 (0)3 81 88 57 14

Abstract

This paper presents a modeling of the dispersion curves for SAW propagation on periodically corrugated substrates in the second stop–band, which involves a strong coupling with BAW propagating both ways, towards– and from the depth of substrate. The proposed modeling uses series of Bloch functions to obey periodic boundary conditions which are integrated over one period of surface profile. The case of finite slope for trapezoidal profile is considered and the effect of anisotropy is taken into account. All equations are written in lagrangian formalism, which will facilitate a farther simulation of static thermal behavior of homogeneous crystalline structures with etched gratings. The paper presents results for energy reflection and conversion coefficients of a SAW–BAW transducer with a finite number of gratings on $Y+\theta$ cuts of quartz.

Introduction

We study the mutual interactions between surface and bulk acoustic waves when the surface of substrate exhibits periodic corrugations with a period Λ very close to the the wavelength λ of SAW. A standard approach in case of corrugated substrates consists of expanding the solutions in terms of Bloch harmonic functions to derive the dispersion curve $f(\delta,\omega)$ where δ denotes the wavenumber component governing the behavior of Bloch sum. Ref.[1] outlines a theoretical analysis and gives closed-form results for a reciprocal coupler when surface waves are of the Rayleigh type and propagate onto an *isotropic* substrate. Boundary conditions are then "projected" onto the average plane of profile by means of Brekhovskikh expansion and combined with Fourier expansion of the surface profile, thereby allowing to take into account the effect of various profiles. In the present paper, we make use of the method previously presented at Ref.[2] instead of Brekhovskikh expansion, *i.e.* we multiply the boundary conditions by orthogonal harmonic functions and we integrate them over one period of the surface profile. Couplings between harmonics are then governed by integrals which can be expanded in terms of powers of the small parameter ε. Mutual couplings are of the second order in terms of ε, whereas they are of the first order for Bragg's reflection. While Ref.[2] focused on Rayleigh waves on isotropic substrates and STW in $Z(YZ)$ directions of trigonal crystals, the present paper studies the case of BAW couplings of Rayleigh waves propagating in $Z(YZ)$ direction of quartz crystal. Due to anisotropy, analytical calculations with help of potential methods are not applicable, and the calculations are carried out numerically.

I. SAW in lagrangian formulation

It is known that for a piezoelectric medium, the incremental balance equations of motion in Lagrange configuration writes as follows :

$$\begin{aligned}\sum_{L=1}^{3}\frac{\partial \tilde{K}_{L\alpha}}{\partial X_L}&=\rho_0\frac{d^2u_\alpha}{dt^2}\qquad \alpha=1\ldots3\\ \sum_{L=1}^{3}\frac{\partial \tilde{D}_L}{\partial X_L}&=0\end{aligned}\tag{1}$$

where Piola-Kirchhoff stress tensor K and material dynamic displacement D respectively indicates the effect of a force or an electric charge on a small element of matter, the surface of which was equal to S_0 in some equilibrium state, prior the force and/or the charge was applied on it :

$$\begin{aligned}n_L^0K_{L\alpha}dS_0&=F_\alpha\qquad \alpha=1\ldots3\\ n_L^0D_LdS_0&=\sigma\end{aligned}\tag{2}$$

In (1), the tilde superscript indicates that the equations pertain to the incremental quantities associated to an infinitesimal vibration superimposed on a static bias. It has been shown that the Piola–Kirchhoff and material displacement can be linearly related to mechanical displacement and electric potential gradients through so–called effective coefficients :

$$\begin{aligned}\tilde{K}_{L\gamma}&=G_{L\gamma M\varepsilon}u_{\varepsilon,M}+R_{K.\gamma L}\varphi_{,K}\\ \tilde{D}_K&=R_{K.\gamma L}u_{\gamma,L}-N_{KL}\varphi_{,L}\end{aligned}\tag{3}$$

Effective constants can be related in a rather cumbersome but quite well–defined way to more fundamental constants which are the coefficients of expansion of a rotationnally

0-7803-5400-1/99/ $ 10.00

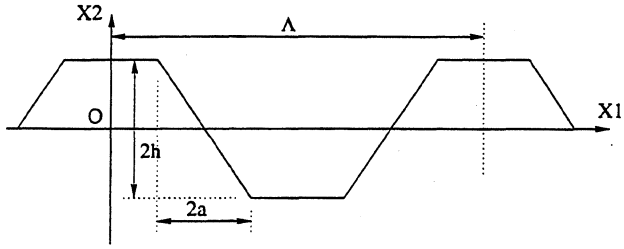


Figure 1: Symmetrical trapezoidal profile

invariant thermodynamic potential with respect to strain and material electric vector. Since the fundamental constants do not immediately appear in the equations of motion, using effective constants is easier provided the studied problem does not involve any inhomogeneous biasing state, such as thermal gradients, for instance. The problem of high-order temperature sensitivity of Rayleigh wave propagation was previously studied with help of effective coefficients at Refs.[3-4].

II. Floquet eigenmodes in periodic gratings

Let us consider a surface with symmetrical trapezoidal profile defined at Fig.1. The heigth of profile $2h$ is supposed to be very small with respect to the period of corrugation Λ. The ratio a/h determines the slope of trapeze edges. Then, assuming $\vec{X}_2$ is the direction of the average normal to substrate (outwardly orientated), and X_1 is the direction of propagation (in suitable rotated axes, of course), the profile is described by the function :

$$y = \xi(X_2) = h\bar{\xi}(X_1) = \varepsilon/Q\bar{\xi}(X_1) \tag{4}$$

(where $-1 \leq \bar{\xi} \leq 1$, and the corrugation wavenumber Q is defined by $Q = 2\pi/\Lambda$), in such a way that the dimensionless parameter $\varepsilon \ll 1$ is small in any case. Since we study a twodimensional problem in the sagittal plane $(X1, X2)$, the mechanical boundary conditions reduce to :

$$n_2 K_{2\alpha} + n_1 K_{1\alpha} = 0 \qquad \forall (X_1, X_2) \in C_\Lambda \tag{5}$$

where C_Λ is the trace of surface profile in the sagittal plane. By simple substitution, we obtain :

$$\begin{aligned} &n_2 \left(G_{2\alpha M\varepsilon} u_{\varepsilon,M} + R_{K.\alpha 2}\varphi_{,K}\right) \\ &+n_1 \left(G_{1\alpha M\varepsilon} u_{\varepsilon,M} + R_{K.\alpha 1}\varphi_{,K}\right) = 0 \end{aligned} \tag{6}$$

Due to the periodic nature of corrugations, Floquet's theorem allows to seek the solutions as sums of Bloch functions :

$$\begin{aligned} u_\alpha &= \sum_{k\in\mathcal{Z}}\sum_{n=1}^{4} C_{k,n} u_\alpha^{k,n} e^{j(q_k X_1 + \beta_k^n X_2 - \omega t)} \\ \varphi &= \sum_{k\in\mathcal{Z}}\sum_{n=1}^{4} C_{k,n} \varphi^{k,n} e^{j(q_k X_1 + \beta_k^n X_2 - \omega t)} \end{aligned} \tag{7}$$

where k is the harmonic rank, α indicates the component and n determines which solution is taken in the determinantal equation obtained when substituting above–mentionned form into the partial derivative equations (1), separately considered for each harmonic. $u_\alpha^{k,n}$ represents the normalized amplitudes for this k-th harmonic and the n-th solution β_k^n of the determinant of PDEs. The small wavenumber component δ governs the whole behavior of Bloch series with respect to ω. Then, associated mechanical boundary conditions can be written as follows :

$$\begin{aligned} 0 = \textstyle\sum_{k\in\mathcal{Z}} \sum_{n=1}^{4} C_{k,n} \Big[&n_2(X_1, \xi(X_1)) K_{2\alpha}^{k,n}(\delta,\omega) \\ &+ n_1(X_1, \xi(X_1)) K_{1\alpha}^{k,n}(\delta,\omega) \Big] e^{j(q_k X_1 + \beta_k^n \xi(X_1)) - j\omega t} \end{aligned} \tag{8}$$

where $K(\alpha, k, n)$ stress terms are defined by :

$$\begin{aligned} K_{2\alpha}^{k,n}(\delta,\omega) &= jq_k \left[\left(G_{2\alpha 1\varepsilon} + G_{2\alpha 2\varepsilon} r_k^n\right) u_\varepsilon^{k,n} \right. \\ &\qquad \left. + \left(R_{1.\alpha 2} + R_{2.\alpha 2} r_k^n\right) \varphi_k^n \right] \\ K_{1\alpha}^{k,n}(\delta,\omega) &= jq_k \left[\left(G_{1\alpha 1\varepsilon} + G_{1\alpha 2\varepsilon} r_k^n\right) u_\varepsilon^{k,n} \right. \\ &\qquad \left. + \left(R_{1.\alpha 1} + R_{2.\alpha 1} r_k^n\right) \varphi_k^n \right] \end{aligned} \tag{9}$$

At the moment, our computation program for SAW propagation on periodic gratings is stable only for purely mechanical problems, so the **R** matrix of effective piezoelectric constants is currently ignored and electric boundary conditions are not yet considered. To solve the boundary problem on periodic uneven surfaces, we arbitrarily multiply the boundary conditions by $exp(-jq_l x)$ with $q_l = \delta + lQ$, l being some arbitrary integer, and then we integrate over one period of profile (more detailed justifications were given in [5]), thereby obtaining the following algebraic system :

$$\begin{aligned} 0 &= \sum_{k\in\mathcal{Z}}\sum_{n=1}^{4} C_{k,n} \left[K_{2\alpha}^{k,n}(\delta,\omega) J_2(m,\beta_k^n) \right. \\ &\qquad \left. + K_{1\alpha}^{k,n}(\delta,\omega) J_1(m,\beta_k^n) \right] \\ J_2(m,\beta_k^n) &= \int_{C_\Lambda} n_2(s) e^{j(mQX_1(s) + \beta_k^n \xi(s))} ds \\ J_1(m,\beta_k^n) &= \int_{C_\lambda} n_1(s) e^{j(mQX_1(s) + \beta_k^n \xi(s))} ds \end{aligned} \tag{10}$$

where $m = k - l$, and C_Λ is the open integration path for one period onto the free surface profile. Remarking that $\exp[j(\beta_k^n \xi(s))]$ can easily be expanded in power series, the integrals appearing in (10) can alternatively be approximated by the following expansions :

$$\begin{aligned} J_2(m,\beta) &\approx \Lambda\delta_{m,0} + j\beta h I_2^1 - \tfrac{\beta^2 h^2}{2} I_2^2 + \ldots \\ J_1(m,\beta) &\approx mQh \left[jI_2^1 - \tfrac{\beta h}{2} I_2^2 + \ldots \right] \end{aligned} \tag{11}$$

with :

$$I_2^1 = \int_{C_\Lambda} \bar{\xi} e^{jmQX_1} dX_1$$
$$I_2^2 = \int_{C_\Lambda} (\bar{\xi})^2 e^{jmQX_1} dX_1 \qquad (12)$$

assuming the notation $\xi(X_1) = h\bar{\xi}(X_1)$. Such expansion is useful for profile functions such that integrating $\exp(j\xi(s))$ is difficult. In case of profiles consisting of broken lines (like trapeze profile), all integrals can be carried out analytically. This approach can be performed in case of finite value of the profile slope, at the difference of Brekhovskikh expansion which requires asymptotically small values of the slope. The generic system of boundary conditions (10a) is readily turned into matrix form $M_{l,\alpha|k,n} C_{k,n} = 0$, where column indexes are obtained by combination of harmonic k and number of EDP solution n, while row indexes come from the combination of stress component α and index $l = k - m$ associated to the projection of boundary conditions onto the basis of harmonic functions. β_k^n values are obtained separately for each harmonic k from equation of dynamics, which yields a polynomial characteristic equation in terms of powers of q_k. In this way, we are able to compute above–mentionned M matrix with any number of harmonics and find the roots of its determinant by means of Newton's method, thereby finding all ratio between $C_{k,n}$ coefficients. Following this approach, we obtained results of Fig. 2, which presents a plot of the real and imaginary parts of small wavenumber δ in the case of trapezoidal profiles, in the second stop band band, *ie* in the vicinity of $\omega_B = QV_R$, where V_R is Rayleigh waves celerity.

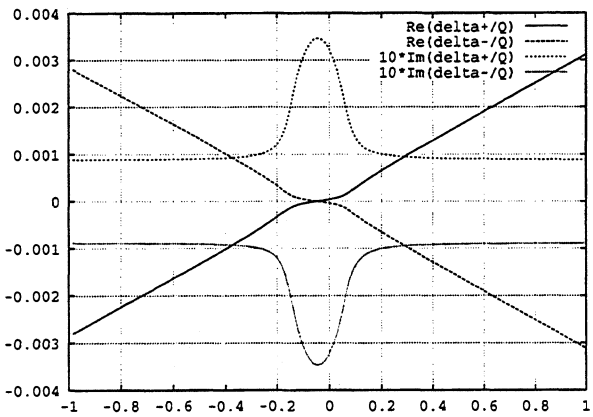


Figure 2: $\delta = f(\delta\omega/(310^{-3}\omega_0)$

Results are obtained with $\varepsilon = 0.05$ and ω_B=155 MHz for an AT32.45–cut. In that case, two branches are obtained, with opposite values of $\delta = f(\omega)$, as can be observed on the figure. The curves were obtained with $a/\Lambda = 0.125$ and $k \in [-4, 4]$. The number of harmonics for convergency of results increases with the slope of trapeze. While keeping the same value of ε dimensionless parameter, a good approximation was obtained with $k \in [-2, 2]$ for sawtooth profile, whereas $k \in [-8, 8]$ is necessary to attain convergency in the case of steep profiles with $a/\Lambda = 0.05$. In our notations, δ^+ indicates the dispersion branch with $Im(\delta) > 0$, which has physical meaning for semi–infinite gratings extending toward $X_1 > 0$, and δ^- indicates the opposite case ($Im(\delta) < 0$), meaningful for semi-infinite gratings extending toward $X_1 < 0$ with incident SAW wave propagating towards $X_1 < 0$. The relation $\delta^- = -\delta^+$ which is observed between both branches in that special case (propagation along X_1 axis of an $Y + \theta$ quartz cut) seems not to hold in case of doubly-rotated orientation, due to SAW–BAW diffraction, making it difficult to obtain both branches. So, at the moment, we stick to presenting numerical results only for sagittal plane containing the X_1 axis of singly rotated cuts, since one branch is not sufficient for the treatment of a transducer with finite length.

III. Coupler with finite length

Tiersten and Mc Carthy [6] have shown that for thermo–electro–elastic analysis carried out in Lagrange configuration, a pertinent expression of the first priciple of thermodynamics is :

$$\frac{d}{dt}\int_{V_0}\left[\rho_0\left(\frac{|\mathbf{v}|^2}{2}+\mathcal{E}\right)+\frac{\varepsilon_0}{2}J|\mathbf{E}|^2\right]dV_0 = \int_{S_0} n_L^0\left[K_{L\alpha}v_\alpha - \mathcal{Q}_L - \varphi\dot{\mathcal{D}}_L\right]dS_0 \qquad (13)$$

where $\mathcal{E}$ is the stored internal energy per mass unit and $\mathcal{Q}$ is the heat flux in Lagrange configuration, which can be set to zero for an adiabatic vibration. Under such circumstances, Poynting's theorem takes the following form :

$$\frac{dW}{dt} = \frac{d}{dt}\int_{V_0}\left[\rho_0\left(E_c+\rho\mathcal{E}\right)+W_e\right]dV_0 + \int_{S_0} n_L^0 \mathcal{P}_L dS_0 \qquad (14)$$

where Poynting's vector measuring the energy flow across a given surface S_0 in reference state is given by :

$$\mathcal{P}_L = -\left(K_{L\alpha}^M + K_{L\alpha}^E\right)v_\alpha + \varphi\dot{D}_L \qquad (15)$$

As a matter of fact, this holds for real values of K, D and v. In the present case, using a complex notation for harmonically varying functions with respect to time complies us to evaluate the real part of the averaged energy flow across any surface S_0 during one period of time, which is equal to :

$$\langle \mathcal{P}_L(X_1)\rangle = -\frac{\omega}{2}\int_{S_0}\Re(K_{L\alpha}(v_\alpha)^*)dS_0 \qquad (16)$$

where $v_\alpha = -j\omega u_\alpha$ with present notations. Two special cases are of interest : a transducer excited by a Rayleigh wave incident at one end, and a transducer excited by a BAW coming from the depth of substrate under normal incidence. In the first case, it is sufficient to consider a combination of both branches δ^+ and δ^- to satisfy the simple physical end–condition : the cumulated amplitude of

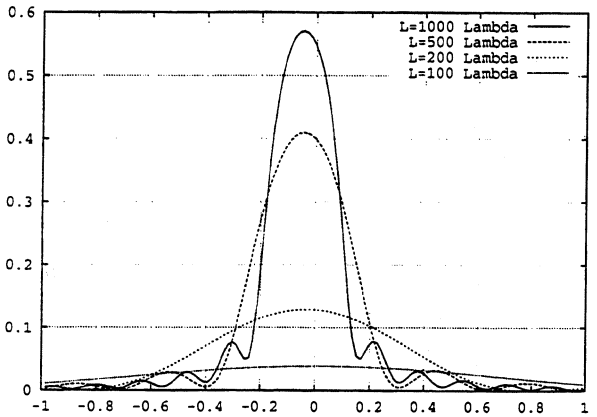


Figure 3: SAW energy reflection *vs* $\Delta\omega/(3.10^{-3}\omega_0)$.

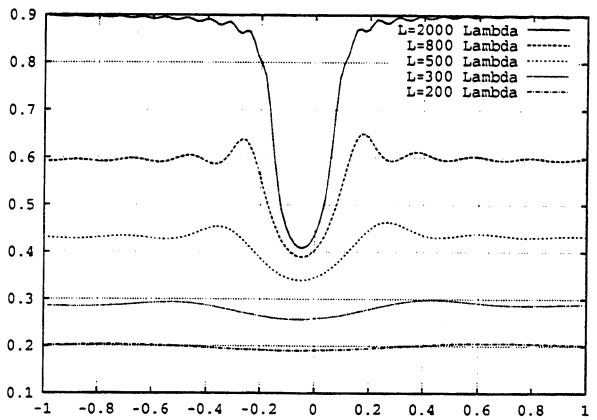


Figure 4: SAW–BAW conversion *vs* $\Delta\omega/(3.10^{-3}\omega_0)$.

SAW harmonics *entering* the coupler's region at the opposite side with respect to the incident SAW should be zero. In the second case, to the combination of both branches we must add the particular Floquet sum with $\delta = 0$ which equilibrates the boundary stress arising from BAW incident wave. Then the relative weights of both eigenmodes are obtained from setting to zero the cumulated amplitudes of all harmonics *entering* the coupler's region at *both ends*. From Eq(16), it can be easily shown that this is sufficient to ensure that no other form of energy than the assumed incident wave enters the coupler's region. In this manner, results of Figs.3, 4 and 5 were obtained for the following parameters : $a/\Lambda = 0.125$, $k \in [-4, 4]$, X_1 axis of AT-cut, $\varepsilon = 0.05$. Figs.3 and 4 respectively present the SAW energy reflection and SAW to BAW energy conversion coefficients for incident wave coming upon one end of coupler, for various coupler lengthes, between 100Λ and 2000Λ. Fig. 5 gives the energy conversion from incident BAW to SAW at one end of the coupler, for various lengthes, in case of either incident extensional (curves with the maximum at $\omega > \omega_0$) or slow shear wave (maximum at $\omega < \omega_0$). In all here-studied cases, an incident fast shear wave generates much smaller SAW than other kinds of BAW waves.

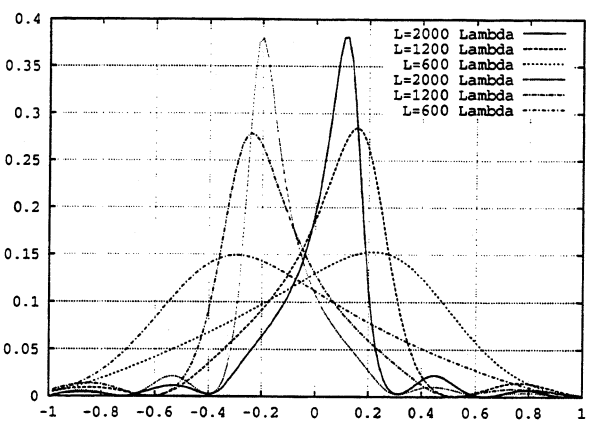


Figure 5: BAW–SAW conversion *vs* $\Delta\omega/(3.10^{-3}\omega_0)$.

IV. Conclusion

The completed work shows that the behavior predicted in Ref.[1] still qualitatively applies for the case of X_1 propagation in singly rotated cuts of quartz or similar symmetry. At least 5 harmonics $k \in [-2, 2]$ must be considered for sawtooth profile, this number quickly increases for the modelling of steep profiles. The behavior of the structural coupler strongly depends on the number of period of gratings and the slope of profile. High order recombinations and rereflections in longer gratings provide with a strongest resonant behavior, narrowing the band of the BAW/SAW transduction, in comparison with shorter transducers. BAW–to–SAW and SAW–to–BAW transduction behave differently : in one case, the dispersive behavior results into a cut band for SAW–to–BAW transduction, in the other case, it results into a pass band for BAW–to-SAW transduction.

References

[1] S.V. Biryukov, Y. Gulyaev, V. Krylov, V. Plessky, Chapter 8, Surface Acoustic Waves in Homogeneous Media, Springer, 1995.

[2] B. Dulmet, H. Watchueng, Proc. EFTF 1999, pp326–331 , Warsaw, March 1998.

[3] B. Dulmet and R. Bourquin, Proc of IEEE Symp on Ultrasonics, pp 331–335, Cannes, Nov 1994.

[4] B. Dulmet, R. Bourquin, S. Ballandras, E. Bigler, Proc. Piezo96' Conf., pp 31–36, Ed. ITR, Varsaw, Oct. 1996.

[5] H. Watchueng, Ph.D thesis, Number 691, University of Franche-Comté, October 1998.

[6] M. Mc Carthy and H. F. Tiersten, Arch. Rat. Mech. An., **(68)**, 1, pp 27–36, 1978.

SCATTERING MATRIX APPROACH TO ONE PORT SAW RESONATORS

Waldemar Soluch
Institute of Electronic Materials Technology
Wolczynska 133, 01-919 Warsaw, Poland

ABSTRACT

The scattering matrix method was used for the derivation of an expression for the reflection coeficient of a one port resonator. This expression was used for the calculations of an input admittance of a synchronous resonator on ST cut quartz. Measured and calculated parameters of the resonator were in very good agreement.

1. INTRODUCTION

The scattering matrix approach was first used for derivation of an analytical expression for the transfer function of a two port SAW resonator [1]. In the case of one port resonator we should derive an expression for the reflection coefficient, then, for a given transmission line, the input admittance and the equivalent circuit parameters of the resonator, can be determined easily.

2. REFLECTION COEFFICIENT AND ADMITTANCE OF THE ONE PORT RESONATOR

The one port SAW resonator (Fig.1) consists of one interdigital transducer (IDT) and two reflectors. We neglect the reflections from the single electrodes of the IDT. We also assume that the IDT is symmetrical and located in any position between the reflectors. There are four normalized amplitudes in the acoustic area: a_{f1} and a_{f2} - forward traveling waves in region 1 and 2, respectively, and a_{b1} and $_{b2}$ - backward traveling waves in the above regions, respectively. At the electrical port

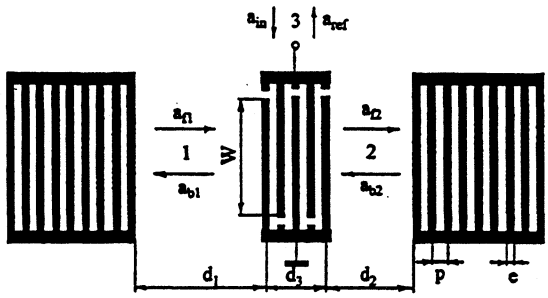


Fig.1. One port SAW resonator.

3 of the IDT we have an incident a_{in} and a reflected a_{ref} wave, respectively.

The one port resonator is described by the following linear equations:

$$a_{f1} = a_{b1} r_1 , \tag{1}$$

$$a_{b1} = a_{in} S_{13} + a_{f1} S_{11} + a_{b2} S_{12} , \tag{2}$$

$$a_{f2} = a_{in} S_{13} + a_{f1} S_{12} + a_{b2} S_{11} , \tag{3}$$

$$a_{b2} = a_{f2} r_2 , \tag{4}$$

$$a_{ref} = a_{in} S_{33} + a_{f1} S_{13} + a_{b2} S_{13} , \tag{5}$$

where

$$r_1 = \Gamma T_i e^{-j\Phi_1} , \tag{6}$$

$$r_2 = \Gamma T_i e^{-j\Phi_2} , \tag{7}$$

$$\Phi_1 = 2\omega\left(\frac{d_3}{2v_t} + \frac{d_1}{v_f}\right), \tag{8}$$

$$\Phi_2 = 2\omega\left(\frac{d_3}{2v_t} + \frac{d_2}{v_f}\right). \tag{9}$$

Here S_{ij} are the scattering coefficients of the IDT, r_1 and r_2 are the reflection coefficients of the reflectors transformed to the center of the IDT, Φ_1 and Φ_2 are the double phase changes from the center of the IDT to the reflectors in the areas 1 and 2, respectively, d_1 and d_2 are the distances between the centers of the IDT electrodes and the edges of the reflectors electrodes, d_3 is the distance between the centers of the first and the last electrodes of the IDT, v_t and v_f are the SAW velocities in the IDT and free surface areas, respectively, Γ is the reflection coefficient of the reflector (two identical reflectors are assumed), and T_i is the loss coefficient.

The reflection coefficient r of a one port resonator is defined as

$$r = \frac{a_{ref}}{a_{in}} , \tag{10}$$

and can be obtained from expression (5) as

$$r = S_{33} + \frac{a_{f1}}{a_{in}} S_{13} + \frac{a_{b2}}{a_{in}} S_{13} . \tag{11}$$

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

The normalized amplitudes can be easily determined from equations (1) - (4), and the following expression for r is obtained:

$$r = S_{33} + \frac{[r_1 + r_2 - 2r_1r_2(S_{11} - S_{12})]S_{13}^2}{(1 - r_1S_{11})(1 - r_2S_{11}) - r_1r_2S_{12}^2}. \quad (12)$$

The first term in the above expression represents a reflection coefficient of the IDT, and the second one represents a reflection coefficient of the reflectors seen at the electrical port 3 of the IDT. If $r_1 = r_2 = 0$, then $r = S_{33}$, as expected.

The expressions for the scattering matrix coefficients S_{ij} of the IDT [2 -4], and for the reflection coefficient Γ of the reflector [1],[3], are presented in the APPENDIX.

Assuming, that $B_L = 0$ and $G_L = Y_0$, where Y_0 is the characteristic admittance of a tranmission line, we can calculate the admittance Y of the one port resonator as [5]

$$Y = Y_0 \frac{1 - r}{1 + r}. \quad (13)$$

All parameters of the one port resonator can be determined from the calculated admittance Y as a function of frequency.

3. CALCULATIONS AND MEASUREMENTS OF ONE PORT SYNCHRONOUS RESONATOR ON ST CUT QUARTZ

The one port synchronous resonator was obtained from the two port one by short circuiting one of the IDTs [6] (Fig. 2). For the IDT located inside the reflector we have: $d_3 = 0$, $d_1 = 0.75p$, $\Phi_1 = 2\omega d_1/v_r$, and $d_2 = n\lambda_f/2 + p/4$, where n is an integer, λ_f is the SAW wavelength for the free surface, and v_r is the SAW velocity in the reflector area.

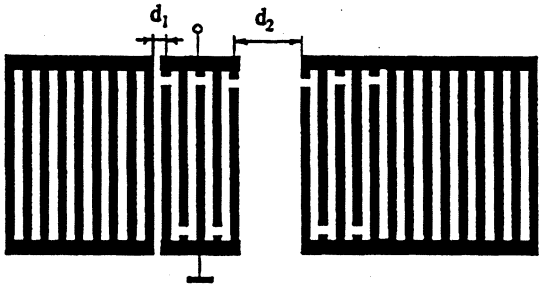


Fig.2. One port synchronous resonator.

For the ST cut quartz, the following data were used [7 - 9]: $v_f = 3158$ m/s, $K^2 = 0.0016$, $\varepsilon_r = 4.5$, $\gamma = -0.5h/\lambda_r$, where K is the electromechanical coupling coefficient, ε_r is the relative dielectric constant, h is the aluminium layer thickness, and $\lambda_r = 2p$.

The following data were used for the resonator at a frequency near 303 MHz: W = 1 mm, p = 5.2 μm, e = 2.6 μm, n = 21, $N_t = 91$, and $N_r = 700$, where N_t is the numer of the IDT electrodes, and N_r is the total number of electrodes in each reflector. Cosine function was used for the IDT apodization to eliminate the transverse spurious modes [6].

Velocity v_r and loss coefficient T_i were determined from the following experimental expressions [6]:

$$v_r = v_f\{1 - e/p[5.8\,10^{-4} + 0.31(h/\lambda_r)] - 5.73(h/\lambda_r)^2\}, \quad (14)$$

$$T_i = \frac{1}{1 + 72.3 \cdot 10^{-12} f_0}. \quad (15)$$

The ratio of the width e to the period p of the electrodes was introduced here to take into account the effect of e/p variation (Fig.1).

The admittance and the static capacitance of the apodized IDT were calculated by the method presented in [10] ($v_t = v_r$). The scattering matrix coefficients S_{ij} of the IDT and the reflection coefficient Γ of the reflector were determined from expressions (A.1) - (A.5). The reflection coefficient Γ of the reflector and the velocity v_r in its area were calculated for the aluminium layer thickness h = 85 nm. The loss coefficient $T_i = 0.978$ was used for the calculations at the frequency 303 MHz. Next, the reflection coefficient r of the resonator was calculated from expression (12), and the real and imaginary parts of the resonator admittance (Y = G +jB) were determined from expression (13) ($Y_0 = 20$ mS). The results of the calculations are shown in Fig.3.

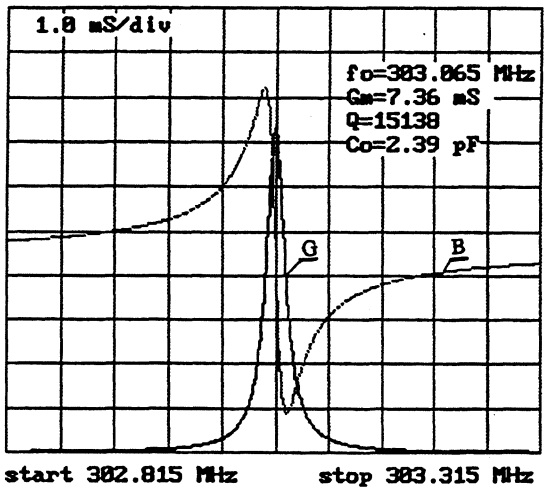


Fig.3. Calculated admittance of the resonator.

Table 1. One port synchronous resonator parameters.

Parameter	f_0 [MHz]	L_1 [mH]	C_1 [fF]	R_1 [Ω]	C_0 [pF]	Q -
Calculated	303.065	1.08	0.26	136	2.4	15 000
Measured	303.106	1.08	0.25	142	2.3	14 500

The one port resonator admittance can be represented by the equivalent circuit shown in Fig.4, and all of the circuit elements can be deduced from the above calculations ($G = G_1$, $B = B_1 + \omega C_0$). The real part G of the equivalent circuit admittance can be written as

$$G = \frac{G_m}{1 + x^2}, \tag{16}$$

where

$$G_m = \frac{1}{R_1}, \tag{17}$$

$$x = 2Q\frac{\Delta f}{f_0}, \tag{18}$$

$$Q = \frac{\omega_0 L_1}{R_1}, \tag{19}$$

$\Delta f = f - f_0$, G_m is the maximum value of G, and Q is the quality factor. For $x = 1$, $G = G_m/2$ and $Q = f_0/\Delta f_{3dB}$,

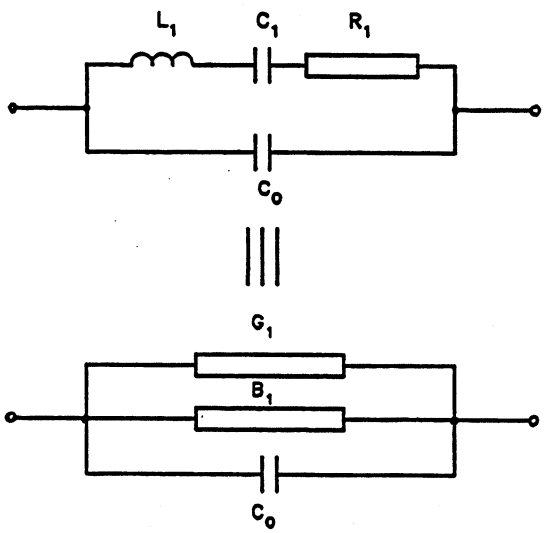


Fig.4. Equivalent circuit of the resonator.

where Δf_{3dB} is the banwidth at the $G_m/2$ level. From the calculated G as a function of frequency (Fig.3), we can determine G_m, Δf_{3dB}, f_0, and Q. Then, the equivalent circuit elements are obtained from the expressions:

$$R_1 = \frac{1}{G_m}, \tag{20}$$

$$L_1 = \frac{QR_1}{\omega_0}, \tag{21}$$

$$C_1 = \frac{1}{\omega_0^2 L_1}. \tag{22}$$

The static capacitance C_0 is equal to that of the IDT. However, some additional capacitance is usually added to take into account the resonator package.

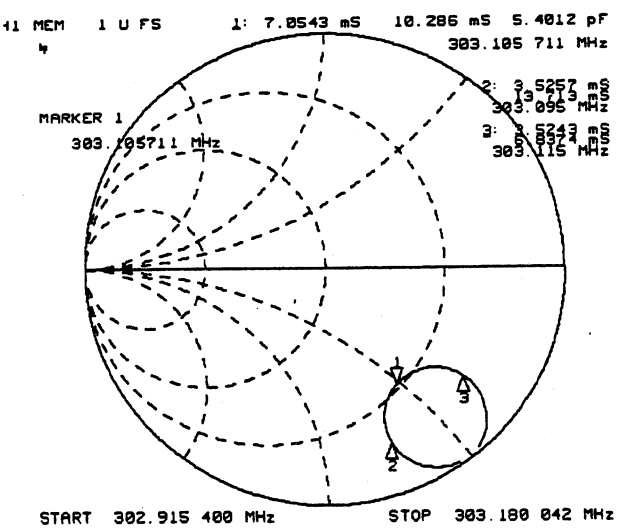


Fig.5. Admittance circle of the resonator.

The resonator was fabricated and measured in a 50 Ω ystem (HP Network Analyzer 8752A in the reflection mode). The mesured admittance circle is shown in Fig.5. Marker 1 is used for the resonance frequency f_0, and markers 2 and 3 for the frequencies at which the condition $G_1=G_m/2$ is satisfied. The results of calculations and measurements are shown in Table 1. The agreement between calculated and measured equivalent circuit parameters of the resonator is very good.

4. CONCLUSION

It was shown, that the scattering matrix approach is a very convenient method for calculation of the parameters of one port SAW resonators. The expression for the reflection coefficient can be used for synchronous and non-synchronous resonators on any piezoelectric substrate.

REFERENCES

[1] L.A.Coldren and R.L.Rosenberg, "Scattering matrix approach to SAW resonators", Proc. IEEE Ultrason. Symp., 1976, pp. 266-271.

[2] S.G.Joshi and P.Sudhakar, "Scattering parameters of interdigital surface acoustic wave transducers", IEEE Trans., Sonics and Ultrason., vol. SU-24, pp. 201-206, May 1977.

[3] R.L.Rosenberg and L.A.Coldren, "Scattering analysis and design of SAW resonator filters", IEEE Trans., Sonics and Ultrason., vol. SU-26, pp. 205-225, May 1979.

[4] T.F.O'Shea and R.C.Rosenfeld, "SAW resonator filters with optimized transducer rejection", Proc. IEEE Ultrason. Symp., 1981, pp. 105-110.

[5] Surface acoustic wave (SAW) resonators. Part 1. General information, standart values and test conditions. IEC 1019-1-2, 1993.

[6] W.Soluch, "Design of SAW resonators on ST cut quartz", Proc. 12th European Frequency and Time Forum, 1998, pp. 198-201.

[7] A.J.Slobodnik, "Surface acoustic waves and SAW materials", Proc. IEEE, vol. 64, pp. 581-595, May 1976.

[8] C.Dundrowicz, "Reflection of surface waves from periodic dicontinuities", Proc. IEEE Ultrason. Symp., 1976, pp. 386-390.

[9] W.J.Tanski and H. van de Vart, "The design of SAW resonators on quartz with emphasis on two ports", Proc. IEEE Ultrason. Symp., 1976, pp. 260-265.

[10]. W.Soluch, "Admittance matrix of a surface acoustic wave interdigital transducer", IEEE Trans. on UFFC, vol.40, pp. 828-831.

APPENDIX

1. Scattering matrix coefficients of the ITD.

The scattering matrix coefficients S_{ij} of the symmetrical IDT can be determined from the following expressions [2 - 4]:

$$S_{11} = \frac{-G_a}{D}, \qquad (A.1)$$

$$S_{12} = \frac{G_L + j(B_a + \omega C_0 + B_L)}{D}, \qquad (A.2)$$

$$S_{13} = -j\frac{\sqrt{2G_L G_a}}{D}, \qquad (A.3)$$

$$S_{33} = \frac{G_L - G_a - j(B_L + B_a + \omega C_0)}{D}, \qquad (A.4)$$

where

$$D = G_L + G_a + j(B_L + B_a + \omega C_0),$$

G_a and B_a are the real and imaginary parts of the IDT admittance, G_L and B_L are the real and imaginary parts of the load admittance, respectively, and C_0 is the static capacitance of the IDT.

2. Reflection coefficient of the reflector

The reflection coefficient Γ, defined at the initial edge of the reflector, can be calculated from the expression [1, 3]

$$\Gamma = \frac{\kappa \tanh(\sigma N_r p)}{\sigma + j\delta \tanh(\sigma N_r p)}, \qquad (A.5)$$

where

$$\sigma = \sqrt{(\kappa^2 - \delta^2)},$$

$$\kappa = \frac{\gamma}{p},$$

$$\delta = 2\pi\frac{f - f_0}{v_r},$$

$$f_0 = \frac{v_r}{2p}.$$

and N_r is the number of strips in the reflector, p is the period of the strips, γ is the reflection coefficient of one strip, v_r is the SAW velocity in the reflector area, and f_0 is the center frequency.

CHARACTERIZATION OF SINGLE-PORT SURFACE TRANSVERSE WAVE RESONATORS IN THE LOWER GHz RANGE

IVAN D. AVRAMOV

Institute of Solid State Physics, 72 Tzarigradsko Chaussee Blvd., 1784 Sofia, Bulgaria

ABSTRACT

This paper presents results from an experimental study on single-port surface transverse wave resonators (STWR) operating in the 1.0 to 2.0 GHz range. The equivalent circuit parameters are extracted in transmission and reflection measurements and compared with data obtained from bulk acoustic wave (BAW) and surface acoustic wave (SAW) based single-port resonators, operating at much lower frequencies. It is shown that, at GHz frequencies, practical single-port STW devices feature extremely low values of the dynamic resistance. On one hand, this results in device Q greatly exceeding the material Q limit for SAW and BAW and on the other hand, it allows the design of simple negative resistance oscillators (NRO) with excellent short term stability, high output power and unprecedented RF/d.c. efficiency.

1. INTRODUCTION

STW based resonant devices for low-noise oscillator applications have enjoyed considerable attention over the last decade because at frequencies around and above 1 GHz they perform much better than their SAW counterparts in terms of device Q, power handling ability and 1/f phase noise [1]. While a lot of work has been dedicated to two-port STW resonators for feedback loop oscillator (FLO) applications, single-port devices have received little attention. This is because many design engineers believe that, because of stray capacitance effects, it is impossible to build a good microwave oscillator with a single-port acoustic device. There are at least four reasons why in many microwave oscillator applications, a single-port STWR should be preferred to a two-port one:

1. Single-port STW resonators have generally higher device Q than two-port ones [2], [3].
2. Since at frequencies above 1 GHz, the active element in the circuit is the dominant source of 1/f noise, the higher Q of the single-port device will reduce the close-to-carrier phase noise, compared to FLO using two-port devices.
3. Although, STW two-port resonators demonstrate extremely low 1/f noise, typically in the -142 to -144 dBc/Hz range at 1 Hz intercept [1], additional improvement is expected with single-port devices since one interdigital transducer (IDT) will generate less 1/f noise than two.
4. STW single-port devices, with their extremely high RF power handling ability, completely eliminate drive power level problems which cause serious aging problems in BAW crystals. This allows the design of fairly simple, power efficient microwave oscillators, with low thermal noise floor, in which the STW single-port resonator operates at a high drive level.

This paper presents and discusses data on high-performance single-port STWR intended for use in NRO with improved close-to-carrier phase noise, high output power and increased RF/d.c. efficiency. The STW resonator performance is compared with BAW and SAW single-port devices operating at much lower frequencies. First results on high-performance NRO using simplified designs are presented.

2. MEASUREMENT METHODS

As shown in [2], in the vicinity of resonance, STW single-port resonators are described by the well known Butterworth-Van Dyke equivalent electrical circuit which is widely used to characterize the resonant behavior of BAW crystals [4]. As evident from Fig. 1, this circuit consists of a series resonant branch (R_d, L_d and C_d), characterizing the motional resonator behavior, which is shunted by the static resonator capacitance C_o. The equivalent circuit parameters can be extracted according to the methods described below.

2. 1. Method M1: The resistive Π-type circuit

If the resonator is connected to a Π-type resistive circuit, as the one shown in Fig. 1, and its transmission characteristic S21 is measured, then a series and parallel resonance at the frequencies of minimum and maximum attenuation f_s and f_p,

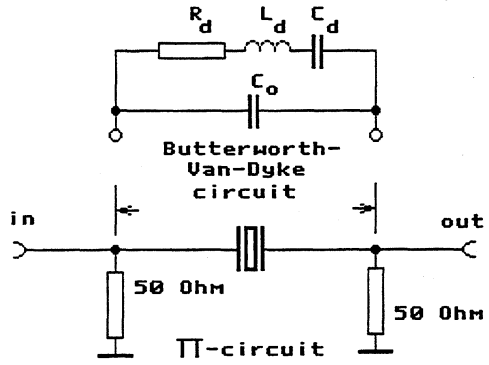


Fig. 1. The Butterworth-Van Dyke equivalent electrical circuit of a single-port acoustic wave resonator, embedded in a Π-type circuit for parameter extraction.

respectively, are clearly observed. The plots in Fig. 2 characterize a 9.14 MHz BAW crystal measured in this way. As shown in [5], from the distance between f_s and f_p which is 19.6 KHz in this measurement, the capacitance ratio C_d/C_o and the device Q can be calculated with a reasonable accuracy as:

$$C_d / C_o = 2(f_p - f_s) / f_s \quad (1)$$

$$Q = 1/[4\pi C_o R_d (f_p - f_s)] \quad (2)$$

The static capacitance is typically measured with a capacitance meter at a low frequency far off resonance and the dynamic resistance R_d can be obtained by replacing the resonator in the Π-circuit with a variable resistor, in parallel with a low-inductance capacitor of the same value as C_o. The variable resistance value, yielding at f_s the same insertion loss as the resonator is the value of the dynamic resistance.

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

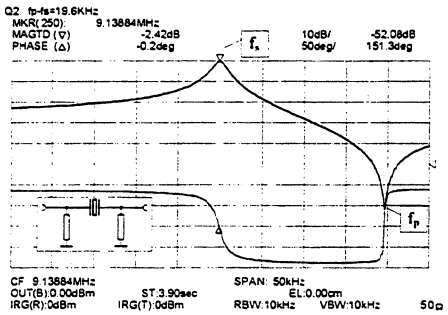


Fig. 2. Frequency and phase responses of a 9.14 MHz fundamental mode BAW crystal. Note that at f_s and f_p the phase is zero meaning that the influence of C_o is negligible.

2. 2. Method M2: Admittance Measurement

If high precision is necessary, then the Q value can be obtained by measuring the complex admittance of the resonator with a vector analyzer and reading characteristic values on the Smith chart. In Fig. 3 this method is illustrated with the BAW crystal from Fig. 2. As shown in [4], f_s is the frequency of maximum conductance G_{max} and the difference between the

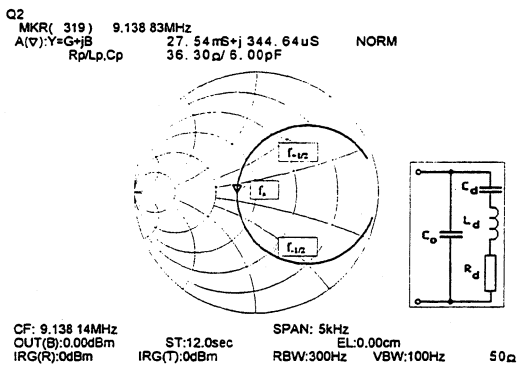


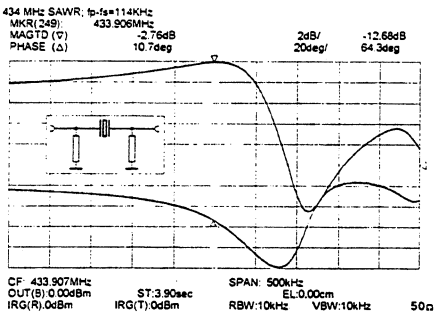
Fig. 3. Admittance plot of the device from Fig. 2.

frequencies $f_{-1/2}$ and $f_{+1/2}$ of maximum susceptance B_{max} is the 3 dB device bandwidth and, therefore, the device Q. According to [6] the frequencies $f_{-1/2}$ and $f_{+1/2}$ correspond to the points at which the conductance is equal to $G_{max}/2$, yielding the same result for the Q-factor. At frequencies, below 100 MHz the influence of C_o and package parasitics is negligible, the admittance plot is symmetric with respect to the real Smith chart axis and crosses it at the frequency f_s where $R_d=1/G_{max}$. This, however, is true only if $R_d<<1/(2\pi f_s C_o)$.

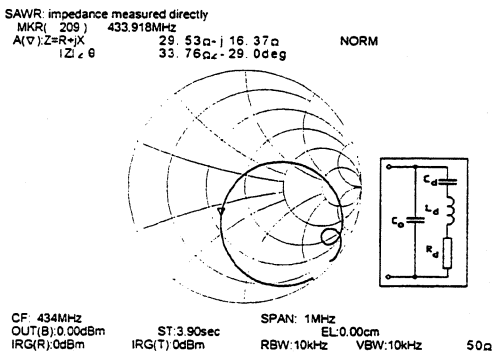
The Q of the 9.14 MHz BAW crystal, characterized with methods M1 and M2, was measured as **21100** and **20770** respectively, indicating an excellent agreement between them.

2. 3. Characterization of SAW single-port resonators

Although SAWR operate at much higher frequencies than their BAW counterparts, methods M1 and M2 apply to SAW single-port resonators too. Figures 4 a) and b) show results obtained from a 433.92 MHz commercially available SAW resonator. This device behaves in the same way as the BAW crystal from Fig. 2 and 3. Figure 4 a) indicates an (f_p-f_s) value of 114 KHz. Under the false assumption that $R_d=1/G_{max}$, from Fig. 4 b), the R_d value is measured as 29.5 Ω. With a C_o value of 2.1 pF, according to equations (1) and (2), the device Q is calculated as 11270 which is in fact too high. The correct Q value is **9170** and is obtained by reading the $(f_{+1/2}-f_{-1/2})$ value from the plot in Fig. 4b) which is 47.3 KHz in this measurement. The reason for the erratic initial Q-measurement was that C_o and package parasitics move the impedance plot into the capacitive part of the Smith chart where $R_d \neq 1/G_{max}$.



a)



b)

Fig. 4. Characteristics of a 433.92 MHz SAW resonator according to a): Method M1 and b): Method M2.

2. 4. Method M3 for correct R_d measurement

If an inductor L_o is connected in parallel with the single-port resonator, and its value is selected in such manner that C_o and all package parasitics are tuned out, then, at series resonance, the resonator will behave like an ohmic resistance equal to R_d. This is illustrated in Fig. 5 where R_d is read directly as 36.41 Ω at the point where the impedance locus crosses the real axis of the Smith chart. If this value is used to calculate the device Q according to Method M1, then the result is **9130** and is in excellent agreement with the value of **9170** measured precisely with Method M2. This means that the

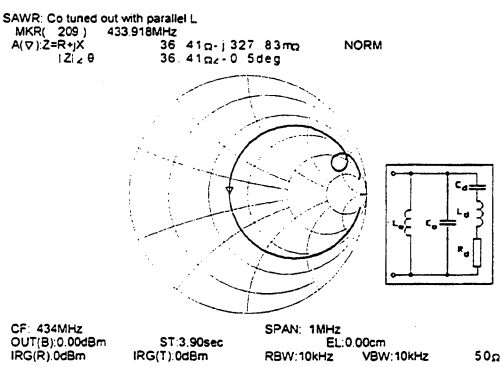


Fig. 5. Impedance of the 433.92 MHz SAW device according to Method M3.

parallel inductor does not compromise the measurement result. Practically, the L_o value does not even need to be very precise because the bandwidth of the tank circuit that it forms with C_o and the package parasitics, is orders of magnitude smaller than the bandwidth of the acoustic single-port resonator.

2. 5. Characterization of STW single-port resonators

Method M2 yields precise Q results if applied to devices with well behaved series resonance, the admittance of which is a nice uniform circle on the Smith chart. Unfortunately, this is rarely the case with STWR. The necessity of locating the resonance at the lower stopband edge to maximize the device Q and minimize its loss, generally yields asymmetric responses. This is evident from Fig. 6 a) and b) characterizing a 2.0 GHz STWR. The asymmetry of the plots makes it difficult to read the frequencies $f_{-1/2}$ and $f_{+1/2}$ and determine the device Q directly from the admittance chart (Fig. 4b)), as shown in Section 2.3. On the other hand, Method M1 provides a sufficiently accurate reading of the frequencies f_s and f_p. For the complete device characterization we need also the C_o value, which is measured with a capacitance meter, and the R_d value which is measured precisely according to Method M3 as described in the previous section. Thus, the characterization of STW single-port devices requires 3 steps:

1. Measure f_s and f_p in a Π-circuit according to Method M1;
2. Tune out C_o with a parallel inductor and read R_d directly from an impedance measurement using Method M3;
3. Measure C_o and calculate the capacitance ratio and device Q using equations (1) and (2).

In Fig. 6 a) and b) the (f_p-f_s) and R_d values are measured as 860 KHz and 8.35 Ω respectively. This yields a device Q of 6800.

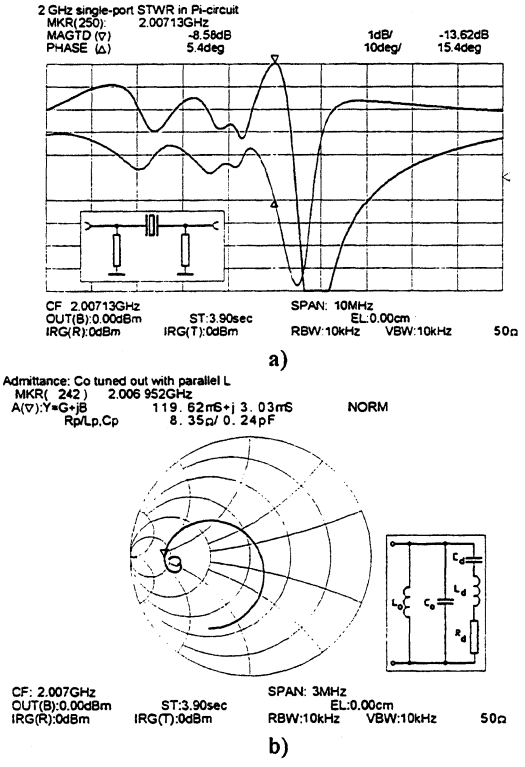


Fig. 6. Data from a 2 GHz single-port STWR measured with: a): Method M1 and b): Method M3.

3. RESULTS FROM EVALUATION OF SINGLE-PORT STW, BAW AND SAW DEVICES

Table 1 summarizes results from different SAW, BAW and STW devices characterized with the methods described in the previous sections. In addition to the fundamental mode 9.14 MHz BAWR, the 434 MHz SAWR and the 2 GHz STW device used in this paper, the table includes data from a 1 GHz extremely high Q STW device from Ref. [3], a 1.244 GHz STWR from Ref. [2] and the best 18.226 MHz fundamental mode BAW crystal that the author had available for this study.

Device	C_d/fF	C_o/C_d	R_d/Ω	Device Q
9.14 MHz BAWR this paper	22.73	233	36.3	21100
434 MHz SAWR this paper	1.103	1904	36.41	9130
2.0 GHz STWR this paper	1.397	**1167**	**8.35**	6800
1.244 GHz STWR Ref. [2]	0.76	**2140**	**9.38**	17950
1.0 GHz STWR Ref. [3]	0.671	**3130**	**5.7**	41550
18 MHz BAWR this work	25.62	232	5.61	60760

Table 1. Performance of GHz range STW single-port resonators compared with SAWR and BAWR devices.

The evaluation of the data in Table 1 allows the following conclusions:

1. Single-port STW resonators in the frequency range 1.0 to 2.0 GHz demonstrate device Q values comparable with the Q of BAW crystals operating below 100 MHz.
2. Both factors: the high capacitance ratio and the low dynamic resistance of STWR contribute to the high device Q which can exceed the material Q limit for SAW on rotated Y-cut quartz by more than 3 times [3].
3. The capacitance ratio of GHz range STWR exceeds by an order of magnitude the capacitance ratio of BAW fundamental mode crystals and is an indication of a low distance between series and parallel resonance. This results in a very high phase slope which is the key for achieving low oscillator phase noise. This is evident from the data in Fig. 6 a) too, (see also References [2] and [3]).
4. Single-port SAWR feature a capacitance ratio comparable to what is achievable with STWR but the SAWR R_d value is generally much too high to achieve high device Q.

4. NEGATIVE RESISTANCE STW OSCILLATORS

As implied in the Introduction, the main idea of this paper was not only to demonstrate the superiority of STW single-port resonators in terms of device Q, but also to show their feasibility for applications in microwave oscillators with improved short-term stability, high output power and increased RF/d. c. efficiency, as well as simplified design. Among RF engineers the NRO is considered to be the best choice for that, especially at microwave frequencies. Single-port STWR seem to be ideal for stabilizing NRO, since they have very low R_d values. On one hand, this makes it easy to ground a transistor terminal, on the other hand, the amount of negative resistance

that has to be generated by the transistor to compensate for the R_d loss is low accordingly. Thus, oscillation starts well and the oscillator does not generate parasitic oscillations.

The first experiments performed by the author used the basic NRO circuits described in Fig. 7 a), b) and c). In the circuits in Fig. 7 a) and b), the resonator operates at series resonance and grounds the base and emitter of the transistor, respectively, while the complex impedances Z1 and Z2 set the phase condition of oscillation. The circuit in Fig. 7 c) uses the popular grounded collector configuration which is known for its low phase noise in voltage controlled oscillator (VCO) applications. With Z2 tunable, the circuit in Fig 7 c) was found to provide about 150 ppm of tuning range when stabilized with a 1 GHz STWR.

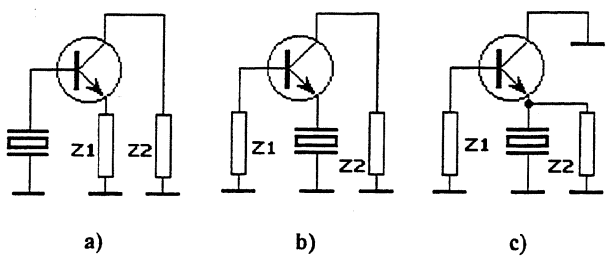


Fig. 7. Simplified schematics of the negative resistance oscillators built and tested in this study.

One great advantage of the NRO is its high power efficiency. Since STWR can stand extremely high incident power levels [1], STW based power NRO can be realized in a very elegant way. The circuit from Fig. 7 a) was used in a 1 GHz grounded base configuration and was run at 23 dBm of output power for several days. The RF/d.c. efficiency was in excess of 30% while the short-term stability was measured as $7x10^{-10}$/s. NRO operating at lower power levels were found to have short-term stability values in the 2 to $5x10^{-10}$/s range.

Most of the 1 GHz STW resonators in this study were in fact two-port devices. To use them as single-port resonators, only one of the IDT was used while the other one was left open. Thus, the same device could be used to evaluate both: FLO and NRO, and compare their short term stability. This approach was used to make sure that the 1/f phase noise contribution of the acoustic device has the same source for both oscillator types. For the same output power, the NRO was found to provide a 2 to 4 times better short-term stability than the FLO using the same STW device in a two-port configuration.

5. FEEDBACK LOOP OSCILLATORS USING SINGLE PORT STW RESONATORS

A further way to improve the overall oscillator phase noise performance while keeping the 1/f noise contribution of the acoustic device at a minimum, is to use it as a single-port device in a capacitive Π-circuit configuration. Although a single-port STWR is used, the Π-circuit behaves as a two-port resonator. Thus, the 1/f noise contribution of a second IDT is eliminated. The frequency response of 1 GHz STWR embedded in such a capacitive Π-circuit is shown in Fig. 8. Increasing the shunt capacitance values improves the sidelobe suppression and increases the circuit Q at the expense of higher circuit loss. Thus, a trade-off between oscillator phase noise and tuning range can be achieved if the circuit is used in a voltage controlled FLO. The lowest phase noise is achieved if the FLO is operated slightly above series resonance where the phase slope is maximum, (see Fig. 6 a)).

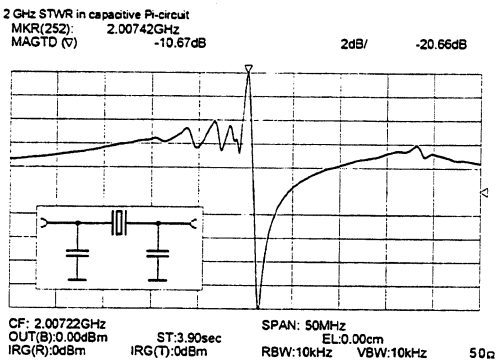


Fig. 8. Frequency response of a 2 GHz capacitive Π-circuit using the single-port STWR from Fig. 6.

6. SUMMARY AND CONCLUSIONS

This study has shown that with respect to device Q, single-port STWR at GHz frequencies are superior to SAWR operating below 1 GHz and comparable with BAW crystals operating below 100 MHz. It has been shown that such GHz range STWR have a great potential for the realization of microwave NRO with improved phase noise, increased output power and efficiency and simplified design. Further studies in this direction are on the way.

Acknowledgments

The author wishes to gratefully acknowledge Vectron Technologies (VTI) for providing the 1 GHz STW devices and RF Monolithics (RFM) for providing the 434 SAW devices used in this study, as well as Fijitsu Laboratories in Akashi, Japan for the opportunity to fabricate the 2 GHz devices there.

REFERENCES

[1.] I. D. Avramov, F. L. Walls, T. E Parker and G. K. Montress, "Extremely Low Thermal Noise Floor, High Power Oscillators Using Surface Transverse Wave Devices", IEEE Trans. Ultrasonics, Ferroelectrics and Freq. Control, Vol. 43. No. 1, January 1996, pp. 20-29.

[2.] I. D. Avramov, "Increasing the Loaded Q of GHz Range Surface Transverse Wave Resonators", Proc. 1997 European Frequency and Time Forum, March 4-7, 1997, Neuchatel, Switzerland, pp. 394-398.

[3.] J. A. Kosinski, R. Pastore and I. D. Avramov, "Theoretical and Experimental Evidence of Superior Intrinsic Q of STW Devices on Rotated Y-cut Quartz", elsewhere in this Proceedings.

[4.] M. Schmid, E. Benes and R. Sedlaczek, "A Computer-Controlled System for the Measurement of Complete Admittance Spectra of Piezoelectric Resonators", Meas. Sci. Technol., No. 1 (1990), pp. 970-975 (UK).

[5.] G. T. Shitikov, P. Y. Cuigankov and O. M. Orlov, "Highly Stable Quartz Oscillators", Moscow, Soviet Radio, 1974, pp. 32-36, (in Russian).

[6.] W. Soluch, "Design of SAW Resonators on ST Cut Quartz", Proc. 1998 European Freq. and Time Forum, March 10-12, 1998, Warsaw, Poland, pp. 198-201.

1999 Joint Meeting EFTF - IEEE IFCS

THEORETICAL AND EXPERIMENTAL EVIDENCE FOR SUPERIOR INTRINSIC Q OF STW DEVICES ON ROTATED Y-CUT QUARTZ

John A. Kosinski and Robert Pastore, U.S. Army CECOM
AMSEL-RD-IW-TI, Fort Monmouth, NJ 07703-5211, USA

Ivan D. Avramov, Institute of Solid State Physics
72 Tzarigradsko Chaussee Blvd., 1784 Sofia, Bulgaria

ABSTRACT

Several advantages of STW devices are by now well known. These include simultaneous moderately high acoustic velocity and zero temperature coefficient of frequency, along with extremely high power handling capability. In this paper we present theoretical and experimental evidence of an additional advantage of STW devices on quasi- AT-cut quartz: the intrinsic Q of such devices is superior to that of ST-cut SAW devices and AT-cut BAW devices, with the STW having $Q \times f \approx 3.25 \times 10^{13}$ Hz

1. INTRODUCTION

Several advantages of STW devices are by now well known. These include simultaneous moderately high acoustic velocity and zero temperature coefficient of frequency, along with extremely high power handling capability. These properties allow reliable fabrication of temperature-stable surface wave devices operating into the low microwave region, with superior phase noise floor as compared to other technologies. However, thus far STW devices have been considered as inferior to SAW and BAW devices with regard to close-in phase noise arising from intrinsic Q considerations.

In this paper we present theoretical and experimental evidence that the intrinsic Q of STW devices on quasi- AT-cut quartz is, in fact, superior to that of ST-cut SAW devices ($Q \times f \approx 1.05 \times 10^{13}$ Hz [1]) and AT-cut BAW devices ($Q \times f \approx 1.35 \times 10^{13}$ Hz [2]). The calculations indicate that the intrinsic Q of the STW mode is 2×-3× larger than those of the other technologies, with the STW having $Q \times f \approx 3.25 \times 10^{13}$ Hz. The theoretical results are validated by experimental data from STW resonators operating in the low microwave region. Measurements of an optimized 1.0 GHz STW single-port device confirm that the material Q limit for STW is approximately $Q \times f \approx 3.6 \times 10^{13}$ Hz.

2. THEORETICAL ANALYSIS

The theoretical analysis can be accomplished in at least two ways; one way employs the acoustic viscosities of quartz and complex elastic stiffnesses in the well known STW propagation equations. This approach is useful in understanding the impact of waveguiding on the overall device Q. However, a substantially simpler approach can be used to evaluate the limiting behavior of the STW. This alternative approach recognizes that the STW is essentially a waveguided pure-shear SSBW mode, for which exact analytic expressions may be written for rotated Y-cut quartz:

$$\left(\varepsilon_{22}^{S}\right)' = \varepsilon_{11}^{S}\cos^2(\theta) + \varepsilon_{33}^{S}\sin^2(\theta) \tag{1}$$

$$(e_{26})' = -e_{11}\cos^2(\theta) - e_{14}\cos(\theta)\sin(\theta) \tag{2}$$

$$\left(c_{66}^{E}\right)' = c_{66}^{E}\cos^2(\theta) + 2c_{14}^{E}\cos(\theta)\sin(\theta) + c_{44}^{E}\sin^2(\theta) \tag{3}$$

$$(\bar{c}_{66})' = \left(c_{66}^{E}\right)' + \frac{\left[(e_{26})'\right]^2}{\left(\varepsilon_{22}^{S}\right)'} \tag{4}$$

$$v = \sqrt{\frac{(\bar{c}_{66})'}{\rho}} \tag{5}$$

$$k = \sqrt{\frac{\left[(e_{26})'\right]^2}{\left(\varepsilon_{22}^{S}\right)'(\bar{c}_{66})'}} \tag{6}$$

$$(\eta_{66})' = \eta_{66}\cos^2(\theta) + 2\eta_{14}\cos(\theta)\sin(\theta) + \eta_{44}\sin^2(\theta) \tag{7}$$

$$\tau = \frac{(\eta_{66})'}{(\bar{c}_{66})'} \tag{8}$$

$$Q \times f = \frac{1}{2\pi\tau} \tag{9}$$

where c, e, ε, ρ, and η represent the elastic, piezoelectric, and dielectric constants, density, and acoustic viscosity.

0-7803-5400-1/99/ $ 10.00

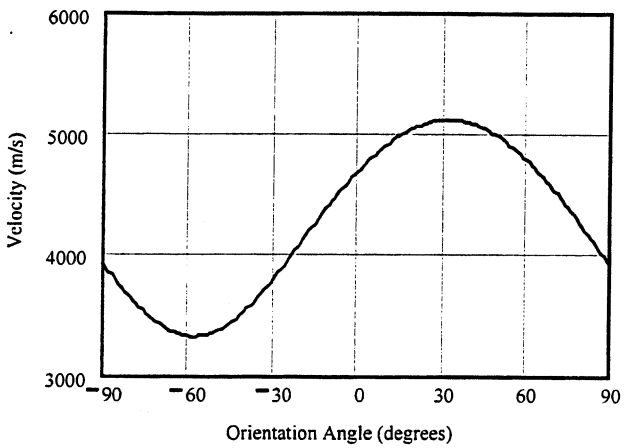


Figure 1. STW velocity versus orientation angle for rotated Y-cut quartz substrates.

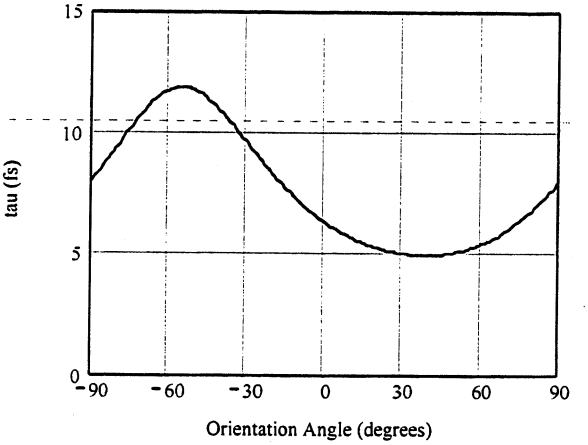


Figure 3. STW motional time constant versus orientation angle for rotated Y-cut quartz substrates.

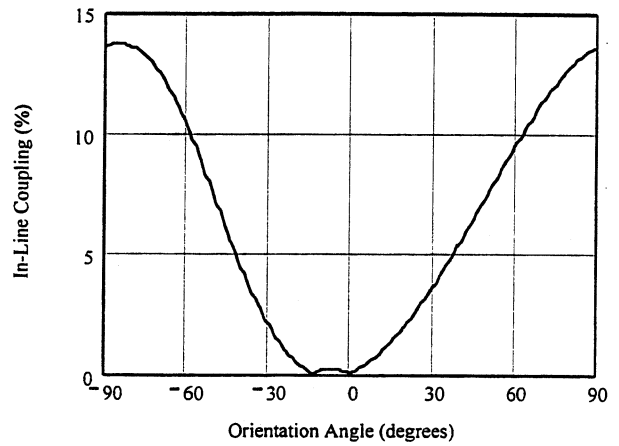


Figure 2. STW in-line piezoelectric coupling versus orientation angle for rotated Y-cut quartz substrates.

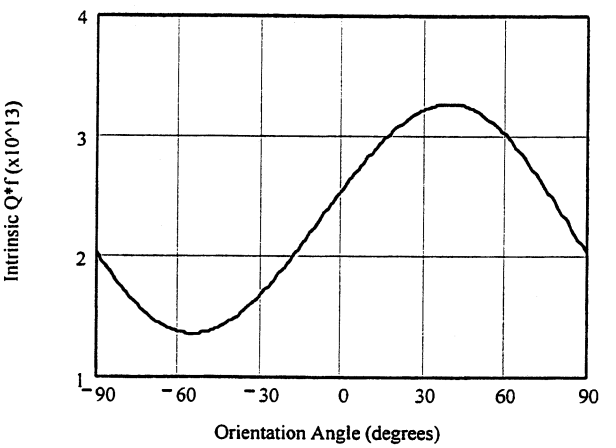


Figure 4. STW Q×f product versus orientation angle for rotated Y-cut quartz substrates.

Figures 1 through 4 show, respectively, the SSBW velocity (5), in-line piezoelectric coupling (6), motional time constant (8), and Q×f product (9) for rotated Y-cut quartz. The calculations indicate that the intrinsic Q of STW devices on quasi- AT-cut quartz is two to three times larger than those of the other technologies, with the STW having $Q \times f \approx 3.25 \times 10^{13}$ Hz. This value is equivalent to that of BT-cut BAW devices, and is in turn superior to that of ST-cut SAW devices and AT-cut BAW devices. The $1/Q^4$ relationship between device Q and close-in phase noise implies a potential 15 dB advantage in close-in phase noise for low noise oscillator applications. It is important to note that the STW device Q varies with frequency, and that the intrinsic Q limit is approached when operating the device below the stopband [3].

3. EXPERIMENTAL VERIFICATION

3.1 Two-Port STW Devices on AT-cut Quartz

The data in Figure 5 were obtained experimentally from a 2 GHz in-line coupled resonator filter (CRF) test structure. The coupling grating intentionally has been made 600 wavelengths long to maximize the cavity length. Since this long coupling grating has the same period as the interdigital transducers (IDTs) and reflectors, it causes strong reflections along the cavity and a very abrupt cut-off behavior. Close to the cut-off frequency, the group delay value is extremely large: 1.932 μs. This corresponds to a loaded Q of 12200 at a frequency of 2008 MHz, which is 81% higher than the material Q for AT-cut BAW. Please note that this type of CRF

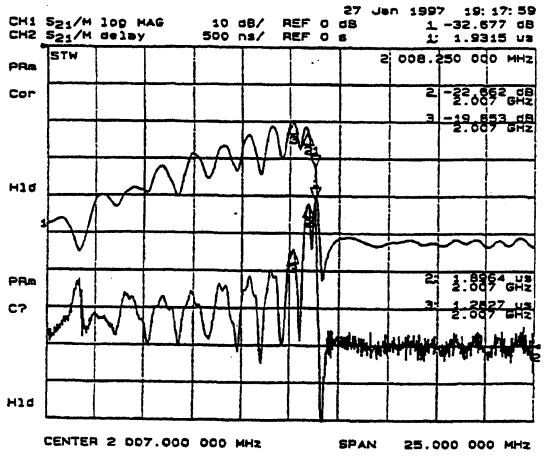


Figure 5. Frequency response (upper curve) and group delay response (lower curve) of a CRF-type test structure with a very long coupling grating.

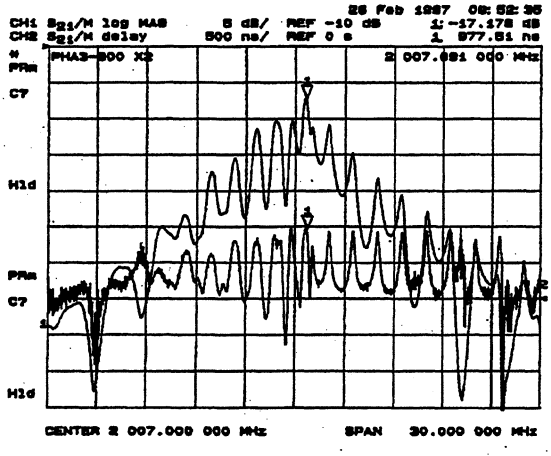


Figure 6. Frequency and group delay responses of a 2 GHz multimode STW resonator.

device was intended to investigate the achievable group delay for STW at 2 GHz. It cannot be used for practical applications since there is no well behaved resonance at the frequency of maximum group delay. If an STW device is to be used in an oscillator with maximized Q, the resonator geometry should be optimized such that it supports one resonance with a magnitude at least 5 dB higher than the adjacent resonances to guarantee stable single-mode oscillator operation. As shown in [4], very high loaded Q values can be achieved with optimized multimode devices which have a long quasitransparent waveguide grating between the IDT to minimize the device loss while keeping the cavity length as long as possible. One such device is characterized in Figure 6. With a group delay value of 978 ns and an insertion loss of 17.2 dB, this 2 GHz device has a loaded Q of 6170 and an unloaded Q of 7160, which is 37% higher than the material Q limit for ST-cut SAW at that frequency.

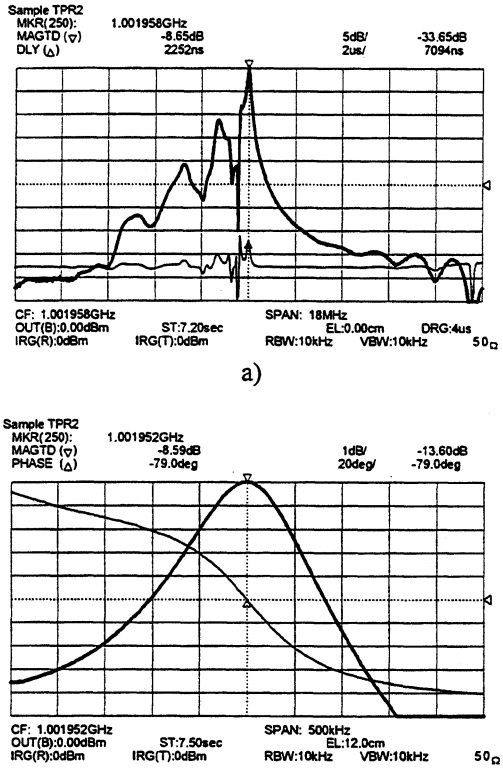


Figure 7. Broadband frequency and group delay responses (a) and narrowband frequency and phase responses (b) of a 1 GHz short-cavity STW resonator.

In the frequency range of 1.0 to 2.0 GHz, unloaded Q values as high as the material Q limit for AT-cut BAW can readily be achieved with practical short-cavity STW devices occupying a very small substrate area. As shown in [4], the Q of such devices is maximized by having two adjacent longitudinal modes coincide at the lower stopband edge. In addition to high loaded Q, this approach also provides low device loss. One such short-cavity device is characterized in Figure 7. The 3 dB device bandwidth is 122 kHz which corresponds to a loaded Q of 8210. With an insertion loss value of 8.6 dB, the unloaded Q is found to be 13080 which is 97% of the material Q limit for AT-cut BAW at this frequency. Similar results also were obtained with 2.0 and 2.5 GHz short-cavity devices [4].

3.2 Single-Port STW Devices on AT-cut Quartz

The two main limiting factors to achieving Q values close to the material limit for STW in two-port resonators, are 1) the ohmic loss of the IDT fingers and 2) the dissipation of wave energy in the bulk when the wave propagates between the IDTs. The effect of both factors is substantially reduced in a single-port resonator configuration. Therefore, single-port STW resonators

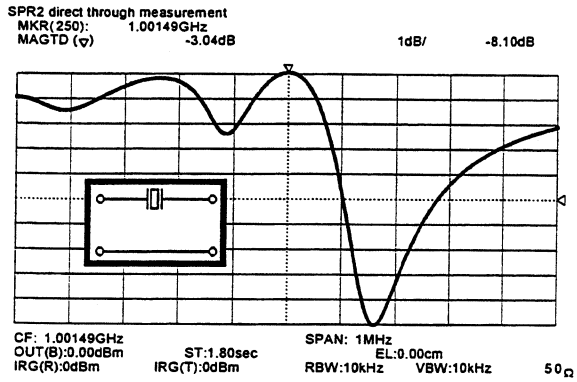


Figure 8. Frequency response of a 1 GHz single-port STW resonator obtained in a through measurement.

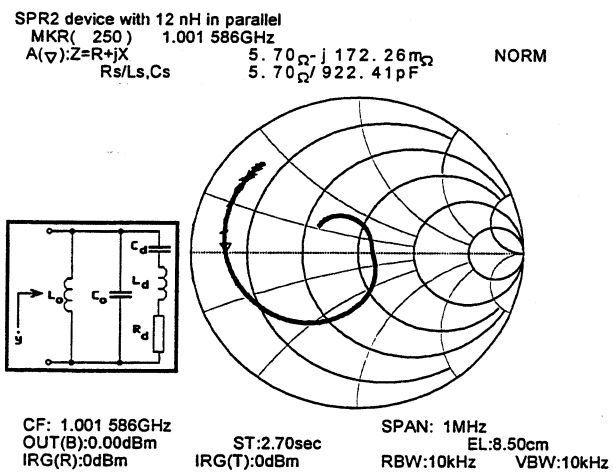


Figure 9. Impedance of a 1 GHz single-port device with the static IDT capacitance C_o tuned out with a 12 nH parallel inductor.

should have higher device Q than two-port devices.

We analyzed experimental data from single-port devices characterized by the method described in [5] that allows the extraction of the equivalent circuit parameters from transmission and reflection measurements. Figure 8 shows the frequency response of a 1 GHz single-port device obtained in a through measurement. The distance between series and parallel resonance, corresponding to the frequencies of minimum and maximum attenuation, respectively, is 160 kHz, yielding a capacitance ratio of 3130. With a static IDT capacitance C_o value of 2.1 pF, this yields a dynamic capacitance C_d of 0.671 fF.

The value of the dynamic resistance R_d is obtained from a reflection measurement, after tuning out C_o with a parallel inductor [5]. The result of this measurement is shown in Figure 9 and indicates a very low dynamic resistance value of 5.7Ω at the frequency at which the complex impedance locus crosses the real axes of the Smith chart. The R_d and C_d values obtained in this way imply a Q value of 41550, and a Q×f product well in excess of 3.25×10^{13} Hz.

4. CONCLUSIONS

Theoretical and experimental evidence demonstrate that the intrinsic Q of STW devices on quasi- AT-cut quartz is superior to that of ST-cut SAW devices and AT-cut BAW devices. The calculations indicate that the intrinsic Q of the STW mode is two to three times larger than those of the other technologies, with the STW having a theoretical $Q\times f\approx3.25\times10^{13}$ Hz. The intrinsic Q limit is approached when operating the device below the stopband. The theoretical results have been validated by experimental data from STW resonators operating in the low microwave region, including an optimized 1.0 GHz STW single-port device. The $1/Q^4$ relationship between device Q and close-in phase noise implies a potential 15 dB advantage in close-in phase noise for low noise oscillator applications.

5. ACKNOWLEDGMENTS

The authors wish to gratefully acknowledge Fujitsu Laboratories in Akashi, Japan for the opportunity to fabricate the 2 GHz STW devices at those laboratories.

6. REFERENCES

[1] A. J. Budreau and P. H. Carr, "Temperature Dependence of the Attenuation of Microwave Frequency Elastic Surface Waves in Quartz," Applied Physics Letters, vol. 18, pp. 239-241, 1971.

[2] A. Ballato, "Doubly Rotated Thickness Mode Plate Vibrators," in Physical Acoustics Vol. XIII, New York: Academic Press, 1977, Chapter 5, p. 145.

[3] D. F. Thompson and B. A. Auld, "Surface Transverse Wave Propagation Under Metal Strip Gratings," in Ultrasonics Symposium Proceedings, 1986, pp. 261-266.

[4] I. D. Avramov, O. Ikata, T. Matsuda, T. Nishihara and Y. Satoh, "Further Improvements of Surface Transverse Wave Two-port Resonator Performance in the 2.0 to 2.5 GHz Range", in Proceedings of the 1997 IEEE International Frequency Control Symposium, May 1997, pp. 807-815.

[5] I. D. Avramov, "Increasing the Loaded Q of GHz Range Surface Transverse Wave Resonators", in Proceedings of the 1997 European Frequency and Time Forum, March 1997, pp. 394-398.

SENSITIVITY OF STW RESONATORS TO RADIAL IN-PLANE STRESS EFFECTS : THEORY AND EXPERIMENTS

S. Ballandras, W. Daniau, E. Henry-Briot, G. Marianneau, G.Martin

Laboratoire de Physique et Métrologie des Oscillateurs du CNRS
associé à l'Université de Franche-Comté
32 Avenue de l'Observatoire - 25044 Besançon Cedex - France
ballandr@lpmo-univ.fcomte.fr

ABSTRACT

The design and fabrication of high spectral purity resonators based on surface waves requires a good understanding of the influence of temperature and mechanical stress effects on their frequency stability. The aim of these paper is to present a perturbation model of Surface Transverse Waves (STW) sensitivity to stresses used to predict the frequency variations of STW resonators built on AT cut quartz plates submitted to radial compression. Comparison between theory and experiments is shown to validate the model and to check the stability of STW properties to the propagation direction of the wave.

I. INTRODUCTION

Surface Transverse Waves (STW) have been studied for more than twenty years for the development of high quality surface wave devices used in ultra-stable oscillators or high frequency filters. However, industrial applications of STW remain rare because of the difficulty to obtain reproducible electrical characteristics with these waves. Nevertheless, very exciting results have been published by Avramov *& al* [1,2] demonstrating the possibility to produce STW resonators exhibiting extremely high unloaded quality factors Q_u at frequencies higher than 1 GHz ($Q_u.F_0$ up to 1.2 10^{13} at 2 GHz). Consequently, the use of STW resonators for the design and fabrication of high spectral purity oscillators still represents an attractive alternative compared to the best classical Rayleigh wave based devices [3]. It is then interesting to check the sensitivity of STW to mechanical perturbations to optimize their packaging or to point out any substrate geometry well adapted to minimize stress effects on their frequency stability. In this purpose, developments have been proposed by Kosinski *& al* [4] and also by Bigler *& al* [5] based on Tiersten's perturbation theory [6] to calculate STW sensitivity to mechanical biases. Two different approaches respectively based on strain [4] and stress [5] sensitivity coefficients have lead to the conclusion that a compensation to symmetrical in-plane stresses exists for a quartz orientation close to the AT cut.

The purpose of this work is to check from both theoretical and experimental points of view the validity of these result and to emphasize the influence of any disorientation of the propagation direction on the compensation effect. In the first part of the paper, general principles of theoretical developments necessary to calculate stress sensitivity coefficients of STW are recalled. It is then shown how to couple the proposed perturbation approach with Finite Element (FE) calculations to improve theoretical predictions. The latter are used to confirm the compensation effect and to refine the angle definition of the stress compensated STW quartz cut. These theoretical results are finally compared to experimental data obtained using 200 MHz STW resonators built on circular AT cut quartz plates. Measurements have been performed for different propagation directions (ψ = 90 ± 0, 2.5 and 5 degrees defined versus X axis). As a conclusion, the validity of the classical criterion for systematically pointing out "stress compensated" STW quartz cuts is discussed

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

II. PERTURBATION THEORY

II.1 Stress sensitivity coefficients for STW

The calculation of these coefficients has been presented first in ref. [5]. It is based on a classical unperturbed propagation model [7] coupled with a perturbation approach inspired from Tiersten's developments [6]. Let us briefly recall the calculation principle. The mechanical wave distribution is given by the following relation :

$$u_1 = \sum_{n=-\infty}^{+\infty}\sum_{r=1}^{2} A_{nr} e^{-\alpha_{nr} a_2} e^{-j\beta_n a_3} e^{j\omega t} \quad (1)$$

where β_n is the wavenumber of the nth spatial harmonic, α_{nr} are the penetration coefficients within the plate of the rth partial wave of the nth spatial harmonic and A_{nr} represents the associated mechanical amplitude. The polarization of the wave is conform to fig.1.

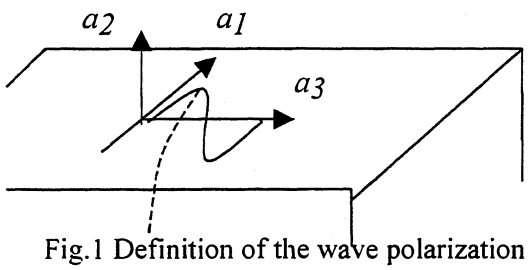


Fig.1 Definition of the wave polarization

All the coefficients of eq. (1) are calculated as shown in ref. [7]. This form is then used to compute the following perturbation equation :

$$\frac{\Delta\omega}{\omega_0} = \frac{\iiint_\Omega \frac{\partial u_i^*}{\partial a_j} \bar{H}_{ijkl} \frac{\partial u_l}{\partial a_k} d\Omega}{2\rho_0\omega_0^2 \iiint_\Omega u_m^* u_m d\Omega} \quad (2)$$

where all the dynamical displacements are assumed unperturbed and only ω is the effective angular frequency variation induced by the perturbation. $\bar{H}ijkl$ is the perturbation tensor relating the fundamental elastic properties of the substrate to the actual quasi-static stress of the plate. Assuming no rigid rotation around a fixed point of the plate (for example $a_1=a_2=a_3=0$) yields the following expression of this tensor :

$$\bar{H}_{ijkl} = (\delta_{ij}\delta_{kl}\delta_{tv} + C_{ijklst}S_{stuv} + C_{ijkt}S_{tluv} + C_{ijtl}S_{ktuv})\bar{T}_{uv} \quad (3)$$

with the Kronecker delta δ_{ij}, the linear and non linear elastic constants C_{ijkl} and C_{ijkluv} and the static stresses $\bar{T}_{uv}$. Inserting eq.(1) and eq.(3) into eq.(2) allows the derivation of the STW stress sensitivity coefficients ${}^s\alpha_{uv}$ as follows:

$${}^s\alpha_{uv} = \frac{1}{2\rho_0\omega_0^2 N} \sum_{n,m=-\infty}^{+\infty}\sum_{r,s=1}^{2} \frac{A_{nr}^* A_{ms}}{\alpha_{nr}^* + \alpha_{ms}} \times (\alpha_{nr}^*\alpha_{ms}\Pi_{1212uv} + \alpha_{nr}^*\beta_m\Pi_{1213uv} + \beta_n^*\alpha_{ms}\Pi_{1312uv} + \beta_n^*\beta_m\Pi_{1313uv}) \quad (4)$$

where N is the numerator of eq.(2), easily deduced from the previous developments and eq.(1). Π_{ijkluv} is simply derived from eq.(3).

II.2 Coupling with finite element

These stress sensitivity coefficients can be used with simple academic stress distribution models but as it has been shown for the case of Rayleigh waves [8], they can also advantageously be coupled with FE calculations to provide precise simulation of surface wave sensitivity to any stress distribution taking anisotropy into account. The frequency variation is then calculated using the following formula :

$$\frac{\Delta\omega}{\omega_0} = \frac{\sum_{e=1}^{E} {}^s\alpha_{uv} <\bar{T}_{uv}(e)> V(e)}{V_{tot}} \quad (5)$$

where E represents the total number of elements used to mesh the STW propagation domain, and $<\bar{T}_{uv}(e)>$ is the stress value at the center of mass of the element e. V_{tot} is the total volume of the above-mentioned domain. From eq.(1) and (2) and considering formal developments of ref. [8], a more rigorous calculation of the frequency variations can be derived. However, since no significant improvement of the theoretical results in comparison with experiments was noticeable with such developments in the case of Rayleigh waves, the simple formula of eq.(5) has been preferred for this work.

III. THEORY VERSUS EXPERIMENTS

The proposed stress sensitivity coefficients have been used first to check the existence of any " stress compensated " quartz cut by computing the term ${}^{s}\alpha_{11}+{}^{s}\alpha_{33}$ which provides a good estimation of the wave sensitivity to symmetrical radial in-plane stresses for Rayleigh waves (see ref. [9]). Figure 2 shows the dependence of this criterion versus cut angle θ and fig.3 is an enlargement of the latter in the vicinity of AT cut, providing a better precision of the angle for which the compensation arises.

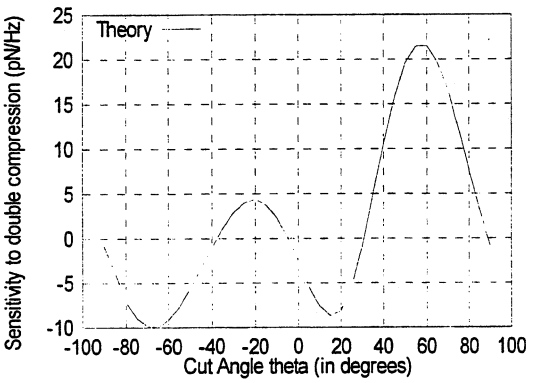


Fig. 2 Sensitivity of STW on quartz to symmetrical compression vs cut angle θ

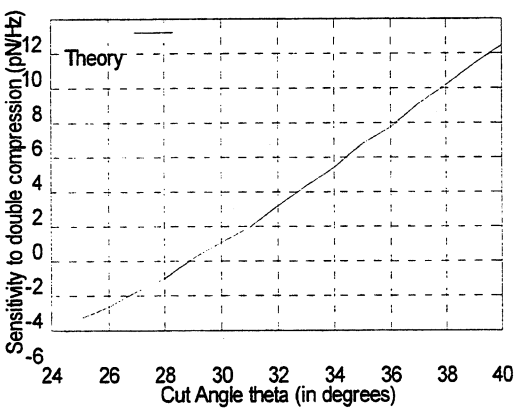


Fig.3 Enlargement of fig. 2 in the vicinity of AT cut

These figures show that different orientations exist for which the STW is small affected by symmetrical compression, and that the STW sensitivity to this effect is rather small close to the AT cut (theoretical compensation at θ=30°). Considering these results, this sensitivity has been experimentally verified by cutting small disks of diameter 21 mm. (1 mm. thick) in an AT cut quartz wafer (Micro-Pierre SA, Besançon) used to fabricate 200 MHz STW resonators. Figure 4 shows the frequency response of a classical STW resonators aligned along Z', and of resonators built with an angle of 2.5° and 5° between the propagation direction and Z'

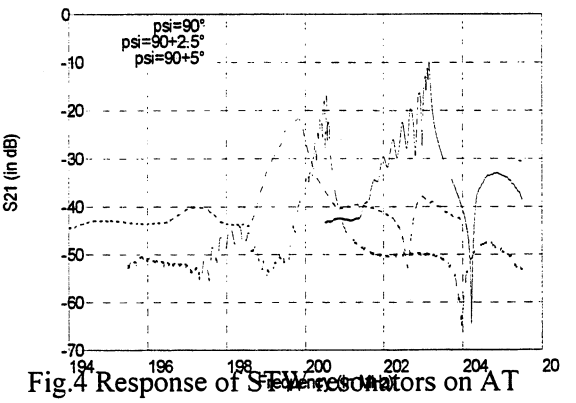


Fig.4 Response of STW resonators on AT quartz disks with $\psi = 90 \pm 0, 2.5°, 5°$

These devices are then submitted to symmetrical radial forces as shown in fig.5

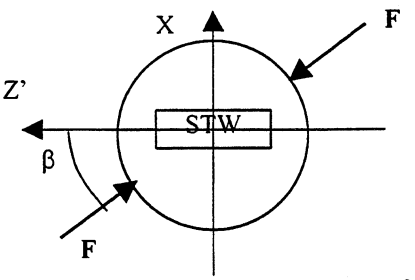
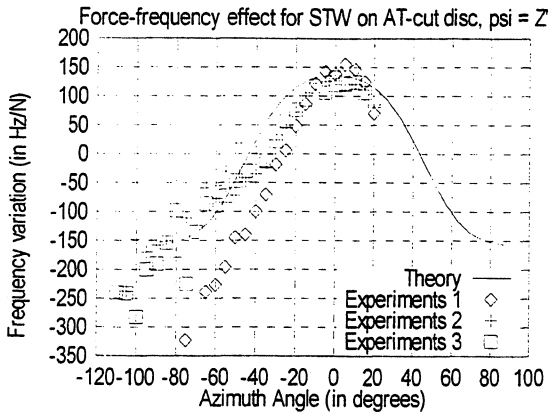
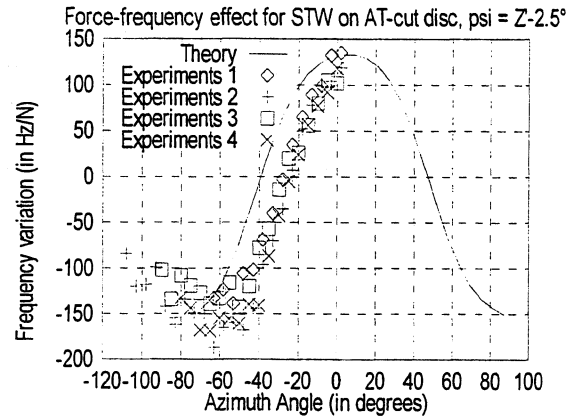


Fig.5 Scheme of the experimental test of STW sensitivity to compression

Figure 5 shows the definition of the azimuth angle β between the propagation and the force directions. Since the coefficient ${}^{s}\alpha_{13}$ is equal to zero for STW, the frequency variation versus β is symmetrical around Z', allowing to measure Δω only for $\pm 90° < \beta < 0°$. Figures 6 (a and b) show the superposition of experimental measurements and theoretical results for the different resonators of fig. 5, except for the ones at $\psi = \pm 5°$ (insertion losses were to important) :



(a)



(b)

Fig. 6 Comparison between theoretical and experimental STW sensitivity to compression.
(a) $\psi = 90°$ (b) $\psi = 87.5°$

It is interesting to point out the fact that the sensitivity of both devices are very close in term of magnitude, but the dependence versus β is influenced by the propagation direction. This phenomenon is due to the loss of symmetry when propagating not exactly along Z'. Moreover, it can be shown by computing the effective permittivity that the wave propagating along Z' $\pm$ ε with ε small but different than 0 exhibit a PSAW like behavior rather than a Bleustein one. Consequently, the proposed model consisting in applying STW calculations to wave propagation not exactly along Z' for singly rotated quartz cuts masks some very fundamental properties of the actual waves, which may explain the discrepancies between theory and experiments. Whatever, it must be underlined that the mean values of both curves of fig. 6 are negative as suggested by fig. 3 (see ref. [9]), which signifies that the criterion ${}^{s}\alpha_{11}+{}^{s}\alpha_{33}$ gives a precise estimation of compressive stress compensated cuts. However, The use of FE based perturbation models is necessary for accurate predictions of stress sensitivity of STW on quartz cuts.

IV. CONCLUSION

A theoretical model of STW sensitivity to stress effects has been implemented and tested by a comparison with experiments performed on STW resonators built on AT cut quartz disks submitted to radial compression. The influence of the direction propagation has been emphasized. The principal conclusion of this work is that the model can be used to estimate properties of wave propagating not exactly along Z', but a more rigorous model of the effective nature of the wave must be used to improve the agreement between theory and experiments.

REFERENCES

[1] I.D. Avramov, M. Suohai, IEEE Trans on UFFC, vol.43, n° 6, pp1133-1135 (1996)
[2] J.A. Kosisnki, R. Pastore, I.D. Avramov, these proceedings, EFTF-IFCS 1999
[3] G.E. Montress, T.E. Parker, M.J. Laboda, Ultr. Symp., IEEE cat 87CH2492-7, pp 47-52
[4] J.A. Kosinski, J.T. Stewart, A. Ballato, R.Almar, IFCS, IEEE cat 95CH35752, pp 486-493 (1995)
[5] E. Bigler, S. Ballandras, IFCS, IEEE cat 96CH35935, pp 422-429 (1996)
[6] H.F. Tiersten, J. Acoust. Soc. Am. 64, 832 (1978)
[7] E. Gavignet, S. Ballandras, E. Bigler, J. Appl. Phys., vol 77, n°12, pp 6228-33 (1995)
[8] S. Ballandras, E. Bigler, IEEE Trans on UFFC, vol 45,n°3 , pp.567-573 (1998)
[9] S. Ballandras, E. Bigler, J. Appl. Phys., vol. 72, n°8, pp. 3272-81 (1992)

1999 Joint Meeting EFTF - IEEE IFCS

DESIGN OF STW RESONATORS ON LANGASITE USING EXPERIMENTALLY DETERMINED COM PARAMETERS

I. S. Mitrofanov*, A. V. Perevalov*, E. Bigler**

*Avangard-Elionica Co., 72 Kondratyevsky av., 195271 St. Petersburg, Russia

**LPMO/CNRS, 32 avenue de l'Observatoire, 25044 BESANCON CEDEX, France

ABSTRACT

Langasite and its modifications, such as langanite and langatate, are very attractive materials for realization of resonators. The first publications about piezoelectric properties of langasite appeared at the beginning of the 80's [1,2]. This material has some advantages over well-known ST-quartz: a higher value of electromechanical coupling, lower acoustic wave propagation losses, absence of phase transitions up to the melting point (1470°C) and absence of twins [3]. However, till now the attempts of design of competitive devices on langasite have not been successful. The imperfection of growing was resulted in a significant dispersion of the physical constants. Good crystals of langasite, possessing stable properties, were grown only a few years ago [4].

This paper presents the design of the Surface Transverse Wave (STW) resonators on langasite. The empirical values of coupling-of-modes (COM) parameters are shown, and a comparison between the theoretical resonator responses computed with these values and measured responses is performed. The resonance and antiresonance frequencies are evaluated as a function of the thickness of metal electrodes. Results for STW resonator response on langasite exhibit very important differences with respect to similar STW devices built on ST and AT quartz.

Finally, temperature sensitivity of the STW langasite resonators has been measured and compared to theoretical predictions based on various sets of published material constants.

1. INTRODUCTION

There is a high interest in studying properties of new piezoelectric materials capable to improve device performance in comparison with conventional ones based on quartz.

This paper presents investigation on resonant STW structures fabricated with submicrometer accuracy, on Y+90° rotated langasite substrates. A tunable design procedure based on high frequency adapted COM formalism is used.

Due to of a lack of references in the literature about COM parameters for acoustoelectronic devices on langasite substrates, as well as the strong dependence of resonator parameters on metal thickness and metallization ratio, an experimental procedure we used to extract COM parameters from several samples of STW resonators with a variation of thickness and metallization ratio, taking into account of their electrical resistance. Based on these data, a simulation procedure yields a model for the electrical response curve of a bandpass filter which is compared to experimental results.

2. DESCRIPTION OF TEST STRUCTURE

The choice of a test structure for the STW langasite resonator was based on a design of a two-port STW resonator previously tested on quartz [5]. The topology of such a resonator structure is presented on Fig.1, main geometry parameters being presented in Table 1.

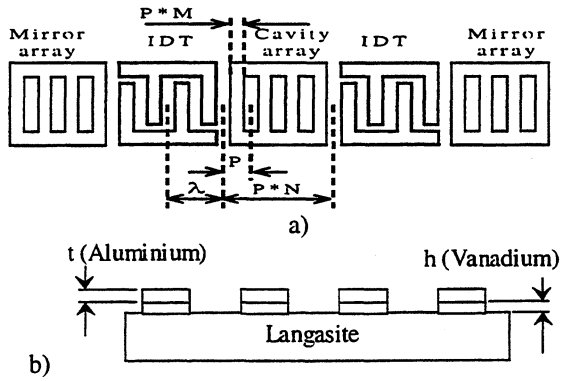


Fig. 1. Layout (a) and crossing (b) of the regular strip grating on the langasite substrate.

TABLE 1 – Geometry of STW langasite resonator.

Strip number N	1000 in each mirror
	150 in each IDTs
	25 in cavity array
Grating (strip) pitch p [μm]	1.69 in mirror arrays and IDTs
	1.67 in cavity array
Bragg wavelength λ	2*p
Waveguide aperture W/λ	90
Metallization ratio M	0.45 in sample #1
	0.25 in sample #2
Aluminium thickness t/λ [%]	0.65 in sample #1
	0.53 in sample #2

3. COM SIMULATION

Metallization of interdigital transdusers (IDT) and shorted strip reflective arrays (Fig.1b) is made of an aluminium conductive layer with vanadium as an adhesive sublayer. The electrical conductivity of vanadium is about ten times lower for vanadium than for aluminium. This configuration can be modeled according to Morgan's relation for phase velocity differences between a metallized (v_m) and free surface of the piezoelectric substrate (v_0) [6]:

$$\frac{V_m^h - V_0^h}{V_m - V_0} \approx \left[1 - \frac{1+\varepsilon_p^T/\varepsilon_0}{1-cth(2\pi h/\lambda)}\right]^{-1} \quad (1)$$

where superscript h indicates the shifted velocities induced by a conductive aluminium layer of height h deposited over the langasite surface, ε_0 is the vacuum dielectric constant, and effective dielectric permittivity

$$\varepsilon_P^T = \sqrt{\varepsilon_{11}^T \varepsilon_{33}^T - \varepsilon_{13}^T \varepsilon_{31}^T} \quad (2)$$

is computed from a combination of dielectric permittivity constants at constant mechanical stress.
In the simulation of a resonator structure, the COM model is represented as a two-port scattering and admittance matrices taking into account electrical loading by the test system (Fig.2).

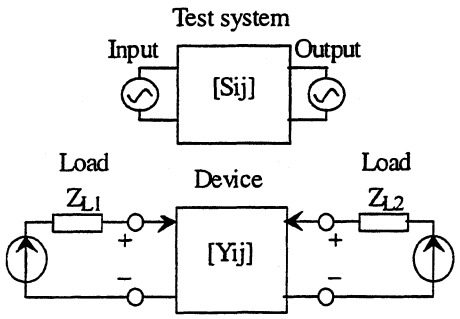


Fig. 2. Test system chart.

Standard scattering parameter S_{21} as forward transmission coefficient for the testing system is

$$S_{21} = \frac{-2\sqrt{R_{L1}R_{L2}}Y_{21}}{(1+Z_{L1}Y_{11})(1+Z_{L2}Y_{22}) - Y_{12}Y_{21}Z_{L1}Z_{L2}} \quad (3)$$

derived from resulting 2-port matrix of admittance parameters $[Y_{ij}]$ for the complete device, taking into account both input and output load impedances Z_{L1} and Z_{L2} respectivitely with their real parts R_{L1}, R_{L2}.
P-matrix parameters can be converted into Y-matrix and its cascade combinations for each constructive elements as IDT and grating arrays are defined on the base of the particular topology and design data for resonator [7,8]. Known COM formalism was extended to high frequency structures, where thin film electrode resistance, pin inductance, case and inter-electrode capacitance are not negligible.
Input, output, and ground connections in the resonator device form a series inductance and a parallel capacitance. The transducer comb structure forms a series electrode resistance and inter-electrode capacitance. All these components define the equivalent circuit (Fig.3), where the series static parameters of the resonance structure were considered as outer load parameters. The static capacitance C_0 can be experimentally evaluated by input capacitance measurements far from resonance frequency.

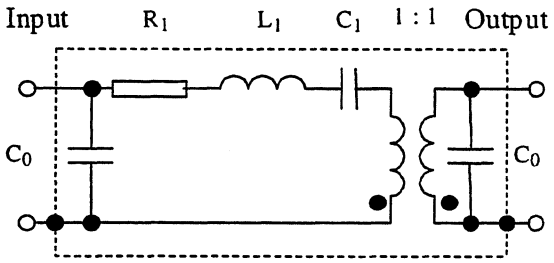


Fig.3. Equivalent circuit for two-port STW resonator on langasite in the vicinity of the main resonance frequency.

Coupling-of-modes formalism analyzes approximately synchronous reflecting and propagating waves in the structure, and yields an estimation of the equivalent circuit parameters in the vicinity of the resonance frequency. This were taken into account when using an experimental method to extract these parameters. [9].

4. EXPERIMENTAL RESULTS

Several sets of STW langasite resonators were designed and tested.
For COM parameter extraction S_{21}, frequency responses of resonators with two different aluminium thickness and metallization ratios were measured at room temperature (Table 2). For simulation purpose we used untuned measurements in a 50 Ω testing system.

TABLE 2 – Two-port langasite STW resonators measured and estimated characteristics.

Parameter	#1	#2
Resonance frequency Fr [MHz]	688.930	688.975
Antiresonance frequency Fa [MHz]	689.823	689.565
Relative resonance spacing ΔFar=(Fa-Fr)/Fr [%]	0.130	0.086
Attenuation at resonance Ar [dB]	7.2	18.2
Loaded quality factor Ql	1330	1180
Unloaded quality factor Qu	2350	1345
Dynamic resistance R_1 [Ω]	130	715
Dynamic inductance L_1 [μH]	70.5	222
Dynamic capacitance C_1 [fF]	0.75	0.24

Experimentally determined main COM parameters are shown in Table 3 with using approximation (4) – (6) [9]. The grating reflectivity upon metallization ratio M and normalized aluminium thickness t/λ is given by
Reflection coefficient :

$$\kappa = \kappa_0 * \sin(\pi * M) * (t/\lambda)^2, \quad (4)$$

electromechanical coupling

$$K^2 = K_0^2 * M * (t/\lambda), \quad (5)$$

STW velocity

$$V = V_0(1 - \Delta_V * K). \quad (6)$$

TABLE 3 – Two-port langasite STW resonators estimated COM characteristics.

Parameter	Value
Reflection coefficient κ_0	120
Electromechanical coupling coefficient K_0^2	0.005
Propagation loss coefficient γ_0 (constant)	0.0008
Free-surface velocity V_0 (m/s)	2332
STW velocity coefficient Δ_V	0.05

Measured and calculated frequency responses of STW langasite resonator with contrast to quartz at close geometry of metallization are shown in Fig.4. Here the referenced resonance frequency value for langasite is 688.930 MHz and for quartz is 1499.875 MHz.

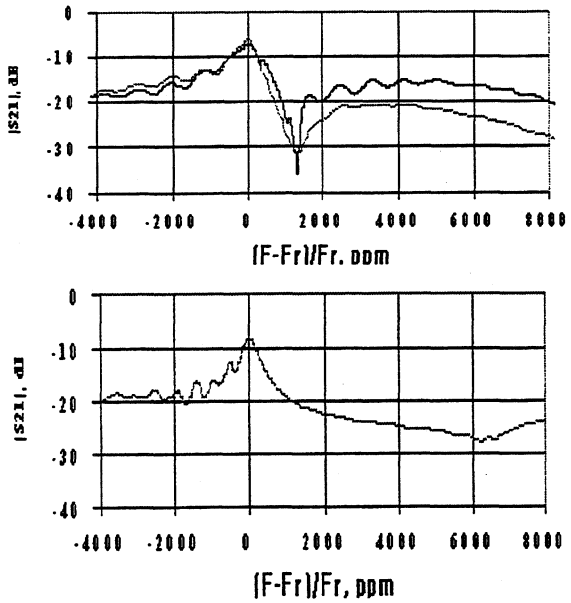


Fig. 4: Measured (solid) and calculated (dashed) frequency response of STW langasite #1 (upper) and quartz (lower) resonator with close geometry for comparison.

Based on experimentally determined COM parameters a set of simulated frequency responses for STW langasite resonator structure with geometry according Fig.1 and Table 1 for metallization ratio 0.5 and aluminium thickness t/λ at (0.6...2.4) % is presented in Fig.5.

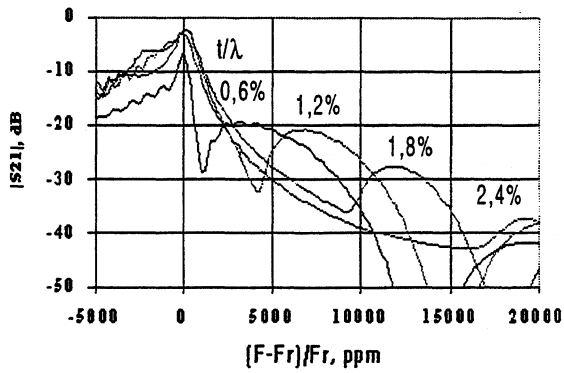


Figure 5: STW langasite resonator resonance spacing changes with metal thickness.

Temperature characteristics for Y+90° rotated langasite were determined. Velocities and temperature effects for the STW structure were simulated by computing the velocity changes versus temperature for a Gulyaev-Bleustein mode propagating under a short-circuited surface with no metal loading. Taking into account the thermal expansion of the substrate in the propagation direction, the first order temperature coefficient has been computed in the vicinity of Euler angles (λ=0, μ=23, θ=90) as a function of the plate cut angle μ.Based on material constants published by Silvestrova et al. [9] and Sakharov et al. [10], a temperature-compensated cut for STWs is expected to be found in the vicinity of Euler angles (0, 23, 90), with a velocity close to 3000 m/s ; the optimum (TCF1=0) being at μ=24 according to [9] or μ=26 according to [10]. Both simulations predict a dependence vs the angle μ for the first-order TCF1 of about -20 ppm/K change every +15 degrees of turning angle μ. Experimental measurements on a STW resonator presented here yields a velocity close to 2330 m/s and positive TCF1 of about +50 ppm/K not compatible with Euler angles (0, 23, 90) but in full accordance with (0,-23, 90) according to the same convention

for the sign of axes in refs [11] and [12], since for this last cut the predicted velocity is close to 2400 m/s with a positive TCF1=+36 ppm/K.

5. CONCLUSION

Experimentally determined COM parameters for STW langasite resonator models were obtained by investigating various samples with several thickness and ratio of electrode metallization. A fairly good agreement between simulated and experimental data were obtained.

In comparison to Y+36° rotated quartz the Y+90° rotated langasite device does not exhibit a frequency shift between resonance and antiresonance peaks proportional to velocity differences between both crystals

6.REFERENCES

[1] B.V. Mill, I.M. Silvestrova, A.A. Kaminski et al, "Piezoelectric, dielectrical, elastical and acoustic properties of La3Ga5SiO14", in Proceedings of the X All Union Acoustic Conference, 1983, pp. 20-29.

[2] I.M. Silvestrova, Yu.V. Pisarevsky, P.A. Senyushenkov, A.I. Krupny, "Temperature dependence of the properties La3Ga5SiO14 single crystalls", Sov. Phys. Solid. Stat., 1986, V28, pp. 1613-1614.

[3] G.D. Mansfeld, "Langasite as a material for piezoelectric devices", in Proceedings of the 12th EFTF, 1998, pp. 61-65.

[4] S.A. Sakharov, A.V. Medvedev, O.A. Buzanov, "HF langasite monolithic filters for GSM Standard", in Proceedings of the 11th EFTF, 1997, pp. 239-242.

[5] A.V. Grouzdev, S.I. Kort, I.S. Mitrofanov, A.V. Perevalov, S. Gawor, A. Milewski, "GHz Range STW Resonators on Quartz and Langasite", in Proceedings of the 12th EFTF, 1998, pp. 332-335.

[6] D.P. Morgan, "Surface Wave-Devices for Signal Processing", New York: ELSEVIER Science Publisher, 1985.

[7] C.S. Hartmann, D.P Chen and J. Heighvay, "Modelling of SAW Transversely Coupled Resonators Filters using Coupling-of-Mode Modeling Technique", in Proceedings of the IEEE Ultrasonic Symposium, 1992, pp. 39-43.

[8] B.P. Abbott and Ken-ya Hashimoto, "A Coupling-of-Modes Formalism for Surface Transverse Wave Devices", in Proceedings of the IEEE Ultrasonic Symp., 1995, pp. 239-245.

[9] S. Krasnikova, B.P. Abbott, R.C. Almar, "COM parameter extraction for STW resonator design", in Proceedings of the 11th EFTF, 1997, pp. 440-444.

[10] I.D. Avramov, "Gigahertz Range Resonant Devices for Oscillator Applications Using Shear Horizontal Acoustic Waves", in Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, vol. 40, No 5, 1993, pp. 459-468.

[11] I.M. Silvestrova, V.V. Bezdelkin, P.A. Senyushenkov, Yu.V. Pisarevsky, "Present stage of La3Ga5SiO14 research", Proc. 1993 I.F.C.S., IEEE cat. 93CH3244-1, pp. 348-350.

[12] S. Sakharov, P. Senushencov, A. Medvedev, Yu. Pisarevsky, "New data on temperature stability and acoustical losses of langasite crystals", IEEE Int. Freq. Cont. Symp. 1995, IEEE cat # 95CH-3575-0

LOW-LOSS DECT-IF SAW FILTERS WITHOUT MATCHING NETWORKS

S.A.Dobershtein, V.A.Malyukhov

ONIIP, Maslennikov str., 231, Omsk, 644009, Russia

ABSTRACT

This paper presents low-loss ring SAW filter on 128^0 YX $LiNBO_3$ with a 3-dB fractional bandwidth of about 1%. Reflective multistrip couplers (RMSCs) with 3 electrodes per λ (λ is SAW wavelength at the center frequency of the filter) and long IDTs with phase weighting were used in the filter. The IDTs have a self-matching effect when the static capacitance of the IDT is compensated by the acoustic radiation susceptance. The ring filter at a center frequency of 110.6 MHz showed an insertion loss of about 3 dB, 3-dB bandwidth of about 1.1 MHz, stopband attenuation of 25, 35 and 50 dB at ±1.1, ±1.7 and ±2 MHz offsets from the center frequency respectively. The filter did not require matching networks, was mounted in SMD package (14.2x8.4x2.7 mm) and was used for IF-filtering in mobile communication (DECT system).

1. INTRODUCTION

There are several IF SAW filter designs for mobile communication providing a narrow bandwidth together with very steep skirts and high stopband attenuation [1,2]. As a rule these filters are not transversal type and have a reduced chip size. But the complexity of a specific construction and topology of such filters result in considerable decrease of the flexibility when developing SAW filters. Moreover matching networks are required to decrease insertion loss. In this connection the use of a ring construction is very perspective. It consists of input/output bidirectional interdigital transducers (IDTs) which are placed in parallel acoustic tracks and 2 reflective multistrip couplers (RMSCs) which provide SAW transmission between these tracks [3-5] (Fig.1). With this design it is easy to realize a high selectivity as in this case the filter frequency is defined by the product of the responses of RMSCs and input/output IDTs. Low insertion losses depend on a structure of the RMSCs. Also the matching networks are eliminated from this filter due to self-matching effect such that the specified real input/output impedances of the filter are obtained when the static capacitance of the IDT is compensated by the acoustic radiation susceptance. Using the RMSCs with 3 electrodes per λ (Fig.1) and the self-matching effect the ring filter can have an insertion loss of about 1 dB [5]. This paper presents low-loss self-matching ring SAW filter on 128^0 YX $LiNBO_3$ with a 3-dB fractional bandwidth of about 1% for IF-filtering in mobile communication (DECT system). The optimization of previously developed SAW ring filter [5] with a 3-dB fractional bandwidth of 2% on 128^0 YX $LiNBO_3$ was provided for achieving a fractional bandwidth of 1% and specified shape of the frequency response.

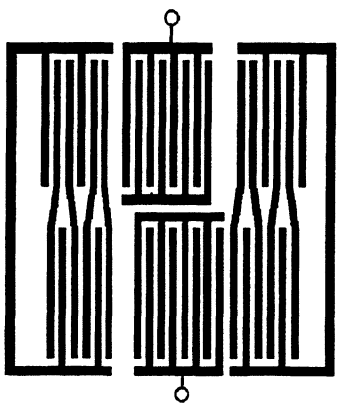


Fig.1. SAW ring filter with RMSCs

2. ANALYSIS OF PREVIOUSLY DEVELOPED LOW-LOSS SAW RING FILTER WITH A FRACTIONAL BANDWIDTH OF 2%

Design and technology aspects of self-matched low-loss ring filters on 128^0 YX $LiNBO_3$ are discussed in detail in our work [5]. According to [5] the fractional bandwidth of the ring filter is proportional to k^2 (k^2 is electromechanical-coupling coefficient) and equal to 2% on 128^0 YX $LiNBO_3$. Using the IDTs with phase weighting [6] the ring filter can have a stopband attenuation above 55 dB with a shape factor (3-40 dB) of 2.5. Our consideration is restricted to the analysis of the factors which influence on such characteristics of the ring filter as a fractional bandwidth and shape of the frequency response. Then we shall find the ways for achieving a specified combination of these characteristics. As shown in [5] the theoretical analysis of the ring filter frequency response was performed under conditions that the RMSC response is specified and optimum (it was determined experimentally for full SAW transmission from one RMSC track to another on the given cut of $LiNBO_3$). Then the calculation for the entire device is reduced to multiplying "the real RMSC response" (i.e., the experimental RMSC frequency response entered into the computer) by the frequency

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

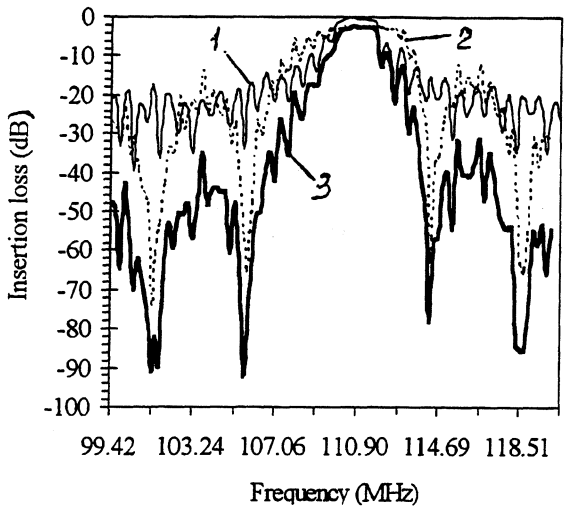


Fig.2. Calculated frequency response of a 2% bandwidth ring filter on 128° YX $LiNbO_3$

response (calculated in terms of the equivalent circuit model) of the ring filter consisting only of input/output self-matched IDT. Fig.2 shows the RMSC experimental frequency response for 128° YX $LiNBO_3$ (plot 1, a fractional bandwidth of 2%, number of RMSC electrodes of 225), calculated frequency response of the SAW ring filter only with input/output self-matched IDT (plot 2), total calculated frequency response of the ring filter with RMSCs and IDTs (plot 3). The latter was obtained by multiplying responses shown on plot 1 and plot 2. Input/output IDTs are unweighted and containe 27 finger pairs. As seen in the figure, in spite of the ripple in the transition band, the filter provides low ripple in a 2% fractional bandwidth and stopband attenuation over 40 dB at ±10 % offset from the center frequency. According to the requirements for DECT IF-filtering IF filter with a center frequency of 110.6 MHz must have low insertion loss, a fractional bandwidth of about 1% and stopband attenuation of 25, 35 dB at ±1.1, ±1.7 MHz offsets from the center frequency respectively. In order for these requirements to be performed on the basis of above discussed self-matched ring filter it is necessary to reduce its bandwidth and obtain a definite shape of the frequency response.

3. NARROWING A BANDWIDTH AND OBTAINING A DEFINITE SHAPE OF THE FREQUENCY RESPONSE FOR THE RING FILTER

The calculations show that for this purpose the Gaussian long phase weighted IDTs (100-120 λ in length) can be used. Then the frequency response of the filter in the transition band will be determined by the frequency response of the weighted IDTs. The high stopband attenuation at great offsets is obtained by the frequency response of RMSCs and IDTs. The final shape of the frequency response for the ring filter is controlled by the RMSCs frequency response (increasing or decreasing the number of RMSC electrodes within low limits); by IDTs frequency response (using a definite length of the weighted IDTs); by the frequency shift between the frequency responses of RMSCs and IDTs (using the different period of the electrodes in RMSCs and IDTs).

4. OBTAINING A LOW INSERTION LOSS AND LOW RIPPLE IN THE PASSBAND FOR SELF-MATCHED RING FILTER WITH LONG WEIGHTED IDTS

It is desirable that a ring filter with long weighted IDTs would be a self-matched filter, i.e. it should provided the low insertion loss and low ripple in the passband without matching networks. Fig.3 shows a calculated frequency response of the ring filter having only input/output IDT with 27 finger pairs on 128° YX $LiNBO_3$ (plot 1). The filter has been connected to 120-Ω loads without matching networks. As seen in the figure the filter provides low ripple in a 2.9% fractional bandwidth. Using the long phase weighted IDTs (100-120 λ in length) the problem of obtaining low ripple and specified pure real input/output impedance of the filter in the passband arises due to reflections from a large number of electrodes. For solution of this problem it is necessary to determine an optimal width and number of the finger pairs in each transverse section of phase weighted IDTs. Results of a large body of calculations are presented in Fig.3 (plot 2). Here the calculated frequency response of the ring filter only with Gaussian phase weighted input/output IDTs is shown. Phase weighted IDTs had a length of 115 λ, electrode width of about 3λ/16 and number of the finger pairs in each transverse section of 26. As seen in the figure the filter provides low ripple in a 1 % fractional bandwidth by self-matching effect.

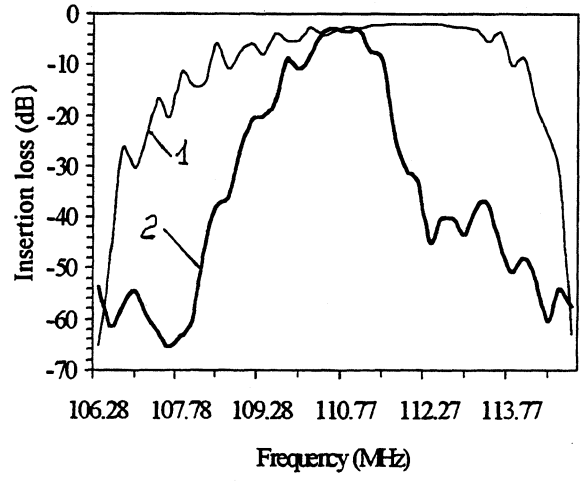


Fig.3. Calculated frequency response of the ring filter only with unweighted and long weighted IDTs

5. DEVELOPMENT OF SELF-MATCHED RING FILTER WITH A 1% FRACTIONAL BANDWIDTH AND SPECIFIED SHAPE OF THE FREQUENCY RESPONSE

The calculation of the ring filter was realized by successive approximations. At first we gained the specified fractional bandwidth and shape of the frequency response on the basis of data from paragraph 3. Then we checked the insertion loss and low ripple within the passband achieved by self-matching effect using data from paragraph 4. If the fractional bandwidth, shape of the frequency response, insertion loss and ripple with self-matching satisfied the specified requirements, we stopped the calculations and determined a topology of the filter (number of RMSC and IDT electrodes; period and width of the electrodes et al.). Otherwise we changed the parameters of the filter topology, carried out some calculated iterations up to realization of all specified requirements for the filter. The calculated frequency response of the ring filter with RMSCs and a tolerance on IF-filtering for DECT system are shown in Fig.4. RMSCs incorporated 220 electrodes, the Gaussian weighted IDTs were 115λ long, the period of the electrodes in RMSCs - to -period of the electrodes in IDTs ratio was about 0.994, width of the electrodes in the weighted IDTs was about 3λ/16, number of the finger pairs in each transverse section of the phase weighted IDTs was 26. As seen from Fig.4 the calculated frequency response of the ring filter satisfies the requirements of IF-filtering in DECT system.

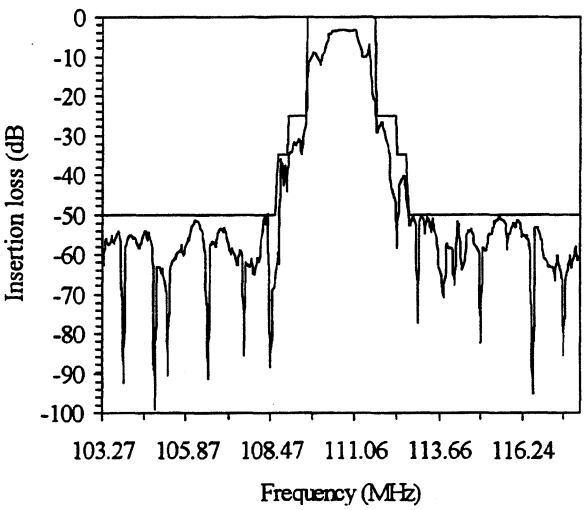


Fig.4. Calculated frequency response of the ring filter with a tolerance on IF-filtering in DECT system

6. EXPERIMENTAL RESULTS

The ring filter was made on 128° YX $LiNBO_3$. The metallization height was 0.5 μm. Measured in a 120-Ω system frequency responses of the ring filter in a broad frequency range and passband are shown in Fig.5 and Fig.6. Fig.4 and Fig.5 show good agreement between the calculated and measured frequency responses. The ring SAW filter at a center frequency of 110.6 MHz showed an insertion loss of about 3dB, 3-dB bandwidth of about 1.1 MHz, stopband attenuation of 25, 35 and 50 dB at ±1.1, ±1.7 and ±2 MHz offsets from the center frequency respectively. These parameters allow to use this filter for IF-filtering in DECT system. The filter did not require matching networks and was mounted in SMD package DCC-14 (14.2x8.4x2.7 mm, Fig. 7).

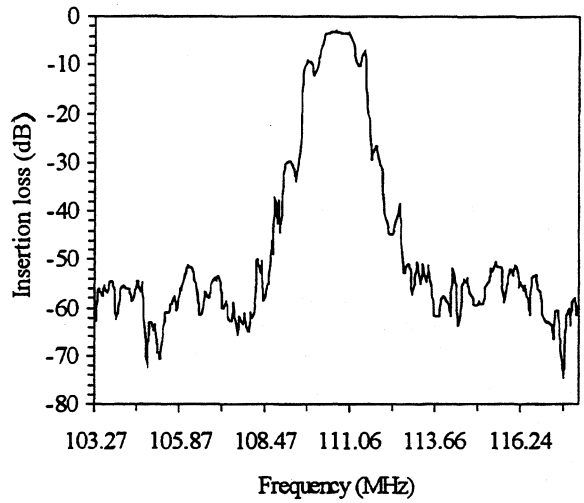


Fig.5. Measured frequency response of the ring filter for IF-filtering in a broad frequency range

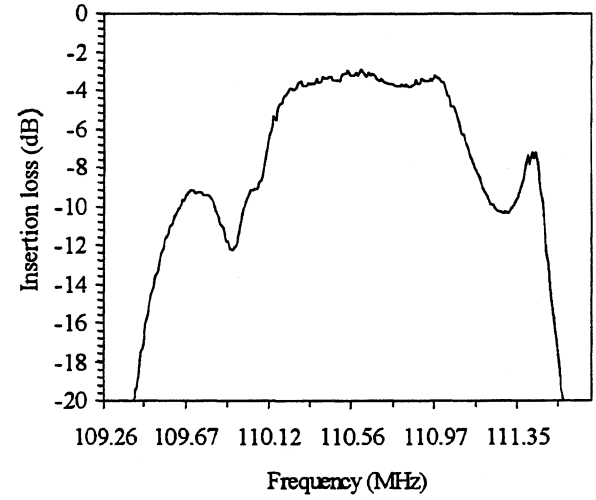


Fig.6. Measured frequency response of the ring filter for IF-filtering in a passband

7. CONCLUSION

We have developed a low-loss ring SAW filter for IF-filtering in the DECT system. In the design of the ring filter we used RMSCs with 3 electrodes per λ, long phase weighted IDTs with self-matching effect. Without matching networks the ring filter provided low insertion

loss (about 3 dB), high stopband attenuation (about 50 dB), narrow bandwidth (about 1%), small size and compatibility with SMT.

8. REFERENCES

[1]. K.Anemogianis et al, "Cordless Phone System Architecture Based on SAW Filters", in Proceedings of the IEEE Ultrasonics Symposium, 1993, pp.85-88.

[2]. R.Dill et al, " A Novel SAW Filter for IF-Filtering in DECT Systems", in Proceedings of the IEEE Ultrasonics Symposium, 1995, pp.51-54.

[3]. F.Sandy and T.Parker, "Surface Acoustic Wave Ring Filter", in Proceedings of the IEEE Ultrasonics Symposium, 1976, pp.391-396.

[4]. W.Pollock et al, "Low-loss SAW Filter Using Single-Phase IDTs and No External Tuning", in Proceedings of the IEEE Ultrasonics Symposium, 1983, pp.87-92.

[5]. S.Dobershtein, V.Malyukhov, "SAW Ring Filter with Insertion Loss of 1 dB", IEEE Transactions on UFFC, vol. 44, pp.590-596, May 1997.

[6]. M.Hikita et al, "Phase Weighting for Low-Loss SAW Filters", in Proceedings of the IEEE Ultrasonics Symposium, 1980, pp.308-312.

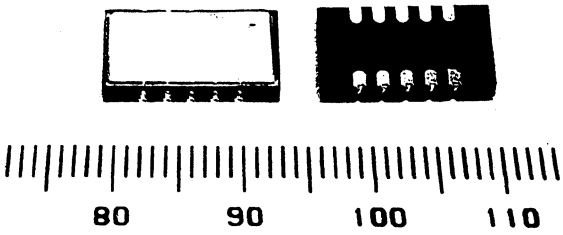


Fig. 7. Photograph of a low-loss ring IF SAW filter in a SMD package

EXPERIMENTAL MEASUREMENTS OF VELOCITIES AND TEMPERATURE EFFECTS FOR SAW ON Y-ROTATED AND X-CUTS OF LANGASITE

E. Henry Briot*, E. Bigler*, S. Ballandras*, G. Marianneau* and M. Solal**
* Laboratoire de Physique et Métrologie des Oscillateurs du CNRS associé à l'Université de Franche-Comté, 32, Av. de l'Observatoire, 25044 Besançon Cedex, France
** Thomson Microsonics, 399 route des Crêtes, BP 232, 06904 Sophia Antipolis Cedex, France

Abstract

A systematic experimental investigation of SAW properties on Langasite is presented on a series of Y-rotated plates. Phase velocities have been measured for many orientations and compared to theoretical predictions based on already published material constants. A good agreement is found between the experiments and computations based on data published by Sakharov *et al.* [1]. Temperature dependance have been also measured and compared to theoretical calculations based on the classical method of variation of elastic constants [2]. Compensated cuts have been found for all three orientations. Additional measurements of SAW propagation on X cut plates will be also presented[1].

Introduction

Recently Langasite (LGS, $La_3Ga_5SiO_{14}$) has attracted a lot of interest for surface acoustic wave applications [3, 4, 5, 6]. Among interesting orientations, singly rotated plates defined by IEEE angles ($\varphi = 0°$, θ, ψ) [7] as well as X cuts ($\varphi = 30°$, $\theta = 0°$, ψ) appear to be very interesting since temperature compensated orientations exist exhibiting both a higher electromechanical coupling factor than the well known (ST, X) cut of Quartz and zero power flow angle.

We present here experimental results on Y+45, Y+50 and Y+55 rotated plates of Langasite. On each plate multiple delay lines at various propagation angles (from $\psi = 0°$ to $\psi = 90°$) have been fabricated and tested. Temperature compensated cuts are found on all three θ orientations for various propagation angles ψ. Parabolic frequency versus temperature dependences have been found and the shift of the turnover temperature versus cut angle θ and propagation angle ψ has been measured.

[1]This work is supported by CNRS (France), Région de Franche-Comté (France) and Thomson Microsonics (Sophia Antipolis, France)

1 Experimental protocol

The experimental devices are delay lines (show figure 1) processed on the surface of the LGS plates. A 1500 Å thick Aluminum film is first deposited on the surface by sputtering technology. Delay lines are then patterned with a classical U.V. photholithography process. Finally, the Aluminum layer is chemically etched to obtain the following transducer structure on the LGS plate.

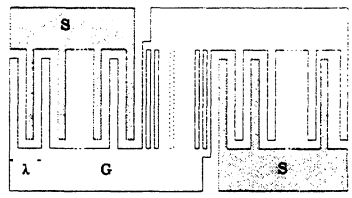


Figure 1: Delay line processed at the surface of the wafer (G: ground, S:signal)

Each device on the plate is processed with an angle ψ (propagation angle) between the flat and the wave's propagation direction. 19 different propagation directions can be then measured on the same 3" wafer with the same technological parameters (Aluminum thickness, times in photolithography process, metallization ratio, ...).

The synchronous frequency of Rayleigh mode is mesured using a network analyser and the velocity V can be roughly derived from both synchronous frequency F and acoustic periodicity λ imposed by the geometry of the device, according to $V = F \times \lambda$.

For temperature measurements, delay lines are submitted to a quasi-static temperature cycle from $-20°C$ to $120°C$. The measure of the frequency and the corresponding temperature is automatically controlled by a computer.

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

2 Experimental results

Figures 2 and 3 present the frequency response for the Y+50 LGS cut and two different propagation angles.

These figures point out the presence of many peaks and higher losses for an orientation exhibiting a bad coupling coefficient ($\psi = 0$, *cf.* figures 2 and 4) than for another one exhibiting a high coupling coefficient ($\psi = 25$, *cf.* figures 3 and 4).

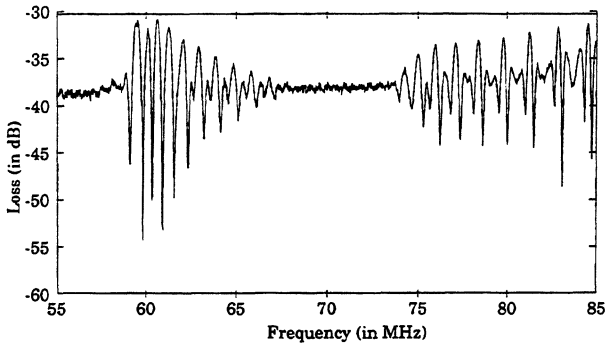


Figure 2: Log. Magnitude of transmission response of a delay line on Y+50 cut of LGS ($\psi = 0°$)

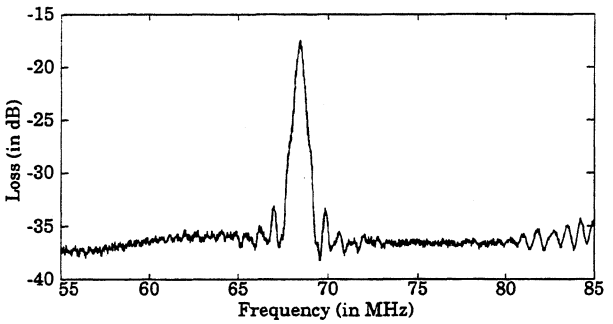


Figure 3: Log. Magnitude of transmission response of a delay line on Y+50 cut of LGS ($\psi = 25°$)

Figure 5 presents the corresponding thermal behaviour for the Y+50 LGS cut ($\psi = 25°$). First (${}^{\theta}\alpha^{(1)}$) and second (${}^{\theta}\alpha^{(2)}$) order Temperature Coefficients of Frequency are then deduced by fitting the experimental data. The turn-over temperature ($T_{turnover}$) is calculated according to :

$$T_{turnover} = 25 - {}^{\theta}\alpha^{(1)}/(2\,{}^{\theta}\alpha^{(2)})$$

3 Comparison with theoretical models

Figures 6, 7 and 8 present a comparison between measured velocities and theoretical predictions based on different sets of elastic constants (LGS1 [8], LGS2 [9],

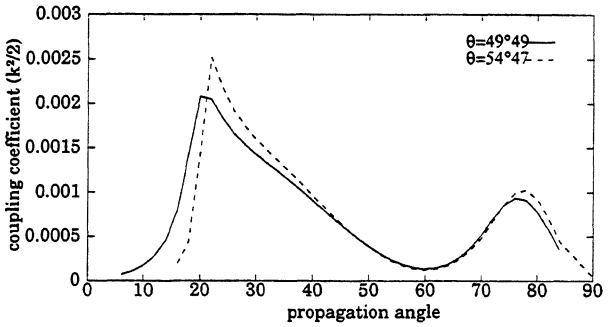


Figure 4: Theoretical coupling coefficient $k^2/2$ for Y+50 and Y+55 langasite cuts

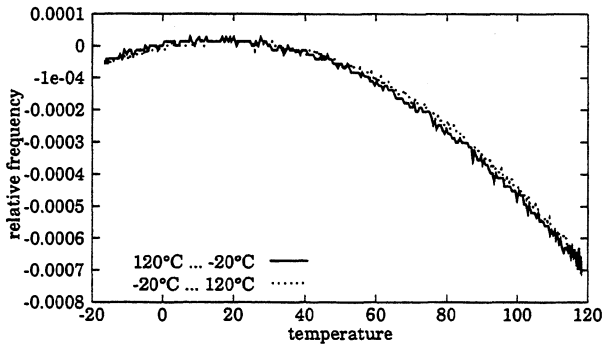


Figure 5: Dependance of the synchronous frequency of a delay line on Y+50 cut of LGS ($\psi = 25°$)

LGS3 [1]). The agreement between theory and experiment is fairly good as far as velocities are concerned, the agreement is better than 1% within the whole set of experimental data.

The last set of constants published by Sakharov [1] seems to be the best one for the prediction of Rayleigh wave velocities on this Langasite wafers. It should be noted that no high velocity mode (except bulk wave radiation) were observed for the 57 measured orientations.

Figures 9, 10 and 11 present the corresponding thermal behaviour in terms of turnover temperature for all three Y-rotated orientations compared to theoretical predictions based on different sets of LGS data (LGS1,2,3). LGS2 data set seems to provide the best prediction of thermal effects compared to our experiments but the experimental ${}^{\theta}\alpha^{(1)}$ dependance on ψ seems to be more similar to the LGS1 set of constants (except the 100°C difference of value). Then, no conclusion could be proposed to define the best set of constants for accurate predictions of thermal behaviour of SAW on LGS before further experiments.

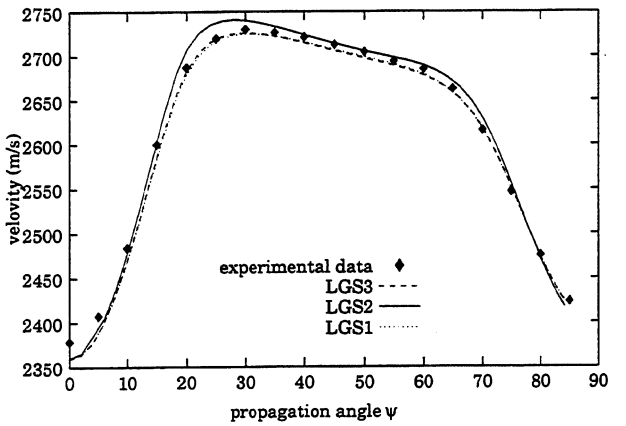


Figure 6: Comparison between theoretical and experimental velocities for a Y+45 langasite cut

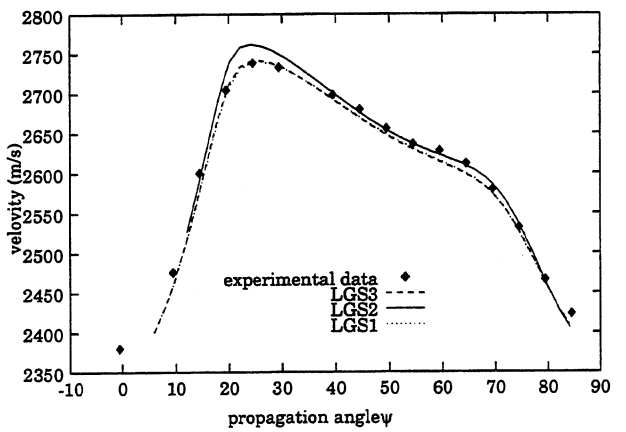


Figure 7: Comparison between theoretical and experimental velocities for a Y+50 langasite cut

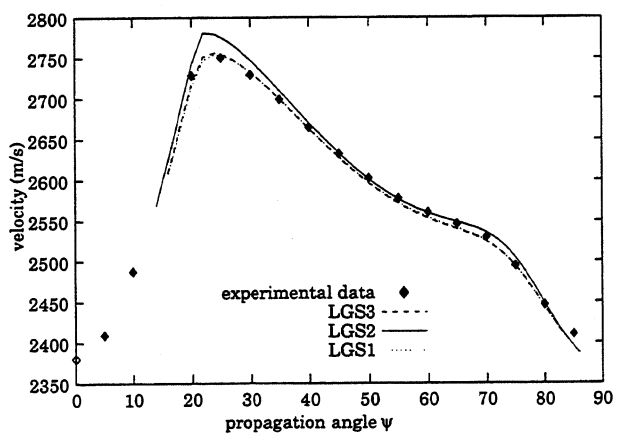


Figure 8: Comparison between theoretical and experimental velocities for a Y+55 langasite cut

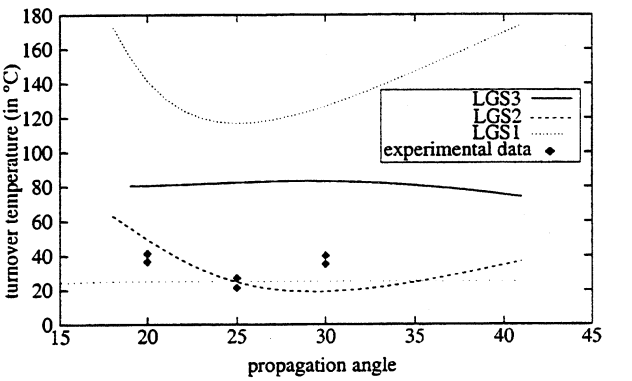


Figure 9: Comparison between theoretical and experimental turnover temperatures for a Y+45 langasite cut

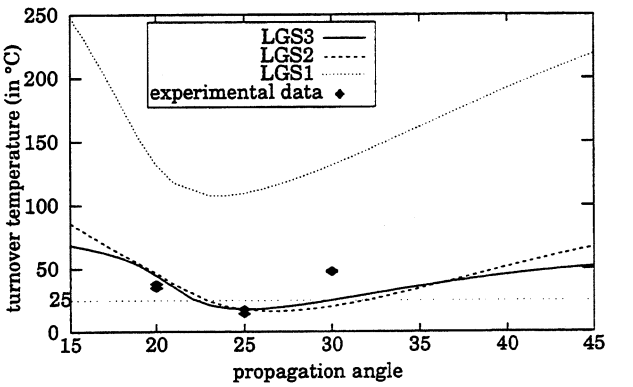


Figure 10: Comparison between theoretical and experimental turnover temperatures for a Y+50 langasite cut

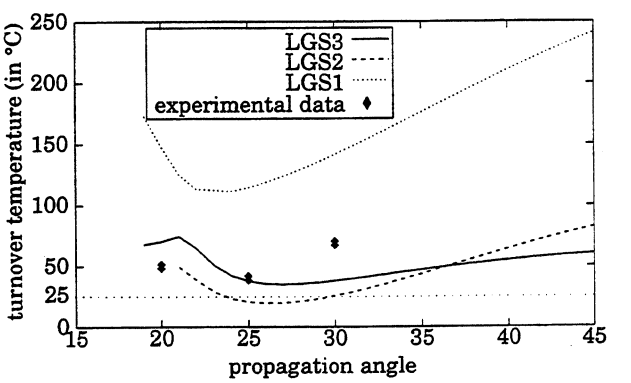


Figure 11: Comparison between theoretical and experimental turnover temperatures for a Y+55 langasite cut

Second order temperature coefficient of frequency were measured between -70 and -100 ppb/°K^2.

Figure 12 presents phase velocities measured on X-cut of LGS compared to theoretical ones for different propagation angles and different sets of constants.

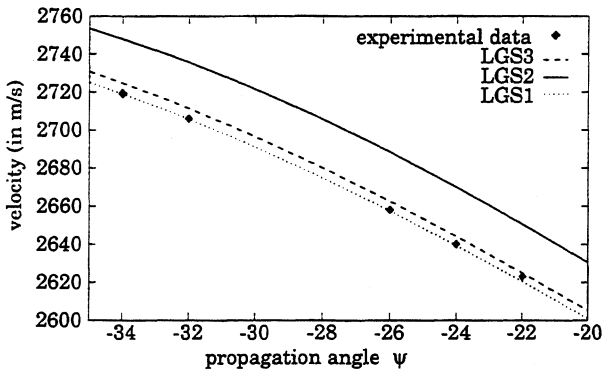


Figure 12: Comparison between theoretical and experimental velocities on X-cut of LGS

In that case, Silvestrova's constants [8] appears as the best set together with Sakharov's ones [1]. Note that the X-cut wafer has been cut in another boule of LGS than the one used to cut the tested Y-cuts.

Conclusion

Measurements on Y-rotated cuts have been presented and compared to theoretical predictions for phase velocities and turnover temperature of SAW propagation. Thermal compensated orientations have been pointed out according to recent publications [4, 5, 6]. The stability of thermal compensation versus propagation angle ψ and cut angle θ is interesting for SAW applications. Finally, if Rayleigh wave velocities can be precisely predicted theoretically, theoretical thermal sensitivity prediction of SAW on LGS remains rather uncertain because of the difficulty to accurately characterize its thermoelastic properties.

References

[1] S.Sakharov, P. Senushencov, A. Medvedev, and Yu. Pisarevsky. New data on temperature stability and acoustical losses of langasite. In *Proc. IEEE International Frequency Control Symposium*, pages 647–651, 1995.

[2] J.J. Campbell and W.R. Jones., A method for estimating optimal crystal cuts and propagation direction for excitation of piezoelectric surface waves. *IEEE Trans. on Sonics and Ultrason.*, SU-15:209–217, 1968.

[3] M. Murota and Y. Shimizu. Theoretical investigation of rayleigh waves on a $la_3ga_5sio_{14}$ substrate. In *Proc. IEEE International Ultrasonics Symposium*, Sendai, Japan, 1998.

[4] K. Inoue and K. Sato. Temperature stability of saw on langasite single crystals. In *Proc. IEEE International Ultrasonics Symposium*, Sendai, Japan, 1998.

[5] A. Bungo, C. Jian, K. Yamagouchi, T. Suetsugu, and Y. Shinohara. Saw filter application of a langasite substrate. In *Proc. 27th Electro-mechanical Symp. in Jpn.*, pages 125–129, may 1998.

[6] M. Kadota, T. Kitamura, and M. Kumatoriya. Surface acoustic wave properties on rotated y-cut $la_3ga_5sio_{14}$. In *Proc. 27th Electro-mechanical Symp. in Jpn.*, pages 117–124, may 1998.

[7] The Institute of Electrical and Electronics Engineers, Inc. *IEEE Standard on Piezoelectricity*, 1949. ANSI/IEEE Std 176-1949.

[8] I.M. Silvestrova, Yu.V. Pisarevsky, V.V. Bezdelkin, and P.A. Senushenkov. New piezoelectric materials. In *Proc. IEEE International Frequency Control Symposium*, pages 351–352, 1993.

[9] I.M.Silvestrova, Yu.V. Pisarevski, P.A. Senushenkov, and A.I. Kruppnyi. Temperature dependances of the elastic properties of $La_3Ga_5SiO_{14}$ single crystals. *Sov. Phys. Solid State*, 28(9):1613–1614, Sept. 1986.

1999 Joint Meeting EFTF - IEEE IFCS

DISPERSION OF SAW VELOCITY AND TRANSFORMATION OF SAW INTO THE BULK WAVES IN REFLECTIVE GRATINGS

V.F. Dmitriev *, I.S. Mitrofanov **

* SP&TC "LENINETZ",212 Moskovsky avenue,196066 St.Petersburg, Russia
** Avangard-Elionica Co.,72 Kondratyevsky avenue, 195271 St.Petersburg, Russia

ABSTRACT

This report deals with the investigation of the SAW velocity dispersion and determination of the coefficients of SAW transformation into the bulk acoustic waves on the base of experimental and theoretical analysis of the SAW propagation in reflective gratings. The reflective gratings consist of the irregular sequence of grooves on the surface of the lithium niobate YZ-cut.

1. INTRODUCTION

A reflective gratings are broadly used in resonators, filters and dispersive delay lines. However, a SAW attenuation and changing of a SAW velocity take place when the SAW is propagating across the reflective grating (RG). Moreover, attenuation of a SAW is the result not only of the losses in the material of the piezoelectric, as well as transformation of the part of the SAW energy into the bulk waves. The process of SAW propagation across the grating may be characterized by the constant of propagation, which took into account the mentioned above effects:

$$\gamma = \beta + \imath\,\alpha, \qquad (1)$$

where, $\beta=\beta_o+\Delta\beta(\omega,h)$, $\alpha=\alpha_o(\omega)+\alpha_1(\omega,h)$, β_o and α_o are the constants of propagation and the coefficient of attenuation of a SAW on the free surface of the piezoelectric; $\alpha_1(\omega,h)$ takes into account the partial transformation of a SAW into the bulk acoustic waves, $\Delta\beta(\omega,h)$ takes into account a changing of a SAW velocity under the grooves. Experimental determination of the $\Delta\beta(\omega,h)$ and $\alpha_1(\omega,h)$ for the event of tilted incidence of a SAW on grooves, has performed on the base of comparison of theoretical and experimental of the amplitude and phase frequency responses a number of dispersive delay lines (DDL) with reflective gratings. A preferred topology in the process of designing the broadband DDL is slanted topology [1]. The equations which have been used for the calculation of the amplitude and phase frequency responses of the slanted DDL was presented in [2]. In the process of the calculating analysis of the amplitude and phase responses such a values of the $\Delta\beta(\omega,h)$ and $\alpha_1(\omega,h)$ were selected, that the small alteration relatively one, has brought about the essential divergence of the results of calculation and experiment.

2. RESULTS OF INVESTIGATIONS

The broadband DDL with big duration of the dispersive delay time are highly perspective for the experimental determination (or verification) of the coefficients of a SAW transformation into the bulk acoustic waves and the values of the changing of a SAW velocity under the RG grooves. The more duration of the DDL dispersive delay time the more is the sensitiveness of the frequency responses of a DDL to the mentioned above effects. Besides, the broader bandwidth of a DDL the more is well-marked manifestation of the mentioned above effects for the different frequencies in the DDL frequency responses. The additional possibilities in the determination of the coefficients are given by some particularities which have been connected with the unperfection of the reflection grating fabrication process. Let's consider the particularities which has been observed in the frequency responses of a DDL-1 with the parameters: F_o=150 MHz, Δf=40 MHz, T=75 μs. Such a slanted DDL on the substrate of 145x30 mm^2 size of the lithium niobate YZ-cut has been designed and manufactured. The topology of a slanted DDL is presented in Fig.1.

Each IDT of the DDL-1 has consists of 215 electrodes and the reflective structure has 10400 grooves. The aperture of the IDT electrodes on central frequency has of 35 wavelength. The equalizing of the unflatness of the DDL frequency response with the negative slopping of the dispersive curve was performed by the linear changing of the grooves depth along RG. This way allows to receive the minimum insertion losses in the DDL and a more flat passband of the frequency response.

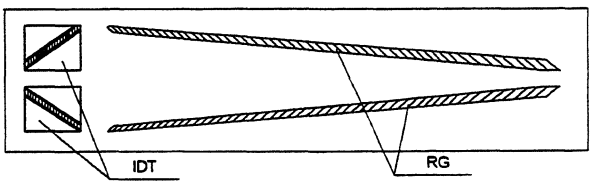


Fig.1. Topology of the slanted DDL.

The calculation of the frequency responses of a DDL was based on the equations, which had been presented in [2]. The attenuation of a SAW amplitude at the passing through the groove number k of the

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

reflective grating with the account of partial transformation into the bulk was calculated according to equation

$$\alpha_1(\omega,h_k) = C_P(\omega)\,(h_k/\lambda)^2, \qquad (2)$$

where h_k is the groove depth of reflective grating, λ is the wavelength of a SAW which falls upon the groove. The calculation method of the frequency-dependent coefficient $C_P(\omega)$ has been presented in paper [3]. For the Poisson's ratio equals to 0.309 and for the reasonable values of the steepness of the grooves edges the frequency dependence of the $C_P(\omega)$ was calculated in accordance with [3]. The frequency dependence of $C_P(\omega)$ used in the following calculations is presented in Fig.2, where f_K is the synchronous frequency of the k-th groove.

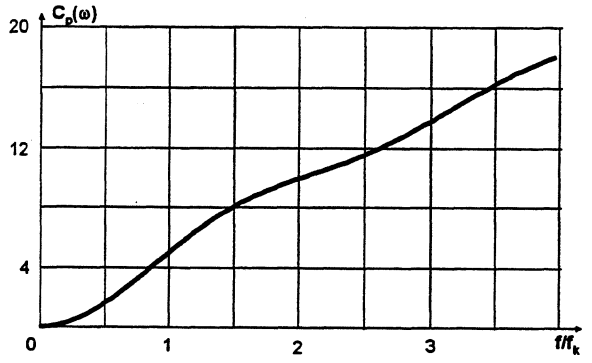


Fig.2.Frequency dependence of the coefficient of a SAW transformations into the bulk.

Additional phase shift at the passing through the groove number k may be calculated as:

$$\Delta\varphi_k = 2\,C_v\,(h_k/\lambda)^2, \qquad (3)$$

where C_v is the coefficient which should be defined in the experiment. Notice that $\Delta\beta=\Delta\varphi_k/p$, where p is a period of the RG in the SAW propagation direction. When the DDL frequency response was being calculated the value of $\Delta\varphi_k$ calculated according to the equation (3) should be added as the additional phase-member in the expression (7) of paper [2]. The problem of the analysis was reduced to the selecting of the coefficients Cp(ω) and Cv so as the distinctive experimentaly observed distortion of the frequency responses of a DDL was to the best agreement with the calculated for the DDL with different parameters. The calculation analysis according to equations [2] shows that the main influence upon the form of a DDL frequency response has three factors connected between itself. The first factor is a correspondence of the grooves depth profile along RG which received in the process of the etching of the substrate to the required linear law. The second factor is a scattering of a SAW into the bulk. The third factor is a changing of a SAW velocity under the RG grooves. The more the depth of grooves, the more are being demonstrated the effects of scatterring into the bulk and the changing of a SAW velocity in the frequency responses. The decreasing of the effects of scatterring into the bulk up to acceptable low-level is possible in the case of the choice of a linear changing of a grooves depth from 0.08 mkm (near the IDT) up to 0.16 mkm (at the end of RG). The results of measurement and calculation for the DDL with a such grooves depth profile by use of the equations which have been described in [2] are presented in Fig.3.

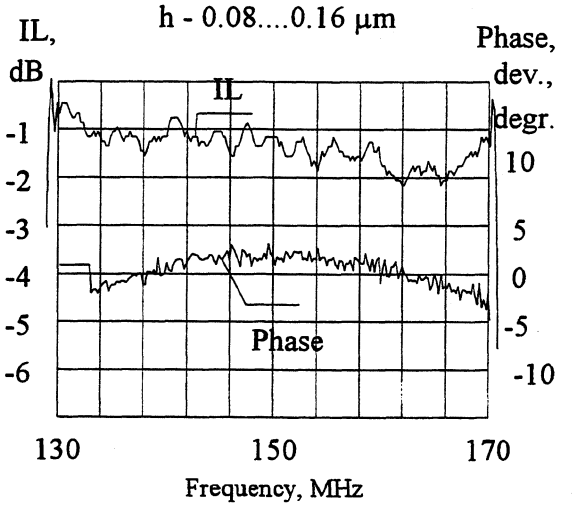


Fig.3,a. Calculated frequency responses

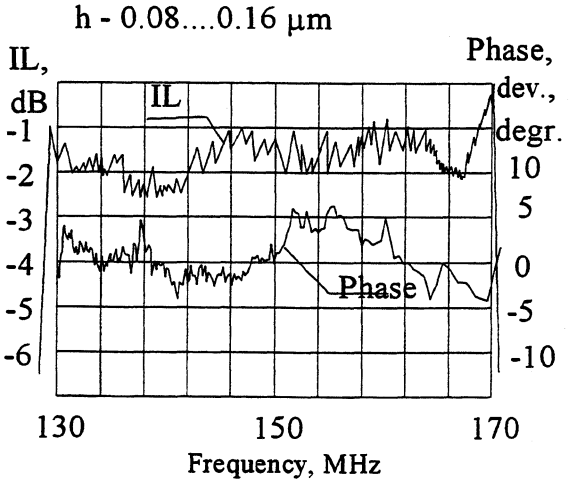


Fig.3,b. Measured frequency responses

The measurements of the grooves depth along RG has showed a linear changing with the deviation from the

required values not more than 5 nm. The value of $C_P(\omega)$ accepted at the calculations in accordance with the curve in Fig.2, C_v=17. The calculated value of the insertion losses is 46.5 dB. In this example on the amplitude and phase responses of the DDL the particularities which deals with the uneven changing of a grooves depth along RG are absent. The uneven changing of a grooves depth have been caused by the imperfection of the technological process of etching and brings about the strong distortion of the DDL frequency responses. However, for these cases are possible the most exact determination of the coefficients $C_P(\omega)$, C_v on the base of the analysis of the calculation and experimental results. The most interesting results, from the point of view of possibility of the determination of the coefficients $C_P(\omega)$, C_v, are presented in Fig.4 and Fig.5.

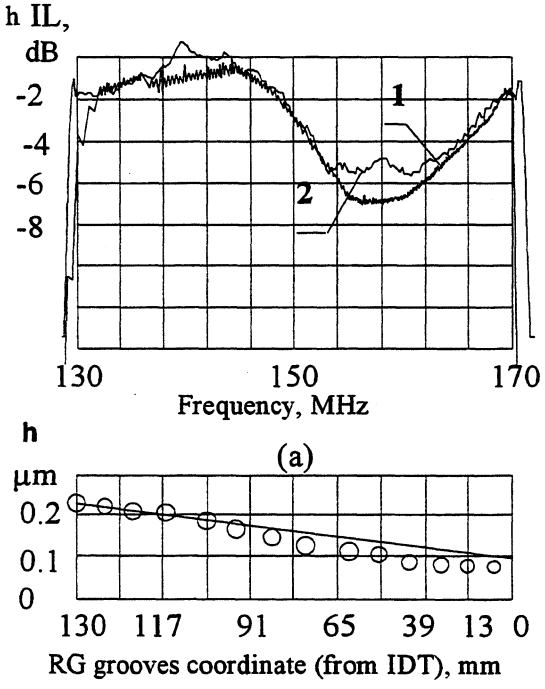


Fig.4. Frequency responses of the DDL-1 no.2

1 - measured, 2 - calculated (a), in comparison with the measured RG grooves depth profile (b). Insertion loss on Fo 42 dB.

The given results has been received on the stage of the technology elaboration of the grooves etching. The required (utter line) and the measured (circles) values of the grooves depth along RG are presented in Fig.4,b and Fig.5,b. The results of measurements of the DDL-1 frequency responses are presented in Fig.4,a and Fig.5,a (depicted by curve 1). The results of calculation are presented in Fig.4,a and Fig.5,a depicted by curve 2. The grooves depth profile along RG is accepted according to measured data. At the DDL frequency response calculation the values of coefficients $C_P(\omega)$, C_v are accepted the same as for the curves in Fig.3.

Bell- type of the DDL frequency response in Fig.5 is due to the sharp deviation from the linear law of the RG grooves depth profile. For the SAW reflected by grooves near this deviation the signal amplitude is increasing (see Fig.5,b). The SAW after the passing through the area of the great deviation of the grooves depth and before being reflected by the RG grooves has a strong attenuation (scattering into the bulk). Such a strong attenuation take place in spite of the greater reflection coefficient of the grooves (h/λ more). The additional losses in the area of sharp deviation of the grooves depth profile from the linear law are connected also with the breaking of the phase

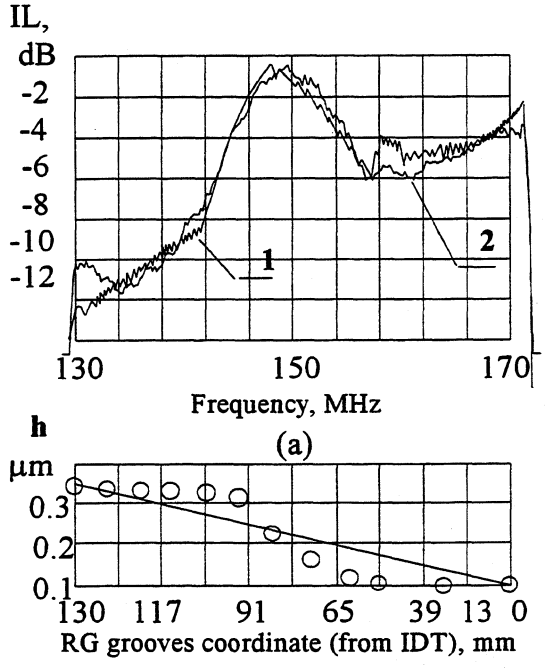


Fig.5. Frequency responses of the DDL-1 no.3.

1 - measured, 2 - calculated (a), in comparison with the measured RG grooves depth profile (b). Insertion loss on Fo 46 dB.

synchronism between the reflected waves from the nearby grooves. Should be noted, that the phase member have changes in accordance with (2). The results of calculation of a DDL frequency responses with the grooves depth profile in accordance with Fig.5 for different values of coefficients $C_P(\omega)$, C_v are presented in Fig.6 and Fig.7. The shift of upper and

lower curves in Fig.6, Fig.7 relatively the curve between them demonstrates the divergence of the calculation and experimental results for the uncorrect choice of coefficients $C_P(\omega)$, C_v. For the convenience in this example have been introduced the additional empirical coefficient K_P such, that the $C_P(\omega) \rightarrow K_P C_P(\omega)$. The frequency responses of a DDL-1 for three different values of coefficient K_P=0; 1; 1.5 (C_v=17) are

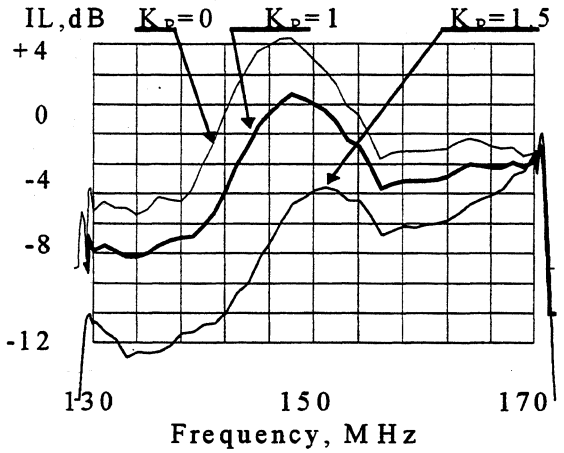


Fig. 6. Calculated frequency responses of a DDL-1 no.3 with the grooves depth profile in accordance with Fig.5,b and for different values of coefficients $C_P(\omega)$ and C_V =17.

presented in Fig.6. The frequency responses of a DDL for three different values of coefficient C_v = 14; 17; 20 (K_P=1) are presented in Fig.7. The best agreement of the calculated and experimental insertion losses of the DDL (the curves in Fig.6 and Fig.7) are received for the values of parameters K_P= 1 and C_v = 17. Note that the curves in Fig.6, Fig.7 have been received for the smaller number of the calculation points, than in the Fig.4 and Fig.5.

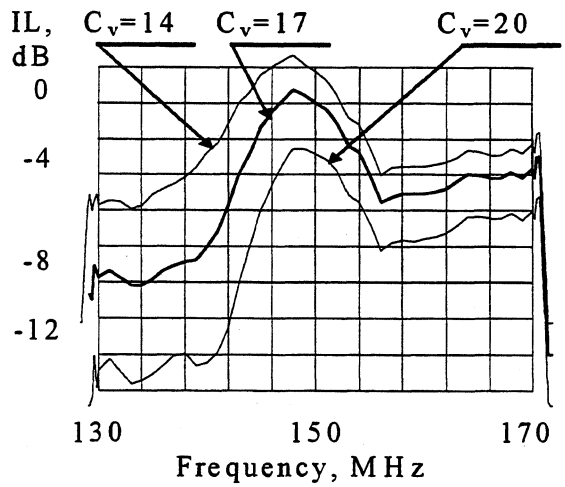


Fig. 7. Calculated frequency responses of a DDL-1 no.3 with the grooves depth profile in accordance with Fig.5,b and for different values of coefficients C_v.

The same calculation have been performed for DDL with parameters:

DDL-2: Fo=125 MHz, Δf=40 MHz, T=30 μs;

DDL-3: Fo=85 MHz, Δf=30 MHz, T=30 μs.

The DDL have been produced on substrates of the lithium niobate YZ-cut with the slanted topology.

The comparison of the frequency responses of the mentioned DDL has show that the best agreement of the results of calculation and experiment is observed at the choice of the values of the coefficients $C_P(\omega)$ in accordance with Fig.2 and C_v=17.

3.CONCLUSION

As a result of conducted investigations on the base of comparisons of the results of calculation and experiments is stated:

1)The process of the SAW conversion into the bulk waves by the irregular sequence of the grooves (ribbs) is being described by the expression (2) with the accuracy sufficient for the practical calculations. The value of the coefficient $C_P(\omega)$ can be accepted in accordance with the curve in Fig.2.

2)For the lithium niobate YZ-cut and in the event of a SAW incidence on the grooves under the corner about 47°, the coefficient of dispersion of a SAW velocity under the RG grooves C_v is equal about 17.

3)The using of a sharp uneven changing of a grooves depth profile along RG (for instance for the equalizing of a DDL amplitude response) carryes in about the degradation of a DDL phase response.

4. REFERENCES

[1] C.Waterkeyn, H.Gautier, "Detailed analysis of slanted device technology", in Proceedings of the IEEE Ultrasonics Symp., 1983, pp.195-199.

[2] V.F.Dmitriev, I.S.Mitrofanov, "SAW conversion in chirp pulse expanders and compressors", in Proceedings of the 12-th EFTF, 1998, pp. 463-468.

[3] V.F.Dmitriev, "Coherent and uncoherent scattering of SAW in the periodical reflective structures", Sov. J. Tech. Phys., vol.65, (8), August, pp.111-123, 1995.

1999 Joint Meeting EFTF - IEEE IFCS

THE DESIGN VERIFICATION MODULE FOR A SAW DESIGN AUTOMATION SYSTEM

M. J. McCOLLISTER and S. M. RICHIE

University of Central Florida, Department of Electrical and Computer Engineering, Orlando, Florida 32816-2450

ABSTRACT

This paper presents a stand-alone version of the design verification module for a SAW design automation system. The automation system that is being developed requires a verification phase that will be used to examine S-parameter data from test devices and determine if a specific device meets design requirements. This paper includes the specification and performance data structures, methods of extracting performance data, and methods of verifying performance versus requirements. The algorithms and techniques for efficiently extracting device performance data from experimental data are presented. Techniques including Fourier and wavelet coefficient reduction for parameter extraction are included. The core of the system contains the specification and performance data structures appropriate for SAW bandpass filter devices. The specifications contain the time response constraints, frequency response magnitude and phase, and delay constraints. Performance data is extracted from measured or calculated S-parameters as a function of frequency. This data is compared with the design requirements to see if a device meets these specifications. If a device fails, the device parameters that caused the failure are identified. All performance failures can be used as a basis for design corrections. The algorithms and programs developed in this paper will be available via the Internet at http://pegasus.cc.ucf.edu/~mjm05082/fcs1999/.

1. INTRODUCTION

The development of a SAW design automation system includes four main parts. These parts are the user specification, rule and model based SAW design, analysis and layout, and performance evaluation and verification [1]. While the end result of a SAW design automation system is the design of a working SAW device, the system must verify that the data derived meets the designer's specifications.

This paper presents the development efforts of SawVerify, the stand-alone version of a SAW design verification module. While this program does not design a SAW device given user specifications, it does take measured data from a real SAW device and extracts various parameters and determines which parameters meet the user specifications. A future completed automated SAW design system will use modeled data as well as measured data from a physical device to determine if a SAW device meets specifications.

2. DATA STRUCTURES

SawVerify was developed using Borland C++ Builder 4 running on Microsoft Windows 95/98. The object oriented paradigm of C++ provided to be the natural way of representing data associated with a measured device. A class, called TMeasDev, was created to encapsulate the data and derived parameters from raw S-parameter data. This class fully represents the SAW device as a "black box". By doing so, the class can either encapsulate data from a measured device or from a device model. This class also separates the GUI from the data and parameter extraction functions so that porting to another compiler requires very little changes to the core processing code.

2.1 C++ and the Standard Template Library

C++ includes the standard template library (STL) which contains predefined classes for complex numbers, vectors and many other data structures. The STL vector class was used extensively, instead of the ANSI C arrays, to simplify memory allocation and vector sizing. This also reduced errors and development time.

2.2 The TMeasDev Class

The TMeasDev class was designed to not only extract parameters from a device, but also to verify that the device meets user specifications. The data sizes necessary to contain the device data of N points include three float vectors of size N, four complex vectors of size N, and one two-dimensional complex vector of size 4×N. Temporary vectors include a complex twiddle array of size N and two float vectors of size N. A small number of local variables are also needed to store values for individual parameters.

Because of the verification, a class called TVals was created to encapsulate the minimum, typical, maximum, and measured parameter values. The TVals class is nested within the TMeasDev and is isolated from the

```
class TMeasDev {  // highly truncated class
private:  // User declarations
   String ImportedFileName;
   vector<Float> RawFreq, Time, WorkFreq;
   vector<Complex> Delay, WorkSParams,
                   Impulse, SmothedMag;
   vector< vector<Complex> > RawSParams;
public:  // User declarations
   TMeasDev();  // constructor
   void SetImportedFileName(String& fn);
   String GetImportedFileName();
   const vector<Float>& GetRawFreq();
   const vector<Float>& GetTime();
   const vector<Float>& GetWorkFreq();
   const vector<Complex>& GetDelay();
   const vector<Complex>& GetWorkSParams();
   const vector<Complex>& GetImpulse();
   const vector< vector<Complex> >&
                          GetRawSParams();
   bool ImportRawData(ifstream& infile);
};
```

Figure 1: TMeasDev Class Fragment

programmer. In order to model such a device, arrays are used to store raw data as well as processed data. None of these arrays are directly accessible to the user and require the user of class member function to access or modify them. *Figure 1* shows a fragment of the TMeasDev class, which holds the data. The ImportRawData() function is used to fill the internal arrays and the various "Get" functions are used to access the data for analysis or plotting.

3. PERFORMANCE DATA EXTRACTION

The current system uses measured S-parameter data and calculates derived performance parameters. S-Parameter data is read from an ASCII text file in any format be it real and imaginary, linear or logarithmic, degrees or radians. None of the user specification data is used in determining the measured parameters, meaning that SawVerify will extract parameters strictly from raw data. Measured parameters of interest include bandwidth, center frequency, lower and upper bandedges, insertion loss at center frequency, minimum insertion loss, magnitude slope, magnitude ripple, center frequency delay, phase ripple and phase maximum deviation from linear. Also the time domain response is calculated and presented.

If SawVerify is presented with raw S-parameter data that has significant RF feedthrough and needs to be windowed, then the user has the ability to time gate the data to remove the RF feedthrough and/or window it using select window types. *Figure 2* shows the S21 raw magnitude response of Carter's 13 chip uniformly weighted stepped chirp code SAW device [2].

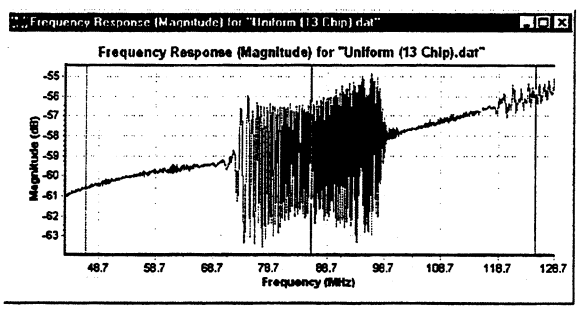


***Figure 2:** Raw Magnitude Data*

By applying the Hanning window and time gating the raw data, the response shown in *Figure 3* is obtained.

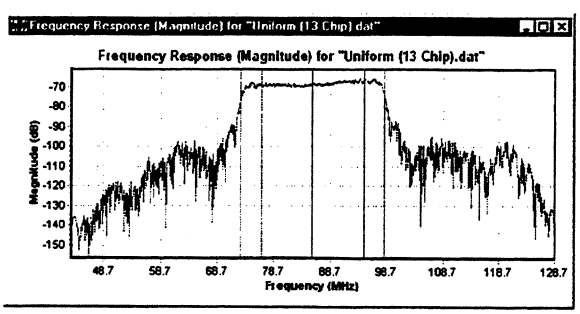


***Figure 3:** Time Gated and Hanning Window*

All of the data in this paper refers to Carter's 13 chip uniformly weighted stepped chirp code SAW device. This device displays the properties of a high loss bandpass SAW filter with high RF feedthrough and was perfect to tax the techniques presented in this paper.

3.1 First Guess for Upper and Lower Bandedges

Initial investigation found the process of calculating the upper and lower bandedges to be more difficult than expected. While it is easy for a human to visually detect the bandpass region, it is not trivial to code an algorithm to do the same, especially with high loss data and high level of RF feedthrough. However, a simple technique, called the top hat method, performs a least-squared error calculation of a pulse function (or top hat) with the magnitude data. This algorithm correlates all possible pulse widths and positions with the data. *Figure 4* shows simplified magnitude data, X, with one possible top hat function, g.

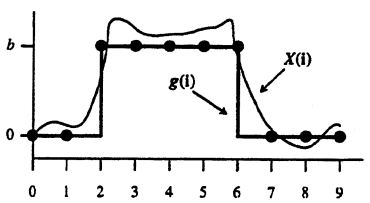


***Figure 4:** First Guess Bandpass Region Using Top Hat*

This technique simply takes the least-squared error of the vector representing the magnitude data, X, and the vector representing the top hat, g. *Equation (1)* shows the dot product of the difference between the two vectors that has to be minimized:

$$f = (X - bg)^T \cdot (X - bg). \tag{1}$$

By taking the partial derivative of f,

$$\frac{\partial f}{\partial b} = -2g^T X + 2bg^T g, \tag{2}$$

and setting it equal to zero, b is found to be

$$b = \frac{g^T X}{g^T g}. \tag{3}$$

Factoring *Equation (1)* and substituting *Equation (3)* for b, yields:

$$f = XX^T - \frac{(g^T X)^2}{g^T g}. \tag{4}$$

To minimize f, then

$$\frac{(g^T X)^2}{g^T g} \tag{5}$$

is maximized to reduce the error and thus finds the best first guess. In other words pick the top hat function, g, that returns the maximum value for *Equation (5)*.

An efficient implementation of *Equation (5)* uses a partial sum vector:

$$S_k = \sum_{i=0}^{N-1} X_i \,. \qquad (6)$$

Since $g^T X$ is the difference of partial sums and $g^T g$ is the top hat width, then *Equation (5)* can be rewritten as:

$$\frac{(S_j - S_i)^2}{j-i}, \qquad (7)$$

which is computationally efficient. The values for i and j from the maximum value of *Equation (7)* give a good first guess position of the lower and upper bandedges.

Once the top hat method determines the first guess upper and lower bandedges, a simple algorithm is used to find the 3 dB bandwidth, or the bandwidth at any level as specified by the user. The outer vertical lines on *Figure 3* show the 10 dB.

3.2 Other Parameters

Most other device parameters are easily extracted from this point on. For example, the center frequency delay is obtained by calculating the slope of the phase at center frequency. The center frequency delay calculated from the phase response of Carter's chirp SAW filter is shown as the vertical line in the middle of the time domain impulse response in *Figure 5*.

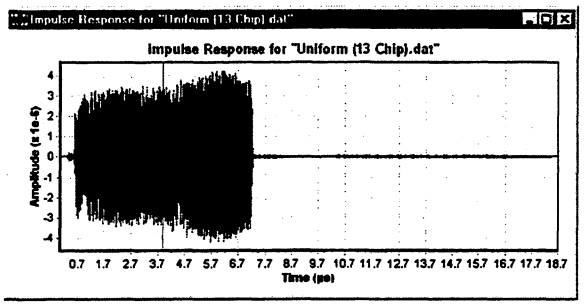


Figure 5: Time Domain Impulse Response

The most difficult parameter to be extracted was the magnitude slope. Many investigated algorithms have a very difficult time finding the local maximums of a "noisy" magnitude response. Wavelet and FFT smoothing techniques were investigated to determine if coefficient reduction of the magnitude response aided in extracting the local maximums of the bandpass region.

4. ALGORITHM COMPARISONS

A wavelet transform is similar to the Fourier transform in that it uses a series of orthogonal functions to transform a function to an alternative domain, as shown by the basic wavelet recursion equation [3] in *Equation (8)*:

$$\psi(t) = \sum_n h(n)\sqrt{2}\psi(2t-n)\,. \qquad (8)$$

Where the scaling coefficients are calculated as follows:

$$h(n) = \sqrt{2}\int \psi(t)\psi(2t-n)dt\,. \qquad (9)$$

Equation (8) and *Equation (9)* show how similar the wavelet transform is to the Fourier transform. The difference is that the wavelet transform uses finite basis functions as opposed to the infinitely oscillating sinusoidal functions used in the Fourier transform [4]. The wavelet transform breaks a function into sections of scale. By removing all but the coefficients that correspond to the high scale (low detail) representation of a function then data can be smoothed yet still retain the overall shape of the data. This is a simplistic method of denoising as presented by Burrus [3].

4.1 Wavelet Smoothing Techniques

There are two methods of coefficient reduction investigated, truncation and thresholding. Truncation is simply removing all but the first N wavelet coefficients. This retains the low detail information. The threshold method removes all but the N largest coefficients thus possibly retaining some higher detail information. *Figure 6* shows a portion of the Daubenchies-20 wavelet transform of Carter's SAW device. The vertical line represents the truncation point in which the first thirty-two wavelet coefficients are used. The horizontal line shows the threshold in which the thirty-two largest coefficients are kept. This uses some higher detail coefficients that the simple truncation did not use.

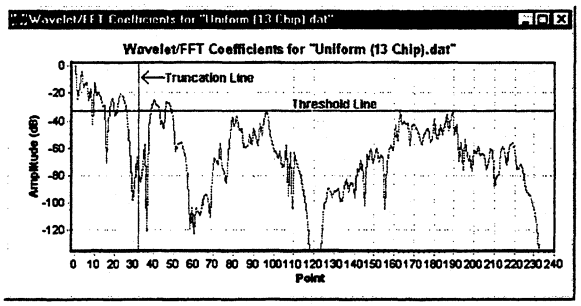


Figure 6: Carter's Daubechies-20 Wavelet Coefficients

4.2 Wavelet Smoothing Results

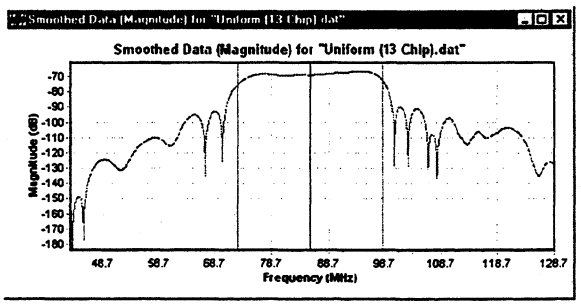


Figure 7: Smoothed, First 32 Daub-20 Wavelet Coef.

Figure 7 shows the smoothed magnitude response of Carter's chirp code SAW device. Smoothing was performed by transforming the magnitude response to the wavelet domain using the Daubenchies-20 wavelet transform [5]. Then all but the first thirty-two coefficients were zeroed. This truncation removed the

low scale (or high detail) wavelet coefficients. Finally the now truncated wavelet coefficients were transformed back to the magnitude domain.

Compare *Figure 3* with *Figure 7* to see that the general shape of the magnitude response is retained. This smoothed data is much easier to extract the local maximums so that magnitude slope can be calculated from the raw magnitude data.

The Daubenchies-20 wavelet is a very smooth wavelet, much like a damped $\sin(x)/x$ function. Truncating all but thirty-two coefficients using a lower order and "rough" wavelet basis function, such as the Daubenchies-4, results in a not so smoothed magnitude response as shown in *Figure 8*. This smoothed version of the magnitude response creates artificial inflection points because the basis function is very rough. This is not a desired result.

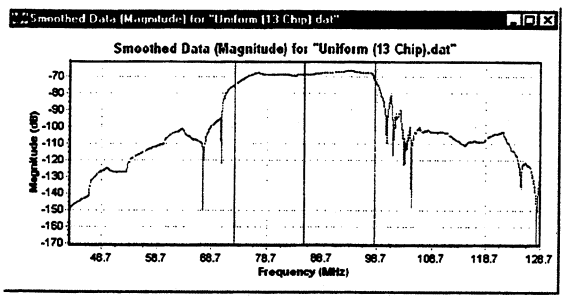


Figure 8: Smoothed, First 32 Daub-4 Wavelet Coef.

The same truncation method was applied using a standard FFT algorithm [6]. The results were very similar to the Daubenchies-20 but faired a little better as shown in *Figure 9*. After testing the wavelet and FFT techniques on various SAW devices from Carter as well as some commercially available devices, it was found that the FFT smoothing method using truncation was the best. Thresholding kept some high frequency or high detailed components that interfered with the algorithm that detected the local maximums of the edges of the bandpass region.

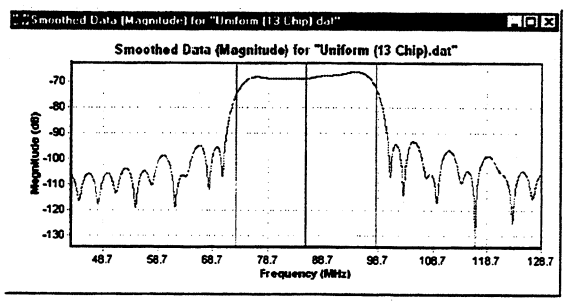


Figure 9: Smoothed, First 32 FFT Coefficients

Parameters from commercially produced SAW filters have been extracted and due to the low loss and low RF feedthrough, the parameter extraction did not require the FFT or wavelet coefficient reduction, as did the high loss filters. SawVerify by default does not use FFT or wavelet coefficient reduction if the minimum insertion loss is less than 40 dB.

5. VERIFICATION METHODS

Verification of a SAW device proved to be quite simple. Because the TVals class contained minimum, typical and maximum user constrains, it was very simple to compare these to the extracted values. For minimum and maximum values, the device parameters that caused the failure are identified "red". However, a percentage range is used to signify how good the typical value is. If it is within 1% then the value is good; if it is within 5% then the value is marked "yellow"; outside 5% the value is marked "red". *Figure 10* illustrates how this is presented to the user.

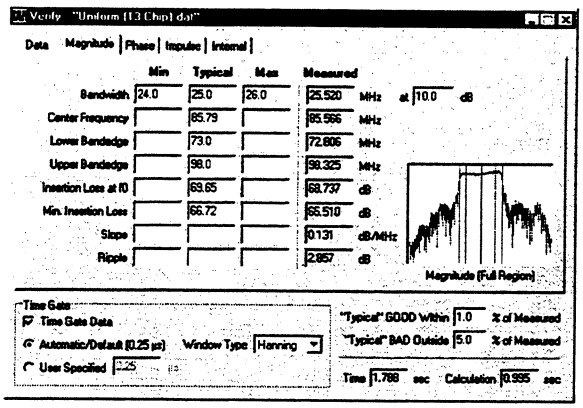


Figure 10: Magnitude Response Information

6. CONCLUSION

A stand-alone version of the design verification module for a SAW design automation system was presented. Methods for extracting device parameters were presented. The top hat method determined a good starting point for parameter extraction. FFT coefficient reduction using truncation proved to be very helpful with magnitude slope extraction. A simple method of presenting parameter failure was also presented.

7. REFERENCES

[1] M. J. McCollister and S. M. Richie, "The Development of a First Pass Verification Module for a SAW Filter Design Automation System", *IEEE Ultrasonics Symposium Proceedings*, vol. 1, pp. 133-136, 1991.

[2] S. E. Carter, "Development and SAW Device Implementation of a New Weighted Stepped Chirp Code Signal for Direct Sequence Spread Spectrum Communications Systems", Ph. D. Dissertation, University of Central Florida, Orlando, Florida, 1998.

[3] C. S. Burrus, R. A. Gopinath, and H. Guo, *Introduction to Wavelets and Wavelet Transforms: A Primer*, New Jersey: Prentice-Hall, ch. 10, pp. 205-211.

[4] M. V. Wickerhauser, *Adapted Wavelet Analysis from Theory to Software*, New York: IEEE Press, 1994.

[5] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*, Second Edition, Cambridge: Press Syndicate, 1992, ch. 13.10, pp. 591ff.

[6] J. G. Proakis and D. G. Manolakis, *Introduction to Digital Signal Processing*, NY: Macmillan, 1988, ch. 9, pp. 698-721.

THE EFFECT OF DIFFRACTION ON THE DISPERSIVE DELAY LINES PERFORMANCE

A. Milewski, S. Gawor

Tele and Radio Research Institute, Ratuszowa 11, 03-450 Warsaw, Poland

ABSTRACT

This paper describes the influence of the diffraction effect on the compressed signal sidelobes level. The analysis was made using Angular Spectrum of Waves (ASoW) model. This model is very flexible and can be used for various SAW devices. Calculations were carried out for many types of dispersive delay lines with one dispersive and one periodical transducer

The iterative synthesis technique for the correction of diffraction in SAW compressor will be described (it will be shown in the first part of the paper, that the influence of diffraction on the expander characteristic is usually negligible). Using the results, various pulse-compression and expansions filters were designed and manufactured. A very good agreement between calculations and measurements was observed.

1. INTRODUCTION

Surface acoustic wave (SAW) dispersive delay lines (DDL) are widely used as signal expanders and compressors. The main problem in the design of such devices is to achieve desired degree of the compressed signal sidelobes suppression. Linear frequency modulated signals are used most commonly and the sidelobes suppression is obtained by amplitude weighting of the compressor. For instance when Hamming weighting is used the theoretical sidelobe level is -42dB below the main peek. But in practice it is very difficult to achieve such good results. It is caused by second order effects such as: SAW diffraction and reflection, bulk wave propagation, direct electromagnetic signal transmission and other. The effect, which has most significant influence on DDL parameters, is the diffraction of SAW propagating between transducers. The method of analysis and design of dispersive delay lines with correction for diffraction effect will be presented.

2. ANALYSIS OF THE SAW DISPERSIVE DELAY LINES

In the simplest model of SAW devices it is usually assumed that surface acoustic wave propagating between transducers is a plain wave. In fact, the amplitude and phase of the wave are significantly distorted due to limited aperture of the lunching transducer (fig. 1). The results were obtained using formulas presented in [1]. The distortion level depends not only on the distance from the transducer and its aperture, but also on the anisotropy of SAW velocity. The anisotropy of SAW velocity may be described using parabolic approximation [2]:

$$\frac{v(\theta)}{v_0} = 1 + \frac{\gamma}{2}\theta^2 \qquad (1)$$

where:

v - SAW velocity,

θ - misorientation angle,

γ - anisotropy parameter.

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

The anisotropy is different for various piezoelectric material and crystallographic orientation. For instance if $LiNbO_3$ YZ substrate is used the diffraction effect is very low. The effect of diffraction has significant influence especially for DDL compressors, because of amplitude weighting and small electrode overlaps.

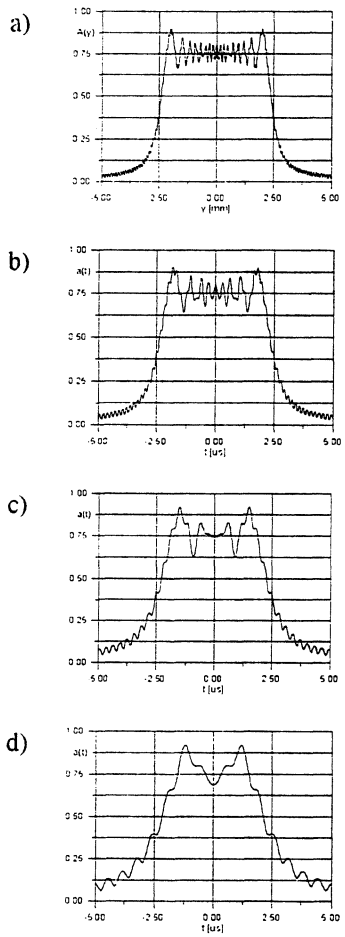


Fig. 1. Amplitude distribution of SAW propagating between transducers. Aperture of the transducer is 5mm. Substrate material is quartz YX. Distance from the transducer: a) 5, b) 10, c) 20, d) 40mm.

The diffraction model for SAW filters was presented in [3]. After few modifications it can be used for the analysis of the dispersive delay lines with nonperiodic transducers. The example of the SAW delay line to by analysed is presented in fig. 2.

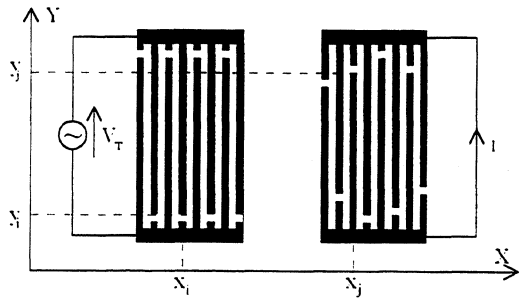


Fig. 2. Analysed SAW delay line structure.

In this model the transadmittance of the delay line can be described using formula

$$Y_{21}(f) = -\frac{\sqrt{2}}{2} 2\pi f \Gamma_s e^{j\frac{\pi}{4}\mathrm{sgn}(1+\gamma)} \cdot \sum_{i=1}^{N^{(1)}-1} \sum_{j=1}^{N^{(2)}-1} \overline{\sigma}_{gi}^{(1)}(k_p)\overline{\sigma}_{gj}^{(2)}(k_p) q_{ij} e^{-jk_0 x_{ij}} \cdot$$
$$\cdot \left[H(A_{ij},1+\gamma) + H(D_{ij},1+\gamma) - H(E_{ij},1+\gamma) - H(F_{ij},1+\gamma)\right] \quad (2)$$

where:

$$q_{ij} = \sqrt{\frac{\pi x_{ij}|1+\gamma|}{k_0}} \quad (3)$$

$$x_{ij} = x_j - x_i \quad (4)$$

$$A_{ij} = \frac{y_{j+1}^{(2)} - y_i^{(1)}}{q_{ij}} \quad D_{ij} = \frac{y_j^{(2)} - y_{i+1}^{(1)}}{q_{ij}}$$

$$E_{ij} = \frac{y_{j+1}^{(2)} - y_{i+1}^{(1)}}{q_{ij}} \quad F_{ij} = \frac{y_j^{(2)} - y_i^{(1)}}{q_{ij}} \quad (5)$$

$$H(x,d) = xE(x,d) + \frac{1}{\pi} e^{-j\frac{\pi}{2}\mathrm{sgn}(d)(x^2+1)} \quad (6)$$

$$E(x,d) = C(x) - j\,\mathrm{sgn}(d) S(x) \quad (7)$$

$i=1..N^{(1)}$, $j=1..N^{(2)}$ - electrode number,

f - frequency,

k_0 - wavenumber,

Γ_s - constant defined in [4],

$\overline{\sigma}_g$ - Fourier transform of the charge density,

C(x), S(x) - Fresnel integrals.

This expression can be used for the analysis of dispersive delay lines with one or two dispersive transducers. To reduce the double sum in equation 2, in a case of DDL with one dispersive transducer, the periodical unweighted transducer may be replaced by a single line source. It leads to reduction of the calculation time without considerable worsening of the accuracy.

3. CALCULATIONS

To evaluate the influence of the diffraction effect on the dispersive delay line performance some computer simulations were made. The assumed DDL parameters are as follow: center frequency 70MHz, bandwidth 6MHz, dispersion time 10μs, acoustic aperture 5mm, distance between transducers 6mm, piezoelectric substrate quartz YX. To reduce the compressed signal sidelobes the compressor amplitude was weighted using Hamming function. Frequency characteristics of DDL expander and compressor were calculated using presented above model. For comparison, calculations using simplified delta function model omitting diffraction effects were also made. For the expander the results are almost identical. Only small differences can be observed for the compressor fig. 2.

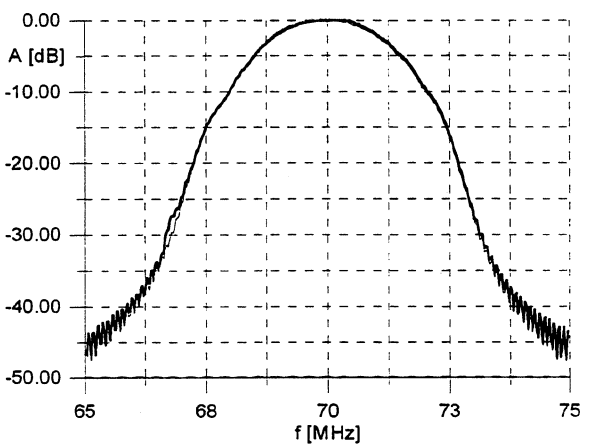


Fig. 2. Frequency characteristics of the compressor calculated including diffraction effect (bold line) and without diffraction (hairline).

Although the influence of diffraction on the DDL characteristics seams to be insignificant it results in considerable increase of sidelobes level in the compressed signal (fig. 3).

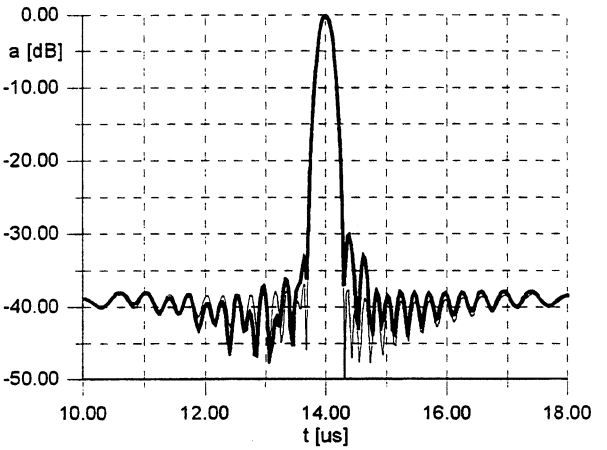


Fig. 3. Signal after compression calculated including diffraction effect (bold line) and without diffraction (hairline).

To evaluate the influence of diffraction on compressed signal sidelobes level, some calculations were made for series of DDL with the dispersion time in the range from 2.5 to 20μs. The other parameters remained unaltered. The dependence of the sidelobes level on the dispersion time is presented in fig. 4.

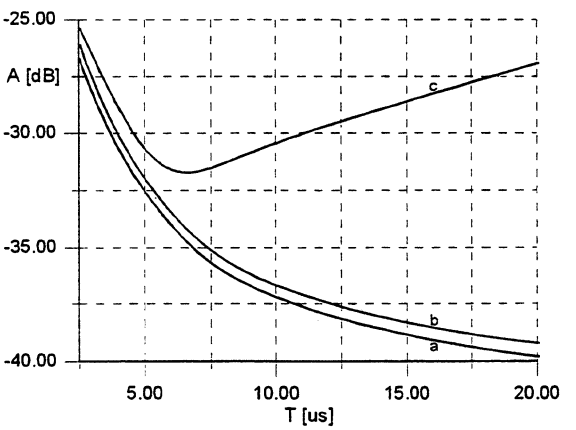


Fig. 4. Dependence of the sidelobes level on the dispersion time: a) calculated without diffraction, b) calculated including diffraction effect in expander, c) calculated including diffraction effect in compressor.

The increase of sidelobes level for small values of the dispersion time is caused by amplitude ripple of the expander characteristic and can be reduced using

reciprocal ripple design technique. If diffraction effect is omitted (curve a) sidelobes level decreases with increase of dispersion time and tends to its theoretical limit (–42dB for Hamming weighting). Including the diffraction effect in the expander analysis (curve b) results in small worsening of the sidelobes level. The influence of diffraction effect is more significant for DDL compressor (curve c).

4. EXPERIMENTS

Using the presented above results a pair of dispersive delay line was designed. The parameters of the lines are as follow: centre frequency 70MHz, dispersion time 10μs, bandwidth 6MHz. To simplify the design process the diffraction effect in the expander was omitted. The effect of diffraction in the compressor was reduced using iterative method. After three iterations –35dB sidelobes level was obtained (fig. 5). The results are in very good agreement with the measurements fig. 6.

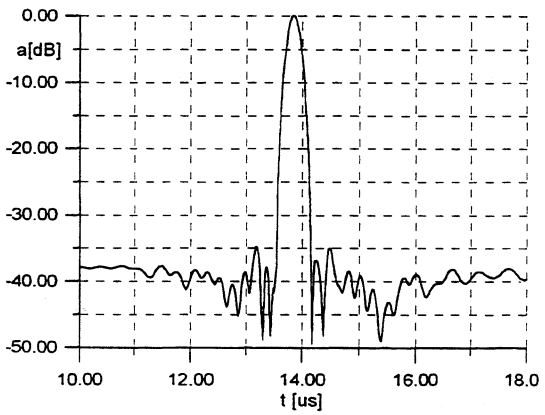


Fig. 5. Signal after compression calculated for the designed dispersive delay lines.

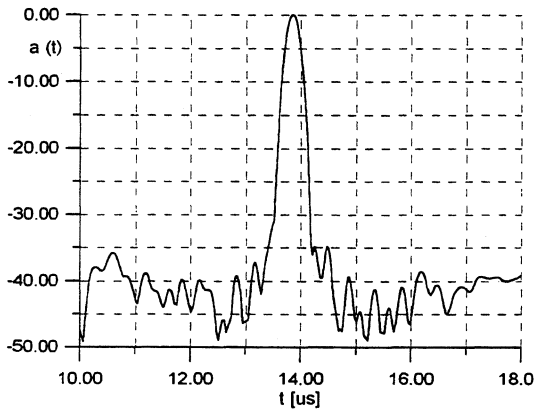


Fig. 4. Measured signal after compression.

4. CONCLUSION

The mathematical method of SAW diffraction analysis was presented. The results are in good agreement with theory. Using the model, a significant improvement in sidelobes suppression may be obtained.

REFERENCES

1. J. H. Hines, B. P. Abbott, M. A. West, D. C. Malocha, A new approach to broadband diffraction theory, Ultrasonics Symposium, 47, (1988).
2. T. Szabo, A. Slobodnik, The effect of diffraction on the design of acoustic surface wave devices, IEEE Trans. on Sonics and Ultrasonics, SU-20, 240 (1973).
3. P. Nagłowski, Kompensacja efektów drugiego rzędu w filtrach z akustyczną falą powierzchniową przy użyciu metod optymalizacji z ograniczeniami, Ph.D. Thesis, Warsaw University of Tech., 1991.
4. C. S. Hartmann, S. Jen, M. A. Domalewski, J. C. Andle, Improved accuracy for determining SAW transducer capacitance and k^2, IEEE Ultrasonics Symposium, 161, (1987).

DESIGN OF SAW FILTERS WITH TWO APODISED IN-LINE TRANSDUCERS ON QUARTZ

J. Renger and B. Wall
Vectron International, Potsdamer Straße 18, D-14513 Teltow, Germany

ABSTRACT

The design of SAW filters on quartz consisting of two apodised in-line interdigital transducers (IDT) without using multistrip coupler is presented. The analytical model is described by taking into account two-dimensional static charge distribution, surface wave diffraction and the influence of the external network. A wideband diffraction compensation method is applied to compensate these second-order effects. The design optimisation criteria of the two apodised in-line IDTs are presented. The theoretical and experimental frequency responses of a filter are shown.

1. INTRODUCTION

A standard design principle is the usage of SAW filters consisting of one apodised and one unapodised IDT. The apodised IDT provides the main contribution to the overall frequency response. But the application of two apodised IDTs has the following advantages. Both IDTs provides similar contributions to the filter response. The admittances of both IDTs are similar. The optimisation of the filter design has additional variables. For highly coupling materials a multistrip coupler (MSC) can be used between both apodised IDTs [1]. A disadvantage of the MSC is the twice as big chip area.

A design method for doubly apodised SAW filter without MSC has been published by Motsoela, et al. [2]. The method [2] does not consider important second order effects, such as diffraction and 2D static charge distribution.

Our paper presents a method for the design of two apodised in-line IDTs on quartz with compensation of second order effects.

2. THEORY

2.1 SAW Filter Analysis Method:

The SAW filter considered here consists of two apodised in-line transducers, the input transducer IDT1 and the output transducer IDT2 (Fig. 1). The gap positions Y_k of the fingers are defined in the figure 1. The design of this filter requires highly accurate simulation to compensate for second-order effects. Our analysis method considers the following effects

- 2D static charge distribution
- surface wave diffraction
- ohmic loss in the bus-bars
- electrical loading effects.

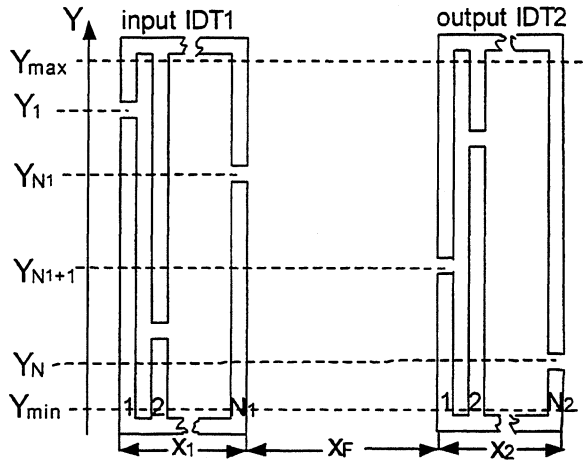


Fig. 1. SAW device with two apodised in-line IDTs

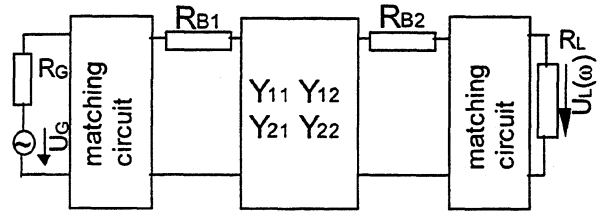


Fig. 2. Equivalent circuit model for the analysis

Our equivalent circuit model for SAW filters with matching circuit (Fig. 2) uses the admittance matrix (Y-parameters) of the filter. R_{B1} and R_{B2} represent the bus-bar impedance of the IDTs to consider ohmic losses of bus bars. The frequency response of the equivalent circuit model is

$$\overline{H}(\omega) = \frac{\overline{U}_L(\omega)}{U_G}. \qquad (1)$$

$\overline{U}_L(\omega)$ is the complex amplitude of the voltage across the load impedance R_L and U_G is the voltage of the generator.

2D-Static Charge Distribution:

The charge distribution of the IDT can be regarded as the source of the generation of SAW. For narrow finger overlaps the exact knowledge of the 2D-charge distribution is necessary. The 2D static charge distribution of a few fingers is represented in Fig. 3. Our analysis of the charge distribution is based on the method of the Green's function of Baghai-Wadji et al. [4]. The method requires a large number of discretisation intervals in the direction of the finger length (y-direction). About 40 non-equidistant intervals in the y-direction are necessary for a good conformance between simulation and measurement result.

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

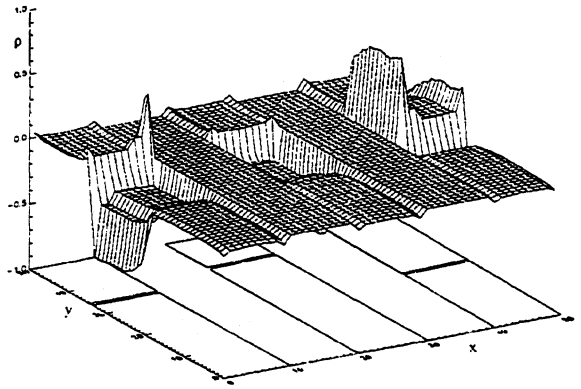


Fig. 3. Charge distribution of a few fingers (normalised)

Admittance Matrix:

Visintini et al. [3] formulated the admittance parameters for the case of 2D SAW propagation. These formulas consider the 2D-charge distribution of the IDT for the generation and detection. Also, the anisotropy of the coupling factor and the diffraction are considered for the SAW propagation. The transfer admittance between the two apodised IDTs corresponds to [3]

$$\overline{Y}_{21}(\omega)=\frac{\omega}{2\pi}\int_{k_{oy}^{\min}}^{k_{oy}^{\max}}K(k_y)\overline{\mathrm{P}}_1(k_y)\overline{\mathrm{P}}_2^{*}(k_y)$$
$$\cdot\exp\left\{-j\left[k_{tx}(k_y)\frac{x_1+x_2}{2}+k_{ox}(k_y)x_F\right]\right\}dk_y. \quad (2)$$

The wave vector components k_{tx} (k_y) and $k_{ox}(k_y)$ are determined by the corresponding slowness curves in both transducer regions and the free surface region, respectively. The input IDT is of length x_1, the output IDT is of length x_2 and x_F is the free surface region between the two transducers. The integration path in (2) is limited to the borders of the slowness curve of the free surface region. A region of ±20° is sufficient for the integration in practice. The expression $K(k_y)$ is the following function

$$K(k_y)=\frac{1}{\varepsilon_0\varepsilon_{\mathrm{eff}}^{\infty}(k_y)}\cdot\frac{k_m^2(k_y)-k_o^2(k_y)}{2k_o(k_y)k_{tx}(k_y)}. \quad (3)$$

$K_m(k_y)$ and $\varepsilon_{\mathrm{eff}}^{\infty}(k_y)$ are the wave vectors on a metalised surface and the effective dielectric constant [3], respectively.

The FT of the static charge density $\sigma_i(x,y)$ of IDTi is

$$\overline{\mathrm{P}}_i(k_y)=\int_{-\infty-\infty}^{+\infty+\infty}\rho_i(x,y)\cdot\exp\{j[k_{tx}(k_y)x+k_y y]\}dxdy \quad (4)$$

for i=1,2.

$\rho_i(x,y)$ is the static charge density of the IDTi, when the unit potential is applied across the transducer. The geometrical reference point is the centre of the IDT. The admittance of the transducer is the following sum

$$Y_{ii}(\omega)=G_{a_i}(\omega)+j\left(\omega C_{T_i}+B_{a_i}(\omega)\right) \quad (5)$$

for i=1,2.

According to [3], the SAW radiation conductance is given by the equation

$$G_{a_i}(\omega)=\frac{\omega}{\pi}\int_{k_{oy}^{\min}}^{k_{oy}^{\max}}K(k_y)\left|\overline{P}_i(k_y)\right|^2dk_y \quad (6)$$

for i=1,2.

The susceptance is computed as the Hilbert transform of $G_{a_i}(\omega)$

$$B_{a_i}(\omega)=-\frac{1}{\pi}G_{a_i}(\omega)*\frac{1}{\omega} \quad (7)$$

for i=1,2.

The transducer capacitance is the sum of charges on the transducer

$$C_{T_i}=\frac{1}{2}\int_{-\infty-\infty}^{+\infty+\infty}\left|\rho_i(x,y)\right|dxdy \quad (8)$$

for i=1,2.

2.2 Compensation Method:

The frequency response of two apodised in-line IDTs is not the product of the two individual frequency responses. Therefore, the compensation of this distortion of the frequency response and the other four second-order effects is necessary. The basis of our compensation method is the wideband compensation algorithm by Naglowski [5]. F. Z. Bi [6] investigated this method as well to compensate the diffraction of TV SAW filters. This compensation method is an optimisation task. The frequency response for the optimisation is the following linear equation

$$\overline{H}_O(\omega_i,\underline{Y},\underline{\Delta Y})=\overline{H}_A(\omega_i,\underline{Y})+\sum_{k=1}^{N}\overline{C}_k(\omega_i,\underline{Y})$$
$$\cdot\Delta Y_k, \quad (9)$$

where

i=1,...,N_f frequency point no.,

ω_i grid of radian frequency,

N=N_1+N_2 sum of finger numbers (IDT1 and IDT2),

$\underline{Y}=\left[Y_1,\ldots,Y_{N_1},Y_{N_1+1},\ldots,Y_N\right]$ vector of gap positions in both apodised IDTs (before optimisation),

Y_k position of gap in finger no. k (see Fig. 1),

$\underline{\Delta Y}=\left[\Delta Y_1,\ldots,\Delta Y_{N_1},\Delta Y_{N_1+1},\ldots,\Delta Y_N\right]$ vector of position displacements (optimisation variable),

ΔY_k displacement of gap position in finger no. k,

$\overline{H}_A(\omega_i,\underline{Y})$ analysed complex frequency response (including second-order effects, normalised) and

$$\overline{C}_k(\omega_i,\underline{Y}) = \frac{\partial \overline{H}_A(\omega_i,\underline{Y})}{\partial Y_k} \tag{10}$$

complex first order derivatives (including second-order effects).

The goal of the optimisation is to minimise the following quadratic error sum

$$\varepsilon(\underline{\Delta Y}) = \sum_{i=1}^{N_f} \left|\overline{E}(\omega_i,\underline{Y},\underline{\Delta Y})\right|^2 \Rightarrow \min, \tag{11}$$

with the constraints for the finger gaps

$$Y_{min} \le Y_k + \Delta Y_k \le Y_{max} \tag{12}$$

for $k=1,\ldots,N$.

Y_{min} and Y_{max} are the lower and upper borders of the gap positions (Fig. 1), respectively. Both values are chosen in consideration of the IDT aperture. The complex error function in (11) is

$$\overline{E}(\omega_i,\underline{Y},\underline{\Delta Y}) = W_R(\omega_i)\left(H_{OR}(\omega,\underline{Y},\underline{\Delta Y}) - D(\omega_i)\right) + jW_I(\omega_i)H_{OI}(\omega,\underline{Y},\underline{\Delta Y}). \tag{13}$$

$H_{OR}(\omega_i,\underline{Y},\underline{\Delta Y})$ and $H_{OI}(\omega_i,\underline{Y},\underline{\Delta Y})$ are the real and imaginary parts of $\overline{H}_O(\omega_i,\underline{Y},\underline{\Delta Y})$, respectively.

The desired frequency response $D(\omega_i)$ has real values for filters with linear phase response. $W_R(\omega_i)$ and $W_I(\omega_i)$ are the weighing functions for the real and imaginary part of the error function, respectively. $D(\omega_i)$ and $W_R(\omega_i)$ depend on the allowed tolerances of the frequency response. $W_I(\omega_i)$ is also dependent on the allowed tolerance of the linear phase response.

The optimisation method needs a start vector $\underline{Y}$ of the gap positions at the beginning. We calculated the initial gaps of both IDTs with the REMEZ algorithm. The new gap positions after one optimisation pass is

$$\underline{Y} := \underline{Y} + \underline{\Delta Y}. \tag{14}$$

A few optimisation passes are necessary for a successful design of the filter.

We used a gradient method (Quasi-Newton Algorithm) for the optimisation. We have written a computer program for the compensation method in FORTRAN. The program is implemented on a HP9000/C160 machine.

3. RESULTS

The following results concern a SAW filter with a centre frequency of 85.38 MHz and an aperture of 2000 µm. It was constructed on ST,X-Quartz. The filter consists of two apodised transducers with 1006 (IDT1) and 606 (IDT2) λ/8 fingers. The matching circuit of both IDTs is a L-C-network. The REMEZ algorithm was used to find the initial apodisation for both single transducers. Fig. 4a and 5a illustrate the initial finger gaps of both IDTs after the REMEZ algorithm.

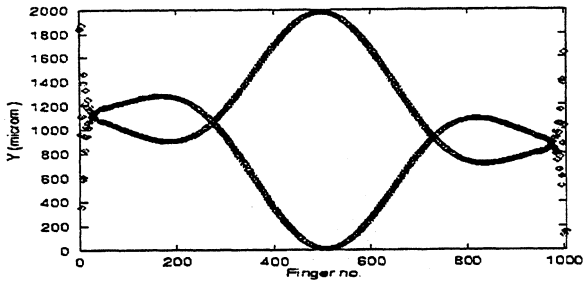


Fig. 4a. Finger gaps of input IDT1 before compensation

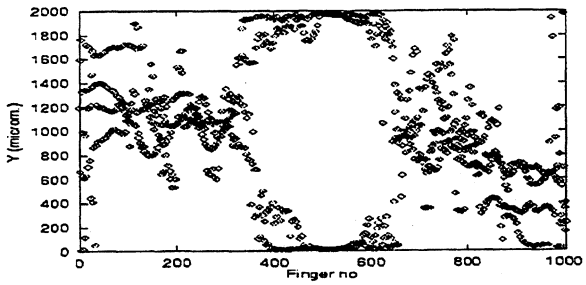


Fig. 4b. Finger gaps of input IDT1 after compensation

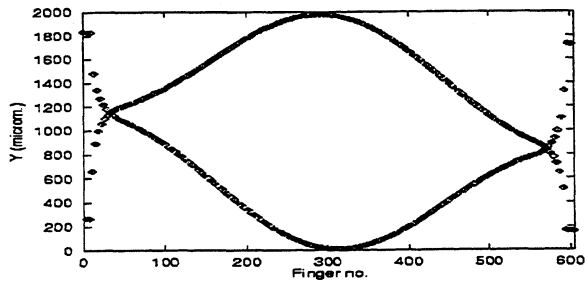


Fig. 5a. Finger gaps of output IDT2 before compensation

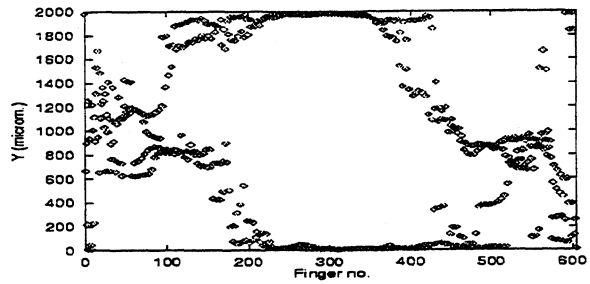


Fig. 5b. Finger gaps of output IDT2 after compensation

The analysed initial frequency response is represented in Fig. 6. The distortions result mainly from the apodisation of both IDTs, diffraction and charge distribution. About 20 compensation passes were necessary to remove all the distortions. The finger gaps of the IDTs after the compensation are showed in Fig. 4b and 5b. The analysed frequency response after compensation is represented in Fig. 6 as well.

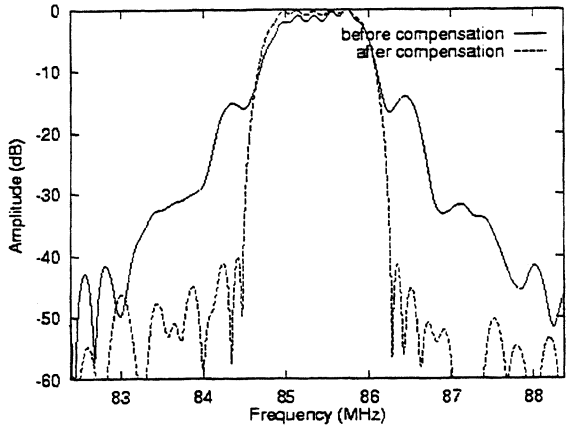


Fig. 6. Theoretical frequency responses before and after compensation

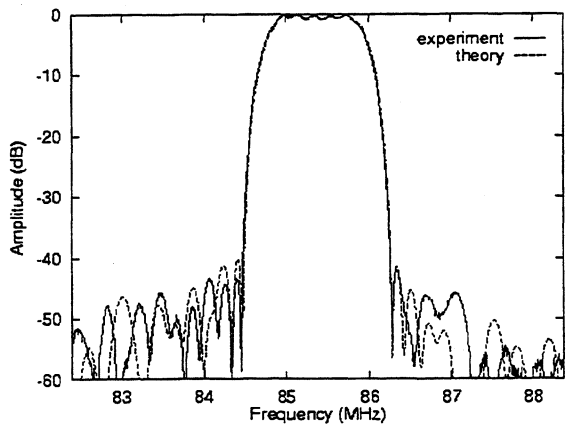


Fig. 7. Experimental and theoretical frequency responses after compensation

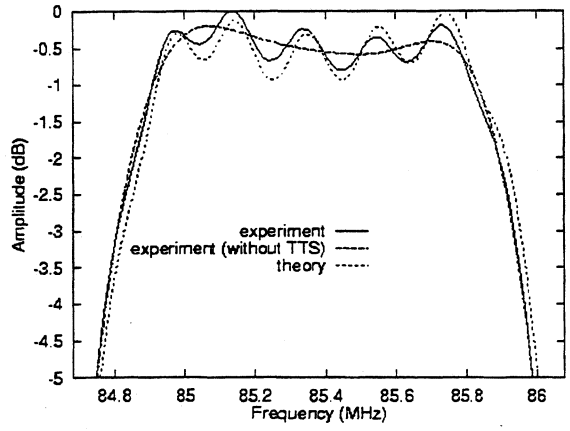


Fig. 8. Experimental and theoretical frequency responses after compensation

The sidelobe suppression on the stop band is increased from only 15 dB to about 40 dB after the compensation. Fig. 7 shows the experimental and analysed frequency responses of the compensated filter. The insertion loss of the filter is about 14 dB. The triple transit signal (TTS) causes a pass band ripple of 0.6 dB with a period of 200 kHz (Fig. 8). The theoretical response corresponds with the experimental response of the filter.

4. CONCLUSION

A method for the design of SAW filter consisting of two apodised in-line IDTs has been described The apodisation of both transducer, charge distribution and diffraction affects the frequency response very strong. The compensation of these effects results in a suitable filter design. By it, the theoretical and experimental responses are corresponding.

5. ACKNOWLEDGEMENTS

The authors wish to thank Dr. M. Weihnacht at IFW Dresden for the presentation of the material parameters and support.

6. REFERENCES

[1] D. Morgan, Surface Acoustic Wave Devices for Signal Processing, Elsevier, 1985

[2] T. Motsoela and P. M. Smith, "Doubly apodized SAW filters," Proc. IEEE Ultrason. Symp., 1995, pp. 95-98

[3] G. Visintini, A. R. Baghai-Wadji and O. Männer, "Modular two-dimensional analysis of surface acoustic wave filters-Part I: theory," IEEE Trans. Ultrason., Ferroelect., Freq. Contr., vol. UFFC-39, pp. 61-72, Jan. 1992

[4] H. Bachl and A. R. Baghai Wadji, "3D electrostatic field analysis of periodic two-dimensional SAW transducers with closed-form formulae," Proc. IEEE Ultrason. Symp., 1989, pp. 359-362

[5] P. J. Naglowski, "Iterative method using constrained optimization techniques for second-order effects compensation in SAW filters," Proc. IEEE Ultrason. Symp., 1989, pp. 125-128

[6] F. Z. Bi and K. Hansen, "Wideband diffraction compensation methods and the angular spectrum of waves (ASoW) model for TV SAW filters," IEEE Trans. Ultrason., Ferroelect., Freq. Contr., vol. UFFC-44, pp. 925-934, July 1997

1999 Joint Meeting EFTF - IEEE IFCS

ACOUSTOOPTIC METHOD FOR MEASUREMENTS OF SAW STEERING BEAM ANGLE AND SAW VELOCITY IN AN ANISOTROPIC SOLID

A.V. Tsarev, Member, IEEE, **E.A. Kolosovsky**, Member, IEEE

Institute of Semiconductor Physics of Siberian Branch of Russian Academy of Sciences, Novosibirsk, 630090 Russia, tsarev@thermo.isp.nsc.ru

Abstract. The application of two optical beams method for measurement of steering beam angle and velocity of surface acoustic wave (SAW) propagated in anisotropic solid is considered. Method is based on results of theoretical and experimental investigation of a simultaneous two light beams diffraction by a SAW in crystals. On a base of carrying out analysis update of measuring technique, algorithm of data processing and a software for the determination of a SAW steering beam angle and also phase and group SAW velocities measurement are carried out. Experimental measurement of steering beam angle for the SAW propagated at 74.8^0 angle with respect to X-axis of $LiNbO_3$ shows good agreement with calculated data.

1. INTRODUCTION

Results of measuring of SAW velocities for different directions of SAW propagation and cuts of crystals are often used for determination of elastic constants of crystals [1]. Reserves of improving accuracy of constants measuring can be find by further development of measuring techniques and by addition in calculated algorithms data's of independent measuring of steering beam angle. Here we describe new measuring algorithm to determine steering beam angle of SAW by acoustooptic interferometric method utilizing two light beams scheme [2].

2. DESCRIPTION OF ACOUSTOOPTIC MEASUREMENTS BY TWO LIGHT BEAMS INTERFEROMETRER SCHEME

Two beams optical interferometer [2, 3] is formed by two parallel semitransparent and opaque mirrors aligned at angle 45^0 respectively to the plane of crystal (see. Fig.1). Let consider coordinate system (X, Y, Z) places at the center of interdigital transducer (IDT) and regard X axis tilts at angle Ψ_0 with respect to crystallographic X' axis of Y-cut $LiNbO_3$. SAW is excited by IDT and propagates along angle $\Delta\Psi$ with respect to X axis on surface of $LiNbO_3$ covered by Al film of thickness 10 nm, were $\Delta\Psi$ is beam steering angle of SAW. IDT has an aperture 0.25 cm and consists of 10 finger-pairs with the length period of 16 mkm.

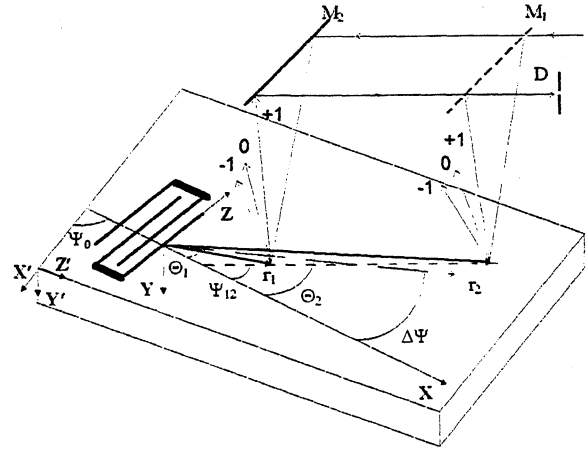


Fig.1. Geometry of two optical beam diffraction.
M_1 - semi-transparent mirror, M_2 - opaque mirror;
0, +1, -1 - are diffracted beams of different orders,
$\Delta\Psi$ - is the beam steering angle of SAW;
Ψ_{12} - is the angle, specifying direction of an interferometer axes;
Θ_l - is the angle, specifying direction at the center of the l-th optical beam;
r_l - is the distance from center of IDT to the center of l-th optical beam.

In the given scheme of experiment two spatially separated parallel optical beam of 1 mm beam width are incident at normal angle to the crystal surface, diffract by the SAW and mix up by interferometer to produce an interference pattern which oscillates with switching of SAW frequency [3]. Oscillations of interferenced (+1) diffracted orders are measured (see Fig.2) by intensity of light passed through the thin split diaphragm (D). SAW steering beam angle as well as the phase and the group velocity of SAW can be determined [2] by the measurement of a period of these oscillations on SAW frequency change.

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

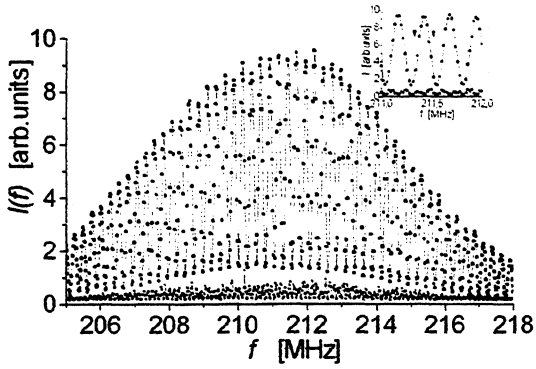


Fig.2. The experimental dependence of diffraction efficiency $I(f)$ on the SAW frequency. • - experimental points, * - noise of the interferometer.

Resolving the task of two light beams diffraction on reflections from SAW, propagating in anisotropic half-space it is possible to show, that in paraxial approximation the diffraction intensity in (±1) diffracted orders is described by expression [2]:

$$I=<|E_1+E_2|^2>=I_1+I_2+2(I_1 I_2)^{1/2}\cos(\gamma), \qquad (1)$$

were $\gamma=\Delta\phi+q_2-q_1$,

$q_l = K_{sl} r_l \cos(\Delta\Psi_l), \qquad l=1, 2,$

$K_{sl} = K(\Psi_{sl})$,

$\Psi_{sl} = \Theta_l - \Delta\Psi_l$,

$\Delta\Psi_l = -arctg(K'_{sl}/K_{sl})$,

Here $I_1$and I_2 are intensities of light diffraction [3] of (+1) orders of beams *1* and *2*, separately, $\Delta\phi$ is optical interferomer phase shift, r_l - is the distance from center of IDT to the center of l-th optical beam, Θ_l - is the angle, specifying direction at the center of the l-th optical beam; $K=2\pi f/v$ - SAW wave number, f and v are frequency and velocity of SAW.

Let consider a change of an oscillations period of $I(f)$ on SAW frequency by introducing a digitization frequency f_m according a condition [1]:

$$I(f_m) = I_1(f_m) + I_2(f_m). \qquad (2)$$

This condition is realized in a point of a maximum slope of $I(f)$ dependence, namely, when $\cos(\gamma)=0$. Then (2) is equivalent to a condition:

$$K_m\Delta x' = \pi M, \qquad m=0,+1,+2,..; \qquad (3)$$

where $K_m=K(f_m)$; $M=m+1/2-\Delta\phi/\pi$, $\Delta x'=\Delta x\bullet A_1 = \Delta r\cos(\Delta\Psi)\bullet A_2$ - *is* effective shoulders of optical interferometer. A_1 and A_2 - geometrical factors due to mismatch of an interferometer axis and direction of a group velocity, Δr is centre to centre distance of optical beams, $\Delta x=\Delta r\cos(\Psi_{12})$. $\Psi_{12}=arctg[(y_2-y_1)/\Delta r]$ is angle defining direction of an interferometer axes, i.e. direction of radius-vector connecting centers of light beams (see Fig.1). As usual all regarded angles are very small and it can be derived expressions for geometrical factors:

$$A_1 \cong 1+[x_1(\Delta\Psi-\Theta_1)^2 - x_2(\Delta\Psi-\Theta_2)^2]/(2\Delta x),$$

$$A_2 \cong 1+\Delta\Psi(\Delta\Psi-\Psi_{12}) - (\Psi_{12}-\Theta_2)(\Psi_{12}-\Theta_1)/2, \qquad (4)$$

From expression (3) it is evident that the experimental dependence f_m from order of number m is defined by the dispersion of SAW velocity. For weak or linear dispersion it is possible to reduce decomposition of a phase velocity on SAW frequency at the first linear term:

$$v \cong v_0(1+\alpha(f-f_0)) \qquad (5)$$

and to derive the dependence $f(M)$ in an obvious form:

$$f(M) = f_0 + (Mv_0 - 2f_0\Delta x)/(2\Delta x - \alpha v_0 M). \qquad (6)$$

The expression (6) defines a dispersion curve $\Omega(K)=2\pi f(\Delta xK/\pi)$ for SAW and make it possible to find its phase velocity and group velocities [3]:

$$v = \Omega/K = 2\Delta f\Delta x\,(1+\alpha(f-f_0))\,A_1,$$

$$v_g = \left|\partial\Omega/\partial K\right| = 2\Delta f\Delta r(1+\alpha(f-f_0))^2 A_2, \qquad (7)$$

where $\Delta f=df/dm$ $(f=f_0)$ is a half-period of oscillations on the SAW frequency, $\alpha=1/(2\Delta f^2)d^2f/dm^2$ $(f=f_0)$.

At first it seems, that geometrical factors A_1 and A_2 decreases acuracy of velocity measurement, but if to look through expression (4) and (7) one can find the possibility to determine beam steering angle $\Delta\Psi$ and to increase measuring acuracy of SAW velocity.

Experimental scheme (see Fig.1) manages to tune coordinate position (x, z) and angle orientation (Ψ_{12}) of optical interferometer. Mismatch of interferometer in X-direction is not sufficient on the velocity measurements. Let regard influence of transverse (Z) tuning and alignment of Ψ_{12} on measurement of group velocity. It can be shown from eq.(4) and (7) that the change of the group velocity has the form:

$$\delta v_g/v_g \cong \Delta\Psi(\Delta\Psi-\Psi_{12}) - x_1/x_2(\Delta\Psi+\Delta z_1-\Psi_{12})^2, \qquad (8)$$

where Δz_1 - is a mismatch of center of the first optical beam from the position corresponding a case of interferometer axis to be aligned along a vector of the group velocity ($\Psi_{12} = \Theta_1 = \Delta\Psi$).

It can be derived from eq. (8) that the slope of $\delta v_g^{max}/v_g$ dependence over Ψ_{12} (see Fig.3) determines a value of beam steering angle:

$$\Delta\Psi = -\partial(\delta v_g^{max}/v_g)/\partial\Psi_{12}. \quad (9)$$

From equations (1-9) it is evident that interference of diffracted orders is strongly dependent on the change of the interferometer phase shift $\Delta\phi$, caused by external conditions, for example, by the temperature deformation of the sample and the interferometer length, etc. Thus we propose to organize experimental measurements in those way that influence of interferometer phase shift $\Delta\phi$ will be self-compensated. Let change SAW frequency during measurement $I(f)$ by special jumping-frequency algorithm, namely,

$$f(n)=f_0+\delta f*(-1)^n \quad , \quad n=1, 2, 3, ...N, \quad (10)$$

ware δf is the frequency step of measuring, N is total number of measured points (~1000). It means that almost at the same time moment and when interferometer phase shift $\Delta\phi$ is not changed we measure two points of frequency response (see Fig. 2.) on opposite sides from the center frequency f_0. Thus influence on γ (see (1)) of the interferometer phase shift $\Delta\phi$ will have the opposite sign (with respect to frequency dependent terms q_i) for the even and odd points of measurement.

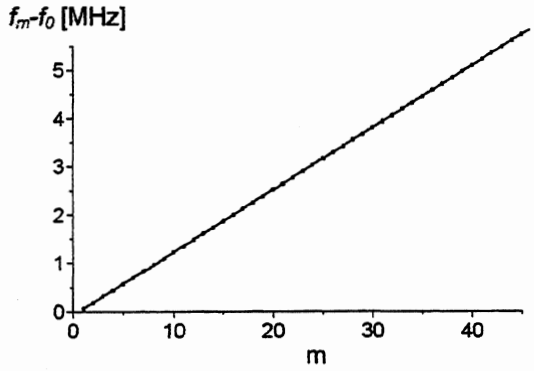


Fig.3. The experimental dependence of digitization frequency f_m.

Lets find digitization frequency $f_{\pm m}$ for frequencies $f>f_0$ and $f<f_0$ and plot dependences $f_m=(f_{+m}+f_{-m})/2$ and $\delta f_m=(f_{+m}-f_{-m})/2$ over m (see Fig.3 and Fig.4). Thus δf_m determines interferometer phase shift $\Delta\phi$ during the experiment and f_m determines according (1)-(9) SAW velocities almost independently on $\Delta\phi$. If it is needed one can add a correction to $f_m(m)$ by the fined $\Delta\phi(f_m)$ dependence.

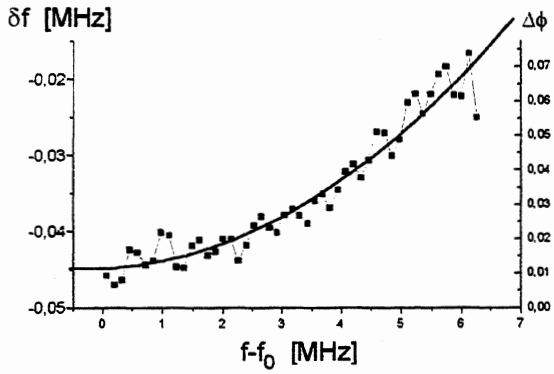


Fig.4. The experimental dependence of digitization frequency difference δf_m and interferometer phase shift $\Delta\phi$ at different measured frequencies $(f-f_0)$.

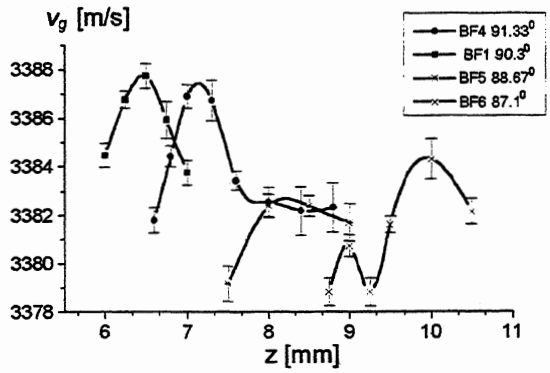


Fig.5. The experimental dependence v_g over the position (z_1) of the first optical beam at different orientation of the interferometer axis.

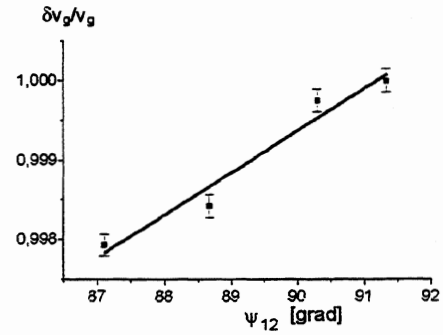


Fig.6. The experimental dependence $\Delta v_g/v_g$ maximum over the orientation of the interferometer axis (Ψ_{12}).

By this procedure we have measured frequency responses $I(f)$ and have determined dependences of group velocities v_g over Δz_1 at different orientations of interferometer axis Ψ_{12} (see Fig. 5). Than we

find extremum $\delta v_g^{max}/v_g$ and plotted it versus Ψ_{12} (see Fig.6). The slope of this dependence gives value $\Delta\Psi = 1.74^0 \pm 0.25^0$. On figure 6 are plotted experimental results and are shown computer calculations of the phase velocities and $\Delta\Psi$ for different SAW propagation directions on Y-cut $LiNbO_3$. It shows that calculated and measured results are in a good agreement with each other.

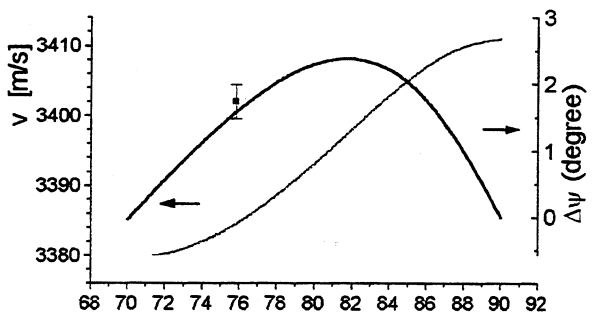


Fig.7. Dependence of velocity and steering angle for different direction of SAW propagation on Y-cut of $LiNbO_3$. Angle Ψ is done with respect to X axis. Experimental point is measured for $\Psi = 74.8^0$.

Accuracy of the of the beam steering angle measurement can be sufficiently improved by further optimization of experimental scheme, for example, (1) by better temperature stabilisation of sample and interferometer, (2) by optimisation of optical beam width respectively to IDT aperture, (3) by taking into account nonlinear dependency of velocity over direction of SAW propagation, etc.

Our analysis gives next recommendations for measurement procedure of the beam steering angle and also the phase and the group velocities of the SAW propagated in anisotropic mediums. Measurement by two optical beam interferometer have to consist of following steps:

- to measure dynamic change of interference partern intensity over change of SAW frequency by self-compensated wave or with taking into account possible change of interferometer phase shift during the measurement duration [2].
- to find the variation (δv_g) of the group SAW velocity dependencies on displacement of optical beams interferometer position (Δz) at different orientations (Ψ_{12}) of interferometer axis respectively to the direction of SAW propagation;
- to plot curves δv_g over Δy and to determine values of its maximum δv_g^{max} for different values of Ψ_{12}.
- to plot curves $\delta v_g^{max}/v_g$ over Ψ_{12} and to determine a beam steering angle ($\Delta\Psi$) by the slop of $\delta v_g^{max}/v_g$ dependence over Ψ_{12}:

$$\Delta\Psi = -\partial(\delta v_g^{max}/v_g)/\partial\Psi_{12}.$$

- to align an interferometer axes along the determined direction of the group velocity vector, and then to measure values of the phase and the group SAW velocities with improved accuracy
- interferometer base Δr can be measured with high accuracy by measurement of SAW velocity on the sample with known velocity value (for example ST-quartz). Then this value Δr can be used in measurement of other crystals with the same interferometer base.

3. CONCLUSION

A novel possibility for the direct measuring of SAW steering angle by two beams optical interferometer is demonstrated. The influence of the interferometer phase change during the measurement is eliminated by special jumping-frequency algorithm that gives self-compensation of interferometer phase shift change on the measured parameters. New measuring procedure is considerably increase possibilities of the given method [2, 3] for a measurement of basic characteristics of acoustic waves in crystals.

The authors thanks to Dr. Rinat Taziev for computer calculations of SAW characteristics in crystals, G.Ostanina and N.Tolstouhov for development of IDT structures.

The work is supported by Grant N 97-02-18458 from Russian Foundation for Basic Research.

4. REFERENCES

[1] R.M.Taziev and I.B.Yakovkin, «Fast algorithm for correction of material constants of piezoelectric crystals on SAW velocity experimental data», in Proceedings of the IEEE International Ultrasonics Symposium, 1994, V.1, pp. 415-419.

[2] A.V. Tsarev, E.A. Kolosovsky, «Two Light Beams Diffraction by a Surface Acoustic Wave in an Anisotropic Solid», presented at 1998 IEEE International Ultrasonics Symposium Sendai, Miyagi, Japan, October 5-8, 1998.

[3] E.A. Kolosovsky, A.V. Tsarev., I.B. Yakovkin, «Precise interferometric AO method for SAW velocity measurements in an anisotropic solid», IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control, vol. 44, No 6, pp. 1219-1223, 1997.

1999 Joint Meeting EFTF - IEEE IFCS

LSAW ATTENUATION IN DOUBLE-LAYER-SYSTEMS

Paul Wallner[1], *Werner Ruile*[1], *Robert Weigel*[2]

[1] *Siemens AG, Corporate Technology, ZT KM 1, 81730 Munich, Germany*
[2] *Universität Linz, Institut für Nachrichtentechnik/Informationstechnik, Altenberger Str. 69, 4040 Linz, Austria*

Abstract –The behavior of Leaky-SAW (LSAW) in layered structures was investigated by new computer calculations. For these investigations *rotYX* $LiTaO_3$ and *rotYX* $LiNbO_3$ crystal cuts were used. Surface acoustic wave propagation is strongly influenced by thickness, density, and elasticity of the layers. For Leaky-SAW both, the velocity and the inherent loss due to bulk wave emission into the substrate depend on these layer properties. Double-layer-systems can be taken into account if Si_3N_4 or Al_2O_3 is added on the metal layer. As a result, this double-layer-system shows new properties: The angle of minimum propagation loss and the velocity strongly depend on distinct parameters and ratios of the layers. This means that in double-layer-systems the behavior of LSAW-loss and velocity differs from single-layer-structures.

1. INTRODUCTION

The technology of SAW devices is based upon the use of layered structures. Common applications employ an aluminum layer or grating on an anisotropic piezoelectric crystal. Layer thickness, materials, and crystal cuts have to be chosen such that they meet different requirements like small attenuation, good temperature stability or high coupling efficiency. Thickness, density, and elastic properties of this metal layer show a strong influence on the wave propagation. Different studies were performed: Farnell and Adler [1] have investigated the behavior of SAW velocity for isotropic single-layer-structures with respect to the layer thickness. They reported for an isotropic layer on an isotropic substrate that the behavior of velocity depends on distinct ratios of elastic constants and densities of the considered materials. More recently, Hashimoto et al. [2] have shown that the inherent LSAW loss due to bulk wave emission into the substrate is also strongly influenced by the layer properties. They reported that an aluminum layer of about h/λ = 10% on a *42rotYX* $LiTaO_3$ substrate shows a smaller insertion loss compared with the *36rotYX* $LiTaO_3$ crystal cut. Furthermore, Hickernell and Adler [3] reported effects on both, velocity and LSAW attenuation for LSAW on glass covered $LiTaO_3$ and $LiNbO_3$ substrates.

In other words, these examples demonstrate that velocity as well as LSAW attenuation are strongly influenced by the properties of this single layer. Therefore, adding a second layer to this structure should change the LSAW properties in a substantial way. A double-layer-systems can be taken into account, if Si_3N_4 or Al_2O_3 is added on the metal layer. This means, that in a double-layer-system with aluminum as first layer and Si_3N_4 or Al_2O_3 as second layer on a $LiTaO_3$ or a $LiNbO_3$ substrate effects on velocity and LSAW attenuation will occur.

2. THEORY

Wave propagation in piezoelectric solids is described by a set of second order partial differential equations with (Eq. 1)

$$\rho u_{j,tt} = c_{ijkl} u_{k,il} + e_{kij} \varphi_{,ki}$$
$$e_{ikl} u_{k,il} - \varepsilon_{ik} \varphi_{,ki} = 0 \qquad i,j,k,l = 1,2,3 \tag{1}$$

where $\boldsymbol{u}$ and φ are the particle displacement and the electrical potential, respectively. Moreover, ρ, $\boldsymbol{c}$, $\boldsymbol{e}$, and $\boldsymbol{\varepsilon}$ describe density, elasticity, piezoelectricity, and dielectrics, respectively. In their general form these material parameters are tensors up to fourth order. For calculations a coordinate system is chosen in such a way that the substrate is located in the lower half space ($x_3 < 0$), whereas the surface acoustic waves propagate

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

along the x_1-axis. The first layer lies between $0 < x_3 < h_1$ and the second layer is located between $h_1 < x_3 < h_1 + h_2$. To obtain a general SAW solution (Eq. 2) will be an appropriate trial function.

$$u_l = \hat{u}_l \exp[jk(\gamma x_1 + \alpha x_3 - vt)] \qquad l = 1...4 \qquad (2)$$

where v represents the velocity, k the wave number, α the decay coefficients, and $\gamma = 1 + j\delta$ the propagation loss. Substituting (Eq. 1) by (Eq. 2) leads to an 8^{th} order polynomial Eigenvalue problem in α.

For isotropic layers it is possible to obtain an analytical expression for the Eigenvalues α and the Eigenvectors $\hat{u}$. Moreover, the whole set of the eight Eigenvalues is selected. In the free space above the layers the Laplace equation can also be solved analytically. In contrast, for anisotropic substrates no analytical expression exists and only four Eigenvalues have to be selected. In order to separate Leaky-SAW and Rayleigh-waves a careful selection has to be performed[4].

For these calculations a continuous shape of the surface was assumed. Due to the coupling of electrics and acoustics both, electrical and mechanical boundary conditions have to be satisfied. Across each interface (substrate:metal, metal:oxide, and oxide:vacuum) the particle displacement $\boldsymbol{u}$ and the normal component of stress T_{i3} must be continuous. In contrast, the mechanical boundary conditions for the top layer are reduced to $T_{i3} = 0$. Furthermore, the electrical boundary conditions have to be satisfied. This means that the parallel component of the electrical field E_1 as well as the normal component of the dielectric displacement D_3 must be continuous across each interface. However, due to electrical shortening of the metal layer the electrical potential is zero there. Taking these boundary conditions into account, the particular solution can be obtained and both, velocity and LSAW attenuation can be calculated. The particular solution for every region (substrate, layers, vacuum) is a superposition of the partial waves weighted by the coefficients A_n:

$$u_l = \sum_n A_n \hat{u}_{ln} \exp[jk(\gamma x_1 + \alpha_n x_3 - vt)] \qquad (3)$$

For these calculations we have used material data from Kovacs [5]. Values for aluminum were taken from Auld [6], whereas appropriate Si_3N_4 data can be found by Hickernell et al. [7]. In addition, the present study is based upon a frequency of $f = 1\ GHz$.

3. Results

LSAW exist only on distinct crystal cuts. As a matter of fact, crystal cuts of appropriate materials are *rotYX* $LiTaO_3$ and *rotYX* $LiNbO_3$. Therefore, the influence of layered substrates on the LSAW will be examined for both materials [8].

At first, the behavior of the LSAW attenuation is investigated for three different crystal cuts of a Si_3N_4:*Al:rotYX* $LiTaO_3$ double-layer-system. Although the Si_3N_4-Al layer thickness of *400nm* remains fixed, the angle of minimum LSAW attenuation depends on distinct ratios of $d_{Al}/(400nm)$ (Fig. 1).

In general, the LSAW loss is a function of more than one parameter. However, if a fixed entire thickness of the double-layer-structure is assumed, the LSAW loss can be presented with respect of the aluminum ratio

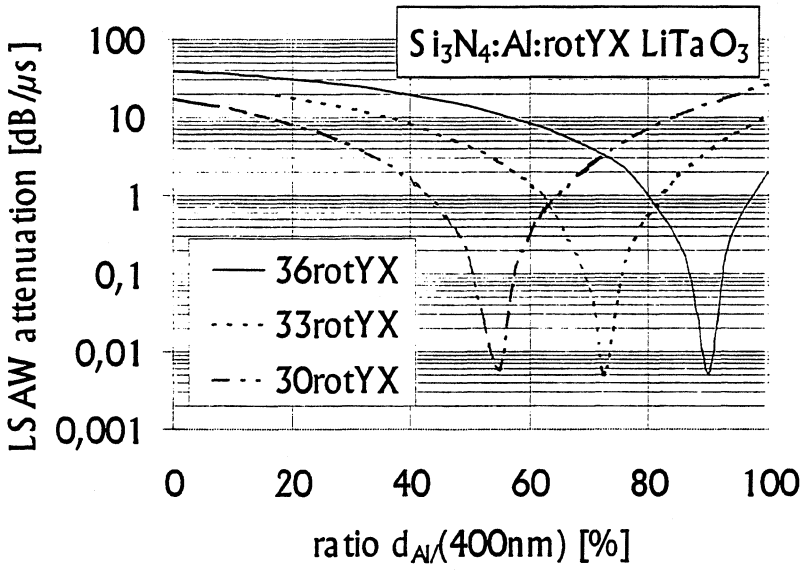


Fig. 1: Behavior of the LSAW loss for three different crystal cuts with respect of an altered aluminum ratio.

and the crystal cut angle. For this calculation a double-layer-system consisting of 400nm aluminum and Al_2O_3 on *rotYX* $LiTaO_3$ was used (Fig. 2). It can be seen that an appropriate low LSAW loss occurs in a distinct channel depending on both, the crystal cut and the layers ratio.

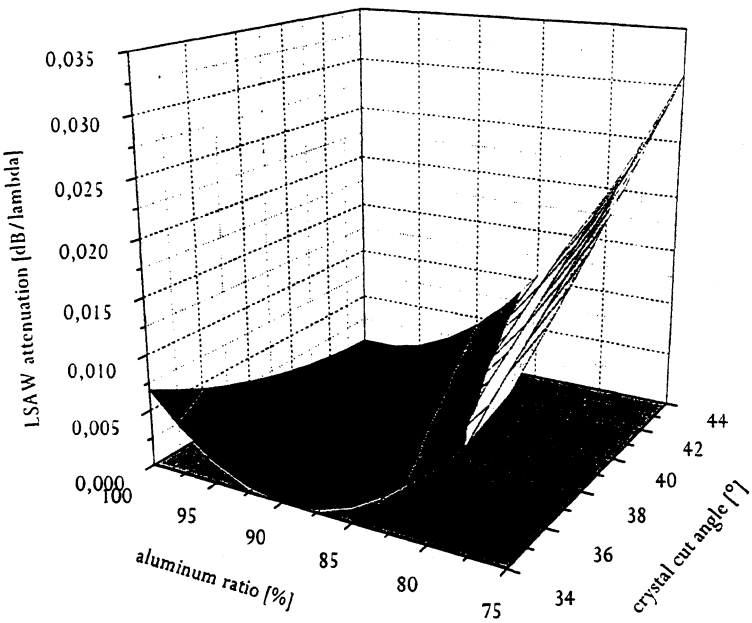


Fig. 2: Behavior of LSAW loss for an Al-Al_2O_3 double-layer-systems with an entire thickness of 400nm as a function of the aluminum ratio and the crystal cut angle.

For $LiTaO_3$ it was found that the LSAW attenuation behavior is strongly influenced by the second layer. Thus, it has to be investigated, if or how strong these effects will occur for other substrates. For this purpose calculations were performed using an aluminum-Si_3N_4 double-layer-system on *rotYX* $LiNbO_3$ crystal cuts.

Again, in Fig. 3 the behavior of the LSAW attenuation is shown for three different crystal cuts (here: *64rotYX*, *60rotYX*, and *57rotYX*) of a Si_3N_4:*Al*:*rotYX* $LiNbO_3$ double-layer-system. Although the Si_3N_4-aluminum layer thickness of *400nm* remains fixed, the angle of minimum LSAW attenuation depends on distinct ratios of d_{Al}*/(400nm)*. However, the behavior of these $LiNbO_3$ crystal cuts is different compared with the double-layer-system of $LiTaO_3$ in Fig. 1: The minimum attenuation of LSAW loss occurs for ratios which are very similar. In contrast, for $LiTaO_3$ larger differences have been the result.

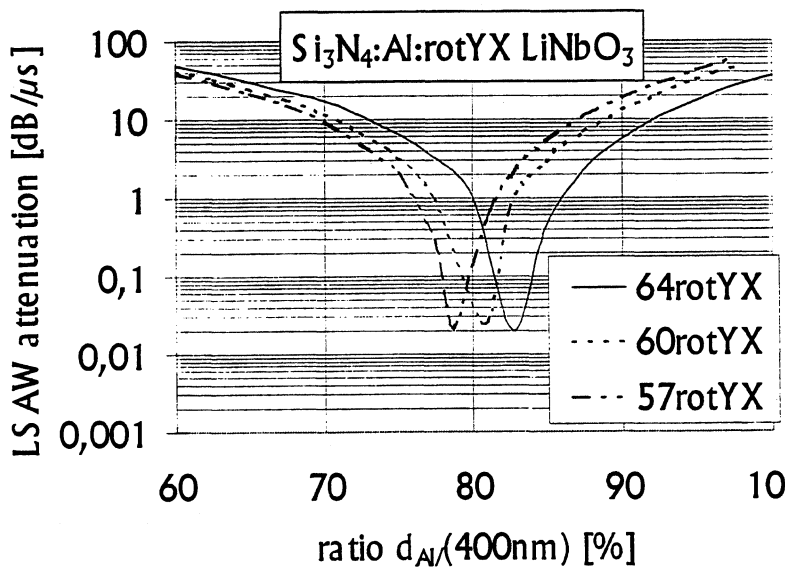


Fig. 3: Behavior of the LSAW loss for three different $LiNbO_3$ crystal cuts with respect of an altered aluminum ratio.

For the previously shown results a fixed entire thickness of 400nm was assumed. A slightly changed situation can be investigated, if a distinct crystal cut is considered and the entire thickness of the double-layer-system is changed. Choosing *64rotYX $LiNbO_3$* as piezoelectric substrate and using an Si_3N_4-aluminum double-layer-system double-layer-system the behavior of minimum LSAW attenuation was calculated for different entire thicknesses with respect of the aluminum ratio (Fig. 4). It can be found that no minimum attenuation occurs for an entire layer thickness of 100nm. Moreover, an increased entire thickness shifts the minimum of LSAW attenuation to smaller ratios.

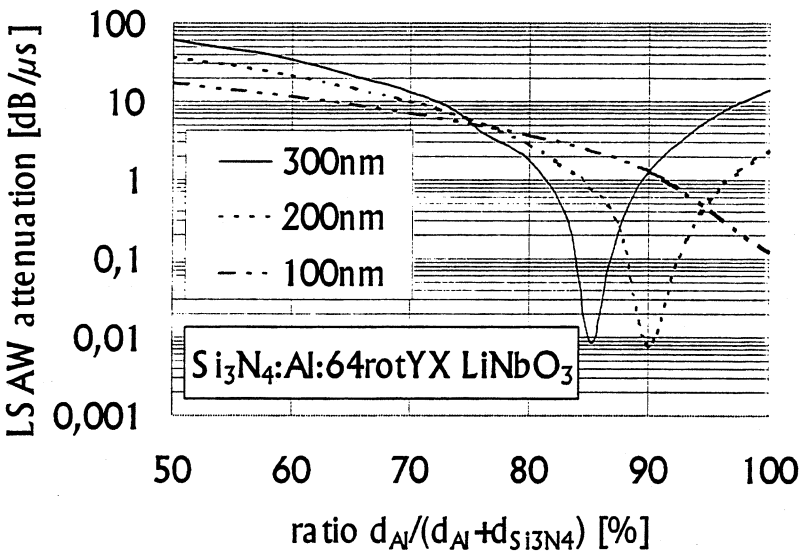


Fig. 4: LSAW loss for three different double-layer-systems on *64rotYX $LiNbO_3$* with an entire thickness of 100nm, 200nm, and 300nm.

4. Conclusions

Calculations for aluminum and Si_3N_4 or Al_2O_3 double-layer-systems on *rotYX $LiTaO_3$* and *rotYX $LiNbO_3$* crystal cuts have been presented. A substantial influence of the second layer on the behavior of the LSAW attenuation has been shown. Moreover, the LSAW loss depends on distinct parameters of the used layers. A more generalized understanding of the LSAW attenuation therefore seems to be possible.

References

[1] G.W. Farnell and E.L. Adler, editet by W.P. Mason, Physical Acoustics, Principals and Methods, vol. 9. New York: Academic Press, 1972, chap. 2

[2] O.Kawachi, G. Endoh, M. Ueda, O. Ikata, K. Hashimoto, and M. Yamaguchi, "Optimum Cut of $LiTaO_3$ for High Performance Leaky Surface Wave Filters", in IEEE Ultrasonics Symposium Proc., 1996, pp. 71 - 76

[3] F.S. Hickernell, E.L. Adler, "Pseudo-SAW Propagation on Layered Piezo-Substrates: Experiments and Theory Including Film Viscosity", in IEEE Ultrasonics Symposium Proc., 1996, pp. 87 - 90

[4] E.L. Adler, "SAW and Pseudo-SAW Properties Using Matrix Methods", in IEEE Ultrasonics Symposium Proc., 1992, pp. 455 - 460

[5] G. Kovacs, M. Anhorn, H.E. Engan, G. Visintini, C.C.W. Ruppel, "Improved Material Constants for $LiNbO_3$ and $LiTaO_3$", in IEEE Ultrasonics Symposium Proc., 1990, pp. 435 - 438

[6] B.A. Auld, Acoustic Fields and Waves in Solids, vol. 1, New York, John Wiley & Sons, 1973

[7] T.S. Hickernell, F.M. Fliegel, F.S. Hickernell, "The Elastic Properties of Thin-Film Silicon Nitride, IEEE Ultrasonics Symposium Proc., 1990, pp. 445 – 448

[8] W. Ruile, U. Rösler, P. Wallner, "Oberflächen-wellen-(SAW-) Bauelement auf insbesondere Lithiumtantalat und –niobat-Substrat, Patent Nr. 19758195.1, 30.12.1997

GHz-Range Surface Acoustic Wave Low Loss Filter at Super Low Temperature

Kazuhiko Yamanouchi*, Hideyuki Nakagawa** and Hiroyuki Odagawa**

*Department of Electronics, Tohoku Institute of Technology, Sendai, 982-0831, Japan
**Research Institute of Electrical Communication, Tohoku University, Sendai 980-8577, Japan

Abstract

The 10 GHz-range surface acoustic wave (SAW) is of great importance in the field of physical acoustics and application of SAW devices, for example, in mobile and wireless communications. The temperature dependency of the propagation attenuation at 10 GHz-range is measured using the three-transducer system with an electrode width of less than 0.1 μm, which is fabricated using an electron beam exposure and lift-off technique. The results show that the attenuation at room temperature is about 1.5 dB/(100 λ) at 10 GHz.

Next, the temperature dependency of the insertion loss of a floating electrode reflection type unidirectional transducer (FEUDT) whose electrodes are respectively fabricated from Al and the superconductor Nb, is investigated experimentally. In the case of Al-FEUDT at 8 G Hz, the insertion loss at a low temperature is about 8.5 dB, which is 3.4 dB lower than the insertion loss at room temperature. The main cause of decrease in the insertion loss is the decrease in the resistivity of the very thin and narrow metal electrodes. The results of Nb-FEUDT show a marked decrease in the insertion loss around the critical temperature Tc=9.23 K

1. Introduction

Surface acoustic wave (SAW) devices are capable of performing many signal processing functions. Therefore, these devices have been very successful due to high volume applications such as intermediate frequency (IF) filters for televisions and RF filters for mobile telecommunications.

The SAW devices have very important properties that are vital for various commercial applications. These properties are as follows: 1) small size and light weight, and 2) suitability for mass production due to their fabrication process, using photo-lithography techniques. Hence, for the development of mobile and wireless communications, the high frequency and low insertion loss SAW devices are required.

For realizing a high frequency and low loss SAW filter, it is important to evaluate the quantitative factor of the insertion loss. In this study, the 10 GHz-range SAW filters are fabricated using the electron beam exposure and lift-off technique, and measured at room temperature and a low temperature. Then, on the basis of the measurement results, the propagation loss of 10 GHz-range SAW is evaluated. An 8 GHz SAW low loss filter using an unidirectional transducer with Al-electrode and a 4 GHz-range low loss filter with a superconductor material Nb-electrode are also described.

2. Measurement of the 10GHz-Range Surface Acoustic Wave Propagation Loss

For GHz-range SAW devices, it is necessary to evaluate the propagation loss in order to obtain low insertion loss SAW devices. The propagation losses are generally due to the following reasons:[1] a) loss due to the interaction with thermal phonons, b) loss due to scattering by crystalline defects, pits and scratches, and c) air loading loss. These losses have a close relationship with temperature, frequency, and air pressure. So, each loss can be evaluated by obtaining the measurements under different conditions. For example, the air loading loss can be evaluated by obtaining the measurements under low pressure conditions. The loss due to the interaction with thermal phonons can be evaluated by the difference in the insertion loss from the room temperature to the low temperature. The remaining losses are due to scattering by crystalline defects, pits and scratches.

In this section, the total propagation loss and the loss due to the interaction with thermal phonons on a 10GHz SAW propagating along the X axis of 128° rotated Y-cut $LiNbO_3$ (128° Y-X $LiNbO_3$) are measured.

$$I = E_{fA} E_{fB} \exp(-\alpha_p L) \qquad (1)$$

2.1 Measurement method for the propagation loss

The total insertion loss I of a two transducer filter is given by the following: [1] Here, E_f is the transducer efficiency, α_p is the surface wave attenuation per unit length and L is the distance between transducers. If the transducer efficiency E_f has no temperature dependency, the attenuation constant α_p can be evaluated by the two-transducer method. However, the transducer efficiency E_f depends on temperature due to the temperature dependency of the electrode resistivity. Therefore, it is necessary to adopt the three-transducer method shown in Fig.1. The SAW excited by transducer A is received by transducers B and C. Then, the ratio of the insertion loss is given by the following:

If the temperature dependency of E_{fC}/E_{fB} is lower than that of α_p, the attenuation constant α_p will be obtained. So, if the temperature dependency of the insertion loss is measured as shown in Fig.2(a), that of the attenuation constant α_p as shown in Fig.2(b) is obtained from the difference of curves C and B. In Fig.2(b), the attenuation constant α_p approaches α_0 at the low

$$\frac{I_C}{I_B} = \frac{E_{fc}}{E_{fB}} \exp[-\alpha_p (L_C - L_B)]. \qquad (2)$$

temperature, which includes the attenuation due to the air loading and the scattering by crystalline defects, pits and scratches. Here, the attenuation α_t due to the interaction with thermal phonons is obtained from $\alpha_p - \alpha_0$.

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

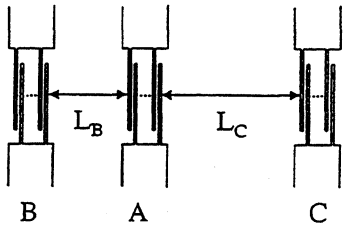


Fig.1 Schematic diagram of three-transducer.

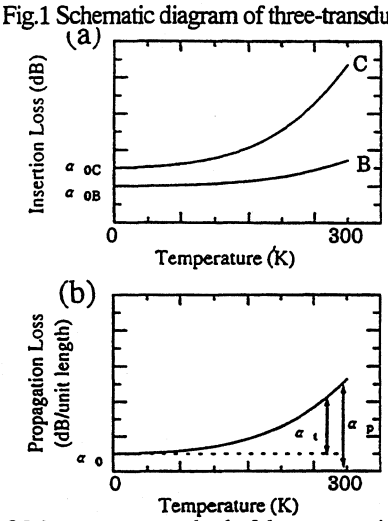


Fig.2 Measurement method of the propagation loss.

2.2 Cryostat and measurement system at low temperature

The cryostat is shown in Fig.3. The cryostat adopts the He flow method. In this method, liquid He from a He dewar is introduced into the heat exchanger through the transfer tube, and the temperature of He is controlled by the temperature controller. The sample is directly cooled by the temperature controlled He. The flow of He is adjusted by a needle valve, and the controlled temperature range is from 3.5 K to 320 K. The schematic of the sample room is shown in Fig.4. The input and output coaxial cables [(1),(2) in Fig.4] and the Cr-AuFe thermocouple [(3) in Fig.4] are introduced from outside the cryostat, and the SAW filter package is connected by the SMC connector as shown in Fig.4. The temperature around the sample is measured exactly by the Cr-AuFe thermocouple, and the filter is measured by a network analyzer. During measurement, the internal pressure of the cryostat is about 20 % of the atomspheric pressure.

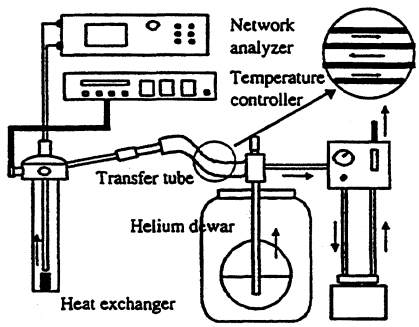


Fig.3 Schematic diagram of the cryostat and measurement system.

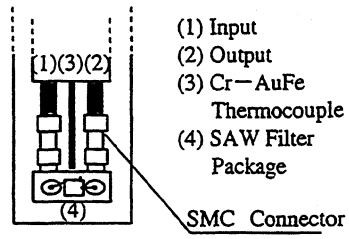


Fig.4 Schematic diagram of the sample room.

2.3 Experimental result

The three-transducer method shown in Fig.1 was used due to the reason given earlier in § 2.1. The three-transducer was fabricated by the electron beam exposure and lift-off technique. Figure 5 shows the scanning electron microscope (SEM) image of the 10 GHz three-transducer resist pattern with a line width of 0.1 μm, which is very fine. The substrate is 128° Y-X $LiNbO_3$. The wavelength λ is 0.4 μm, and the propagation length is L_B=105 λ and L_C=300 λ, shown in Fig.1. The number of pairs is 10, the aperture is 20 λ, and the film thickness of Al electrodes is 30 nm.

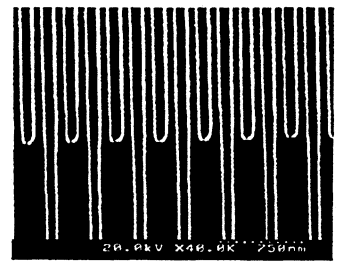


Fig.5 SEM image of 10 GHz three-transducer resist pattern. (line width = 0.1 μm)

The insertion loss of the fabricated 3-transducer at room temperature (293 K) and low temperature (7 K) is shown in Table 1. The measurement at low temperature (7 K) was carried out after the sample was stabilized at 7 K for 1 h., and again, after the sample was warmed to room temperature (293 K) in a period of 3−4 h, and measured at 293 K.

The frequency dependency of the propagation loss at room temperature is shown in Fig.6. [2] A new point at 10 GHz is calculated using the values of Table 1. The attenuation is almost proportional to 1.25 power of frequency, hence it is 1.5 dB/ (100 λ) at 10 GHz, which is not so large. The frequency dependency of the loss due to the interaction with thermal phonons is shown in Fig.7. A new point at 10 GHz is also calculated on the basis of the measurement method elucidated in § 2.1, using the values of Table 1. The loss is almost proportional to the square of frequency. The reasons why some points of the plot are outside the squared line are due to the fact that the receiving interdigital transducers (IDTs) of three-transducer are not completely symmetrical and the characteristics of the sample are degraded

symmetrical and the characteristics of the sample are degraded by impedance mismatching on the packaging filter with Al bonding.

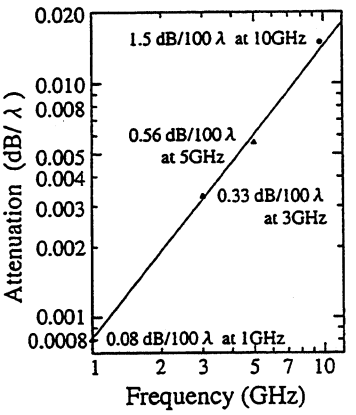


Fig.6 Frequency dependency of the propagation loss on 128° Y-X $LiNbO_3$ at room temperature.

Table 1. Insertion loss of three-transducer at 293 K and 7 K.

Frequency (GHz)	Propagation length (λ)	Insertion loss at 293 K (dB)	Insertion loss at 7K (dB)
9.63	105	31.3	28.3
9.70	300	34.2	30.1

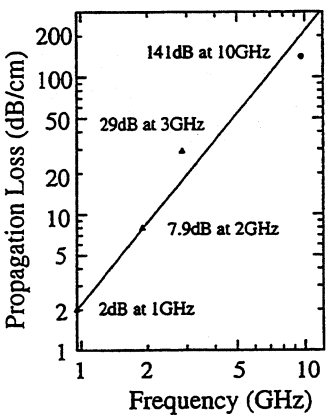


Fig. 7 Frequency dependency of the loss due to the interaction with thermal phonons on 128° Y-X $LiNbO_3$ at room temperature.

3. 10 GHz-Range Low Loss Filter

Conventional IDT has an inherent minimum insertion loss of 6 dB because of bidirectionality and strong pass band ripple due to secondary effects. However, uni-directional IDT (UDT) can ideally be designed for less than 1 dB insertion loss, and suppress the influence of the triple transit echo (TTE). Thus, the UDT is effective in low loss filters. Figure 8 shows the λ/10-type floating electrode reflection-type UDT (FEUDT), which is an UDT using an internal reflection with the floating electrode of an open metal strip.[3)4)] It has the following features:

(i) The second harmonic is strongly excited.
(ii) It is easy to fabricate, because the electrodes and gaps are at regular intervals.

So, the λ/10-type FEUDT is suitable for the high frequency and low loss filters.

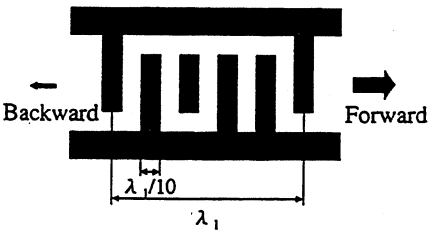


Fig.8 Configuration of λ/10-type FEUDT. (λ_1 is a fundamental wavelength)

3.1 Experimental results

3.1.1 8 GHz λ/10-type FEUDT

8 GHz λ/10-type FEUDT was fabricated by the electron beam exposure and lift-off technique. Figure 9 shows the SEM image of a 8 GHz λ/10-type FEUDT resist pattern with a line width of 0.09 μm, which is very fine. The substrate is 128° Y-X $LiNbO_3$.

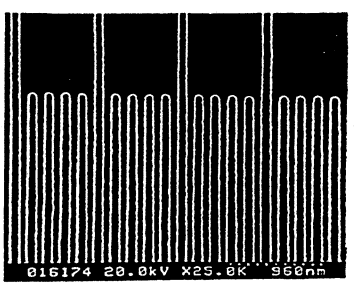


Fig.9 SEM image of λ/10-type FEUDT resist pattern (line width = 0.09 μm).

The wavelength on second harmonic λ_2 is 0.45 μm, and the propagation length is $46\lambda_2$. The number of pairs is 20, the aperture is $46\lambda_2$, and the film thickness of Al electrodes is 30 nm. The measurement method is the same as described in § 2.2. Figures 10(a) and (b) respectively show the frequency responses of 8 GHz λ/10-type FEUDT at room temperature (300 K) and low temperature (7 K), which include TTE. Com-paring both the minimum insertion losses, the insertion loss at low temperature is about 8.5 dB, which is 3.4 dB lower than the insertion loss at room temperature. The reasons why the insertion loss at the low temperature is lower than that at the room temperature are the decrease in resistivity of IDT electrodes, the mismatch loss, and the decrease in loss due to the interaction with thermal phonons. However, if the propagation length is about $46\lambda_2$ at 8 GHz, the loss due to the interaction with thermal phonons is only about 0.26 dB from Fig.7. So, the loss due to the resisitivity of IDT electrodes and the mismatch loss are very serious in 10 GHz-range SAW filters.

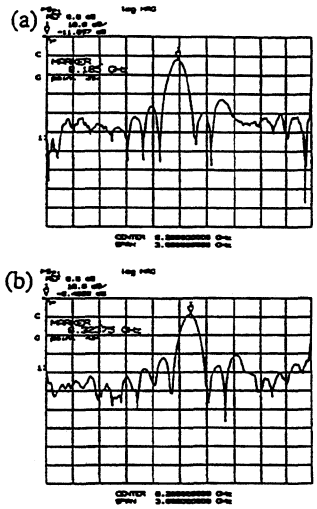


Fig.10 Frequency responses of 8 GHz λ/10-type FEUDT. (a) at 300 K, (b) at 7 K.

3.1.2 4 GHz-range Nb λ/10-type FEUDT

The resistivity of IDT electrodes is a very important factor as noted in § 3.1.1, so we propose using a superconductor as a material for IDT electrodes. 4 GHz-range Nb λ/10-type FEUDT is fabricated by the electron beam exposure and lift-off technique. The substrate is 128° Y-X $LiNbO_3$. The wave-length on the second harmonic λ_2 is 0.8 μm, and the propagation length is 40 λ_2. The number of pairs is 20, the aperture is 40 λ_2, and the film thickness of Nb electrodes is 100 nm. Figures 11 (a) and (b) respectively show the frequency responses of 4 GHz-range Nb λ/10-type FEUDT at room temperature (293 K) and low temperature (9 K). Comparing both the minimum insertion losses, the insertion loss at the low temperature is about 6.7 dB, which is 9.7 dB lower than the insertion loss at the room temperature.

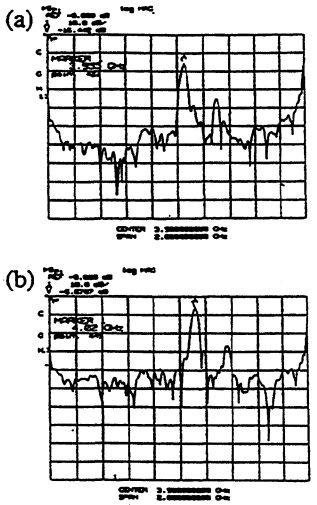


Fig.11 Frequency responses of 4 GHz-range Nb λ/10-type FEUDT. (a) at 293 K, (b) at 9 K.

The plots in Fig.12 show the temperature dependency of the 4 GHz-range Nb λ/10-type FEUDT. The broken line shows the relative resistivity of the Nb film. The lower the temperature, the lower is the insertion loss, and finally, it changes markedly by about 2.2 dB at 10 K because the Nb resistance changes from normal to super-conducting. The changes between the insertion loss and relative resistivity correspond well at the lower temperature.

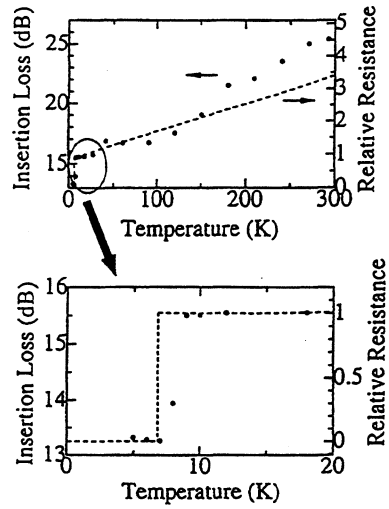


Fig.12 Temperature dependency of the insertion loss of a 4 GHz-range Nb λ/10-type FEUDT and relative resistance of the Nb film

4. Conclusion

The total propagation loss and the loss due to the interaction with thermal phonons are measured. The total propagation loss is 1.5 dB/(100 λ) at 10 GHz, which is low for 10 GHz SAW devices. Then, the temperature dependencies of the insertion losses of FEUDTs whose electrodes are fabricated from Al and Nb are investigated experimentally. The results of Nb λ/10-type FEUDT showed a marked decrease in the insertion loss around the critical temperature T_c=9.23 K. We also observed an insertion loss of about 8.5 dB at 8 GHz at the low temperature. Therefore, the unidirectional IDTs with thick electrodes are very useful for low loss filters.

Acknowledgement

We would like to express our thanks to Yamashita Laboratory of Research Institute of Electrical Communication, Tohoku University for fabricating the Nb thin films.

References

1)A.J. Slobodnik, P.H. Carr and A.J. Budreau: J. Appl. Phys. Vol. 41 (1970) 4380.

2)K. Yamanouchi: Proc. 1994 IEEE Ultrason. Symp. (1994) 421.

3)K. Yamanouchi, CH. S. Lee, K. Yamamoto, T. Meguro and H. Odagawa: Proc. 1992 IEEE Uiltrason. Symp. (1992) 139.

4)K. Yamanouchi and T. Furuyashiki: Elec-tron. Lett. Vol. 20 (1984) 989.

1999 Joint Meeting EFTF - IEEE IFCS

FREQUENCY-TEMPERATURE EFFECTS IN THREE-DIMENSIONAL SURFACE ACOUSTIC WAVE PERIODIC STRUCTURES.

Yook-Kong Yong[†] *and Shigeo Kanna*[‡]

[†]*Civil & Environmental Engineering Dept., Rutgers University, NJ, U.S.A. yong@jove.rutgers.edu*
[‡]*Seiko Epson Corporation, 3-3-5, Owa, Suwa-shi, Nagano-ken, 392 JAPAN*

Abstract

Numerical results on the resonant frequencies and their frequency-temperature behavior in a three-dimensional surface acoustic wave (SAW) periodic structure are presented. These results are compared with those of the quartz SAW resonator . The 3-d fem models of SAW periodic structures are useful in the analyses of IDT's geometry and crystal cuts of SAW resonators which have power flows transverse to the SAW phase velocity. They are also useful for analyzing transversely coupled SAW resonators.

Piezoelectric, Lagrangian equations for the frequency-temperature behavior of quartz are presented. Three-dimensional finite element results for a half wavelength SAW periodic structure are compared with measured data from SAW resonator . The frequency characteristics of the fundamental SAW mode and its transverse overtones as a function of the electrode length W in the transverse direction are studied. Generally good results are obtained by the 3-d fem models when compared with measured values. Their frequency-temperature behavior are also studied and are reasonably well predicted by the 3-d SAW periodic structure. The 2-d fem models are shown to yield results consistent with the 3-d fem models.

I. Introduction

Two- and three-dimensional finite element analysis of periodic structures of surface acoustic wave (SAW) waveguides has been useful in analyzing the effects of the interdigital transducers (IDT's) geometry on SAW phase velocity. [1,2,3,4,5,6,7]. Hasegawa and Koshiba[7] presented the use of three-dimensional finite elements for the analysis of periodic surface acoustic waveguides.

An important attribute for a good SAW resonator is its frequency-temperature stability. The use of finite elements for analyzing the frequency-temperature stability of the SAW resonator is a natural extension of the use of finite elements for the analysis of periodic SAW waveguides. Yong and Kanna[8] presented results which showed that the two-dimensional straight crested wave model of a SAW periodic structure predicted fairly accurately the frequency-temperature behavior of quartz SAW resonators. However, the 2-d fem models cannot capture the length effects of the IDT's, width of the SAW resonator and power propagation in the direction transverse to the direction of the SAW propagation.

We introduce in this paper the third dimension in our finite element models of SAW periodic structures. These 3-d fem models are useful in the analyses of IDT's geometry and crystal cuts in SAW resonators where the power flow of the SAW mode is not in the same plane as its phase velocity. They are especially useful for analyzing transversely coupled SAW resonators. Since the 3-d models are numerically intensive, sparse matrix storage technology along with an efficient eigensolver must be used. Adaptive meshing is used to reduce the size of the finite element model.

II. Piezoelectric Lagrangian Equations for the Frequency-Temperature Behavior of Quartz.

The classical model for the frequency-temperature behavior of quartz crystal resonators is based on the infinitesimal theory of elasticity or piezoelectricity. In order to account for the homogeneous temperature field, the density, resonator geometry, and hence the reference frame, are allowed to change with temperature. This is rather confusing because these temperature coefficients of elastic constants and density are then referred to a reference frame which is itself a function of temperature. Furthermore, the thermal shear strains are not taken into account in this ad hoc treatment. In order to overcome this confusion, and to ensure consistent results for anisotropic materials, we present a Lagrangian formulation for the frequency-temperature behavior of SAW resonators. This formulation is similar to those, for example, in references 9 and 10, but we prefer to use the Piola-Kirchhoff stress tensor of the second kind for reasons of symmetry. The formulation was previously developed in references 11 and 12. The following incremental Lagrangian equations which includes the piezoelectric effect was presented in reference 8, and we show them here again for the convenience of the reader.

IIa. Strain-displacement, and electric field-potential relations.

$$s_{ij} = \frac{1}{2}(\beta_{kj}u_{k,i} + \beta_{ki}u_{k,j}), \qquad \text{Eq.(1a)}$$

$$E_i = -\phi_{,i} \qquad \text{Eq.(1b)}$$

The terms s_{ij}, $u_{k,i}$, E_i and ϕ are, respectively, the incremental strains, the first partial derivative of incremental displacement, incremental electric field and incremental electric potential. T_0 is the reference temperature set at 25° C, δ_{ik} and $\alpha_{ik}^{(n)}$ denote the

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

Kronecker delta and nth order temperature coefficients of thermal expansion, respectively.

IIb. Equations of motion and electrostatics.

$$\beta_{ik}\sigma_{kj,j} = \rho\ddot{u}_i \text{ in } V, \text{ and} \quad \text{Eq.(2a)}$$

$$D_{i,i} = 0 \text{ in } V, \text{ where} \quad \text{Eq.(2b)}$$

$$\beta_{ik} = \delta_{ik} + \alpha_{ik}^{(1)}\theta + \alpha_{ik}^{(2)}\theta^2 + \alpha_{ik}^{(3)}\theta^3 \quad \text{Eq.(2c)}$$

$$\theta = (T - T_o) \quad \text{Eq.(2d)}$$

where σ_{kj} is the incremental stress tensor, ρ the mass density, u_i the incremental mechanical displacement and D_k the incremental electric displacement vector.

IIc. Initial and boundary conditions.

A unique solution is guaranteed by specifying u_i, $\dot{u}_j$, and E_j at each point of the body at an initial time $t = t_o$, and by specifying at all times one member of each of the following products on the crystal surface: p_1u_1, p_2u_2, p_3u_3, and $q\phi$ where $p_i = n_j\beta_{ik}\sigma_{kj}$ is the component of the surface traction vector, $q = n_iD_i$ is the surface charge density and n_i is the unit outward normal to the surface.

IId. Constitutive relations.

$$\sigma_{ij} = D_{ijkl}s_{kl} - e_{kij}^{\theta}E_k, \quad \text{Eq.(3a)}$$

$$D_i = e_{ijk}^{\theta}s_{jk} + \varepsilon_{ik}^{\theta}E_k \quad \text{Eq.(3b)}$$

$$D_{ijkl} = C_{ijkl} + D_{ijkl}^{(1)}\theta + D_{ijkl}^{(2)}\theta^2 + D_{ijkl}^{(3)}\theta^3 \quad \text{Eq.(3c)}$$

$$e_{ijk}^{\theta} = e_{ijk}(1 + Te_{ijk}^{(1)}\theta + Te_{ijk}^{(2)}\theta^2 + Te_{ijk}^{(3)}\theta^3) \quad \text{Eq.(3d)}$$

$$\varepsilon_{ik}^{\theta} = \varepsilon_{ik}(1 + T\varepsilon_{ik}^{(1)}\theta + T\varepsilon_{ik}^{(2)}\theta^2 + T\varepsilon_{ik}^{(3)}\theta^3) \quad \text{Eq.(3e)}$$

where C_{ijkl} is the elastic stiffness tensor measured at constant electric field and temperature, $D_{ijkl}^{(n)}$ are the nth order thermoelastic constants, e_{ijk} is the piezoelectric stress constants tensor measured at constant strain and temperature, $Te_{ijk}^{(n)}$ are the nth temperature coefficient of piezoelectric stress constants, ε_{ik} is the dielectric permittivity constants tensor measured at constant strain and temperature and $T\varepsilon_{ik}^{(n)}$ are the nth temperature coefficient of dielectric permittivity constants.

IIe. Material Constants for Quartz

There are a few sets of material constants used by researchers and engineers for the quartz resonator. The most commonly used values for the elastic stiffness constants, dielectric permittivity and piezoelectric stress constants are those by Bechmann[13]. The coefficients of thermal expansion are from a paper by Bechmann, Ballato and Lukaszek[14]. The thermoelastic constants are from reference 11. We presently use the first temperature coefficients of piezoelectric constants and dielectric permitivities from reference 15 and 16 because we do not have their effective material temperature coefficients. (This will not contribute too much inaccuracy into our models because quartz has very weak piezoelectric couplings.) The following Tables 1 to 4 show the quartz (Y-cut) material constants used in this paper (density of quartz 2649 kg/m^3:

Table 1: Coefficients of thermal expansion [14]

ij	$\alpha_{ij}^{(1)}$, 10^{-6}/C	$\alpha_{ij}^{(2)}$, 10^{-9}/C^2	$\alpha_{ij}^{(3)}$, 10^{-12}/C^3
11	13.71	6.5	-1.9
22	13.71	6.5	-1.9
33	7.47	2.9	-1.5

Table 2: Elastic constants and their temperature constants

pq	Elastic constants [13] C_{pq} (10^9 Pa)	1st thermo-elastic constants [11] $D^{(1)}_{pq}$ (10^6 Pa/C)	2nd thermo-elastic constants [11] $D^{(2)}_{pq}$ (10^3 Pa/C^2)	3rd thermo-elastic constants [11] $D^{(3)}_{pq}$ (Pa/C^3)
11	86.74	-5.9217	-10.119	-5.3702
13	11.91	-8.3983	-12.176	-0.51647
14	-17.91	-1.5087	0.97458	14.615
33	107.2	-21.177	-16.918	10.044
44	57.94	-10.780	-15.777	-2.6439
66	39.88	6.3083	6.0238	-9.5420

Table 3: Piezoelectric stress constants[15]

$e_{i\lambda}$ Piezoelectric stress constants, C/m^2		$Te_{i\lambda}^{(1)}$ 10^{-6}/C
$e_{11} = -e_{12} = -e_{26}$	0.171	-16
$e_{14} = -e_{25}$	- 0.0406	-144
$e_{13} = e_{35} = e_{36}$	0	0

Table 4: Dielectric permittivity[16]

ε_{ij}, Dielectric permittivity, 10^{-12} C/V-m		$T\varepsilon_{ij}^{(1)}$, 10^{-6}/C
ε_{11}	39.21	28
ε_{22}	39.21	28
ε_{33}	41.03	39

III. Principles of periodic analysis

In order to achieve a periodic structure in the X direction, the wave propagation is assumed to be periodic in the X direction. This assumption is to a practical limit valid in

instances where there are steady state vibrations in the regions of the SAW resonator which are structurally periodic.

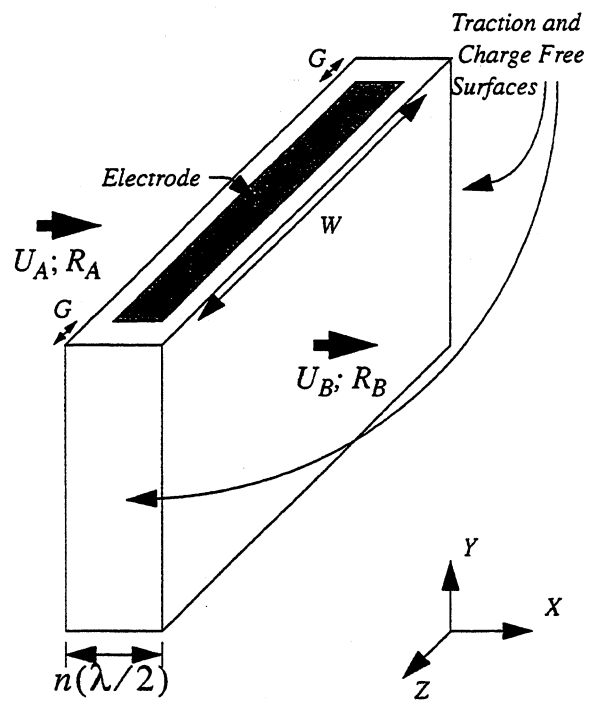


FIG. 1: Symbolic representation of the finite element model of a periodic SAW structure.

Figure 1 above shows a representation of a periodic SAW structure where the displacement vector U_A (including electric potentials) and the reaction force vector R_A (including point charges) from the removed structure are acting on the X surface of the left hand side. Similar displacements and forces are shown on the X surface of the right hand side. The width of the periodic structure is $n(\lambda/2)$ where n is an integer and λ is the wavelength of the SAW in the X direction. In order for the fem model to represent the periodic structure removed from a series of repeating structures, the following constraints must be imposed:

$$\begin{aligned} U_A &= pU_B \\ R_A &= -pR_B \end{aligned} \tag{4}$$

where $p = (-1)^n$. The finite element matrix equation for the periodic structure of Fig. 1 may be written as:

$$\begin{bmatrix} K_{II} & K_{IA} & K_{IB} \\ K_{IA}^T & K_{AA} & K_{AB} \\ K_{IB}^T & K_{AB}^T & K_{BB} \end{bmatrix} \begin{bmatrix} U_I \\ U_A \\ U_B \end{bmatrix} = \omega^2 \begin{bmatrix} M_{II} & M_{IA} & M_{IB} \\ M_{IA}^T & M_{AA} & M_{AB} \\ M_{IB}^T & M_{AB}^T & M_{BB} \end{bmatrix} \begin{bmatrix} U_I \\ U_A \\ U_B \end{bmatrix} + \begin{bmatrix} 0 \\ R_A \\ R_B \end{bmatrix} \tag{5}$$

where the subscript I denotes internal nodal degrees of freedom. When the constraints of eqn.(4) are imposed on the matrix eqn.(5), the second and third row of the matrix equation can be combined to form

$$\begin{bmatrix} K_{II} & (K_{IA} + pK_{IB}) \\ SYM & (K_{AA} + K_{BB} + pK_{AB} + pK_{AB}^T) \end{bmatrix} \begin{bmatrix} U_I \\ U_A \end{bmatrix} = \omega^2 \left(\begin{bmatrix} M_{II} & (M_{IA} + pM_{IB}) \\ SYM & (M_{AA} + M_{BB} + pM_{AB} + pM_{AB}^T) \end{bmatrix} \begin{bmatrix} U_I \\ U_A \end{bmatrix} \right) \tag{6}$$

The eigenpair solutions to eqn.(6) yield the frequencies and mode shapes of resonant modes.

IIIa. Treatment of the electrode finger and the unelectroded regions.

The electrode finger is assumed to be perfect conductors. The elastic properties of aluminum are used. For a shorted condition, the electric potential at the interface of the finger and quartz substrate is set to zero. In an open circuit condition, the electric potentials at all nodal points along this interface are constrained to be equal thus forming an equipotential surface.

At the unelectroded regions of the top Y surface, the surface traction and surface charge density are set to zero. The surfaces of the electrode finger are traction free. The Z end surfaces of the periodic structure are traction free and charge free. The bottom Y surface may be fixed or traction free since the energy of the SAW is concentrated at the top few wavelengths of SAW. We prefer to keep the bottom Y surface traction free and charge free.

IIIb. Finite element models

A 20-node serendipity brick element is used in our analysis. There are two layers of elements per electrode finger. Only rectangular electrodes with a width in the X direction of one quarter wavelength are used in this study. A half-period (half wavelength) structure is modeled, and the wavelength is 10 microns.

The finite element mesh used is adaptively finer at the top where the SAW of interest is located. The adaptive mesh may not be as accurate for the BAW modes but at present the BAW modes are not our primary concern. We seek only an accurate representation of the SAW modes along with a rough estimate of the couplings with the BAW modes. In order to maintain accuracy, we use at least three elements per half wavelength of the resonant wave mode shape. The progressive and gradual change in the element size of the adaptive mesh minimize spurious wave reflections from the mesh discretization itself, and hence reduces spurious wave reflections from the bottom. The mesh allows for superior savings in computer resources.

IV. Three-Dimensional Finite Element Analysis of Periodic SAW Structures.

An obvious three-dimensional effect in SAW resonators is the use of transverse modes in transversely coupled SAW resonators filters[17-21]. Here the SAW propagates essentially in the *X-Y* saggital plane but is coupled transversely in the *Z* direction to another SAW propagating essentially in the same saggital plane.

IVa. 3-d fem analysis of the fundamental SAW mode and its transverse overtones in the Z-direction.

We study the transverse SAW modes of the periodic structure of Fig. 1. One half wavelength of the SAW in the *X* direction is modeled. The wavelength of the SAW is 10 microns. The substrate is quartz ST-cut with Euler angles 0/123/0, and a thickness of 100 microns, while the electrode is aluminum with a length $W = 32$ wavelengths, an *X*-width of one quarter wavelength and a height of 0.3 microns. The length of the unelectroded section *G* is five wavelengths. A finite element mesh of 4x30x30 elements with an adaptive factor of 0.98 was used. The electrode is short circuited. Fig. 2 shows the first five modes of SAW. The first mode is the fundamental SAW mode, while modes 2 to 5 are the transverse overtones of the fundamental SAW mode.

We check the accuracy of the SAW mode and its transverse overtones, by comparing their frequencies with those of the measured frequencies of a SAW resonator in Fig. 3. Only modes 1, 3 and 5 are shown in the figure because the SAW resonator was designed not to have antisymmetric modes. The calculated frequencies of the transverse overtones of SAW match reasonably well with the measured data.

The changes in frequencies of the transverse overtones of the fundamental SAW modes as a function of the electrode length *W* are compared with the measured experimental data in Fig. 4. In the figure, the transverse overtone frequencies are normalized with respect to the fundamental SAW mode frequencies. We observe a good fit of fem frequencies with experimental data.

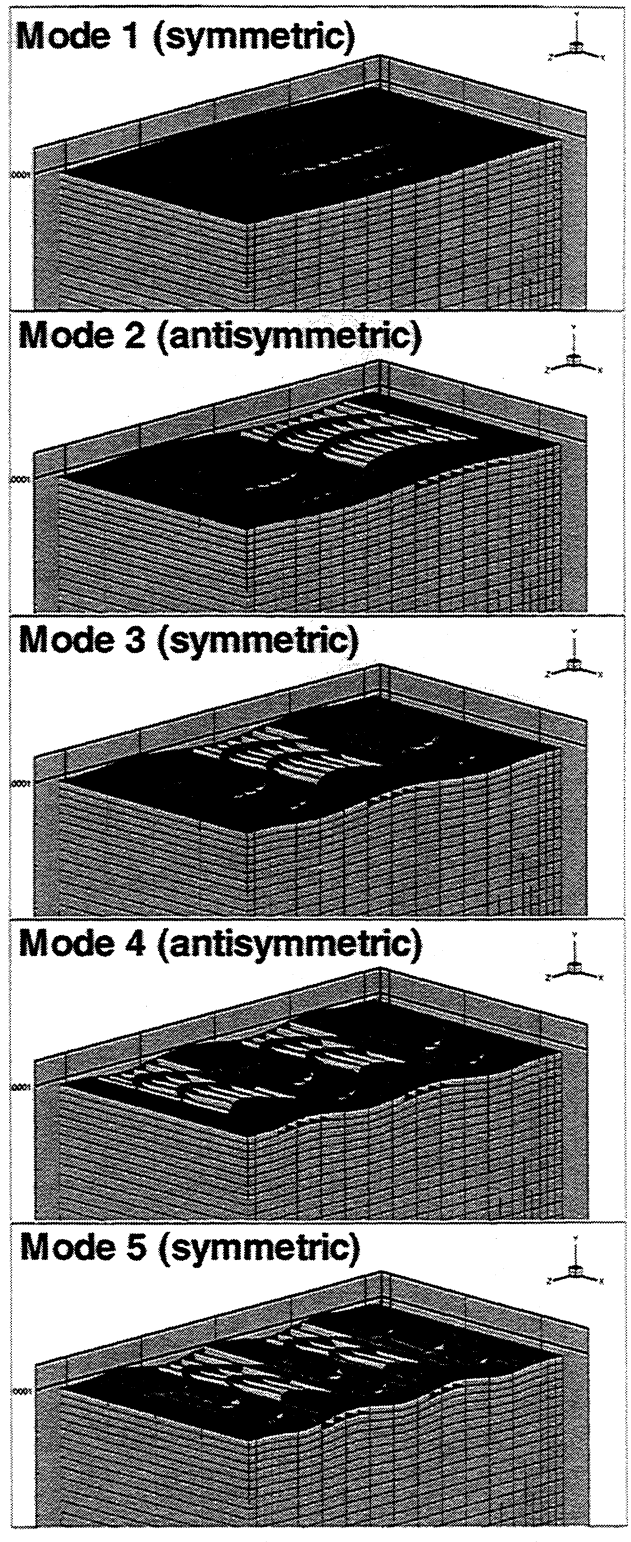


FIG. 2: Fundamental SAW mode (Mode 1) and transverse overtone of the fundamental SAW mode (Mode 2 to 5)

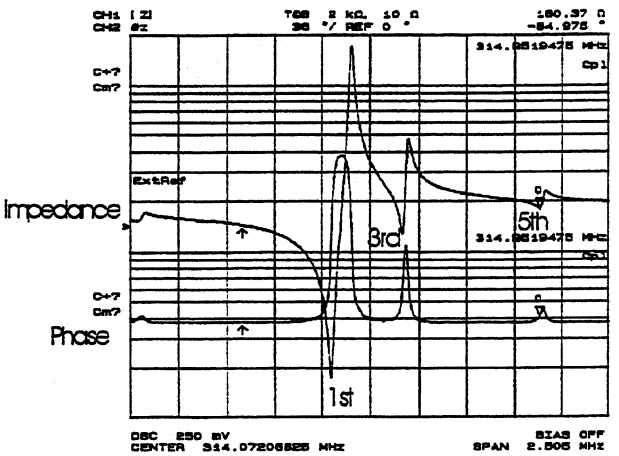


FIG. 3: Comparison of fem frequencies of fundamental SAW mode and its transverse overtones with measured data

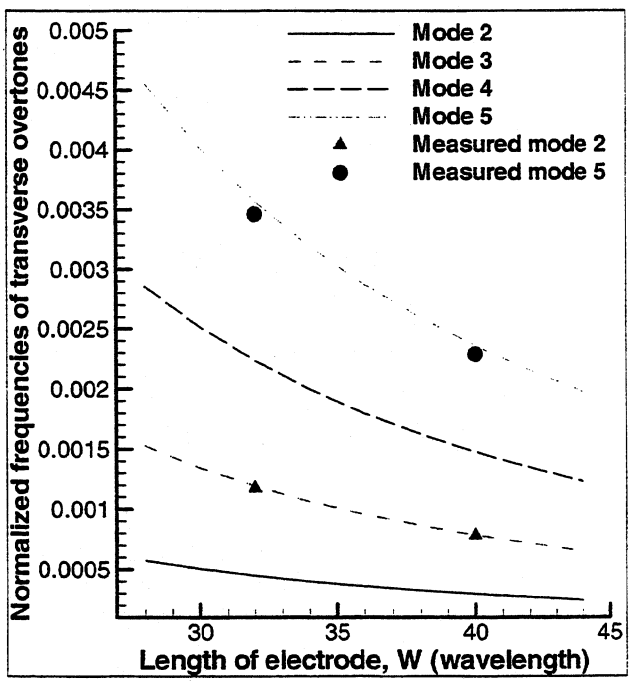


FIG. 4: Comparison of fem frequencies of the transverse overtones of the fundamental SAW mode with measured data. (Frequencies are normalized with respect to the fundamental SAW mode frequencies)

IVb. Frequency-temperature results from the 3-d fem analysis of the fundamental SAW mode and its transverse overtones in the Z-direction.

The frequency-temperature behavior of a SAW resonator filter is usually an important attribute. In order to gain further confidence in our fem models, we calculated the frequency-temperature characteristics of the SAW periodic structure of Fig.1, and compared the results with measured data and also with the results from the two-dimensional fem programs of our previous work[8].

Fig. 5 shows the 3-d fem frequency-temperature curve of the fundamental SAW mode (wavelength = 10 microns) compared with the measured curve and the results from our 2-d fem program. The electrode length W is 32 wavelengths, height is 0.02796 wavelength. The 2-d curve does not take into account of the electrode length but as can be seen from the figure, it coincides well with the 3-d curve. The measured data fits reasonably well with the calculated results.

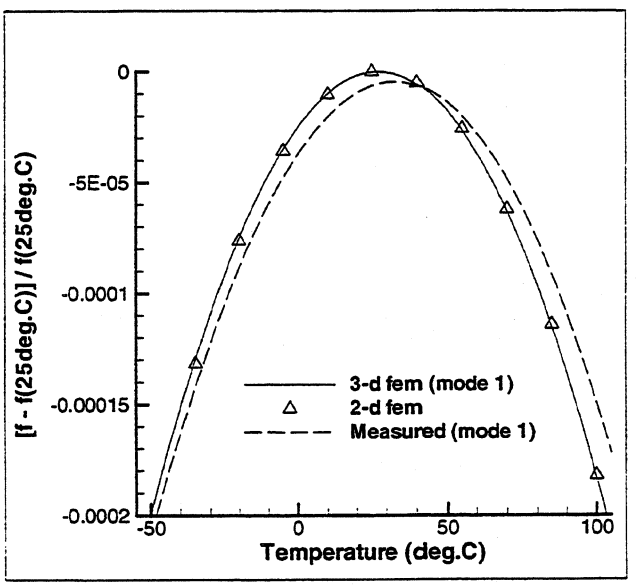


FIG. 5: Frequency-temperature curves of the fundamental SAW mode (mode 1) from the 3-d fem, 2-d fem and experimental data.

The frequency-temperature behaviors of the transverse overtones of the fundamental SAW mode are also studied. A summary of the results is provided in Tables 5 and 6 which tabulate, respectively, the turnover temperatures and the second order temperature coefficients of the frequency-temperature curves. Results for two electrode lengths, W=32 wavelengths and W=40 wavelengths, are shown.

In Table 5, we see that the calculated turnover temperatures are off by about 5 to 9 degrees Celsius compared to the measured values. We think that these are reasonably good comparisons because the turnover temperatures are quite sensitive to crystal cuts and electrode heights. Reasonably good comparisons are also seen in Table 6 for the second order temperature coefficients. In both tables, the

changes due to a change in electrode length W from 32 wavelengths (height=0.2959microns) to 40 wavelengths (height=0.2796 microns) are captured reasonably by the 3-d fem model.

Table 5: Turnover temperatures (degrees C)

Electrode:	W=32 wavelengths height=.2959 microns			W=40 wavelengths height =.2796 microns		
Mode:	mode 1	mode 3	mode 5	mode 1	mode 3	mode 5
Measured:	30.4	30.4	29.92	32.09	32.06	33.14
3-d fem:	21.09	21.77	23.25	26.66	27.04	27.88

Table 6: 2nd order temperature coefficient ($x10^{-8}/C^2$)

Electrode:	W=32 wavelengths height=.2959 microns			W=40 wavelengths height =.2796 microns		
Mode:	mode 1	mode 3	mode 5	mode 1	mode 3	mode 5
Measured:	-2.87	-2.87	-3.06	-3.30	-3.30	-3.30
3-d fem:	-3.44	-3.43	-3.41	-3.39	-3.38	-3.37

V. Summary

Piezoelectric, Lagrangian equations for the frequency-temperature behavior of quartz was presented. Results of three-dimensional finite element models of a half wavelength SAW periodic structure are compared with measured data from SAW resonator . The frequency characteristics and their frequency-temperature behavior of the fundamental SAW mode and its transverse overtones are studied. Generally good results are obtained by the 3-d fem models when compared with measured values. The effects of changes in the electrode length W are also reasonably well predicted by the 3-d periodic structure. The 2-d fem models of reference 8 are shown to yield results consistent with the 3-d fem models.

Acknowledgment

Support by Seiko Epson Corporation is gratefully acknowledged

References

1. Y-K. Yong, R. Garon, S. Kanna and K-Y. Hashimoto, "Effects of Periodically Missing Fingers and Periodically Shifted Fingers on SAW Propagations in Quartz Resonators," *Proceedings of the 1998 IEEE International Frequency Control Symposium*, 1998, pp. 461-469.
2. M. Koshiba, K. Hasegawa and M. Suzuki, "Finite Element Analysis of Finite Periodic Waveguides for Surface Acoustic Waves," *Japanese Journal of Applied Physics*, Vol. 24 (1985) Supplement 24-1, pp. 112-114.
3. M. Koshiba, S. Mitobe, and M. Suzuki, "Finite Element Solution of Periodic Waveguides for Acoustic Waves," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, Vol. UFFC-34, No. 4, July 1987, pp. 472-477.
4. M. Koshiba and K. Ohbuchi, "An Analysis of Surface Acoustic Wave Devices Using Coupling of Mode Theory and Finite Element Method," *Japanese Journal of Applied Physics*, Vol. 30(1990) Supplement 30-1, pp. 140-142.
5. P. Bauerschmidt, R. Lerch, J. Machui, W. Ruile and G. Visintini, "Reflection and Transmission Coefficients of SAW in a Periodic Grating Computed by Finite Element Analysis," *Proceedings of the 1990 IEEE Ultrasonics Symposium*, 1990, pp. 421-423.
6. M. Buchner, W. Ruile, A. Dietz, and R. Dill, "FEM Analysis of the Reflection Coefficient of SAWS in an Infinite Periodic Array," *Proceedings of the IEEE 1991 Ultrasonics Symposium*, 1991, pp. 371-375.
7. K. Hasegawa and M. Koshiba, "Finite Element Analysis of Periodic Surface Acoustic Waveguides," *Japanese Journal of Applied Physics*, Vol.34 (1995) pp. 2642-2645.
8. Y-K Yong and S. Kanna, "IDT Geometry and Crystal Cut Effects on the Frequency-Temperature Curves of a SAW Periodic Structure of Quartz", *Proceedings of the 1998 IEEE Ultrasonics Symposium*, 1998, pp. 223-228.
9. B. K. Sinha and H. F. Tiersten, "First Temperature Derivatives of the Fundamental Elastic Constants of Quartz", *Journal of Applied Physics*, 50(4), 1979, pp. 2732-2739.
10. B. Dulmet and R. Bourquin, "Application of Lagrangian Effective Material Constants to the Study of the Thermal Behavior of SAW Propagation in Piezoelectric Crystals", *Proceedings of the 1994 IEEE Ultrasonics Symposium*, 1994, pp. 331-336.
11. P.C.Y. Lee and Y-K. Yong, "Frequency-Temperature Behavior of Thickness Vibrations of Doubly Rotated Quartz Plates Affected by Plate Dimensions and Orientations", *Journal of Applied Physics*, 60(7), 1986, pp. 2327-2342.
12. Y-K. Yong, "Three-Dimensional Finite Element Solution of the Lagrangian Equations for the Frequency-Temperature Behavior of Y-Cut and NT-Cut Bars", *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, Vol. 34, No.5, Sept. 1987.

13. R. Bechmann, "Elastic and Piezoelectric Constants of Alpha-Quartz", *Phys. Rev.*, Vol. 110, June 1958, pp. 1060-1061

14. R. Bechmann, A.D. Ballato, and T.J. Lukaszek, "Higher Order Temperature Coefficients of the Elastic Stiffnesses and Compliances of Alpha-Quartz", *Proceedings of the IRE*, Vol. 50, no. 8, 1962, pp. 1812-1822.

15. Landolt-Bornstein: Numerical data and functional relationships in science and technology, vol.2. Bechmann R., Hearman R. F., Kurtz S. K.: Elastic, piezoelectric, piezooptic constants, and nonlinear dielectric susceptibilities of crystals. Springer Verlag, Berlin 1969

16. Landolt-Bornstein: Numerical data and functional relationships in science and technology, vol.2. Bechmann R., Hearman R. F., Kurtz S. K.: Elastic, piezoelectric, piezooptic constants, and nonlinear dielectric susceptibilities of crystals. Springer Verlag, Berlin 1979

17. C.S. Hartmann, P.P. Abbot, S. Jen and D.P. Chen, "Distortion of Transverse Mode Symmetry in SAW Transversely Coupled Resonators Due to Natural SPUDT Effects", *Proceedings of the 1994 IEEE Ultrasonics Symposium*, 1994, pp. 71-74.

18. J. Tsutsumi, O. Ikata and Y. Satoh, "Transversely Coupled Resonator Filters with 0.1% fractional Bandwidth in Quartz", *Proceedings of the 1996 IEEE Ultrasonics Symposium*, 1996, pp. 65-69.

19. D.P. Morgan, S. Richards and A. Staples, "Development of Analysis Techniques for SAW Transverse-Coupled Waveguide Resonator Filters", *Proceedings of the 1996 IEEE Ultrasonics Symposium*, pp. 177-181.

20. D.C. Molocha, B.P. Abbott and S.M. Knapp, "Theoretical Predictions and Experimental Results for Transverse Modes in SAW Cavities", *Proceedings of the 1997 IEEE Ultrasonics Symposium*, 1997, pp. 1-4.

21. M. Solal, J. Desbois, "A New Low Impedance Balanced Drive Structure for SAW Transversely Coupled Resonator Filters", *Proceedings of the 1997 IEEE Ultrasonics Symposium*, 1997, pp. 83-88.

Swept Frequency Acoustic Time Domain Reflection Measurements

Scott K. Frederick, Mitch Chou, Donald C. Malocha
Consortium for Applied Acoustoelectronic Technology
University of Central Florida, Orlando, Florida 32816-2450

Abstract- An inverse Fourier transform of frequency domain S11 reflection measurements obtained from a cube shaped piezoelectric material produces highly accurate time domain reflection data with the aid of mathematics and processing. A completely automated system has been developed to produce time domain reflection measurements over a temperature range of 200°C. The target of the process is the analysis and extraction of elastic constants of new man made piezoelectric materials such as langasite and langatite.

Introduction

Traditionally, velocity measurements in solids have been made using time-based pulse-echo techniques [1,2,3,4,5]. These techniques used a short, wide-band pulse, or a gated RF CW signal, to excite a transducer and create an acoustic wave in a solid. The reflected waves are then measured and recorded in real time. The techniques have limitations due to the required time-base accuracy and have typically been accomplished by manual measurements. Calibration of delay of the experimental setup is often difficult and needs great care. Dynamic range and signal to noise is limited due to the finite input pulse length.

This paper presents an approach to time domain echo analysis by using swept frequency measurements and then signal processing using Fourier transform techniques. The approach is efficient and accurate over any frequency range of interest, and can be fully automated using computer control and network analyzers.

Principle of Operation

The use of swept frequency measurements for time domain echo analysis is achieved by using Fourier transform techniques. The approach discussed uses measurements of S11 over frequency of a transducer bonded to a material under test. The approach can also be extended for full two port measurements if two transducers are used to make transmission measurements. Since the transducer is band limited, a finite and rather small number of frequency data points need to be taken in order to meet all the necessary Nyquist criteria for sampling. Since these are swept frequency measurements, the automatic network analyzer can be set for optimum settings in power, IF bandwidth, and sweep time, to obtain signal to noise ratios of approximately 100 dB. Because of the limited bandwidth of the transducer, the measured S11 out of band is approximately unity. This allows the use of data padding of thousands of frequency points without any frequency measurements. Since time resolution is determined by the highest frequency in a data set, padding allows for enhanced resolution in the time domain, via the inverse Fourier transform.

Matlab is used as the data processing platform and is used to extend the data frequency bandwidth and consequently improve time resolution when the Inverse Fast Fourier Transform is performed. Data processing in the frequency domain is used to eliminate the direct electrical reflection; leaving only the acoustic echo responses. Data processing is also used in the time domain to obtain the echo delay times.

Experimental Setup

For this presentation, precisely cut and measured 1cm $\pm 10um$ cubes are mounted with a 10 MHz center frequency coax style $LiNbO_3$ transducer, using a sub-micron-thick layer of photoresist as the bonding agent. The transducer style allows for electrical connection to a single side of the transducer. An HP 8753 automatic network analyzer is used to obtain swept frequency S11 measurements. The electrical connection is made via two pogo pins inserted in a fabricated 50 ohm coaxial line. The electrical assembly is calibrated using short, open, and 50 ohm terminations to the ends of the pogo pins to remove delay and measurement errors. LabWindows based software is used to interface and control the network analyzer and to take a data run of 1601 points in the frequency domain from 1 MHz to 19 MHz. A separate unit is used to control temperature and provide a temperature output reading to the computer. Data files are saved over temperature and input into a Matlab program for processing.

A custom designed solid copper test fixture with a cavity for the crystal/transducer combination was manufactured. The fixture contains miniature pogo pins to contact the transducer. Two part construction of the fixture allows for adding/removing the cube. The fixture has cylindrical holes to hold OMEGA cartridge heaters and a final hole to feed wires and coax for measurement of data and provide thermocouple access.

Aluminum "compression" fixtures were also designed to mount transducers to cubes. HR-120 photoresist is spun

This work is supported under a contract from the U.S. Army, Fort Monmouth, NJ, Contract #N66001-97-C-8634.

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

onto a face of the cube. A small aluminum piece fits directly over the cube. This piece has a 5mm hole, and allows for centering of the transducer. The transducer/cube is then inserted into the spring-loaded compression fixture and baked for approximately 30 minutes at 150 °C.

Subsequently, the transducer/cube is inserted into the test fixture and is attached to cabling. The analyzer uses an industry standard GPIB interface to communicate with a computer controller. A Yokogawa UP-550-01 temperature controller is also connected to the computer using an RS-232 serial interface.

Temperature profiles are downloaded to the Yokogawa controller, which runs the profiles independently of the computer controller. Profiles run slowly to allow the cube to saturate for 5 minutes at a given temperature. Only the temperature output of the Yokogawa controller is used by the computer controller.

National Instrument's LabWindows v5.0 is the software used with the computer controller. A modified version of an existing 8753 driver controls the data acquisition. When doing a "data-run" over temperature, the LabWindows program monitors the temperature output of the Yokogawa, then acquires and saves S11 data according to an algorithm designed to take data only when the temperature has been stable for several minutes.

Calibrated data from the network analyzer is transferred across the GPIB interface and automatically saved by the computer controller. The data contains real and imaginary components at each frequency. The network analyzer has the IF bandwidth set at the minimum 300 Hz and typical input power is 0 dBm. At this time, the data from the network analyzer is not averaged, and none of the smoothing features on the network analyzer are used. A typical plot of S11 is shown in *Figure 1*.

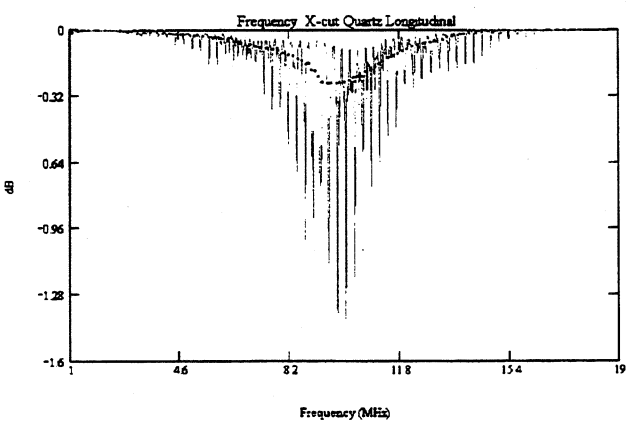


Figure 1: *S11 measured data in dB vs. frequency (solid line). S11 smoothed data in dB vs. frequency (dotted line).*

Data Processing

Referring to *figure 1*, the large ripples are due to the acoustic echo reflections. If the cube were infinite in dimension, there would be no ripple (acoustic reflection) and there would be a smooth S11 profile corresponding to the transducer bandwidth when bonded to the cube. This electrical reflection can be approximated by smoothing the measured data. This was accomplished by convolving, in the frequency domain, the measured data with a Hamming window. The resulting smoothed curve is almost identical to the S11 that would be obtained with a cube of infinite length (no reflection), leaving only the electrical S11 (see *figure 1*). This smoothed curve is then subtracted from the original raw data to yield the pure acoustic frequency response curve.

The smoothed, limited bandwidth data files are then expanded by "padding" out to a required frequency which is a factor of 2. This is accomplished by gently rolling up to 1dB from both ends; the first frequency point at 1 MHz down to zero frequency, and the last frequency point at 19 MHz up to maximum frequency. This allows for the IFFT to be performed. *Table 1* shows a summary of the time and velocity resolutions that can be obtained for our 1.0 cm cubes from the various bandwidths created by padding with different numbers of frequency data points.

# of Points	Bandwidth	Time Resolution 1/BW	Equivalent Velocity Resolution	Equivalent Length Resolution	% Error
1601	18 MHz	56.6 ns	34 m/sec	277.6 um	0.680 %
2^15	368.63 MHz	2.71 ns	1.7 m/sec	13.6 um	0.034 %
2^16	737.27 MHz	1.36 ns	0.8 m/sec	6.8 um	0.016 %
2^17	1.475 GHz	0.68 ns	0.4 m/sec	3.4 um	0.008 %
2^18	2.95 GHz	0.34 ns	0.2 m/sec	1.7 um	0.004 %
2^19	5.9 GHz	0.17 ns	0.1 m/sec	0.85 um	0.002 %

Table 1: *Time resolution obtained for various numbers of frequency data points. Resolution is the expressed in equivalent velocity and length resolution for the 1.0 cm cubes tested.*

The resultant IFFT of padded frequency data has resolution in time that is inversely proportional to the bandwidth obtained. As an example, data with 2^{15} points in the frequency domain corresponds to a velocity error of only 1.7 m/s for a 2 cm total path length. *Figure 2* shows the plots of the corresponding time data from the measured and smoothed S11 frequency data, respectively, after the IFFT. Note that the reflection magnitude in time, starting

at t=0, is reduced tremendously for the pure acoustic data. The peak is reduced approximately 3 orders of magnitude.

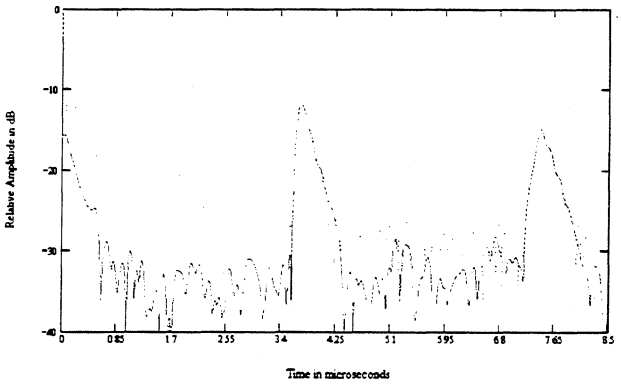


Figure 2: IFFT of raw data in dB vs. frequency (dotted line). IFFT of pure acoustic data in dB vs. time (solid line). First two reflections only.

Multiple data runs taken over a temperature from -50 to +150°C are automatically input into a Matlab program. A 3-D plot of magnitude vs. time over temperature is shown in *figure 3.*

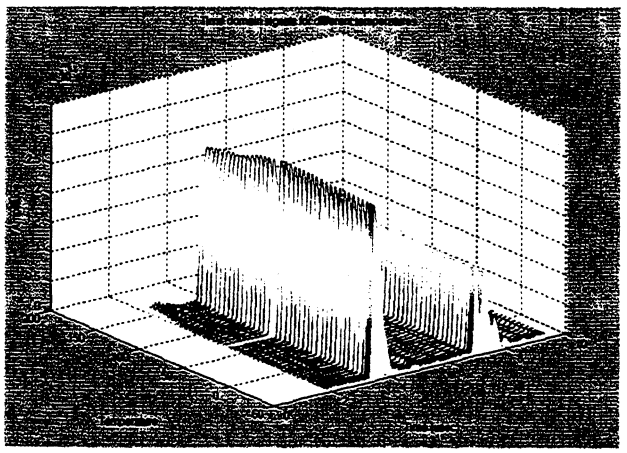


Figure 3: 3-D IFFT plots of pure acoustic data (longitudinal mode, X-axis quartz) in linear magnitude vs. time over temperature format. First two reflections only.

Once the data is in the system, a wide range of signal processing is possible. As an example, using the derivative of the curves in *figure 3* produces a useful plot for finding time delays between reflections, since it provides the point of maximum slope in the leading pulse edge. This data can be used to find the relative time between pulses with a simple search algorithm. If the pulse shapes are not well behaved, the slope method works reasonably well. The 3-D plot of just the first two echoes shows the uniformity of the time echo data at 172 measured temperatures. The plot is easily rotated to show a useful 2-D plot of delay times vs. temperature.

Experimental Results

Time resolution can be increased to the point where the quantization error is masked by the thermal noise of the network analyzer. Having evaluated the resolution of frequency data files with 2^{17} points, quantization error appears to still dominate at this resolution of 0.68 ns.

In time, the arrival of the first pulse is determined by a combination of magnitude and phase delay, essentially looking for a definite start in either curve, and assuming the first significant change to actually signify the beginning of an incoming reflected wave. Once a method of obtaining a true reflection time from the time data is found, relative measurements of peaks and/or derivatives of time domain data can be used and referenced to the absolute point. This will greatly simplify analysis of data for materials that support multiple modes of propagation of waves travelling at similar speeds.

Ultimately, velocity data is used to extract the elastic constants of the piezoelectric material under study. For complete elastic constant derivation, velocity measurements must be made on the X, Y, Z, and 45° axis of precisely cut cubes of the material under test.

Figure 4 shows a typical plot of the extracted velocity versus temperature for a quartz cube measured at 172 different temperatures. Thermal expansion of the cube is taken into account. The data was obtained using the point of steepest slope on the leading edge of the first reflected pulse as the marker for delay time when processing. Time delays are then adjusted to the initial pulse start, which represents the actual delay time. The IFFT contained 2^{16} points.

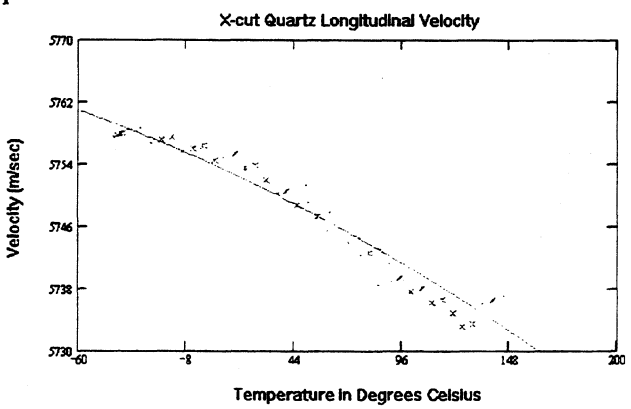


Figure 4: Plot of longitudinal acoustic wave velocity along the X-axis of a quartz cube over temperature.

Figure 5 is a plot of the elastic constant C11 versus temperature for a quartz cube based on the velocity data of *figure 4.* This represents one of the first automatic data runs. At high temperature, the data has more scattering than expected and the cause may be in the actual measurement error, properties of the adhesive at high

temperature, or may be due to the processing of the data. Further work is required to find the error sensitivities and to optimize the system.

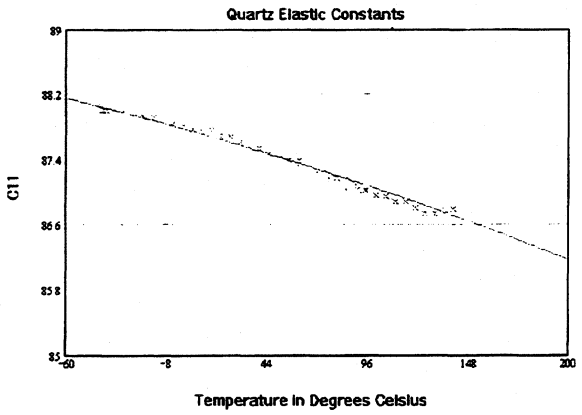


Figure 5: Experimentally obtained plot of the C11 elastic constant for quartz over temperature.

Quartz data is used to verify the accuracy of the measurement methods discussed [1,6,7]. The slope of the curves in figures 4 and 5 is somewhat steeper than that found in published data, with lower temperature velocity data being more accurate. Data in both figures 4 and 5 compares within 0.2% of results obtained by B.J. James [1].

Conclusion

This paper has presented a method for measuring time domain pulse echoes using swept frequency measurements and Fourier transform techniques. The system can be fully automated and signal processing can be used to enhance the measured and processed data. Results of the fully automated system are shown, and the first results obtained with the data acquisition are shown. Further work is required to optimize the data processing and to maximize accuracy in measuring the pulse time delays.

Acknowledgements

The authors are grateful for the support from the U.S. Army, FT. Monmouth, NJ. We appreciate the material and technical support furnished by Piezo Technology Incorporated of Orlando, FL. Key individuals include: Gary Archibald, who assisted us with the temperature controller, Robert Smythe, on technical issues, and Robert Helmboldt who fabricated excellent quality cubes. The materials studied are provided by Crystal Photonics, Inc., Orlando, FL.

References

[1] B. J. James, Thesis, "Determination of the Elastic and Dielectric Properties of Quartz," Royal Holloway and Bedford New College, London Univ., Spring 1987.

[2] B.J. James, "A New Measurement of the Basic Elastic and Dielectric Constants of Quartz," *42nd Annual Frequency Control Symposium*, pp. 146-154.

[3] G. L. Petersen, B. Chick, W. Junker, "Error Correction Methods in Gated Amplifier Absolute Velocity Measurements and Comparison with the Pulse-Echo-Overlap Technique," *IEEE Ultrasonics Symp. Proc. 1975*, Cat. #75 CHO 994-4SU.

[4] W. Soluch, "Measurements of Elastic and Piezoelectric Constants of $Li_2B_4O_7$ Crystal," *11th European Frequency and Time Forum*, Neuchatel, pp. 382-385, Mar. 1997.

[5] M.A. Breazeale, J.H. Cantrell Jr., J.S. Heyman, "Ultrasonic Wave Velocity and Attenuation Measurements," *Methods of Experimental Physics, vol. 19, pp. 67-135*, Academic Press, Inc., 1981.

[6] J. C. Brice, "Crystals for Quartz Resonators," *Reviews of Modern Physics, vol. 57, no. 1*, pp.105-146, Jan. 1985.

[7] D. Salt, *Hy-Q Handbook of Quartz Crystal Devices*, Van Nostrand Reinhold Co. Ltd., Berkshire, England, 1987.

1999 Joint Meeting EFTF - IEEE IFCS

DESIGN AND FABRICATION TECHNOLOGY OF LOW LOSS AND HIGH FREQUENCY SAW DEVICES FOR MOBILE RADIO PHONES

JUN YAMADA

Semiconductor & Integrated Circuits Division, Hitachi Ltd.,
6-2, Otemachi 2-Chome, Chiyoda-ku, Tokyo 100-0004 Japan
Tel: +81-3-5201-5017,Fax: +81-3-3270-5120
E-mail: jun-yamada @ denshi.head.hitachi.co.jp

ABSTRACT

This paper describes designs for IIDT and band-eliminate filters with insertion loss of less than 3dB, and fabrication technology for aluminum electrodes enduring some watt-class electrical power for high frequency filters in mobile radio phones.

1. INTORODUCTION

In recent years the subscribers of mobile radio phones have been rapidly increasing in many countries, and R&D activities for the small-sized and light-weighted portable phones are strongly accelerated. In 1983 the phone had the volume of about four liters and weight of four kilograms. Since then they have been gradually miniaturized, and they are enough handy to carry at the present time. The volume and the weight are 68 milliliters and 68 grams respectively. They are about one-sixtieth of the previous ones for the last fifteen years. High integrated RF/IF, digital baseband LSIs and fine ASIC technology contribute the miniaturization of mobile phones. SAW devices also reduce the size of the transmitter and receiver circuits in RF and IF portions.

Figure 1 shows the RF and IF circuits in cellular radio. The meshed blocks can be replaced by SAW devices. They are duplexer, RF and IF filters. This paper describes the front filters in the duplexer. The main requirements of SAW filters are low insertion loss for power reduction, high power durability in watt-order operation, and high selectivity for spurious signal suppression.

The duplexer has transmitter and receiver filters in it. The target specifications of them are shown in table 1. Insertion loss of the transmitter filter is required less than 2dB, and needs to endure the power of more than one watt. The resonator filters with IIDT or band-eliminate structure are suitably designed to achieve low insertion loss for receiver or transmitter filter Aluminum compound electrodes sputtered by a DC magnetron improve power durability of SAW devices. And fine fabrication process with reactively ionized dry-etching attains highly precise pattern replication.

2. DESIGN OF LOW LOSS FILTERS

2.1 Improved Circuit Model

The equivalent circuit model for the plain resonator has been re-examined before the concrete design of the resonator filter. The stored energy effect B which is caused at the electrode edge is the important parameter for the resonator design. And the conventional model considers the effect as a parallel susceptance at the electrode edge, which is shown in Figure 2 (a). However, when the electrodes are thick enough against the wavelength in the UHF range, the model is not sufficient for the experimental results.

The improved circuit model has been proposed which contains the stored energy effect factor B in the transmission phase angle, shown in Figure 2 (b). Then the improved model is varified by the experiments. Figure 3 shows the experimental results of single mode SAW resonators which have the peak frequency of 668 MHz. The resonators have 60 paired normal IDTs and 500 reflectors on the ST-quartz. The electrode thickness is

0-7803-5400-1/99/ $ 10.00

0.104 μm and the duty factor is 45 %. The improved circuit model sufficiently explains the experimental results, compared with the conventional model.

2.2 Design of IIDT Filter

The low insertion loss and high outband suppression are required for the receiver filter in the duplexer. The IIDT filter has been investigated to meet the requirements, and the many design parameters are optimally determined.

The IIDT filter shown in Figure 4 has seven IDTs and they are symmetrically lined up at different intervals. And two identical IIDT filters are cascaded to suppress outband responses sufficiently. In this case 28 design parameters should be suitably determined to satisfy the specification. They are maximum overlap-length in all IDT, six intervals between IDTs, and each IDT has three parameters-finger numbers, center frequency and weighting function, which are shown in the table 2. And the initial and the final values for the parameters are also listed in the table. Here, the weighting function is defined that it is equal to Hamming window when the value is 1. Optimization process for the parameters is as follows, which is shown in Figure 5.

(1) give same values for every IDT
(2) numerical calculation by improved circuit model
(3) determine error functions between calculated values and the specifications
(4) square the error functions and sum them up along frequency
(5) change values of the parameters to reduce the summed error functions
(6) iteratively calculate by the circuit model
(7) stop the calculation when the summation is small enough

Frequency characteristics of the IIDT are shown in Figure 6. Experimental and calculated values are expressed by the solid and the dotted lines respectively. The 36 degrees Y-X $LiTaO_3$ substrate is used here, and the insertion loss of 2.5dB and outband suppression of 30dB have been obtained.

2.3 Design of Band-Eliminate Filter

Extremely low insertion loss of less than 2dB is required for the transmitter filter in the duplexer. The SAW band-eliminate filter shown in Figure 7 is investigated for the purpose. The filter is constituted of many series connected SAW resonators. They have passband characteristics near their resonant frequencies, and have stopband characteristics near their antiresonant frequencies. The resonant frequencies of the respective IDTs are slightly different one another. The frequency characteristics of the transmitter filter is shown in Figure 8. Insertion loss of 1.3dB in passband is achived in this case. The calculated and experimental values are in good-agreement on many points.

3. FABRICATION TECHNOLOGY

3.1 High Power Durability

High power durability of more than 1 watt is required for the transmitter filter in the duplexer. When we apply some watt-class electrical power on the UHF SAW devices with pure aluminum electrodes, we observe electrode-damage and degradation of device performance by two stages. At first, many hillocks appear on the electrodes and they grow periodically up on them in SAW propagation path. Next, electrodes adjoined each other are short-circuited when the hillocks grow larger, then the electrodes are melted, or occasionally the wafer is cracked. Two causes on these phenomena are considered. One is mechanical stress which damages directly the electrodes. The other is large density of electric current, which accelerates destruction as temperature rising.

Various electrode materials constituted as plain SAW resonators are studied at high power operation, which are shown in Figure 9. The vertical axis means high power durability, and it is defined as time of resonant frequency shift. The test is performed at the temperature of

120℃, and the failure time is largely accelerated. The horizontal axis means SAW stress, which is calculated from loading power, resonant Q and cavity length. This figure shows sputtered aluminum-0.8% titanium film is more durable than sputtered pure aluminum, aluminum-copper film and evaporated aluminum films. Besides the DC magnetron sputtering has better reproducibility of film composition than the conventional e-gun evaporation, and it has been verified by ICPS analysis on additive concentration in the films.

3.2 Dry Etching for Fine Electrodes

Dry etching technology is indispensable for stable fabrication of SAW electrodes with less than 1μm. Several aluminum etching techniques have been developed, and they are compared from the viewpoints of SAW application, as follows. CF_4 gas in the barrel-type plasma etching reacts chemically silicon and silicon dioxide, but it rarely responds to metal film. The physical reaction in the sputter etching and the ion milling have strong anisotropy of etching rate, and it causes easily substrate damage. So they are inapplicable to SAW device fabrication. Therefore the etching technology of RIE or microwave plasma ECR which executes both chemical and physical reaction at the same time has been investigated for fine SAW fabrication. The microwave plasma ECR is relatively expensive because it needs higher vacuum chambers, so we inquire process optimum of reactive ion etching technology.

However some problems remain in RIE process, shown in table 3, and countermeasures are considered as follows. CF_4 gas is mixed into BCl_3 gas to eliminate accumulated aluminum chloride ($AlCl_3$) which adheres to water and vapor, because CF_4 gas convents hygroscopic aluminum chloride ($AlCl_3$) to sublimative aluminum fluoride (AlF_3). The mixture achieves high etch-rate and no electrode-corrosion at the same time. And low bomband energy of gas particles avoids substrate damage. The results are confirmed by emission spectrum, RHEED and SAW velocity dispersion.

SEM photographs of SAW electrodes fabricated by wet and dry etching are shown in Figure 10. Many small tails and trapezoids in cross-section are observed in wet etching. Compared with them, dry etched electrodes have perpendicular cliffs and no tails.

4. CONCLUSION

The SAW filters, required low insertion loss, high power durability and fine electrode pattern, have been studied for the miniaturized mobile phones.

The IIDT and band-eliminate filters by SAW resonators are optimally designed with an improved equivalent circuit model, and satisfy the target specifications of insertion loss and outband rejection in the front filters.

Ti added Al electrodes by DC magnetron sputtering have obtained both higher power-durability and better reproducibility than conventional e-gun evaporated Al-Cu electrodes.

Al-reactive ion etching has achieved highly precise electrode-pattern fabrication with damage free and no corrosion for high frequency SAW devices.

REFERENCES

[1] J.Yamada, " Technical trends and future evolution of surface acoustic wave devices ", IEEE 1996 Freq.Cont.Symp.Proc., pp.183-187, 1996

[2] M.Hikita, H.Kojima, T.Tabuchi and Y.Kinoshita, " 800 MHz high performance SAW filter using new resonant configuration ", IEEE Trans., Microwave Theory Tech., vol. MTT-33, No. 6, pp.510-518, 1985

[3] J.Yamada, T.Shiba, N.Hosaka and T.Toyama, " High stability SAW resonators in UHF range ", IEEE 1986 Ultrasonics Symp., Proc., pp.267-272, 1986

[4] J.Yamada, T.Shiba, N.Hosaka and T.Buma, " Equivalent circuit model analysis of SAW resonators in UHF range ", Report of the 150 th Committee on Elastic Wave Device Technology, pp.23-28, March 1986 (in Japanese)

[5] H.Ohnuki, N.Hosaka, J. Yamada and T.Kobayashi, "Computer aided design of low-loss SAW filters emplying IIDT for high performance mobile radio duplexer ", IEEE 1990 Ultrasonics Symp.Proc., pp.123-127, 1990

[6] J.Yamada, N.Hosaka, A.Yuhara and A.Iwama, "Sputtered Al-Ti electrodes for high power durable SAW devices ", IEEE 1988 Ultrasonics Symp., Proc., pp.286-290, 1988

[7] J.Yamada and A.Yuhara, " Dry process technology for high frequency SAW devices", IEEE 1989 Ultrasonics Symp., Proc., pp343-349, 1989

[8] J.Yamada and A.Yuhara, " Precise fabrication technology of high frequency and high power durable SAW filters for mobile radio communication ", Proc. Int. Symp. SAW Devices for Mobile Comm., Japan Society for the Promotion of Science, pp.148-158, 1992

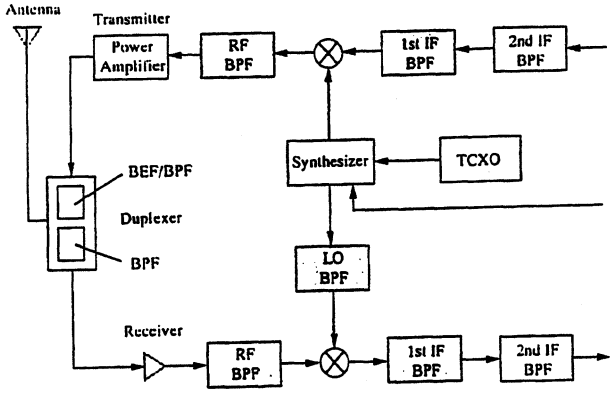
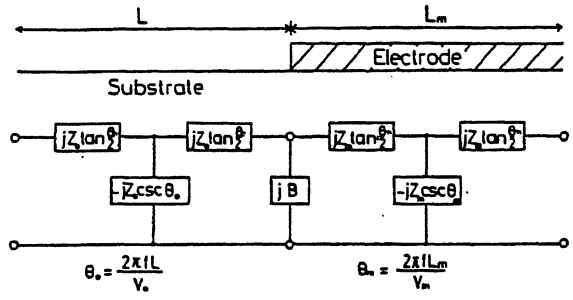


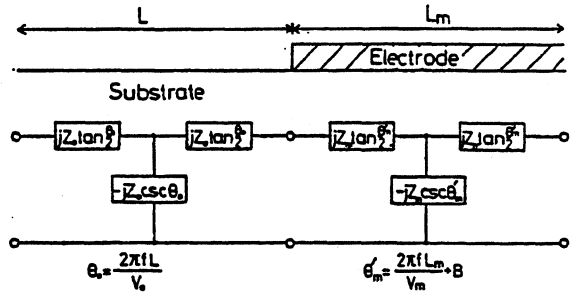
Fig. 1 RF Circuit in Mobile Radio Phones

Table 1 Target Specifications of Front Filters in Duplexer (Japanese Digital Cellular, class Ⅲ)

	Transmitter	Receiver
Center Freq. (MHz)	948 (= f_T)	818 (=f_R)
Bandwidth (MHz)	16	16
Loss (dB)	≤ 1.8	≤ 2.5
Outband Rejection (dB)	≥ 20 (f_R) ≥ 35 ($2f_T$)	≥ 22 (f_T) ≥ 40 (f_R + 2 IF)
High - Power Durability (W)	≥ 1.0	——



(a) Conventional Model



(b) Improved Model

Fig. 2 Equivalent Circuit Model for SAW Reflectors, Allowing for Stored Energy Effect

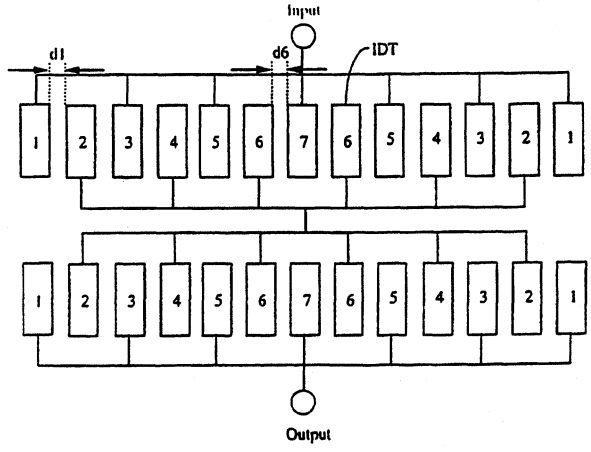
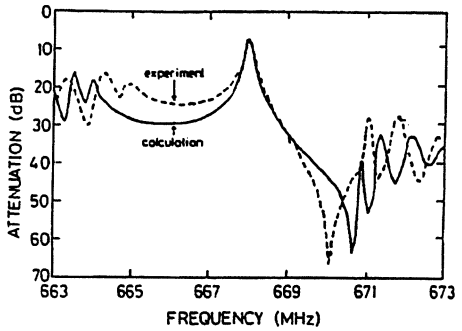
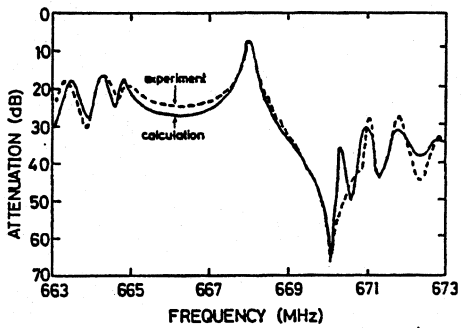


Fig. 4 Constitution of IIDT Filter



(a) Conventional Model



(b) Improved Model

Fig. 3 Comparison between Experiment and Calculation

Table 2 Design Parameters and Values

Parameter		Initial	Final
Max.overlap length (μm)		110	90
Distance d_1		63.0	67.7
Distance d_8		63.0	67.6
IDT 1	Nmber of fingers	45	35
	Center frequency (MHz)	827.5	815.5
	Weighting	1.2	18
⋮			
IDT 7	Nmber of fingers	45	49
	Center frequency (MHz)	827.5	827.5
	Weighting	1.2	5.0

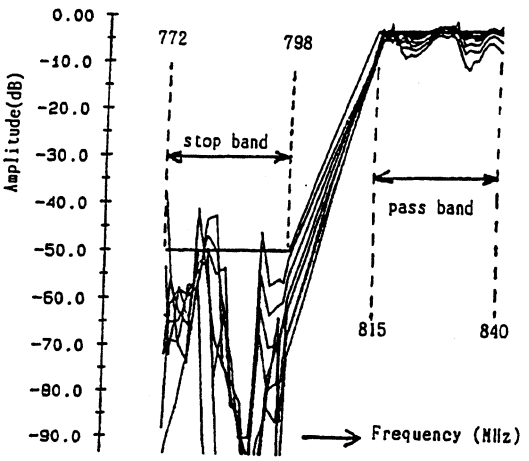


Fig. 5 Transition of Frequency Characteristics in Optimization Process

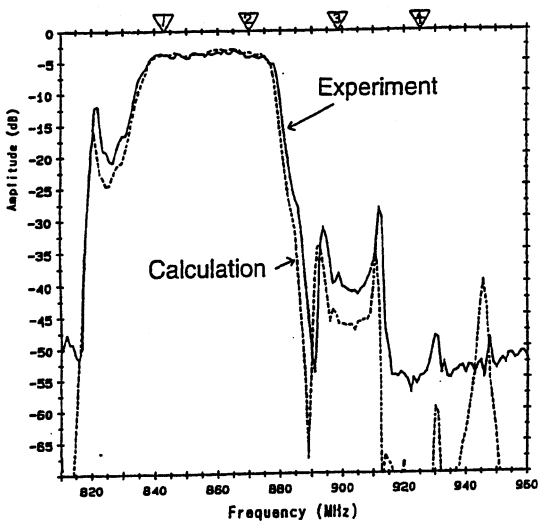


Fig. 6 Frequency Characteristics of IIDT Filter (Receiver Filter)

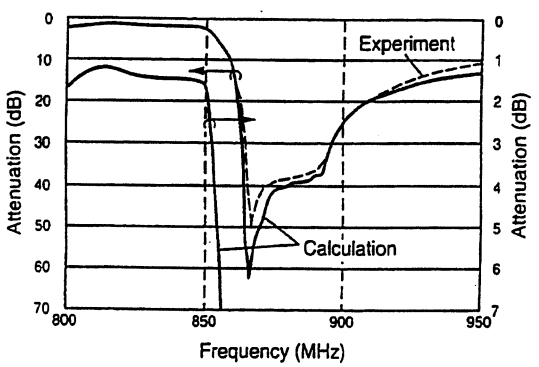
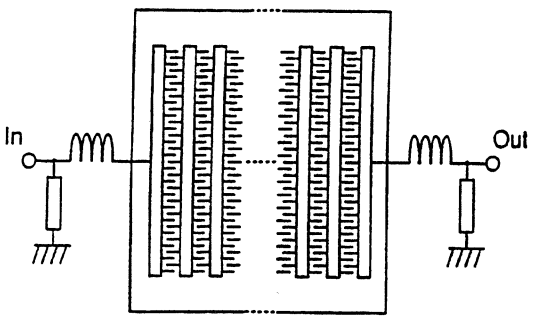
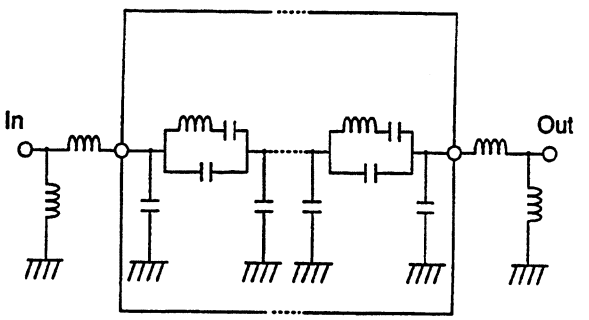


Fig. 8 Frequency Characteristics of Band-Eliminate Filter (Transmitter Filter)



(a) electrode constitution



(b) equivalent circuit model

Fig. 7 SAW Band-Eliminate Filter

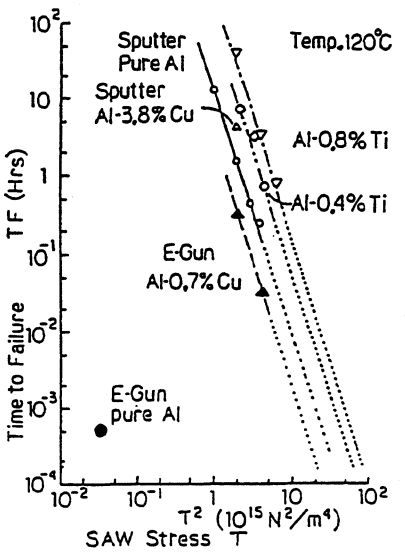
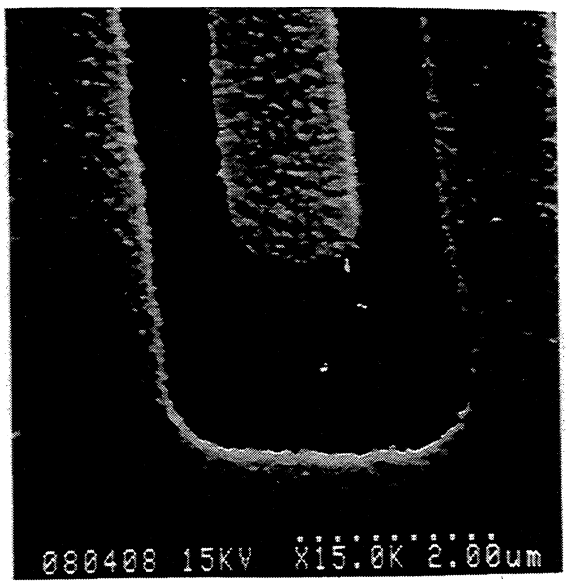


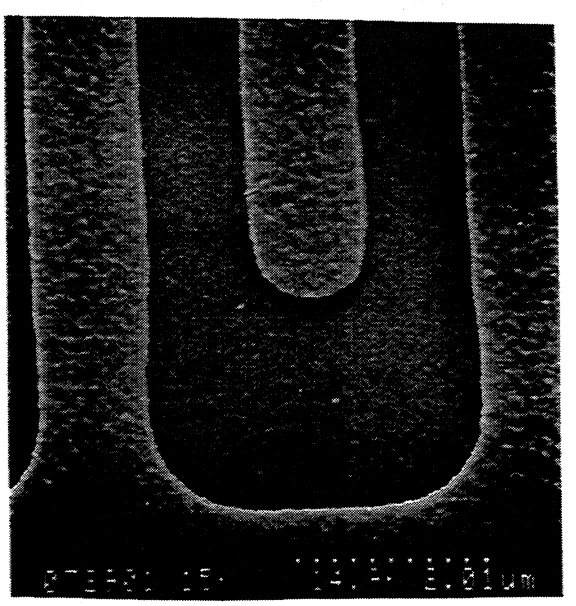
Fig. 9 High Power Durability

Table 3 Problems and Countermeasures of RIE

Problems	Countermeasure	Assessment
Etch Rate	BCl_3+CF_4	
Corrosion	CF_4, O_2 Sputter	Emission Spectrum
Damage	Low Particle Energy	RHEED SAW Velocity Dispersion
Cost	High Yield High Throughput	



(a) Wet etching



(b) Dry etching

Fig. 10 SEM Photographs

REFLECTOR-FILTER USING AN SAW WAVEGUIDE DIRECTIONAL COUPLER FOR IF APPLICATIONS IN CDMA SYSTEM

Jun Tsutsumi, Takashi Matsuda, Osamu Ikata and Yoshio Satoh

FUJITSU LABORATORIES LTD., 64, Nishiwaki, Ohkubo-Cho, Akashi 674-8555, Japan

Abstract

The reflector-filter design is an attractive technique to obtain steep skirt characteristics in a short chip length, because the SAW propagation path is folded and the frequency response is synthesized by utilizing both IDT and reflector responses. In this paper, we propose the new reflector-filter structure using an SAW wavuguide directional coupler to realize small-size CDMA IF filters.

For the implementation of proposed reflector-filter, the key technology is the design of SAW waveguides. The Al grating, which has the stopband kept away from that of IDTs, is first employed to form a low-loss waveguide. The directional coupler is composed of two Al grating waveguide, and the coupling property is verified.

Using the proposed reflector-filter structure, PCS-CDMA IF filters are fabricated on quartz substrates. The filter exhibits an insertion loss of 8.5 dB and fulfills the specified frequency response. The package size is 4.8×9.1 mm and half the size of a conventional transversal filter is achieved.

1. Introduction

IF(Intermediate Frequency) filters for CDMA(Code Division Multiple Access) system have to satisfy stringent requirements, including high shape factor and phase linearity. One of the conventional solutions is the transversal filter design using two strictly weighted IDTs. In this filter design, however, a large number of electrode fingers are required to fulfill the specified frequency response. As a result, the filter size becomes extremely large, specifically in the SAW propagation direction.

The reflector-filter design makes the filter size much smaller, because the SAW propagation path is folded and the frequency response can be synthesized by utilizing both IDT and reflector responses. The key technology to develop the reflector-filter is to separate the reflected SAWs from the incident waves without energy loss and deterioration of temperature stability. Several types of the reflector-filter structure have been reported so far[1-3]. However, it seems that they are not sufficient to obtain the practical performance for CDMA applications.

In this paper, we propose the new reflector-filter structure that realizes the separation of reflected waves using an SAW waveguide directional coupler. At first, the principle will be explained and then the implementation of the directional coupler is described. At last, the experimental results of PCS-CDMA IF filters fabricated on quartz substrates will be presented and the advantage of proposed reflector filter structure will be shown.

2. Principle of the proposed reflector-filter

The structure of proposed reflector-filter is shown in Fig.1. The filter consists of two tracks. An IDT, a waveguide and a reflector are placed on each track. The tracks are coupled with each other by two closely placed waveguides composing a directional coupler. Here, the propagation length in the directional coupler is set to half the complete coupling length.

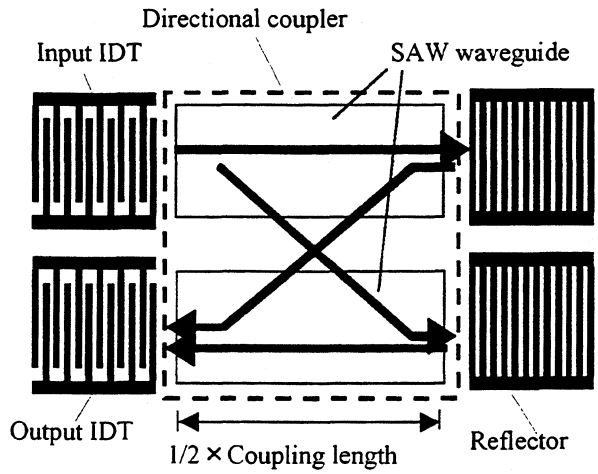


Fig.1 Structure of the proposed reflector-filter

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

In the configuration, SAWs excited by the input IDT on the upper track travel through the directional coupler, and an equal amount of the SAW energy reaches the reflector in both tracks. The reflected waves in both tracks travel again through the directional coupler. As a result, the whole SAW energy is transferred to the lower track because the total propagation length in the directional coupler is equal to the complete coupling length, and detected by the output IDT. Hence, this reflector-filter has no excess loss except the bi-directional losses of IDTs theoretically. Besides, temperature stability is conserved because the orientation of SAW propagation doesn't change.

In the proposed reflector-filter structure, the waveguide directional coupler has been employed because the coupling length can be controlled by the waveguide parameters, including the waveguide width and the gap width. This means that the compact directional coupler can be formed even on a low electromechanical coupling substrate, such as quartz. Therefore, the filters with the high temperature stability are obtained.

Although the multi-strip coupler may also be used as a directional coupler, the coupling length is inversely proportional to the electromechanical coupling coefficient[4]. Therefore, it can never be used on quartz.

3. Designs of a waveguide and a directional coupler

In the implementation of the proposed reflector-filter, the design of the directional coupler is the key procedure for the miniaturization of the total filter size. We investigated the directional coupler on ST-quartz to realize the temperature stable CDMA IF filters.

First, we investigated the SAW waveguide composing the directional coupler. In order to form the core region of the waveguide where the velocity is slower, the SAW velocity under various surface conditions on ST-quartz were examined. Fig.2 shows the SAW velocity under free surface, the uniform Al film and the Al grating. We have proposed to employ the Al grating as the core region for the flexibility of waveguide design. With the core of Al grating, one can choose either free surface or the uniform Al film for the cladding region because the velocity under the Al grating is the slowest. Furthermore, the velocity under the Al grating can be tuned by controlling the line width of the grating fingers. Besides the advantage of design flexibility, the Al grating waveguide doesn't need any additional fabrication process. From these considerations, it is confirmed that the Al grating is suitable for the core region of the waveguide on quartz.

Although the Al grating causes the reflection of waves, it can be avoided by keeping the stopband of the waveguide away from the passband of the filters.

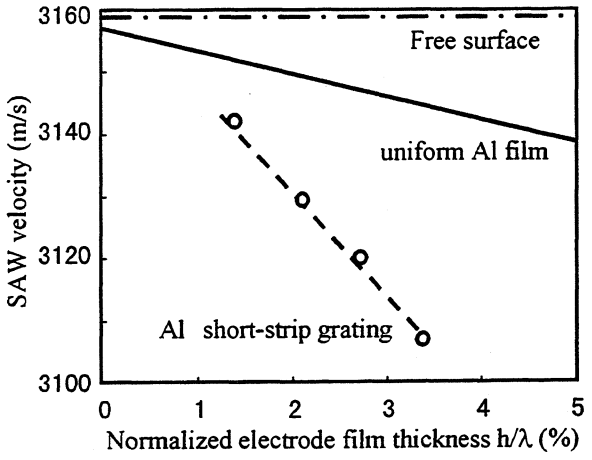


Fig.2 SAW velocity on ST-quartz

In order to evaluate the Al grating waveguide on ST-quartz, we measured the propagation loss using test samples shown in Fig.3. In this case, the cladding regions were formed with the uniform Al film. The Al grating waveguide was placed between the input and output IDTs, and the variations of the insertion loss were measured. The electrode period λ of the IDT was 15 μm and the period of Al grating was set to 5.7 μm to avoid the reflection of the waves.

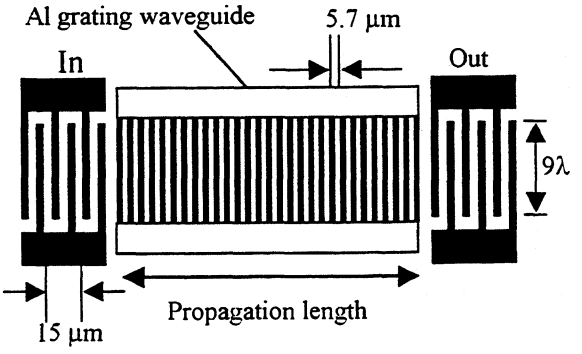


Fig.3 Test sample for the evaluation of SAW waveguide

Fig.4 shows the dependence of the insertion loss on the propagation length. The insertion loss under free

surface is also shown with open circles for comparison. It is seen that the insertion loss is almost constant with the Al grating, while the insertion loss increases drastically under free surface, as the propagation length becomes longer. Moreover, the value of the insertion loss with the Al grating was smaller by more than 3 dB at the propagation length of 200λ. From these experimental results, it is verified that the Al grating has the guiding effect and can be used as a low-loss SAW waveguide.

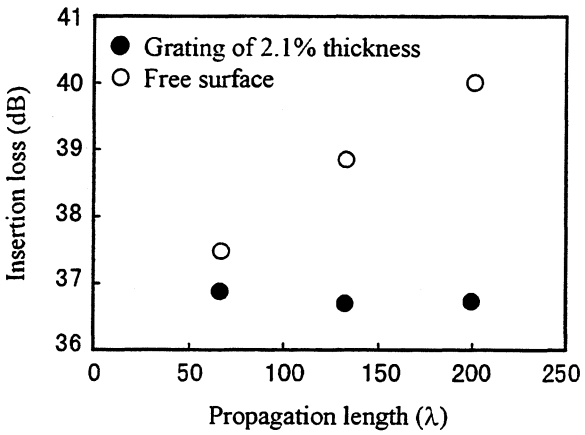


Fig.4 Propagation loss of Al grating waveguide

Next, we fabricated the directional coupler using two Al grating waveguides, and evaluated the coupling property. Fig.5 shows the structure of test samples. The input IDT was placed on the upper track and the output IDT was placed either on the upper or on the lower track.

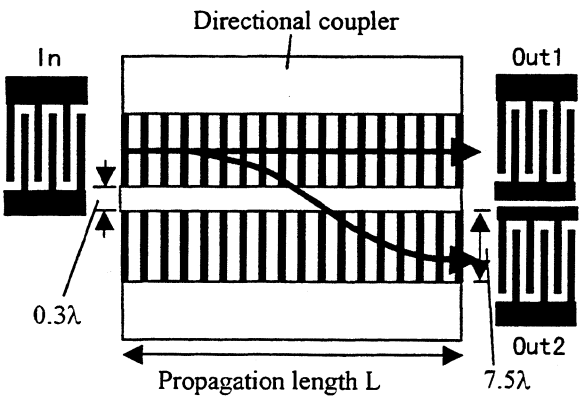


Fig.5 Test sample for the evaluation of directional coupler

Here, the waveguide width was 7λ and the gap width was 0.3λ.

Fig.6 shows the variations of insertion loss versus the propagation length. The insertion loss increased in the case that the output IDT was placed on the upper track. On the other hand, the insertion loss decreased for the lower track. From these results, we found that the coupling between the waveguides occurred and the SAW energy was transferred from the upper to the lower track. Moreover, the experimental plots agree well with the theoretical curves for the coupling coefficient between the waveguides $\kappa=250\ m^{-1}$ and the coupling length is estimated approximately 400λ. The coupling length of 400λ is much shorter than that of the multi-strip coupler, and the compact directional coupler was realized on quartz substrate.

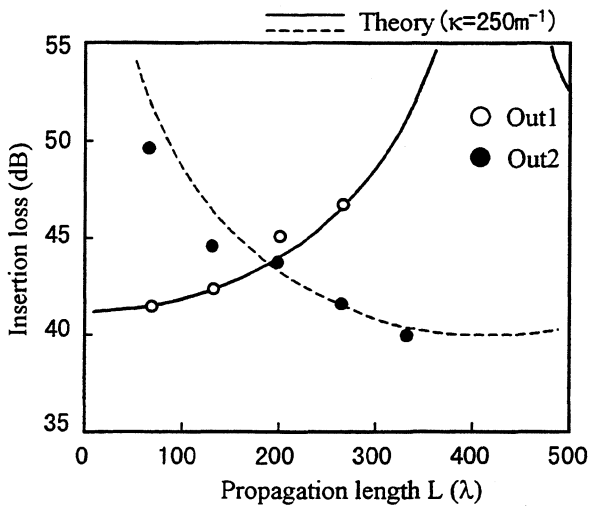


Fig.6 Coupling property of the directional coupler

4. Experimental results of PCS-CDMA IF filters

Based on the above experimental results, we designed PCS-CDMA IF filters. The unidirecitonal transducers were used for both of the input and the output IDTs. To fulfill the specified frequency response, only the reflectors were weighted. Fig.7 shows the frequency response. The measurements were performed with the external matching circuits. The 5 dB band width of 1.45 MHz and 33 dB band width of 2.39 MHz were obtained and the requirement for the shape factor has been achieved. The insertion loss was 8.5 dB and the group delay ripple was less than 500 ns. Fig.8 shows the time response of the filter. The suppression level of the triple transit echo (TTE) signal was 34 dB. Because of the low TTE level, all

the frequency responses were obtained without the time gating procedure. The package size of the developed filters was 4.8×9.1 mm and half the size of 6.5 × 13.3 mm package used for the conventional transversal filter has been achieved. From these experimental results, it is confirmed that the proposed reflector-filter structure is effective to miniaturize the filter size.

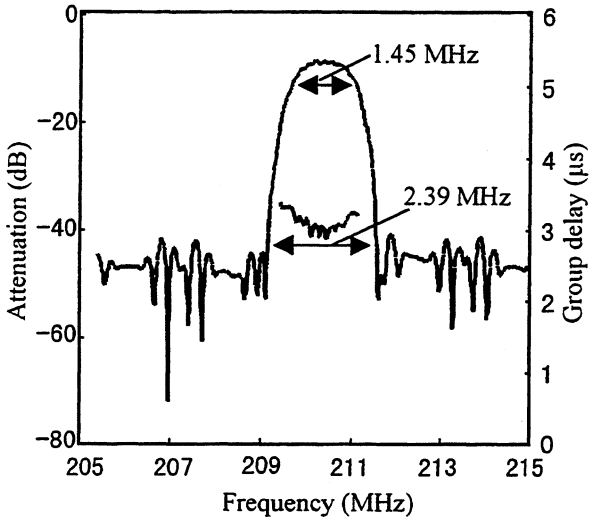


Fig.7 Frequency response of PCS-CDMA IF filter

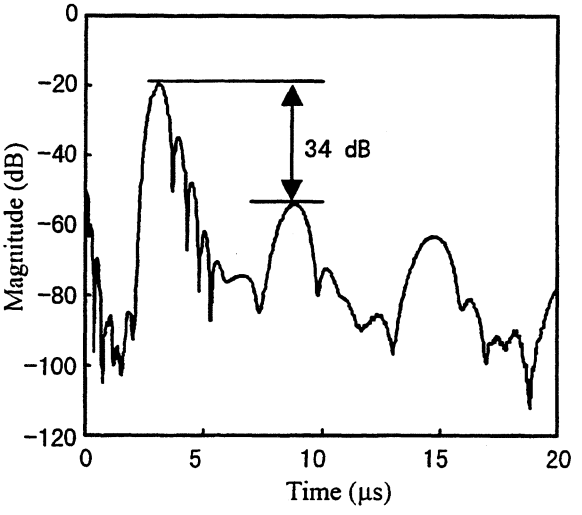


Fig. 8 Time response of the filter

5. Conclusion

The new reflector-filter structure has been presented to realize the small-size CDMA IF filters. The problem of the separation of reflected waves has been overcome using the SAW waveguide directional coupler. The compact filter can be obtained even on low electromechanical coupling substrate with the proposed reflector-filter structure, because the coupling length of the waveguide directional coupler can be controlled by the waveguide parameters, instead of electromechanical coupling coefficient.

The Al grating waveguide has been first proposed to form the directional coupler on quartz substrate and the coupling property of the directional coupler has been verified. PCS-CDMA IF filters have been fabricated and the requirements for the frequency response have been satisfied. The package size was 4.8 × 9.1 mm and half the size of the conventional transversal filter design has been achieved.

6. Acknowledgement

The authors wish to thank Prof. M. Yamaguchi, Prof. K. Hashimoto and Dr. T. Omori for many valuable discussions.

Reference

[1] G. Riha, H.R. Stocker, R. Veith and W. Bulst, "RAC-filters with position weighted metallic strip arrays", IEEE Proc. Ultrason. Symp., 1982, pp.83-87.

[2] J.Machui and W.Ruile, "Z-path IF-filters for mobile telephones", IEEE Proc. Ultrason. Symp., 1992, pp.147-150.

[3] A. Bergmann, H. Machui, K. Wagner and R. Weigel, "Two-track-reflector-filters for CDMA mobile telephones", IEEE Proc. Ultrason. Symp., 1996, pp.57-60.

[4] F.G.Marshall and E.G.S.Paige, "Novel acoustic-surface-wave directional coupler with diverse applications", Electron. Lett., vol.7, 1971, pp.460-464.

4-POLE SAW-COUPLED-RESONATOR FILTERS WITH SOLELY ACOUSTIC PROXIMITY COUPLING

Thomas W. Johannes

Siemens AG, Corporate Technology, 81730 Munich, Germany,
email: thomas.johannes@mchp.siemens.de

ABSTRACT

Transverse coupled SAW resonator filters (TCF) are widely used as interfrequency filters in the range between approximately 100MHz and 500MHz. They offer a high stopband rejection, steep skirts and insertion losses well below 10dB. Common bandwidths range between 500ppm and slightly above 1000ppm. In this paper operation of TCFs with relatively high 3dB-bandwidths is considered.

For the realized 4-pole filters resonator coupling is performed entirely by evanescent acoustic coupling between adjacent resonator tracks as proposed by Coldren in 1979 [1].

Thereby the detuning of the centre resonators and the resulting deterioration of the passband shape in conventional electrically coupled 2-pole TCF structures is avoided.

In electrically coupled TCFs the sensitivity to the inductivity value of a required external coil increases with increasing bandwidth and thus limits wideband operation of these filters. Since in entirely proximity coupled TCFs the coupling strength is determined by photolithography the sensitivity problem is avoided.

Two filters are presented: one with a 3dB-bandwidth of 292kHz and the other with 118kHz corresponding to relative bandwidths of 1400ppm and 3500ppm. The minimum insertion losses are 2.5dB and 3.0dB respectively for filters mounted on printed circuit boards (PCBs) and matched to 50Ω. A selectivity of 29dB has been achieved for the narrowband filter driven unsymmetrically. The chip length is only 6.8mm for both devices for centre frequencies of 83MHz.

In the paper the following items are addressed: electrically coupling of SAW-TCFs, the sensitivity on discrete external coupling elements, electrical crosstalk and the influence of weak resonances of the uncoupled resonators.

1. INTRODUCTION

Surface-Acoustic-Wave (SAW) resonator filters are the current choice for interfrequency application in mobile phones. This is essentially a consequence of the resonant operation which leads to small chip sizes and insertion losses.

In this paper transverse coupled filters (TCFs) are investigated, a class of SAW coupled-resonator filters with compact size in comparison to other techniques. Current state-of-the-art TCFs provide stopband rejection levels up to 70dB. Their passband skirts are usually determined by a four-pole transfer function and by the commonly used Butterworth design, i. e. with a flat passband.

Their operation principle relies on the fact that there is a variation in the phase velocity perpendicular to the propagation direction. In SAW TCFs the phase velocity is lowest in the resonator's grating region which is due to the energy storage effect. Similar to optical dielectric waveguides as e. g. in optical fibres acoustical waves are confined to the low velocity region of this waveguide. In fact acoustical plane waves are entirely reflected at the boundary between a homogeneous grating and homogeneously metallized busbars provided the angle between the wave vector and the boundary line is below the criticle angle for total reflection.

Interdigital transducers (IDTs) acting as waveguides are placed between acoustical reflectors so that all the acoustical energy is confined within the resonator. Consequently acoustical radiation losses can be made arbitrarily small which is the reason for the low losses in these filters.

Identical resonators are placed side by side so that there is an overlap between the acoustical evanescent fields and energy may be transfered from one track to the other. The coupling scheme is depicted in Fig. 1 together with the corresponding modes' shapes and the velocity profile. Coupling increases with decreasing distance resulting in an increasing splitting of resonances [2]. The amount of resonance splitting determines the filter bandwidth directly. It is however limited by the difference in phase velocity between the grating and the surroundings.

To get a flat passband the input and output impedance has to supply a characteristic positive real impedance level supposed the transducers static ca-

0-7803-5400-1/99/ $ 10.00

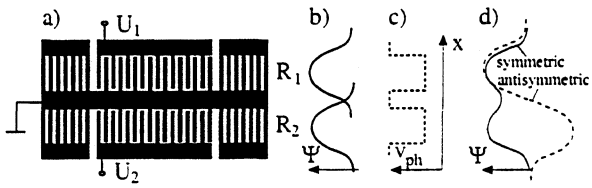


Figure 1: a) schematic representation of a TCF 2-track substructure comprising two identical resonators R1 and R2; b) unperturbed mode shapes; c) velocity profile; d) coupled mode shapes

pacitance is compensated. This characteristic impedance level increases proportionally to the resonance splitting and therefore to the bandwidth supposed the motional elements as defined by Fig. 2 remain the same. To achieve a high relative bandwidth beside a high velocity difference also small resonator apertures have to be chosen. Otherwise unwanted additional resonances will occur. As a consequence for high relative bandwidths a small aperture is necessary and hence the impedance level is additionally increased. This limits the possible range of operation to low frequencies where the influence of parasitic capacities is relatively small. In Fig. 2 the

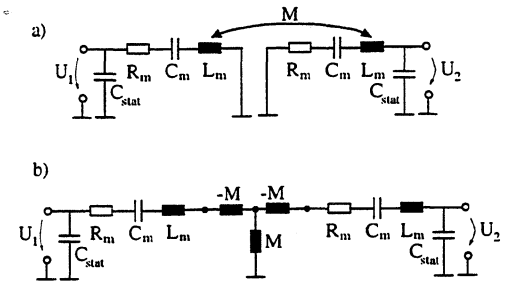


Figure 2: a) Equivalent electrical circuit of a TCF 2-track substructure with magnetic coupling; b) magnetic coupling resolved by inductivities.

equivalent electrical circuit for the 2-pole substructure from Fig. 1 valid in the vicinity of the passband is shown. One resonator is modeled by its motional inductance L_m, its capacitance C_m and its resistance R_m. The static capacitance of the interdigital transducer is denoted by C_{stat}. The coupling between two identical resonators is modeled by the mutual inductivity M, which is assumed to be positive. Modeling the magnetic coupling by its equivalent T-network yields the representation in Fig. 2 b).

2. ELECTRICAL COUPLING OF SAW-RESONATORS

By appropriate matching at the input and output port of the TCF-substructure shown in Fig. 2 a Butterworth 2-pole filter characteristic may be achieved. For this case the transfer function has two poles for positive frequencies.

To achieve steeper skirts and higher stopband rejection identical 2-pole TCF-substructures are usually cascaded to four-pole structures. To adjust the coupling strength between those substructures either a capacity or inductivity is additionally connected which is also shown in Fig. 3 for the case of an inductivity.

If not compensated by an inductivity the maximum impedance of the coupling stage is determined by the static capacitance as $Z = -j/2\omega C_{stat}$. Using an additional capacity the impedance and consequently the coupling strength is further reduced allowing only for small fractional bandwidths. To achieve a high bandwidth compensation is achieved by using a coil as determined by the following eq. (1). The correct coupling strength between the two TCF-

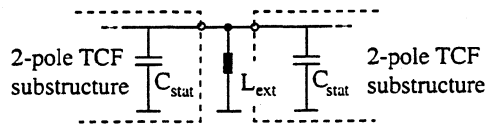


Figure 3: Equivalent circuit for coupling of two TCF substructures with an external inductivity.

substructures with $M = M_{12}$ and $M = M_{34}$ given by $M = M_{23}$ has to be set appropriately. Therefore the coupling stage given by the T-network for the proximity coupling (Fig. 2.b)) and the parallel resonant circuit for the electrical coupling scheme (Fig. 3) are compared:

Firstly, it follows that the external inductivity L_{ext} must be chosen as

$$L_{ext} = \frac{M_{23}}{1 + \omega^2 M_{23} C_\Sigma} \qquad (1)$$

to compensate C_Σ, the sum of the static capacity C_{stat} of the transducers and of the package C_{pack} at operation frequency $f = \omega/2\pi$.

Secondly, one realizes that the two inductivities with a value of $-M$ are missing. Thus the motional inductivities of the resonators are increased by $+M$. Consequently, in the case of the described electrical coupling the adjacent resonators become slightly detuned compared to the coupling scheme for an ideal Butterworth filter with $M_{34} = 1.55 M_{23} = M_{12}$ and equal resonators, i. e. equal motional elements [3]. As a consequence the resonances are no longer symmetric to the centre frequency and an undesired deviation from the ideal Butterworth characteristic is the result.

In Fig. 4 magnitudes of calculated transfer admittances $|Y_{21}|$ and transfer functions for the matched state are shown for an electrically cascaded 4-pole TCF-structure with a relatively high bandwidth

around 3500ppm. Solid lines refer to the ideal coupling scheme $M_{34} = 1.55M_{23} = M_{12}$ mentioned above and dashed lines refer to a reduced coupling with $M_{34} = 1.27M_{23} = M_{12}$ to improve matching over the whole passband. Obviously an ideal Butterworth characteristic cannot be achieved in either case and dips within the passband of the transfer function occur. It is obvious and advantageous that

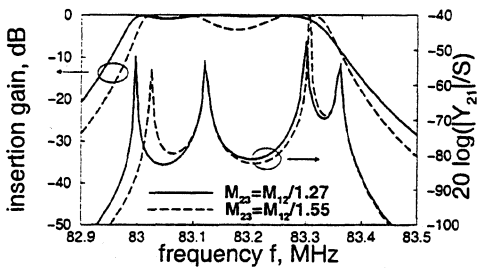


Figure 4: Transfer function for matched, electrically coupled 4-pole TCFs with $M_{23} = M_{12}/1.27$ (solid) and $M_{23} = M_{12}/1.55$ (dashed). Also shown are the corresponding transfer admittance magnitudes $|Y_{21}|$. $R_m = 500\Omega$, $C_m = 0.76$fF, $L_m = 4.8$mH and $M_{12} = 13.7\mu$H.

this deviation can be avoided if all 4 resonators are coupled by proximity coupling alone. However, for useful filter bandwidths and consequently relatively high coupling strengths the coupling can no longer be determined by overlap integrals between the field distributions of the unperturbed resonators [2]. Instead the acoustic behavior of the TCF structure has to be determined as a whole, i. e. transverse modes extending over all resonator tracks have to be determined. Appropriate calculation schemes have been described in literature to a great extent. It is referred to the simulation tool used in this investigation [4]. In the high coupling regime the shape of the modes differs extremely and the coupling-of-modes (COM) coupling parameter $\kappa = r/p$ is reduced [4] with increasing mode order. Here, r denotes the reflectivity per mechanical period p of the grating. For higher modes this results in a higher group delay of the acoustical reflection coefficients within the stopband for IDTs and reflectors. As a consequence modes of higher order shift to higher frequency in addition to the shift from modal dispersion. This again results in an asymmetry of the resonance frequencies.

3. SENSITIVITY TO VARIATIONS IN THE EXTERNAL COUPLING INDUCTIVITY

The relative sensitivity of the effective mutual coupling inductivity M_{23} to a change in the external coupling inductivity L_{ext} for electrically cascaded 4-pole TCFs is derived by derivation of eq. (1):

$$\frac{\Delta M_{23}}{M_{23}} = \eta \frac{\Delta L_{ext}}{L_{ext}} \tag{2}$$

with

$$\eta = \frac{L_{ext}}{M_{23}} \frac{\partial M_{23}}{\partial L_{ext}} = 1 + \omega^2 M_{23} C_\Sigma \tag{3}$$

It can be shown that by scaling all linear dimensions by an amount of s and also the frequency by $1/s$ the elements of the admittance matrix remain unchanged (it is also assumed that the acoustical attenuation per wavelength remains constant). To compare filters with identical relative bandwidth the product ωM_{23} is therefore kept constant. As follows from Fig. 5 already for relatively low frequencies e. g. at 83MHz the relative sensitivity $\eta = 8.7$ is poor. Furthermore, due to the detuning of the central resonators mentioned above proper matching is not possible and the sensitivity of the transfer function relative to M_{23} is also high. Thus, the overall high sensitivity to changes of the coupling inductivity will avoid practical filter realizations in general. Since in the case of an entirely proximity coupled device the coupling strength is determined by photolithography the sensitivity problem does not occur in this case. Losses in the external coupling element are also avoided.

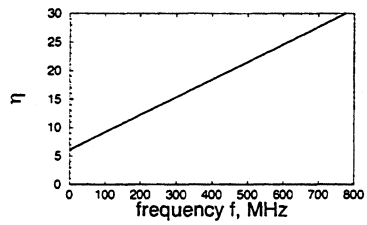


Figure 5: Relative sensitivity of effective coupling mutual inductivity M_{23} relative to a variation in the external inductivity for constant capacity $C_{pack} =$ 1pF but constant product $\omega 2C_{stat} = 4\pi \cdot 83\text{MHz} \cdot$ 1pF and fixed product $\omega M_{23} = 2\pi \cdot 83\text{MHz} \cdot 9.4\mu$H, which is a measure for the relative bandwidth. The sensitivity is given as a function of $f = \omega/2\pi$.

4. ELECTRICAL CROSSTALK

Since the impedance level of resonator filters with fixed resonator number and design increases with increasing bandwidth capacitive crosstalk has a stronger influence on the stopband rejection level for wideband filters.

Due to the adjacent placement of the resonators the crosstalk capacity from the input to the output terminal scales linearly with increasing length for typical TCFs. For common designs also the maximum

values of the admittance matrix elements grows only slightly more than linearly with the length. It is therefore necessary to keep the crosstalk capacity per unit length small to avoid a high crosstalk level.
In Fig. 6 the metallisation pattern of an entirely proximity coupled TCF for unsymmetric operation is shown. To avoid crosstalk as far as possible an ad-

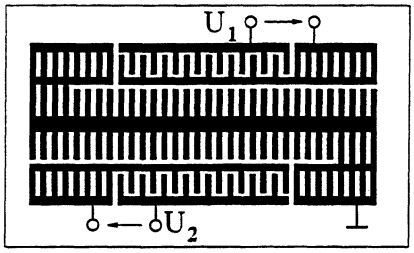


Figure 6: Schematic representation of the filter layout of an entirely proximity coupled TCF.

ditional ground shield is placed between input and output transducers by connecting the centre resonators to ground. The input and output transducers are separated by around 10 acoustical wavelengths. In the case of an electrically coupled filter the central busbars have to carry the signal and therefore have a high impedance to ground. Thus the overall electrical coupling from input to output should be of comparable size or even higher than for purely acoustically coupled TCFs.
In the case of TCFs, which use only two tracks but additionally also multiple longitudinal resonances to achieve a four-pole transfer function [5], the crosstalk level is even higher. This is a consequence of the high capacity between input and output transducers which are separated by a distance of the order of one acoustical wavelength only.

5. EXPERIMENTAL RESULTS

The transfer admittance of an entirely proximity coupled four-pole filter has been presented by Martin et. al. [6]. Here, two designs of 4-pole entirely proximity coupled TCFs are presented together with their transfer functions in matched state. They differ in bandwidth as a consequence of different resonator track separations. Both filters are driven unsymmetrically and the selectivity is limited due to electrostatic crosstalk. The selectivity could however be increased by approximately 20dB using symmetrical signals at in and output. In Fig. 7 the matched transfer function of a filter with a relative bandwidth of 3500ppm is shown. The reflectors comprise 94 and 52 fingers, while the IDT comprises 209 fingers. The metallisation height is 1100nm. The aperture per track is 4.65 wavelengths. As a consequence of the relatively high bandwidth the characteristic impedance is $Z_c \approx 6\text{k}\Omega$. There is a dip in the pass-

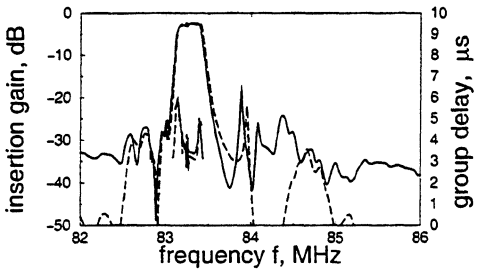


Figure 7: Transfer function and group delay for the filter with 3500ppm 3dB-bandwidth (measurement solid, simulation dashed). No gating has been performed.

band accompanied by a distortion of the group delay. The origin of this dip will be explained in the next section. The resonance associated with a fifths mode is visible at around 83.9MHz as well. The simulation is performed with a tool described in [4]. In the measurement acoustical and electrical crosstalk is visible which was not included in the simulation. The essential parameters like frequency and bandwidth show a good agreement.
In Fig. 8 the matched transfer function of the narrowband filter with a relative bandwidth of 1400ppm is shown. The geometry is the same as for the wideband device however with an increased track separation to reduce coupling. The characteristic impedance is $Z_c \approx 2.3\text{k}\Omega$. Again centre frequency

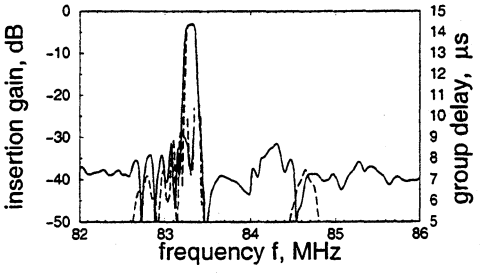


Figure 8: Transfer function and group delay for the filter with 1400ppm 3dB-bandwidth (measurement solid, simulation dashed). No gating has been performed.

and bandwidth is in agreement with the simulation. A selectivity of 29dB has been achieved although the filter is driven unsymmetrically.
Although the filters are symmetric with respect to an interchange of input and output terminal an unsymmetry in the input and output admittance has been observed. Namely the 5th mode which is supported by the wideband filter is excited with different amplitudes by input and output transducer. This difference is attributed to natural SPUDT effects related to the symmetry properties of the phase

of the electrical potential of Rayleigh waves as has been discussed by Hartmann et. al. [7].

6. INFLUENCE OF RESONATOR SIDELOBES

In this section the origin of the dip in the wide-band filter shall be explained with the help of Fig. 9. There the magnitude of the measured transfer admittance is shown together with the contribution from the 4th transverse mode as determined with our calculation tool [4]. According to the simulation

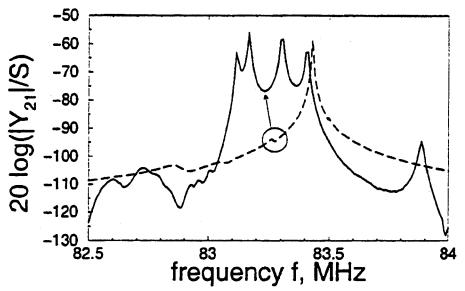


Figure 9: Measured magnitude of the transfer admittance $|Y_{21}|$ (solid) and simulated contribution from the 4th transverse mode (wide bandwidth structure).

procedure the admittance elements are the superposition of the contributions from several transverse modes, which differ in frequency and absolute value but have a similar shape. In the simulation result in Fig. 9 for the 4th mode a weak sidelobe is visible. Although it is not pronounced and cannot be easily observed in the measured transfer admittance it is the origin for the dip in the matched response. This follows from the evaluation of the simulation result shown in Fig. 7.

These sidelobes even occur in uniform transducers without a reflector present but they become even more pronounced when reflectors are added to decrease radiation losses. Thus, the occurrence of sidelobes limits the maximum bandwidth if a flat passband and a groupdelay without major distortions is required.

7. CONCLUSION

Low-loss transverse mode coupled resonator filters (TCFs) with relatively high bandwidths have been presented with solely proximity coupled SAW-resonators. A bandwidth of 3500ppm has been achieved while the insertion loss is 2.5dB for a filter matched on a printed circuit board. The chip size is only 6.8mm due to the resonant operation principle. The stopband rejection is mainly governed by electrostatic crosstalk for the presented unsymmetrically driven filters. It could however be increased by around 20dB using symmetric signals at input and output terminals.

The advantages of pure proximity coupling in TCFs comprise a precise determination of coupling between resonators by photolithography. Thereby a high sensitivity to variations in the inductivity of an external coil is avoided, which is needed in conventional electrical coupled TCFs with relatively high bandwidths. Additionally losses in the external coil are avoided.

The electrical crosstalk level is at least comparable to electrical coupled TCFs, while resonator sidelobes may lead to distortions in the passband in entirely proximity coupled filters as well as in electrically coupled filters with high relative bandwidths.

8. REFERENCES

[1] L. A. Coldren and R. L. Rosenberg, "Surface-Acoustic-Wave Resonator Filters", Proc. of the IEEE, pp. 147–158, 1979.

[2] R. V. Schmidt and L. A. Coldren, "Thin Film Acous-tic Surface Waveguides on Anisotropic Media", IEEE Transactions on Sonics and Ultrasonics, vol. SU-22, pp. 115–122, 1975.

[3] M. Dishal, "Alignment and Adjustment of Synchronously Tuned Multiple-Resonant-Circuit Filters", Proc. of the I. R. E., pp. 1448–1455, 1951.

[4] G. Scholl, W. Ruile and P. H. Russer, "P-Matrix Modelling of Transverse-Mode Coupled Resonator Filters", in IEEE Ultasonics Symp. Proc. , vol. 1, pp. 41–46, 1993.

[5] G. Martin and B. Wall, "A Four-Pole Saw Resonator Filter Combining Transverse And Symmetrical Longitudinal Modes", in IEEE Ultasonics Symp. Proc. , vol. 1, pp. 37–40, 1997.

[6] G. Martin, B. Wall, R. Kunze and M. Weihnacht, "Four Modes Waveguide Resonator Filters", in IEEE Ultasonics Symp. Proc. , vol. 1, pp. 35–39, 1993.

[7] C. S. Hartmann, B. P. Abbott, S. Jen and D. P. Chen, "Distortion of Transverse Mode Symmetry in SAW Transversely Coupled Resonators Due To Natural SPUDT Effects", in IEEE Ultasonics Symp. Proc. , vol. 1, pp. 71–74, 1994.

1999 Joint Meeting EFTF - IEEE IFCS

AN SAW FILTER EMPLOYING WEIGHTED REFLECTORS AND A MULTI-STRIP COUPLER

[1]Y. KANEDA, [2]M. TAJIMA, [2]T. OMORI, [2]K. HASHIMOTO,
[2]M. YAMAGUCHI, [3]J. TSUTSUMI, [3]O. IKATA and [3]Y. SATOH

[1]FUJITSU MEDIA DEVICES LIMITED, [2]Chiba University
[3]FUJITSU LABORATORIES LTD.

Abstract

This paper deals with a miniaturised surface acoustic wave (SAW) filter employing two IDTs, two weighted reflectors and a multi-strip coupler (MSC).

In the structure, the width-weighted reflectors are to be responsible for the frequency response of the filter. The reflectors are optimally designed to realise the specified frequency response by taking account of the SAW reflectivity of width-weighted electrodes.

The filter having 3dB bandwidth of 4.0MHz was designed at the centre frequency of 200MHz and fabricated on 128°YX-$LiNbO_3$. The total acoustic length of the filter was reduced to less than 5mm. The experimental result obtained was in good agreement with the theoretical prediction.

1 Introduction

It is of current interest to realise IF surface acoustic wave (SAW) filters with well-controlled frequency and linear-phase responses, which could be applied to such digital mobile communication systems as CDMA. In addition to their sophisticated device responses, further miniaturisation of the devices is required owing to the rapid reduction in total size of hand sets.

Although SAW filters of the transversal type may be applied to this purpose[1, 2], optimally weighted interdigital transducers (IDTs), which are basically responsible for realising precise and sophisticated responses, become very large in acoustic length. This would be one of the most significant obstacles to the device miniaturisation.

A resonant single-phase unidirectional transducer (R-SPUDT)[3] is also one of the promising candidates for the present purpose. Although the resonance enables to expand an effective acoustic length, it should be noted that this also causes distortions in phase responses simultaneously.

Bergmann, et al. proposed a reflector-based filter, in which the frequency response is substantially realised by appropriately weighted reflectors instead of IDTs[4]. Since SAWs make a round propagation in the reflector, the effective acoustic length of the reflector can roughly be regarded as two times the acoustic length of IDTs used in transversal filters. This would be most promising for miniaturising filters.

In [4], two parallel-connected filters with different acoustic path lengths were employed so that the reflected SAWs are selectively detected. However, slightly different IDT characteristics cause "acoustic" feed-through, which significantly deteriorates device responses. From this reason, it does not seem very easy to achieve good out-of-band rejection by this structure.

From this point of view, this paper discusses a filter structure as shown in Figure 1, where the two wide bandwidth IDTs are unweighted, while the two identical reflectors are electrode-width-weighted. The multi-strip coupler (MSC) is optimally designed so that the SAW energy impinged upon the upper track may be divided in half between the upper and lower tracks.

The basic behaviour of such a filter is as follows,

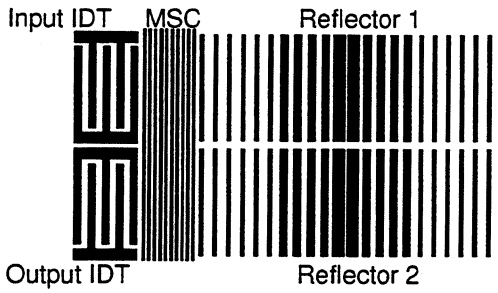
a) the SAW energy launched from the input IDT is divided equally between two tracks by the MSC, and impinges upon the two identical Reflectors 1 and 2,

b) the equal amount of SAW energy reflected by the reflectors is added up together through the MSC, and detected by the output IDT.

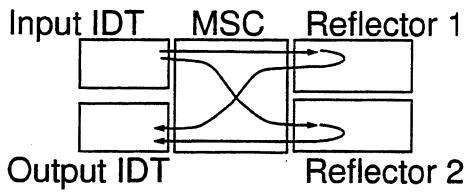
In this structure, therefore, the electrode-width-weighted reflectors are to be responsible for the frequency response of the filter.

The filter having 3dB bandwidth of 4.0MHz was designed at the centre frequency of 200MHz and fabricated on 128°YX-$LiNbO_3$. The total acoustic length of the filter was reduced to less than

0-7803-5400-1/99/ $ 10.00



(a) Configulation of the reflector based filter



(b) Diagram of SAW path

Figure 1: Basic filter structure.

5mm. The experimental result obtained was in good agreement with the theoretical prediction.

2 Electrode-width-weighting for SAW grating reflector

In the structure shown in Figure 1, two reflectors are substantially responsible for the frequency response of the filter, when both input and output IDTs are of wide bandwidth. This means that the reflectors should be properly weighted to realise required device responses.

According to the coupling of mode (COM) analysis, the behaviour of reflectors are expressed by,

$$\begin{cases} \dfrac{dU_+(X)}{dX} = -j\theta_u U_+(X) - j\kappa_{12} U_-(X) \\ \dfrac{dU_-(X)}{dX} = j\kappa_{12}^* U_+(X) + j\theta_u U_-(X) \end{cases}, \quad (1)$$

where $U_\pm$ denotes the amplitude associated with travelling SAWs in the $\pm X$ direction, respectively, and θ_u is a detuning factor. Reflectors could be weighted by modulating κ_{12} having appropriate weighting functions of X.

In the present paper, the electrode-width-weighting is employed[6] for practical devices where continuous weighting is required to realise precise frequency responses.

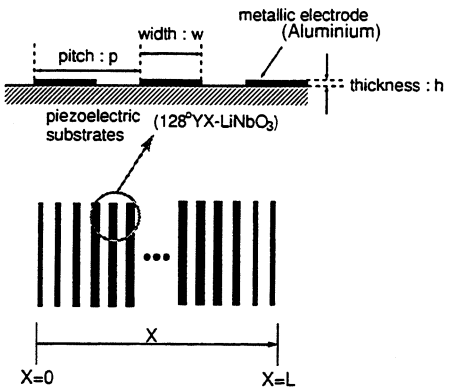


Figure 2: Geometrical parameters for the weighted reflector.

An open-circuited grating reflector having bandwidth of 4MHz is designed at the centre frequency of 200MHz. In the design, it is assumed that metallic aluminium (Al) grating electrodes are placed on 128°YX-$LiNbO_3$ substrate as shown in Figure 2, where w and p, and h are the width, pitch and thickness of the grating electrodes, respectively.

The reflector is weighted by modulating the reflectivity (per one electrode) $\kappa_{12}p$ for each grating electrode as,

$$\kappa_{12}p = \kappa_{12}p|_{w/p=0.6} \cdot W(X), \quad (2)$$

where $W(X)$ is a weighting function. Let $W(X)$ be

$$W(X) = A\left[a + b \cdot \cos\left(2\pi\frac{X}{L}\right) + c \cdot \cos\left(4\pi\frac{X}{L}\right)\right] \quad (3)$$

where the acoustic length L of the reflector is assumed to be $200p$, and A is a determined as,

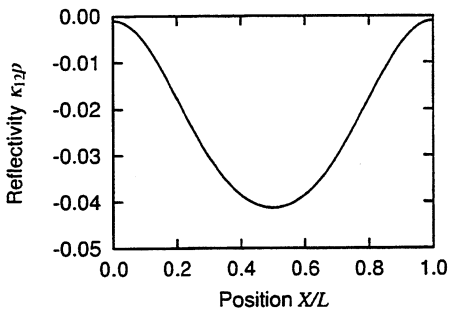
$$\max[W(X)] = 1. \quad (4)$$

The coefficients a, b and c are determined in the optimisation.

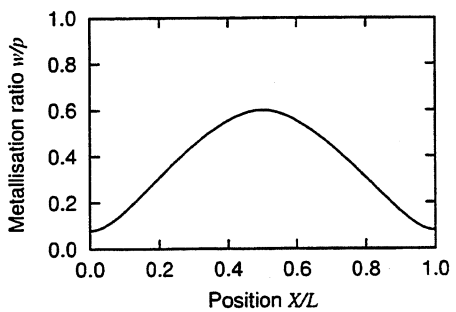
With an optimally determined $W(X)$, $\kappa_{12}p$ is shown as a function of X/L in Figure 3(a). When the electrode thickness h is constant, $\kappa_{12}p$ is basically determined only by the metallisation ratio (w/p) as shown in Figure 4 [7]. This means that as a function of X/L, the value of w, which gives $\kappa_{12}p$ shown in Figure 3(a), is easily determined as

shown in Figure 3(b). The reflectivity of the optimally designed reflector is compared with that of the unweighted reflector in Figure 5. It is clearly seen that the sidelobes can markedly be suppressed by properly weighting the reflector.

When applying the electrode-width-weighting to practical devices, one should take the account of the effect of the SAW velocity change under the grating electrodes with different w. Since this velocity fluctuation causes the deterioration of the reflection response, the electrode pitch p should be locally adjusted so that the delay within each p is constant for a whole reflector.



(a) Distribution of reflectivity in the open-circuited grating reflector.



(b) Distribution of electrode width.

Figure 3: Electrode-width-weighting.

3 Experimental result

A filter was designed at the centre frequency of 200MHz and its design parameters are shown in Table 1.

According to the result in 2, the reflectors were electrode-width weighted, while the split finger IDTs were unweighted. The total acoustic length of the filter was reduced to less than 5mm and it was mounted on a surface mount package with $7 \times 5\text{mm}^2$.

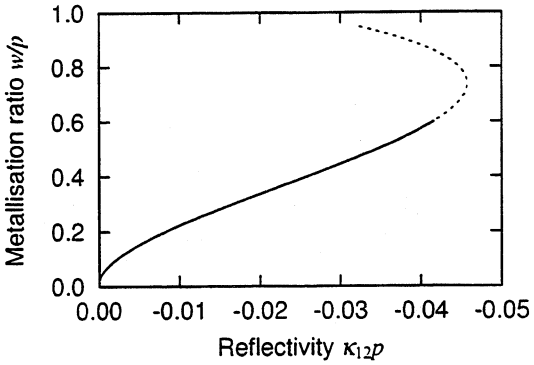


Figure 4: Metallisation ratio as a function of reflectivity. (Open circuited-grating : $h/p = 0.03$)

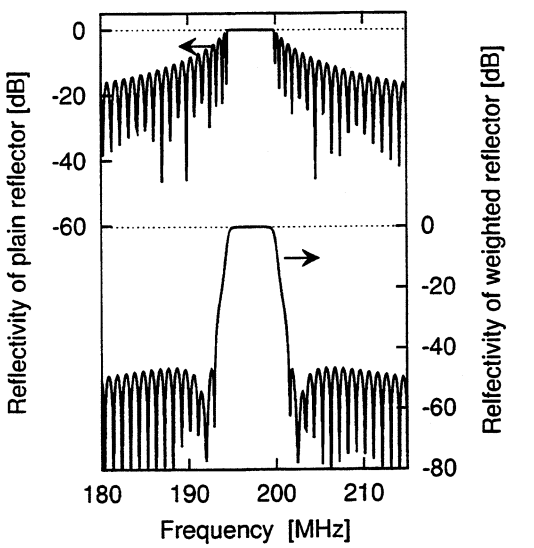


Figure 5: Reflectivity of the weigted reflector.

Figure 6 compares the frequency response of the fabricated filter with the result of theoretical calculation. Except the frequency region far from the centre frequency, the experimental result is, as a whole, in good agreement with the theoretical prediction. The disagreement in the out-of-rejection may be caused by the electrical feed-through.

The out-of-band rejection of the filter is nearly 40dB at $\pm$5MHz from the centre frequency, and the insertion loss of the pass band is around 13dB. Most of the insertion loss in the pass band may be caused by the bidirectional loss of IDTs and the source miss-matching loss.

Table 1: Design parameters of the filter.

centre frequency	200MHz
bandwidth	4MHz
packaging (SMD)	$7 \times 5\text{mm}^2$
substrate	128°YX-$LiNbO_3$
electrode	aluminium
thickness of electrodes (h/p)	0.03
Input & output IDTs	
period ($2p$)	19.68μm
number of finger pairs	10 pairs
aperture	40λ
Weighted reflector	
period (p)	10μm (*typ.*)
number of electrodes	200
metallisation ratio (w/p)	0.06 ~ 0.6

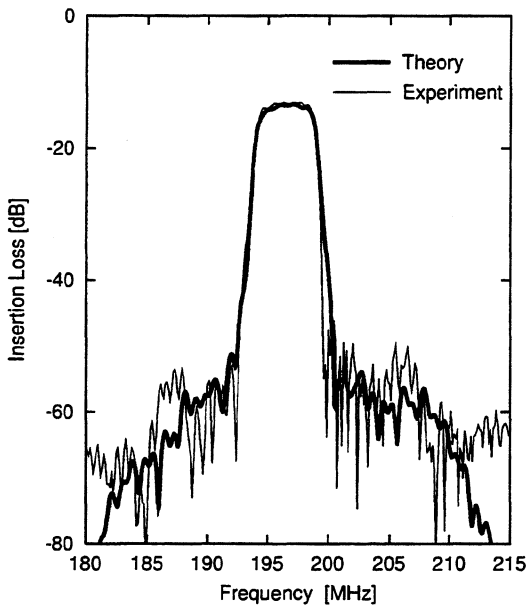


Figure 6: Frequency response of the filter.

4 Conclusion

This paper has discussed a miniaturised filter consisting of unweighted wide bandwidth input and output IDTs, two weighted reflectors, and an MSC. The reflectors were weighted by changing the electrode width in the direction of the acoustic length (electrode-width-weighting) which is one of the practical methods to realise precise responses. The filter having 3dB bandwidth of 4MHz at the centre frequency of 200MHz was designed and fabricated. The total acoustic length of the filter was reduced to less than 5mm, and then it was mounted on $7 \times 5\text{mm}^2$ SMD packages. The experimental results showed good agreement with the theoretical prediction.

References

[1] http://www.sawtek.com

[2] http://www.microsonics.thomson-csf.com

[3] P. Ventura, M. Solal, P. Dufilie, J. Desbois, M. Doisy and J.M.Hode:"Synthesis of SPUDT Filters with Simultaneous Reflection and Transduction Optimization", Proc. IEEE Ultrason. Symp. (1992) pp.71-75.

[4] A. Bergmann, J. Machui, K. Wagner and R. Weigel, "Two-track-reflector-filter for CDMA mobile telephones.", Proc. IEEE Ultrason. Symp. (1996) pp. 57-60.

[5] F. G. Marshall and E. G. S. Paige: "Novel Acoustic-Surface-Wave Directional Coupler with Diverse Applications.", Electron. Lett., **7** (1971) pp.460-464.

[6] G. Müller, G. Sindic, F. Seifert and R. Weigel: "Synthesis method for broadband dispersive 180° reflecting arrays in SAW devices.", Proc. IEEE Ultrason. Symp. (1993) pp. 233-236.

[7] K. Ibata, T. Omori, K. Hashimoto and M. Yamaguchi: "Polynomial expression for SAW reflection by aluminium gratings on 128°YX-$LiNbO_3$.", Proc. IEEE Ultrason. Symp. (1998) *to be published.*

1999 Joint Meeting EFTF - IEEE IFCS

A MIXED TRANSVERSE MODES AND ANGULAR SPECTRUM OF WAVES MODEL FOR THE ANALYSIS OF SAW TRANSVERSELY COUPLED RESONATORS FILTERS

Marc SOLAL

THOMSON MICROSONICS
SOPHIA ANTIPOLIS, FRANCE

ABSTRACT

Transversely-coupled SAW resonator filters commonly employ two or more sections in cascade to obtain the required number of poles in the response. As shown in [3], it is possible design 3 or 4 pole filters in a single section, resulting in potentially lower insertion loss, smaller chip size, and elimination of the need for an external-coupling component. A 71 MHz 3 pole filter and a 200 MHz 4 pole filter were designed using this approach. Unfortunately, unexpectedly large parasitic coupling was observed on the high-frequency side of the transfer function. This coupling could not be described using only the classical waveguide model.

If discrete transverse modes are identified, under the assumption that the energy distribution vanishes outside the gratings, a continuum of solutions exists in the case of propagating waves outside the grating (similar to angular spectrum of waves). By looking at the mode profiles, it was shown that a large part of the source excitation was not coupled to the transverse modes. To take this remaining energy into account, a decomposition of the continuum was performed. Our P-matrix based model for transversely coupled structures was improved to take into account this continuum.

A comparison of the measured and simulated frequency responses is shown for different filters and different metalization thicknesses. The excellent agreement demonstrates the need to take into account the coupling continuum in a TCF, even for a filter with only 2 cascaded sections.

1. INTRODUCTION

One major application of the technology is for IF mobile phone filters. Higher IF's make coupled-resonator filters well suited for the application. Transversely-coupled SAW resonators offer the dual advantages of low insertion loss and small size, both important in mobile phone applications. Classically, a cascade of two 2 pole filters is used to make a 4 pole filter. When used for designing GSM IF filters around 200 MHz, an extra inductor is generally necessary to couple the two stages.

Several design approaches have been proposed in the literature for implementing single-stage filters. One approach uses devices with more than two tracks to increase the number of useful transverse modes from two to three or four [3]. Another approach combines two longitudinal modes and two transverse modes to achieve a four-pole response. Up to now, published results do not show good rejection for either kind of filter.

We followed a similar approach, trying to design multi-track filters. A 71 MHz three-pole filter and a 200 MHz four-pole filter were designed and manufactured. Unfortunately, large parasitics on the high-frequency side of the transfer function were observed. These parasitics were not expected using our classical waveguide model. To describe these parasitics correctly, a new model had to be developed to take into account not only the guided modes (characterized by an exponential decay of energy outside the structure) but also propagating modes outside the structure (similar to angular spectrum of waves theory).

2. FILTER DESCRIPTION

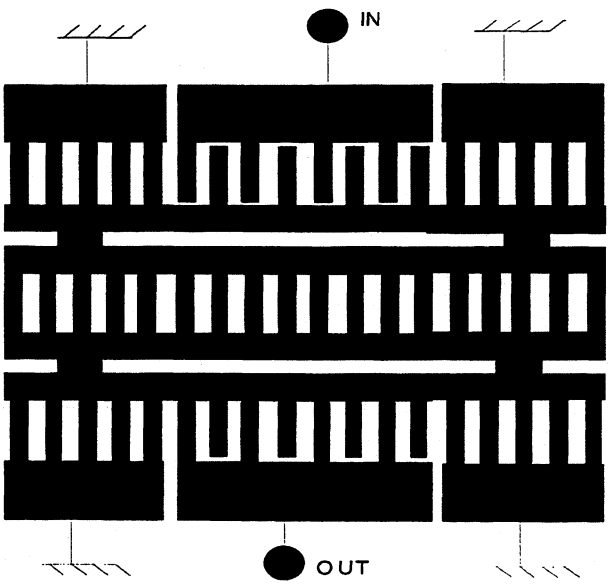


FIGURE 1: PRINCIPLE OF THE 71 MHz THREE-POLE FILTER

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

2.1. - 71 MHz Filter

The first filter designed using the multi-transverse modes technique was a 71 MHz three-pole filter. A schematic of the filter is shown in Fig. 1.

The filter comprises three coupled tracks. The center track is only a grounded grating. The upper and lower tracks are the input and output resonators. The filter was designed to fit in a 9×7 mm^2 package.

Fig. 2 shows a comparison between the simulated response and the measured response in 3 kΩ system. The simulated insertion loss is less than 5 dB and the theoretical rejection is better than 35 dB. The measured transfer function is consistent with the simulation except for a -30 dB shelf on the high-frequency side.

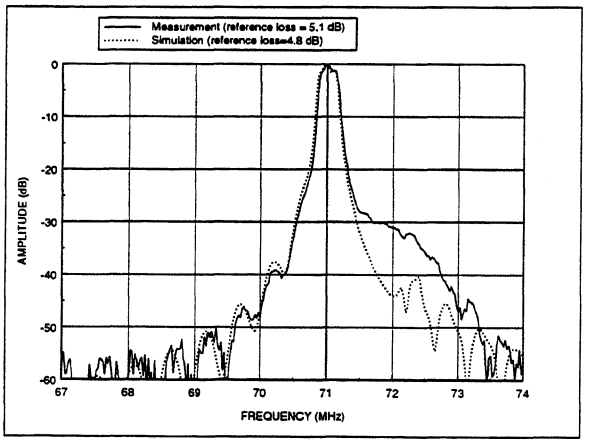


FIGURE 2: COMPARISON OF THE 71 MHz FILTER SIMULATED AND MEASURED TRANSFER FUNCTIONS (guided modes only)

2.2. - 200 MHz filter

A similar approach was used to design a 200 MHz filter. Instead of three tracks, the device had four tracks and employed four modes (two symmetric and two antisymmetric). Fig. 3 shows a comparison between the simulated and measured transfer functions.
The theoretical response has a rejection of 25 dB for low frequencies and 30 dB for high frequencies which was sufficient for system needs. Unfortunately, similar to the 71 MHz filter, it has a large spurious shelf on the high frequency side with a level of only –20 dB.

To try to understand the origin of this parasitic coupling, the same device was manufactured with double the metal thickness. Fig. 4 shows the simulated and the measured transfer functions. Due to the difference in velocities, the bandwidth is narrowed. The new bandwidth is in good agreement with the simulated one but the parasitic coupling is still present at a level of around -30 dB.

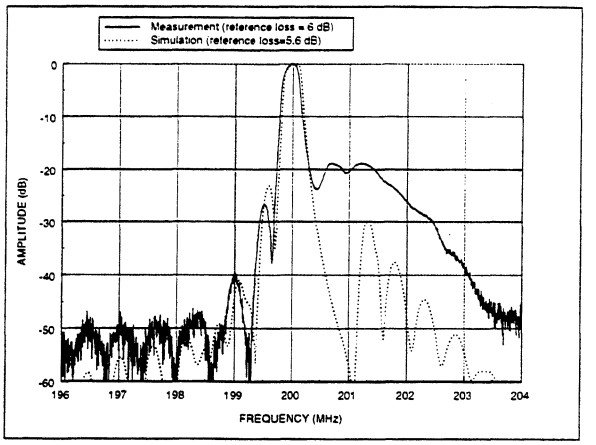


FIGURE 3: COMPARISON OF THE 200 MHz FILTER SIMULATED AND MEASURED TRANSFER FUNCTIONS (guided modes only)

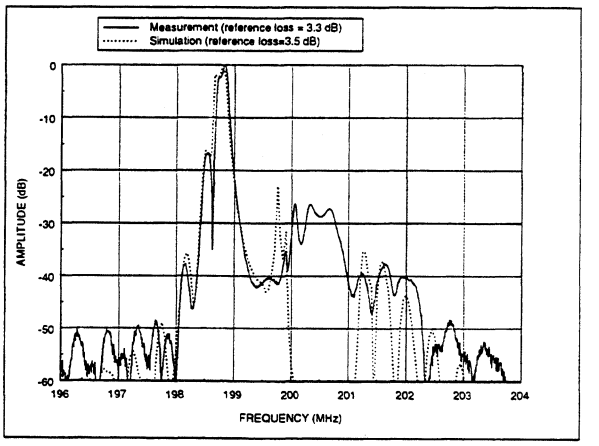


FIGURE 4: 200 MHz FILTER WITH DOUBLE METAL THICKNESS (guided modes only)

2.3. - Discussion of results

Parasitic coupling was observed experimentally on the high-frequency side of transfer function for both the 71 MHz and 200 MHz filters. This coupling was strongly dependent on the metalization thickness and thereby the waveguiding conditions. To investigate the origin of this coupling, admittance measurements were made showing good agreement between simulated and measured values for the modes region, but parasitics are also visible on the admittances. Fig. 5 shows the decomposition of the rectangular source into the guided modes for the 200 MHz filter. It can be seen that a large part of the energy is not coupled to these modes. It was hypothesized that this remaining energy goes through the grating from the input resonator to the output resonator, and is responsible for the parasitic coupling. A new

model was developed to take this energy coupling into account.

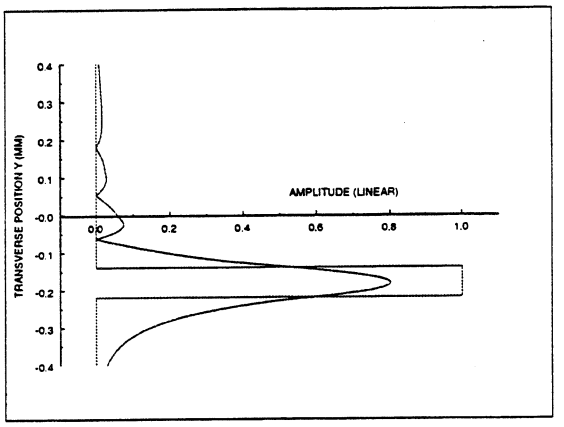


FIGURE 5: SOURCE EXCITATION AND DECOMPOSITION UNDER MODES SHAPES

3. THE MIXED WAVEGUIDING- ANGULAR SPECTRUM THEORY MODEL

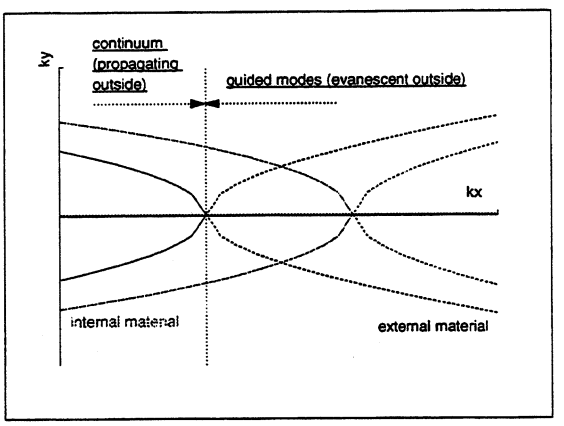


FIGURE 6: SLOWNESS CURVES FOR THE SIMPLE GUIDE

3.1. - Waveguided modes and continuum

For illustration, we examine here a very simple waveguide with two materials corresponding to two different wavenumbers k_{in} and k_{ext}. The external material is infinite. For simplicity, the two materials are considered isotropic, though the approach is also valid for anisotropic materials. For waveguiding, the velocity must be slower inside the structure than outside ($k_{in} > k_{ext}$).

3.1.1. - Waveguided modes dispersion equation

Classically, the dispersion equation is written assuming evanescent behavior outside the guide. The equations are very simple. For a given mode wave number k_x, the mode shape inside the grating is (for symmetrical modes):

$$g(y) = A\cos(k_{in,y}y) \quad \text{where: } k_{in,y} = \sqrt{k_{in}^2 - k_x^2}$$

and outside the grating :

$$k_{ext,y} = \sqrt{k_{ext}^2 - k_x^2} \quad \text{where: } k_{ext,y} = \sqrt{k_{ext}^2 - k_x^2}$$

We assume here that k_x is larger than k_{ext} so that $k_{ext,y}$ is pure imaginary. Only the determination of $k_{ext,y}$ corresponding to an evanescent wave has a physical meaning. The mode shape is then:

$$g(y) = B\exp(-jk_{ext,y}(y - W/2))$$

where the determination of the square root is such that : $-jk_{ext,y} = -b$, and b is real positive.

The dispersion equation is obtained from the continuity requirement of g(y) and dg(y)/dy at the interface:

$$\begin{bmatrix} \cos k_{in,y} W/2 & -1 \\ -k_{in,y} \sin k_{in,y} W/2 & b \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} = 0$$

For these two equations in two unknowns, a solution only exists when the matrix has a determinant equal to 0. This leads classically to a discrete number of modes determined by the k_x giving a null determinant.

3.1.2- Continuum

Fig. 6 shows the slowness curves for the two materials and their imaginary branches. The region of k_x examined to find the guided modes corresponds to imaginary $k_{ext,y}$.

It is also possible to examine the region where $k_{ext,y}$ is real, corresponding to propagating waves outside the grating. In this region, both solutions for $k_{ext,y}$ are valid. Thus, the wave in the external material is the combination of two waves, corresponding to the two possible determinations of $k_{ext,y}$:

$$g(y) = B\exp(-jk_{ext,y}(y - W/2)) + C\exp(jk_{ext,y}(y - W/2))$$

The mode shape inside the guiding remains unchanged. As before, we can write the continuity of g(y) and of dg(y)/dy at the interface, leading to the new equations:

$$\begin{bmatrix} \cos(k_{in,y} W/2) & -1 & -1 \\ -k_{in,y} \sin(k_{in,y} W/2) & jk_{ext,y} & -jk_{ext,y} \end{bmatrix} \begin{bmatrix} A \\ B \\ C \end{bmatrix} = 0$$

Here we have a system of two equations in three unknowns. This system can always be solved by choosing one unknown arbitrarily and solving for the two others. This significantly changes the situation. Previously, for any wavenumber k_x, there was a unique mode shape g(y) satisfying the equations. Now that the number of modes is no longer discrete and there thus exists a continuum of modes.

This result can be seen as a generalization of the angular spectrum of waves theory, classically used for diffraction analysis. The wave generated by a source is divided into two parts. The first part is converted into the waveguided modes and the second part follows classical diffraction laws. An interesting result is that the continuum corresponds to small k_x,i.e. high frequencies, which is consistent with measurement results.

3.2. - Device analysis

To compute the continuum mode shape, we extended the stack matrix theory [9] that we already use to compute the waveguided modes for the case of real $k_{ext,y}$. This was straightforward and allowed us to make the computation of the continuum for an arbitrary geometry taking into account a large number of different layers (typically 10 for usual devices).

The rectangular source excitation on the jth active aperture W_j is decomposed into the waveguided modes and the continuum:

$$rect(y,W_j) = \sum_i \alpha_{i,j} g_i(y) + \int_{k_{ext,y}} \alpha_j(k_{ext,y}) g(y,k_{ext,y}) dk_{ext,y}$$

$$\text{where } \alpha_{i,j} = \frac{\int_{W_j} g_i^*(y)dy}{\int |g_i(y)|^2 dy}$$

$$\text{and } \alpha_j(k_{ext,y}) = \frac{\int_{W_j} g^*(y,k_{ext,y})dy}{\int |g(y,k_{ext,y})|^2 dy}$$

To more easily manage the continuum, we chose to replace the integration by a discrete summation on a sampling of the continuum with a weighting of α.
To extend our P matrix model [7], the only modification necessary was to consider the continuum samples exactly in the same way as we consider the guided modes. The basic principle of the model is to divide a transducer in several unit cells (usually the unit cell is a single period), to compute the P matrix of a single unit cell and then to cascade the P matrices to obtain the overall P matrix of the transducer. The continuum is included in the model directly by increasing the number of acoustical ports.

4. RESULTS

Our model was used to analyze the 71 MHz and the 200 MHz filters. Fig. 7 shows to convergence of the model with the number of samples for the continuum. As seen on the figure, for small numbers of samples, the effect of each sample is to add a sharp resonance peak to the transfer function at a frequency corresponding to the velocity of the sample. For larger numbers of samples a smoother curve is obtained. Further increasing the number of samples has no significant effect on the transfer function.

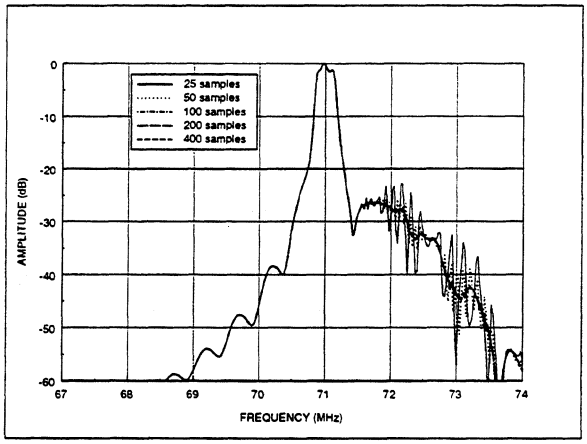


FIGURE 7: ANALYSIS OF THE 71 MHz FOR DIFFERENT SAMPLES NUMBERS

Figures 8-10 show comparisons of measurements and the results of our new model for the 71 MHz and the 200 MHz filters. Good agreement is demonstrated.
The new model was used to analyze the transfer function of a 175 MHz filter using a cascade of 2 three-track structures as described in [7]. It can be shown that same kind of parasitic coupling still exists in the transfer function and is well described by our model. Due to the cascade, the level of these parasitics is acceptable for the application. This explains why the results of waveguide-only model are sufficient to design cascaded filters.

5. CONCLUSION

Unexpectedly large high-side parasitic coupling was observed when trying to design transversely coupled resonators filters using only one stage. This parasitic coupling could not be modeled using a classical waveguide only model. We found that this parasitic coupling is due to the part of the excited wave which is not converted into the guided modes. This remaining energy goes from the input resonator to the output resonator by a diffraction-like phenomenon. This can be described by including a continuum of propagating waves outside the guides, similar to the angular spectrum theory of waves. Our P matrix based model was extended to take into account for this continuum. Comparison of experimental results and analysis results are in good agreement and demonstrate the validity of the new model.

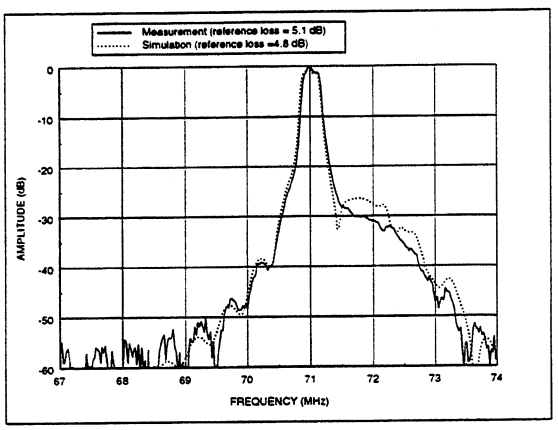


FIGURE 8: MEASUREMENT AND SIMULATION WITH CONTINUUM FOR THE 71 MHz FILTER

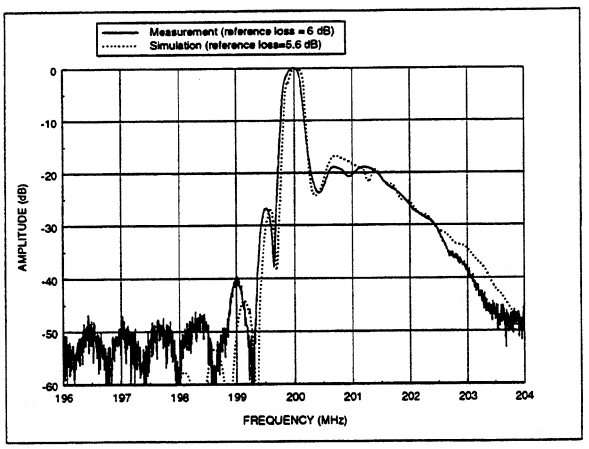


FIGURE 9: MEASUREMENT AND SIMULATION WITH CONTINUUM FOR THE 200 MHz FILTER

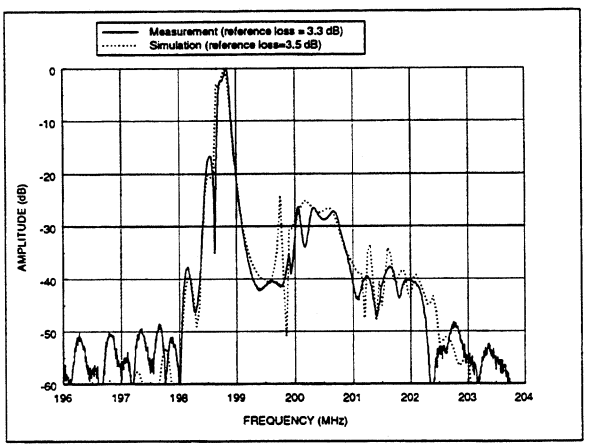


FIGURE 10: MEASUREMENT AND SIMULATION WITH CONTINUUM FOR THE 200 MHz FILTER WITH DOUBLE METALIZATION THICKNESS

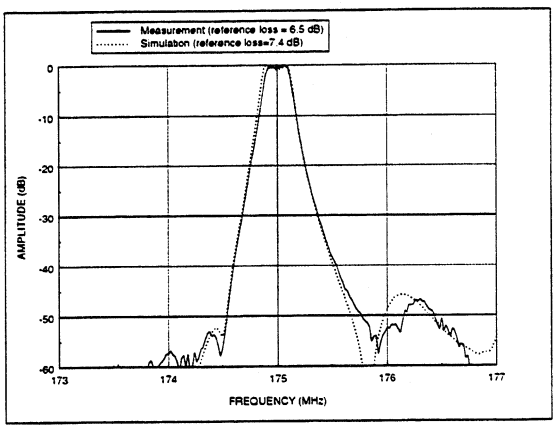


FIGURE 11: MEASUREMENT AND SIMULATION FOR A CASCADE OF TWO 2 POLE FILTERS

6. REFERENCES

[1] H. TIERSTEN and R. SMYTHE, Appl Phys Letters, 28, 111-113 (1976).

[2] M. TANAKA et al., 38th Frequency Control Symposium, 1984, p. 286-293.

[3] G MARTIN, B. WALL, R. KUNZE, M. WEIHNACHT "Four modes waveguide resonator filters", 1993 IEEE U.S Proc., pp. 35-39.

[4] Y. YAMAMOTO and R. KAJIHIRA, 1993 IEEE Ultrasonics Symp. Proc., pp. 47-51.

[5] G. MARTIN, B. WALL, "A SAW resonator filter using longitudinal and transverse modes", 1996 IEEE Ultrasonics Symposium Proc, pp. 53-55.

[6] HAUS, J. Appl. Physics, Vol 48, No 12. Dec 77, pp. 4955-4961.

[7] M. SOLAL, J. DESBOIS, "A new low impedance balanced drive structure for SAW transversally coupled resonator filter", 1997 IEEE Ultrasonics Symp. Proc., pp. 83- 88.

[8] D.P. MORGAN, S. RICHARDS, A STAPLES, "Development and analysis of analysis techniques for SAW transverse-coupled waveguide resonator filters", 1996 IEEE US. Proc., pp. 177-181.

[9] W.D. HUNT et al., 1992 IEEE US Proc., pp 45-50

[10] S. BIRYUKOV et al., 1995 IEEE UFFC Trans. vol 42, no. 4, July 95.

THE USE OF DESIGN OF EXPERIMENTS FOR THE OPTIMIZATION OF DEPOSITED GLASS ON SAW FILTERS

F. S. Hickernell[1] and H. D. Knuth
Motorola Systems Solutions Group
Scottsdale, Arizona 85257 USA

ABSTRACT

Vacuum deposited glass layers on piezoelectric crystals have been effectively used to improve the performance properties of surface acoustic wave (SAW) filters. These include, reducing the temperature coefficient of frequency, increasing the coupling factor, frequency trimming, and surface passivation for improved wafer yield. Using an orthogonal design of experiments (DOE) procedure, based upon a set of 9 experiments, the influence of process parameters on SAW properties were determined for RF diode sputtered glass on 128° Y-X lithium niobate. The glass was sputtered to a thickness around 600 nm for each experiment. There were four dependent sputter system variables, substrate temperature, background sputter gas pressure, RF power to the target, and gas flow rate. Three independent properties were measured, film stress, SAW velocity, and SAW propagation loss. The work is illustrative of a technique for optimizing the deposition conditions for films on SAW substrates using a minimum number of experiments. The DOE procedure can be applied to any combination of film and substrate with the results dependent upon the limits of the vacuum equipment and deposition technique used.

1. INTRODUCTION

Thin-film glass has been used to improve the temperature characteristics of high coupling factor SAW substrates because of its strong positive temperature coefficient of frequency (TCF), approximately 80ppm/°C, compared to the high negative coefficients of substrates such as lithium niobate ($LiNbO_3$) and lithium tantalate ($LiTaO_3$) [1]. The 128° Y rotated X propagating cut of $LiNbO_3$ was identified early [2] as an important substrate for development of SAW devices. The SAW properties of sputtered SiO_2 on 128° Y-X lithium niobate have been investigated [3-4] showing both TCF reduction and improved coupling efficiency. The resulting SAW properties are strongly dependent on the acoustic properties of the thin-film glass which in turn is dependent on the properties of the deposition system.

In this paper is described an orthogonal design of experiments (DOE) procedure, based upon a set of 9 experiments used to determine optimum processing parameters for RF diode sputtered glass on 128° Y-X lithium niobate. The glass was RF diode sputtered to a thickness around 600 nm for each experiment. There were four dependent sputter system variables, substrate temperature, background sputter gas pressure, RF power to the target, and gas flow rate. Three independent properties were measured, film stress, SAW velocity, and SAW propagation loss. Velocity and loss were measured with a series of SAW interdigital electrode patterns on the glass surface. The SAW velocities varied from 3880 to 3910 m/s for a film-thickness to acoustic-wavelength ratio of 0.1 and the variation in SAW propagation loss was between 18 dB/cm and 40 dB/cm in the 600 MHz region. A multilinear regression of the data, at a significance level where there is less than a 10% chance that a variable was left in the model when it has no effect, gave three polynomial expressions for stress, velocity, and propagation loss with their dependencies on the significant sputtering parameters. From these expressions best deposition conditions could be determined.

2. SUBSTRATE PROCESSING

The substrates used were standard three inch diameter 20 mil thick 128° Y rotated X propagating lithium niobate ($LiNbO_3$). The glass was RF diode sputter deposited on the $LiNbO_3$ substrates from a 6-inch pyrex glass target using a 90:10 argon:oxygen gas mixture. For the experiments the substrate temperature, background sputter gas pressure, RF power to the target, and gas flow rate were varied. The gas flow rate was adjusted by placing circular aperture constrictions of 4, 2, and 1-inch diameter in the throat of the vacuum system. A

[1] Fred S. Hickernell, 5012 E. Weldon, Phoenix,AZ 85018 USA, E-mail: f.hickernell@ieee.org

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

typical run time of 200 minutes gave film thickness in the range between 500 and 700 nm. Prior to placing the harmonic interdigital transducer (IDT) electrode array patterns on the surface of the glass, the stress induced from the glass deposition was measured.

To determine the best deposition conditions, a set of 9 experiments using an orthogonal DOE with the 4 sputter system parameters at three levels was used [5]. An orthogonal DOE has the property that each possible pair of levels for any two factors appears exactly the same number of times.

3. SAW MEASUREMENTS

The SAW measurement techniques used in this investigation required the photolithographic patterning of linear arrays of thin-film aluminum interdigital electrodes. The thin film aluminum was 100 nm thick. The electrodes were oriented for the excitation of surface acoustic waves along the (100) direction. The center-to-center spacing between transducers was 3.81 mm. Each transducer had 10.5 split-electrode finger pairs with a periodicity of 100 micrometers, and a 2.54 mm aperture. Individual electrodes within a transducer were 15 micrometers wide, and the spaces between adjacent electrodes were 10 micrometers. This transducer pattern facilitated the excitation of a wide range of harmonic waves and permitted the delineation of SAW velocity and loss characteristics over an extended frequency region. Frequency and loss measurements between transducer pairs were made using a network analyzer test set with time gating and signal averaging to enhance signal sensitivity at the higher frequencies. In general, measurements were made over the frequency range from 40 MHz to 1.4 GHz.

For each film sample the center frequencies of the fundamental and the higher harmonic responses were measured. The measurements were made on a HP 8753 network analyzer with a HP 85046 S-parameter test set. From the frequency information the SAW phase velocity could be calculated and from the insertion loss measurements the transducer conversion loss and propagation loss could be determined. SAW velocities were calculated from the frequency measurements using a small frequency scan in the peak region, and a velocity dispersion characteristic was developed as a function of glass film thickness-to-wavelength ratio. Insertion loss was measured between transducer pairs with different separations, and the SAW propagation loss was determined as a function of frequency from these measurements. Eight pairs of transducers with sequentially increasing separation were used. The linear change in loss with increasing transducer spacing is dominated by the propagation loss. A least-squares linear regression of the data points was used to establish a dB/cm value at each frequency

4. MEASUREMENT DATA

From the network analyzer scans the frequency and loss data were tabulated for each of the nine glassivated substrates. The SAW velocity value, calculated from the frequency using the wavelength periodicity fixed by the interdigital pattern, was then plotted against the ratio of film-thickness to acoustic-wavelength ratio (h/λ). The SAW propagation loss was determined at each frequency using the least-squares linear regression of the data points from the sequential increase in separation of the transducers. The conversion loss between adjacent transducers as a function of frequency was determined from the insertion loss intercept value from the SAW propagation loss regression at a single frequency. The stress value was a single measure for each substrate determined from the wafer bow before and after deposition.

To obtain a single matrix of values for analysis, a mid-range film-thickness to wavelength-ratio value was chosen to set the velocity value, and a mid-range frequency was chosen to set the representative propagation loss value. The matrix data is shown in Table 1. The temperature is in degrees Centigrade, the pressure is in milliTorr., the power is in Watts, and the aperture is diameter inches. The stress is in dynes/cm squared times 10 to the minus ninth power, the SAW velocity, taken at a film thickness to wavelength ratio of 0.1, is in meters per second, and the SAW propagation loss taken at 430 MHz is in decibels per centimeter.

5. ANALYSIS

The multilinear regression of the data for the sputtered glass (~550 nm) on 128 Y-X Lithium niobate to determine the best deposition conditions was based upon a set of 9 experiments using an orthogonal design procedure. The regression was done with a 10% significance level; i.e. there is less than 10% chance that a variable was left in the model when it really has no effect. There were four dependent sputter system variables: average substrate temperature,

background sputter gas pressure, RF power to the target, and aperture constriction for controlling the gas flow rate. There were three independent properties measured: film stress, surface acoustic wave (SAW) velocity, and SAW propagation loss. In order to carry out the multilinear regression it is advantageous to normalize the variables as coded variables: m=(u_max + u_min)/2, s=(u_max - u_min)/2 x=(u-m)/s. This results in the following dependent variable set as shown in Table II.

Table 1. Matrix of Values for Analysis

No	AVETEMP u1	PRESSURE u2	POWER u3	APERTURE u4	STRESS y1	VELOCITY y2	PLOSS y3
1.	162.5	7	200	2	1.863	3905	9.8
2.	169	8	250	4	2.900	3922	11.8
3.	157	4	150	1	1.474	3885	16.2
4.	262.5	7	250	1	2.738	3883	14.8
5.	254	8.5	150	2	1.305	3898	12.3
6.	258	4	200	4	2.007	3913	11.8
7.	69.5	7	150	4	2.210	3918	23.7
8.	79	7.5	200	1	1.321	3888	16.4
9.	88	4	250	2	1.557	3906	10.8

Table II Normalized Variables

No	(u1-m1)/s1 x1	(u2-m2)/s2 x2	(u3-m3)/s3 x3	(u4-4)/s4 x4
1	-0.0363	0.3333	0	-0.3333
2	0.0311	0.7778	1.0000	1.0000
3	-0.0933	-1.0000	-1.0000	-1.0000
4	1.0000	0.3333	1.0000	-1.0000
5	0.9119	1.0000	-1.0000	-0.3333
6	0.9534	-1.0000	0	1.0000
7	-1.0000	0.3333	-1.0000	1.0000
8	-0.9016	0.5556	0	-1.0000
9	-0.8083	-1.0000	1.0000	-0.3333

The stepwise regressions for a second degree polynomial in x1 - x4 gave the following:

1. y1 = 1.9371 + 0.3641 x3 + 0.2659 x4 + 0.5131 x1 x3
(Significance: 0.10)

2. y2 = 3909.3261 - 3.1463 x1 + 16.1558 x4 - 7.8315 x4 x4
(Significance: 0.06)

3. y3 = 14.0043 - 2.0165 x1 - 2.3595 x3 + 4.1477 x1 x3
(Significance: 0.10)

A summary of the numerical data indicates the following:

- The stress seems to depend mostly on power, aperture and somewhat on temperature as it interacts with power.
- The velocity depends mostly on aperture and somewhat on temperature.
- The propagation loss depends on temperature and power.
- Pressure does not seem to be important in this case to affect any of the factors.

Figures 1-3 show the stress, velocity, and loss as they relate to the dominant process factors. For low stress films a low flow rate (small aperture) and low power are the best conditions. The velocity characteristic can be tailored by the flow rate with a wide aperture producing a high velocity. The propagation loss is a more complex function with the low loss under conditions of low temperature and high power and high temperature and low power. The transducer conversion loss represented by the total insertion loss between adjacent transducer pairs was lowest under conditions of high power and lower temperature.

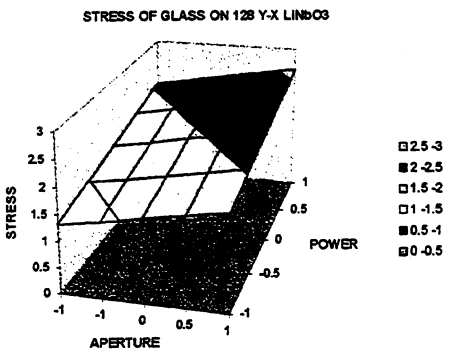


Fig. 1. Relationship of film stress to temperature and power for sputtered glass on 128° $LiNbO_3$.

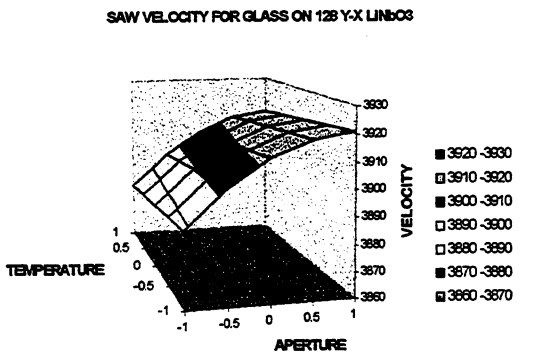


Fig. 2. Relationship of SAW velocity to temperature and aperture for sputtered glass on 128° Y-X $LiNbO_3$.

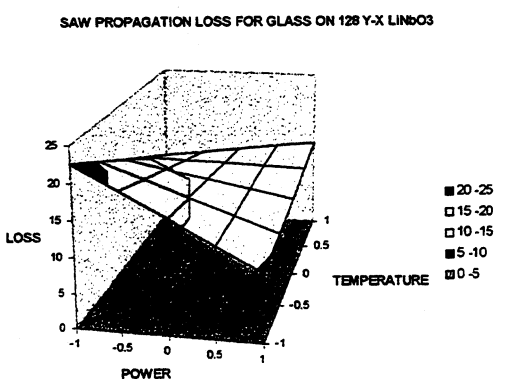


Fig. 3. Relationship of SAW propagation loss to temperature and power for sputtered glass on 128° Y-X $LiNbO_3$.

6. ACKNOWLEDGMENTS

The authors thank Fred J. Hickernell of Hong Kong Baptist University for suggesting the particular DOE used and for data analysis and Bob Dablemont of Motorola for measurement assistance.

7. CONCLUSIONS

Design of experiment techniques with multilinear regression analysis has been applied to determine the sputter system parameters affecting film stress and the SAW performance parameters for glass on 128° Y-X lithium niobate. The basic DOE technique is applicable to any film/substrate combination and can reveal interesting information which is important in the application of films to SAW device development.

7. REFERENCES

[1] T. E. Parker and H. Wichansky, "Temperature-compensated surface-acoustic-wave devices with SiO_2 film overlays," J. Appl. Physics, vol. 50, pp. 1360-1369, 1979.

[2] K. Shibayama, K. Yamanouchi, H. Sato, and T. Meguro, "Optimum cut for rotated Y-cut $LiNbO_3$ used as the substrate of acoustic-surface-wave filters," Proc. IEEE, vol. 64, pp. 595-597, 1976.

[3] K. Yamanouchi and S. Hayama, "SAW properties of SiO_2/128° Y-X $LiNbO_3$ structure fabricated by magnetron sputterng technique," IEEE Trans. Sonics and Ultrason., vol. SU-31, pp.51-57, 1984.

[4] K. Yamanouchi, H. Satoh. T. Meguro, and W. Wagatsuma, " High temperature stable GHz-range low-loss wide band transducers and filter using $SiO_2/LiNbO_3$, $LiTaO_3$," IEEE Trans. Ultrason., Ferroelectrics, and Freq. Ctrl., vol. 42, pp. 392-396, 1995.

[5] F. J. Hickernell, R. X. Yue, and F. S. Hickernell, "Statistical modeling for the optimal deposition of sputtered piezoelectric films," IEEE Trans. Ultrason., Ferroelectrics, and Freq. Ctrl., vol. 44, pp. 615- 622, 1997.

1999 Joint Meeting EFTF - IEEE IFCS

NONEQUIDISTANT UNIFORM-BEAM-PROFILE IDTS IN BROADBAND SAW FILTERS*

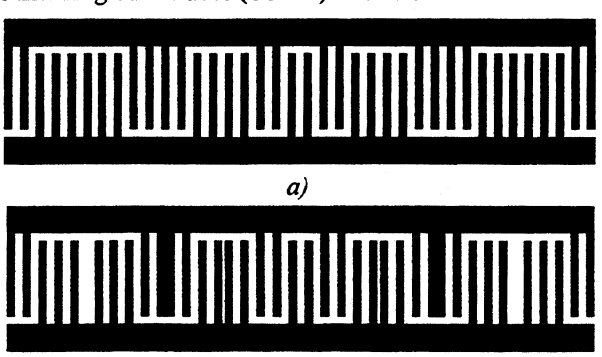
Evgeniy V. Bausk and Evgeniy A. Kolosovsky
Institute of Semiconductor Physics of Russian Academy of Sciences
630090 Novosibirsk, Russia

ABSTRACT

An optimization algorithm has been developed for the design of nonequidistant SAW transducers with uniform electrode length. Equidistant transducers weighted by assignment of polarities of individual electrodes are used as initial ones for this procedure. A set of SAW filters with bandwidth of 2-7%, shape factor of 1.2-1.5 has been designed using the suggested algorithm. The improvement of out-of-band rejection for filters with nonequidistant transducers is 3-6 dB in wide stopbands or 5-10 dB in most important narrow frequency regions as compared with equidistant analogs. In addition, the proposed algorithm can synthesize specified nonsymmetric frequency responses.

1. INTRODUCTION

The most compact surface acoustic wave (SAW) filter layouts need so small ratio between the aperture and the length of interdigital transducers (IDT) that any apodization (overlap weighting) is undesirable because of diffraction effects. Weighting by assignment of polarities of individual electrodes (polarity weighting) is a good technique for improvement of selectivity of IDTs with uniform electrodes [1,2], however, it is not so flexible as apodization: according the delta-function model, SAW sources amplitudes in polarity weighted (PW) transducers can have values of 1,0,-1 only. As a result, equidistant PW IDTs can provide frequency responses (FR) with good shape factor and close-in rejection by the price of a high level of far-out sidelobes that is the important disadvantage for filters with both PW IDTs on ST-quartz because of presence of surface skimming bulk waves (SSBW) in this cut.



a)

b)

Fig.1. An initial equidistant (a) and a final nonequidistant (b) PW IDTs

* This work was supported by SAWTEK Inc.

This paper is devoted to the next step of improvement of uniform-beam-profile IDT selectivity: optimization of electrode positions, that means optimization of discrete nonperiodic equal-amplitude SAW sources allocation to meet frequency response specification in the best way. Initial equidistant PW structures for this procedure (Fig.1a) are synthesized by algorithms described in [1].

2. OPTIMIZATION OF SAW SOURCES ALLOCATION IN IDTS

Within the scope of the delta-function model, frequency response of PW transducer can be written as:

$$H(f)=\sum_{n=2}^{N}(P_n - P_{n-1})/2\ \exp(i2\pi f x_n/v) \qquad (1)$$

where $P_n = \pm 1$ is the polarity of the n-th electrode; x_n is the coordinate of the n-th SAW source located in the middle of a gap between neighboring electrodes; v is the SAW velocity, N is the total number of electrodes in the IDT. We'll consider here only symmetrical transducers with N odd. Due to the fact that amplitudes of SAW sources located between electrodes of the same polarity are equal to 0 (i.e. any source isn't present here), the expression (1) for symmetrical PW IDTs can be simplified:

$$H(q)=2\sum_{m=1}^{M}(-1)^m \sin(q\,x_m), \qquad (2)$$

where $q=2\pi f/v$ is the normalized frequency, M is the number of SAW sources (with amplitudes not equal to 0) in a half of IDT. By specifying the desired values of the frequency response at M frequency points one can get the following set of equations:

$$H(q_j)=2\sum_{m=1}^{M}(-1)^m \sin(q_j\,x_m), \qquad (j=1,\ldots,M) \qquad (3)$$

In principle, the expression (3) allows to find a vector of coordinates $\{x_m\}$ for resolving the optimization task. Since the task is nonlinear and the amount of varying parameters $\{x_m\}$ is very large, one can not obtain an optimum set of $\{x_m\}$ by a simple and straight way. We suggest a method that transforms solving the nonlinear system of equations (3) to an iterative process of solving in turn a new set of linear equations. Let's consider it in more detail.

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

At small deviation of SAW sources coordinates $\{\delta x_m\}$ the response (3) has the following increment:

$$\delta H(q_j) = \sum_{m=1}^{M} (-1)^m q_j \cos(q_j x_m) \delta x_m . \quad (j=1,...,M) \qquad (4)$$

I.e. in the case of small difference between initial and desired frequency responses, one can, in principle, meet FR specifications at M frequency points. (M is the number of SAW sources in IDT) by solving the expression (4). However, the direct numerical solution is possible only if the number of equations (and the number of SAW sources) M is not larger than 25-30, otherwise the system becomes improperly stipulated as its matrix is expanded by new approximately equal terms. The problem of stability is known as a basic problem of optimization of systems with a large amount of parameters (up to a thousand). In order to solve it we use the known procedure of accumulation [3]: instead of the exact solving the exp.(4) for M frequency points we approximate the desired function $\delta H(q_j)$ by the expression

$$\sum_{m=1}^{M} (-1)^m q_j \cos(q_j x_m) \delta x_m \qquad (j=1,...,J)$$

in a greater number of points J corresponding to those parts of FR that exceed the demanded corridor of specifications. Let's consider the functional:

$$F(\delta x_1,..., \delta x_M) = = \sum_{j=1}^{J} \left| \sum_{m=1}^{M} (-1)^m q_j \cos(q_j x_m) \delta x_m - \delta H(q_j) \right|^2 , \qquad (5)$$

where $J>M$, usually J=1000-2000. According to the rule of the least-squares method:

$$\partial F(\delta x_1,..., \delta x_M)/\partial(\delta x_m)=0, \qquad (m=1, ..., M). \qquad (6)$$

As a consequence, the following system of $M \times M$ linear equations with respect to the vector $\{\delta x_m\}$ is defined:

$$\sum_{m=1}^{M} \left(\sum_{j=1}^{J} A_{jm} A_{jk} \right) \delta x_m = \sum_{j=1}^{J} (\varepsilon_j A_{jk}), \qquad (7a)$$

where $k=1,..,M$ and

$$A_{jm} = (-1)^m \{q_j \cos(q_j x_m) - q_{max} \cos(q_{max} x_m) H(q_j)/H(q_{max})\}/H(q_{max}) \qquad (7b)$$

Here $H(q_j)/H(q_{max})$ is the frequency response normalized by its maximum value, $\varepsilon_j = \delta H(q_j)/H(q_{max})$ is the negative residual of the normalized FR at a frequency of q_j. The vector $\{\delta x_m\}$ can be found by solving the set of linear equations (7a). Hence, in the case of small difference $\Delta H = H'(q) - H(q)$ between initial $H(q)$ and desired $H'(q)$ frequency responses, one can determine new SAW sources coordinates $\{x'_m\} = \{x_m\} + \{\delta x_m\}$ that provide the desired values of $H'(q)$ at J frequency points. When the difference is large, one can partition ΔH into many pieces δH and divide the way to target into small steps. At each step the results of the previous step ($\{x_m\}$ and $H(q)$) are used, i.e. an iterative process takes place.

The linear system of equations (7a) is symmetric and has one more important feature: a positive defined matrix. In our case it provides numerical stability of a solution [4]. Utilizing the matrix symmetry, we increase more than in 2 times the efficiency of algorithm at the stage of solution of (7a). The modified idea of Choletsky in factorization of matrixes is used, but without the square root operation [5].

In the case of a large number of SAW sources (more than 30-35) our experience shows that the system like (7a) is sensitive to a computer precision. This is the restriction of principle. That is why we have developed an original technique: the system of equations is divided into a set of subsystems with about 10-20 sources in each one. Each subsystem of equations is solved exactly. The splitting into subsystems can be made rather arbitrary. We have utilized the following variant: the s-th subsystem includes sources numbered: s, $S+s$, $2S+s$,..., $M-S+s$, where S is the number of sources in each subsystem. The total number of subsystems includes all M sources.

Thus, using the PW IDT frequency response analytical representation and the idea of the least-squares method, the iterative scheme to implement an FR optimization procedure has been constructed. The solution of non-linear system (3) with a large amount of equations has been reduced to the multi-repeated strict solutions of linear subsystems with moderate rank of matrix. As a result, the computation at PC/AT-P225 for the IDTs with 400 and 2500 electrodes (160 and 1080 SAW sources) takes about 10-15 minutes and 3-4 hours, respectively.

The iterative scheme can be easily generalized for the case of the total filter response optimization. The FR of other selective elements is regarded constant and is defined as $T(q_j)$. Expression (7b) should be substituted by:

$$A_{jm} = (-1)^m \{q_j \cos(q_j x_m) - q_{max} \cos(q_{max} x_m) T(q_{max}) \bullet \bullet H(q_j)/H(q_{max})\} \bullet T(q_j)/[H(q_{max})\, T(q_{max})]. \qquad (8)$$

The case of parallel optimization of both input and output PW IDTs can be reduced to (7a), (8) as well. Two linear systems for two transducers are resolved in turn, namely, all subsystems of the first transducer, then all subsystems of the second one, then again the first IDT and so on. The parallel optimization gives better results compared with the optimization in turn of input and output PW IDTs.

3. OPTIMIZATION OF IDT TOPOLOGY

The described optimization procedure is based on the simple delta-function model that is rather exact for initial equidistant structures where widths of all electrodes and gaps are equivalent. However, in final nonequidistant IDTs the amplitude of a SAW source placed in a narrow gap is not equal to the amplitude of a source in a much wider one. The account of this effect

would resulted in a radical complication of the optimization algorithm and in a great increase of the calculation time.

To provide an acceptable simulation accuracy while the simple delta-function model is used, we have developed a special procedure for the design of topology that corresponds to the optimized SAW source distribution in the best way. This procedure does not remember the electrode structure of the initial equidistant IDT and builds a new structure on the base of the final nonequidistant SAW sources distribution. A gap of a width $\lambda_0/8$ is formed on the position of each source. After that the adjacent electrodes of the same width are formed, then the next gaps and so on. As a result, the most electrodes and gaps in such nonequidistant IDT have the width $\lambda_0/8$ to suppress internal SAW reflections. Forming the IDT structure is directed inside regions between each two adjacent SAW sources. Only one electrode or gap in the middle of each such region has a non-standard width (Fig.1). The width of the middle electrode or gap takes a value between $\lambda_0/16$ and $3\lambda_0/16$. Since the most non-standard electrodes (gaps) are placed far from active gaps, they don't influence considerably on the SAW source amplitudes. The final nonequidistant structure usually has another number of electrodes compared with the initial equidistant transducer.

4. RESULTS

A number of nonequidistant polarity weighted IDTs and SAW filters were designed on the base of the described optimization algorithm.

4.1 Quasisymmetric frequency responses

At the beginning of the filter design procedure the software described in [1] optimizes structures of equidistant PW IDTs. Then the proposed here algorithm forms nonequidistant structures to provide quasisymmetric FRs with improved out-of-band rejection compared with equidistant analogous. If both input and output IDTs have PW structures, the algorithm optimizes both transducers simultaneously. If only one transducers is polarity weighted, it is optimized with taking into account the FR of the second IDT. In the last case the structure of the second IDT is not changed.

Fig.2 shows calculated frequency responses of the SAW filter with the bandwidth of 2.5%. The broken line is FR of the filter with two equidistant PW transducers (335 split electrodes in each IDT), their structures are initial ones for the optimization procedure. The final nonequidistant IDT structures give about 5 dB improvement for rejection in the wide stopband. Fig.3 demonstrates good coincidence between simulated and measured characteristics.

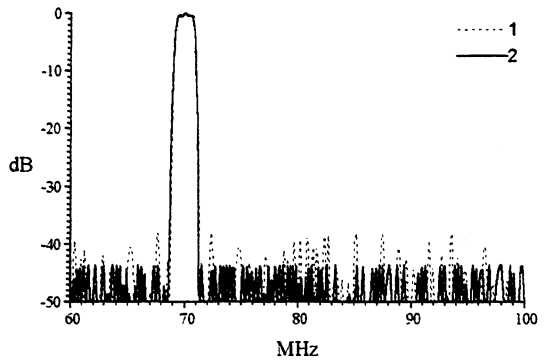


Fig.2. Calculated frequency responses of SAW filters with bandwidth of 2.5%:
1 - two equidistant IDTs; 2- two nonequidistant IDTs.

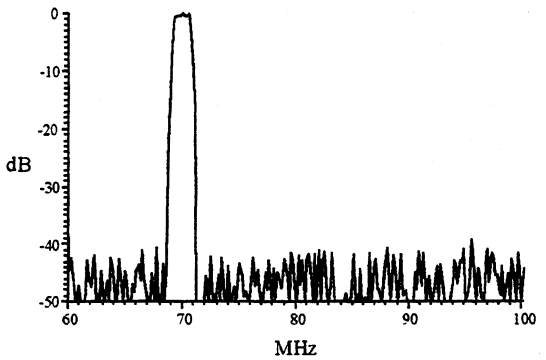


Fig.3. Measured FR of SAW filter with nonequidistant IDTs.

A more broadband filter with good shape factor of 1.2 is shown in Fig.4. Its response was designed to provide maximum close-in rejection at the expense of an increase of far-out sidelobes that can be suppressed by matching circuits. However, the measured response (Fig.5) doesn't have so good close-in rejection. The more accurate simulation (broken line in Fig.4) shows that the main reason of it is the presence of surface skimming bulk waves (SSBW) in ST-quartz.

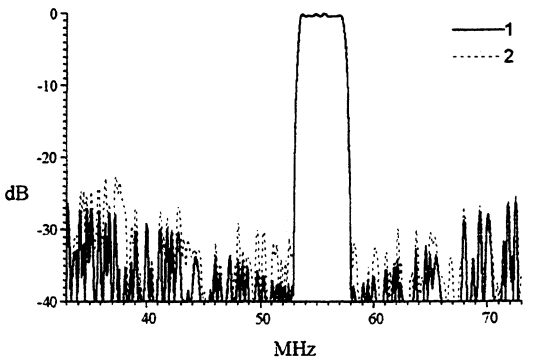


Fig.4. FRs of a SAW filter with bandwidth of 7%:
1 - simulated by the simple delta-function model,
2 - simulated with accounting the longitudinal SSBW.

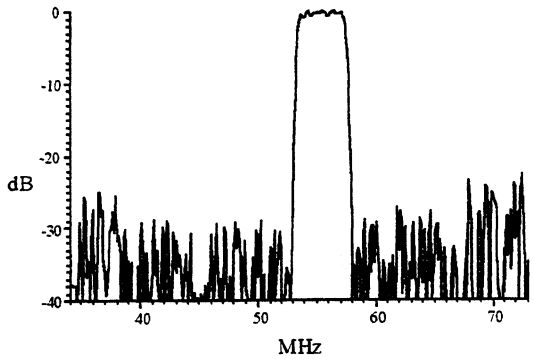


Fig.5. Measured FR of the 7% bandwidth filter with two nonequidistant PW IDTs

4.2 Nonsymmetric frequency responses

Nonequidistant IDTs can be used not only to improve selectivity of equidistant transducers but to form FRs with an asymmetry. Again, at the first stage an equidistant IDT with a symmetric FR is designed (Fig.6). Fig.7 shows good coincidence between simulated and measured characteristics.

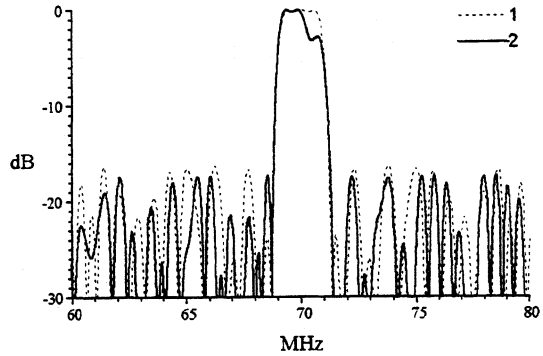


Fig.6. FRs of a single PW IDT with bandwidth of 2.5% simulated by the delta-function model: 1 - equidistant; 2 - nonequidistant.

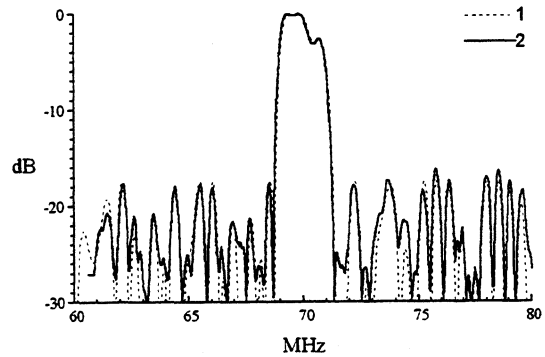


Fig.7. FRs of a single PW IDT with bandwidth of 2.5% 1 - simulated; 2 - measured.

All devices presented here were manufactured on ST-quartz.

Our software gives the possibility to compensate a difference between calculated and measured responses by making a predistortion for the synthesized one.

5.CONCLUSION

The original algorithm and the software were developed for optimization of discrete quasi-periodic equal-amplitude SAW sources allocation. It can be used for realization of nonequidistant uniform-beam-profile polarity weighted interdigital transducers with specified frequency responses. Although the algorithm is based on "quick" delta-function model, it includes possibility of compensation of some second-order effect by the special procedure of forming IDT topology and the mechanism of frequency response predistortion.

The subject of optimization can be either a single transducer or a total filter structure.

A number of IDTs and SAW filters with quasisymmetric and nonsymmetric frequency responses were designed using the proposed algorithm. Calculations show that the improvement of out-of-band rejection in filters with nonequidistant PW IDTs compared with equidistant analogs is about 3-5 dB in broad stopbands and 5-10 dB in the most important narrow regions.

ACKNOWLEGEMENTS

The authors thank Dr. Anatoliy Kozlov for his software providing IDT simulation by the more precise quasistatic charge distribution model. This software was used for testing effectiveness of the described technique.

REFERENCES

[1] E.V.Bausk and L.P.Solie, "Optimization of broadband SAW transducers with polarity weighting", in Proceedings of European Frequency and Time Forum, March 1998, pp.189-192.

[2] E.V.Bausk, "Floating electrodes: additional opportunities for optimization of broadband polarity weighted SAW transducers, in Proceedings of Frequency Contr. Symp., 1998, pp.473-476.

[3] D.D.McCracken and W.S.Dorn, Numerical methods and FORTRAN programming with application in engineering and science. John Wiley and Sons, Inc., New York, London, Sydney, Wiley International Edition, Second printing, 1965.

[4] J.H.Wilkinson, C.Reinsch, Handbook for automatic computation. Vol. IV, Linear Algebra. Heidelberg, NewYork, Springer-Verlag Berlin, 1971.

[5] S. Pissanetzky, Space matrix technology. London: Academic Press Inc., 1984.

1999 Joint Meeting EFTF - IEEE IFCS

LEAKY SAW BRANCHES COUPLED WITH OBLIQUE ACOUSTIC AXES IN TRIGONAL CRYSTALS

V. G. Mozhaev, F. Bosia,* and M. Weihnacht*

Faculty of Physics, Moscow State University, 117234 Moscow, Russia
*Institut für Festkörper- und Werkstoffforschung, Helmholtzstraße 20, D-01069 Dresden, Germany

ABSTRACT

The relationship between leaky surface acoustic wave (SAW) branches and acoustic axes for bulk waves is studied in the case of trigonal crystals. Three types of inclinations of acoustic axes are considered: 1) axes lying in the mirror planes of acoustic properties, oblique to the Z axis, 2) axes oblique to the mirror planes, 3) inclinated axes with respect to the boundary plane of the crystals. A new leaky SAW branch is found on the X cut of quartz. This branch is coupled with an acoustic axis lying in the boundary plane at an angle of about 66.3° from the Z axis. The other border of its existence region is determined by an acoustic axis oblique to the mirror planes. The theory by Khatkevich (1962) of acoustic axes of general position is revised and the results for trigonal crystals are corrected. It is shown that the relationship under study still holds when small deviations from the boundary plane are introduced in the acoustic axes. However, the permissible angles of deviation are smaller, as a rule, than the angles expected when the degeneracy solution is interpreted in terms of a bulk-wave-reflection problem.

1. INTRODUCTION

Leaky surface acoustic waves (SAWs) may be preferable to normal (non-attenuating) SAWs for device applications in specific cases when they have a higher electromechanical coupling coefficient or better thermal stability. The higher phase velocity of leaky SAWs is an additional advantage for high-frequency applications. However, the computational search for leaky SAWs in crystals involves the difficulties caused by restrictions on their existence region and a wide range of probable complex values of the wave number. The direct relationship between leaky SAW branches and acoustic axes (directions along which phase velocities coincide) for bulk acoustic waves in crystals has been recently indentified and explained by us [1,2]. In particular, new leaky SAW branches coupled with the acoustic axis coincident with the Z-axis of trigonal crystals have been found in berlinite and langasite. The objective of the present paper is to study the possibilty of using the oblique acoustic axes as reference points to search for new leaky SAW branches in trigonal crystals. We restrict our study to the case of trigonal crystals since they are the main substrate materials for SAW devices. The term "oblique axes" is used here in three different senses. Firstly, we consider leaky SAW branches coupled with acoustic axes oblique to the Z-axis when these axes lie in a symmetry plane of elastic properties of the crystal. Leaky SAWs on the YZ plane of quartz coupled with such an axis are found at the azimuthal angle about 66.3° from the Z-axis. Secondly, we study acoustic axes of general position which are oblique to symmetry planes. The previous statement on the absence of such axes in trigonal crystals [3], as we show, is incorrect. Correct equations for angles determining directions of acoustic axes oblique to symmetry planes are derived. Such oblique acoustic axes and associated leaky SAW branches are present in berlinite and quartz. Thirdly, the possibility of a relationship between leaky SAW branches and acoustic axes deviating from the boundary plane is examined.

2. LEAKY WAVES ON X-CUT QUARTZ

Quartz is a standard piezoelectric material for surface acoustic waves. It has been widely used both in ultrasonic devices and in physical investigations for many years. The X cut of quartz is one of the most

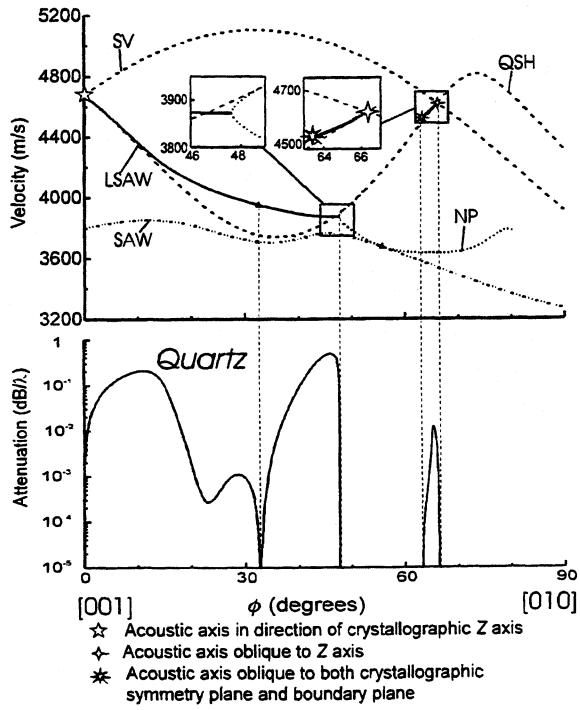


Fig. 1. Phase velocity and attenuation of leaky SAWs (LSAW) on X-cut quartz as a function of azimuthal angle ϕ. Velocity curves for bulk acoustic waves (SV and QSH), surface acoustic waves (SAW), and nonphysical solutions (NP) are also given. The inset on the right-hand side of the figure shows an additional leaky SAW branch on an enlarged scale.

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

popular cuts. The boundary plane for this cut is coincident with the mirror plane for elastic and acoustic properties of the crystal. Leaky SAWs on this cut exist in the sector of azimuthal angles adjacent to the Z crystallographic axis [4]. As shown in our previous papers [1,2], the leaky SAW solution for the propagation direction along the Z axis on this cut degenerates into a peculiar solution of a bulk-wave-reflection problem. Such a reflection problem was considered for the first time in Ref. [5]. The incident wave in this case impinges on the boudary plane at glancing angle and is called incident only due to the anisotropic deviation of its power flow from the wave vector to the boundary. This well-known leaky SAW branch exists when the propagation direction deviates from the Z axis no more than approximately 47.65°. Besides, there is an additional acoustic axis in the YZ plane of quartz. This axis is inclined with respect to the Z axis by 66.3°. The relationship under study between leaky SAW branches and acoustic axes and also corresponding results for lithium tantalate [6] allow us to assume that this additional acoustic axis can give rise to an additional leaky SAW branch on the X cut of quartz. This assumption is confirmed by our calculations (Fig. 1). Material constants for calculations are taken from Ref. [7]. The sector of existence of this new leaky branch is rather small (about 2°). The other end point of this branch also corresponds to an acoustic axis, this time slightly tilted to the bulk. In the next section, we consider in detail the theory of such acoustic axes of general position with arbitrary oblique orientations with respect to the symmetry planes of crystals.

3. ACOUSTIC AXES OUT OF THE SYMMETRY PLANES

The general theory of acoustic axes was developed by Khatkevich [3]. Khatkevich noticed that the orientations of acoustic axes lying out of the symmetry planes of crystals may be found from the condition that all cofactors, A_{ik}, of matrix $a_{ik} \equiv (\Gamma_{ik} - \rho v^2 \delta_{ik})$ be equal to zero

$$A_{ik} = 0, \quad i,k = 1, 2, 3. \tag{1}$$

Here Γ_{ik} is the Green-Christoffel tensor, ρ is the mass density of the crystal, δ_{ik} is Kronecker delta, v is the phase velocity of bulk waves. This condition is much simpler than the one obtained from the analysis of the secular equation. It is given in the paper by Khatkevich [3] without derivation, with reference to the book by Fedorov [8]. However, the formulae of Ref. [8] can give the condition (1) only if one assumes that the solution for the displacement vector is not unique in this case. Although such an assumption is not erroneous, it is not explicitly proved in the mentioned references and so rigorous proof of Eqs. (1) is absent there. Arguments in support of the validity of this condition presented in the other book by Fedorov [9] are rather tangled. Besides, only symmetrical three-dimensional matrices are discussed by Fedorov [8,9]. We present below a simple and general proof of Eqs. (1). This proof is valid for matrices of arbitrary dimension and their symmetry is not essential in this case. The problem of calculating the phase velocities of bulk acoustic waves in crystals is equivalent to the following eigenvalue problem

$$E_i \equiv (\Gamma_{ik} - \lambda \delta_{ik})\, u_k = 0, \tag{2}$$

where $\lambda = \rho v^2$ is an eigenvalue of Γ_{ik}, u_i is the particle displacement (eigenvector of Γ_{ik}). A nontrivial solution of Eqs. (2) is obtained if

$$D(\lambda) \equiv \det(\Gamma_{ik} - \lambda \delta_{ik}) = 0. \tag{3}$$

When Eq. (3) does not have coincident roots, these roots correspond to the points of sign change in the curves $D(\lambda)$, $E_i(\lambda)$. It is evident from geometrical analysis that there is no change of sign in these curves at the points of coincidence of two roots. Such contact points correspond, obviously, to local extrema of the functions which implies

$$\partial D/\partial \lambda = \partial E_i/\partial \lambda = 0. \tag{4}$$

This synchronous behaviour of several functions at the degeneracy points is in fact a true reason of great simplification of degeneracy conditions. Substituting E_i into Eqs. (4) produces

$$(\Gamma_{ik} - \lambda \delta_{ik})\, u'_k = u_i. \tag{5}$$

Eqs. (5) can be transformed into the form

$$D\, u'_k = A_{jk} u_j. \tag{6}$$

The well-known Cramer solution is obtained from Eqs. (6) dividing them by D under the condition that $D \neq 0$. However, in our case $D = 0$ and therefore

$$A_{jk} u_j = 0. \tag{7}$$

Eqs. (2) and (7) should be fulfilled simultaneously. Decomposing the determinant of Eqs. (2) in arbitrary i-th line gives

$$D = \sum_j a_{ij} A_{jk} = 0. \tag{8}$$

There is no summation over i in Eqs. (8) and (9). On the other hand, if we replace one of Eqs. (2) with fixed subscript i by one of Eqs. (7) with the same subscript ($k = i$), then the condition of existence of nontrivial solution of this combined system of equations takes the form

$$D = \sum_j A_{jk}^2 = 0. \tag{9}$$

From this it follows that Eqs. (1) should be fulfilled if the matrices under consideration are real. Only three

among Eqs. (1) may be considered as independent due to additional relations imposed by Eq. (3). As such independent equations, it is convenient to consider the cofactors of elements a_{23}, a_{13}, a_{12}

$$\begin{bmatrix} \Gamma_1-\lambda & \Gamma_6 \\ \Gamma_5 & \Gamma_4 \end{bmatrix} = \begin{bmatrix} \Gamma_6 & \Gamma_2-\lambda \\ \Gamma_5 & \Gamma_4 \end{bmatrix} = \begin{bmatrix} \Gamma_6 & \Gamma_4 \\ \Gamma_5 & \Gamma_3-\lambda \end{bmatrix} = 0$$

Excluding λ from these equations, we obtain two conditions derived by Khatkevich [3]

$$(\Gamma_1-\Gamma_2)\Gamma_4\Gamma_5+\Gamma_6(\Gamma_4^2-\Gamma_5^2)=0. \quad (10)$$

$$(\Gamma_1-\Gamma_3)\Gamma_4\Gamma_6+\Gamma_5(\Gamma_4^2-\Gamma_6^2)=0. \quad (11)$$

Analysing these equations, Khatkevich concluded that all the acoustic axes in trigonal crystals lie only in a symmetry plane, that their number is 4 or 10, and that previous numerical data for quartz available in the literature [10] are incorrectly calculated. This discrepancy and also the more recent numerical finding of out-of-plane acoustic axes in $LiTaO_3$ [11] and quartz [12] has stimulated us to recalculate Khatkevich's results on the basis of the general equations given in his paper. It has been found that the analytical results obtained by Khatkevich for trigonal crystals and therefore his conclusion mentioned above are incorrect. The correct expressions have the form

$$\tan^2\theta=\frac{(c_{33}+c_{44}+2c_{13})(c_{11}+c_{12})-2(c_{13}+c_{44})^2}{(c_{11}-c_{44})(c_{11}+c_{12})-2c_{14}^2}, \quad (12)$$

$$\sin 3\varphi=\frac{2(c_{33}+c_{44}+2c_{13})-(2c_{11}+c_{13}-c_{44})\tan^2\theta}{c_{14}\tan^3\theta}, \quad (13)$$

where θ and φ are the polar and azimuthal angles, respectively, and c_{ij} are elastic constants.

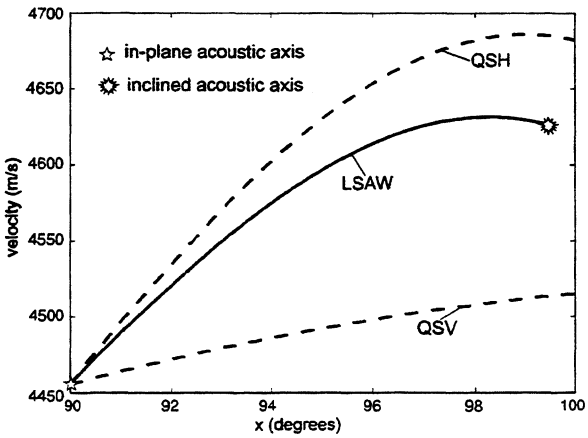


Fig. 2. Leaky SAW branch derived from nonsymmetric acoustic axis in quartz.

Eqs. (12) and (13) give real directions oblique to the symmetry planes for some trigonal crystals like quartz and berlinite: φ = 19.13°, θ = 63.04° for quartz and φ =25.94°, θ = 65.46° for berlinite. These values are confirmed by numerical solutions with account of piezoelectricity which give: φ = 19.64°, θ = 63.41° for quartz and φ = 25.37°, θ = 65.48° for berlinite. It is of interest to note that according to Eqs. (10) and (11) oblique acoustic axes should also exist in lithium niobate when φ = -21.40°, θ = 55.60°. However, numerical calculations of phase velocities of bulk waves in this crystal taking into account piezoelectricity reveal a small but finite spacing between two sheets of the slowness surface rather than their contact in the corresponding cone of directions. Thus piezoelectricity has a pronounced effect on the existence of acoustic axes in this special case.

The oblique acoustic axis in quartz has been used to find an associated leaky SAW branch (Fig. 2). The orientation of cut and the propagation direction are defined by the Euler angles (109.6361°, -26.5870°, x). The acoustic axis is in the boundary plane at x = 90°. At the termination of the leaky SAW branch, x = 99.49°, there is a degeneracy of leaky mode associated with an acoustic axis slightly tilted to the bulk.

4. ACOUSTIC AXES OUT OF THE BOUNDARY PLANE

Two configurations suitable to search for leaky SAWs with the aid of acoustic axes have been found in Ref. [1]. For one of them an acoustic axis should be in the plane of the surface. For the other an acoustic axis lying in the sagittal plane is directed into the bulk and one of the bulk waves propagating along it should simultaneously be a limiting wave (erroneously called exceptional wave in Ref. [1]). The interpretation of the degeneracy solution as a solution of a bulk-wave-reflection problem allows us to propose that all the orientations of acoustic axes confined between the two previously specified directions are permissible. To test this hypothesis, we have investigated some cuts of different crystals like berlinite, langasite, and potassium niobate. In all the cases studied, it is possible to tilt acoustic axes into the bulk and still maintain their coupling with the leaky SAW branch. The maximum angle of deviation corresponds to the case when one of

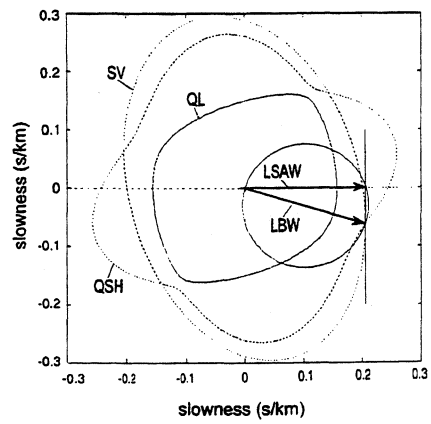


Fig. 3. Observed case of maximum deviation angle of acoustic axis from the boundary for constructing a leaky wave. Quartz, Euler angles (61.56°, 72.83°,-89.99999°). (LSAW = leaky surface acoustic wave, LBW = limiting bulk wave, SV = shear vertical wave, QSH = quasi shear horizontal wave, QL = quasi longitudinal wave)

the bulk acoustic waves propagating along the acoustic axis is simultaneously limiting bulk wave. This situation has been realized, in fact, only in special cases like rotated *X*-cut langasite for acoustic axis oblique to the *Z*-axis [13] and rotated *Y*-cut quartz. The best coincidence between results and predictions is observed in the case of quartz with the Euler angles (61.56°, 72.83°, -89.99999°), see Fig. 3. For comparison, the expected angle derived from the slowness curves is 72.82°. The circle in the figure represents the surface projections of slowness along the direction of acoustic axis for various angles of tilt of acoustic axis into the bulk (the sagittal plane is held constant in this case). The slowness-surface cross-section under consideration coincides practically with the *YZ* plane. Nevertheless, a small deviation from this plane is required to transform the degeneracy solution into leaky SAWs. Due to this small deviation, the direction mentioned as acoustic axis in Fig. 3 is not an acoustic axis in the strict sense, but it is very close to such an axis. The exact positioning of acoustic axis in the sagittal plane in this case results in the degeneracy of leaky SAWs into the so-called "no-motion" solution identified for the first time by Taylor [14].

Thus, our expectations concerning the limiting angle of tilt of acoustic axes coupled with leaky SAW branches are confirmed in the case shown in Fig. 3. However, the permissible angle of deviation is substantially smaller than expected one in many other studied cases. This is because of other types of leaky wave degeneracy being essential during changing the Euler angles after starting from a LSAW solution at a certain crystal orientation.

ACKNOWLEDGEMENTS

This research is supported in part by INTAS (project 96-441) and by Deutsche Forschungsgemeinschaft (project We 2065/1).

REFERENCES

[1] V.G. Mozhaev and M. Weihnacht, "Search for leaky SAWs in crystals with the aid of acoustic axes for bulk waves," in Proceedings of the 1997 IEEE Ultrasonics Symposium, 1997, pp. 267-273.

[2] M. Weihnacht, F. Bosia, and V. Mozhaev, "Acoustic axes and leaky surface acoustic waves," in Proceedings of the 24. Jahrestagung der Deutschen Akustischen Gesellschaft (DAGA 98), 1998, pp. 612-613.

[3] A. G. Khatkevich, "The acoustic axes in crystals," Kristallografiya, vol. 7, pp. 742-747, 1962 [Sov. Phys. Crystallography, vol. 7, pp. 601-604, 1963].

[4] G.W. Farnell, "Properties of elastic surface waves," in Physical Acoustics, vol. VI, edited by W. P. Mason and R. N. Thurston (Academic, New York, 1970), pp. 109-166.

[5] E.N. Koshkina, V.E. Lyamov, and T.A. Mamatova, "Reflection of acoustic waves in the case of conical refraction," Kristallografiya, vol. 23, pp. 1274-1277, 1978 [Sov. Phys. Crystallography, vol. 23, pp. 721-723, 1978].

[6] K. Hashimoto and M. Yamaguchi, "Non-leaky, piezoelectric, quasi-shear-horizontal type SAW on *X*-cut $LiTaO_3$," in Proceedings of the 1988 IEEE Ultrasonics Symposium, 1988, pp. 97-101.

[7] R. Bechmann, "Elastic and piezoelectric constants of alpha-quartz," Phys. Rev., vol. 110, pp. 1060-1061, 1958.

[8] F.I. Fedorov, Optics of Anisotropic Media, Izd-vo AN BSSR, Minsk, 1958, p.140 (in Russian).

[9] F.I. Fedorov, Theory of Elastic Waves in Crystals, Plenum: New York, 1968.

[10] W.G. Cady, Piezoelectricity, McGraw-Hill Book Company, New York, 1946.

[11] V.S. Bondarenko, V.I. Klimenko, N.V. Perelomova, and A.A. Blistanov, "Anisotropy of elastic wave propagation in lithium tantalate crystals," in Fizika dielektrikov i poluprovodnikov}, Mezhvuzovskii tematicheskii sbornik (Volgograd inzhenerno-stroitel'nyi institut, Volgograd, 1981) pp. 28-39 (in Russian).

[12] N.F. Naumenko, "Application of exceptional wave theory to materials used in surface acoustic wave devices," J. Appl. Phys., vol. 79, pp. 8936-8943, 1996.

[13] V.G. Mozhaev, F. Bosia, and M. Weihnacht, "Types of leaky SAW degeneracy," in Proceedings of the 1998 IEEE Ultrasonics Symposium, 1998, pp. 143-148.

[14] D.B. Taylor, "Surface waves in anisotropic media: the secular equation and ist numerical solution", Proc. R. Soc. Lond. A, vol. 376, pp. 265-300, 1981.

ON THE EXCITATION OF SURFACE AND PSEUDO-SURFACE QUASI-BULK WAVES

A.N.Darinskii[a] and N.F.Naumenko[b]
[a] Institute of Crystallography, Acad. Sci. Russia, Leninskii pr. 59, Moscow, 117333 Russia
[b] Moscow Steel and Alloys Institute, Leninskii pr. 4, Moscow, 117936 Russia

Abstract An investigation is made of the specific features of the excitation of (pseudo) quasi bulk surface waves. A particular attention is focused on the analysis of wave fields associated with "non-physical" plane wave solutions. As illustrations, numerical computations have been performed for SH-waves and quasilongitudinal leaky waves in quartz.

1. INTRODUCTION

The SAW devices designed on the basis of deeply penetrating quasi-bulk waves, such as Gulyaev-Bleustein waves and high-velocity low-attenuated leaky waves, are of considerable engineering interest. They can operate at high frequencies, exhibiting strong piezoelectric coupling and low temperature coefficient of frequency.

The phase velocity of a quasi-bulk wave is close to the limiting, or cutoff, velocity of bulk waves. This markedly complicates the analysis of the excitation of quasi-bulk waves since an IDT is then capable of efficiently generating both inhomogeneous and homogeneous modes. The excitation and detection problem for quasi-bulk waves was investigated in a number of works. In particular, much attention has been paid to the analysis of the Green function describing the displacement fields on the surface of the substrate and of the effective permittivity function corresponding to the Gulyaev-Bleustein and high-velocity leaky waves [1-4]. Notice that in [5] some numerical results regarding the amplitude dependence of the wave field on the distance from the surface have recently been obtained.

In the present paper, using a simplified model, neglecting, in particular, the impact of the IDT on the wave propagation in the crystal, we analytically consider the specific features of the distribution of the quasi-bulk wave fields produced by the IDT. We are also concerned with the excitation problem for so-called "non-physical" surface and pseudo surface waves.

2. QUASI-BULK MODES

If the velocity $v = \omega/k$, where k is the tangential projection of the wavevector, equals the limiting velocity $\hat{v}$ (Fig. 1). a pair of plane wave solutions of the type

$$\mathbf{u}_\alpha(\mathbf{r},t),\ \varphi_\alpha(\mathbf{r},t) \propto e^{ik[(\mathbf{m}+p_\alpha\mathbf{n})\mathbf{r}-vt]} \qquad (1)$$

to the wave equation, for example, $\alpha = 3$ and $\alpha = 6$, coalesce to one solution, $\alpha = d3$, of which the parameter p_{d3} is real. The mode $\alpha = d3$ describes the limiting bulk wave. Its group velocity is parallel to the surface of the substrate (Fig. 1).

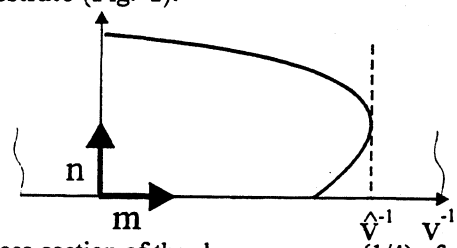


Fig. 1 Cross-section of the slowness curve (1/4) of a bulk wave.

From the examination of Fig. 1 one can infer that near $\hat{v}$ the parameters p_α of the coalescent modes $\alpha = 3, 6$ depend on v as $p_\alpha(v) - p_{d3} \propto (v^2 - \hat{v}^2)^{1/2}$. A more rigorous analysis of plane modes in the vicinity of $\hat{v}$ involves the application of perturbation theory to the eigenvalue problem (see, e.g., [6]) $\hat{\mathbf{N}}(v)\boldsymbol{\xi}_\alpha = p_\alpha\boldsymbol{\xi}_\alpha$, $\alpha = 1,\ldots,8$, where $\hat{\mathbf{N}}(v)$ is a real 8×8 real matrix of which the component are the contractions involving the material tensors of the medium and the vectors m and n; $\boldsymbol{\xi}_\alpha$ is an eight-component vector-column, $\boldsymbol{\xi}_\alpha = \left(\mathbf{A}_\alpha, \Phi_\alpha, \mathbf{L}_\alpha, D_\alpha\right)^t$. Here $\mathbf{A}_\alpha$ is the polarization vector, $\mathbf{L}_\alpha$ is the traction, Φ_α is the electric potential, and D_α is the normal projection of the electrical induction.

At $v = \hat{v}$ the matrix $\hat{\mathbf{N}}$ becomes non-semisimple degenerate. The matrix then possesses only seven eigenvectors, and the generalized eigenvector $\boldsymbol{\xi}_{d6}$ must be introduced in order to have a complete set of eight linearly independent vectors. Near $\hat{v}$ the coalescent eigenvectors and eigenvalues $\alpha = 3, 6$ depend on v as [7,8]

$$\begin{aligned} \boldsymbol{\xi}_{3,6} &= \boldsymbol{\xi}_{d3} \mp \Delta p\boldsymbol{\xi}_{d6} + \textstyle\sum_{n=2}^{\infty}(\mp\Delta p)^n\boldsymbol{\psi}_n, \\ p_{3,6} &= p_{d3} \pm \Delta p \pm \textstyle\sum_{n=2}^{\infty} m_n(\pm\Delta p)^n, \end{aligned} \qquad (2)$$

where the upper sign corresponds to $\alpha = 3$, m_n are real constants, $\boldsymbol{\psi}_n$ are real vectors to be found as linear combinations of the eigenvectors of $\hat{\mathbf{N}}(\hat{v})$ and the vector $\boldsymbol{\xi}_{d6}$; $\Delta p = \sqrt{\rho}|\mathbf{A}_{d3}|f(v)$ or $\Delta p = i\sqrt{\rho}|\mathbf{A}_{d3}|f(v)$, depending on whether the curvature of the slowness curve is positive (like in Fig. 1) or negative at $\hat{v}$, respectively, ρ is the density,

$$\begin{aligned} f(v) &= i\sqrt{\hat{v}^2 - v^2}, \quad v < \hat{v}; \\ f(v) &= \sqrt{v^2 - \hat{v}^2}, \quad v \geq \hat{v}. \end{aligned} \qquad (3)$$

It can be shown that independently of the curvature sign $\alpha = 3$ is either an inhomogeneous mode decaying with depth into the substrate ("physical" mode) or a reflected bulk wave. Accordingly $\alpha = 6$ is either an inhomogeneous mode growing with distance

0-7803-5400-1/99/ $ 10.00

from the surface ("non-physical" mode) or an incident bulk wave. Note that the non-degenerate eigensolutions change with v as $\delta\xi_\alpha$, $\delta p \propto v - \hat{v} \propto (\Delta p)^2$ in the vicinity of $\hat{v}$.

Due to (2) near $\hat{v}$ the determinant of boundary conditions $F(v)$ becomes $F(v) = \sum_{n=0}^{\infty} s_n(\Delta p)^n$, where s_n are constants. A root v_0 of the equation $F(v) = 0$ near $\hat{v}$ corresponds to a quasi-bulk wave. This wave is a linear combination of several modes. The main constituent is the mode $\alpha = 3$. Its amplitude is commonly much larger than the amplitude of other modes. Therefore the resultant amplitude of the quasi-bulk wave and its penetration depth are essentially determined by the appropriate characteristics of the mode $\alpha = 3$.

The quasi-bulk wave may come about either as a surface wave, in which case v_0 is real, $v_0 < \hat{v}$, and Δp is purely imaginary valued, or as a pseudo surface wave. In the latter instance, v_0 is complex valued, $v_0 = v_l' - iv_l''$, in accordance with (1) v_l'' being positive. For complex velocities, it appears convenient to consider that the function $f(v)$ involved in Δp is the analytical continuation of the function (3) onto the complex plane with cut-off branch along $[\hat{v}, \hat{v} - i\infty)$ so that the argument of v ranges from $-\pi/2$ to $3\pi/2$. Accordingly, the argument of Δp ranges from $-\pi/4$ to $3\pi/4$ (henceforth we assume the curvature of the slowness curve to be positive).

In the lowest approximation $F(v) \approx \Delta p - \gamma$, where $\gamma = s_1/s_0$. For the quasi-bulk solution to exist it is required that $|\gamma| \ll 1$ and $\pi/2 \leq arg(\gamma) \leq 3\pi/4$ or $-\pi/4 \leq arg(\gamma) < 0$. Due to the law of energy conservation the argument of $arg(\gamma)$ cannot belong to the interval $0 \leq arg(\gamma) < \pi/2$. On the other hand, it may be that $arg(\gamma)$ falls into the interval from $3\pi/4$ to $-\pi/4$. In this instance the function $F(v)$ does not vanish at small Δp. However, if the solution to the boundary-value problem is formally sought in the form of a linear combination involving, instead of $\alpha = 3$, the mode $\alpha = 6$, then, instead of $F(v)$, one obtains $\tilde{F}(v) = \sum_{n=0}^{\infty} s_n(-\Delta p)^n$, i.e., Δp is merely replaced with $-\Delta p$. Accordingly, $\tilde{F}(v) \approx -\Delta p - b$ and the equation $\tilde{F}(v) = 0$ is solvable in the neighborhood of $\hat{v}$. The linear combination satisfying the boundary conditions is termed a "non-physical" solution. In particular, if $\tilde{F}(v)$ becomes zero at a real $\tilde{v}_0$, there appear the so-called "non-physical" surface wave; this wave involves the mode $\alpha = 6$ of which the amplitude increases with distance from the surface. The roots of the equation $\tilde{F}(v) = 0$ lie on the non-physical sheet of the Riemann surface for the function $F(v)$ but close to the cut-off branch.

Consider now two examples of quasi-bulk waves. Fig.2 shows the velocity and parameter γ for SH-waves propagating on quartz with the Euler angles $(0°, \theta, 90°)$. In this case, γ is purely imaginary valued and $arg(\gamma) = \pm\pi/2$. The "physical" surface wave exists provided $Im(\gamma) > 0$. This condition is always satisfied on the metalized surface. On the free surface, however, the "physical" SAWs (Bleustein-Gulyaev waves)

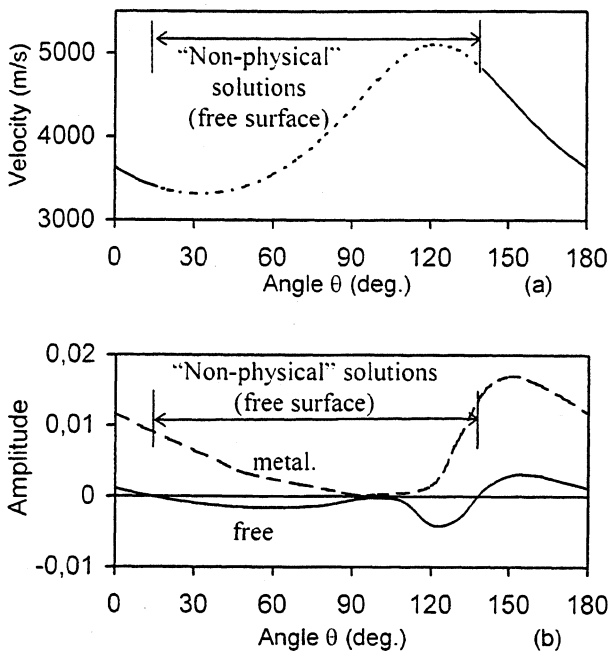


Fig.2. Velocities (a) and Im(γ) (b) for SH-polarized surface waves propagating on free and metallized surface in orientations of quartz with Euler angles (0°,θ,90°), as functions of angle θ.

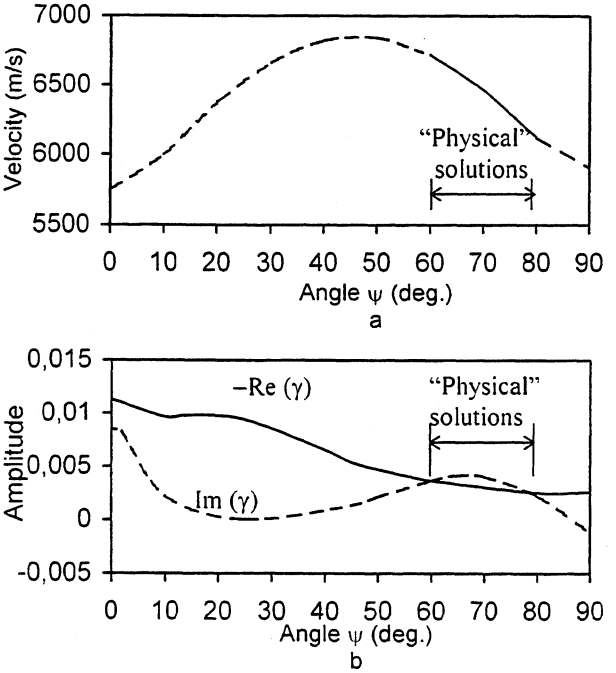


Fig.3. Velocities (a) and γ (b) for quasi-longitudinal leaky waves propagating on metallized surface of ST-quartz, Euler angles (0°,132.75°, ψ), as functions of angle ψ.

exist only in the intervals $0 \le \theta < 15^o$ and $138^o < \theta \le 180^o$. These results agree with the previous investigations of SH-waves in quartz. In particular, the "non-physical" solutions called "improper SAWs" have been found in the interval $15^o < \theta < 138^o$ with velocities below those of the limiting SH-polarized bulk waves [1]. These solutions involve the inhomogeneous mode growing with depth.

The second example is the quasi- longitudinal pseudo-surface wave in ST-cut quartz with the Euler angles $(0^o, 132.75^o, \psi)$ [9]. The results of calculations are shown in Fig.3. In this case, γ is complex valued with $Re(\gamma) < 0$ for any ψ. The "physical" pseudo-surface waves can therefore exist if $\pi/2 \le arg(\gamma) \le 3\pi/4$ and, hence, $Im(\gamma) \ge |Re(\gamma)|$. From Fig.3 we infer that this condition is fulfilled only within the interval $60^o < \psi < 78^o$. Outside it $3\pi/4 \le arg(\gamma) \le 5\pi/4$. Such γ's correspond to the solutions with velocity exceeding the velocity of the limiting quasi-longitudinal bulk mode. These solutions involve the incident bulk wave.

Notice that the interval of the existence of "physical" pseudo-surface waves shown in Fig. 3 is a part of the region specified by the Euler angles $(0^o, \theta, \psi)$, in which solutions of such a type exist (see Fig.5 in ref.[9]).

3. EXCITATION OF QUASI-BULK WAVES

Let an IDT having length $2L$ and finger periodicity d be deposited on the surface of the crystal. The transducer occupies the region $|x| \le L$. We want to estimate the spatial distribution of the amplitude $|u(x,y)|$ of the quasi-bulk wave that the transducer generates when a harmonic signal is applied to it.

The displacement $u(x,y)$ is computed as follows.

$$u(x,y) = \sum_\alpha \int H(k) \frac{F_\alpha(k)}{F(k)} e^{i[kx + k_{\alpha y} y]} dk, \qquad (4)$$

where the summation is carried out over all modes that can be involved in a "physical" quasi-bulk solution, $H(k) = \sin[(k-k_0)L]/(k-k_0)$ is the spatial-frequency characteristics of the IDT with the "central" wavevector $k_0 = (2m+1)\pi/d$. The function $F(k)$ is the determinant of the boundary conditions expressed in terms of the tangential component of the wavevector. The function $F_\alpha(k)$ characterizes the amplitude of the partial mode α.

As has been mentioned, the amplitude of the mode $\alpha = 3$ is commonly much larger than that of the other modes. Hence, we may leave in (4) only the addend $\alpha = 3$. We consider that $\hat{k}L \gg 1$ and that k_0 is close to $\hat{k}$. This allows only two first terms to be kept in $F(k)$. We may also set $F_3(k) \approx const$ and then

$$u(x,y) \propto \int \frac{H(k)}{\sqrt{\hat{k} - k - \beta}} e^{i[kx + k_{3y} y]} dk, \qquad (5)$$

where $\beta = -\gamma\sqrt{\hat{k}}/\sqrt{2}a$ and $a = \sqrt{\rho}|\mathbf{A}_{d3}|\hat{v} \propto 1$. The integrand may still further be simplified for y such that $y/L \ll 4\sqrt{\hat{k}L}/a(3\sqrt{2}-4) \propto 25\sqrt{\hat{k}L}$. Owing to (2) $k_{3y} \approx a\sqrt{2\hat{k}(\hat{k}-k)} + 2a^2 m_2(\hat{k}-k)$, where m_2 is the constant before $(\Delta p)^2$ in (2). Upon integration, we obtain

$$|u(x,y)| = |\Phi(l_2, y/L) - \Phi(l_1, y/L)| \qquad (6)$$

for $x \ge L + 0.5 m_2 a^2 y$ and

$$|u(x,y)| = |\Phi(l_2, y/L) - \Phi(0, y/L)| \qquad (7)$$

for $max(L, 0.5m_2a^2y - L) \le x \le L + 0.5m_2a^2y$, where $l_1 = (x - 1 - 0.5m_2a^2y)/2a^2L$ and $l_1 = (x + 1 - 0.5m_2a^2y)/2a^2L$. In (6) and (7)

$$\Phi(X,Y) = f(X,Y) + 0.5[g_1(X,Y) + g_2(X,Y)], \qquad (8)$$

where

$$f(X,Y) = \frac{\gamma e^{i\Theta^2[(\delta-\gamma^2)X + \gamma Y]}}{\delta - \gamma^2} \cdot erfc\{\frac{\Theta}{\sqrt{i}}(\frac{Y}{2\sqrt{X}} - \gamma\sqrt{X})\}, \qquad (9)$$

$$g_{1,2}(X,Y) = \frac{e^{\pm i\Theta^2\sqrt{\delta}Y}}{\sqrt{\delta} \mp \gamma} erfc\{\frac{\Theta}{\sqrt{i}}(\sqrt{\delta X} \mp \frac{Y}{2\sqrt{X}})\}, \qquad (10)$$

where $\delta = 2a^2(\hat{k} - k_0)/\hat{k}$ and $\Theta = \sqrt{\hat{k}L}$.

The asymptotic behavior of $\Phi(X,Y)$ depends on the type of the solution. If the solution is "physical", then $|arg(-e^{-i\pi/4}\gamma)| \ge \pi/2$. and we find that on the surface of the crystal $(y = 0)$ $\Phi(X,0)$ changes with distance x as

$$\Phi(X,0) \approx \frac{2\gamma e^{i\Theta^2(\delta-\gamma^2)X}}{\delta - \gamma^2} - \frac{e^{i\Theta^2\delta X}}{2\gamma^2\delta\Theta^3}\frac{\sqrt{i}}{X^{3/2}}. \qquad (11)$$

The first term in (11) corresponds to the (pseudo) surface wave. It brings in (6) either an exponential decay if the wave is pseudo surface ($Re[-i\gamma^2] < 0$) or a constant if the wave is pure surface ($\gamma = i\sigma$, $\sigma > 0$). The second term can be viewed as the contribution of bulk modes. In accordance with (6), the contribution from the bulk modes to the resultant wave field decays as $x^{-3/2}$.

When a "non-physical" solution appears in the vicinity of $\hat{v}$, i.e., $|arg(-e^{-i\pi/4}\gamma)| \le \pi/2$, the function $\Phi(X,0)$ involves only the second term in (11) so that the displacements generated by the transducer decrease as $x^{-3/2}$ with distance along the surface.

As to the y-coordinate asymptotic dependence of $\Phi(X,Y)$, it is not difficult to obtain that

$$\Phi(X,Y) \approx -\frac{4\sqrt{i}X^{3/2}}{\sqrt{\pi}\Theta Y^2} e^{i[\frac{\Theta^2}{4}(\frac{Y^2}{X} + 4\delta X)]} \qquad (12)$$

independently of $arg(-e^{-i\pi/4}\gamma)$; Eq. (12) assumes $y/L \gg |\alpha x|, |\sqrt{\delta}x|$. So, in any case $|u(x,y)|$ decreases as y^{-2} with depth into the substrate.

In Fig. 4 the graph is plotted displaying the dependence of the amplitude $u(x,y)$ on x and y for the quasi-longitudinal leaky wave on ST,X+70°-cut quartz ($\gamma = -0.004 + i0.006$) radiated by IDT with N_p=120

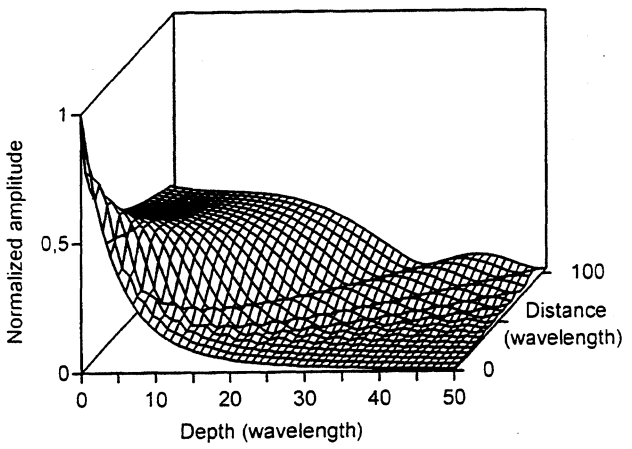


Fig.4 Coordinate dependence of "physical" quasi-longitudinal leaky wave amplitude in orientation of quartz with Euler angles (0°,132.75°,70°), N_p=120.

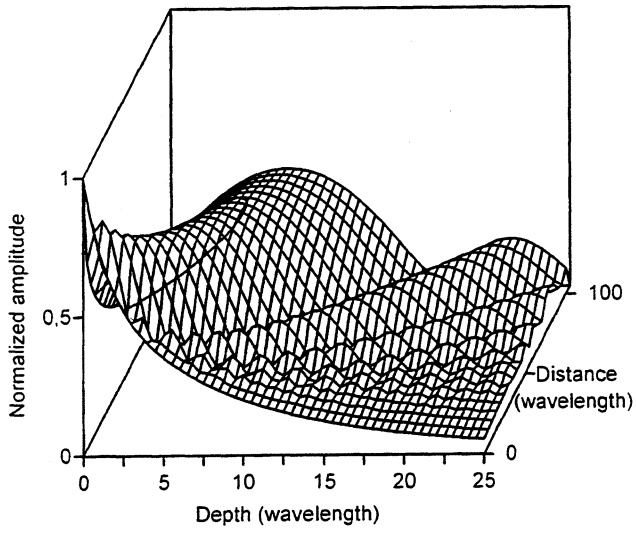


Fig.5 Coordinate dependence of "non-physical" SAW amplitude if $|\gamma\Theta| \approx 1$.

electrode pairs. This field distribution does not change appreciably if a "non-physical" leaky wave is considered (e.g. on ST,X-cut quartz) unless the product $\gamma\Theta$ becomes of the order of unity. Since for SH surface waves and quasi-longitudinal leaky waves under consideration γ does not exceed 0.016 in magnitude (see Fig.2 and Fig.3), the condition $|\gamma|\Theta = 1$ implies $N_p > 1000$, which is far beyond the practical range, N_p=20-100. So, Fig.4 illustrates a typical behavior of a quasi-bulk wave amplitude whether it is "physical" or "non-physical".

Fig.5 shows the coordinate dependence of a "non-physical" surface wave amplitude at $\gamma = -0.07i$ and N_p=200, i.e., for $|\gamma|\Theta \approx 1$ ($\sqrt{\delta} = -\gamma$). One sees that the "non-physical" vibrations decrease abruptly enough with distance along the surface. There also appears a pronounced maximum below the surface. This maximum corresponds to the bulk waves emitted into depth of the substrate.

4. CONCLUSIONS

An analysis of surface and pseudo-surface waves with quasi-bulk structure has been performed. A simple criterion has been suggested to distinguish between the "physical" and "non-physical" plane wave solutions. This criterion has been applied to SH-polarized surface waves and quasi-longitudinal pseudo-surface waves in quartz. The wave fields associated with "physical" and "non-physical" solutions have been investigated. In particular, it appears that the "non-physical" vibrations can be efficiently excited by the IDT with bandwidth of less than 1% provided the "gap" between the velocity of the solution and that of the limiting bulk waves is small enough.

5. ACKNOWLEGMENTS

A.N.Darinskii is grateful to Russian Foundation for Basic Research (98-02-16077) for partial financial support and to the Organizing Committee for granting his participation in 1999 Joint Meeting 13th EFTF - 1999 IEEE IFCS. N.F.Naumenko acknowledges financial support from SAWTEK Inc.

REFERENCES

[1] F.Josse and D.L.Lee,"Analysis of Excitation, Interaction, and Detection of Bulk and Surface Acoustic Waves on Piezoelectric Substrates", IEEE Trans. Sonics Ultrason. v.SU-29(5), 261-274 (1982).

[2] M.Goodberlet and D.L.Lee, "The excitation and Detection of Surface-Generated Bulk Waves", IEEE Trans. Sonics Ultrason. v.SU-31(2), 67-76 (1984).

[3] Y.Zhang and M. Planat, "Effect of piezoelectricity on the excitation and radiation of acoustic waves in rotated Y-cut quartz", J.Appl.Phys. 67(5), 2257-63 (1990).

[4] L.Boyer et al., "Theoretical determination of the pseudo surface acoustic wave characteristic parameters", 1991 IEEE Ultrason. Symp. Proc. 353-358.

[5] S.Biryukov and M.Weihnacht, "Real-space field of surface sources and the problem of fast leaky wave generation in a piezoelectric half-space", J.Appl.Phys. 83(6), 3276-87 (1998).

[6] J. Lothe and D. M. Barnett, "Further development of the theory for surface waves in piezoelectric crystals", Physica Norvegica 8(4), 239-254 (1976).

[7] A.N.Darinskii, "Leaky waves and the elastic wave resonance reflection on a crystal-thin solid layer interface. II. Leaky waves given rise to by an exceptional bulk wave", J. Acoust. Soc. Am. 103(4), 1845-54 (1998).

[8] A.N.Darinskii, "Quasi-bulk Rayleigh waves in semi-infinite media of arbitrary anisotropy", Wave Motion 27(1), 79-93 (1998).

[9] N.F.Naumenko, "The behavior of quasi-longitudinal leaky surface waves in crystals", Proc. 1996 IEEE Ultrason.Symp., pp.107-110.

THE DISSIPATIVE QCM-D TECHNIQUE: INTERFACIAL PHENOMENA AND SENSOR APPLICATIONS FOR PROTEINS, BIOMEMBRANES, LIVING CELLS AND POLYMERS

F. Höök[1,2], M. Rodahl[3], C. Keller[1], K. Glasmästar[1], C. Fredriksson[3], Patrik Dahlqvist[3] and B. Kasemo[1]
[1.] Department of Applied Physics, Chalmers University of Technology and Göteborg University, SE-412 96 Göteborg
[2.] Department of Molecular Biology, Göteborg University, SE-412 96 Göteborg
[3.] Q-Sense AB, Holtermansgatan 1, 412 92 Göteborg

Corresponding author: Bengt Kasemo, Fysikgränd 3, Dep of Appl Phys, Chalmers Univ of Tech and Göteborg Univ, E-mail: kasemo@fy.chalmers.se, Tel: +46-31-7723464, Fax: +46-31-7723134

ABSTRACT

Biological substances in contact with solid, non-biological materials, is a situation of broad scientific interest and technological importance, and there is a growing need for new tools to study these interactions. Among many different properties of the biological films formed at the interfaces, the visco-elastic properties are of central interest, since these properties can be used as a discriminator in bio sensing and in the study of polymer films.

We have developed a sensor system based on the traditional quartz crystal micro balance (QCM) technique, but where both the resonant frequency (f) and the energy dissipation (D) are measured simultaneously for a non-driven (freely oscillating) sensor crystal. This provides accurate and precise measurements of f and D in the gaseous and liquid phases. The performance of this so called QCM-D system is illustrated by measurements of protein adsorption and antibody-antigen reactions, biomembrane formation on surfaces from vesicles in solution, cell attachment experiments, and polymer cross-linking kinetics.

1. INTRODUCTION

Our original motivation for developing a QCM-D sensor capable of fast and precise measurements of both f and D in the liquid phase, was the lack of methods for time resolved measurements of both the adsorbed amount and the conformation (visco-elastic properties) of biofilms (proteins, lipid bilayers or cells) on biomaterials for medical implants. This is an area that has undergone a very fast development in the last decade [1,2]. It soon became obvious that such a sensor would also have a more direct use for biosensing and for studies of polymer films. Application areas include screening of biomaterials [2], immuno assays [3], biofouling [4], polymer processing [5] and cell-surface interactions [2, 4].

2. EXPERIMENTAL

The schematics of the original measurement system [6] is shown in Fig. 1.

The key components are *(i)* the sensor crystal (QCM) mounted in a measurement chamber with facilities for batch mode or flow measurements in liquid or gas, *(ii)* the drive electronics (relay and signal generator) and *(iii)* the recording electronics (probe, reference frequency, filter) including data-handling and software (analog-to-digital converter and computer). The principle of the measurement is to periodically switch on and off the driving power to the sensor and to record the output voltage from the freely decaying oscillator.

This causes the crystal to switch to either its parallel or series resonant frequency, depending on the settings of the recording electronics [7]. The amplitude of oscillation (usually) decays exponentially in time while the driver circuit is disconnected.

The decay of the QCM oscillation is recorded on a digitizing oscilloscope using a high (>10 MΩ) or low resistant probe, which ensures that the crystal decays in the parallel or series mode, respectively [8]. The decaying voltage output from the crystal, which has a frequency given by the resonance frequency (f_0) of the crystal (in our system most often ~5 MHz), is mixed prior to reading with a constant reference frequency

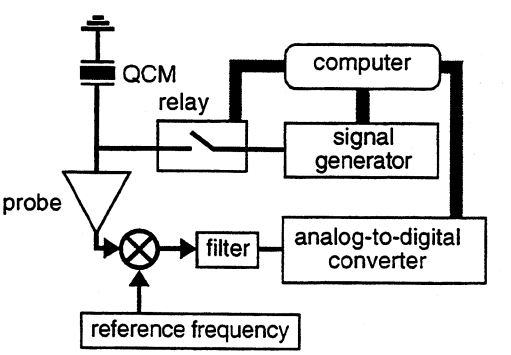


Figure 1 Schematic illustration of the measurement system

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

(f_r), which is about 100 kHz lower than f_0, and filtered in a low-pass filter with a cut off frequency of ~500 kHz. The recorded signal, $A(t)$, is then transferred to a computer, where a numerical fit to an exponentially damped sinusoidal is performed:

$$A(t) = A_0 e^{t/\tau} \sin(2\pi f t + \alpha) \qquad (1)$$

where f (=f_0-f_r) is obtained directly and D via the relation:

$$D = \frac{1}{\pi f \tau} \qquad (2)$$

The energy dissipation is a dimensionless quantity defined as:

$$D = \frac{1}{Q} = \frac{E_{dissipated}}{2\pi E_{stored}} \qquad (3)$$

where Q is the quality factor, E_{stored} is the energy stored in the oscillating system and $E_{dissipated}$ is the energy dissipated during one period of oscillation.

In this way simultaneous measurements of both f and D are done with a repetition rate of typically 1-3 Hz. The resolution in f and D in liquid environment is ±0.1 Hz and $1 \cdot 10^{-7}$, respectively, while in vacuum or air it is a factor of ~10 lower. As a comparison, the shifts in f and D, for a typical protein, vesicle or cell adsorption measurement is 10 to 100 Hz and 0.5 to $4 \cdot 10^{-6}$, respectively (see Fig. 2-7 below) and generally a factor of 10 larger for thick (>100 nm) polymer films (see Fig. 8 below).

The measurement chambers for liquid phase measurements are constructed with special attention to e. g., temperature stability and mechanical stress induced on the crystal when mounted. Thereby temperature- and pressure-induced transients in f and D can be avoided when e. g. one liquid is replaced with another. In order to obtain reliable data in conducting salt solutions it is important to pay attention to capacitive leakage over the crystal. This is either taken care of by measuring in the series mode (see above) or by having the side of the crystal facing the liquid completely covered with the electrode material (most often gold) [8].

The original set up briefly described above has later been improved in a number of ways in a commercial instrument by Q-Sense AB, Gothenburg, Sweden, that we are now using in most of our measurements. The measurements described in the next paragraph have been made with several different set ups with slightly different characteristics. Therefore noise levels and errors may differ in the presented data, even for seemingly identical conditions.

2.1 The frequency to mass conversion

A general comment is appropriate on the conversion of frequency to mass for non-rigid, visco-elastic overlayers. Below, we have applied the Sauerbrey equation [9], well aware that it may not apply in situations when there (for instance) is a risk of surface slip or high dissipation [5, 10, 11]. Special resonance situations may also occur [12]. Another important comment is that any Δf to mass conversion may include considerable amounts of trapped water, as demonstrated previously [10, 13, 14] and below. Rather than adding reservations in each experimental case we make this general reservation here, with the implicit understanding that "mass" really means a f-shift converted to coupled mass (including water) using the Saurbrey relation.

3. RESULTS AND DISCUSSION

3.1 Protein adsorption and antibody-antigen reactions

There is a general tendency for proteins to accumulate spontaneously at aqueous interfacial regions [4]. In many practical situations this is made use of, like in e. g., biosensor applications [3]. In other situations, however, it is an unwanted process, as in e. g., fouling processes during medical treatments [2] or on marine equipment [15]. If the influence from the surface is large enough, the conformational-free-energy minimum for a protein attached on a surface might correspond to a conformation that differs from that of the native protein. It is thus likely that a protein-surface interaction affects the conformation, and hence the function of the proteins, which in turn might cause severe problems in many practical situations [4, 16]. The QCM-D technique offers a possibility investigate changes in the visco-elastic properties to in real time, which are directly related to the conformation of the adsorbed proteins. Examples of how simultaneous measurements of f and D contribute with such information are presented below (Figs 2-5).[i]

As a first example we choose to show adsorption measurements of hemoglobin[ii] (Hb) at pH 7.5 and low salinity (< 10 mM NaCl) and ferritin[iii] at pH 7.5 and high salinity (= 200 mM NaCl). The f- and D-shifts of these molecules versus time are shown in Fig. 2 a. The dissipation shifts seem in both cases to follow the frequency shift, but not exactly. This is illustrated Fig. 2 b, which display the f-shift ($\Delta f(t)$) vs the D-shift ($\Delta D(t)$). We have found this to be an extremely valuable way of displaying combined $\Delta f(t)$ and $\Delta D(t)$ data. Absolute slopes and changes in slopes in these D-f plots provide unique information about different kinetic regimes, phase transitions, and conformational changes [10, 17-19, 20] (see also below).

[i] All measurements in Section 3.1 were made on a hydrophobic, electrically inert methyl (-CH_3) terminated thiolated gold surface.

[ii] Hemoglobin is a spherical protein with a diameter of ~5 nm and a molecular weight of ~65 kD.

[iii] Ferritin is a spherical protein with a diameter of ~12.5 nm and a molecular weight of ~630 kD.

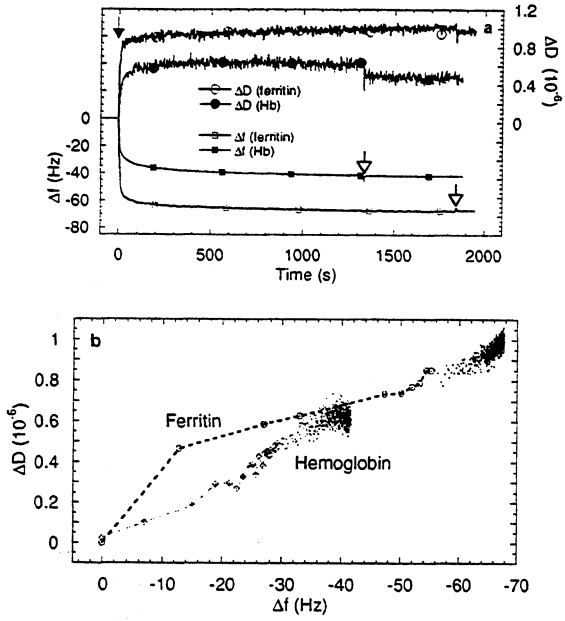


Figure 2 (a) Δf and ΔD vs time and (b) ΔD vs Δf for Hb and ferritin adsorption. In A, the filled arrow indicates addition of protein, the open arrows indicate exchange to pure buffer.

The following two examples illustrate how surface close packing of these proteins is modified by salt concentration and pH. Figure 3 shows how the mass-uptake ($|\Delta f|$) and $\Delta D/|\Delta f|$ of Hb at saturation varies with pH (measured at low ionic strength <10 mM NaCl). This commonly observed bell-shaped curve for the mass-uptake of proteins is primarily due to electrostatic repulsion between the adsorbed molecules on either side of the isoelectric point (pI, which for Hb is ~6.7) due to excess of $-H^+$ or $-OH^-$ at the exterior of the protein molecules, at low and high pH, respectively [16]. It also interesting to note that the adlayers formed at high and low pH, as well as close to pI induce the largest dissipation per adsorbed mass (Δf). We attribute this primarily to differences in the water shells around the proteins, and its consequences for the energy dissipation. Note, however, that the nature of the hydrodynamic coupling between the adsorbed protein molecules and the surrounding liquid depends directly on the conformation of the protein molecules and their distribution in the adlayer. At low and high pH the large $\Delta D/|\Delta f|$ values are attributed to *(i)* the increase in flexibility of individual protein molecules known to occur as the repulsion between their surface charges increases and *(ii)* to the tendency of protein molecules that carry a net charge to adsorb as separated molecules. In contrast, the large $\Delta D/|\Delta f|$ values close to pI is attributed to bilayer formation due to surface induced denaturation of the protein molecules directly attached to the surface, allowing further adsorption on top. At this pH bilayer formation is not hindered by electrostatic repulsion, since the net-charge of the protein molecules is close to zero. This interpretation was further supported by both the high mass uptake and the fact that the ratio

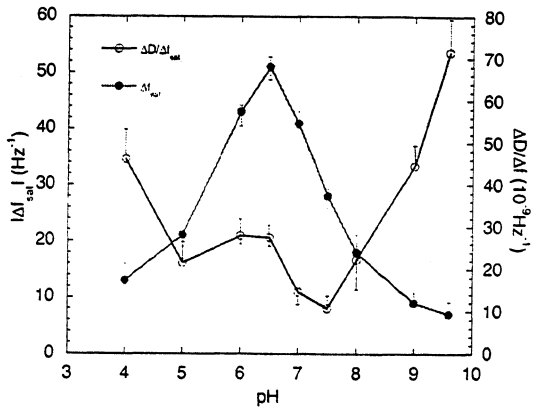


Figure 3 The absolute value of Δf (filled circles) and $\Delta D/\Delta f$ (empty circles) at saturation for the adsorption of Hb vs pH.

between $\Delta D(t)$ and $\Delta f(t)$ was actually not linear during the adsorption (see reference [17] for details).

Figure 4 shows for ferritin, how the surface coverage at saturation (at constant pH=7.5, which is ~2 pH units above its pI) varies strongly with salt concentration [19]. This behavior is also due primarily to electrostatic effects: The role of increased salt concentration is increased screening of the electrostatic repulsion between proteins (the Debye screening length is reduced). At low salinity, this results in virtually now adsorption. Above 100-200 mM solution of NaCl, the amount adsorbed is actually 1.4 times higher than for

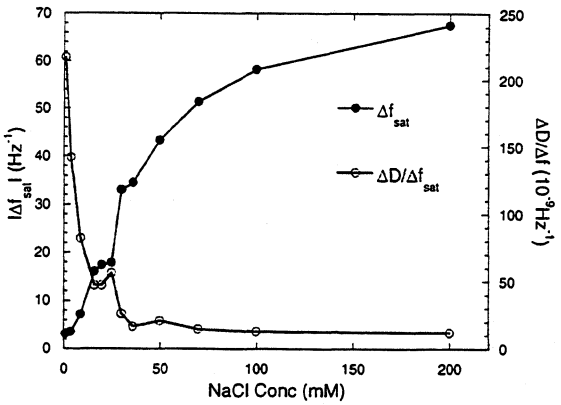


Figure 4 The absolute value of Δf (filled circles) and $\Delta D/\Delta f$ (empty circles) at saturation for the adsorption of Ferritin vs NaCl concentration

a monolayer of proteins when the overestimation due to hydrodynamically coupled water is also taken into account (not shown). (Monolayer means here that obtained assuming irreversible random sequential adsorption (RSA), which corresponds to ~54 % of a close packed monolayer[1].) Interestingly, the dissipation per adsorbed mass is much larger at low coverage in low salinity, compared to high coverage/high salinity as also shown in Fig. 4. We attribute this difference primarily to clustering of ferritin molecules at the surface at high salinity, which as a consequence

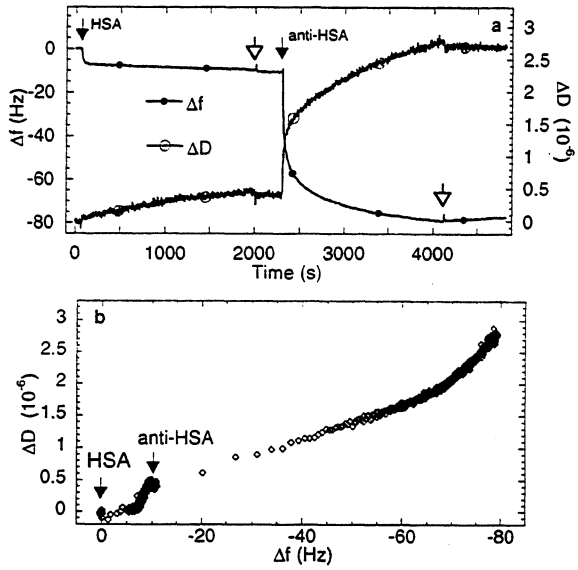


Figure 5 Δf and ΔD vs time (a) and ΔD vs Δf (b) for HSA adsorption followed by exposure to anti-HSA.

decreases the flexibility of the adlayer and reduces the coupling to the surrounding water.

We now turn to an example of an antibody-antigen reaction. The chosen model system is Human Serum Albumin[iv] (HSA) and anti-HSA[v].The experiment was done by sequentially exposing the methyl-thiol coated gold surface first to HSA, then to anti-HSA including exchange to a pure buffer solution after saturated adsorption of HSA and anti-HSA. Figure 5 shows (a) $\Delta f(t)$ and $\Delta D(t)$ and (b) the D-f plot for these sequences. In both panels it is clearly demonstrated that the antibody-antigen reaction can be detected with good S/N. Moreover, the D-shift induced by anti-HSA adds information beyond that obtained directly through $\Delta f(t)$, as demonstrated by the two linear regimes that are displayed in the D-f plot. Without going into a detailed interpretation of the structural rearrangements that are likely to be responsible for this observation (for details see reference [18]), we emphasize the added value from this kind of multivariable analysis for sensing of specific immuno processes, or recognition events in general. For instance, the $\partial D/\partial f$ value during antibody recognition generally varies for different antigens and/or antibodies (which will be published elsewhere [20]).

3.2 Vesicle adsorption, decomposition and fusion to supported membranes.

The formation of supported membranes [21] is an important component in surface functionalization for biosensors, future medical implants and protein-repelling surfaces. A common way of making such supported membranes is to use vesicles or liposomes that upon adsorption decompose/fuse forming extended membranes covering the surface. The interaction is surface specific, i. e. with otherwise constant conditions, different surfaces give rise to different end results [22]. The building blocks of vesicles and biomembranes are linear amphiphilic molecules: one end (the head group) is hydrophilic, the tail and the other end are hydrophobic. These chemical properties make the molecules self assemble in water into structures that hide the hydrophobic groups from, and expose the hydrophilic end groups to, water. As a result bilayer structures are formed, such as flat membranes or spherical vesicles.

The question posed here is: What happens to a vesicle that comes in contact with a solid surface? Figs. 7 a-c show experiments that demonstrate three distinctly different results for the adsorption of egg-phosphatidylcholine vesicles (25 nm in diam.) onto three different surfaces. We can explain these results by assigning a different structure to each of the adsorbed films; these are indicated schematically in each figure (for details see [22]). The interesting points in the present context are the following:

(i) On the hydrophobic, methyl-terminated, thiolated gold surface (Fig. 7a) the vesicles spontaneously (faster than the time scale of the experiment) decompose on the surface and cover it with a monolayer of lipid molecules. This is the thermodynamically expected result, driven by the hydrophobic interaction between the surface and the phospholipid molecules. The frequency shift (-13 Hz) is of the magnitude expected for a close packed monolayer, and the dissipation shift is barely detectable, indicating a quite rigid, nonslipping monolayer.

(ii) On the oxidized (by ozon exposure [23]) very hydrophilic gold surface (Fig. 7b), the frequency and dissipation shifts are very large, and the Δf value is consistent with a monolayer of intact vesicles [22]. (The application of the Sauerbrey equation may in this case be questioned in view of the large D-shift.) The large dissipation increase is attributed to the viscous structure of the vesicle layer; upon shear oscillation it deforms and gives rise to internal friction within the layer of vesicles and hydrodynamically coupled water.

(iii) On hydroxylated, hydrophilic SiO_2 (Fig 7 b) the vesicles are initially causing changes in f and D very similar to those on oxidized gold (c. f. Fig 7 c). However, at a critical vesicle coverage, f suddenly shifts towards higher values (apparent mass loss) and D decreases. Since vesicles are known to form bilayers on SiO_2 [24], the interpretation is that intact vesicles populate the surface up to the critical coverage (minimum in $\Delta f(t)$), after which they fuse to a bilayer membrane. This interpretation is further supported by the fact that the f-shift at completion was exactly twice that of the monolayer of lipids (c. f. Fig. 7 a). The apparent mass loss is most likely a real mass loss due to the loss of the trapped water inside and between the vesicles.

[iv] HSA is a cylindrically shaped protein with a molecular weight of ~65 kD.

[v] anti-HSA is a Y-shaped protein with a molecular weight of ~150 kD.

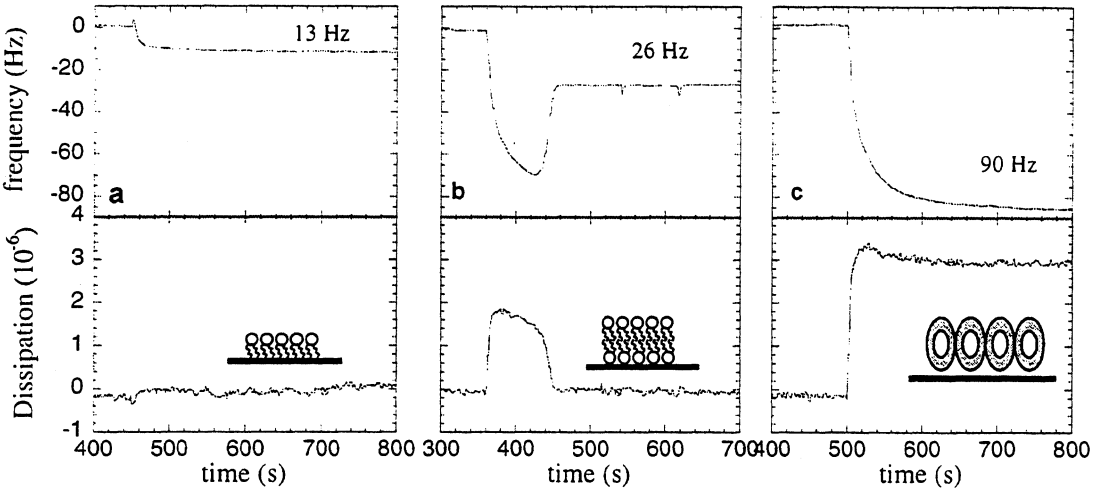


Figure 6 $\Delta f(t)$ (upper panels) and $\Delta D(t)$ (lower panels) for addition of small unilaminellar vesicles on a (a) hydrophobic methyl terminated, thiolated gold surface, (b) a hydrophilic SiO_2 surface, and (c) a hydrophilic oxidized gold surface.

A more extensive account of these results will be published elsewhere [25].

3.3 Cell-surface interactions and cell processes

When applying the QCM-D technique to study biological cells, one should realize that this represents a much more complex situation, compared to proteins, vesicles etc. Cells consist of plasma membranes (5-10 nm thick) much like the phospholipid bilayers discussed above, equipped with a variety of transmembrane molecules for attachment, signaling and surface recognition. Cells also have very complex internal structure, such as their cytoskeletal network, and furthermore they are "alive".

The size of a cell (1-100 μm depending on type and shape etc.) means that the QCM-D technique essentially samples the properties of the fraction of the cell closest to the surface (due to the limited shear wave extinction depth). In spite of these complications, our initial results have shown very encouraging results concerning the dynamics of cell adhesion [26-28]. In particular, these studies have shown that the attachment of a cell layer causes substantial energy dissipation due to active intracellular mechanisms.

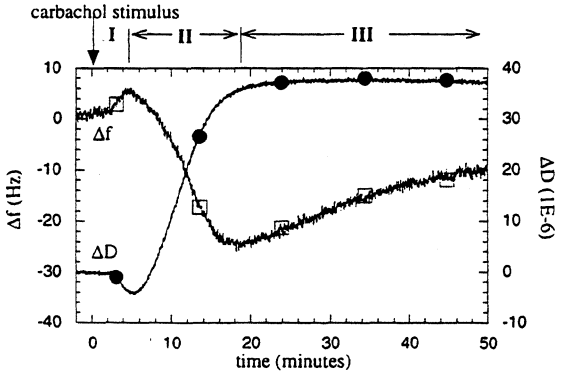


Figure 7 Diagram showing the multiphase response response in Δf (open squares) and ΔD (filled circles) of a confluent cell layer to added secretion stimulus. The mucussecreting human epithelial (colon) cells were cultured directly on a tissueculture prepared sensor surface.

Changes in the internal cell structure due to e.g., external stimuli, can therefore be monitored over time via this dissipation change. Hence, the technique can be useful sensor for pharmacological, toxicological or radiation treatment purposes etc.

The single example shown here is displayed in Fig. 7. It shows results from a mucus secreting human epithelial cell layer, cultured directly on a tissue-culture quality sensor surface [29]. The signals show that stimulation of these cells to secrete mucus, using an established secretagogue (carbachol), causes a multiphase response. An initial (phase I) decrease of mass and dissipation appears to be a quite general feature of secretion by exocytosis. The second phase might tentatively be correlated to the amount and/or quality of the secreted mucus depositing on the sensor, since there are indications that its magnitude correlates with the amount of mucus secreted, as measured by retrospective collection of the secreted material. No interpretation is yet suggested for the third phase.

There are to the best of our knowledge no other method that so sensitively can monitor *in vitro* the mucus secretion process of a significant area of cells under physiological conditions in real time.

2.4 Cross-linking in polymer films.

QCM-D has the potential to be a valuable tool to study curing kinetics in thin polymer coatings. Curing kinetics is difficult to monitor with conventional techniques, which is a drawback considering the vast number of applications for polymer coatings. One example is environmentally friendly nonfouling coatings for marine vessels based on fluorine enriched silicone. By combining the mobility and flexibility of silicone with the low surface energy of fluorine one hopes to make a coating that is resistant to marine biofouling [30].

We have studied cross-linking in a coating based on silanol terminated polydimethylsiloxane (PDMS) which is cross-linked by (tridecafluoro1,1,2,2,tetrahydrooctyl) triethoxysilane

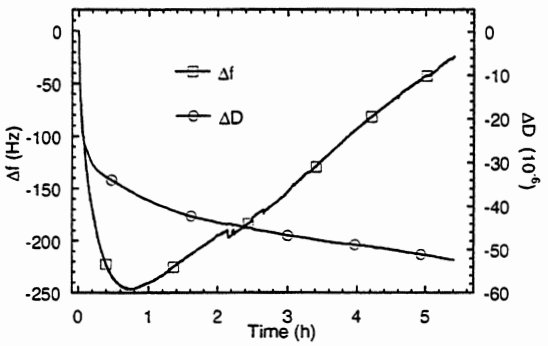


Figure 8 Δf (squares) and ΔD (circles) upon cross-linking of PDMS by FTEOS.

(FTEOS). The cross-linker, with its fluorine rich tail, has a tendency to migrate to the surface during curing rendering the coating hydrophobic. The cross-linking is catalyzed by dibutyltindiacetate. Cross-linking starts with the hydrolysis of the triethoxysilane ends of the cross-linker. For each water molecule used one ethanol molecule is abstracted.

In this study, the silicone coating was spread on a 10 MHz crystal and the resulting frequency shift was approximately 10 kHz and the dissipation shift was approximately 450×10^6. The crystal was mounted in a measurement chamber that was held at room temperature and 100% relative humidity.

Figure 8 shows how the curing process can be followed, in real time, with the QCM-D technique. Initially, there is a rapidly decreasing resonant frequency indicating a net uptake of mass by the coating. This mass uptake is most likely due to water absorption, which activates cross-linking in the film. After approximately 50 minutes, the resonant frequency starts to increase indicating that ethanol desorption dominates over water adsorption. During the whole curing process, the dissipation factor decreases continuously as the viscous losses in the film decrease.

4. CONCLUDING REMARKS

The examples above demonstrate that the present QCM-D system provides sensitive and well controlled measurements of f and D-shifts in real time ($\leq$ 1s) in the liquid phase and, most important, the simultaneous measurement of the f and D shifts provides superior insight into overlayer processes involving biofilms and other viscous overlayers, compared to measurements of f shifts alone. The D-f plots eliminate time as an explicit parameter, and are a useful way of identifying transitions between different kinetic regimes or real structural phase transitions. The present and other similar data form an important database for ongoing theoretical development [31] to make more quantitative interpretations possible, with regard to the shear visco-elastic properties, and internal and interfacial processes in the overlayers. From the collected data set and comparison with other measurement methods, it is already evident that trapped water plays an important role in QCM-D studies of biological monolayers and biofilms on surfaces.

ACKNOWLEDGMENTS

This work has been and is financially supported by Swedish Research Council for Engineering Sciences (TFR), and the Biomaterial Consortium and the Biocompatible Materials Program, both funded by the Foundation of Strategic Research (SSF).

REFERENCES

[1] J. J. Ramsden, "Review of New Experimental Techniques for investigating Random Sequential Adsorption," J of Stat Phys, vol. 73, pp. 853, 1993.

[2] B. D. Ratner, A. S. Hoffman, F. J. Schoen, and J. E. Lemons, "Biomaterials Science; An Introduction to Materials in Medicine," . San Diego: Academic Press, 1996; J Gold and B Kasemo " Implant Surfaces and Interface processes" Adv in Dental Research (in press 1999) & B. Kasemo and J. Lausmaa, "The Biomaterial Biosystem Interface" in Surface Characterization: A User´s Sourcebook, Ed H Brune (WileyVCH, Weinheim 1997).

[3] A. F. Collings and F. Caruso, "Biosensors: Recent Advances," Rep. Prog. Phys, vol. 60, pp. 1397-1445, 1997.

[4] J. L. Brash and P. W. Wojciechowski, "Interfacial Phenomena and Bioproducts," in Bioproc Tech, W. C. McGregor, Ed. New York, Basel, Hong Kong: Marcel Dekker, Inc., 1996.

[5] H. L. Bandey, A. R. Hillman, M. J. Brown, and S. J. Martin, "Viscoelastic Characterization of Electroactive Polymer Films at the Electrode/Solution Interface," Faraday Disc 107: Acoustic waves and Interfaces, Lester UK, vol 107, pp 105-122. 1997.

[6] M. Rodahl, F. Höök, A. Krozer, P. Brzezinski, and B. Kasemo, "Quartz Crystal Microbalance Set Up for Frequency and Qfactor Measurements in Gaseous and Liquids environments," Rev of Sci Instr, vol. 66, pp. 3924-3930, 1995.

[7] M. Rodahl and B. Kasemo, "A Simple Setup to Simultaneously Measure the Resonant Frequency and the Absolute Dissipation Factor Of a Quartz Crystal Microbalance," Rev of Sci Instr., vol. 67, pp. 3238-3241, 1996.

[8] M. Rodahl, F. Höök, and B. Kasemo, "QCM Operation In Liquids: an Explanation Of Measured Variations In Frequency and Q Factor With Liquid Conductivity," Anal Chem, vol. 68, pp. 2219-2227, 1996.

[9] G. Saurbrey, Z Phys., vol. 155, pp. 206-222, 1959.

[10] M. Rodahl, F. Höök, C. Fredriksson, C. Keller, A. Krozer, P. Brzezinski, M. Voinova, and B. Kasemo, "Simultaneous Frequency and Dissipation Factor QCM Measurements of Biomolecular Adsorption and Cell Adhesion," Faraday Disc 107:

Acoustic waves and Interfaces, Lester UK, vol. 107, pp. 229-246, 1997.
[11] M. D. Ward and D. A. Buttry, "In Situ Interfacial Mass Detection with Piezoelectric Transducers," Science, vol. 249, pp. 1000, 1990.
[12] K. Yoonkee, J. R. Vig, and A. Ballato, "Sensing the Properties of Liquids with Doubly Rotated Resonators," presented at Proceedings of the 1998 IEEE International Frequency Control Symposium, New York, USA, 1998.
[13] M. Muratsugu, F. Ohta, Y. Miya, T. Hosokawa, S. Kurosawa, N. Kamo, and H. Ikeda, "Quartz Crystal Microbalance for the Detection of Microgram Quantities of Human Serum Albumin: Relationship Between the Frequency Change and the Mass of Protein Adsorbed," Anal Chem, vol. 65, pp. 2933, 1993.
[14] F. Caruso, D. N. Furlong, and P. Kingshott, "Characterization Of Ferritin Adsorption Onto Gold," J of Coll Interface Sci, vol. 186, pp. 129, 1997.
[15] J. H. Waite, "Nature's Underwater Adhesive Specialist," Int J of Adhesion and Adhesives, vol. 7, pp. 914, 1987.
[16] W. Norde, "Driving Forces for Protein Adsorption at Solid Surfaces," Macromol. Symp., vol. 103, pp. 581, 1996.
[17] F. Höök, M. Rodahl, B. Kasemo, and P. Brzezinski, "Structural Changes in Hemoglobin During Adsorption to Solid Surfaces: Effects of pH, Ionic Strength, and Ligand Binding," Proc Nat Am Sci, vol. 95 (21), pp. 12271-12276, 1998.
[18] F. Höök, M. Rodahl, P. Brzezinski, and B. Kasemo, "Energy Dissipation Kinetics for Protein and Antibody-antigen Adsorption under Shear Oscillation on a Quartz Crystal Microbalance," Langmuir, vol. 14 (4), pp. 729-734, 1998.
[19] F. Höök, M. Rodahl, P. Brzezinski, and B. Kasemo, "QCMmeasurements of Ferritin Monolayers on MethylThiolated Gold; Dependence of Energy Dissipation and Saturation Coverage on Salt Concentration," J Coll and Interface Sci, vol. 208, pp. 263-267, 1998.
[20] F. Höök and et. al, To be published.
[21] E. Sackmann, "Supported Membranes: Scientific and Practical Applications," Science, vol. 271, pp. 438, 1996.
[22] C. A. Keller and B. Kasemo, "Surface Specific Kinetics of Lipid Vesicle Adsorption Measured with a Quartz Crystal Microbalance," Biophys J, vol. 75 (3), pp. 1397-1402, 1998.
[23] A. Krozer and M. Rodahl, " Xray Photoemission Spectroscopy Study of UV/ozone of Au under Ultrahigh Vacuum Conditions," J. Vac. Sci. Tech. A, vol. 15, pp. 1704-1709, 1997.
[24] J. Salafsky and et. al, "Architecture and Function of Membrane Proteins in Planar Supported Bilayers: a Study with Photosynthetic Reaction Centers" Biochemistry, vol. 35, pp. 14773, 1996.
[25] C. Keller, K. Glasmästar, and B. Kasemo, To be published.
[26] C. Fredriksson, S. Kihlman, M. Rodahl, and B. Kasemo, "The Piezoelectric Quartz Crystal Mass and Dissipation Sensor: A Means of Studying Cell Adhesion," Langmuir, vol. 14, pp. 248-251, 1998.
[27] C. Fredriksson, S. Kihlman, B. Kasemo, and D. M. Steel, "*In Vitro* RealTime Characterization of Cell Attachment and Spreading," J of Material Sci: Materials in Medicine, vol. 9, pp. 785-788., 1998.
[28] G. Nimeri, C. Fredriksson, H. Elwing, L. Liu, M. Rodahl, and B. Kasemo, "Neutrophil Interaction with Protein Coated Surfaces Studied by an Extended Quartz Crystal Microbalance Technique," Coll and Surfaces B: Biointerfaces, vol. 11, pp. 255-264., 1998.
[29] D. M. Steel and e. al., To be published.
[30] T. M. Chapman, R. Benrashid, K. G. Marra, and J. P. Keener, Determination of Low Critical Surface Tensions of Novel Fluorinated Poly(Amide Urethane) Block-Copolymers"Macromolecules, vol. 28, pp. 331-335., 1995.
[31] M. V. Voinova, M. Jonson, and B. Kasemo, "Dynamics of Viscous Amphiphilic Films Supported by Elastic Solid Substrates," J.Phys.: Condens. Matter, vol. 9, pp. 7799-7808, 1997.

SURFACE ACOUSTIC WAVE-BASED SENSORS USING MODE CONVERSION IN AN ARRAY OF PERIODIC GRATINGS

Florian Bender, Reiner Dahint and Fabien Josse*
Angewandte Physikalische Chemie, Universitaet Heidelberg,
Im Neuenheimer Feld 253, 69120 Heidelberg, Germany.
*Microsensor Research Laboratory and Department of Electrical and Computer Engineering
Marquette University, P.O. Box 1881, Milwaukee, WI 53201-1881

ABSTRACT

Novel surface acoustic wave (SAW) sensor platforms with electrode-free sensing surface are being investigated for use in applications in which the sensing environment may be conducting or corrosive. It is shown that, by using a system of four grating arrays, the IDT generated SAW can be directed to the opposite surface where it interacts with a chemical sensing medium before being directed back to the electroded surface. The design principles of these grating-based SAW sensor platforms are discussed. The signal quality and the mass sensitivity of the devices are analyzed and compared to those of a direct path SAW using chemical vapor deposited thin polymer films. Experimental results are presented for the detection of nitrogen dioxide (NO_2) in sub-ppm concentrations on devices fabricated on YZ-$LiNbO_3$ plates.

1. INTRODUCTION

Perhaps of all the sensor technologies, surface acoustic wave (SAW) sensors offer the opportunity of realizing robust, low cost sensitive detectors for chemical environments [1]. However, if the sensing environment is conductive or corrosive, care has to be taken to protect the transducers. This is the case, for example, in liquid-phase applications utilizing shear horizontally polarized (SH-SAW) as the sensing wave. An alternative approach is to use acoustic plate mode (APM) devices, because the chemical analyte can then be detected on the surface opposite to the transducers.

Although APM sensors have proven to be a promising tool for liquid-phase analysis, the implementation of high-sensitivity devices still provides several difficulties. The difficulties in APM sensor design are twofold. First, in order to increase mass sensitivity it is desirable to operate APMs at relatively high frequencies. However, under these conditions the mode density becomes very high, and mode interference may deteriorate the sensor response and worsen reproducibility. Second, the acoustic wave energy is distributed across the thickness of the crystal and not confined to the surface as in the case of SAWs. Thus, the mass sensitivity is intrinsically lower. In order to overcome these drawbacks but still take advantage of a non-electroded sensing surface, devices based on mode conversion in periodic gratings could be investigated and tested for their ability to be used in gas and liquid phase analysis.

In the present work, the design, characterization and mass sensitivity of a novel SAW sensor platform with electrode-free sensing surface utilizing an array of periodic gratings are presented.

2. GRATING COUPLING BETWEEN SAWS AND APMS

Periodic gratings on the surface of a piezoelectric substrate can cause a propagating SAW to radiate its energy into the plate. To propagate as plate modes, the generated bulk modes must satisfy the plate boundary conditions. The coupling between the SAW and the bulk modes via the periodic gratings can be appropriately described in terms of the coupling of mode equations. Thus, the wave solution can be obtained as a superposition of acoustic plate modes (APM) and SAWs which match the boundary conditions in the grating surface.

The analysis shows that a SAW of wavelength λ incident on a grating of period $p \approx \lambda$ produces a series of narrow bandwidth bulk modes. These modes have been identified as acoustic plate modes. SAW mode conversion into bulk modes occurs if the following coupling condition is satisfied [2]:

$$k_{SAW} = k_{APM} + k_G \qquad (1)$$

with $k_{SAW} = 2\pi f / v_{SAW}$, $k_G = 2\pi / p$ and

$$k_{APM} = \sqrt{(2\pi f)^2 / v_B^2 - (m\pi)^2 / t^2}$$

In the above equations, k_{SAW} represents the propagation constant of the SAW, k_{APM} is the propagation constant of the APM and k_G, a vector of the reciprocal lattice, is related to the grating period. v_{SAW} and v_B are the phase velocities of the SAW and the bulk mode, respectively; f denotes the frequency of operation, m the order of the APM, and t the thickness of the crystal plate. The bulk mode can have either longitudinal or transverse particle displacement. For $p \approx \lambda$, the bulk mode is launched into the substrate at almost 90° to the surface. This mode can then propagate through the plate by multiple reflections.

Using eq. (1), a formula describing the frequencies of the bulk modes generated in a given grating can be derived as [2]:

0-7803-5400-1/99/ $ 10.00

$$f(m)=\frac{v_{SAW}}{p}\cdot\frac{1-(v_{SAW}/v_B)\cdot\sqrt{1-(mv_B p)^2/(2v_{SAW}t)^2\cdot(1-v_{SAW}^2/v_B^2)}}{1-v_{SAW}^2/v_B^2} \quad (2)$$

Based on this concept, new sensor platforms can be developed using mode conversion in etched gratings. One of these sensors utilizes all the advantages of the commonly used SAW devices (with the sensing surface located on the transducer surface) in addition to isolating the transducers from the chemical medium. This, in turn, will reduce or eliminate signal drifts and instability, a major problem with SAW chemical sensors that can affect the detection limit.

3. SENSOR DESIGN AND PRINCIPLES

APM grating devices

The two sensor configurations investigated are shown in Fig. 1. In the first configuration, as shown in Fig. 1a, two arrays of periodic gratings etched into the lower crystal surface are utilized to convert a SAW into an APM and vice versa. For this configuration, the detection process is still mediated by an APM; however, the waves are now excited by a resonant process. As a result, mode separation is significantly improved compared to conventional APM devices.

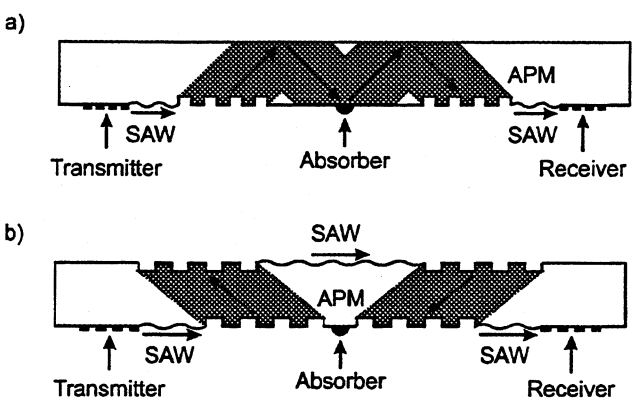


Fig. 1: *Mode conversion in periodic gratings (schematic). a) Excitation of a narrowband APM with improved mode separation. b) Implementation of a SAW device with electrode-free sensing surface.*

To prove the advantages of the new design principle, eq. (2) has been utilized to calculate the mode spectrum of a grating device on a 0.5 mm thick YZ-$LiNbO_3$ plate with a grating period of 23.23 µm. The results of the calculation are shown in Fig. 2b. For this crystal orientation the phase velocities of the SAW, longitudinal and shear bulk modes are 3485 m/s, 4444 m/s and 6883 m/s, respectively. A spectrum of well-separated modes is obtained around a SAW frequency of 150 MHz. Up to now no theory exists to accurately predict the strength of the individual modes. For comparison, a conventional APM spectrum is shown for the same frequency regime in Fig. 2a. The material parameters are selected as above and an IDT period of 30 µm was chosen to excite a high coupling PSAW-APM at about 150 MHz.

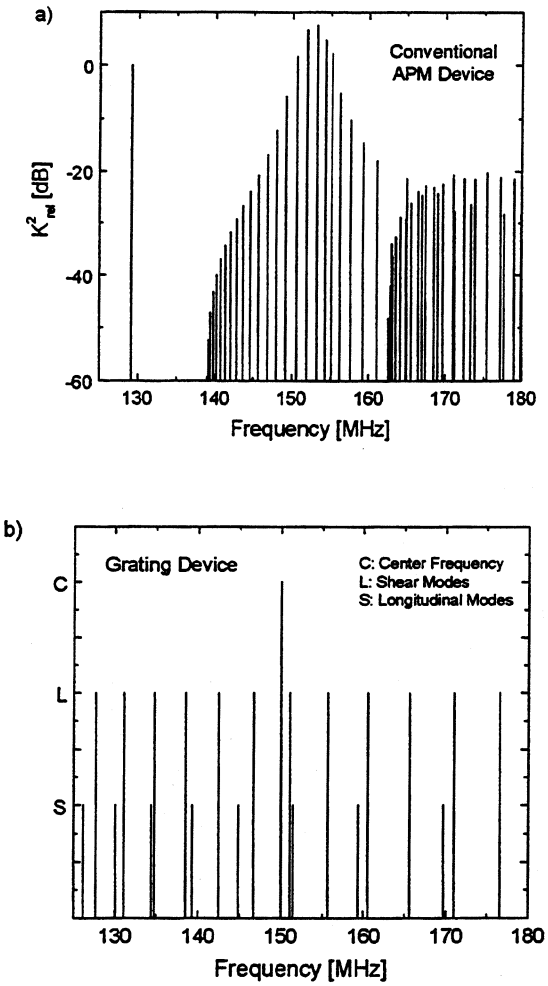


Fig. 2: *Comparison of the mode spectra for conventional and grating-based APM sensors at about 150 MHz. a) Conventional APM device on 0.5 mm ZX-$LiNbO_3$. b) Grating device on 0.5 mm YZ-$LiNbO_3$. The improved mode separation of the new sensor design is clearly revealed.*

SAW grating devices

A schematic diagram of the second grating sensor design is shown in Fig. 1(b). In this new type of SAW device, the SAW is generated by a conventional interdigital transducer (IDT), directed to the opposite surface where it interacts with the analyte, and finally directed back to the electroded surface where it is received by a second IDT. In this geometry, four grating arrays are needed on each acoustic delay line.

Dual delay line devices were designed and fabricated on YZ-$LiNbO_3$ plates of thickness t = 0.5 mm. Each grating consists of 200 grooves with a depth of about 0.4 µm and a grating period of p = 16 µm. Such values have been found

to allow for efficient mode coupling on this material [2, 3]. In order to avoid strong reflection of the SAW in the grating, the IDT period and thus the SAW wavelength, λ, must be approximately equal to p (or an odd multiple, or a fraction of p). On the other hand, problems with reconversion of the excited waves suggest that a small difference between λ and p is needed. This is due to the following facts: If $p \approx \lambda$, it follows from eq. (1) that $k_{APM} \approx 0$ and thus the phase velocity $v_{APM}=2\pi f/k_{APM}$ of the APM will be very large. The resulting mode can be considered to be a superposition of bulk waves of the same frequency $f=\omega/2\pi$ as the SAW, but with their wavefronts almost parallel to the surfaces of the plate. This results in a very large phase velocity v_{APM} of the APM while its group velocity $v^G_{APM}=\partial\omega/\partial k_{APM}$ will be very small.

These facts will affect the bandwidth of the APM, which is calculated as

$$\frac{\Delta f}{f}=\frac{\Delta\lambda}{f}\frac{\partial f}{\partial\lambda}=\frac{\Delta\lambda}{f}\frac{\partial f}{\partial\frac{1}{\lambda}}\frac{-1}{\lambda^2}=\frac{-\Delta\lambda}{\lambda}\frac{\partial f}{\partial\frac{1}{\lambda}}\frac{1}{f\lambda}=\frac{-\Delta\lambda}{\lambda}v^G_{APM}\frac{1}{v_p} \quad (3)$$

If $k_{APM} \approx 0$, the case of a standing wave is approached, much like in a quartz crystal resonator (QCR), and the resulting bandwidth $\Delta f/f$ will be very small. Since the plate thickness, t, and thus the APM phase velocity may vary slightly along the delay line, the exact mode coupling frequencies of two gratings may differ by a small amount which must not be larger than the bandwidth of the mode. Otherwise, a mode generated in the first grating cannot be reconverted in the second grating. For this reason, a 10% mismatch between the grating and IDT periods has been chosen, ensuring that k_{APM} and thus the bandwidth $\Delta f/f$ will not be too small.

Also, as seen from eq. (1), k_{APM} can be positive or negative, depending on whether k_{SAW} is larger or smaller than k_G. In other words, the APM can be scattered in forward or reverse direction [3]. Backscattering appears to be more advantageous. Based on these considerations, an IDT period of 17.6 µm has been chosen for the design.

4. RESULTS AND DISCUSSION

APM grating devices

In order to demonstrate that the concept (i.e. the periodic grating device) is suitable for liquid sensing, devices originally designed to investigate the influence of mode conversion in surface acoustic wave filters [2], are used for ion detection in aqueous solutions. The devices were fabricated on 1.27 mm thick YZ-$LiNbO_3$ crystal plates with extraordinary parallel surfaces with a grating period about 20 µm. Each array consists of 200 grooves with a depth of 0.45 µm. Fig. 3 shows the response of the device to varying concentrations of KCl aqueous solution. The selected frequency of operation is 180 MHz. The frequency shifts and loss curves obtained are in good agreement with previous experiments using conventional APMs. As the SAW is confined to the lower crystal surface, no interaction between the SAW and the liquid medium occurs.

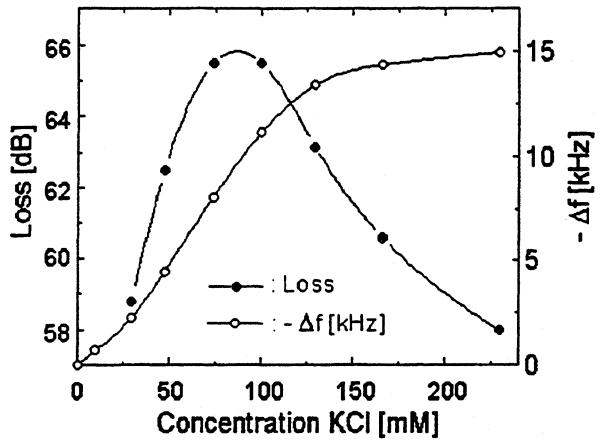


Fig. 3: *Loss and frequency change of a 180 MHz APM grating sensor on 1.27 mm YZ-$LiNbO_3$ as a function of aqueous KCl concentration. In the experimental set-up, the liquid contacts the electrode-free surface. The strong response towards acoustoelectric interactions confirms the successful mode conversion.*

In biosensing applications, dual delay line configurations with almost identical propagation paths are preferred in order to reduce spurious sensor response. Although the design parameters were nominally the same, the propagation characteristics of the available grating devices were too different to be used in such a setup. The observed differences are most likely due to slight variations in crystal thickness which change the coupling conditions at the gratings, as indicated by eq. (2). Thus, it appears that practical implementation of liquid-phase biosensors based on the above concept will be extremely difficult, as even slight variations in the thickness of the substrate drastically change the coupling conditions in the gratings. As a result, a second sensor design and geometry was investigated.

SAW grating devices

Fig. 4(a) shows the insertion loss of the device as a function of frequency. If nothing but the IDTs were at the electroded surface, a SAW with a nominal bandwidth of 10 MHz would be seen at a center frequency of about 195 MHz. Due to the presence of the gratings, narrow-bandwidth absorption bands (*stop bands*) occur in the SAW signal as indicated by the dotted line in Fig. 4(a). The absorber attenuates the SAW by 40–50 dB. The remaining narrow bandwidth transmission peaks (solid line) mark the frequencies at which efficient coupling to an APM occurs in the gratings. The modes are well separated, with an adjacent mode spacing of about 2 MHz. Thus, the new sensor design proves to be advantageous compared to traditional high-frequency APM devices, where gating

techniques or special transducer geometries [4] are often required to clearly resolve the individual modes.

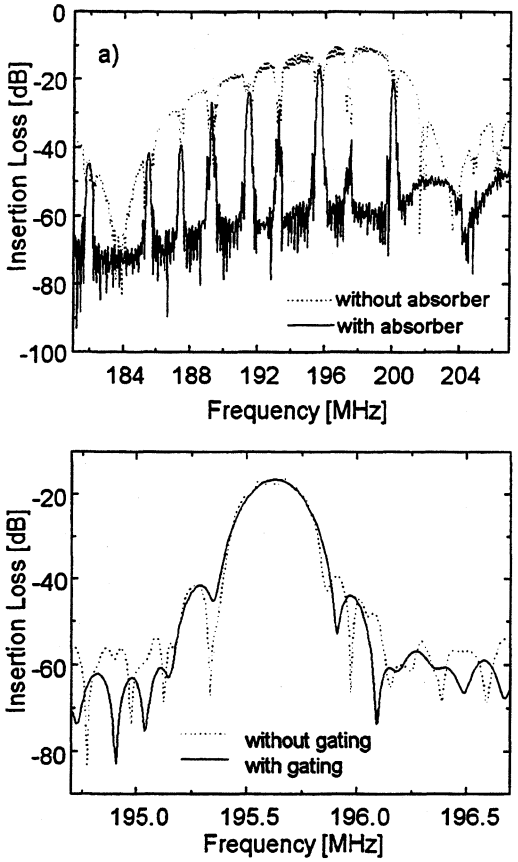


Fig. 4: *(a) Transmission curve of the grating sensor (b) Detailed view of the dominant mode at 195.6 MHz. Filtered and non-filtered transmission curves are almost identical.*

The SAW travelling along the direct path on the bottom of the device reaches the receiver after about 5 µs. The remaining time domain response is dominated by a signal observed in a time window from 6 to 12 µs, which relates to the converted modes. Thus, the mean travelling time for these modes from IDT to IDT is about 9 µs. The broad time window for the propagation of the converted modes is due to both the backscattering in the extended arrays and the multitude of the modes which have different phase velocities, while the enhanced mean travelling time is due to the longer path of propagation.

A more detailed view of the dominant mode is given in Fig. 4(b) indicating that suppression of electromagnetic crosstalk and neighboring modes is high enough to abstain from time-consuming filtering routines.

Mass Sensitivity of SAW Grating Device

Several SAW grating sensors were fabricated using the design parameters described above. In order to characterize the mass sensitivity of the devices, thin polyamic acid (PAA) polymer films are evaporated onto one of the two delay lines by chemical vapor deposition (CVD). During film growth, both the individual responses of the two lines and the difference frequency are monitored. The film thickness is determined using atomic force microscopy (AFM).

Fig. 5 shows the sensor response (frequency shift) during the deposition of about 29 nm of PAA on the electrode-free surface of one of the delay lines, i.e. the sensing line. In order to avoid any changes in the mode conversion characteristics, the film was exclusively deposited *in between* the two gratings on the sensing surface. As can be seen, the slope in the difference frequency is quite linear. As the response becomes non-linear for film thicknesses far in excess of 30 nm, only experiments with film thicknesses up to about 40 nm have been evaluated. The results are shown in Fig. 6, together with a linear fit which reveals a sensitivity of $c = (-1.75 \pm 0.16)$ kHz/nm.

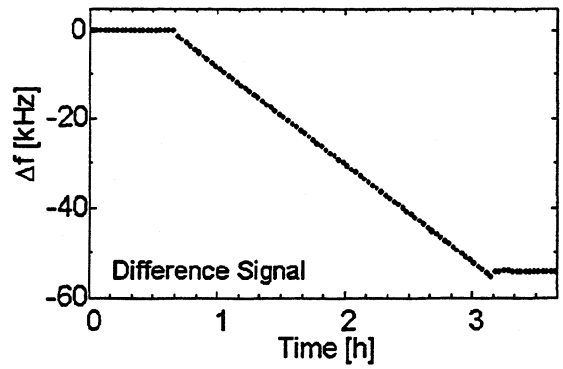


Fig. 5: *Frequency shifts of the grating sensor during deposition of about 29 nm of PAA. Deposition started after 40 min and continued for 150 min.*

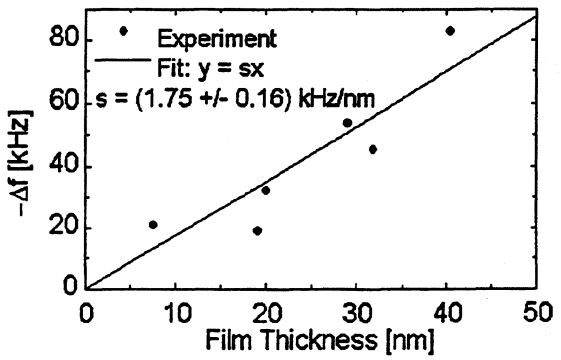


Fig. 6: *Frequency decrease, $-\Delta f$, of the grating sensor as a function of PAA film thickness, determined using AFM. The linear fit reveals an absolute sensitivity $c=(1.75\pm0.16)$ kHz/nm.*

In the case of pure mass loading, the frequency shift for the SAW on YZ-$LiNbO_3$ is expected to be given by $\Delta f = k f^2 h \rho$ where $k = -5.505 \cdot 10^{-7}$ cm^{2}s/kg, and h and ρ are thickness and density of the film, respectively. With f = 195.6 MHz and ρ = 1.42 g/cm^3 for PAA, the absolute mass

sensitivity is expected to be $\Delta f / h = -2.99$ kHz/nm. However, this value has to be corrected since only part of the SAW path is exposed to the coating. The center-to-center IDT spacing is 18 mm, corresponding to a delay time of 5 μs. Since the indirect path along the opposite surface corresponds to a mean propagation time of 9 μs, the effective sensitivity of the device is then expected to be c = (9.5/32.4)·(−2.99 kHz/nm) ≈ −0.88 kHz/nm, which is only about 50% of the experimental value.

One among several possible explanations of this discrepancy may be an underestimating of the film thickness, h. A second measurement of h taking into account small holes in the PAA film indicates a sensitivity of c=-1.0 kHz/nm. Also using a profilometer, a sensitivity of c = (−0.9 ± 0.4) kHz/nm has been found. Both of these results are close to the value expected from theory. Another possible explanation for the discrepancy may be due to contributions from viscoelastic properties of the film. Also, exposing the gratings to the coating is not as critical as exposing an IDT to the analyte.

5. GAS SENSING APPLICATIONS

In an example of gas sensing applications using this type of SAW sensor platform, the detection of nitrogen dioxide (NO_2) using copper phthalocyanine (CuPc) thin films was conducted.

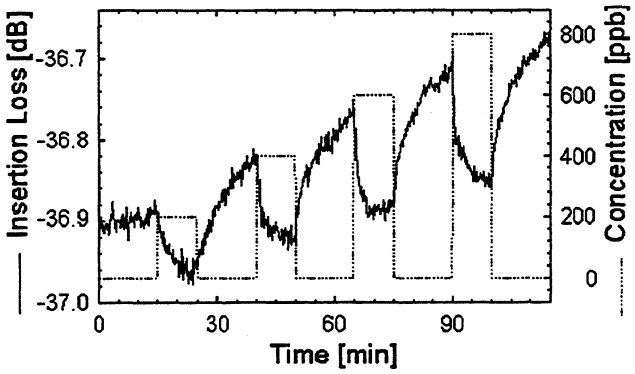


Fig. 7: *Insertion loss (—) and NO_2 gas concentration (···) as a function of time during gas sensing experiment. Only the response of the CuPc covered delay line is shown.*

In the present experiments, the sensing line of a dual delay line device has been coated with a CuPc film of 15 nm thickness. Again, the film was only deposited in between the gratings at the non-electroded surface. The device was then heated to 130 °C in order to shorten sorption time and thus reduce the response times. Fig. 7 shows the response of the CuPc covered delay line during gas exposure.

First, the measurement reveals that the device significantly responds to sub-ppm gas concentrations of NO_2. Second, the device is very sensitive to temperature variations. When the heating was turned on, the insertion loss of the device increased rapidly by more than 10 dB, but decreased slowly during the next hours. This drift can be seen in Fig. 7.

The observed phenomenon is most likely due to the fact that the temperature gradient imposes different coupling conditions on the two surfaces of the device, either by substrate thermal expansion or by a shift in the sound velocities. Thermal expansion of the substrate at the operating temperature is such that the grating period, p, has deviated from the initial design value of 16 μm. However, the observed drift and noise in the sensor signal may also be due in parts to the effect of temperature on the physical stability of the Pc film.

6. CONCLUSION

Mode conversion in periodic gratings can be used for the design of novel SAW sensor platforms to detect chemical analytes on electrode-free surfaces. Using thin mass layers, in CVD experiments, no evidence of interaction mechanisms other than mass loading has been found. The observed discrepancy between theory and measurements for the mass sensitivity is most likely due to an underestimation of the layer thickness. The transmission curve of the device is shown to be characterized by narrow bandwidth modes that are well separated in frequency. The excellent mode separation together with the moderate insertion loss are promising conditions for integration of the device into an oscillator circuit. While only a gas sensor experiment was demonstrated in the present work, this concept can also be utilized in the implementation of liquid-phase sensors in which a predominantly SH-SAW is utilized as the sensing wave.

7. ACKNOWLEDGMENTS

The authors wish to thank M. Grunze and J. Melngailis for helpful technical discussion.

8. REFERENCES

1. H. Wohltjen: Mechanism of Operation and Design Considerations for Surface Acoustic Wave Device Vapour Sensors, Sens. Actuators, **5** (1984) 307-325.
2. D.-P. Chen, J. Melngailis, and H.A. Haus: Filters Based on Conversion of Surface Acoustic Waves to Bulk Plate Modes in Gratings, IEEE Ultrason. Symp. Proc. (1982) 67-71.
3. J. Melngailis and R.C. Williamson: Interaction of Surface Waves and Bulk Waves in Gratings: Phase Shifts and Sharp Surface-Wave/Reflected Bulk Wave Resonances, IEEE Ultrason. Symp. Proc. (1978) 623-629.
4. R. Ros Seigel, P. Harder, R. Dahint, M. Grunze, F. Josse, M. Mrksich, and G. M. Whitesides: On-Line Detection of Nonspecific Protein Adsorption at Artificial Surfaces, Anal. Chem., **69** (1997) 3321-3328.

ACOUSTIC EMISSION SENSOR FOR IDENTIFICATION OF CHEMICAL REACTIONS

R. M. Lec, P.A. Lewin, M. Musavi*, S.W. Bang, S. Goel, S. Kwoun and E. Radulescu
The School of Biomedical Engineering, Science and Health Systems, and Department of Electrical and Computer Engineering, Drexel University, 3141 Chestnut Street, Philadelphia, PA 19104, USA
*Department of Electrical and Computer Engineering, University of Maine, Orono, Maine 04469, USA

Abstract

This paper describes a novel acoustic wave sensor designed for identification of chemical reactions. The principle of operation of the sensor is based on the acoustic emission (AE) phenomena. Chemical changes taking place during reactions are accompanied by energy transfer with some of the energy converted into acoustic waves. The AE waves frequency spectra exhibited unique patterns for the studied chemical reactions, and were used for their identification. A backpropagating neural network was used to implement the recognition of the AE chemical pattern. Potential applications of the AE chemical sensor include process control in the chemical, food, medical and pharmaceutical industries.

1. Introduction

This is a growing evidence that many chemical dynamic systems are acoustically active [1-7]. Chemical changes taking place during reactions are accompanied by energy transfer with some of the energy converted into acoustic waves. Some of these changes which produce AE include liquid-solid and solid-liquid transitions, dissolution, hydration and gelation [1]. Betterdge and co-workers [2] examined AE spectra from several different chemical systems. They found a correlation between chemical and acoustic events, although some acoustic responses were difficult to interpret. Sawada and co-workers [3] indicated that AE signals were influenced by changes of volume, heat balance, and reaction rate. Wentzel and Wade [4] analyzed the frequency spectra of AE signals from several chemical systems. They showed that the individual physical processes give different contributions to the AE response, e.g. gas release resulted in low-frequency signals, while crystal fracture produced higher frequency components. The purpose of the present work is to study correlation between the AE spectra and the type of the chemical reaction and to utilize it for the development of a prototype AEC sensor that could be used for identification of chemical reactions.

2. AEC Sensor System

A block diagram of the AEC sensor shown in Figure 1 consists of the following sections: the measurement cell with ultrasonic transducer (1); the signal conditioning unit (2); the time domain data acquisition unit which includes digital (3) and analog (4) oscilloscopes, a mixer (5), signal generator (6), audio amplifier (7) and a speaker (8); the frequency domain data acquisition with a spectrum analyzer (9); the pattern recognition unit with a neural network system (10) and a PC computer for overall control and data acquisition (11).

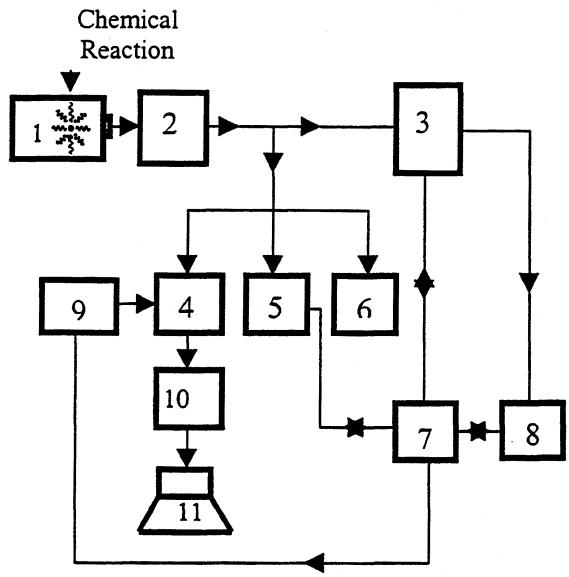


Figure 1. A block diagram of the acoustic emission chemical (AEC) sensor measurement system.

An AE signal generated by a chemical reaction propagates through the cell is received by an ultrasonic transducer and converted into an electrical signal. The received AE signals are amplified and analyzed in the frequency and time domains. Subsequently, the AE spectra are processed by a neural network pattern recognition unit. The operating frequency range of the AEC sensor is from 90 to 200 kHz and its sensitivity is of the order of picowatts of acoustic power.

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

3. AE Signals from Various Chemical Processes: Results and Discussion

Several chemical processes representing different physico-chemical mechanisms responsible for the generation of the AE signals have been studied. These mechanisms include microbubbles formation, phase transitions, gel formation, precipitation, and ion-ion exchange.

3.1 Microbubles Formation – Dissolution of Alka Seltzer.

Dissolution process of Alka Seltzer is strongly accompanied by the generation of the gas bubbles. Here, a solid form sodium bicarbonate $NaHCO_3$ is converted into sodium ion Na^+, hydroxide ion OH^- and carbon dioxide CO_2 gas:

$2CH_3CO_2C_6H_4CO_2H(s) + NaHCO_3(s) + H_2O \rightarrow$
$2CH_3CO_2C_6H_4CO_2H(s) + Na^+(aq) + OH^-(aq) + CO_2(g)\ I + H_2O$

Where the (s) marks a solid phase, the (aq) a liquid phase and the (g) a gas phase, respectively.

The AE frequency spectra taken from 90 to 200 kHz at different elapsed times t=0, 20 sec, and 1 and 2 minutes are shown in Figure 2. The ordinate and abscissa represent the amplitude and frequency of the signal respectively. The graph 2-a taken at t=0 represents the baseline, i.e. the measurement taken with distilled water present in the cell. At time t=0+ Alka Seltzer was added to water and the next three graphs represent the AE response to the reaction of Alka Seltzer with water. The AE signals are clearly present and their spectra evolve in time. The frequency spectrum of AE signals ranges from 90 to 180 kHz and exhibits three peak activities at 110, 155 and 170 kHz. The reaction is relatively rapid and continues with a slow decay of the AE amplitude of the AE spectra over next 30 minutes. At that time the AE activities are almost at the level of the baseline

3.2 Solid-Solid Transition in Liquid Phase – Hydration of Calcium Oxide

The hydration of calcium oxide CaO takes place when solid CaO is added to water. In this reaction, a phase transition from one solid form CaO to another solid form $Ca(OH)_2$ occurs:

$CaO(s) + 2\ H_2O \text{ ---> } Ca(OH)_2(s) + H_2O$ (2)

The AE frequency spectra taken at different time intervals t=0, 20 sec and 1 and 2 minutes are shown in Figure 3. Again, the graph 3-a taken at t=0 represents the baseline, i.e. the measurement taken with water only is identical with the one shown in figure 2a. At time t=0+ calcium oxide CaO was added to water and the next two graphs 3b,c,d represent the AE response to the reaction of CaO with water. Whereas, the AE signals are relatively weak and noisy, they clearly exhibit some AE activity in the frequency range from 90 to 110 kHz. The reaction appears to be slow and continues with little change in the AE spectra over next 60 minutes.

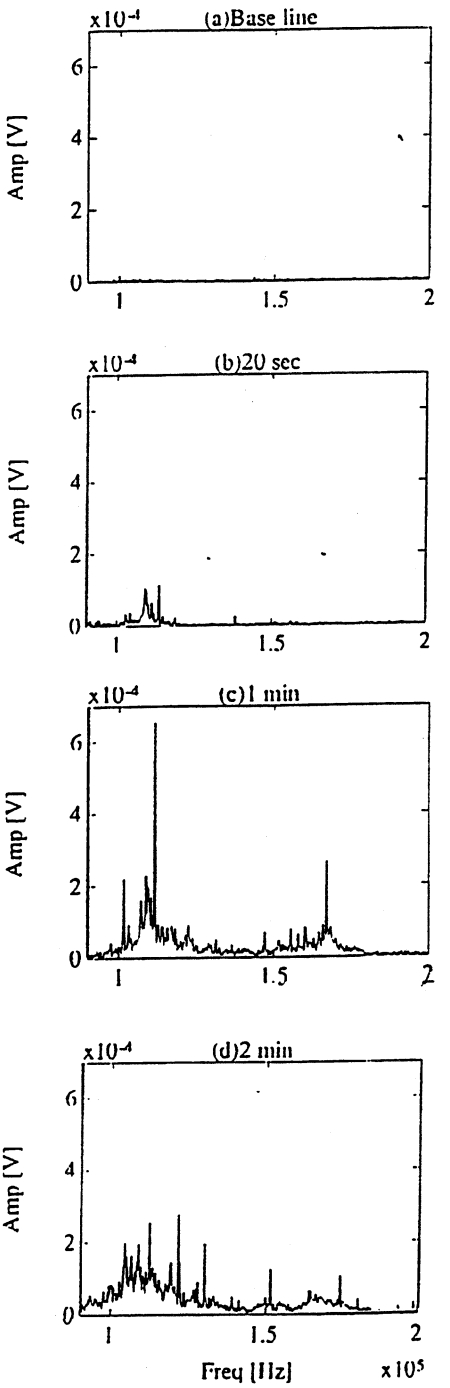


Figure 2. The AE frequency spectra representing a dissolution of Alka Seltzer taken at different elapsed times.

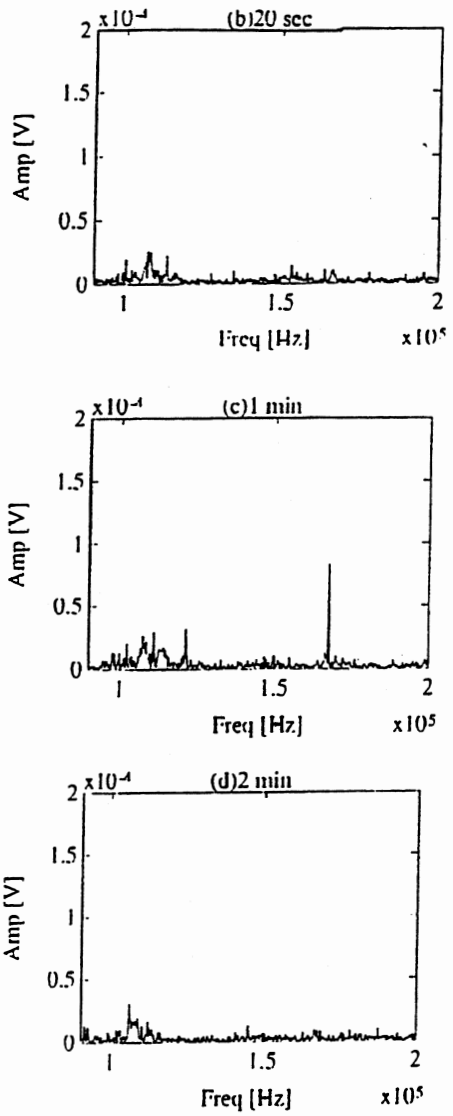


Figure 3. The AE frequency spectra representing a hydration of calcium oxide CaO taken at different elapsed times.

3.3 Phase Transition from Solid to Liquid Phase: Dissolution of sodium carbonate Na_2CO_3

In this reaction sodium carbonate undergoes dissolution from a solid to an ionic solution according to the equation:

$$Na_2CO_3(s) + H_2O \rightarrow$$
$$2Na^{+}(aq) + CO_3^{-2}(aq) + H_2O \qquad (3)$$

Figure 4 shows the AE frequency spectra from 90 to 200 kHz obtained by adding sodium carbonate to water. The AE frequency spectra were taken at different elapsed times t=0, 20 seconds, 1 and 2 minutes. The baseline taken from the cell filled with distilled water is similar to that shown in figure 2a. At time t=0+ sodium carbonate $Na_2CO_3(s)$ was added to water and the next three graphs represent the AE response to the reaction of $Na_2CO_3(s)$ with water. It is clear that the AE signals are relatively strong and the reaction proceeds rapidly, and is completed within two minutes. The maximum amplitude of AE signals were recorded just after adding sodium carbonate to water, i.e. in the elapsed time range from 20 to 30 seconds. The frequency content of AE signals is in the frequency range from 95 to 115 kHz.

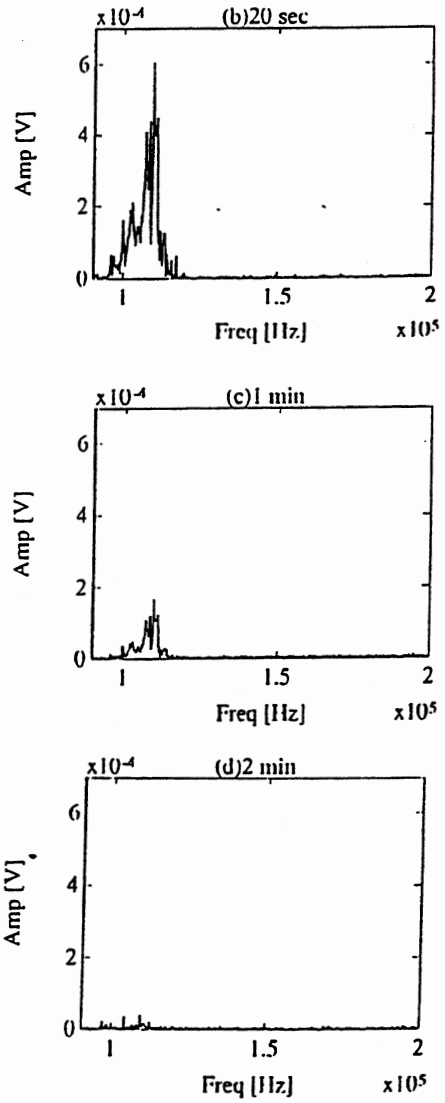


Figure 4. The AE frequency spectrum representing dissolution of solid calcium carbonate Na_2CO_3 taken at different elapsed times.

3.4 Ion-Ion Exchange Process: Causticization Reaction

To examine this reaction a completely dissolved sodium carbonate was added to a suspension of calcium hydroxide particles in water suspension. As a result, calcium carbonate was formed in a solid form according to the equation:

$$2Na^{+}(aq) + CO_3^{-2}(aq) + H_2O + Ca(OH)_2(s) \rightarrow CaCO_3(s) + 2Na^{+}(aq) + 2(OH)^{-}(aq) + H_2O \quad (4)$$

AE spectra characteristic for this reaction were taken at different times t=0, 20 second, 1 and 2 minutes, and are shown in Figure 5. At time t=0+ dissolved sodium carbonate $Na_2CO_3(aq)$ was added to the $Ca(OH)_2$ suspension and the next three graphs represent the AE response to this reaction. The baseline for this experiment was similar to that shown in figure 2a. Figure 5.b shows the AE spectrum taken at t=20 seconds after the prepared sodium carbonate solution was added to the calcium hydroxide suspension already in the testing cell. As a result, a strong AE peak appeared at 115 kHz. In Figure 5-c one can find that not only the amplitude of the first peak increased significantly, but also two additional peaks appeared at the frequency 150 and 170 kHz. The presence of these two new peaks can be interpreted as an appearance of an additional AE mechanism occurring during the causticization reaction.

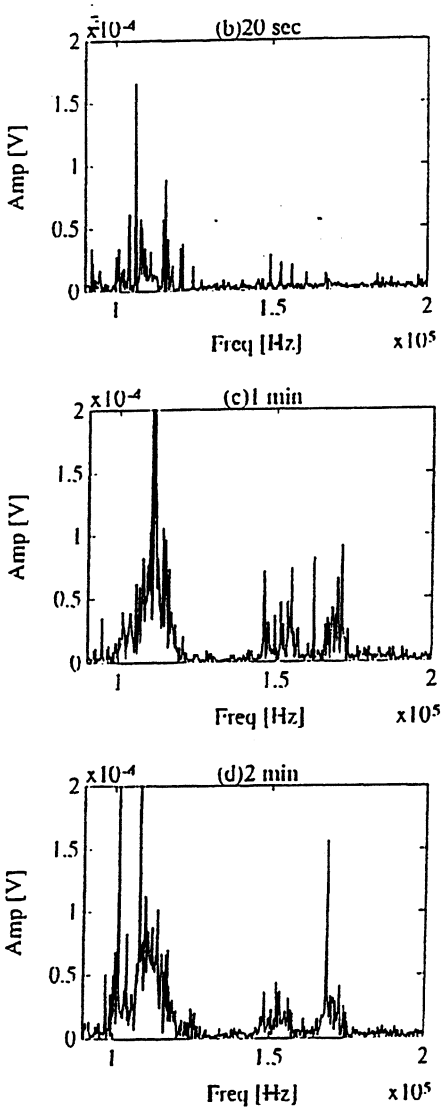


Figure 5. The AE frequency spectra representing the causticization reaction taken at different elapsed times.

4. Use of Neural Networks for Classification of AE Chemical Systems

As evidenced above the AE spectra of the chemical systems considered exhibit distinct signatures. This is of importance because the AE spectra can be used for identification of a specific reaction. Identification was carried out using artificial neural networks. Two types of neural networks, a backpropagating and radial basis function (RBF) were employed. The performance of the backpropagating neural network (BNN) was better then that based on the RBF and the results obtained for this network are presented here. The architecture of the BNN is shown in figure 6. It consists of the input and

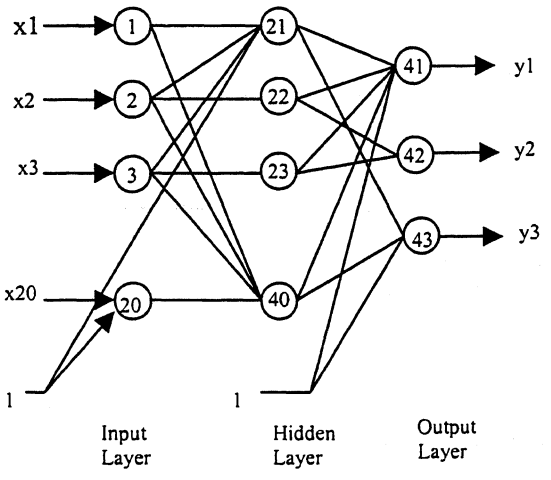


Figure 6. Backpropagating Neural Network architecture.

the output layer and one hidden layer. The number of nodes in the input and hidden layers was equal to 20 and the output layer included 3 nodes.

In Table 1 five various chemicals with different number of AE patterns available for training and testing the BNN are listed. In addition to the four used in the above-discussed chemical reactions, the data for dissolution of aluminum chloride are also included.

Table 1: AE Patterns Available for Different Chemical Systems.

Chemical Systems	Class	Number of AE Patterns
Alka Seltzer	0	17
Aluminum Chloride	1	8
Calcium Oxide	2	8
Sodium Carbonate	3	20
Caustization Reaction	4	17

For the BNN training, twenty spectral lines equally spaced at 5.5 kHz and selected over the frequency range 90-200 kHz were taken for each AE chemical pattern and subsequently were applied to the input layer of the BNN. The output of the BNN having three nodes code represented five classes of the studied chemicals by assigning 000,001, etc. code for the each chemical.

The BNN was trained and tested using the different sets of training and testing patterns. The total number of the AE patterns used was 70. For example, the set No.1 as it is shown in Table 2 included 36 AE patterns for training and 34 for testing.

Table 2: Testing Results of Backpropagation Network trained by given AE patterns sets.

Set No.	No. of Mis-classifications	No. of Testing patterns	Error %
1.	24	34	70
2.	19	28	67
3.	10	22	45
4.	6	20	30
5.	8	70	11

The results presented in Table 2 show that the BNN was able to perform classification of the studied chemical systems using the AE spectra. The classification error decreases with the number of AE patterns used for the training.

5. Summary, Conclusions and Future Work

An acoustic emission chemical (AEC) sensor able to classify chemical reactions was designed, built and tested. The acoustic emission (AE) spectra were measured for five different reactions representing four different mechanisms of chemical processes. The AE frequency spectra and their time evolution were characteristic for each reaction indicating that each of these reactions had its own AE signature. These features of the AE spectra were used subsequently for the identification of chemical reactions using the backpropagating artificial neural network. Also, it was demonstrated that the AEC sensor was capable to monitor the kinetics of these reactions. Future work will be focused on the measurements of AE spectra over wider frequency range (up to a few MHz) and on using a wavelet based patterns recognition approach to provide more efficient identification procedures.

References

1. A.P. Wade, D.B. Sibbald, M. Bailey, R.M. Belchamber, S. Bittman, J. A. McLean and P.D. Wentzel, Anal. Chem., vol. 63, p.497A, (1991)
2. D. Betteridge, M.T. Joslin and T. Lilley, Anal. Chem. vol. 53, p.1064,(1981)
3. T. Sawada, Y. Goshi, C. Abe and K. Furuya, Anal. Chem. , vol.57, p. 1743, (1985)
4. P.D. Wentzel and A.P. Wade, Anal. Chem., vol.61, p. 2638, (1989)
5. P.D. Wentzel, S.J. Vanslyke and K.P. Bateman, Anal. Chem. Acta, vol. 246, p.43 (1991)
6. T. Sawada and Y. Goshi, Anal. Chem. vol. ,57, p.p.366 (1985)
7. R.M. Belchamber, D. Betteridge, M.P. Collins, T. Lilley, C.Z. Marczewski and A.P. Wade, Anal. Chem., vol. 58, p. 1877, (1986)
8. R. Lec, M.Musavi, H. Pendse and W. Ahmed, ,Smart Structures and Materials '93, SPIE, vol. 1918, pp. 441-448, (1993),

EFFECTS OF A LIQUID LAYER ON THICKNESS-SHEAR VIBRATIONS OF RECTANGULAR AT-CUT QUARTZ PLATES

P. C. Y. Lee and R. Huang
Department of Civil Engineering & Operations Research
Princeton University, Princeton, N.J. 08544

Abstract

Thickness-shear vibrations of a rectangular AT-cut quartz plate with one face in contact with a layer of Newtonian (linearly viscous and compressible) fluid are studied. The governing equations for vibrations of piezoelectric crystal plates given in a previous article by Lee, Yu and Lin are employed in the present study. One-dimensional solutions for the normal and shearing stresses at the bottom of a liquid layer are used as approximations to the stresses of the liquid layer exerting on the crystal surface in the plate equations.

Closed form solutions are obtained for both free and piezoelectrically forced thickness-shear vibrations of a finite AT-cut quartz plate in contact with a liquid layer. From the present solutions, a simple and explicit formula is deduced which includes the effect of compressional wave in the liquid layer and that of the thickness-to-length ratio of the plate. The formula reduces to the well-known frequency equation obtained by many previous investigators for infinite plates. The resonance frequency of a rectangular AT-cut quartz resonator, computed as a function of the thickness of the liquid layer, agrees closely with the experimental data measured by Schneider and Martin.

1 Introduction

Thickness-shear vibrations of quartz plates of infinite extent in contact with a liquid layer have been studied both analytically and experimentally by many investigators for a long time[1-4]. More recently, the influence of compressional wave in liquid layer due to the non-uniform thickness-shear motion in a finite quartz plate have been considered [5-7].

In the present article, the non-uniform thickness-shear vibrations of a finite, rectangular, AT-cut quartz plate in contact with a liquid layer is studied.

The recently derived 2D equations for the thickness-shear and flexural vibrations[8] are used for the AT-cut quartz plate. The normal and tangential stresses in the liquid layer subject to uniform oscillating boundary conditions are obtained from simplified governing equations for the Newtonian fluid by assuming small perturbation in fluid density and pressure[9]. By requiring the continuity of displacements and stresses at the interface of the crystal plate and the liquid layer, the face tractions of the plate are expressed in terms of boundary stresses of the liquid layer. The resulting equations are further simplified by Tiersten's "thickness-shear approximation"[10]. Closed form solutions are obtained for both free and piezoelectrically forced thickness-shear vibrations and the effects of the liqud layer are discussed.

2 Two-dimensional First-order Equations

Consider a rectangular AT-cut quartz plate with thickness $2b$, length $2a$ and width $2c$ ($c \gg b$) as shown in Figure 1. The surfaces of the plate at $x_2 = \pm b$ are fully covered by electrodes of thickness $2b'$. A liquid layer of thickness h_L is in contact with the upper surface of the plate at $x_2 = b + 2b'$.

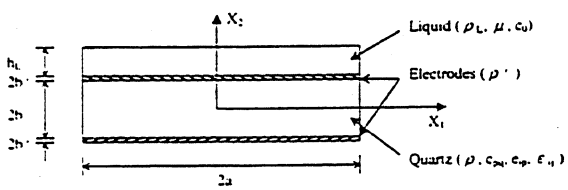


Figure 1. An AT-cut quartz plate with electrodes and in contact with a liquid layer

A set of two-dimensional first-order equations for vibrations of piezoelectric crystal plates with electroded faces was derived in [8]. For AT-cut quartz plates, the equations for thickness-shear and flexural vibrations

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

are reduced as follows:

$$
\begin{aligned}
&u_1(x_1,x_2,t) = -u_{2,1}^{(0)}x_2 + u_1^{(1)}\cos\left[\frac{\pi}{2}(1-\frac{x_2}{b})\right] ,\\
&u_2(x_1,x_2,t) = u_2^{(0)} ,\\
&\phi(x_1,x_2,t) = \bar{V}_0 + \bar{V}_1\frac{x_2}{b} + \phi^{(2)}\sin(\pi\frac{x_2}{b}) ,
\end{aligned}
\tag{1}
$$

and

$$
\begin{aligned}
&2c_{66}u_{1,1}^{(1)} + \mathcal{F}_2^{(0)} = 2(1+R)\rho b\ddot{u}_2^{(0)} ,\\
&c_{11}^{(1)}u_{1,11}^{(1)} - \left(\frac{\pi}{2b}\right)^2 c_{66}u_1^{(1)} - \frac{8b}{\pi^2}c_{11}^{(1)}u_{2,111}^{(0)} + \frac{8}{3\pi}e_{11}^{(1)}\phi_{,11}^{(2)}\\
&\quad -\frac{2\pi}{3b^2}e_{26}\phi^{(2)} - \frac{2}{b^2}e_{26}\bar{V}_1 + \frac{1}{b}\mathcal{F}_1^{(1)}\\
&\quad = (1+2R)\rho\ddot{u}_1^{(1)} - \frac{8}{\pi^2}\rho b\ddot{u}_{2,1}^{(0)} ,\\
&-\frac{2b}{\pi}e_{11}^{(1)}u_{2,111}^{(0)} + \frac{8}{3\pi}e_{11}^{(1)}u_{1,11}^{(1)} - \frac{2\pi}{3b^2}e_{26}u_1^{(1)} - \epsilon_{11}^{(1)}\phi_{,11}^{(2)}\\
&\quad + \left(\frac{\pi}{b}\right)^2\epsilon_{22}\phi^{(2)} = 0 ,
\end{aligned}
\tag{2}
$$

where

$$
\begin{aligned}
&\mathcal{F}_j^{(n)} = T_{2j}(x_2 = b+2b') - (-1)^n T_{2j}(x_2 = -b-2b') ,\\
&\bar{V}_1 = \frac{1}{2}[\phi(x_2=b) - \phi(x_2=-b)] ,\\
&\bar{V}_0 = \frac{1}{2}[\phi(x_2=b) + \phi(x_2=-b)]
\end{aligned}
\tag{3}
$$

and

$$
\begin{aligned}
&c_{pq}^{(1)} = c_{pq} - \frac{c_{p2}c_{2q}}{c_{22}} , \quad e_{ip}^{(1)} = e_{ip} - \frac{c_{p2}e_{i2}}{c_{22}} ,\\
&\epsilon_{ij}^{(1)} = \epsilon_{ij} + \frac{e_{i2}e_{j2}}{c_{22}} , \quad R = \frac{2\rho' b'}{\rho b} .
\end{aligned}
\tag{4}
$$

3 Effects of the Liquid Layer

For plates in contact with a liquid layer, the coupled thickness-shear and flexural vibrations will generate both shear and compressional waves in the liquid layer. By assuming small disturbances in the liquid density and pressure and neglecting the convective part of acceleration, the governing equations of linear compressible viscous liquids are reduced to linear displacement equations of motion as follows:

$$
\rho_L\ddot{\mathbf{u}} = \rho_L c_0^2\nabla(\nabla\cdot\mathbf{u}) + (\mu+\lambda)\nabla(\nabla\cdot\dot{\mathbf{u}}) + \mu\nabla^2\dot{\mathbf{u}} + \rho_L\mathbf{g} ,
\tag{5}
$$

where ρ_L is the density of the liquid, c_0 the sound velocity in the liquid, μ,λ the first and second coefficients of viscosity and $\mathbf{g}$ the vector acceleration of gravitiy or other kinds of body forces.

A transverse shear wave in the liquid layer generated by a horizontally oscillating plane boundary at bottom is considered and the stress at the bottom of the liquid layer is obtained as

$$
\tilde{T}_{21} = (1-i)\omega^{3/2}\sqrt{\frac{\rho_L\mu}{2}}\tanh[(1+i)\kappa_s h_L]\bar{u}_1 e^{i\omega t} ,
\tag{6}
$$

where ω is the oscillation frequency, $\bar{u}_1$ is the amplitude of displacement in x_1 direction at the bottom of the liquid layer, and κ_s is the attenuation coefficient of the shear wave, which is defined as

$$
\kappa_s = \sqrt{\frac{\rho_L\omega}{2\mu}} .
\tag{7}
$$

where ω is the oscillation frequency.

Similarly, a compressional wave generated by a plane boundary oscillating in x_2 direction is also considered. For a liquid layer with thickness h_L and traction-free top surface, the normal stress at the bottom of the liquid layer is

$$
\tilde{T}_{22} = \frac{\rho_L\omega^2}{\zeta_c}\tan(\zeta_c h_L)\bar{u}_2 e^{i\omega t} ,
\tag{8}
$$

where $\bar{u}_2$ is the amplitude of displacement in x_2 direction at the bottom and

$$
\zeta_c = \frac{\omega}{c_0}\left[1 - i\omega\frac{2\mu+\lambda}{2\rho_L c_0^2}\right] .
\tag{9}
$$

Alternatively, if the liquid layer has a rigid plane boundary at top surface, the normal stress at the bottom becomes

$$
\tilde{T}_{22} = -\frac{\rho_0\omega^2}{\zeta_c}\mathrm{ctan}(\zeta_c h_L)\bar{u}_2 e^{i\omega t} .
\tag{10}
$$

Equations (6) and (8) or (10) can be used as approximations of the stresses applied to the vibrating crystal plate due to the liquid layer, although the shear wave and compressional wave in the liquid layer are actually coupled. By requiring continuity in displacements and stresses at the interface of liquid and crystal, we obtain

$$
\begin{aligned}
&\mathcal{F}_2^{(0)} = \mathcal{L}_c\rho_L h_L\omega^2 u_2^{(0)} ,\\
&\mathcal{F}_1^{(1)} = (1-i)\tanh[(1+i)\kappa_s h_L)\omega^{3/2}\sqrt{\frac{\rho_L\mu}{2}}(u_1^{(1)} - bu_{2,1}^{(0)}) ,
\end{aligned}
\tag{11}
$$

where

$$
\begin{aligned}
&\mathcal{L}_c = \frac{\tan(\zeta_c h_L)}{\zeta_c h_L} \quad \text{for traction free top surface} ,\\
&\mathcal{L}_c = -\frac{\mathrm{ctan}(\zeta_c h_L)}{\zeta_c h_L} \quad \text{for rigid top surface} .
\end{aligned}
\tag{12}
$$

The effects of the liquid layer on vibrations of the crystal plate are thus taken into account by substituting (11) into the plate equations (2).

4 Thickness-shear Approximations

For a plate of infinite extent in x_1 and x_3 directions, the plate equations in (2) can be reduced to one single equation for simple thickness-shear vibrations of AT-cut quartz plates:

$$\left[1+\frac{16}{9\pi^2}\frac{e_{26}^2}{c_{66}\epsilon_{22}}-(1-i)\mathcal{L}_s\left(\frac{\omega}{\omega_1}\right)^{3/2}\right]\left(\frac{\pi}{2b}\right)^2 c_{66}u_1^{(1)} +(1+2R)\rho\ddot{u}_1^{(1)}+\frac{2}{b^2}e_{26}\bar{V}_1=0\,, \tag{13}$$

where

$$\mathcal{L}_s=\frac{2}{\pi}\sqrt{\frac{\rho_L\mu\omega_1}{2\rho c_{66}}}\tanh[(1+i)\kappa_s h_L]\,,\quad \omega_1=\frac{\pi}{2b}\sqrt{\frac{c_{66}}{\rho}}\,. \tag{14}$$

A complex resonance frequency is obtained from (13):

$$\Omega=\frac{\omega}{\omega_1}=1-R+\frac{8}{9\pi^2}\frac{e_{26}^2}{c_{66}\epsilon_{22}}-\frac{1}{2}\mathcal{L}_s(1-i)\,. \tag{15}$$

For plates of finite extent in x_1 direction, following Tiersten's idea of thickness-shear approximation[10], we eliminate $u_2^{(0)}$ from plate equations and reduce three equations in (2) to two equations:

$$\begin{aligned}&\left[c_{11}^{(1)}+(1-i)\mathcal{L}_s c_{66}\right]u_{1,11}^{(1)}-\left(\frac{\pi}{2b}\right)^2 c_{66}u_1^{(1)}[1-(1-i)\mathcal{L}_s]\\ &+\frac{8}{3\pi}e_{11}^{(1)}\phi_{,11}^{(2)}-\frac{2\pi}{3b^2}e_{26}\phi^{(2)}-\frac{2}{b^2}e_{26}\bar{V}_1\\ &=(1+2R)\rho\ddot{u}_1^{(1)}+\frac{32\rho b^2}{\pi^4(1+R)}\ddot{u}_{1,11}^{(1)}\left(1-\frac{\mathcal{L}_c R_L}{1+R}\right)\,,\\ &\frac{8}{3\pi}e_{11}^{(1)}u_{1,11}^{(1)}-\frac{2\pi}{3b^2}e_{26}u_1^{(1)}-\epsilon_{11}^{(1)}\phi_{,11}^{(2)}\\ &+\left(\frac{\pi}{b}\right)^2\epsilon_{22}\phi^{(2)}=0\,,\end{aligned} \tag{16}$$

where

$$R_L=\frac{\rho_L h_L}{2\rho b}\,. \tag{17}$$

5 Free Vibrations

By setting $\bar{V}_1$ to zero in (16) and solving for free vibrations with shorted electrodes, we obtain a dispersion relation, which only consists of two branches corresponding to $u_1^{(1)}$ and $\phi^{(2)}$ respectively while the other branch corresponding to $u_2^{(0)}$ is neglected for predominant thickness-shear vibrations.

For plates with fixed displacement and charge-free edges, we require

$$u_1^{(1)}=\bar{D}_1^{(2)}=0\,,\quad \text{at } x_1=\pm a\,. \tag{18}$$

By substituting solutions of $u_1^{(1)}$ and $\phi^{(2)}$ into (18), a frequency equaiton in transcendental form is obtained. By further approximations, an explicit formula for the resonance frequency of fundamental thickness shear mode is obtained as follows:

$$\begin{aligned}\Omega=1-R+&\frac{8}{9\pi^2}\frac{e_{26}^2}{c_{66}\epsilon_{22}}-\frac{1}{2}(1-i)\mathcal{L}_s\\ &+\left(\frac{4}{\pi^2}+\frac{1}{2}\frac{c_{11}^{(1)}}{c_{66}}+\frac{16}{9\pi^2}\frac{e_{26}e_{11}^{(1)}}{c_{66}\epsilon_{22}}\right)\frac{b^2}{a^2}\\ &-\frac{4}{\pi^2}\frac{b^2}{a^2}\frac{\mathcal{L}_c R_L}{1+\frac{8}{\pi^2}\frac{b^2}{a^2}\mathcal{L}_c R_L}\,.\end{aligned} \tag{19}$$

(19) includes the effect of the thickness-to-length ratio of the plate and that of compressional wave in the liquid layer, which is also dependent upon the thickness-to-length ratio of the plate. For plates with infinte length, $a\to\infty$, (19) reduces to (15).

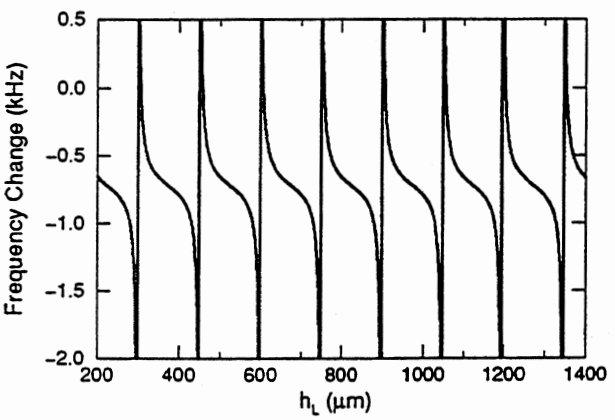


Figure 2. Predicted changes in thickness-shear frequency vs liquid layer thickness h_L with $2a=0.4in$, $2b=0.33mm$, $\rho_L=998.2kg/m^3$, $\mu=1.002mPa\cdot s$, $c_0=1500m/s$, and rigid top surface.

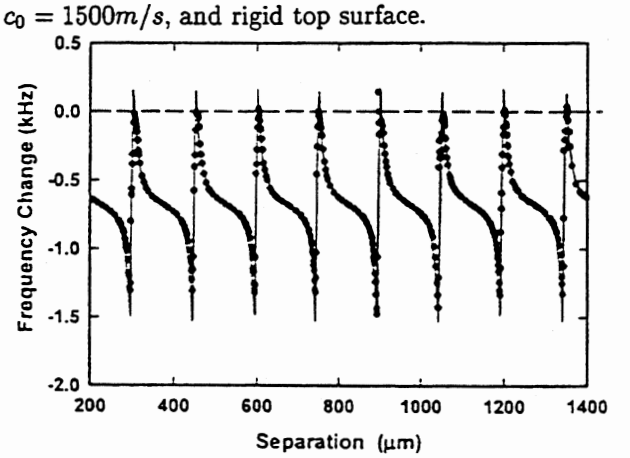


Figure 3. Change in resonance frequency vs the liquid layer thickness. Points are measured by Schneider and Martin[6].

We see in (19) that the effect of liquid layer thickness h_L on the resonance frequency is contained in the factor $\mathcal{L}_c R_L$ which are defined in (12) and (17). The frequency change of the fundamental thickness shear mode is computed for a rectangular AT-cut of quartz with $2a = 0.4in$, $2b = 0.33mm$ and in contact with a water layer ($\rho_L = 998.2kg/m^3$, $\mu = 1.002mPa \cdot s$, $c_0 \approx 1500m/s$) with rigid top surface. The predicted result is shown in Figure 2, which can be compared with the experimental result for a circular disk of AT-cut quartz measured by Schneider and Martin[6] as shown in Figure 3. The agreement is very close.

6 Piezoelectrically Forced Vibrations

The complete solution of (16) for piezoelectrically forced vibrations, i.e., for $\bar{V}_1 = \phi_0 e^{i\omega t}$ is also obtained. Motional capacitance and admittance are calculated as functions of forcing frequency. Distributions of displacement, electric potential and surface charge at various frequencies are examined.

Due to space limitation, detailed results are not included in the present article.

7 Conclusions

Non-uniform thickness-shear vibrations of a finite, rectangular AT-cut quartz plate in contact with a liquid layer are studied. Solutions are obtained for free and piezoelectrically forced vibrations with fixed displacement and charge-free edge conditions.

A simple and explicit formula for thickness-shear resonance frequency is deduced, which includes the effect of compressoinal wave in the liquid layer and the effect of the thickness-to-length ratio of the plate. The computed resonance frequency as a function of the thickness of the liquid layer agrees very closely with the experimental data by Schneider and Martin[6] although they were measured for circular disks of AT-cut quartz.

Acknowledgment

This work was supported by Grant No. DAAH 04-95-0102 from the U.S. Army Research Office.

References

[1] K. K. Kanazawa and J. G. Gordon II, "The Oscillation Frequency of A Quartz Resonator in Contact with a Liquid", *Anal. Chim. Acta*, vol. 175, pp. 99-105, 1985.

[2] C. E. Reed, K. K. Kanazawa, and J. H. Kaufman, "Physical description of a viscoelastically loaded AT-cut quartz resonator", *J. Appl. Phys.*, **68**(5), pp. 1993-2001, 1990.

[3] F. Josse, Z. A. Shana, D. E. Radtke and D. T. Haworth, "Analysis of Piezoelectric Bulk-Acoustic-Wave Resonators as Detectors in Viscous Conductive Liquids", *IEEE Trans. Ultrasonics, Ferroelectrics and Frequecny Control*, **37**(5), pp. 359-368, 1990.

[4] M. Thompson and G. Hayward, "Mass Response of the Thickness-shear Mode Acoustic Wave Sensor in Liquids as a Central Misleading Dogma", *Proc. 1997 IEEE International Frequency Control Symposium*, pp. 114-119, 1997.

[5] Z. Lin and M. D. Ward, "The Role of Longitudinal Waves in Quartz Crystal Microbalance Applications in Liquids", *Anal. Chem.*, vol. 67, pp. 685-693, 1995.

[6] T. W. Schneider and S. J. Martin, "Influence of compressional wave generation on thickness-shear mode resonator response in a fluid", *Anal. Chem.*, vol. 67, pp. 3324-3335, 1995.

[7] R. Thalhammer, S. Braun, B. Devcic-Kuhar, M. Groschl, F. Trampler, E. Benes, H. Nowotny and M. Kostal, "Viscosity Sensor Utilizing a Piezoelectric Thickness Shear Sandwich Resonator", *IEEE Trans. Ultrasonics, Ferroelectrics and Frequecny Control*, **45**(5), pp. 1331-1340, 1998.

[8] P. C. Y. Lee, J. D. Yu and W. S. Lin, "A new two-dimensional theory for vibrations of piezoelectric crystal plates with electroded faces", *J. Appl. Phys.*, **83**(3), pp. 1213-1223, 1998.

[9] R. B. Lindsay, *Mechanical Radiation*, McGraw-Hill, Inc., New York, 1960.

[10] H. F. Tiersten, *Linear Piezoelectric Plate Vibrations*, Plenum Press, New York, 1969.

1999 Joint Meeting EFTF - IEEE IFCS

SIGNAL AMPLIFICATION WITH MULTILAYER ARRANGEMENTS ON CHEMICAL QUARTZ-CRYSTAL-RESONATOR-SENSORS

Ralf Lucklum, C. Behling, and P. Hauptmann
Otto-von-Guericke-University, Institute for Measurement Technology and Electronics (IPE)
D-39016 Magdeburg, Germany
phone: 49 (391) 671 8308, Fax: 49 (391) 671 2609, e-mail: Ralf.Lucklum@E-Technik.Uni-Magdeburg.DE

ABSTRACT

Viscoelastic properties of chemically sensitive coatings can amplify the mass sensitivity of quartz-crystal-microbalance-(QCM)-sensors. We developed a sensor concept which is based on a double layer arrangement. Acoustic amplification and chemical sensitivity are separated. With a proper selection of materials the chemical sensitive layer acts pure gravimetrically while the first layer realizes a constant acoustic amplification of the mass effect. Major sensor design parameter are the shear modulus and the thickness of the first layer. They are sufficiently independent from the design of the chemically active layer. Parallel measurement of the resonant frequency of the quartz crystal and its motional resistance in the equivalent circuit can be exploited to check the acoustic amplification.

1. INTRODUCTION

Recently it has been shown that sensitive films used as chemical interfaces for acoustic-wave-based devices may undergo significant deformation across the film thickness, accompanied by a significant phase shift of the acoustic wave across the film [1]. This effect leads to a dependence of the sensor response not only on surface mass changes. Material parameters of the coating also gain influence on the electrical impedance/admittance of the sensor. The latter is responsible for viscoelastic contributions to the frequency shift [2] and gives way to an acoustic amplification of the mass effect [3]. The modulus generated frequency shift cannot be separated from the mass generated frequency shift with a simple frequency measurement, hence the amplification factor is usually unknown. The positive effect of sensitivity enhancement and the problematical effect of an unknown sensor transfer function are the two sides of one medal. If coating thickness or accumulated mass are to be calculated, the sensor must be calibrated with separate experiments.

In the following we analyze a sensor concept based on a double layer arrangement where acoustic amplification and chemical sensitivity are separated. We will concentrate on quartz crystal resonators although the physical background has general validity.

2. MODEL

The theoretical analysis is based on a one-dimensional model of wave propagation in the system quartz-coating-analyte. To calculate the overall systems characteristics, the chain matrix technique is used. Nonpiezoelectric layers are modeled with a transmission line with two acoustic ports, the quartz crystal additionally with one electrical port to include the piezoelectricity of the quartz. With this technique, the coating can be either a single layer or a multilayer arrangement. A gaseous analyte acts acoustically as a short cut at the front acoustic port, a liquid analyte (Newtonian liquid) as a viscose load. The central role in the acoustic analysis of the coating properties (incl. mass changes) plays the (surface) acoustic load at the interface between the quartz and the coating. This acoustic load carries all information of the system coating-medium. With some approximations the imaginary part of the acoustic load, $\mathrm{Im}(Z_L)$, is proportional to the negative frequency shift, Δf, while the real part of the acoustic load, $\mathrm{Re}(Z_L)$, is proportional to the resistance change, ΔR, [4,5]:

$$\Delta f \propto -\mathrm{Im}(Z_L), \qquad \Delta R \propto \mathrm{Re}(Z_L) \qquad (1).$$

We therefore analyze the acoustic load first. In the case of a pure thickness-shear-mode resonator, as it is with an AT-cut quartz crystal, the acoustic load, Z_L, is:

$$Z_L = Z_{c1}\frac{Z_2 + \mathrm{j}Z_{c1}\tan\left(\omega(\rho_1/Z_{c1})h_1\right)}{Z_{c1} + \mathrm{j}Z_2\tan\left(\omega(\rho_1/Z_{c1})h_1\right)} \qquad (2).$$

$Z_{c1} = \sqrt{\rho_1 G_1}$ is the characteristic acoustic impedance of the first layer adjacent to the acoustic device, ρ_1, h_1 and $G_1 = G'_1 + \mathrm{j}G''_1$ are the density, the thickness and complex shear modulus of the first layer, respectively. G' is the real part, G'' the imaginary part of G. Z_2 represents the acoustic load acting on top of the first coating. The load may be generated from a second film, a multilayer or a semi-infinite material. In the case of a single layer with $Z_2 = 0$ eq 2 simplifies to:

0-7803-5400-1/99/ $ 10.00

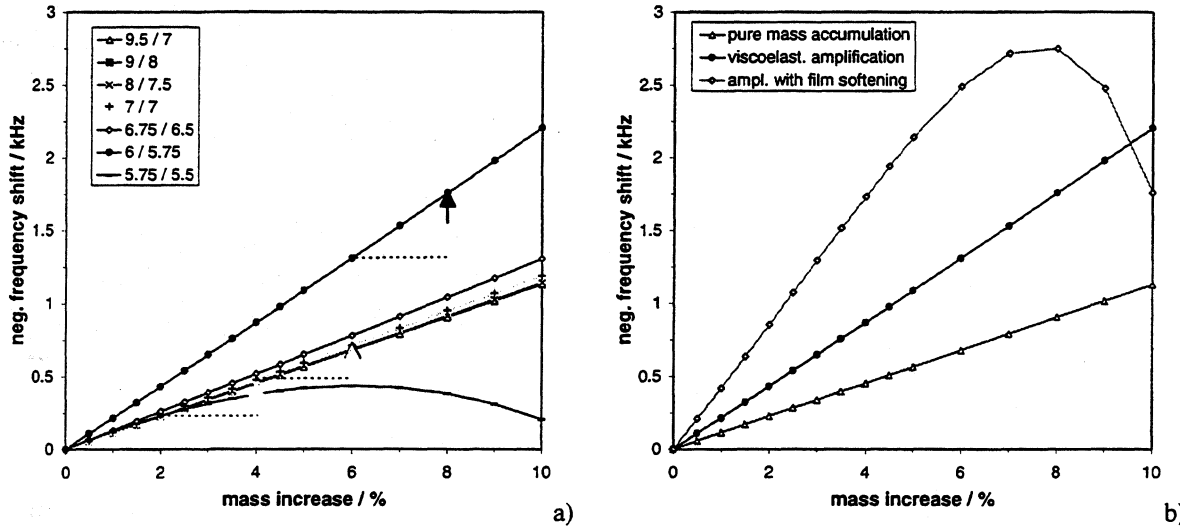


Fig. 1: Frequency shift during analyte sorption without and with film softening ($f_0 = 10$ MHz, initial $h_1 = 500$ nm).

$$Z_{L1} = jZ_{c1}\tan\left(\omega(\rho/Z_{c1})h_1\right) = jZ_{c1}\tan(\varphi_1) \quad (3a)$$

which can be rewritten as follows:

$$Z_{L1} = j(\omega\rho_1 h_1)\frac{\tan(\varphi_1)}{\varphi_1} = jM_1 * V_1 \quad (3b).$$

This notation [4] makes the effect of the acoustic amplification, V, of the mass effect, M, most obvious. For thin rigid layers the acoustic phase shift within the coating, φ_1, is very small, hence $\tan(\varphi_1) \approx \varphi_1$ and $V_1 = 1$. This is the gravimetric regime, where the acoustic device acts as a pure microbalance, $\Delta f \propto -\mathrm{Im}(Z_{L1}) = -\omega\rho_1 h_1$, and where Sauerbrey's equation [6] is valid.

With increasing phase shift V increases and amplifies the mass effect on the acoustic load and hence the mass sensitivity of the acoustic device. As long as the shear modulus does not change during measurement and as long as the mass increase (per area) is small compared to $\rho_1 h_1$, the change in V is small compared to the experimental accuracy, except near acoustic film resonance. V acts as an almost constant amplification factor. This is demonstrated in Fig. 1a for a few examples. The legend in the diagram is noted as log(*G'*/Pa)/log(*G"*/Pa), the slash represents a separator, not a quotient. The first three examples represent a polymer in the glassy state, the last three examples a polymer in the rubbery state. The "7/7"-example (*G'*=10 MPa, *G"*=10 MPa) may represent a polymer in the transition range. The gray arrows symbolize the frequency shift due to mass increase, the black arrows the viscoelastic contribution. As expected, glassy polymers do not exhibit significant viscoelastic contributions. By contrast, rubbery polymers exhibit a significant viscoelastic contribution. The curves are almost linear (except the "5.75/5.5"-example), the slope of $-\Delta f$ *vs.* $\Delta m/m$ is much higher than that of glassy films.

If analyte sorption is accompanied by a swelling induced analyte softening [2] the modulus change generates its own, unknown contribution to the frequency shift, Fig. 1b. In our example we assume a slight analyte softening within the rubbery state from log(*G'*/Pa)=6, log(*G"*/Pa)=5.75 to log(*G'*/Pa)=5.75, log(*G"*/Pa)=5.5 after 10% mass increase. As long as the film is sufficiently far away from acoustic resonance, $-\Delta f$ *vs.* $\Delta m/m$ is still almost linear (up to 5% mass increase in our example). The acoustic amplification could be calculated if the initial modulus and that after analyte sorption are available. Unfortunately, neither of these is actually known. To go round that problem, the frequency shift is usually related to the analyte concentration in the medium, consequently the sensor must be calibrated in a preceding experiment.

Using the notation of eq (3b) eq (2) can be rewritten for a two layer arrangement as follows:

$$Z_{L12} = \frac{M_1V_1 + M_2V_2}{1+V_{12}} = (M_1V_1 + M_2V_2) * V_M \quad (4).$$

$$V_M = \frac{1}{1+V_{12}} = \frac{1}{1+(M_1V_1M_2V_2)/{Z_{c1}}^2} \quad (5)$$

is the mixed acoustic factor.

The mixed acoustic factor can be estimated from the frequency shift, f_1, after preparing the first layer, the frequency shift, f_2, of the second layer alone (measured e.g. on a separate blank quartz sensor) and the frequency shift, f_{12}, after finishing the double layer arrangement:

$$V_M \approx \Delta f_{12}/(\Delta f_1 + \Delta f_2) \quad (6).$$

3. METHOD

Mixed term acoustic amplification requires both a $V_1 > 1$ and a small Z_{c1} while V_2 can be 1. Hence the acoustic amplification also works with a rigid second layer or a thin film. The coating thickness, h_1, of layer 1 is the most important sensor design parameter. The shear modulus, G, is the most important material selection parameter. It must be small, about 1 MPa. We work with a rubbery polymer as first layer. Note, that the material must be in the rubbery state for the high frequency acoustic wave. The literature values cannot be applied. The shear modulus "seen" from the high frequency acoustic wave differs significantly from static values or those measured with low frequency mechanical measurements. Some rubbery polymer coatings behave acoustically rigid, hence not every polymer which is rubbery at static conditions fulfills the requirement of a small Z_{c1}. The design of the second layer (in terms of acoustics) is rather uncritically. In Fig. 2 we compare the frequency shift due to mass accumulation of a single 500 nm sensitive coating with two double layer arrangements with a rubbery 220 nm first layer. The amplification is 1.9 or 2.9, respectively.

The practical application faces *three* major problems. *First*, the first film must be uniform and homogeneous. We found spin coating followed by a thermal treatment of the film as the best preparation method. However, under unpropitious preparation conditions the film surface is rather rough and the film has inhomogeneities or even microcracks. These irregularities do not significantly influence traditional quartz crystal microbalance. They become a serious problem with the double layer arrangement

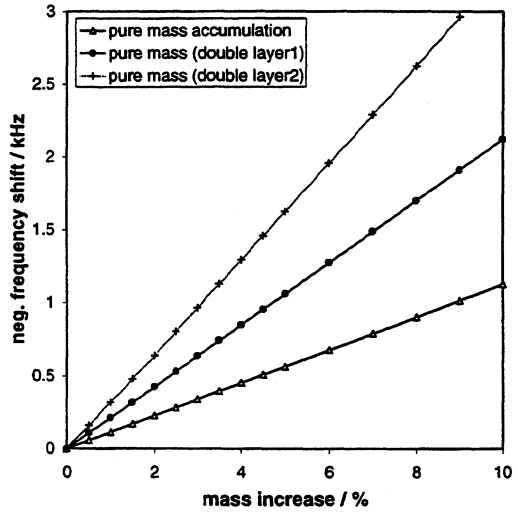


Fig. 2: Acoustic amplification with a double layer arrangement. The amplification is realized with the first layer (two different 220 nm rubbery films, ● or +), the 500 nm sensitive layer realizes pure mass accumulation in both cases (Δ).

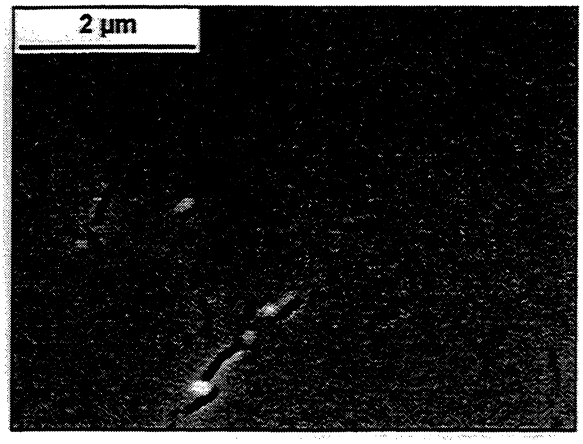


Fig. 3: REM-picture of a polymer film prepared with spin coating from a 15% solution

because, e.g., the second film partly traps the rubbery material. In consequence, the complex shear modulus of these "mixed" layer parts effective for the acoustic wave increase and the acoustic amplification vanishes. The irregularities are hardly recognizable with optical microscopy. Fig. 3 shows a micrograph made with a raster electron microscope.

Second, the preparation of the second layer must not attack the first layer. This problem can be solved with a proper selection of solvents or solvent mixtures and a fast preparation cycle, including the thermal treatment. Under certain circumstances it is advantageous to introduce a separation layer to overcome this and the next problem.

Third, if the analyte diffuses from the chemically sensitive layer into the acoustic amplification layer, the material properties of that layer may change. Consequently the sensor would exhibit an unwanted and uncontrolled change of the amplification factor. This material property change is also accompanied by an increase of acoustic energy dissipation in the double layer arrangement. It is usually expressed by a change of the resistance in the standard Butterworth-van Dyke equivalent circuit. In our sensor concept we determine this value from the amplification necessary to sustain oscillation with constant electrical oscillation amplitude. This new feature of our oscillator is realized by adding a monitored automatic level control (ALC). It is realized with a RF-precision rectifier, comparison of desired/instantaneous value, PI-control and an analog multiplier to get a negligible influence on amplitude response of the oscillator circuit. The feed back of the ALC to oscillator frequency is minimal. Details can be found in [7].

Our analysis shows that the resistance increase, ΔR, is not the best measure to control acoustic amplification. We found the quotient $\Delta R/\Delta f$ as a better adapted value. Although $\Delta R/\Delta f$ is not independent of the mass increase as found for a single layer [5], $\Delta R/\Delta f$ increases less than ΔR alone.

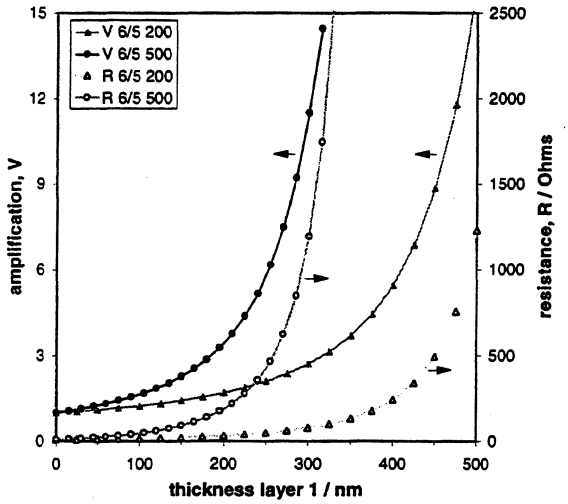


Fig. 4: Acoustic amplification, V, and resistance, R, in dependence of the thickness of the first layer with a shear storage modulus of 1 MPa and a shear loss modulus of 0.1 MPa. The thickness of the second layer is 200 nm or 500 nm, the mass increase is 5%.

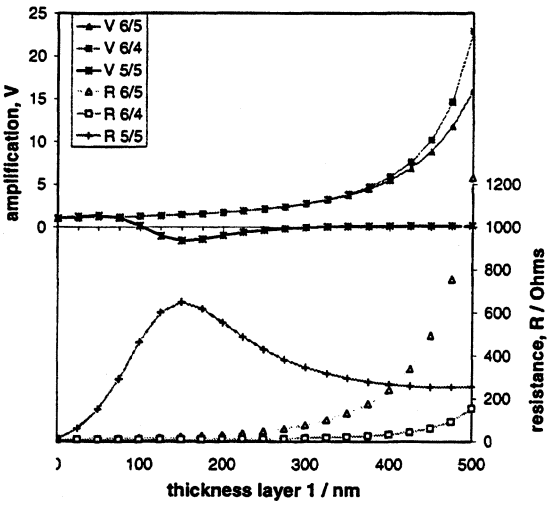


Fig. 5: Acoustic amplification, V, and resistance, R, in dependence of the thickness of the first layer for different complex shear storage modulus. The thickness of the second layer is 200 nm, the mass increase is 5%.

Taking the values of Fig. 2, $\Delta R/\Delta f$ increases 1.4% (double layer1) or 1.6% (double layer2) with a mass increase of 1% and 15% or 16% with a 10% mass increase, respectively.

Fig. 4 summarizes acoustic amplification, V, and the motional resistance, R, for two double layer arrangements in the case of a 5% mass increase. Both values vary with the thickness of the first 'amplifying' layer and the thickness of the second layer (legend parameter). In theory, an amplification of the mass effect of the second layer of more than a factor of 10 is possible. In practice, the resistance increase limits this possibility. We found a resistance up to about 500 Ω acceptable with our oscillator concept. The respective theoretical amplification is about 9 for a 450 nm first and a 200 nm second layer.

Fig. 5 finally shows acoustic amplification, V, and the motional resistance, R, for three double layer arrangements where the shear modulus of the first layer (legend parameter) is varied. Although the difference in acoustic amplification between the "6/5"-example and the 6/4-example are not big, a material with the smaller shear loss modulus is better adapted due to the smaller acoustic energy dissipation. The accessible amplification factor would be higher even with a smaller resistance limitation. However, such film modulus is hardly to realize in practice. The very soft and viscose material of the last example ("5/5") is not usable. The maximum in the resistance at 150 nm is related to film resonance. At about 300 nm the acoustic wave is almost completely absorbed and therefore any mass change of the second layer is not recognizable.

4. REFERENCES

[1] S.J. Martin, G.C. Frye, S.D. Senturia, "Dynamics and Response of Polymer-Coated Surface Acoustic Wave Devices: Effect of Viscoelastic Properties and Film Resonance ", Anal. Chem., 66, pp. 2201-2219, 1994.

[2] J.W. Grate, S.N. Kaganove, V.R. Bhethanabotla, "Examination of mass and modulus contributions to thickness shear mode and surface acoustic wave vapour sensor responses using partition coefficients" Faraday Disscuss., 107, pp. 259-283, 1997.

[3] R. Lucklum, C. Behling, P. Hauptmann, "Gravimetric and non-gravimetric chemical quartz crystal resonators", Techn. Digest 7th Int. Meeting on Chemical Sensors; Beijing, China, July 27-30, 1998, pp. 121-123.

[4] C. Behling, R. Lucklum, P. Hauptmann, "The non-gravimetric quartz crystal resonator response and its application for determination of polymer shear modulus" Meas.Sci.Techn., 9, pp. 1886-1893, 1998.

[5] R. Lucklum, C. Behling, P. Hauptmann, " The role of mass accumulation and viscoelastic film properties for the response of acoustic-wave-based chemical sensors" Anal. Chem., 1999, in print.

[6] Sauerbrey, G., "Verwendung von Schwing-quarzen zur Wägung dünner Schichten und Mikrowägung", Z. Phys. 155, pp. 206-222, 1959.

[7] F. Eichelbaum, R. Borngäber, J. Schröder, R. Lucklum, P. Hauptmann, "Interface circuits for quartz-crystal-microbalance sensors", Rev. Sci. Instr., 70, 1999, in print.

GAS PHASE CHEMICAL DETECTION WITH AN INTEGRATED CHEMICAL ANALYSIS SYSTEM

Stephen A. Casalnuovo, Gregory C. Frye-Mason, Richard J. Kottenstette, Edwin J. Heller, Carolyn M. Matzke, Patrick R. Lewis, Ronald P. Manginell, Susan L. Hietala, W. Kent Schubert, Vincent M. Hietala, Darryl Y. Sasaki, and John L. Reno

Sandia National Laboratories
P. O. Box 5800, MS-1425
Albuquerque, NM, USA 87185-1425

ABSTRACT

Microfabrication technology has been applied to the development of a miniature, multi-channel gas phase chemical laboratory that provides fast response, small size, and enhanced versatility and chemical discrimination. Each analysis channel includes a sample concentrator followed by a gas chromatographic separator and a chemically selective surface acoustic wave detector array to achieve high sensitivity and selectivity. The performance of the components, individually and collectively, is described. The design and performance of novel micromachined acoustic wave devices, with the potential for improved chemical sensitivity, are also described.

1. INTRODUCTION

Numerous chemical detection scenarios, for example, industrial process control and public safety applications, impose challenging requirements on the performance of chemical detection systems. These applications require detection of trace levels of specific target analytes in real-world environments that may contain more than 1000-fold higher concentrations of potentially interfering compounds. In addition, rapid analysis, instrument portability, and low rates of compound misidentification often are critically important. In an effort to address the requirements of these applications, Sandia's μChemLab™ program has a goal to develop small (palm-top computer sized), lightweight, and autonomous systems that provide rapid (1 min), sensitive (1-10 parts per billion (ppb)), and selective detection of target analytes. Although the μChemLab™ program includes gas and liquid phase analysis systems, only the performance of the gas phase components, being optimized for chemical warfare (CW) agent detection, is described herein.

Figure 1 shows the system design utilizing multiple analysis channels to provide enhanced chemical discrimination and very low false alarm rates. Sensitive and highly selective detection is achieved using a small adsorbent sample concentrator connected to a gas chromatographic (GC) column that feeds a low dead volume surface acoustic wave (SAW) detector array. Commercially available batch microfabrication processes are employed to produce these three components, pictured in Fig. 2. Among the advantages of this approach are size and cost minimization for individual devices. The design and performance of the current component set is described below, along with the development of novel, potentially more sensitive, acoustic wave chemical sensors.

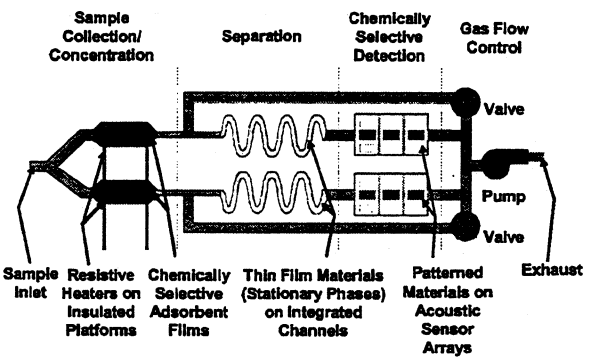


Fig. 1: Schematic of the gas-phase μChemLab™ system. The system incorporates Sandia designed and fabricated concentration, separation, and detection components (see Fig. 2) and commercially available diaphragm pumps and miniature valves.

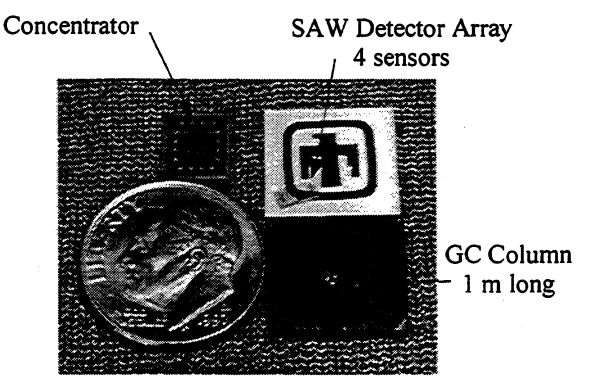


Fig. 2: Clockwise from above U.S. dime used for scale (18 mm diameter): microfabricated sample concentrator; four-element, 380 MHz SAW detector array, and 1 m long gas chromatograph column.

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

2. CONCENTRATOR

The concentrator stage collects target analytes from the air stream over an extended time and then releases them in a rapid, concentrated pulse into the GC column. This device, in essence, is a microfabricated hotplate (see Figs. 3 and 4). A chemically selective layer, typically a microporous oxide, adsorbs the analytes from the environment. After sufficient analyte has been collected, an embedded Pt heater rapidly raises the temperature of the concentrator to desorb the analytes. Concentration enhancement factors of greater than 100 have been achieved after 40 s adsorption periods. The desorption pulse width is 200 ms (full width at half maximum).

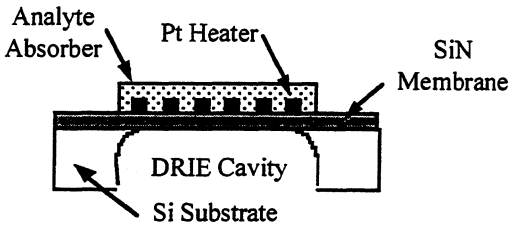


Fig. 3: Schematic cross section of the microhotplate concentrator stage.

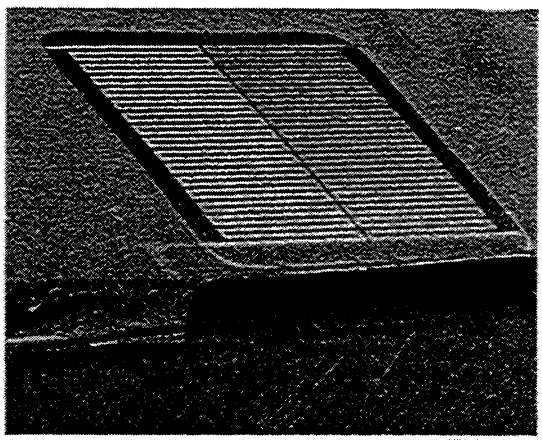


Fig. 4: Scanning electron micrograph (SEM) of the concentrator stage showing the micromachined SiN membrane and Pt heater. Analyte adsorbing layer is not shown.

The heater is fabricated on a 0.5 μm thick silicon nitride (SiN) membrane suspended over a cavity etched through a Si substrate. The 400 μm deep cavity is produced using deep reactive ion etching (DRIE) [1]. The low thermal mass and good thermal isolation of the SiN membrane are critical for quickly heating the concentrator with a minimum of electric power. As shown in Fig. 5, the microhotplate requires approximately 20 ms to reach a steady-state temperature of 200°C. This steady-state temperature is sustained by 105 mW of electrical power. Thermal modeling has been used to optimize the concentrator design [2].

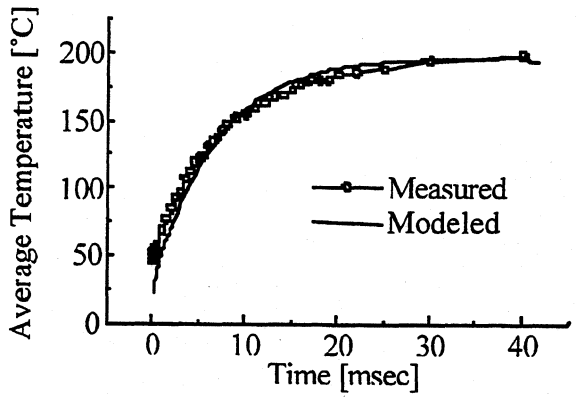


Fig. 5: Comparison of measured and modeled heating rates of the microhotplate concentrator during analyte desorption. The concentrator can be ramped from room temperature to 200°C in 20 ms, drawing 105 mW at steady-state.

For prototype testing, the concentrator is mounted in a standard integrated circuit package that provides the electrical interconnections. A glass lid with machined gas flow channels is glued directly to the Si concentrator substrate. Microcapillary tubes are inserted into through holes machined in this lid to provide interconnection to a miniature pump, valve, and the GC column.

3. GC COLUMN

A miniature GC column provides temporal separation of the analytes and any interferants that may pass through the concentrator stage. A Si DRIE process, similar to that used to fabricate the concentrator, is employed to produce the GC column [3]. Because of the exceptionally high aspect ratio and anisotropy of the DRIE process, closely spaced, narrow gas flow channels can be etched into the Si substrate to a depth many times the channel width. This approach affords good GC performance in a small footprint while maintaining short transit times through the column. A typical column is a 1 m long spiral with 40 μm wide channels separated by 40 μm thick walls etched to a depth of 300 μm (see Fig. 2). It occupies a 1 cm^2 area. The close-up of a channel cross-section in Fig. 6 illustrates the capability of the fabrication process. After the channels are etched, the Si substrate is thermally oxidized to produce a thin, glasslike layer on the surface of the channels in order to facilitate stationary phase deposition (see below).

Closed channels are produced by anodically bonding a Pyrex lid to the top surface of the Si substrate [3]. Since the bonding process is carried out at elevated temperatures, Pyrex is used because it closely matches Si's thermal expansion coefficient.

As with the concentrator, microcapillary tubes are inserted into through holes machined in this lid to provide gas interconnection.

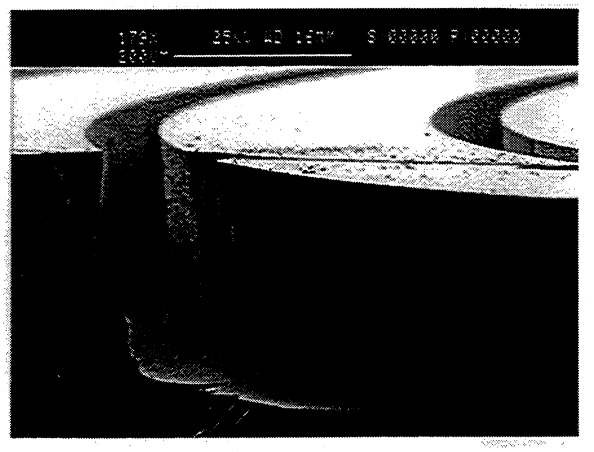


Fig. 6: SEM of a cross-section of a spiral GC column etched into Si using DRIE. The etch process results in deep, smooth, vertical channels.

After the channels are sealed, GC stationary phase materials are deposited on the walls using conventional polymer solution techniques as well as sol-gel coating technology. The retention of the analytes in the stationary phase produces a separation in time of the analytes arriving at the array of acoustic wave detectors. This temporal separation provides an additional means of distinguishing analytes from one another and from interferants, as shown in Fig. 7, aiding in analyte identification.

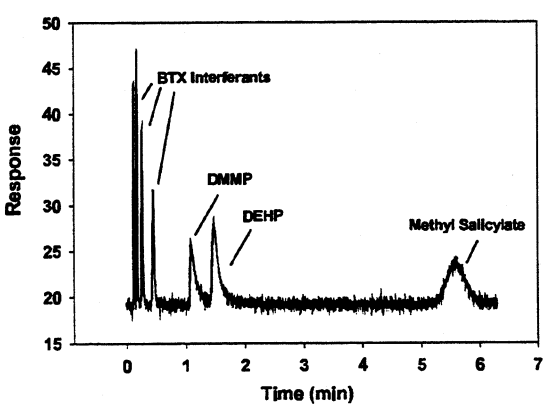


Fig. 7: Microfabricated GC column elution times for a mixture of dimethylmethylphosphonate (DMMP), diethylhydrogenphosphate (DEHP), methyl salicylate, all simulants for CW agents, and benzene (B), toluene (T), and xylene (X), interferants present in gasoline vapor.

As illustrated in Fig. 1, the μChemLab™ system will have two gas analysis channels that incorporate GC columns, each with a different stationary phase. The two channels can be optimized for different analytes, increasing the functional range of the unit. The channels can also be optimized to provide two analyses for an analyte, improving the reliability and lowering the false alarm rate, as demonstrated in Fig. 8. The stationary phase selection, in this case, non-polar versus polar, reorders the elution times of the components.

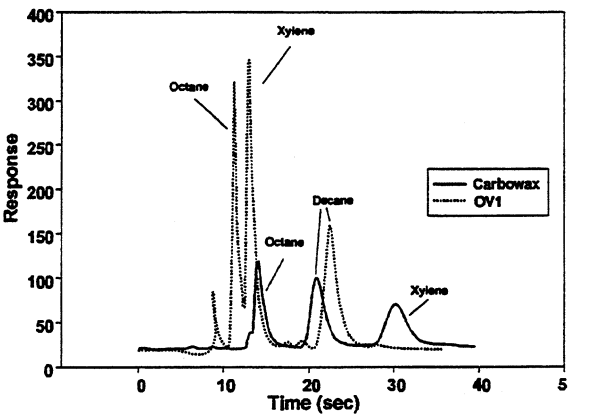


Fig. 8: Elution times for two microfabricated GC columns with different stationary phases (Carbowax and OV1). Reordering of elution times in the two channels provides confirmation of results.

4. SAW CHEMICAL SENSOR ARRAY

The use of SAW delay lines as sensitive chemical mass sensors is well known [4]. The μChemLab™ employs a 4-element quartz array in each gas analysis channel, as shown in Fig. 9. A center input interdigitated transducer (IDT) launches a SAW in both directions. Four smaller output IDTs, two on either side of the input IDT, reconvert the acoustic wave to an electric signal. Three of the SAW delay lines are coated with different chemically sorbent materials, each optimized for the analytes of interest. The fourth delay line serves as a reference. Delay lines operating between 100 MHz and 700 MHz have been designed and tested. Delay line frequency is determined by the spacing between adjacent IDT fingers and the acoustic velocity in the material.

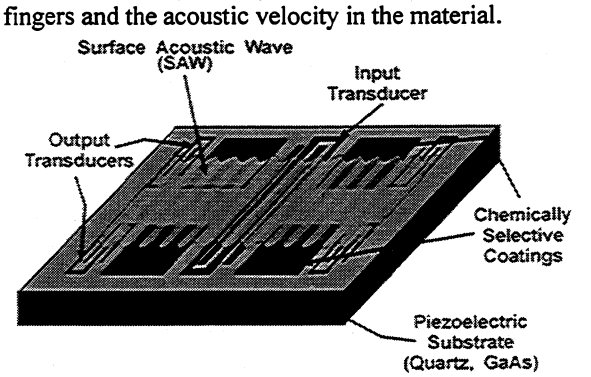


Fig. 9: Schematic representation of a 4-element SAW detector array.

The delay line frequency is a critical parameter in the design and operation of the SAW sensor array. When a delay line is configured in an oscillator circuit, the sensitivity of the oscillator frequency, f, to added surface mass density (mass/area) increases as f^2 [4]. In addition, the size of the delay line decreases as $1/f^2$. The net result is that the sensitivity to the total mass of the analyte scales as f^4, which is a strong argument for going to higher frequency. These trends are illustrated in Fig. 10. The scaling of the detection limit for a delay line sensor, that is, the smallest amount of added mass that can be detected, depends on how the system noise scales with frequency. The frequency dependence of sensor noise is currently being studied.

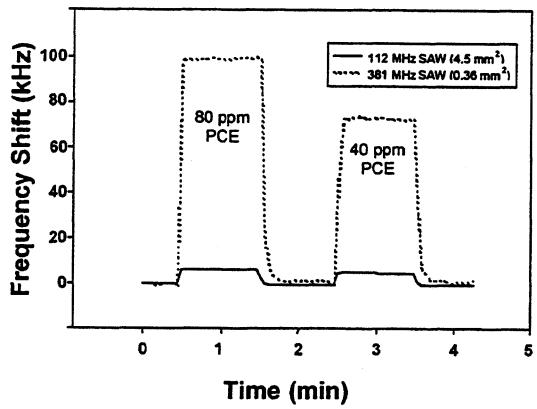


Fig. 10: Sensitivity data for two SAW devices with identical coatings when exposed to perchloroethylene vapors. The higher frequency SAW is much smaller, but exhibits a much larger response.

All three microfabricated components, the concentrator, GC column, and SAW detector array, have been assembled in a laboratory test system to evaluate their collective performance. An example of the system performance, showing the temporal separation of target analytes and the varied response of the three elements in the SAW detector array, is given in Fig. 11. The initial gas sample contained DMMP and DHEP at a concentration of 100 ppb. Analyte gas flowed over the concentrator for 3 minutes prior to thermal desorption. SAW response was measured with a phase demodulation technique.

Within the actual prototype unit, the SAW sensor array is driven by custom, high frequency GaAs integrated circuits (IC). The reference delay line is wired into the feedback path of an IC amplifier. The phase delay of each of the three sensor delay lines is compared to the reference by an IC phase comparator. This IC produces a DC voltage proportional to the phase difference between the sensor delay line and the reference delay line. As analytes are sorbed by the coatings on the three sensor delay lines, the added surface mass decreases the velocity of the SAW, producing a change in phase in the delay line. With this approach, sub-picogram detection limits are possible.

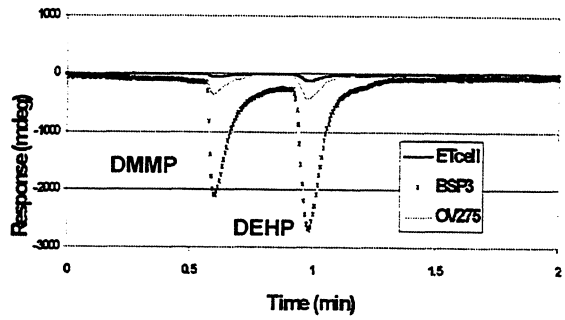


Fig. 11: SAW sensor array response to mixture of DMMP and DEHP after passing through the concentrator and GC column. Sensor array elements were coated with ethyl cellulose (Etcell), a H-bond acid modified polymer (BSP3) [5], and a cyano-modified polysiloxane (OV275).

The attraction of higher frequency SAW sensors is reduced somewhat by the difficulty associated with packaging high frequency components. This problem is minimized if all high frequency components, including the SAW delay line, can be monolithically integrated onto a single IC requiring only DC electrical power as input and providing a DC output. As a first step in this approach, we have integrated a high frequency amplifier and a single SAW delay line onto a GaAs substrate [6] (Fig. 12). This is possible because GaAs is both semiconducting and piezoelectric. This oscillator circuit requires only a 3.5 VDC power supply to operate. Frequency response for the integrated oscillator is shown in Fig. 13. While this circuit requires only a DC input, it still yields a high frequency output. The next step in this development is to integrate a second delay line and the phase comparator circuit for complete DC in/DC out, single chip operation. This effort is currently underway.

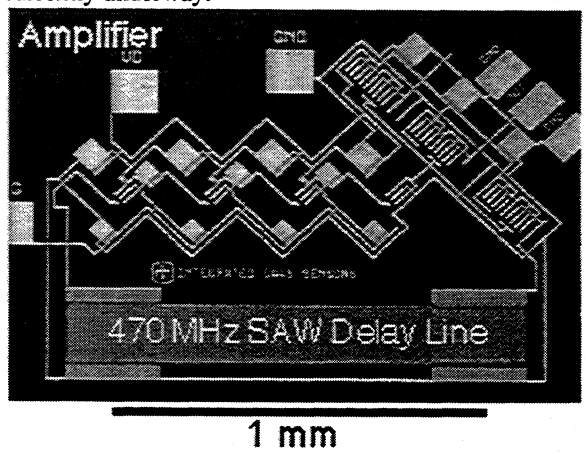


Fig. 12: Micrograph of an integrated GaAs SAW delay line oscillator with a SAW delay line in the feedback loop of a high frequency amplifier. The oscillator operates at 470 MHz.

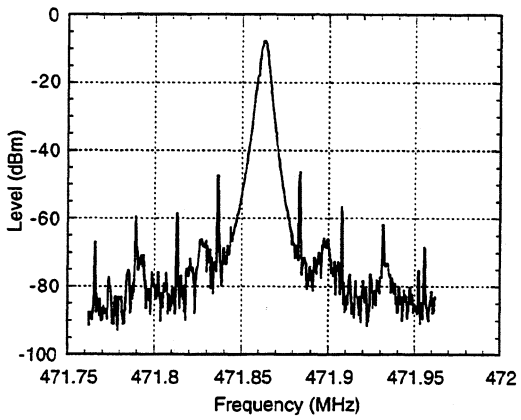


Fig. 13: Frequency response of integrated GaAs oscillator shown in Fig. 12

5. ACOUSTIC WAVE DEVICE DEVELOPMENT

As a sensor material, GaAs has advantages beyond the integration of acoustic wave transducer and microelectronics. With the application of micromachining processes analogous to those found in Si micro-electro-mechanical systems (MEMS) technology, it is possible to produce freely suspended GaAs membranes that retain all the crystalline properties of bulk GaAs [7]. We have used this approach to fabricate acoustic wave devices on piezoelectric GaAs membranes as thin as 0.5 μm [6]. There are two advantages to this approach [4]. The first is that it is possible to launch acoustic modes in these membranes, specifically, flexural plate waves (FPW) and thickness shear modes (TSM), that will propagate when the sensor is submerged in a liquid. The SAW mode is highly damped in a liquid and does not propagate along the sensor substrate. The second advantage is that the sensitivity to added mass for each of these modes increases with decreasing membrane thickness. Potentially, FPW and TSM acoustic wave chemical sensors can be made much more sensitive than technically realizable SAW sensors.

FPW delay lines are fabricated much the same as SAW delay lines, except that the IDTs are placed on a membrane whose thickness is less than the wavelength of the FPW. Acoustic mode wavelength is determined by the spacing between adjacent IDT fingers for both FPWs and SAWs. Since IDTs are fabricated before the membrane is formed, it is possible to test both SAW and FPW propagation with the same device. An example of this is shown in Fig. 14. Prior to the release etch that forms the membrane, the SAW mode exhibits a peak in its frequency response at 320 MHz. After the release etch, the peak in the frequency response shifts to 220 MHz. This shift to lower frequency is one characteristic of FPW modes and results from the reduced acoustic phase velocity for FPW modes relative to SAW modes.

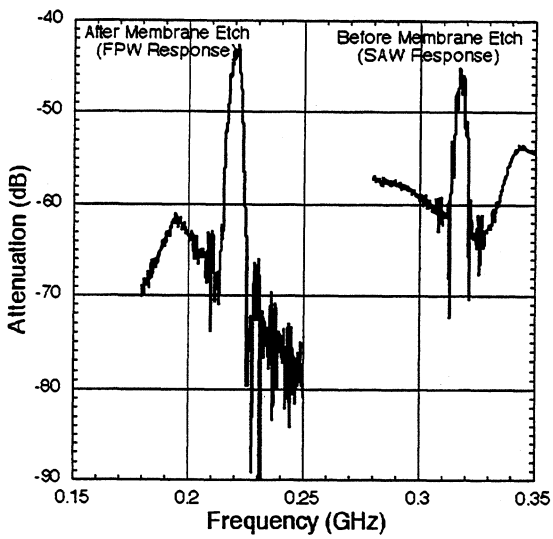


Fig. 14: Comparison of SAW and FPW frequency response. Measurements are taken on the same device before (SAW) and after (FPW) the membrane release etch. The shift to lower frequency is characteristic of the FPW mode, the result of a reduced acoustic phase velocity.

FPW phase velocity scales with t/λ, where t is the membrane thickness and λ is the FPW wavelength. FPW velocity goes to zero for small t/λ and approaches the SAW value in the limit of large t/λ. For a given material, FPW velocity falls on a single curve for all values of t and λ, when plotted against t/λ. This dependence on t/λ is a clear indication of FPW mode propagation. A comparison of FPW phase velocity data to the theoretical prediction is shown in Fig. 15 for several membrane thicknesses. Phase velocity is determined by multiplying the frequency corresponding to the maximum response (as shown in Fig. 14) by the wavelength (determined by the IDT spacing). This is just a restatement of the relationship $v = f\lambda$. The theoretical curve is completely determined by GaAs materials properties. As such, there are no adjustable parameters to fit to the data. As seen in the Fig. 15, the agreement between theory and experiment is excellent. FPW resonators are now being constructed.

It is also possible to create a TSM resonator in the GaAs membrane, where the two surfaces of the membrane act as acoustic "mirrors". For this mode, the resonant frequency is given by $f = 1.67/t$ GHz, where t is measured in microns [8]. This is in good agreement with the measured resonance in 0.5 μm thick GaAs membranes (see Fig. 16). Multiple resonances were observed for 1.0 μm membranes, making identification of the fundamental mode difficult. No resonance was found in the 3.0 μm membranes. Additional study of these devices is underway.

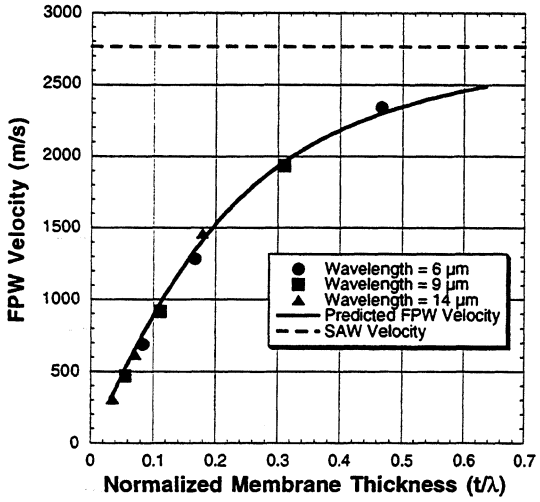


Fig. 15: Comparison between measured FPW phase velocity and theoretical prediction for three membrane thicknesses (0.5, 1.0, and 3.0 μm) and three acoustic wavelengths. No adjustable parameters were used to fit the data.

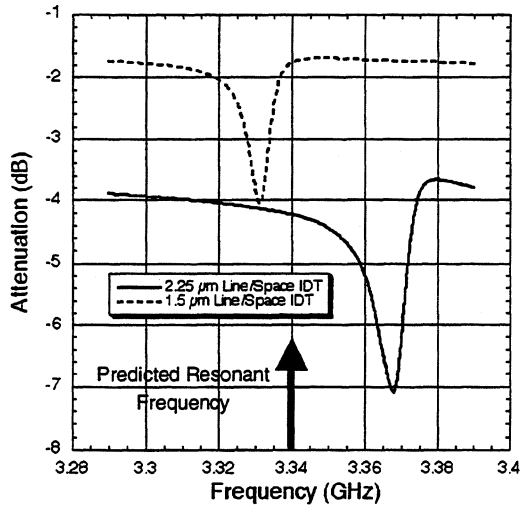


Fig. 16: TSM resonances for two 0.5 μm thick GaAs membranes. Predicted resonant frequency is 3.34 GHz.

6. CONCLUSION

Trace level detection of chemical warfare agent simulants has been accomplished using microfabricated chemical analysis components. A sample concentrator, a gas chromatographic column, and a 4-element surface acoustic wave sensor array have been combined to detect target analytes and interferants with detection limits as low as 10 parts per billion at the system inlet, depending upon the analyte. Pattern recognition algorithms applied to temporally separated detector array responses provide highly reliable identification of a number of analytes. Integration of the detectors with drive and signal conditioning electronics is being explored to improve the performance of high frequency sensors and to simplify packaging. Micromachined acoustic wave devices are being studied because of their potential for greater sensitivity.

7. ACKNOWLEDGEMENTS

Sandia National Laboratories is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under Contract DE-AC04-94AL85000.

8. REFERENCES

[1] R.P. Manginell et al., "Microfabrication Of Membrane-Based Devices by HARSE and Combined HARSE/Wet Etching", in Proceedings of the SPIE Conference on Micromachining and Microfabrication IV, vol. 3511, 1998, pp. 269-276.

[2] R.P. Manginell et al., "Finite Element Modeling of a Microhotplate for Microfluidic Applications", to be presented at the International Conference on Modeling and Simulation of Microsystems (MSM'99), San Juan, Puerto Rico, April 19-21, 1999.

[3] C.M. Matzke et al., "Microfabricated Silicon Gas Chromatographic Micro-Channels: Fabrication and Performance", in Proceedings of the SPIE Conference on Micromachining and Microfabrication IV, vol. 3511, 1998, pp. 262-268.

[4] D.S. Ballantine et al., Acoustic Wave Sensors: Theory, Design, and Physico-Chemical Applications, San Diego, Academic Press, 1997, ch. 3.

[5] Provided by Jay Grate, Pacific Northwest National Laboratory.

[6] S.A. Casalnuovo et al., "Acoustic Wave Chemical Microsensors in GaAs", in Proceedings of the SPIE Conference on Micromachined Devices and Components IV, vol. 3514, 1998, pp. 103-110.

[7] K. Hjort, "Gallium Arsenide Micromechanics: A Comparison to Silicon and Quartz", Alta Frequenza-Rivista di Elettronica, vol. 6, pp. 48-54, 1994.

[8] J. Soderkvist and K. Hjort, "The Piezoelectric Effect of GaAs Used for Resonators and Resonant Structures", Journal of Micromechanics and Microengineering, vol. 4, 1994, pp. 28-34.

THE USE OF DEXTRAN AS AN INTERMEDIATE LAYER: A NEW APPROACH TOWARDS SAW BASED BIOSENSORS

N. Barié*, M. Rapp*, H. Sigrist#

* Forschungszentrum Karlsruhe GmbH, Institut für Instrumentelle Analytik
P.O. Box 36 40, D–76021 Karlsruhe, Germany, Fax: +(49) 7247-82-4618

Centre Suisse d'Electronique et de Microtechnique
Jaquet-Droz 1, CH-2007 Neuchâtel, Switzerland, Fax +(41) 32-7205-720

ABSTRACT

We present a new method for covalent binding of dextran as an intermediate layer on surface acoustic wave (SAW) devices. The SAW devices were originally developed for use in modern telecommunications and are thus available as series products at low costs.
For biosensing applications these devices must be coated with a shielding layer to prevent corrosion effects of the aluminium structures in aqueous media. Thin films of polyimide and parylene, respectively, showed good shielding properties and were used as a base for further immobilization.
Dextran immobilization dextran to the polymer coated surfaces is achieved by a photoimmobilization process. A aryldiazirine-functionalized protein (T-BSA) serves as a multifunctional light-activated linking agent (photolinker polymer). Dextran and the photolinker are mixed and photobonded to the sensor surface.
Immobilization of proteins to the dextran layer via carbodiimide chemistry is exemplary demonstrated with anti-urease antibodies and the feasibility of specific immunosensing is investigated using SAW sensors connected to a fluid handling system.

1. INTRODUCTION

In the past decade SAW devices have been more and more considered as an appropriate tool for biosensing. To create a biosensor these devices are coated with a sensitive layer of receptive molecules which enables specific biological interactions. These interactions are detected by determining changes in surface wave velocity, predominantly caused by mass adsorption or viscosity changes.

In the above mentioned studies the SAW devices used were self-made and based on special developments. SAW devices used in this study are widely used in telecommunication systems and thus they are commercially available as a mass product at low cost. They base on horizontal polarized shear waves with lithiumtantalate as substrate material. The operating frequency is 380 MHz and the typical attenuation during operation is 2 dB in air and 4 dB in water.

For biosensing purposes these SAW devices have to be coated with a suitable shielding layer in order to protect the aluminium structures (transducer structures and bonding wires) on the surface [1]. Two different polymers were used to prevent corrosion effects: polyimide and parylene. They both were found to provide exceptional passivation properties, enable serial production and suit subsequent surface bioengineering.

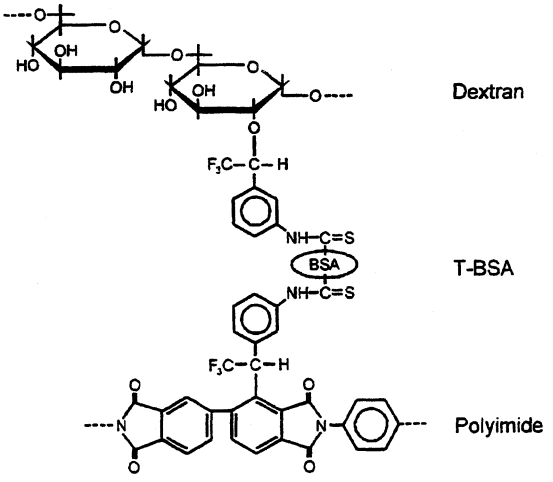


Fig. 1: Schematical scheme for the layer structure of photolinker-mediated covalent immobilization of dextran to the shielding layer (not restricted to polyimide).

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

This study introduces a new approach to produce very versatile SAW biosensor platforms. This is achieved by single step, covalent photobonding of an intermediate dextran layer to the shielded sensor surface. Surface-bound dextran permits experimentally facile, rapid and reproducible immobilization of almost any receptive molecule with known techniques [2].
The dextran layer is covalently bound to the polymer coated surface by means of photo-immobilization [3]. Bovine serum albumin, multiply derivatized with light-sensitive trifluoro-methyl-aryldiazirine (T-BSA) [4], serves as a linker between the shielding layer and dextran (fig. 1). T-BSA and dextran are mixed and applied to the polymer coated sensor surface. After solvent removal, the coated surfaces are exposed to activating light. Aryldiazirines absorb light of wavelength 350 nm and form highly reactive carbenes, which insert into chemical bonds [5]. The high degree of substitution of BSA with photoreagents (6-10 mol diazirine per mol BSA) enables the simultaneous insertion of photo-generated carbenes into both dextran and the shielding layer. Thus covalent coupling of dextran to the surface is achieved.

In view of the optimization of the dextran binding process, biotinylated dextran and carboxy-methylated dextran were used in this study. Biotinylated dextran served to quantitate the photoimmobilization step by recording the retention of radiolabeled streptavidin. Carboxy-methyl-modified dextran was used as a chemical base for covalent binding of sensing biomolecules. Receptive molecules are covalently immobilized to dextran applying chemistries which confer with the nature of the ligand and the purpose of the immobilization.

2. MATERIALS AND INSTRUMENTATION

SAW devices, type SAF380T with an operating frequency of 380 MHz, were purchased from *Murata Company, Japan*. The devices are commercially available low-loss-filters based on horizontal polarized shear waves (HPSW, STW) on a 36° rotated $LiTaO_3$ XY cut substrate. For analyte detection modified devices were placed in a round chamber, 8 mm diameter (TO 39). Mounted in teflon adapters the sensors were connected with a continuos flow system with a peristaltic pump. Analyte binding measurements were carried out in an oscillator circuit, which was purchased from the *University of Heidelberg, Germany*. Resonance frequencies were detected by a Hewlett Packard frequency counter (*HP-PM 6680*), scanner (*HP-PM 2327*) and switch (*HP-PM 2301*). The instruments were controlled by a personal computer via a GPIB interface.

For irradiation a Stratalinker 2400 UV Cross-linker (*Stratagene GmbH, Germany*) was used.

Phosphate buffer contained either 0.5, 20 or 50 mM potassium phosphate. The pH was adjusted to 7.2 with KOH.
Modified surfaces were washed with 50 mM phosphate buffer containing 0.02 % (v/v) Tween 20 (phosphate buffer/Tween).
Phosphate buffered saline (PBS) used for immunoreactions consisted of 20 mM phosphate buffer (pH 7.2) containing 140 mM NaCl.

[^{35}S]-labeled streptavidin was detected by scintillation counting utilizing *Ultima Gold (Packard No. 6013329)* scintillation fluid (5 ml each sample).

3. METHODS

3.1 Surface passivation

Polyimide was applied as a shielding layer which was covalently attached to the hydrophilic SAW surfaces after silanisation [1]. Surface silanisation was done with dimethyl-aminopropyl-ethoxy-silane. The polyimide consisted of the monomers p-phenylene-diamine and 3,3',4,4'-biphenyltetra-carboxylic dianhydride.
Parylene C (poly-(2-chloro-p-xylylene) was deposited in a factory setup by *Comelec SA, Switzerland*. The thickness of the parylene film was 0.8 ± 0.3 µm.

3.2 Photoimmobilisation of dextran

T-BSA and dextran were mixed in phosphate buffer (pH 7.2). Passivated surfaces were coated with the T-BSA / dextran solution, dried at ambient temperature and irradiated for the indicated length of time with UV light source. After photobonding, the surfaces were washed repeatedly in chaotropic / detergent containing solutions.
For optimization of the coating procedure, parylene coated glass substrates were modified with biotinylated dextran. An aqueous solution of [^{35}S]-streptavidin was applied to the coated surface and incubated at room temperature. After washing with phosphate buffer / tween and water

bound [^{35}S]-streptavidin was quantified by liquid scintillation counting.
At least 3 supports were modified identically, and triplicate samples were analyzed. Identically treated non-irradiated samples served as controls.

3.3 Protein immobilization

Aqueous solutions of NHS (100 mM) and EDC (400 mM) were mixed. The reagent mixture was applied on the sensor surface (coated with carboxymethylated dextran) and incubated for several minutes. After activation of the dextran carboxyl functions, protein solution in acetate buffer (pH 5.0) was applied and incubated at room temperature for 20 min. Excess NHS ester was deactivated with ethanolamine (1 M). Unbound protein was washed off by repeated rinsing of the sensor surface with phosphate buffer. Non-specific binding sites were saturated with BSA (4% w/v, 10 min).

3.4 Monitoring of the immunoreaction

The coated SAW devices were mounted in the flow system and rinsed with PBS until the resonance frequency. Solutions containing the antigen or non-specifically interacting (immuno-) reagents, dissolved in PBS were injected in the carrier buffer. Time dependent resonance frequency changes were monitored under constant flow conditions.

4. RESULTS AND DISCUSSION

4.1 Photoimmobilization of dextran

Photoimmobilization of dextran to the polymer coated sensor surface was achieved by co-immobilization with T-BSA as described above.
In the co-immobilization of binary systems the molar or mass ratio of both molecules is of decisive importance. To determine the optimum ratio in the co-immobilization of dextran with T-BSA, parylene-coated surfaces were covered with various mixtures of biotinylated dextran (MW 10,000) and T-BSA. Then sampling was performed with [^{35}S]-streptavidin.
The optimized dextran coating solution consists of 25 mol% dextran and 75 mol% T-BSA. The photoreaction was completed after 45 min irradiation (fig. 2).

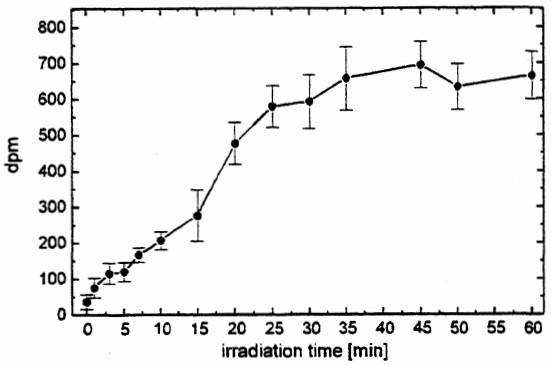


Fig. 2: Optimization of irradiation time. Parylene passivated glass substrates were coated with a mixture of 25 mol-% biotinylated dextran and 75 mol-% T-BSA, dried and exposed to light for the indicated length of time. Bonding was determined by measuring the remaining radioactivity after reaction with [^{35}S]-streptavidin.

After photoimmobilisation under optimized conditions 7 % of the applied dextran was recovered on the surface. This corresponds to a surface density of 1 ng/mm² dextran.
The reproducibility of the coating procedure was determined by coating of 10 parylene coated substrates under identical conditions. A coefficient of variation of 8% was found.

4.2 Immobilization of proteins to dextran using carbodiimide chemistry

The carboxymethylated dextran coated SAW device was rinsed with 20 mM phosphate buffer until a stable base line was obtained. A mixture of NHS and EDC was injected to activate the carboxymethylated dextran. Due to different conductivities of both solutions a sensor response

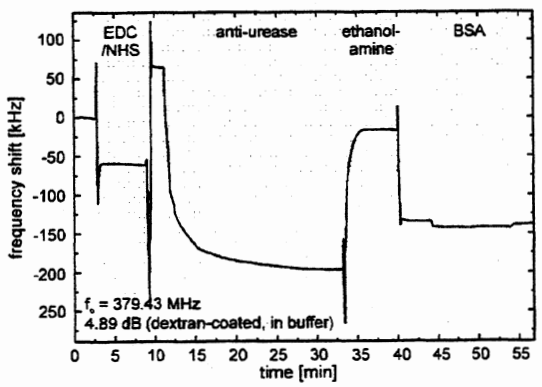


Fig. 3: Sensor response during protein immobilization on carboxymethylated dextran. Anti-urease antibodies were immobilized by reacting with in situ generated NHS esters.

was observed. Subsequently, the protein solution (in this case: anti-urease antibodies) was injected. The resonance frequency decreased indicating covalent antibody binding. Ethanolamine was injected to deactivate remaining NHS esters. Non-specific protein binding sites in the system were blocked with BSA (see fig. 3) [2,6-8].

4.3 Specific and non-specific binding

After the immunoreaction with urease (500 µg/ml in PBS) subsequent rinsing with PBS slightly reversed the signal (release of loosely adsorbed urease), leaving a net frequency shift of 110 kHz, representing the irreversible specific binding of urease to the antibodies (see fig. 4).
The reagents non-complementary to anti-urease led to small and nearly reversible frequency shifts. In addition, a SAW device coated with polyimide and carboxymethylated dextran but without immobilized antibodies, was treated with urease.

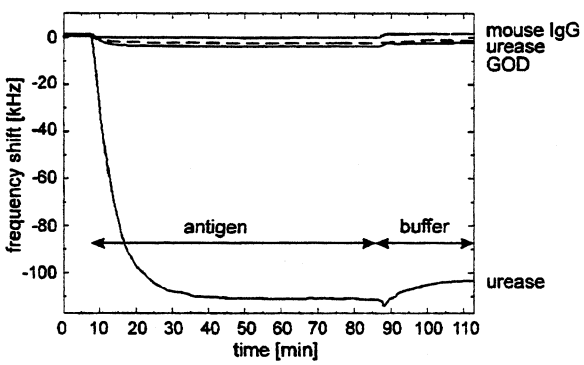


Fig. 4: Sensor responses of identically prepared sensors (coated with monoclonal anti-urease antibodies on carboxymethylated dextran) to specific and non-specific antigens. In addition, the response of a dextran coated sensor (without immobilized antibodies) to urease is shown (dotted line).

These minor effects are due to non-specific adsorption of either reagent applied. In contrast, the specific reaction of urease led to a significant and irreversible frequency shift.

4.4 Reproducibility of sensor responses

For studying the reproducibility of sensor signals, five dextran coated SAW devices were modified with anti-urease antibodies and treated with a solution of urease (100 µg/ml). A mean frequency shift of 110 kHz was obtained with a coefficient of variation of 12%.

SUMMARY AND CONCLUSION

A new method for biofunctionalization of SAW sensor surfaces was presented. After application of a polymer coating the commercially available and inexpensive SAW devices are suited for use as chemical or biochemical sensors. Covalent photo-bonding of dextran to polymer coated surfaces was achieved by photolinker-mediated immobilization with aryldiazirine-functionalized bovine serum albumin. The use of dextran as an intermediate layer for biosensor applications was shown by immobilization of monoclonal anti-urease antibodies on carboxymethylated dextran. Thus, the engineering of a very versatile SAW biosensor has been attained. The method is applicable for simple, rapid and reproducible immobilization and detection of biomolecules.

References

[1] Wessa, T.; Barié, N.; Rapp, M.; Ache, H.J. *"Polyimide, a new shielding layer for sensor applications"* Sens. Act. **B 53** (1998) 63-68

[2] O'Shannessy, D.J.; Brigham-Burke, M.; Peck, K. *"Immobilization chemistries suitable for use in the BIAcore surface plasmon resonance detector"* Anal. Biochem. **205** (1992) 132-136

[3] Sigrist, H.; Collioud, A.; Clémence, J.F.; Gao, H.; Luginbühl, R.; Sänger, M.; Sundarababu, G. *"Surface immobilization of biomolecules by light"* Optical Engineering **34** (1995) 2339-2348

[4] Sigrist, H.; Gao, H.; Wegmüller, B. *"Light-dependent, covalent immobilization of biomolecules on inert surfaces"* Bio/Technology **10** (1992) 1026-1028

[5] Dolder, M.; Michel, H.; Sigrist, H. *"3-(Trifluoromethyl)-3-(m-isothiocyanophenyl)diazirine"* J. Prot. Chem. **9** (1990) 407-415

[6] Johnsson, B.; Löfas, S.; Lindquist, G. *"Immobilization of proteins to a carboxymethyl dextran-modified gold surface for biospecific interaction analysis in SPR sensors"* Anal. Biochem. **198** (1991) 268-277

[7] Johnsson, B.; Löfas, S.; Lindquist, G.; Edström, A.; Müller-Hillgren, R.M.; Hansson, A. *"Comparison of methods for immobilization to carboxymethyl dextran sensor surfaces"* J. Mol. Recogn. **8** (1995) 125-131

[8] Löfas, S.; Johnsson, B.; Tegendal, K.; Rönnberg, I. *"Dextran modified gold surfaces for SPR sensors"* Coll. Surf. B: Biointerfaces **1** (1993) 83-89

1999 Joint Meeting EFTF - IEEE IFCS

SENSING OF ORGANIC VAPOR ADSORPTION ON GOLD USING A TEMPERATURE INSENSITIVE MICROBALANCE

Yoonkee Kim* and Daniel E. Pierce**
* U.S. Army Communications-Electronics Command
Fort Monmouth, NJ 07703 (ykim@ieee.org)
**Department of Chemistry and Physics
William Paterson University, Wayne, NJ 07470

Abstract

The adsorption of organic vapor molecules onto the gold electrode surface of a dual mode 10MHz 3rd overtone SC-cut quartz crystal microbalance (QCM) was investigated. Using this dual mode QCM, the separation of mass and temperature change can be made in real time, with resolutions of 56 pg and better than 0.01K, respectively. The QCM experiments were performed in a UV/Ozone chamber so that the gold electrodes could be maintained in a clean state. Organic vapors or liquids (e.g., benzene, hexane, methanol, ethanol, isopropanol, or butanol) were injected into the chamber while the two modes of the dual mode QCM were monitored and the mass-frequency and temperature changes were extracted. This paper demonstrates the ability of the sensor to determine in a matter of seconds, sub-ppm (ppm≡parts per million) gas concentrations of benzene and hexane while operating in an elevated temperature environment without any temperature stabilization. In addition, the mass changes for a series of adsorbed methanol, ethanol, isopropanol, and butanol were found to be proportional to molecular weights.

Introduction

Quartz crystal oscillators are some of the most widely used acoustic devices with a long history of development as sensors of temperature, stress, acceleration, viscosity, mass, and are universally used to measure time [1-5]. Due to the effects of mass loading, the quartz crystal microbalance (QCM) works as a sensor of mass, with a capability of sensing considerably less than a single molecular layer of adsorbed material. Most QCM's utilize AT-cut, which has a temperature coefficient of 0 ppm/K at room temperature. Though working well at the specified design temperature, as the temperature deviates from this "turnover temperature" (usually near room temperature), the temperature coefficient of frequency increases, which makes the temperature induced mass sensing errors also increase. Depending on the application, precise temperature regulation may be necessary to eliminate temperature change effects. For example, a QCM typically used for film thickness monitoring in a thin film deposition system is cooled by circulated water. Temperature control, however, imposes increased power, size, and complexity requirements. Moreover, temperature control has its own limitations, e.g., due to thermal gradients among the temperature sensor, heater (or cooling means), and the resonator. Temperature control is not always practical in certain applications.

We have demonstrated a temperature insensitive quartz microbalance (QCM) technique [6], which employs dual mode frequencies, the fundamental mode and 3rd overtone, to discriminate between mass change effects and temperature change effects. This technique requires neither temperature control nor an independent temperature sensor, only the resonator itself. Dual mode excitation of SC-cut resonators can be used for highly accurate self-temperature sensing over wide temperature ranges. SC-cut resonators derive their "SC" designation from the fact that the cut is "stress compensated". This property also results in thermal-transient compensation. These unique properties of the SC-cut allowed for the development of a temperature compensated microbalance. Detailed discussion of the method of separating temperature and mass effects can be found in Ref. [6]; and is not reproduced here.

Experiments

The resonator used in the experiments was a plano-convex SC-cut 10 MHz 3rd overtone (3.3 MHz fundamental mode) resonator developed for a microcomputer compensated crystal oscillator (MCXO) [7]. The diameters of the plate and gold electrode were 14 mm and 5.0 mm, respectively, and the plate contour was 3 diopters. The resonator frequencies were measured using a dual mode crystal oscillator (DMXO), which was originally developed for an MCXO, then modified for an opened resonator. A resonator in a vacuum-sealed metal can was calibrated for temperature effects in a temperature controlled chamber, in which temperature

was ramped from 10°C to 85°C in steps of 4°C. At each temperature step the oven was held for 30 minutes after reaching the target temperature. This was to ensure that the resonator and the quartz thermometer (used to measure the oven temperature) were as close to the same temperature as possible. Fig. 1 shows the fundamental (or c1) mode f_{c1} and the 3rd (or c3) overtone f_{c3} versus the beat frequency f_β (defined by $f_\beta = 3f_{c1} - f_{c3}$). It can be seen that both the f_{c1} and f_{c3} curves are nonlinear, with the f_{c3} having a turnover at room temperature, so f_{c1} or f_{c3} alone would not be useful for measuring temperature when mass change is present or visa versa. The beat frequency (0.166 MHz), on the other hand, decreases monotonically and nearly linearly as temperature increases. The slope of the beat frequency was about 91 ppm/K or 15 Hz/K.

By using two counters controlled by a personal computer, both frequencies were simultaneously measured with 0.1 sec gate time and 0.06 sec dead time for data transfer and calculation. The hysteresis of the DMXO with the unopened resonator was measured to be 1.5×10^{-8}. Allan deviation [8] of the DMXO with the opened resonator was obtained to be 2.2×10^{-10} and 2.3×10^{-10} for f_{c1} and f_{c3}, respectively, at uncontrolled laboratory ambient temperature. The peak to peak noise of the mass frequency (or the resolution) of the DMXO with the opened resonator was 2 ppb (parts per billion).

Since the QCM was exposed to air, the gold electrode surfaces were quickly contaminated by hydrocarbons existing in the ambient air. Such organic contamination could be removed to a submonolayer level using an UV-ozone cleaning chamber. The UV/O_3 chamber was a commercially available model and its volume was ~2.7×10^4 cm^3. When the unit is turned on, the mercury lamp emits UV primarily at 185nm and 254nm. The 185nm radiation produces ozone from oxygen in the air and the UV and ozone work in concert to clean the gold electrode surfaces [9]. The resonator's temperature rose as a result of the UV photon flux (~30 mW), the transfer of heat from the lamp to the chamber air, and the exothermic reactions occurring in the chamber and at the surface. Typically the maximum temperature was 65°C. When the UV lamp was turned off, immediate temperature drop induced transient effects as 0.1~0.2 ppm changes of the mass frequency and contamination began to collect on the surface. After 30 minutes from the point the lamp was turned off, the surface contamination from the ambient completely blocked the sensing of injected organic vapor. The same amount of vapor could be easily detected if the UV lamp was left on continually cleaning the surface. To circumvent problems of surface contamination and the effects of the thermal transient, the UV lamp was kept on during the injection experiments. This procedure introduces some questions about the possibility of sensing UV-ozone decomposed molecules in the vapor, rather than intact sample molecules. Nevertheless, whatever was adsorbed, different injected molecules produced different responses, and, qualitatively, the responses varied with the molecular weights, which suggests that, with further development, one may be able to obtain not only high sensitivity but also selectivity. With better contamination control a clean sensor surface would be maintained for a longer time and experiments could be performed with the UV lamp off.

Results and Discussion

Fig. 2 shows the results for consecutive benzene vapor injections of 0.3 ml and 0.6 ml. Figs. 2(a)(b) are the raw data of f_{c1} and f_{c3}, respectively. Note that from the raw data it is not clear when the benzene was injected. In Fig. 2(c), however, by applying the algorithm that separates mass and temperature effects, the two benzene peaks from the injections are clearly identifiable. The units of mass frequency changes are shown in ppm.

In Fig. 2(d) the temperature is shown to be about 62.6°C with a precision on the order of thousandths of a degree. The majority of the frequency changes in Fig. 2(a) and Fig. 2(b) are due to temperature. When sensing small mass change, small drifts in temperature, which are unavoidable, can obscure the measurement of mass. The advantage of the SC-cut QCM in removing this effect is obvious here. Note that the temperature was far above the temperature that would be required for an AT-cut QCM to work properly without temperature control.

In Fig. 2(c) adsorption occurred in less than a half second. The adsorbed molecules were completely removed within about 5 minutes. This removal occurred in part by the UV-ozone cleaning process, but also by the lowering of the benzene vapor pressure due vapor phase UV-ozone reaction and by leakage out from the chamber as the chamber was not hermetically sealed. In a short time the clean surface was regenerated and this allowed for consecutive injections. The 0.3 ml benzene vapor was equivalent to a maximum of 1.3 ppm (by volume) benzene in the chamber. The mass frequency change for 0.3 ml injection was 13 ppb, which is six times greater than the balance's resolution of 2 ppb. The sensitivity is on the order of (2/13) × 1.3 ppm = 0.2 ppm (by volume) for benzene vapor. The microbalance with an unmodified gold surface is a better detector for benzene than the human nose, which has the odor detection threshold for benzene vapor of 1.5 ppm (by volume) [10]. In addition, the detection occurs instantly.

Fig. 3 shows the results for three consecutive hexane vapor injections of 0.6 ml, 0.4 ml, and 0.6 ml. Figs. 3(a)(b) are the raw data of f_{C1} and f_{C3}, respectively. In the raw data the injection points are identifiable. In Fig. 3(c), applying the separation algorithm of mass and temperature effects increased the signal to noise. As in the benzene case, the adsorption happened very quickly, but the adsorbed molecules were also completely removed in 2 minutes, which was faster than that of benzene. The 0.4 ml hexane vapor was equivalent to a maximum of 3.7 ppm (by volume) hexane in the chamber. The mass frequency change for 0.4 ml injection was 44 ppb, which is 22 times greater than the balance's resolution of 2 ppb. The sensitivity is on the order of (2/44) × 3.7 ppm = 0.17 ppm (by volume) for hexane vapor. The balance thus, has a detection limit of about 0.13% that of the 130 ppm (by volume)[10] odor detection threshold for hexane vapor. Again the detection happens very rapidly on the clean gold surface. Also shown is Fig. 3(d), the temperature was about 55.8 °C.

Fig. 4 shows the result of the injection of 0.01 ml of liquid methanol, ethanol, propanol, and butanol. Each liquid was injected into the chamber some distance away from the actual resonator and the liquid vaporized in the warm chamber. The adsorption of the alcohols was less than for benzene and hexane, so small liquid samples were used instead of vapors in order to increase the partial pressure of the vapors. The increased noise compared to the benzene and hexane experiments was due to the airflow in the chamber. The previous two experiments were performed with the resonator housed in a metal can with only the top electrode surface exposed to the lamp. This was to minimize air turbulence effects. The trade-off was that the exposed surface only was cleaned. In an effort to clean the both surfaces (*i.e.* to double the sensing area), the fully exposed resonator without the can was put in a position perpendicular to the lamp for the alcohol experiment. This caused the resonator to be more susceptible to the airflow. It is seen that the mass frequency changes are approximately proportional to the molecular weights of the alcohols.

The thickness of a 10 MHz 3rd overtone resonator is ~0.53 mm. This, with the density of quartz (2.65g/cm^3), gives an areal density of ~0.14g/cm^2. The microbalance has a sensitivity of ~140 ng/cm^2-ppm. The measurement resolution of 2 ppb represents 280 pg/cm^2. Since the electrode area was about 0.2 cm^2, the mass resolving ability is on the order of 56 pg (assuming a simple thin film model). This sensitivity can also be considered in terms of monolayers. When large quantities of hexane liquid were injected into the chamber with a fully exposed resonator (data not shown), the frequency mass change approached a saturation value of about 1.85 ppm. This change equates to 259 ng/cm^2 or 1.8×10^{15} molecules/cm^2 considering both electrodes and 0.9×10^{15}/cm^2 for each electrode. The 280 pg/cm^2 resolution equates to about 1×10^{12} hexane molecules on each gold electrode surface.

Summary and Conclusions

A microbalance utilizing dual mode excitation of an SC-cut resonator, has been demonstrated to sense concentrations of benzene and hexane below 1 ppm (by volume), by separating mass from temperature effects. The method improves the accuracy of a microbalance, especially when operating over wide temperature ranges. The method also eliminates the need for controlling the microbalance temperature. A mass change resolution of 56 pg has been demonstrated at elevated temperatures and under uncontrolled transient temperature conditions.

Acknowledgements

The authors wish to thank Mr. B. Rose of Q-tech for providing DMXO and Dr. J. R. Vig of US Army CECOM for helpful advice.

References

[1] D. S. Ballantine, R. M. White, S. J. Martin, A. J. Ricco, E. T. Zellers, G. C. Frye, and H. Wohltjen, Acoustic wave sensors, theory, design, and physico-chemical applications, Academic Press, N.Y., 1997.

[2] E. P. EerNisse, R. W. Ward and R. B. Wiggins, "Survey of quartz bulk resonator sensor technologies," IEEE Trans. Ultrason. Ferroelect. Freq. Contr., Vol. 35, pp. 323-330, 1988.

[3] G. G. Guilbault and J. M. Jordan, "Analytical uses of piezoelectric crystals: A review," CRC Critical Reviews in Analytical Chemistry, Vol. 19, Issue 1, pp. 1-28, 1988.

[4] J. J. McCallum, "Piezoelectric devices for mass and chemical measurements," Analyst, Vol. 114, pp. 1173- 1189, Oct. 1989.

[5] C. Lu and A. W. Czanderna, Eds., Applications of piezoelectric quartz crystal microbalances, Elsevier, New York, 1984.

[6] D. E. Pierce, Y. Kim, and J. R. Vig, "A temperature insensitive quartz microbalance," IEEE Trans. Ultrason. Ferroelect. Freq. Contr., Vol. 45, pp. 1238-1245, 1998.

[7] R. L. Filler and J. R. Vig, "Resonators for the microcomputer-compensated crystal oscillator," Proc. 43rd Ann. Symp. Freq. Contr., pp. 8-15, 1989.

[8] E. S. Ferre-Pikal, J. R. Vig, J. C. Camparo, L. S. Cutler, L. Maleki, W. J. Riley, S. R. Stein, C. Thomas, F. L. Walls, and J. D. White, "Draft revision of IEEE STD 1139-1988 standard definitions of physical quantities for fundamental frequency and time metrology-random instabilities," Proc. 1997 Int. Freq. Contr. Symp., pp. 338-357, 1997.

[9] J. R. Vig in, Treatise on Clean Surface Technology, edited by K. L. Mittal , Plenum, New York, 1997.

[10] United States Environmental Protection Agency's internet web site, http://www.epa.gov.

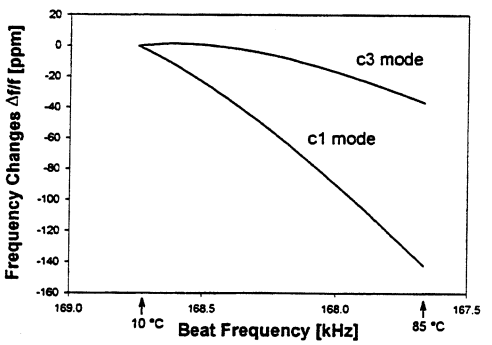


Fig.1: Frequency changes as a function of beat frequency for a dual mode SC-cut resonator oscillator.

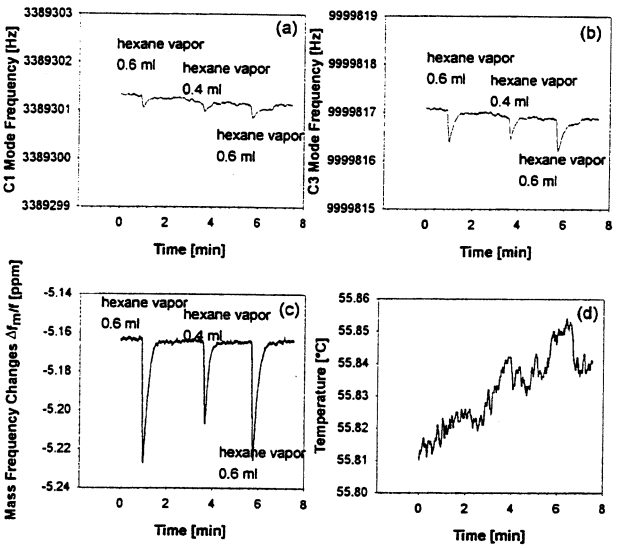


Fig. 3: Hexane vapor injection: (a) c1 mode, (b) c3 mode, (c) mass frequency changes, and (d) temperature.

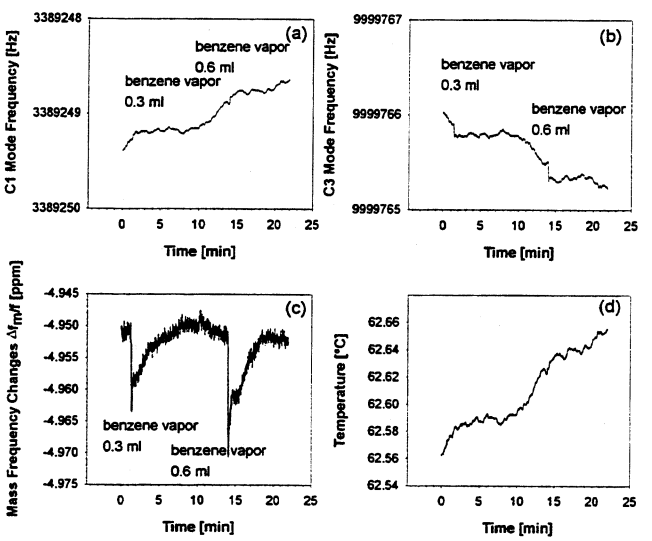


Fig. 2: Benzene vapor injection: (a) c1 mode, (b) c3 mode, (c) mass frequency changes, and (d) temperature.

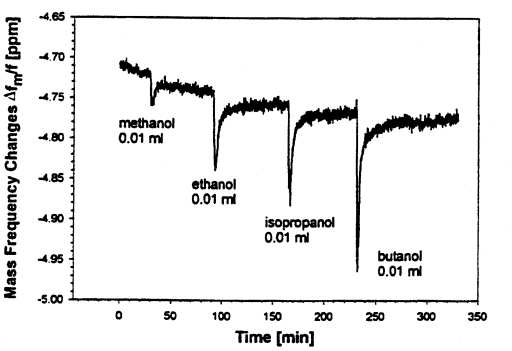


Fig. 4: Methanol, ethanol, isopropanol, and butanol injections.

LATEX PIEZOELECTRIC IMMUNOASSAY: IT'S APPLICATION FOR CLINICAL & ENVIRONMENTAL ANALYSIS

*[1,2]SHIGERU KUROSAWA, [3]MAKOTO MURATSUGU, [1]CHIKASHI NAKAMURA, [1]HIDENOBU AIZAWA, [2]NORIHIKO MINOURA, [1]JUN MIYAKE, [4]MINORU YOSHIMOTO and [5]NAOKI KAMO
[1]National Institute for Advanced Interdisciplinary Research (NAIR), 1-1-4 Higashi, Tsukuba 305-8562, Japan
[2]National Institute of Materials and Chemical Research (NIMC), 1-1 Higashi, Tsukuba 305-8565, Japan
[3]Osaka Prefectural College of Health Sciences, Habikino 585-8555, Japan
[4]Kagoshima University, Kagoshima 890-0065, Japan
[5]Hokkaido University, Sapporo 060-0812, Japan
*Corresponding author: Shigeru Kurosawa, TELFAX: +81-298-54-4746, E-mail: kurosawa@nimc.go.jp

ABSTRACT

Latex piezoelectric immunoassay (LPEIA) is a new immunoassay method that requires no immobilization of antigen or antibody on the electrode surface of quartz crystal, in contrast to previous immunoassays in which a piezoelectric crystal is used as a microbalance and immobilization is essential. The frequency change was observed during the aggregation of antibody- or antigen-coated latex particles. This method was used for the detection and is sufficiently sensitive for clinical applications. We make prototype LPEIA sensor for remote sensing and diagnosis to care on handicap person in home.

1. INTRODUCTION

An increase in the level of some proteins including some hormones, leukocytes, tissues and viruses in human blood or urine is an indicator of disease or infection. To monitor these levels, various good diagnostic techniques have been developed, among the most sensitive and selective of which involve immunoreactions. For clinical assays, the latex agglutination test is usually employed. When an antigen is added to a suspension of antibody-coated latex particles, immunoreaction causes agglutination of the particles. These reactions are measured in terms of the degree of light scattering or magnitude of change in absorbance of light. The method for optical measurement of the agglutination is called latex photometric immunoassay (LPIA). Many kits including antibody-coated or antigen-coated latex particles are commercially available for LPIA: some latex reagent kits for the determination of various substances such as C-reactive protein (CRP), antistreptolysin O (ASO), rheumatoid factor (RF), ferritin, immunoglobulins (IgG, IgA, IgM) and α-fetoprotein (AFP). LPIA is sensitive, precise, gives reproducible results, and is widely used clinically for serum testing [1]

A piezoelectric quartz crystal functions as a microbalance (QCM): adsorption on the surface of the crystal changes the oscillation frequency. The adsorption of about 1ng of material decreases the frequency by 1Hz when a 9 MHz of AT-cut quartz crystal is used [2]. Many attempts have been made to use such a highly sensitive, convenient and portable apparatus for immunoassay. In almost all studies, a thin film was formed on the quartz crystal surface to which antibodies or antigens were fixed, usually by chemical treatment [3]. In contrast, we developed a new immunoassay method, combining antibody or antigen-coated latex particles and a piezoelectric quartz crystal. This method does not require the formation of films on quartz crystal for the fixation of the antibody or antigen. We named this method as the latex piezoelectric immunoassay (**LPEIA**) [4]. We success to determine the diagnosis concentration of CRP and ASO in the human serum. We make prototype **CRP sensor 1** for remote sensing and diagnosis to care on handicap person in home and to use it on small clinical usage.

2. EXPERIMENTAL SECTION

2.1 Materials and Methods

Ultrapure water with a specific resistance of more than 18 MΩ/cm was prepared using Milli-Q (Millipore Ltd., Tokyo, Japan). Antigen- or antibody-coated latex particles such as ones coated with anti-CRP antibody (Seratestam CRP-H) and streptolysin O (Seratestam ASO-E) were gifts from Hitachi Chemical Co., Ltd (Tokyo, Japan). The latex to which the antigen or antibody was adsorbed was made of polystyrene with particle size of approximately 0.1 μm. CRP standard serum (Seratestam S CRP, 5.4 mg/dl), ASO standard serum (Seratestam S ASO, 204 IU/ml) and the serum of high ASO titer (1,040 IU/ml) were gifts from Hitachi Chemical Co., Ltd. (Tokyo). Normal human serum containing no CRP (with total protein concentration of 5.2 g/dl) was obtained from I.I.C. Japan and diluted 6-fold before use with 10 mM phosphate-buffered saline (PBS), pH 6.5, containing 5 % BSA, 16.5 mM NaN_3 and 135 mM NaCl. AT-cut piezoelectric quartz crystals (9MHz resonance frequency, 8×8×0.15 mm) were purchased from Yamagata Crystal Co. (Yamagata, Japan) and a silver electrode was deposited on each surface. In order to achieve stable oscillation in saline, one side of the crystal was sealed with silicon sealant (Sealant-45, Shin-Etsu Kagaku, Tokyo) as described previously [2]. An electronic circuit was made from TTL gates

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

(SN7400). The resistance and capacitance of this circuit were the same as those of circuit II described in a previous work [2]. The signal was fed to a universal counter (Model 7202, Iwatsu, Tokyo) and the data generated from changes in oscillation frequency were stored in a microcomputer (NEC PC9801, Tokyo).

3. RESULT AND DISCUSSIONS

3.1 Detection of CRP

We modified the original LPEIA for CRP (flow-cell [4]) to reduce the assay volume using a batch-cell [5]. The assay volume was used to 1,200 µl, and a one-side-sealed quartz crystal was placed in a small cuvette containing a small stirring bar. The entire apparatus except the microcomputer was accommodated in an air chamber maintained at 20±0.1 °C. The experimental procedure was essentially the same as the previous one [4]. The entire apparatus (except the microcomputer) was accommodated in an air chamber maintained at 20±1 °C, unless otherwise noted. The latex particles were suspended in PBS with the concentration being 5.56×10^{-3} % (100 µl of the stock latex suspension was added to 1,100 µl of PB). The suspension (1,200 µl) was agitated using magnetic stirrer bar. The frequency was stabilized in about 30 to 60 min and the stabilized oscillation frequencies were differed from one crystal to another. This variation may come from the one-side sealing procedure: the variation of amounts of silicon sealant that work as a foreign mass on the crystal changes the frequency. It is noted that the deflection of frequency by addition of samples was taken as response. Typical data on the frequency decrease caused by the addition of CRP (see Figure 3 on ref. [4]). At t=0, 50 µl of CRP stock solution (final concentration, 56 µg/dl) was added and a reduction in frequency (ca. 570 Hz) was observed. About 60 min after the addition, the frequency became constant and this frequency change (ΔF) was taken as the magnitude of response. The reproducibility was examined; ΔF upon addition of CRP (final concentration of 56 µg/dl and final latex concentration of 5.56×10^{-3} %) was measured for five independent crystals. The mean value and the standard deviation were 570 and 120 Hz, respectively, and the coefficient of variation was 29 %. As described above, sealing one side of the crystal was necessary for this prototype, and it seemed probable that this sealing altered the sensitivity, which varied from one crystal to another. Figure 1 shows the dependence of the frequency decrease on CRP concentration that is expressed in terms of -ΔF. A linear relationship was observed up to 56 µg/dl final concentration of CRP.

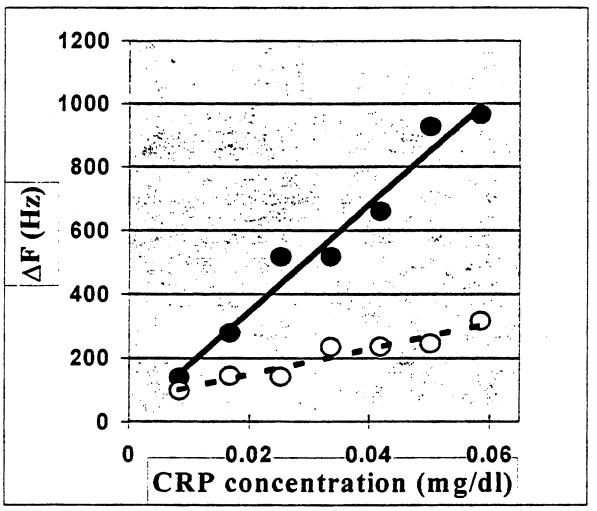


Fig. 1 Calibration curve for determination of CRP. Open circles show the data from serum without CRP. Filled circles show the data form serum with various concentration of CRP. Each data is the mean value of five times of experiment on different crystals (the coefficient of variation within 30 %).

Next, we investigated the specificity of frequency change. A frequency change of only 20 Hz was detected upon addition of 100 µl of normal human serum (essentially free from CRP). Upon addition of 200 µl of normal human serum, only 25 Hz was observed. These results show that the present system is specific for CRP whose antibody is immobilized on latex particles. Another control experiment was performed without any latex particles. Addition of CRP or normal human serum (100 µl) resulted in a change of 250 to 300 Hz. Comparison of this value with 20 Hz observed in the presence of the latex particles suggests that a large part of proteins added adsorbs to the latex particles. This adsorption to latex particle, however, does not evoke the frequency change and large changes in frequency are observed only for the proper combination of antigen and antibody. In another words, latex particle prevents the nonspecific adsorption of proteins to the crystal. In clinical analysis, serum CRP levels of 1,000 to 10,000 µg/dl usually needs to be monitored. Thus, the addition of 10 µl of the patient's serum to the present system should be adequate to give a working range of 10 to 100 µg/dl. Although the prototype LPEIA was successfully used to measure the concentration of the antigen, several modifications should be made, namely reduction of assay time and increase in ΔF. The magnitude of response increase with increase of latex concentration (data not shown). Therefore, some means of reducing the assay volume are necessary to reduce the amounts of latex required. About 60 min. was required for testing this system. It will be ideal to shorten this assay time.

3.2 Improvement of assay time: initial rate method

The ASO antibody test is the most valuable serologic test used to detect infection by group A β-hemolytic streptococci. Streptolysin O produced by most strains of the streptococci stimulates the production of a specific antibody, ASO. To reduce the assay time, we employed the rate method; the time-dependent frequency change was differentiated with time using computer software and the constant rate of frequency change was obtained. A one-side-sealed crystal was dipped in 1,200 μl of suspension containing SO-coated latex to which 20 μl of ASO serum (1,040 IU/ml) was added. After addition of the serum, the first derivative changed markedly and then became almost constant. We determined the initial rate of gradual frequency change as the slope of the solid line [4]. After the addition of serum, steady values of initial rate were obtained within 2 to 3 min, which represents a great reduction of the assay time. To investigate the possibility of clinical application, we used clinical specimens for measurement of ASO with LPEIA. After addition of each serum from 24 individuals was added to the SO-latex suspension, initial rate was determined and then compared with ASO values determined using the conventional turbidimetric agglutination method. The coefficient of correlation between initial rate and the ASO value was 0.950 ($P < 0.01$), indicating that the values obtained using LPEIA correlated well with those obtained using the conventional method.

3.3 Mechanism of the frequency change: increase with increasing sensitivity

Two factors can affect the oscillation frequency in solutions: one is ρη change of the solution, and the other is adsorption of molecules from the solution onto the surface. The agglutination may induce the magnitude of the ρη change and first this was considered to be a main reason for the frequency change, but the following observations may rule this out. After stabilization of the frequency change due to agglutination ($-\Delta F = 570$ Hz), the assay solution was gently removed from cell and fresh phosphate buffer was added. If the ρη change due to the agglutination was the main reason for the frequency change, this replacement is anticipated to restore the frequency value before the agglutination. This was not the case however: only a small frequency change of 20 Hz was observed. Thus, the binding of agglutinated latex particles is an important factor in the change. Also, the frequency change in LPEIA depended mostly on properties of the electrode surface. In addition, LPEIA responses were roughly proportional to the frequency change by the addition of latex particles before agglutination. These suggest that the surface properties that facilitate the binding increase the response. We previously found that the amount of antibody adsorbed is increased 4-fold when the substrate is coated with a plasma-polymerized allylamine [6, 7]. Such coating may increase the LPEIA response, and other surface modifications are investigated in progress.

3.4 Comments on further use of LPEIA

LPEIA is not limited to clinical use; it can also be used in other fields such as in the analysis of compounds of environmental interest. For determination of trace amounts of pesticides in water, soil, plant and food samples, a GC/MS or HPLC/MS method is generally used. Compared to these methods, the present method has some distinct advantages such as allowing more rapid screening of samples at portability and with no need for sophisticated laboratory equipment. Monitoring of endocrine disrupter exist environmental soil, water, plants, animal, and human is serious problem in Japan. Thus, construction of simplified sensing system for endocrine disrupter is emergence problem. Dioxin is one of the most widely yielded products from an incinerator of dust in the world. Due to the importance of this compound, a number of analytical methods have been developed to screen for and quantitative this chemicals in water and other environmental samples. Therefore, we are currently planning the feasibility of using anti-dioxin antibody coated latex for LPEIA to detect dioxin.

3.5 Development of portable LPEIA sensor

Quartz crystal microbalance is required for monitoring the levels of chemicals in the environments, especially for the field work. Figure 2 show the portable LPEIA sensor (laboratory made; **Prototype CRP sensor 1**). The equipment of LPEIA consist with a universal counter, a TTL circuit with quartz crystal and the measuring cell, whole size of which is less than 15×15×10cm and 1kg including AC power adapter. Quartz crystal microbalance costs the equivalent of only $1 in Japan. The cost of an entire experimental LPEIA system is only about **$200**. The portability and the low cost of this method make it suitable for analysis in field studies of pollutants.

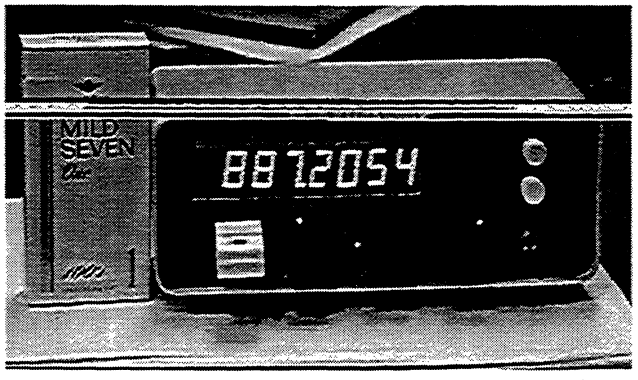


Fig. 2 The photograph of prototype of **CRP sensor 1**.

ACKNOWLEDGMENT

We express special thanks for the finantial support of research program on "Basic Research for Life and Society Project In Japan (FY1996-1998): Studies on a participation Support Technology for a Comfortable Life and Aged and Handicapped People/ Development of New Sensing Devices for the Health Management of Aged and Handicapped People/ Devices for Sensing Protein Makers in Blood".

REFERENCES

[1] S. Kurosawa, M. Muratsugu, H. O. Ghourchian and N. Kamo, ACS Sympo. Series 657: Immunochemical Technology for Environmental Application, New York: American Chemical Society, 1997, ch. 15, pp. 185-196.

[2] S. Kurosawa, E. Tawara, N. Kamo and Y. Kobatake, "Oscillating Frequency of Piezoelectric Quartz Crystal in Solutions," Anal. Chim. Acta, vol. 230, pp. 41-49, 1990.

[3] M. Muratsugu, F. Ohta, Y. Miya, T. Hosokawa, S. Kurosawa, N. Kamo and H. Ikeda, "Quartz Crystal Microbalance for the Detection of Microgram Quantities of Human Serum Albumin: Relationship between the Frequency Change and the Mass of Protein Adsorbed," Anal. Chem., vol. 65, pp. 2933-2937, 1993.

[4] S. Kurosawa, E. Tawara, N. Kamo, F. Ohta and T. Hosokawa, "Latex Piezoelectric Immunoassay: Detection of Agglutination of Antibody-Bearing Latex Using a Piezoelectric Quartz Crystal," Chem. Pharm. Bull., vol. 38, pp. 1117-1120, 1990.

[5] M. Muratsugu, S. Kurosawa and N. Kamo, "Detection of Antistreptolysin O Antibody: Application of an Initial Rate Method of Latex Piezoelectric Immunoassay," Anal. Chem., vol. 64, pp. 2483-2487, 1992.

[6] M. Muratsugu, S. Kurosawa and N. Kamo, "Adsorption and Desorption of F(ab')2 Anti-hIgG on Plasma-Polymerized Allylamine Thin Film: The Application of the Film to Immunoassay;" J. Colloid. Interface Sci., vol. 147, pp. 378-386, 1991.

[7] S. Kurosawa, N. Kamo, T. Arimura, A. Sekiya and M. Muratsugu, "Close-Packed Adsorption of F(ab')2 Fragment of Immunoglobulin G on Plasma-Polymerized Allylamine Film," Jpn. J. Appl. Phys., vol. 34, 7B, pp. 3925-3929, 1995.

1999 Joint Meeting EFTF - IEEE IFCS

Chemical Identification Using Internal Friction and Frequency Measurements on a Polymer Coated Resonator

R.A. Kant, C.L. Daly[1] and Huey-Daw Wu[2]
Naval Research Laboratory
Washington DC 20375-5345
Communications: kant@nrl.navy.mil

Abstract

We demonstrate the use of high resolution internal friction and frequency measurements to distinguish between different types of analytes by determining the loss per unit mass of absorbed analyte. In this study we determined the selectivity and sensitivity of this sensing technique using AT-cut, polymer coated, quartz crystal resonators. In addition, we report the mechanical response changes of the resonator due to coating alone and with an analyte present.

1. The Method

The technique we describe here takes advantage of the fact that an analyte changes both mass of a polymer absorber and its mechanical properties. In effect, our sensing technique determines the ratio of the change in the dissipation of mechanical energy to the mass of absorbed analyte. Since this ratio is peculiar to each analyte/film combination, we should be able to use it to identify an analyte using only one resonator.

Our instrument uses internal friction measurements to determine the changes in energy dissipation rate of a vibrating analyte/coating combination and it uses resonant frequency measurements to reflect the change in mass. The internal friction loss is the equivalent of the reciprocal of the quality factor Q of the composite resonator. For the measurements described here we determined the loss Q^{-1} as function of the analyte concentration in the atmosphere surrounding coated resonators while also measuring their resonant frequency.

In Principle, our technique is similar to several measurement techniques that are commonly used in circuit analysis to determine resonator properties including : the lumped circuit elements of a circuit that is equivalent a resonator[1], the S-parameters of a resonator[2], or the distributed parameters that describe an equivalent transmission line[3]. In practice, however, our method is relatively simple to interpret and easy to implement. One major advantage is that it avoids the expense and complexity of a network analyzer that is normally required to implement the above circuit analysis techniques.

2. Experimental Equipment

There are several established methods used to measure internal friction and the method of choice depends upon frequency and loss (Q^{-1}) of the system under study[4] The method we selected for these measurements is closely related to that which determines Q from the width of a resonance peak. We determine both loss and resonant frequency from a fourth order fit to the inverse square of the resonator current for several frequencies about a series resonance peak.

The entire measurement system is computer controlled. A sample chamber holds an array of quartz crystals at any temperature between 0 C and 110 C, with temperature variation of less than 0.01 C. In order to reduce the effects of thermal gradients the rate of temperature change is kept low during thermal cycling (less than 10 C per hour). A Microsensor Systems VG400 vapor generator supplies volume controlled concentrations of any one of 8 analytes. The VG400 uses 3 stages of duty cycle dilution to produce concentrations ranging from 10^{-6} of saturation to N_2 fully saturated with analyte at a flow rate of 100 standard cm^3/minute. We examined the response of the coated resonators to changes in vapor concentration and/or changes in temperature by exciting each resonator in an open loop configuration. An HP signal generator supplies drive signals at frequencies that span each crystal's resonance while a home-made detection circuit measures resonator current at a each frequency.

In practice, our array of resonators can consists of up to six crystals that we selected such that none of their resonance peaks overlap. Thus,

---

1 NRC Post Doc.
2 SFA, Inc., Largo, MD

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

because we operate in an open loop configuration, all crystals can be wired in parallel and then each crystal can still be interrogated individually by using a range of excitation frequency that affect only it.

Sample preparation: We used an aerosol spray method to deposit common chemically-selective polymer films on both sides of 3 MHz and 5 MHz AT-cut quartz crystals. The accumulated film thickness was monitored during deposition by measuring the resultant frequency shift and the final film thickness was nominally 900 nm. This thickness assumes that the density of each polymer was its bulk value and that all of the measured frequency shift was due to mass loading. The coatings used in this study were fluoropolyol (FPOL) and polyepichlorohydrin (PECH).

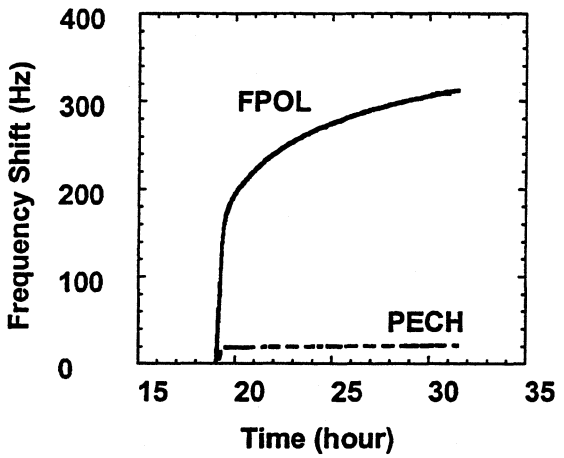


Figure 1. Frequency shifts of freshly coated resonators held at 55 C.

After coating, the resonate frequency of each composite resonator continues to shifts due to the release of the volatile solvents used in the aerosol spray. Baking the coated resonators under flowing nitrogen at temperatures ranging from 60C to 100 C served to increase rate of release of these solvents. Figure 1. shows that while the PECH films stabilized in less than 1 hour, the FPOL films were still not stable after more than 12 hours of baking. The results described below employed stable films.

3. Overview of the Experiments

The experiments consisted of three major phases. First, we measured the temperature dependence of loss and frequency for the bare crystals. Then we measured the temperature dependence of the loss and frequency signals for coated resonators with and without analyte exposure. For the above set of experiments, the temperatures ranged from 15 C to 55 C. The atmosphere in the sample chamber was either N_2 gas or N_2 plus an analyte at constant concentration. The temperature dependence of the loss and frequency shifts due to the coatings alone was obtained by subtracting the bare crystal data from that for the coated crystal. Similarly, the net effect of each analyte on each coating was obtained from the difference between the runs with the coated crystals with and without analyte. The analyte concentration for each run was constant. The concentrations were large enough to ensure reasonable signal to noise ratio and thus varied from analyte to analyte. The typical fraction of saturation ranged from about on 1/8 to 1/128. We used the data from these experiments to select the "optimum" temperature at which to measure the loss and frequency shifts as a function of analyte concentration.

4. Results

Because of space restriction, the results described in this paper are limited mostly to the case of the PECH coatings exposed to bischloroethylether (CEE). Additional details for this case and PECH and FPOL coating exposed to toluene and dimethyl-methylphosphonate (DMMP) will be published elsewhere.

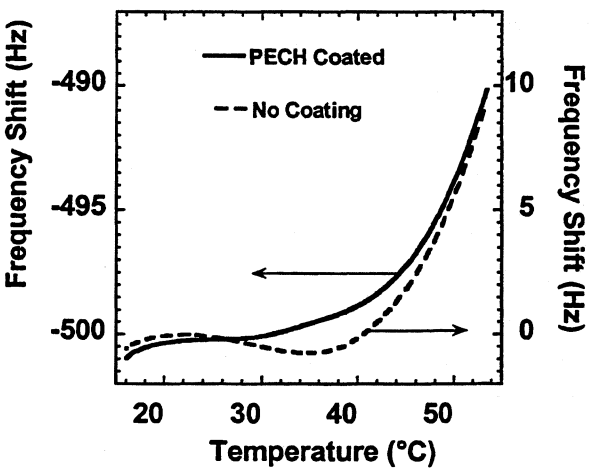


Figure 2 Effect of coating on the temperature dependence of frequency.

Figure 2. shows that the effect of adding a PECH coating on the temperature dependence of the resonant frequency is to nearly remove the peak near 23 C seen for the bare crystal. That peak corresponds to the temperature of zero temperature coefficient of frequency. The significance of this result is that a coating that is only 100 nm thick can shift or possibly remove, a point of stable operation (i.e., the temperature at which the resonant frequency is

insensitive to small temperature changes). This result has important implications for any resonant sensor system that derives its output from the difference between two resonators (one is typically coated while the other is not). Such common mode rejection techniques are used to cancel of environmental effects (such as temperature shifts) common to both crystals, but, to be useful the two crystals must have the same thermal characteristics. The result indicates that this may be difficult or impossible to achieve.

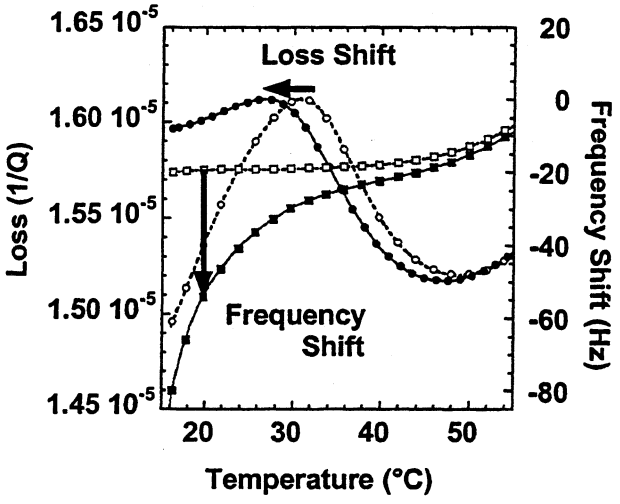


Figure 3. The effects of CEE on PECH are to decrease the resonant frequency and to shift the loss peak. Closed symbols indicate analyte and open indicate no analyte.

Figure 3 illustrates the effects of analytes on the loss and frequency shifts for a PECH coating exposed to CEE. CEE produces a large decrease in the resonant frequency at lower temperatures and the size of this downward shift decreased with increasing temperature. This temperature dependence of the resonant frequency is consistent with the expected increase in thermally-induced desorption of analyte with increasing temperature.

There is a clearly defined loss peak in figure 3. that shifts toward lower temperature when PECH is exposed to CEE. In addition, loss at the low temperature side of this peak increases significantly. At present, the source of the peak is under study. One possible source is a thickness resonance phenomenon in which the time for an acoustic wave to traverse the film is a multiple of the half the vibration period of the resonator. However, a preliminary examination suggests that such a thickness resonance would require physically unrealistically low shear modulus and/or large decrease in film density.

Figure 4. is a plot of the net loss and frequency response of the PECH film to three different analytes for temperatures ranging from 15 C to 55 C. The largest effects on frequency are the frequency decreases observed at the lowest temperature. It is reasonable to attribute most of this response to an increased mass loading. As noted above, this net increase in film mass is expected to be larger at low temperatures because of a corresponding drop in the rate of thermal desorption of analyte. This temperature dependence is consistent with the that of the partition coefficient for this case.

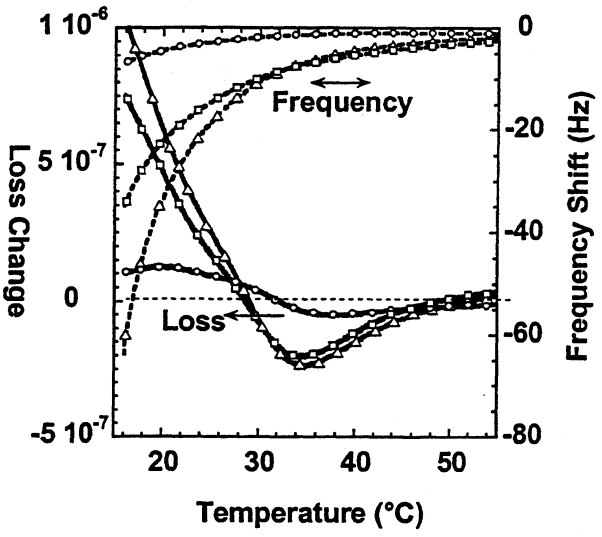


Figure 4. Loss and Frequency changes for a PECH coating due to each of three analytes. Dashed curves are frequency and solid curves are loss shifts. Circles indicate exposure to DMMP, squares to toluene, and triangles to CEE.

The net change in the loss curves is the result of a shift in the position of the loss peak and an increase in loss on the temperature side of the peak. The importance of these curves is that they provide the information needed to select an operating temperature which we can use to either enhance or discriminate against the response due to a particular analyte. Note, for example, that there is a temperature at which the loss response to each analyte is zero. Thus, if one operates a sensor at one of these temperatures, then the sensor will discriminate against the corresponding analyte while remaining sensitive to the others.

Figure 5 is a plot of the change in loss vs. change in frequency for two different coatings exposed to three different analytes at low concentration. This figure shows that the slope of each curve serves to clearly differentiate the response of each film to different analytes. Since mass loading causes the frequency shift, the horizontal axis might also be labeled mass increase. Thus the slope of each

curve is a measure of the loss per unit mass of absorbed analyte. Since, as we show here, each analyte/coating combination exhibits a different slope, this parameterization offers a simple and direct means to distinguish between analytes.

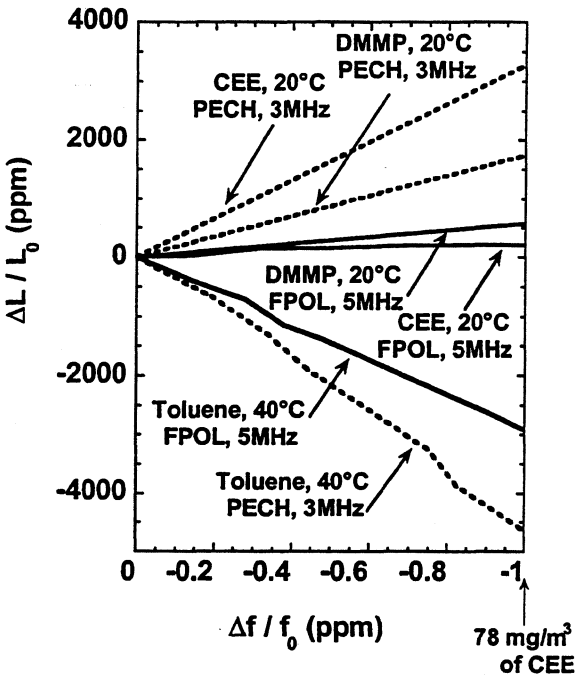


Figure 5. The relative loss change is plotted against the relative change in frequency for a collection of analyte/coating combinations.

It is significant that the change in loss is about four orders of magnitude larger than the corresponding frequency response for the same analyte exposure. Since the detection limits for loss change are nearly the same as for frequency shifts, this means that if significant improvements in loss measurements could be achieved, loss measurements could easily produce greater sensitivity than obtained with frequency based sensing.

5. Discussion and Conclusions

We have described a method for chemical sensing that uses a combination of internal friction loss and frequency shifts to both detect and identify analytes. The data we have presented show that we achieved chemically identification by adding internal friction loss measurement capabilities to conventional frequency-based chemical sensing technology. The data presented above clearly demonstrate that this method provides a way to distinguish between analytes. Moreover, although the absolute uncertainty in the loss signal is large, the fractional change in the loss signal is several orders of magnitude larger than the corresponding change in the frequency signal. The net result is that the detection sensitivity of loss shift is nearly as good as that of frequency shift. However, this work is still embryonic and there is ample opportunity to increase the sensitivity and selectivity of the technique. There is a large parameter space yet to be explored; including are about seven orders of magnitude of frequencies to be examined, there are different modes of vibration, and there is a practically unlimited variety of polymer and other films to be explored which can be expected to provide enhanced loss response.

We expect that a systematic study of the effects on loss due to the choice of frequency, vibration mode, intrinsic crystal loss, and coating technology will lead to significant improvements in loss-based sensors. For the experiments described here we were limited to common commercially prepared 3 MHz and 5 MHz quartz crystals that we coated with compounds that have already been used in resonant chemical sensors. These compounds had been selected for those sensors either without regard for their potential loss response or because they exhibited minimal "unexplained" frequency shifts. However, such unexplained frequency shifts might well indicate a strong response in our loss parameter. Therefore, work is currently underway at NRL to identify classes of polymers expected to exhibit large loss response.

References

1. Virgil E. Bottom, "Introduction to Quartz Crystal Unit design", p. 101-133, Van Nostrand Reinhold Company, New York, (1982).

2. Russell Chung, R. Andrew McGill, Paul Mathews, "Phase Noise Characterization of Polymer Coated-SAW Gas Sensors: Implications For the Performance of an Oscillator", Proc. 1997 *IEEE International Frequency Control Symposium,* 169-174, (1997)

3. R.W. Cernosek, S.J. Martin, A.R. Hillman, and H.L. Bandey, "Comparison of Lumped-element and Transmission-Line Models for thickness-Shear-Mode Quartz Resonator Sensors", *Proc. 1997 IEEE International Frequency Control Symposium*, 96-104

4. See, for example: A. S. Nowick and B. S. Berry, "Anelastic Relaxation in Crystalline Solids, New York Academic Press, 1972

1999 Joint Meeting EFTF - IEEE IFCS

WIRELESSLY INTERROGABLE ACOUSTIC SENSORS

F.SEIFERT[1], A. POHL[1], R STEINDL[1], L. REINDL[2], M.J. VELLEKOOP[3], B. JAKOBY[3]
[1]University of Technology 359/2, Vienna, Austria,[2]Siemens Corporate Research, Munich, Germany
[3]Delft University of Technology, DIMES, The Netherlands

A survey is given on wirelessly interrogable passive surface acoustic wave (SAW) sensors. Three types of SAW devices are discussed and classified as to their function in remote sensing of measurands (e.g. temperature, mechanical strain, stress ec.): Type I: transit time sensors, type II: SAW delay lines for wireless evaluation of remote conventional sensors and type III: SAW resonators. Efficient and successfully proven methods of wireless interrogation and evaluation are presented and some applications briefly mentioned.

1. INTRODUCTION

More than thirty years bulk wave and surface acoustic wave (SAW) devices are successfully applied to the processing of electrical signals. A large variety of electrical delay lines, filters and oscillators in the frequency range between some kHz and several GHz were designed and are widely in use now. Their superiority to other analog signal processing methods is based on two facts:
(a) Acoustic devices are robust in respect to temperature and mechanical stress, reliable over more than the lifetime of other electronic devices, and light weight in a small volume package. (b) Due to the fact, that sound waves in solid state materials are about five orders of magnitude slower than electromagnetic waves the signal processing capability, measured as the time bandwidth product TB can be high. Here T stands for the total electrical delay time of acoustical signal interaction, for example in the transversal filter structure of a SAW filter. $T.v_s$ is the sum of the length of input and output interdigital transducers (IDTs) on the piezoelectric filter substrate, v_s is the SAW velocity. The bandwidth B (more accurately speaking: the transfer function H(f)) is given for most devices mainly by the design of the acoustoelectrical coupling elements in the SAW filter i.e. by the geometry of the IDTs [1, 2]. The influence of temperature and mechanical stress (e.g. acceleration) on early SAW filters has been observed but the attempt to use these effects for sensor applications began in the seventieth [3] and lead to SAW filter sensors connected to signal generators and amplifiers by wires [4.5.6].
Wireless interrogation of reflective SAW devices was proposed 1975 [7] and resulted in identification (ID) applications [8,9,10] of passive SAW IDtags. Combining wireless readout with sensors in 1992, wirelessly interrogable passive SAW sensors were invented [11,12]. Since then different measurement methods for these devices were developed, among others, by the Corporate Research Laboratory Siemens Munich and our group in Vienna.The sensors incorporate the advantages of passive SAW devices given above in (a) and (b) with the applicability in measurement situations, where a wire or collector ring connection from the signal generating and evaluation unit to the measurement site is not appropriate and batteries or storage capacities must not be used.
For wireless measurements, the interrogation and evaluation unit (EU) generates a radio frequency (RF impulse) signal at the sensor's center frequency f_o. Over its antenna the EU radiates this signal to the sensor antenna, which electrically is connected to an IDT at one end of the piezoelectric sensor. The acousto-electrically excited SAW propagates along the sensor and is reflected by one or several reflectors on the sensor substrate. Thus a delayed electrical response signal is excited in the IDT. Over the connected sensor antenna the response is sent back to the EU's antenna and evaluated by the EU in terms of the measurand.
Following this general principle of measurement three different structures of sensors have been interogated by short (large bandwidth) or long (wide bandwidth) RF impulses:

- A type I sensor is shown in fig.1a. One IDT at L_o is connected to the sensor antenna AA'. The difference in transit time of the SAW signal between the reflectors at L_1, L_2 and L_3 is determined by the measurand (e.g. substrate temperature, mechanical strain ec.).
- A type II sensor is sketched in fig.1b. One IDT at L_o is connected to the antenna AA'. At least one of the reflective IDTs has electrical connections to 2 contact pins at the outside of the device's package. These pins are connected to conventional impedance sensors. Here a sensor is assumed which changes its capacitance by the measurand (e.g. geometrical position, compression length of a spring ec.) Inductive or resistive sensors can be implemented, too.
- A type III sensor is shown in fig.1c. It is a SAW resonator sensor which wirelessly can be excited over the center IDT connected to the antenna AA'. Similar to type I transit time sensors, measurands (e.g. temperature, mechanical strain) determine the resonance frequency f_o of the SAW resonator through the periodicity of the reflective arrays.

0-7803-5400-1/99/ $ 10.00

The next three chapters give a survey on the three sensor types and their interrogation methods. Chapter five presents sensor configurations with correlative signal processing. The paper concludes with a discussion of possible wireless acoustic sensors.

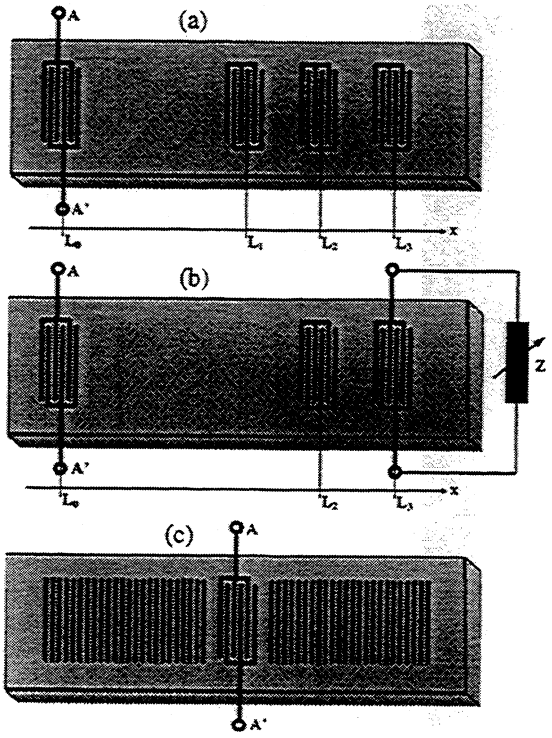


Fig.1 : Three types of SAW wireless sensors, drawn without absorbers at the edges. (a) type I transit time, (b) type II SAW delay line with external conventional senor, (c) resonator sensor

2. TYPE I: TRANSIT TIME SENSORS

If the sensor fig.1a is interrogated by a short impulse we observe the EU's receiver signal, sketched in fig.2. At $t = T_0$ the large interrogation impulse with length Δt and bandwidth $B \approx \Delta t^{-1}$ appears. As a response to this impulse we observe three smaller impulses at $T_1 = 2L_1/v_s$, $T_2 = 2L_2/v_s$ and $T_3 = 2L_3/v_s$. Any variation of length of sensor or of the SAW velocity v_s causes a variation of T_1, T_2 and T_3 which determines the measurand, (e.g. temperature, mechanical strain). We have not taken into account here the electromagnetic wave delay ΔT from the double travel between EU antenna and transponder sensor antenna. If the sensor is attached to a moving part (e.g. a rotating wheel) the time dependence of ΔT causes a measurement error, if the interrogating signal of the EU is taken as the reference time T_0. So, only in non moving measurement scenarios with stable radio channel for Type I sensors T_0 can be taken as a time reference and only one reflector (e.g. at L_1) is necessary. All other measurements have to evaluate the time difference T between T_1, T_2 and T_3.
It can be seen from fig.2 that the measurement accuracy depends on sensor design and signal processing:

- **Measurement timeT:** It should be as long as possible to allow the measurand to alter the time delay between the most distant reflections clearly observable (e.g. in fig.2 $T = T_3 - T_1$ should be chosen). Very long SAW subtrates are expensive in expensive packages, so for economical high accuracy long time delay sensor structures the acoustical path is bent by 180° using two 45° tilted reflectors, doubling T for one 180° bend.
- **Measurement bandwidth B:** It should be as wide as possible to produce short response impulses which accurately can be timed in the EU. For B sincere constraints exist: on the one hand standardized bandwidth in the usable industrial, scientific and measurement (ISM) RF-bands is modest. On the other hand IDTs with very few fingers for very short SAW impulses cannot be designed with good coupling efficiency. Short response impulses (e.g.< 50 ns) have to be traded off by large sensor insertion loss.
- **Advanced signal processing:** From the discussion above the difficulties in the determination of the difference time T from amplitude measurements can be seen. Especially for weak and time variable (fading) radio channels an exact time measurement between noisy response peaks is not possible.

Precise sensor readout cannot be compared with the binary decision in the EU for an ID tag (reflector: set or not set). In a road pricing device the first and last (e.g. 32.) reflector is set in all tags and used as a time frame for all other possible bit codes (e.g. of length 30). In order to achieve processing gain without summation in the time domain for the readout of ID tags the well known method of linearily frequency modulated continuous wave (FMCW) radar is applied. Here a slow interrogating FM chirp is multiplied with the acoustically delayed tag response (i.e. homodyne mixing). For a linear chirp with bandwidth B` and a sweep time T` an acoustical delay time step $(T_{n+1}-T_n)$ between two set reflectors yields a frequency step $(f_{n+1}-f_n)$ in the baseband output signal

$$f_{n+1}-f_n=(T_{n+1}-T_n).\ B`/T` \qquad (1)$$

So the IDtag code is evident from the frequency composition of baseband.
The FMCW radar method depends on the linearity of the chirp generator in the EU, which is a precondition of eq.(1) over the correlation time T´. A slight nonlinearity over the chirp length T´ would cause a shift in the frequency and a broadening of spectral lines in the baseband. To gain linear chirps over T´ a precise and stable clock and a coherent digital chirp generation in the EU is necessary. For sensor interrogation FMCW may be used also, if the measurand keeps a constant value over T´, which is in the range of ms [13].

A precise and fast method to read out type I sensors is sampling of response signal (fig. 2) and coherent digital evaluation of difference time between response peaks. Here too, a precise and time stable clock is required in

the EU which coherently generates the carrier of a short (typically 100 ns) RF interrogation gated impulse. The response impulses at T_1, T_2, T_3 are defined from amplitude but evaluated in terms of the phase of the continuous EU carrier. Digitally the phase is determined at the maxima of the response impulses by down conversion of the response signal as usual into an inphase (I) and quadrature (Q) signal. A short time resolution of about 3 degrees phase angle has been achieved in measurements by A.POHL [14,15,16] which corresponds to a time resolution of about 10ps at f_o = 860 MHz carrier frequency. But it should be mentioned here that such excellent results demand the stability of EU clock over a time interval which is long compared to the observed variation of measurand. Also integrative digital methods can be applied if several measurements can be averaged to one result.

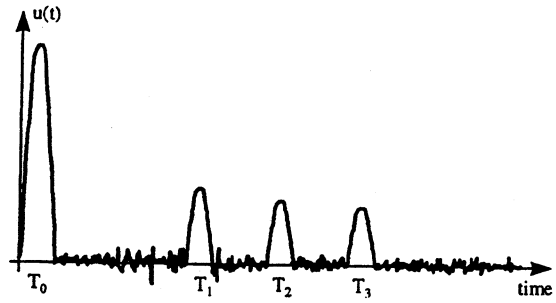


Fig.2 : Impulse response of type I sensor of fig. 1a.

In completion to this broadband short impulse interrogation of transit time sensors a narrow band accurate evaluation method was invented also by A.POHL [17,18]. Here a readout signal, about of length T_1 in fig.2 is used. In order to avoid electromagnetic pollution over a large bandwidth its transients may be flattened by time domain filtering. This interrogation impulse together with the response signal due to the first two reflectors in fig.1a is sketched in fig.3. Starting at T_1 the reflection at L_1 is observed . From T_2 to T_3 the reflection at L_2 is added and causes (destructive in fig.3)

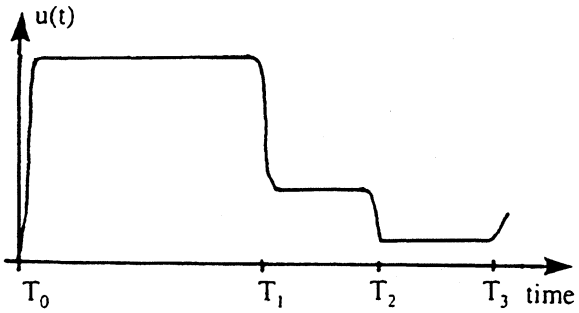


Fig.3 : Narrowband interrogation and destructive intersymbol interference response of type I sensor.

interference whose amplitude allows to determine the relative phase of the SAW propagating between L_1 and L_2. So e.g. by sweeping the interrogation frequency and observing the frequency of response notches between T_2 and T_3 in fig.3 the measurand being effective between L_1 and L_2 is determined. This method of measurement yields high definition but requests much more time than a sampling one shot evaluation.

3. TYPE II: EXTERNAL CONVENTIONAL SENSOR ACOUSTICAL DELAY DEVICES

The invention of these wirelessly interrogable sensors resulted from the investigation of electrically programmable ID tag reflectors in railroad applications by L.REINDL and W.RUILE [19,20]. Usually in ID tags the reflecting IDTs are not accessible. But connecting a load admittance Y_L to the IDT changes its SAW reflection coefficient ρ given by :

$$\rho = \rho_{SC} + \frac{P_{13}^2}{Y_{IDT} + Y_L} \qquad (2)$$

Here ρ_{SC} and Y_{IDT} are the acoustical reflection coefficient for short cut IDT and its admittance, respectivly. P_{13} is the acoustoelectrical interaction (coupling element) expressed in the acoustoelectric scattering P matrix.
So, by switching the busbars of a reflecting IDT over connecting pins to an appropriate capacitor or varactor diode the IDT ρ goes from strong SAW reflection in open circuit to nearly total transmission with ρ = 0. This happens if the IDT electrically sees a short, and no acoustoelectrically excited voltage is effective at its fingers. The electrical short is given for series resonances at f_o of the outside capacitor with its inductive wire connections to the IDT.

At present R.STEINDL [21] is exploiting this resonance behaviour for measurements of external conventional sensors. If an external variable capacity as a sensor is biased near the resonance value, a small deviation of measurand from its bias causes a large change (more than 30 dB) of the reflectivity of the IDT which easily can be evaluated. For this the sampling method as discussed above for type I sensors can be modified taking amplitude and phase data simultaneously. In the acoustic delay line together with the sensor connected IDT at least one unconnected IDT has to be present as a reference of reflection amplitude and phase (This reference doesn't change since the SAW substrate is thermally and mechanically stable.) External mechanical sensors show decisive advantages compared to mechanical SAW sensors, where mechanical stress is applied to the substrate. Mechanical overload (e.g. in high speed tire profile deformation measurements [14]) sometimes destroys the SAW substrate, but capacitive or inductive conventional sensors appropriately designed can withstand all overload with the SAW

delay line in a robust, hermetically tight package. Also resistive (potentiometric) external sensors could match this condition but in wireless applications with type II SAW delay lines their sensitivity and accuracy is poor due to the lack of series resonance behaviour.

4. TYPE III: SAW RESONATOR SENSORS

SAW resonators on ST-cut quartz substrates widely are used in clock circuits above 100 MHz as temperature stable frequency reference. Under mechanical stress their resonance frequency f_0 changes with the same sensitivity like type I transit time sensors of the same material and crystal cut. So the use of SAW resonators in sensor and ID devices is indicated.

Connecting the input IDT of a one port resonator to an antenna, as shown in fig. 1c, two principles of wireless resonance frequency f_0 interrogation are applicable :

a) In a short distance to the EU antenna, where the sensor antenna is in its near field, the electromagnetic energy feedback makes possible the swept frequency measurement of f_0 by the EU's antenna radiation admittance. At f_0 a sharp dip in the admittance appears.

b) In the far field of the EU antenna the admittance of the sensor antenna cannot be detected. Here, the passive resonator acts as a frequency selective energy storage. By interrogation near f_0 with a sufficiently long (some resonator decay times) narrow band impulse or a strong and short one oscillation energy is stored. After switching off the pump impulse the resonator decays with its natural frequency f_0.

Since the objectives are long range measurements method b) is used. After a comprehensive discussion of wireless SAW applications, W.BUFF et al. present a straight-forward method of resonance frequency determination within an accuracy of 10 kHz at 189,2 MHz for a $LiNbO_3$ temperature sensor [22,23] by sweeping the excitation frequency manually, the amplitude of the response decay is maximized and so the resonance f_0 found.

We employ an adapative search of the EU interrogation impulse frequency [24]. During the resonator's decay the EU receives its ring out frequency f_0 over a gain controlled logarithmic amplifier. The interrogation impulse frequency is adaptively synchronized and at last locked to f_0. This lock follows f_0 for variations of the measurand. Since resonance sensors need minimal bandwidth for interrogation and response signal, several resonators can be placed in one ISM band. This makes them useful as ID tags with a few bits. Another feature of resonator sensors is the relatively low insertion loss compared to type I and II sensors. Thus a maximal interrogation range R can be achieved with type III sensors. E.g. with 20 dBm (100 mW) interrogation pulse power, and quarterwave dipole antennas an R = 10 m is possible.

5. ADVANCED SENSORS AND SYSTEMS

A unique feature of SAW devices, applied from the late sixtieth up to now is the capability of pulse compression in chirped or coded phase (PSK) and/or amplitude modulated IDT structures. These structures can be employed to improve sensitivity and accuracy of type I sensors and to design sensor families where each member can be electronically interrogated separately [12,22].

Using a chirp structure, from the time shift of the amplitude of the compressed impulse, enhanced by the factor TB, a resolution of 1 mK of a $LiNbO_3$ SAW thermometer could be achieved experimentally [25].

Industrial applications were the objective of the development of an all quartz package gas pressure gauge. It started with the PHD thesis of H.SCHERR [26], which was improved by L. REINDL, C.RUPPEL et al. in Munich [13]. In fig. 4 a cross-section of the quartz pressure gauge is sketched. The cavity with hight $d_{SL} \approx 30$ µm can be filled by an inert gas at reference pressure or can be evacuated. The IDTs are positioned along bending lines of the elastic membrane, the SAW delay between them is calibrated in terms of outside gas pressure. Busbars for antenna connections are extended through the adhesive to the outside. Using an elaborate design technique for chirp transducers [27] the insertion loss and sensor qualitiy of the intracavity SAW delay lines could be improved recently [13]. Using two IDTs with the same dispersion a linear phase was achieved. This gauge was used to monitor the tire pressure of cars up to 100 km/h [28].

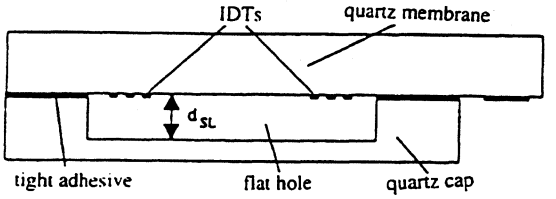


Fig.4: All quartz package SAW pressure gauge

Employing a specific arrangement of three SAW resonators on a 35° y cut quartz pressure gauge membrane in the 434 MHz ISM band and evaluating their difference frequencies W. BUFF et al. were able to measure either pressure or temperature or both simultaneously [23].

Some attempts have been made by OSTERMAYER to simultaneously selectively evaluate families of ID-tag type I sensors with coded response [29,30,31]. These are based on the principle of code division multiple access (CDMA) systems. If the adressing interrogation signal is exactly the time inverse of the sensor's impulse response the sensor will answer with its autocorrelation function (ACF). By changing the time scale of the interrogation signal the maximum of the ACF indicates

the effective measurand of the adressed sensor. 13 bits Barker codes were used. The applicability of this CDMA system is restricted to equal distance sensors. since the radio channel for this passive SAW sensor as transponder is damped with at least the fourth power of distance. But for single sensor readout in an electromagnetically polluted area these correlative methods bring considerable advantages [29,33].

6. CONCLUSIONS AND OUTLOOK

A survey on wirelessly requestable SAW sensors was given in a limited range of applications, mainly for temperature and mechanical measurands with the three types of sensors under investigation at present. The feasability of RF sensor interrogation by a EU, storing or delaying the signal acoustically and retransmitting the sensor modulated signal to the EU for evaluating is proven for many different situations.

But this principle is applicable to many more acoustical devices. Bulk acoustic wave resonator [34] would have sufficiently long time constants of decay, but hardly are fabricable for resonance frequencies above 100 MHz (λ = 3 m), thus dipole antennas for remotely placed bulk wave devices would be to big. Eventually small ferrites with coils could be tried as sensor antenna.

More hope to find new areas of application lies in wireless request adaptations of SAW and Love wave sensors for chemical gas analyses and liquids as developed by A.VENEMA, M.VELLEKOOP and B.JAKOBY in Delft [35,36,37] . In fig. 5 a Love wave device for wireless read out is sketched. The main energy of the LOVE mode is guided in a guiding layer, IDTs are used for excitation and reflection. Recent results show the applicability of these modes for both, viscosity sensing and low loss propagation in a frequency range suitable for wireless request. Gas sensing and chemical analysis is possible with specific layers. Taking into account the additional losses, a quantitative model of wireless read out has to be used, as given by A.POHL in a following contribution [38].

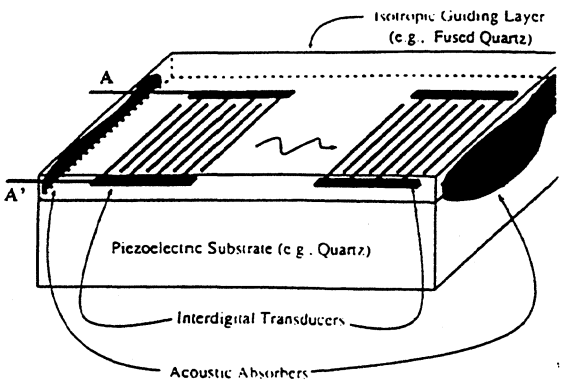


Fig.5: Reflective Love wave device. [37].

ACKNOWLEDGEMENT

We thank Siemens Corporate Research ZT KM1, namely E.BULST and C.C.W. RUPPEL and all former leaders of the SAW group for making possible our SAW research since 1983.

REFERENCES

[1] D.P. Morgan, *Surface Wave Devices for Signal Processing*, Elsevier, Amsterdam, 1985.

[2] C.C.W. Ruppel, R. Dill, A. Fischerauer, G. Fischerauer, W. Gawlik, J. Machini, F. Müller, L. Reindl, W. Ruile, G. Scholl, I. Schnapp and K.Ch. Wagner, SAW devices for consumer communication applications, *IEEE Trans. Ultrason. Ferroelectr. Freq. Control, UFFC-40*(5) (1993) pp.438-452.

[3] T.M. Reeder and D.E. Cullen, Surface acoustic wave pressure and temperature sensors, *Proc. IEEE, 64* (5) (1978) pp.754-756.

[4] D. Hauden, S. Rousseau, G. Jaillet and R. Coquerel, Pressure and temperature measurements with SAW sensors, *Proc. 36th IEEE Ann. Freq. Control Symp., 1982,* pp. 284-289.

[5] R.M. White, Surface acoustic wave sensors, *Proc. IEEE Ultrasonics Symp. San Francisco, CA, USA, 1985,* pp. 490-494.

[6] A. Venema, E. Nieuwkoop, M.L. Vellekoop, W.J. Gnijsen, A.W. Barendsz and M.S.Nieuwenhuizen, NO_2 gas concentration measurement with a SAW-chemosensor, *IEEE Trans. Ultrason. Ferroelectr. Freq. Control, UFFC*-34 (2) (1978) pp.148-155.

[7] D.E.N. Davies, M.J. Withers and R.P. Claydon, Passive coded transponder using an acoustic-surface-wave delay line, *Electron. Lett., 11* (Apr.) (1975) pp.163-164.

[8] P.A. Nysen, H. Skeie and D. Armstrong, System for interrogating a passive transponder carrying phase-encoded information, *US Patent Nos. 4 725 841; 4 625 207; 4 625 208*; (1983-1986).

[9] A. Wolfe, SAW filters replace bar codes in ID systems, *Electronics Week,* (May 1985) p.28.

[10] Editorial Technology Focus: Norway implements r.f. vehicle identification, *Microwave RF Eng.*, (Feb. 1989) pp.42-43.

[11] F. Schmidt, O. Scesny, L. Reindl, V.Mágori, Remote Sensing of Physical Parameters by Means of Passive Surface Acoustic Wave Devices (ID-TAG), *Proc.IEEE Ultrasonics Symp.* 1994, pp.589-592.

[12] F. Seifert, W.E. Bulst, C.C.W. Ruppel, Mechanical sensors based on surface acoustic waves, *Sensors and Actuators, A 44* (1994) pp.231-239.

[13] L. Reindl, C.C.W. Ruppel, K. Riek, T. Pankratz, R. Weigel, A wireless AQP Pressure sensor using chirped SAW delay line structures, *Proc.IEEE Ultrasonics Symp.*, Sendai 1998, E3 in print.
[14] A. Pohl, F. Seifert, Wirelessly Interrogable SAW Sensors for Vehicular Applications, *IEEE Trans. on Instrumentation and Measurements*, Vol.46 (1997), No.4, pp.1031-1038.

[15] A. Pohl, A. Springer, L. Reindl, F. Seifert, R. Weigel, New Applications of Wirelessly Interrogable Passive SAW Sensors, *IEEE Trans.on Microw.Theory and Technique.* Vol. 46 (1998) No.12, pp.2208-2212.

[16] A. Pohl, R. Steindl, L. Reindl, F. Seifert, Wirelessly Interrogable Sensors for Different Purposes in Industrial Radio Channel, *Proc. IEEE Ultrasonics Symp.*, Sendai 1998, E1 in print.
[17] A. Pohl, G. Ostermayer, L. Reindl, F. Seifert, Notch Sensors – A New Signal Processing Method for Interrogation of Passive SAW Sensors, *Proc. IEEE Ultrasonics Symp.*, Toronto, 1997, pp.355-358.

[18] A. Pohl, A Low Cost High Definition Wireless Sensor System Utilizing Intersymbol Interference, *IEEE Trans. on Ultrasonics, Ferroelectrics and Frequency Control*, Vol.45, No.5, Sept.1998, pp.1355-1362.

[19] L. Reindl. W. Ruile, Programmable Reflectors for SAW-ID-Tags, *Proc. IEEE Ultrasonics Symp.* 1993, Vol. pp.125-130.
[20] L. Reindl, T. Ostertag, W.. Ruile, C.C.W. Ruppel, A. Lauper, R. Bächtiger, H. Ernst, Hybrid SAW-Device for a European Train Control System, *Proc.IEEE Ultrasonics Symp.* 1994, pp.175-178.
[21] R. Steindl, A. Pohl, F. Seifert, SAW Delay Lines for Wirelessly Requestable Conventional Sensors, *Proc. IEEE Ultrasonics Symp.*, Sendai 1998, E2 in print.

[22] W. Buff, F. Plath, O. Schmeckebier, M. Rusko, T. Vandahl, H. Luck, F. Möller, D.C. Malocha, Remote Sensor System Using Passive SAW Sensors, *Proc.IEEE Ultrasonics Symp.* 1994, pp.585-588.
[23] W. Buff, M. Rusko, M. Goroll, J. Ehrenpfordt, T. Vandahl, Universal Pressure and Temperature SAW Sensor for Wireless Applications, *Proc.IEEE Ultrasonics Symp.* 1997, pp.359-362.

[24] A. Pohl, G. Ostermayer, F. Seifert, Wirelessly Sensing Using Oscillator Circuits Locked to Remote High-Q SAW Resonators, *IEEE Trans. on Ultrasonics, Ferroelectrics and Frequency Control.*, Vol.45, No.5, Sept.1998, pp.1161-1168.

[25] F. Seifert, A. Pohl, G. Ostermayer, G. Berger, Wireless Sensors and Data Links Based on SAW Devices, *Proc. of the IEEE MTT/AP and UFFC Chapter Workshop* 1997 (invited), pp.27-34.

[26] H. Scherr, G. Scholl, F. Seifert, R. Weigel, Quartz Pressure Sensor Based on SAW Reflective Delay Line, *Proc. IEEE Ultrasonics Symp.* 1996, pp.363-367.

[27] C.C.W. Ruppel, L. Reindl, K.Ch. Wagner, Optimum Design of Low Time-Bandwidth Product SAW Filters, *Proc. IEEE Ultrasonics Symp.* 1994, pp.61-64.
[28] A. Pohl, G.Ostermayer, L. Reindl, F. Seifert, Monitoring the Tire Pressure at Cars Using Passive SAW Sensors, *Proc. IEEE Ultrasonics Symp.*, Toronto, 1997, pp.471-474.
[29] A. Pohl, F.Seifert, L. Reindl, G. Scholl, T. Ostertag, W. Pietsch, Radio Signals for SAW ID-Tags and Sensors in Strong Electromagnetic Interference, *Proc. IEEE Ultrasonics Symp.*, 1994, pp.195-198.

[30] G. Ostermayer, A. Pohl, C. Hausleitner, L. Reindl, F.Seifert, CDMA for Wireless SAW Sensor Applications, *Proc. IEEE International Symp. on Spread Spectrum Techniques & Applications, ISSTA* 1996, pp.795-799.
[31] G. Ostermayer, A. Pohl, L. Reindl, F. Seifert, Multiple Access to SAW Sensors Using Matched Filter Properties, , *Proc. IEEE Ultrasonics Symp.*, Toronto, 1997, pp.339-342.
[32] G. Ostermayer, A. Pohl, R. Steindl, F. Seifert, SAW Sensors and Correlative Signal Processing – a Method Providing Multiple Access Capability, *Proc. ISSTA* 98, South Africa, pp.902-905.

[33] A. Pohl, G. Ostermayer, C. Hausleitner, F. Seifert, L. Reindl, Wavelet Transform with a SAW Convolver for Sensor Application, *Proc. IEEE Ultrasonics Symp.* 1995, pp.143-146.
[34] E. Benes, M. Gröschl, F. Seifert, A. Pohl, Comparison between BAW and SAW Sensor Principles, *IEEE Trans. on. Ultrasonics, Ferroelectrics and Frequency Control.* Vol. 45, No. 5, Sept.1998, pp.1314-1330.

[35] A. Venema, E. Nieuwkoop, M.J. Vellekoop, W.J. Ghijsen, A.W. Barendsz, M.S. Nieuwenhuizen, NO_2 Gas-Concentration Measurement with a SAW-Chemosensor, *IEEE Trans. on Ultrasonics, Ferroelectrics and Frequency control*, Vol. UFFC-34, No.2, March 1987.
[36] B. Jakoby, M.J.Vellekoop, Vicosity sensing using a Love-wave device, *Sensors and Actuators A68*, 1998, pp. 275-281.
[37] B. Jakoby, M.J. Vellekoop, Properties of Love waves: applications in sensors, *Smart Mater. Struct.6*, 1997, pp.668-679.

[38] A. Pohl, State of the art in signal processing for wireless SAW sensing, *Proc. Joint Meeting 13th Europ. Frequ. and Time Forum, 1999 IEEE Internat. Frequ. Control Symp.*, this issue.

1999 Joint Meeting EFTF - IEEE IFCS

LONG-TERM STABILITY AND PERFORMANCE CHARACTERISTICS OF CRYSTAL QUARTZ GAUGE AT HIGH PRESSURES AND TEMPERATURES

N. MATSUMOTO, Y. SUDO, B. SINHA* and M. NIWA

Schlumberger K.K. Fuchinobe 2-2-1, Sagamihara-shi, Kanagawa-ken 229-0006, Japan
*Schlumberger-Doll Research, Old Quarry Road, Ridgefield, CT 06877-4108, U.S.A.

ABSTRACT

Good long-term stability of high precision quartz pressure sensors are necessary for various applications ranging from pressure transient analysis to permanent monitoring systems for optimal reservoir management in the petroleum industry. A Crystal Quartz Gauge (CQG) is a dual-mode thickness-shear quartz pressure sensor that has been used in the oil-field services for the past eight years. High accuracy, resolution, and fast response time of this sensor enable a reliable estimate of formation permeability and helps in reducing the overall cost of oil and gas production. Production statistics of manufactured CQGs show its yield exceeding 85% with excellent performance repeatability. The pressure reading errors of manufactured gauges are less than 6.89 kPa (1 psi) (plus 0.01% of the reading because of the uncertainty of the dead-weight tester). The pressure resolution is better than 20.7 Pa (0.003 psi) over 1 second gate time. An extremely effective dynamic compensation algorithm yields corrected pressure readings with a response time as short as a strain-gauge based pressure transducer while retaining high performance of a quartz gauge. Recent long-term stability tests of CQGs show a negligibly small drift on the order of a few-tenths of a psi (0.1 psi = 689 Pa) at 103 MPa (15 kpsi) and 175 °C for a period of more than a year. These results confirm that the CQG characteristics exceed the demanding specifications for both the well-tests and permanent monitoring systems.

I. INTRODUCTION

High precision pressure and temperature sensors are, generally, made of crystalline quartz because of its excellent stability under temperature and pressure cyclings; high material Q; and minimal hysteresis under loading and unloading [1-2]. While the mechanical strength and stability of the material are critically dependent on its surface preparation, such as mechanical polishing and chemical etching [3], it is known that crystalline quartz is approximately 24 times stronger in compression than in tension. Given this mechanical property of quartz, it is desirable to design probe structure and sensing element so that the entire structure is largely in compression.

A thickness-shear quartz resonator frequency changes as a function of biasing stresses and temperature gradients in its vibrating element [1-2, 4-8]. Different crystalline orientations of quartz exhibit significantly different stress-frequency and temperature-frequency characteristics. A dual-mode thickness-shear quartz pressure sensor is designed with a crystalline orientation such that the B-mode is largely sensitive to the temperature and the C-mode is largely sensitive to the biasing stresses of the resonator element [4-8]. These characteristics of the two thickness-shear modes enable the inversion of changes in the two thickness-shear resonant frequencies for the biasing stresses and temperature of the resonator element. The biasing stresses in the resonator element are caused by the applied hydrostatic pressure to the exterior of a probe structure. The probe structure design allows some optimization of the transformation ratio of applied pressure to the stresses produced in the resonator element.

Figure 1 shows a schematic diagram of the CQG probe structure. This probe structure consists of three cylindrical parts: the central portion has a rectangular plate that is an integral part of the cylindrical shell; and the two end caps that have hollowed interior and are hermetically sealed to the central portion. The rectangular plate is the sensing element of the probe structure. A doubly-rotated orientation of crystalline quartz is used to design a dual-mode thickness shear resonator with different pressure and temperature characteristics for the two modes.

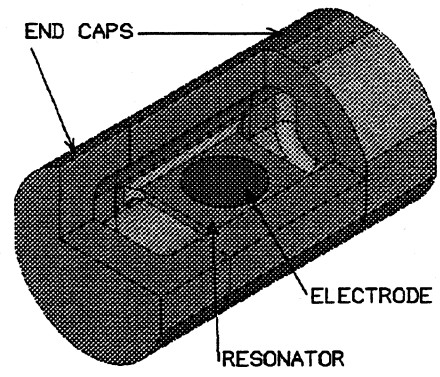


Figure 1: Schematic diagram of the CQG

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

II. MANUFACTURING STATISTICS

The pressure sensor is calibrated over a pressure range of 750 psi (5.17 MPa) to 15 kpsi (103 MPa) and a temperature range of 35°C to 175°C. The lowest pressure is set by the available mass of the dead-weight tester. Under static conditions, both the externally applied pressure P_s and probe temperature T_s can be retrieved from two calibration polynomials with two arguments as shown below.

$$P_S = \sum_{i,j=0}^{n} A_{ij} f^i{}_P f^j{}_T \quad (1)$$

$$T_S = \sum_{i,j=0}^{n} B_{ij} f^i{}_P f^j{}_T \quad (2)$$

where f_P and f_T are the frequencies of the primarily pressure sensitive (C-mode) and primarily temperature sensitive (B-mode), respectively. Both coefficients A_{ij} and B_{ij} have 15 nonzero elements. The largest exponent n of f_P is chosen to be 4 for P_S, and 3 for T_S; while that of f_T is chosen to be 5 for both P_s and T_s.

In Figure 2 we show average calibration or fitting errors for both the load and unload cycles for 404 manufactured CQGs at four temperatures of 35°, 82°, 129°, and 175°C, respectively. The error bars defined by 3σ, (where σ is the standard deviation) is always less than 1 psi. The ppm scale refers to the fractional deviation of the pressure reading from the maximum pressure range of 103 MPa (15 kpsi).

Figure 3 shows the Allan variance [9] of pressure reading of 61 gauges that were calibrated with 10-digit frequency counters to see the short term stability of the gauges more accurately. This test was performed at atmospheric pressure and at various temperatures. A rather small variance indicates the mean pressure stability to be better than 0.003 psi.

III. DYNAMIC PERFORMANCES

Pressure transient analyses are routinely used for estimating reservoir permeability. In highly permeable reservoirs, it becomes necessary to detect a small duration pressure pulse of small magnitude. This requires a pressure sensor with a fast response time in the draw-down and build-up analyses of the pressure transient tests. Since a large pressure step causes adiabatic heating of the silicone oil surrounding the quartz probe structure, the sensing resonator is subject to a time varying temperature gradients that results in a time varying pressure reading. During a draw-down and build-up analysis, the probe is subjected to pressure steps ranging from 69 kPa to 13.8 MPa (10 to 2000 psi) and pressure transients may last from 6 to 30 seconds. To circumvent the transient behavior of the pressure gauge, we have developed a dynamic compensation algorithm that corrects the pressure reading drift caused by the temperature gradients in the sensing resonator.

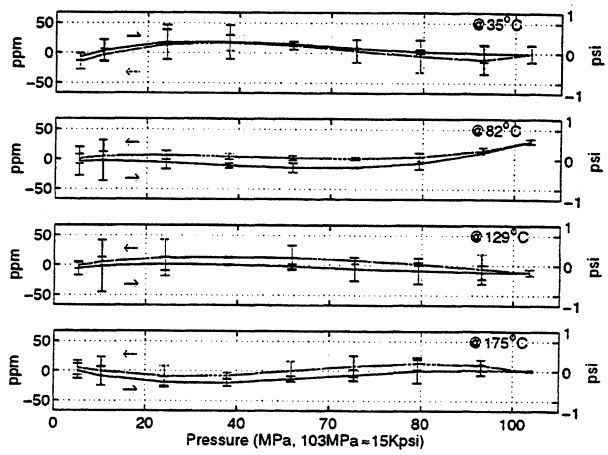


Figure 2: Average errors for 404 gauges with increasing and decreasing pressures.

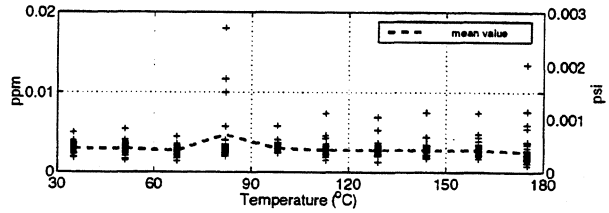


Figure 3: Allan variance of pressure reading at various temperatures

Figure 4 shows response of the CQG to a pressure step of 87 MPa (approximately 12.6 kpsi) to atmospheric pressure. The CQG pressure reading after dynamic compensation stabilizes more than 10 times faster than a conventional quartz gauge. In addition, the maximum undershoot in the CQG pressure reading is approximately 2 psi compared to about 18 psi for a conventional quartz gauge.

When the pressure probe is lowered into an oil well, the probe is subject to a temperature ramp that would also cause temperature gradients in the sensing resonator and the pressure reading drifts until thermal equilibrium is established. When the tool carrying the pressure probe is lowered into a well at a typical speed of 61 m/min (200 ft/min), it is subjected to a temperature ramp of 2° to 4°C per minute. At a normal logging speed of 9 to 18 m/min (30 to 60 ft/min), the tool experiences a temperature ramp of 0.5 to 2°C. A tool entering a gas zone may be subject to a temperature step of up to 30°C.

To enable reliable pressure measurements in the presence of temperature gradients in the sensing resonator, we apply the dynamic compensation algorithm that corrects for the temperature gradient effects on the pressure reading. Figure 5 shows response of the CQG pressure reading when the probe housing is subject to a temperature step of 125° C to 45° C and then back to 125° C at an ambient pressure of 10.47

MPa (1517.7 psi) applied to the probe. Because of the thermal impedances between the temperature bath and the sensing resonator, the temperature reading from the CQG probe shows a negative and then a positive ramp before stabilizing to a steady-state value. Even when the CQG temperature reading is varying with time, the pressure reading is essentially unaffected. In contrast, a conventional quartz gauge shows pressure errors of up to 0.2 MPa (approximately 30 psi) when the sensing resonator temperature is rapidly varying with time.

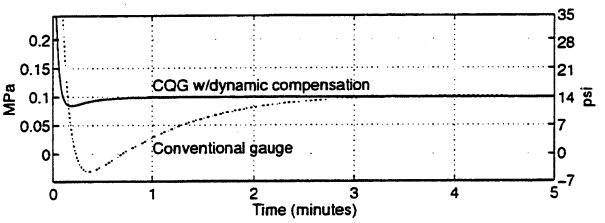


Figure 4: Response of CQG to a pressure step of 87 MPa to atmospheric pressure at 125 °C.

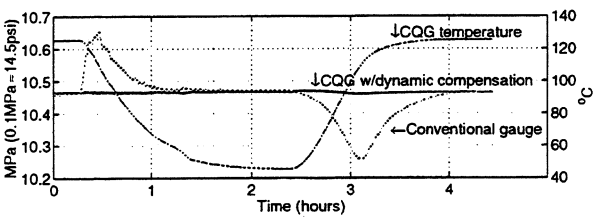


Figure 5: Response of the CQG to temperature steps of 125° C to 45° C then back to 125° C applied to the temperature bath.

IV. LONG-TERM STABILITY

In Figure 6, we show a schematic diagram of the measurement setup used in the long-term stability tests. These tests are essential for qualifying a pressure sensor for permanent monitoring systems. While the pressure sensor environment is relatively stable, these sensors are kept at high pressure and temperature for their deployment life of up to 5 years or longer. Long-term stabilities of the CQG were tested by placing four probes in a constant pressure and temperature baths for a period of 1 year. Four gauges A, B, C and D were placed in temperature controlled ovens. The temperature was kept at 150 °C for the gauges A and B, and 175 °C for gauges C and D. A constant pressure of 103 MPa (15 kpsi) was applied by RUSKA dead-weight tester 2452 with an accuracy of 0.01% of reading. The piston position was controlled so that accurate pressure could be maintained for the whole test duration. These gauges were continuously vibrated with oscillators driven at 1 micro-watt. The oscillator circuit was kept outside the oven so that the long-term stability of sensing resonator can be observed. The frequency data was recorded every hour. We measured atmospheric pressure and temperature of the piston of the dead-weight tester for removing their effects on the pressure reading. At each data acquisition, 100 frequency readings were taken with a sampling rate of 1.5 second. The counter gate time was set to 1 second. We averaged these 100 samples for removing a periodic noise in the pressure output from the dead-weight tester caused by its mass rotation – the magnitude of this correction was not negligible.

Figure 7 shows the long-term stability of the pressure reading from these four probes. The * denotes the instances when the pressure system trouble was detected. All of the four probes exhibited pressure reading stability of well within 6.89 kPa (1 psi). Figure 8 displays long-term drifts in the temperature reading from the four probes. The temperature stability is well within 0.1°C of the bath temperatures 150°C and 175°C.

CONCLUSIONS

We have demonstrated that the Crystal Quartz Gauge (CQG) based on a dual-mode thickness-shear resonator design exhibits pressure accuracy of better than 6.89 kPa (1 psi); pressure resolution of better than 20.7 Pa (0.003 psi) for a 1 second counter gate time over a calibrated pressure range of 103 MPa (15 kpsi) and a temperature range of 175°C. A significant advantage of the dual-mode design is its dynamic performance. An *in-situ* measurement of the resonator temperature allows an implementation of an extremely effective dynamic compensation algorithm that yields corrected pressure readings with a response time as short as a strain-gauge based pressure sensor while retaining the high accuracy of a quartz gauge under static conditions. The CQG response times to transient pressure pulses and temperature ramps after dynamic compensation for removing temperature gradient effects are faster than any of the requirements in the oil and gas industry. A fast response time of pressure sensors is essential for reliably estimating reservoir high permeability (higher than 1 darcy) from flow rate and pressure transient analysis. The CQG overall performance specifications exceed the requirements of downhole pressure measurements for various applications. More importantly, we have also shown that the long-term stability of CQG at 15 kpsi and 175°C is on the order of a few-tenths of a psi over a period of 1 year. To our knowledge, this is the first long-term stability test results that ensures a successful application of this gauge in permanent monitoring systems in the petroleum industry. Because of the extraordinary stability of crystalline quartz, these long-term stability results are better than other pressure sensor technologies.

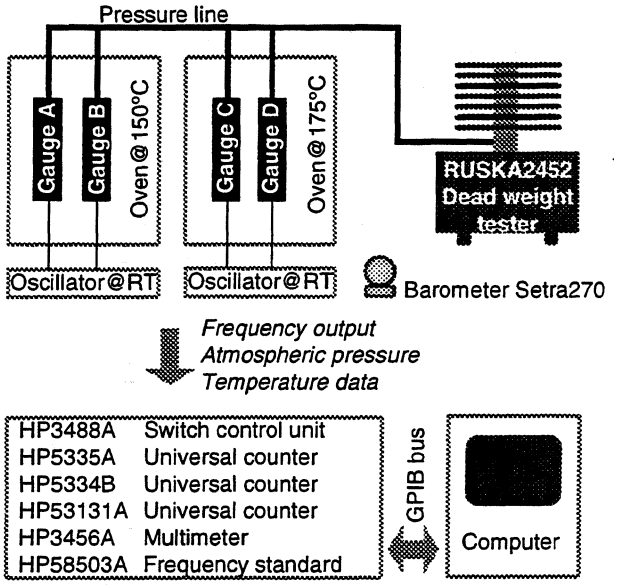


Figure 6: Schematic diagram of the set up for the long-term stability tests.

ACKNOWLEDGEMENTS

We wish to thank Michel Berard, Jean-Claude Ostiz and Pierre-Marie Petit of Schlumberger-Riboud Product Center for help with the commercialization and providing data for the dynamic response of the CQG probe in the WCQR tool. We would also like to thank Charles Adams and Jack Kusters of Hewlett-Packard Company for many helpful discussions on the design, fabrication and manufacturing of quartz pressure sensors.

REFERENCES

[1] H.E. Karrer and J. Leach, "A quartz resonator pressure transducer," *IEEE Trans. Ind. Electron. Contr. Instru.*, **IECE-16**, 44-50 (1969).

[2] R.W. Ward and E.P. EerNisse, "A reduced hysteresis, extended range quartz pressure transducer," *Proc. 41st Annu. Freq. Contr. Symp.*, 334-349 (1987)

[3] J.R. Vig, J.W. LeBus, and R.L. Filler, "Chemically polished quartz," *Proc. 31st Annu. Freq. Contr. Symp.*, 131-143 (1977).

[4] B.K. Sinha, "Stress compensated orientations for thickness-shear quartz resonators," *Proc. 35th Annu. Freq. Contr. Symp.*, 213-221 (1981).

[5] B.K. Sinha, "Stress compensated quartz resonators," U.S. Patent 4,419,600, December 6, 1983.

[6] M. Valdois, B.K. Sinha, and J-J. Boy, "Experimental verification of stress compensation in the SBTC-cut," *IEEE Trans. Ultrason. Ferroelec. Freq. Cont.*, **36**, 643-651 (1989).

[7] M. Valdois, P. Maitre, R. Besson, and J-J. Boy "Piezoelectric pressure and/or temperature transducer," U.S. Patent 4,547,691, October 15, 1985.

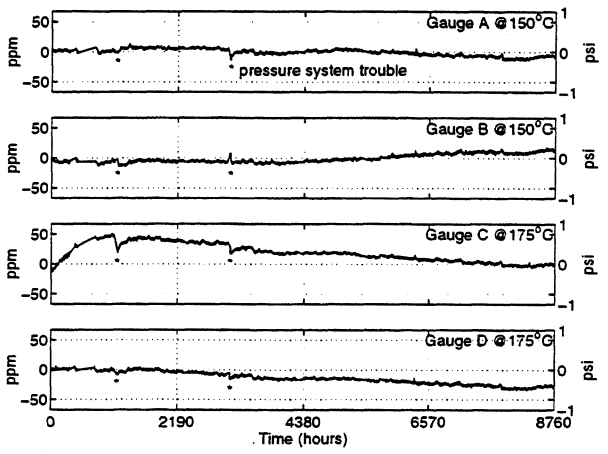


Figure 7: Long-term drift of pressure readings at 103 MPa (15 kpsi) for the two gauges at 150° C and the other two gauges at 175° C.

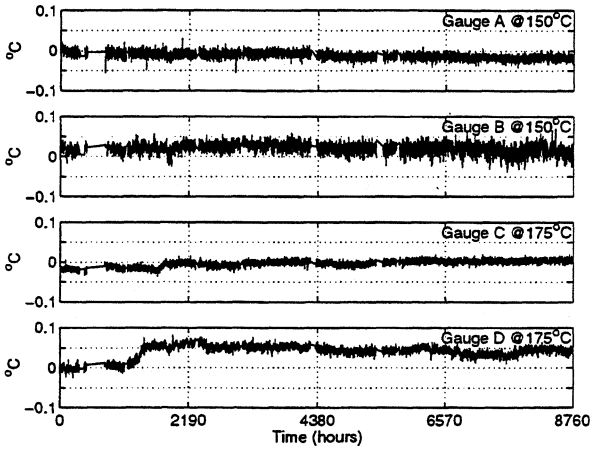


Figure 8: Long-term drift of the difference of temperature readings from the set temperatures of 150°C and 175° C at 103 MPa (15 kpsi).

[8] R.J. Besson, J-J. Boy, B. Glotin, Y. Jinzaki, B.K. Sinha, and M. Valdois, "A dual-mode thickness-shear quartz pressure sensor," *IEEE Trans. Ultrason. Ferroelec. Freq. Cont.*, **40**, 584-591 (1993).

[9] J.R. Vig, "Introduction to quartz frequency standards," *SLCET-TR-92-1, U.S. Army Laboratory Command, Fort Monmouth, NJ 07703-5601*, unpublished.

Schlumberger K.K.
Phone: 81-42-759-2111, Fax: 81-42-759-0905
nmatumot@fuchinobe.skk.slb.com
sudo@fuchinobe.skk.slb.com
niwa@fuchinobe.skk.slb.com

Schlumberger-Doll Research
Phone: 1-203-431-5515, Fax: 1-203-438-3819
sinha@ridgefield.sdr.slb.com

SOLVING THE CABLE PROBLEM BETWEEN CRYSTAL SENSOR AND ELECTRONICS BY USE OF A BALANCED BRIDGE OSCILLATOR CIRCUIT

E. Benes, M. Schmid, M. Gröschl, P. Berlinger, H. Nowotny+, K.C. Harms*

Institut f. Allgemeine Physik, Vienna University of Technology, Wiedner Hauptstraße 8/134, A-1040 Wien, Austria; +Institut f. Theoretische Physik, Vienna University of Technology, Wiedner Hauptstraße 8/136, A-1040 Wien, Austria; *AVL Ges. f. Verbrennungskraftmaschinen und Meßtechnik mbH., A-8020 Graz, Kleiststr. 48

ABSTRACT

The balanced bridge oscillator circuit presented here perfectly compensates for the negative influence of the cable. The oscillator circuit is characterized by two almost equal bridge branches; the first one contains the sensor crystal connected via the sensor cable, the second one contains an identical cable terminated by a capacity equal to the resonator's static capacity C_0. The cable compensation performance of the balanced bridge oscillator has been justified by a respective circuit analysis and by measurements of its key specifications in comparison with those of a conventional oscillator.

1. INTRODUCTION

In sensor applications, some distance between the sensing resonator and the electronics usually is required (Fig.1). In addition, the demands on oscillator electronics for sensor applications of piezoelectric crystal resonators differ significantly from those on oscillator electronics for frequency/time control applications:

1) Because the utilized resonance frequency changes with varying measurement quantity, a frequency range (in the thin film thickness quartz sensor example, 25%) instead of a fixed frequency has to be allowed for by the oscillator circuit. In the equivalent circuit of the sensor crystal shown in Fig.2 the varying resonance frequency is mainly determined by a varying motional inductance L_m. To less extent also the motional capacitance C_m – that is directly related to the active vibration area – may change.[1]

2) The Q-value of composite resonators often varies significantly within the measurement range. This is especially true for heavily loaded sensor crystals in thin film thickness monitors [2] and for sensor crystals in contact with a liquid.[3] The change of the Q-value is directly related to a respective change of the motional resistor R_m. As a consequence, the amplifier gain must vary over a wide range (in the thin film thickness quartz sensor example, 250:1) to keep the loop gain of the oscillator circuit equal to 1 for maintaining harmonic excitation. An interesting approach to handle this wide dynamic range of the resonator's motional loss resistance is, e. g., the so-called lever oscillator.[4]

3) In sensor applications usually remote electronics is desired. For a temperature sensor, e. g., it is advantageous to locate the sensor remotely in order not to restrict the temperature range by the respective limits of the electronics and to avoid heating of the sensor by the electronic circuit's dissipation. While SAW sensors offer - because of their significantly higher operating frequency - the possibility of a wireless interrogation via antenna coupling,[5]

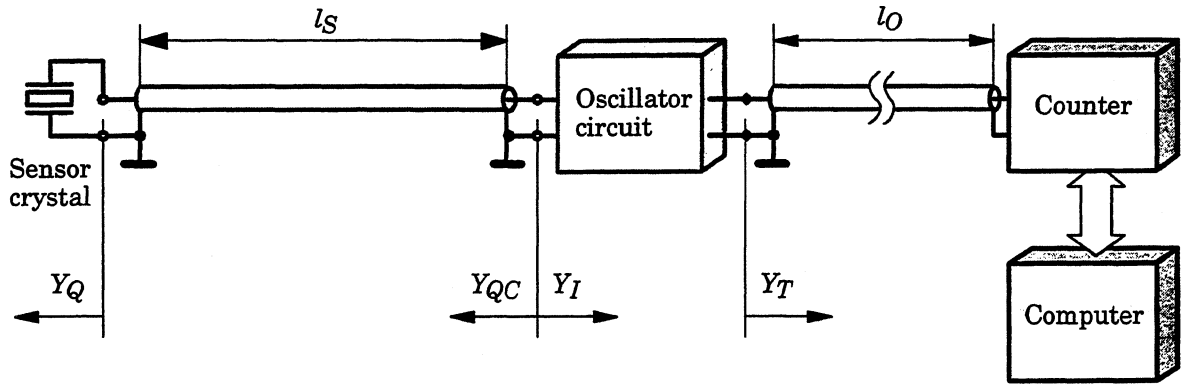


Fig.1. Coaxial cables between sensor crystal and oscillator circuit (length l_S) and between oscillator and counter (length l_O). Y_Q ...Admittance of sensor crystal; Y_{QC} ...Admittance of sensor crystal connected via sensor cable; Y_I ...Input admittance of oscillator circuit; Y_T ...Admittance of oscillator cable terminated with its characteristic impedance (e.g. 50 Ω)

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

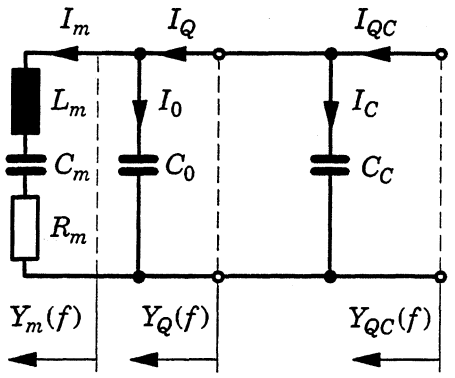


Fig.2. Lumped element equivalent circuit of sensor crystal and sensor cable in the vicinity of a crystal resonance frequency (e.g. thickness shear fundamental).
L_m, C_m, R_m ...Motional branch inductance, capacitance and loss resistance of the sensor crystal, respectively
C_0 ...Parallel capacity of the crystal
C_C ...Capacity of the sensor cable
$Y_m(f)$, $Y_Q(f)$, $Y_{QC}(f)$...Admittance of the crystal's motional branch, of the total crystal and of the crystal connected via the cable, respectively

BAW sensors usually need a coaxial cable connection to the exciting electronics.[6] In the higher frequency range above about 25 MHz a coaxial transmission line with a cable length of an integer multiple of a half wavelength ($n\lambda/2$) of the electrical signal in the cable can be used to allow for a distance between the sensor and the electronic oscillator circuit. Such a cable transforms the electrical admittance of the resonator 1:1 to the position of the oscillator electronics. Thus, the oscillator circuit can be designed as for the case where the resonator is directly connected to the electronics. For example, by use of a 50 MHz AC-cut quartz crystal as temperature sensor, a coaxial cable (Teflon® insulator, $\varepsilon = 2$) with a length corresponding to one wavelength $\lambda = 4.23$ m has been successfully used.[7] In this application the frequency change in the covered temperature range from -50 to 200 °C is less than 1%, therefore the λ-condition remains sufficiently fulfilled.

In the lower frequency range, e.g. for the quartz crystal thin film thickness sensor, a cable with a length l sufficiently smaller than a quarter wavelength is used ($l << \lambda/4$). This is because a $\lambda/4$-cable transforms a low impedance load, e.g. at the sensor's series resonance, into a high impedance load at the oscillator circuit input port, thus the cable's influence would be maximum.

For a cable with a length l sufficiently smaller than a quarter wavelength ($l << \lambda/4$), the cable performs essentially as a capacity C_C that adds to the resonator's static capacity C_0 (Fig.2). However, sometimes the cable capacity C_C causes an impedance so low that the oscillator circuit oscillates even without a sensor crystal connected. Even if this extreme condition does not apply, a high influence of the cable capacity C_C on the oscillator frequency remains. In principle, for a single frequency, the compensation of the cable capacity can be achieved by connecting an inductor of the same absolute admittance value as C_C in parallel. But an inductor has different and less perfect properties (especially with regard to temperature dependence and losses) than a capacitor. Furthermore, for the example of the thin film thickness monitor sensor, where the sensor frequency changes typically from 6 to 4.5 MHz with increasing film load, compensation effective over the entire measurement frequency range is desired. In the following Section 2 a typical conventional series resonance oscillator is described and its limitations with respect to the sensor cable influence are indicated. The balanced bridge oscillator as appropriate problem solution is presented in Section 3.

2. CONVENTIONAL CRYSTAL SENSOR OSCILLATORS

To keep the characteristic impedances in the high frequency range of acoustic wave crystal sensors are in the order of 50 Ohms, the resonator sensors are usually excited at a low impedance resonance, preferably at one of the thickness shear mode series resonances, e.g. the fundamental. A series resonance circuit like the motional branch in Fig.2 shows the most pronounced resonance effect if it is driven with a "stiff" voltage amplitude (zero source impedance) and the current through the crystal is used as resonance indicating signal.

Fig.3 shows a preferred series resonance oscillator concept, where the circuit function is that of an amplifying current to voltage converter with stiff output voltage fedback to the sensor crystal and controlled by the current through the sensor crystal.

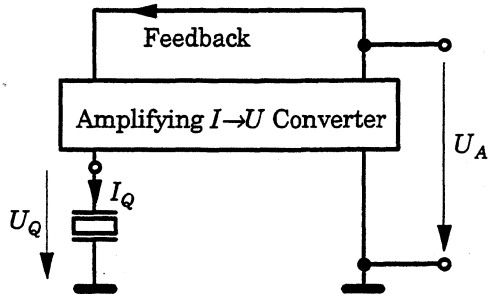


Fig.3. Series resonance oscillator principle.

A typical realization of this concept is shown in

Fig.4, where the connection of the crystal via the cable is already taken into consideration, that is instead of $I_Q \rightarrow I_{QC} = I_Q + I_C$ is regarded. Here and throughout this paper sinusoidal time dependence $e^{j\omega t}$ is considered and all quantities have the respective complex meaning.

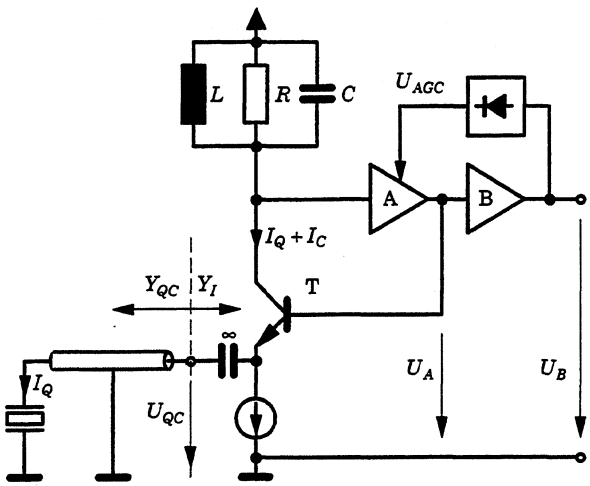


Fig.4. Typical conventional quartz crystal sensor oscillator circuit with automatic gain control

Introducing the forward transfer impedance Z_G we obtain for the loop gain g_L

$$Z_G = \frac{U_A}{I_{QC}}, \quad g_L = \frac{U_A}{U_{QC}} = \frac{I_{QC} Z_G}{U_{QC}} = Y_{QC} Z_G$$

The condition for harmonic excitation is $g_L = 1$, the negative input admittance of the oscillator has to be equal to the sensor+cable admittance

$$-Y_I = Y_{QC} = 1/Z_G$$

Assuming simplified transistor properties (no frequency dependencies, negligible base current) we obtain

$$U_A = g Z_{LRC} I_{QC} = g Z_{LRC} Y_{QC} U_{QC}$$

$$g_L = g Z_{LRC} Y_{QC} = 1, \quad Y_{QC} = Y_Q + Y_C = Y_{LRC} \frac{1}{g}$$

$$\text{where } Y_{LRC} = \frac{1}{Z_{LRC}} = \left(\frac{1}{R} + j\omega C + \frac{1}{j\omega L} \right)$$

In Fig.5 both, the locus of the crystal+cable admittance with the frequency as parameter, as well as the locus of admittance of the negative oscillator input impedance with the gain g of the gain controlled amplifier A, is plotted. The crossing point with the lowest required gain is the operating point for a crystal resonator with a given damping (R_m value). From Fig.5 the extremely unfavorable influence of the cable capacitance C_C together with that of the crystal's parallel capacitance C_0 can be seen. For highly damped crystals (high motional resistance R_m) excitation is only possible for a steep angle φ, however, that means that, especially for the less loaded crystal, the operating frequency is dragged far off the series resonance, resulting in a low effective Q-value and thus decreased frequency stability.

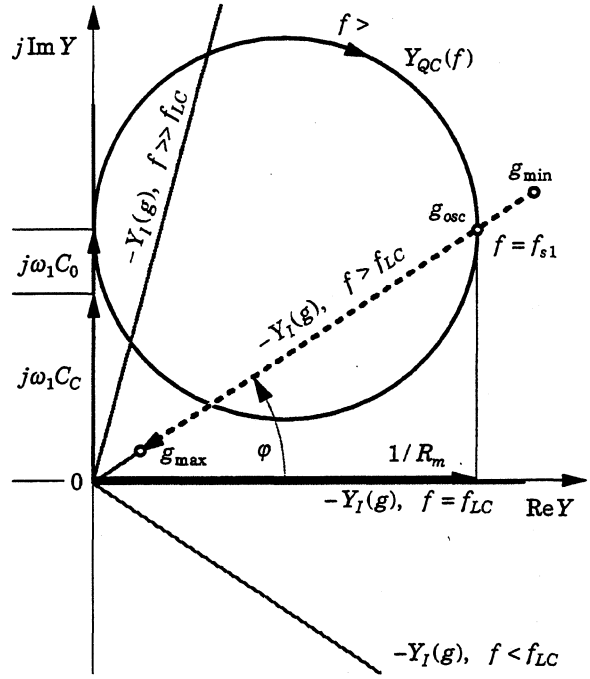


Fig.5. Locus of admittance $Y_{QC}(f)$ of the crystal connected via the sensor cable and locus of negative input admittance $-Y_I(g)$ of the oscillator with the frequency f and the amplifier gain g as parameters, respectively. $C_0 \approx 15\text{pF}$, $C_C \approx 45\text{pF}$ $R_m = 10 \ldots 2800\Omega$

3. THE BALANCED BRIDGE OSCILLATOR

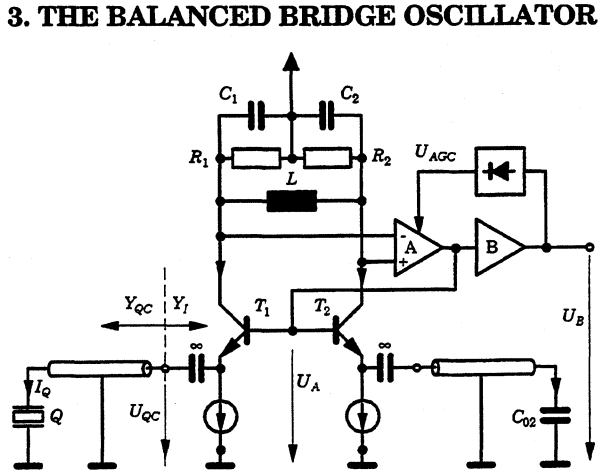


Fig.6. Principle of the balanced bridge sensor oscillator circuit with automatic gain control

The balanced bridge oscillator circuit is characterized by two almost equal bridge branches; the first one contains the sensor crystal connected via the sensor cable, the second one contains an identical cable terminated by a capacity C_{02} equal to the resonator's static capacity C_0. Here the input voltage of the differential amplifier A is proportional to the difference of the currents through the two cable branches. Thus, only the motional branch of the crystal remains as frequency controlling circuit.

$$\Delta I = I_m + I_0 + I_C - I_{02} - I_{C2} = I_m$$

This effectively shifts the origin of the oscillator admittance curves $-Y_I(g)$ in Fig.5 upwards to the leftmost point $j\omega(C_0+C_C)$ of the circle representing $-Y_{QC}(f)$.

4. COMPARISON OF OBTAINED SPECIFICATIONS

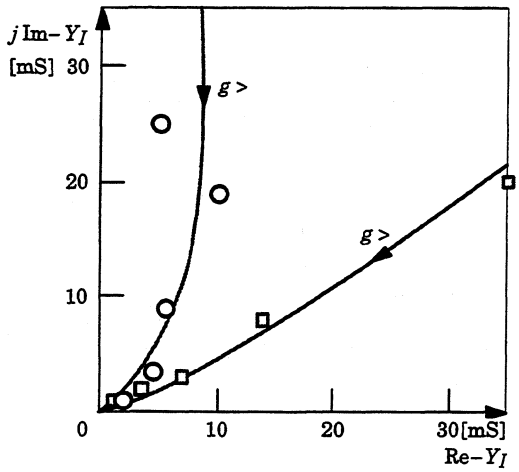


Fig.7. Locus of the negative input admittance of the oscillator circuit $-Y_I$ for two differently dimensioned collector admittances Y.

Table 1. Frequency stabilities $\Delta f/f$ for varying temperature ΔT and supply voltage ΔU_b; maximum motional resistance R_m range of a conventional (Inficon) and of the balanced bridge (Balbri) oscillator utilising a 75 Ω cable of 70 cm length.

Oscillators:	Inficon	Balbri
$(\Delta f/f)/\Delta T$ [K^{-1}]	0.003/Q	0.0002/Q
$(\Delta f/f)/\Delta U_b$[V^{-1}]	0.1/Q	0.01/Q
R_m range	0...200 Ω	0...2800 Ω

In Fig.7 the measured negative input admittance of the oscillator circuit is plotted for two different collector admittances Y, whereby the flatter curve is much more advantageous. Note that - due to the compensation of $C_0 + C_C$ - the curves start now at the zero point of the complex plane. The adjustment of the appropriate g values was obtained by insertion of crystals with different load and respective Q-values. Table 1 shows a comparison between some key specifications of a conventional and of the balanced bridge oscillator. For these measurements instead of the sensor crystal a lumped equivalent circuit according to Fig.2 was used, however with L_m and C_m values of 0.01 and 100 times the crystal's values, resepctively. The resulting lower Q-value allows for an easier to measure, more pronounced influence of the oscillator circuit's contribution.

5. SUMMARY AND CONCLUSIONS

Since the sensor frequency may vary over a relative frequency range of up to 25%, a compensation of the negative influence of the cable effective over the entire measurement frequency range is desired. Such a perfect compensation can be achieved with the balanced bridge oscillator. The cable compensation performance of the balanced bridge oscillator has been analized and demonstrated by measurement of characteristic performance parameters. The key specifications of the balanced bridge oscillator are typically 10 times better than those of a conventional state of the art series resonance oscillator.

ACKNOWLEDGEMENTS

Part of the work was financially supported by "Österreichischer Fonds zur Förderung der wissenschaftlichen Forschung" (Austrian Science Foundation, Project No. P 7198), by Leybold AG Hanau, BRD, and Inficon Inc., Syracuse, USA.

1 M. Schmid, E. Benes, W. Burger and V. Kravchenko, Motional capacitance of layered piezoelectric thickness-mode resonators, *IEEE Trans. Ultrasonics, Ferroelectr. Freq. Control, UFFC 3* (1991) 199-206.

2 Benes, Improved quartz crystal microbalance technique, *J. Appl. Phys., Vol.56, No.3* (1984) 608-626

3 Thalhammer, S. Braun, B. Devcic-Kuhar, M. Gröschl, F. Trampler, E. Benes, H. Nowotny, M. Kostal, M. Hruskovic, J. Hribik, Viscosity sensor based upon an angular momentum compensated piezoelectric thickness shear sandwich resonator, *IEEE Trans. Ultrason., Ferroelect., Freq. Contr. 45* (1998) 1331 - 1340E.

4 K.O. Wessendorf, The lever oscillator for use in high resistance resonator applications, *1993 IEEE International Frequency Control Symposium* (1993) 711-717.

5 A. Pohl, F. Seifert, Wirelessly Interrogable SAW Sensors for Vehicular Applications, *IEEE Trans. Instr. Measurement, 46* (1997) 1031-1038R.

6 Ewald Benes, Martin Gröschl, Franz Seifert, Alfred Pohl, Comparison between BAW and SAW Sensor Principles, *IEEE Trans. Ultrason., Ferroelect., Freq. Contr. 45* (1998) 1314 - 1330

7 E. Benes,Verwendung eines AC-Schwingquarzes als Temperaturfühler, *Elektronikschau, 10* (1979) 31-32.

IDENTIFICATION WITH SAW DEVICES IN PASSIVE REMOTE TELEMETRY SYSTEMS

W. Buff, J. Ehrenpfordt, St. Klett, M. Rusko and M. Goroll

Technical University of Ilmenau, Institute of Solid State Electronics, D-98684 Ilmenau, Germany

ABSTRACT

Identification devices (ID-tags) for remote telemetry systems in the European ISM standards require the use of narrow-band solutions. SAW ID-tags are characterized by particularly good coding and correlation capabilities which is used in many filter applications.

To solve the narrow-band problem of ISM-bands two principles for passive SAW ID-Tags were introduced, the Multiple Apodized Self-Correlating Delay Line (MASCoT) and the „On-Chip-Correlation" (OCC) device. The advantage of the devices is an „impulse-free" interrogation so that the used bandwidth does not extend the allowed bandwidth and the European ISM standard can be fulfilled. The design, simulation and experimental results were introduced and discussed.

1. INTRODUCTION

SAW elements have been proofed to be applicable to identification systems with passive identification devices [1]-[6]. However, to detect simultaneously more than one ID-tag in the interrogation area is not a trivial problem. All responses overlaying each another will be received by the interrogation unit at the same time, must be separated and assigned to the corresponding identification tag. This problem can be solved by using structured signals created by the identification tag while interrogated. PSK (phase shift keyed) signals [4] are well suitable for this aim. Well-known solutions for such SAW ID-tags usually operate with impulse interrogated PSK coded SAW filters or reflecting delay lines [5].

The main disadvantage of these systems is the use of an impulse radio transmission between the interrogation unit and the ID-tag which needs a wide bandwidth in the radio channel. A modern philosophy of the „frequency economy" and the frequency standards are hardly implemented. The new solution presented in this paper overcomes the difficulties and meets the conditions of the frequency standards by using a correlation on the SAW chip. This type of devices also allows to create a system of passive sensors with self-identifying capabilities.

2. PRINCIPLE OF "MASCoT" DEVICE

MASCoT is an ID-Tag based on a multiple apodized self correlated tapped delay line [6]. The device is derived from conventional tapped delay lines and is optimized for wireless interrogation systems as a two-terminal ID-tag. This design allows low-bandwidth and low-power interrogation using long, unmodulated sine bursts. The device responds with a PSK-code sequence suitable for a new wireless CDMA-based identification system that allows a simultaneous interrogation of multiple ID-tags by using orthogonal codes.

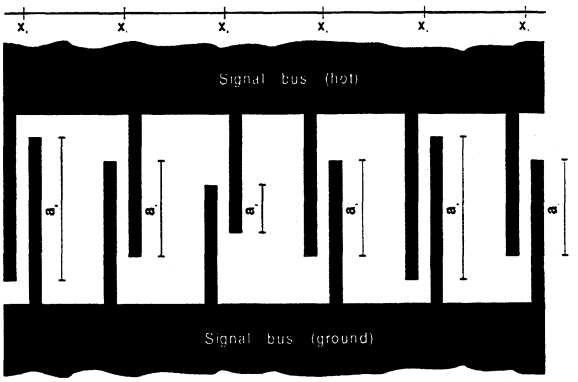


Fig. 1: Layout of the six tap MASCoT device

Fig. 1 shows the layout of a six-tap device for a five bit PSK code sequence. It consists of six (n=6) interdigital transducers (IDTs, taps) with different apertures a_i that are connected in parallel via two bus bars. The working principle is based on superposition of the correlation signals of multiple apodized transducer pairs. In the device, every transducer (tap) works as both, receiver and transmitter of surface acoustic waves. In [6] a mathematical model for the first order approximation of the response of the device to a long sine burst in time domain has been published. It allows the design of the apertures in dependence on the code sequence (binary PSK). The apertures are calculated by using a restricted tree-search algorithm.

The measured S_{11} versus the frequency is presented in Fig. 2. For the measurements test devices have been designed on the YX-128°-cut-$LiNbO_3$ substrate with split finger IDTs. The center frequency f_0 of 428.5 MHz was measured.

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

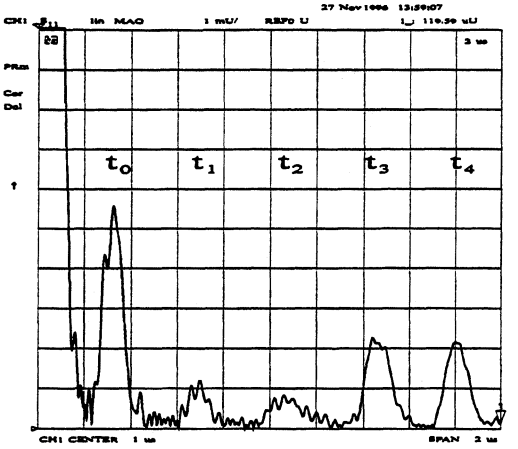


Fig. 2: Measured time response of MASCoT device $|IFFT(S_{11})|$ [6]

Fig. 2 shows the IFFT transformed S_{11} response into time domain for illustration of the working principle. The result closely resembles the impulse response of the device. Due to the electrical impedance of the network analyzer a large peak at zero time position is present. The trace shows triangular artefacts at discrete time steps t_0, t_3 and t_4 which are evidence of the correlation of the generated SAW with the receiving IDT. The signals at t_1 and t_2 should be zero because of cancellation of SAWs.

Compared to the conventional solutions [5] the MASCoT device can operate in narrow-band telemetry systems with the simultaneous access to several ID-Tags in detection area of the interrogation unit. The separation of the signals of the different sensors uses CDMA technique. The disadvantage of this method is a rather limited number of realizable codes due to the interrogation with an orthogonal code [6].

3. PRINCIPLE OF "ON-CHIP CORRELATION"

The particular disadvantage of MASCoT can be solved by using the OCC-devices. In the case of "On-Chip-Correlation" (OCC) device the correlation function is used to create a peak in form of a SAW on the chip. It is ensured that the interaction between different transducer codes does not falsify the output signal.

The correlation function is often used in the communication technology as a method for the separation of different codes (sensors) [8]. If the source code corresponds to the device code a typical correlation peak results at the device output.

Conventional solutions [5] with passive SAW reflective ID-Tags operate with the impulse interrogation of a PSK device and gets a coded ID-Tag response back. This principle has the disadvantage of an impulse radio transmission. Mentioned impulse has a form of an interrogation impulse for the coded SAW structure or of a peak response of the delay line ID-Tag. However both versions violate the band limitation of the ISM-Band.

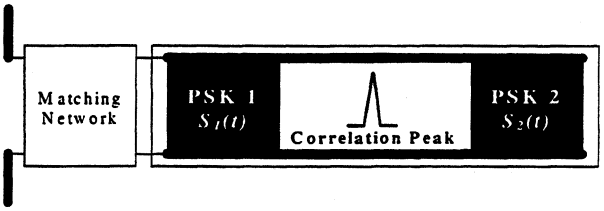


Fig. 3: The design of OCC-ID-Tag

In this case the device (Fig. 3) consists of two PSK coded SAW transducers PSK1 and PSK2 which are connected via a common signal bus and a matching network to an antenna. The PSK1 transducer is coded with a source code *S1(t)*. This part of the structure provides the interrogation of the second PSK2 transducer with a different code *S2(t)*. The device interrogation by the PSK signal with the *S1(-t)* code generates in the PSK1 transducer a high energy correlation peak on the substrate propagating in the direction to the PSK2 transducer. This impulse causes an response of the PSK2 transducer with *S2(t)* code. Due to the commutative properties of the convolution the forward and backward signal path are equal. That's why the connection between PSK1 IDT and PSK2 IDT does not influence the output signal negatively. The principle can be written by Eq. (1).

Forward:

$$\text{Output(t)}=S1(-t)*S1(t)*S2(t) \tag{1a}$$

Backward:

$$\text{Output(t)}=S1(-t)*S2(t)*S1(t) \tag{1b}$$

$$S1(-t)*S1(t)*S2(t)=S1(-t)*S2(t)*S1(t) \tag{1c}$$

Condition:

$$S1(-t)*S1(t)\approx\delta(t) \tag{1d}$$

4. CORRELATION CODE

For the generation of a correlation/interrogation impulse a binary sequence $s(i)\in\{+1, -1\}$, $i=1:N$, with a suitable correlation function Φ_{ss} was chosen. The evaluation criteria are the peak to sidelobe ratio *(PSR)* for characterizing the energy, the merit-factor *(MF)* as a noise measure in the sidelobes and the sequence length *N*. The *PSR* and *MF* calculation show Eq. (2) and (3) [8].

$$PSR=\frac{\Phi_{ss}(0)}{\max\left|\Phi_{ss}(i)\right|},\forall i\neq 0 \tag{2}$$

$$MF=\frac{\Phi_{ss}^2(0)}{2\cdot\sum_{i=1}^{N}\left|\Phi_{ss}(i)\right|^2} \tag{3}$$

According to well known tables of Lindner-codes [9] a sequence with a length N=28, PSR=14, MF=7.86 was chosen for the PSK1 transducer as *S1(t)*.

Fig. 4 shows the calculated correlation function. The code notation is hexadecimal.

The correlation peak reaches a maximum value of 28 and the sidelobes never cross the value of 2. At the bottom the peak has time width of two bits. That means if one bit is modulated with n subbits (sine burst), an identification code $S2(t)$ with $2n$ subbits is necessary.

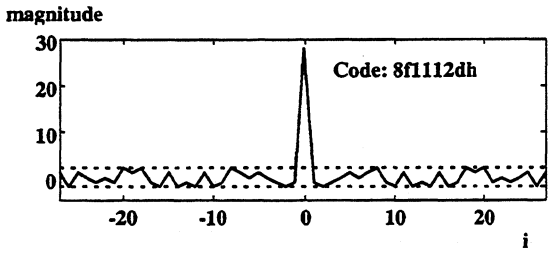


Fig. 4: The correlation peak of 28-bit Lindner sequence

5. EXPERIMENTAL SETUP WITH 80 MHz

To prove the OCC idea a first test device with center frequency of f_0=80 MHz was used. The setup (Fig. 5) consists of two PSK coded transducers with an additional wide-band IDT in-between.

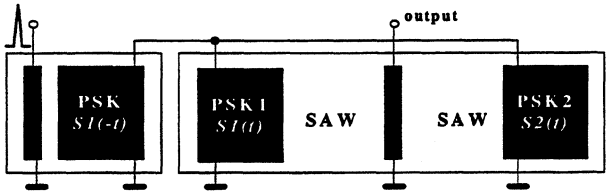


Fig. 5: The design of 80 MHz „On-Chip Correlation" ID-Tag with in-between wide-band transducer.

Two parallel connected PSK1 and PSK2 transducers have different codes $S1(t)$ and $S2(t)$. The time inverted interrogation signal $S1(-t)$ was prepared in an additional PSK device by stimulation with a short pulse (10V, 6ns). Fig. 6a shows the generation of the interrogation code $S1(-t)$. This sequence was linked to the common bus electrodes of PSK1 and PSK2. A convolution of $S1(-t)$ and $S1(t)$ results in a correlation peak on the chip in form of a SAW which is needed for an interrogation of the PSK2 transducer. The interaction between $S1(-t)$ and $S2(t)$ results in a cross correlation function with a neglected magnitude which does not disturb the identification output signal of the OCC device.

The mentioned auxiliary signal at the output pin of an „in-between" wide-band transducer was amplified and measured as shown in Fig. 6b. After switching off the interrogation signal $S1(-t)$ a clear correlation peak follows which confirms the idea on OCC.

Unfortunately the identification signal $S2(t)$ disappears in the noise. A better SNR setup of the experiment is necessary to filter the output signal.

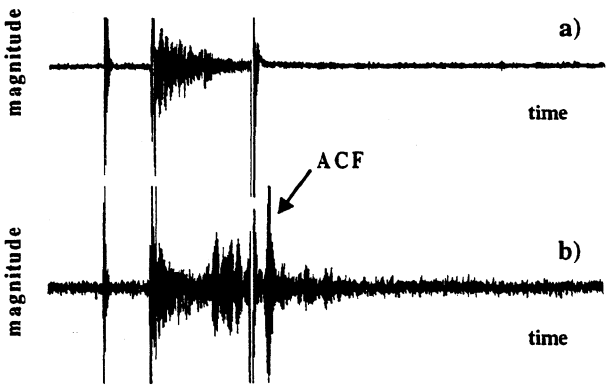


Fig. 6: An impulse response of the „On-Chip Correlation" device (generation of an interrogation PSK signal S1(-t)(a), output IDT(b))

6. EXPERIMENTAL SETUP WITH 433.92 MHz

A new device for 433.92 MHz with an improved setup of two parallel connected PSK transducer on common signal bus (which was prepared for direct connection to an antenna) was designed. The „interrogation" transducer PSK1 was coded with a 28 bit code $S1(t)$ with 4 subbits each bit. The identification information is hidden in the PSK2 transducer with a 14 bit $S2(t)$ binary code. With respect to device symmetry, its electrical parameters and the good signal-fitting 8 subbits in $S2(t)$ were designed. A compromise between the device/code length has to be kept to hold the acceptable device geometry and its static capacitance.

The simulation of the device function was made with Matlab. The basis for the calculations was the modeling of an elementary signal/wave in form of a [0 1 0-1] sequence. Fig. 7 shows the simulation results of the OCC device. Fig. 7a represents a PSK coded signal $S1(-t)$ in a simplified form as was used for the narrow-band „interrogation" of the OCC device.

The computation of the autocorrelation function ACF of two coded sequences in Matlab uses the function $CONV$.

$$ACF = CONV(FLIPLR(S1), S1) \tag{4}$$

$FLIPLR(S1)$ function shapes a time inverted form of $S1(t)$ sequence to $S1(-t)$. These two sequences are necessary for the generation of the interrogation correlation peak (see Fig. 7b). An interaction of the correlation peak with the $S2(t)$ coded second transducer is the output signal of the device as was simulated (see Eq. 5 and Fig. 7c).

$$OUTPUT = CONV(ACF, S2) \tag{5}$$

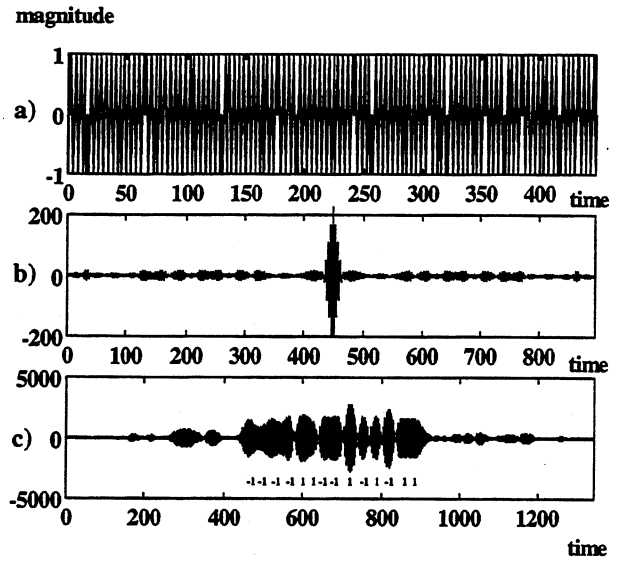


Fig. 7: Identification simulation (exciting PSK signal (28 bit /4 subbit) (a), *ACF* code (b), output as 14 bit ID code (c))

Fig. 8 shows the comparison between the simulated (Fig. 8a) and the measured (Fig. 8b) impulse response of the 433.92 MHz OCC device. There is a good agreement in the phase shifts which are clearly to see in both sequences.

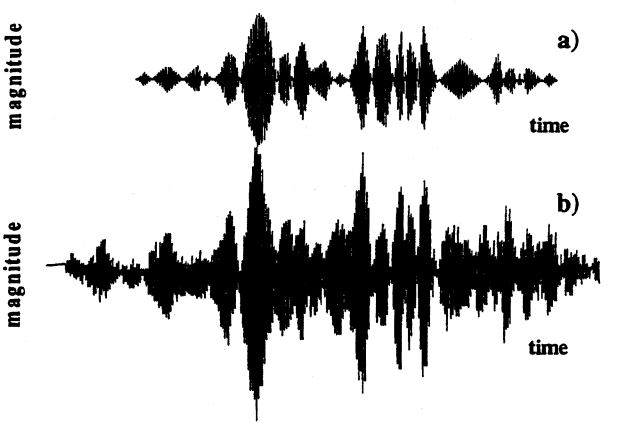


Fig. 8: An impulse response of the 433.92 MHz „On-Chip Correlation" device. A comparison between the simulation (a) and the measurement (b)

7. CONCLUSION

Two new ID-tags for passive narrow-band ISM identification systems were presented. Their special design allows impulse free interrogation and multiple access (MASCoT) in the interrogation area. The choice of the „interrogation" code is significant for the device function. The principle of the OCC was verified by simulation and measurements with an 80 and 433.92 MHz test setup under laboratory conditions. There is a good agreement between the simulation and experimental results.

8. ACKNOWLEDGMENT

The presentation of this work has been supported by the EU Contract BRPR-CT96-0361 (REMSENSE). The authors thank Prof. H. Hofmann from HTW-Dresden and Dr. Wall (Vectron Telefilter tft GmbH) for their device support and fabrication

9. REFERENCES

[1] Buff, W.: *SAW Sensors*, Sensors and Actuators, 17 (1989), pp. 55-66

[2] Buff, W., Goroll, M.: *Wireless Remote Sensing and Identification by Using Passive SAW Devices*, Proc. Mechatronics'98, 1998, pp. 553-558

[3] Cole, P.H., Vaughan, R.: *Electronic Surveillance System*, U.S. Patent 3,706,094, Feb. 26, 1970

[4] Möller, F. et. al.: *Differential Phase Shift Keying direct Sequence Spread Spectrum SAW Based Correlator Receiver*, Proc. IEEE Ultrasonics Symp., 1994, pp. 189-193

[5] Reindl, L.; Ruile, W.: *Programmable Reflectors for SAW ID-Tags*, Proc. IEEE Ultrasonics Symp., 1993, pp. 125-130

[6] Vandahl, Th. Et. al.: *Multiple Apodized Self-Correlating Tapped Delay Line for ID-Tag and Sensor Applications*, Proc. IEEE Ultrasonics Symp., 1997, pp. 349-354

[7] I-ETS 300 220: Radio Equipment and Systems (RES); *Short range devices Technical characteristics and test methods for radio equipment to be used in the 25 MHz to 1000 MHz frequency range with power levels ranging up to 500 mW*, ETSI, France, Oct. 1993

[8] Lüke, H. D.: *Korrelationssignale*, Springer-Verlag, Berlin Heidelberg, 1992

[9] Lindner, J.: *Binary sequences up to length 40 with best possible autocorrelation function*, Electronics Letters, Vol. 11, 1975, No. 21, p. 507

Corresponding author:

Prof. Dr.-Ing. habil W. Buff
Technical University of Ilmenau
Institute of Solid State Electronics
D-98684 Ilmenau, Germany

☎ +49 3677 693124
fax +49 3677 693132
e-mail: werner.buff@e-technik.tu-ilmenau.de

1999 Joint Meeting EFTF - IEEE IFCS

STATE OF THE ART IN SIGNAL PROCESSING FOR WIRELESS SAW SENSING

Alfred Pohl

University of Technology Vienna, E3592, Applied Electronics Laboratory,
Gusshausstrasse 27, A-1040 Vienna, Austria. alfred.pohl+e3592@tuwien.ac.at

Wireless measurement systems with passive SAW sensors offer the new and exciting perspectives for remote monitoring of rotating parts, or in harsh environments. This paper concentrates on the signal processing required for those sensor systems with regard to the sensors, the system hardware and the radio transmission. Different directly effected delay lines, impedance loaded devices and resonators have been used as SAW sensors. Wideband time domain interrogation and narrowband frequency domain interrogation have been investigated theoretically and experimentally in combination with the radio channel and the evaluation method. We focus on the occurring errors and the origins of these effects. Actual estimations for the different errors are given. We achieved a resolution of 10^{-5} to 10^{-6}, the measurement's bandwidth is up to 10 kHz for conventional sensors and up to a few hundreds of kHz for the new transponder type. A read out distance without averaging of several meters is achieved.

1. Introduction

The current paper focuses on the radio request of passive surface acoustic wave (SAW) devices, published only a few years ago for radio sensor purposes. Initially, these circuits have been invented for the passive identification of animals [1]. The first industrial application is a road pricing system for the Norwegian Highways around Oslo. Up to now, a wide range of actually performed applications have been published, e.g. [2], [3], [4].

Interrogated by an RF radio signal, a linearily distorted version of the requesting signal is retransmitted. The distortion is effected by the measurand. Due to the energy storage in the SAW, a time separation between radio request signal and sensor response is achieved. The received response signal is afflicted with noise and interference. Errors occurring during transmission yield additional measurement errors inseparable from the sensor effect. In Table 1, actually employed SAW radio sensors are listed. Utilizing DL and DDL, a delay ΔT has to be evaluated. For DLT (delay line transponders, loaded by an impedance), also the amplitude of the reflected signal can be used. Requesting SAWR, the frequency of the decaying resonator and decaying amplitude is measured.

	delay lines		resonators
	wideband	dispersive	
one port	DL	DDL	SAWR
two port	DLT	-	-

Table 1: SAW devices actually used for radio sensors

In chapter two, the applied methods of radio request are presented. After considering the radio channel in the third chapter, in the fourth, the deterministic and stochastic errors occurring at sensor response detection are investigated. An estimation of measurement performance versus the distance range and the error parameters is given in chapter five. The contents of the paper is summarized in a brief conclusion.

2. Radio request of passive SAW sensors

Regardless of which type of passive radio sensor is used, a wireless one port response measurement with time division has to be performed. As for RADAR systems, the receiver usually is located near the transmitter, coherent detection is feasible.

2.1 Radio request methods

In general, the measurement can be performed in time and frequency domain.

2.1.1 Time domain sampling (TDS), wideband or full band sampling

For TDS, the sensor is requested by a radio request signal transmitted, covering the total system bandwidth at once (full band). The duration of the radio request signal is given for non spread spectrum signals to be $T_{signal,TDS} \leq 1/2B_{sensor}$, with the sensor bandwidth B_{sensor}. Each radio request signal causes one (for a single SAWR) or a number of (for DL) response signals. A wideband sampling is done in the receiver. TDS is a single scan measurement method, the whole sensor response can be recorded in one radio request cycle.

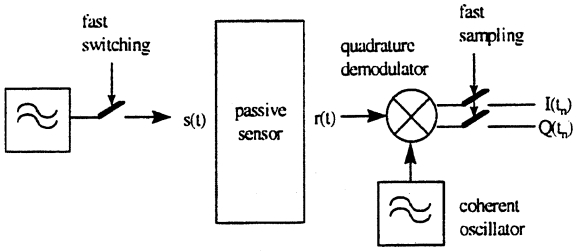


Figure 1: TDS transmitter and receiver

The energy contents $E_{TDS,0}$ of one request is the time integral of the transmitted signal power, due to the

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

wideband impulses, commonly it is small. The convolution of the signal with the sampling window has to be noticed. The sampling theorem applies. To avoid aliasing, the sampling has to be performed with at least twice the bandwidth of the baseband signals.

2.1.2 Frequency domain sampling (FDS), narrowband or partial band sampling

FDS or partial band sampling means scanning of the total bandwidth in M steps in the frequency domain. To achieve high resolution, the bandwidth of one step must be low, requiring a relative long duration $T_{signal.\ FDS} \geq 1/2B_{res}$ (with the resolution bandwidth $B_{res} = B/M$) of the radio signal and enhancing the total measurement time as well as the total energy used for detection. In the receiver, magnitude and phase of the narrowband response is detected. As for the TDS system, for coherent detection, the RF signals can be derived from one oscillator.

FDS is a multiscan measurement. To achieve the information of a number of M points in time, M frequencies have to be scanned. The total measurement in minimum lasts M times the minimum measurement cycle of TDS.

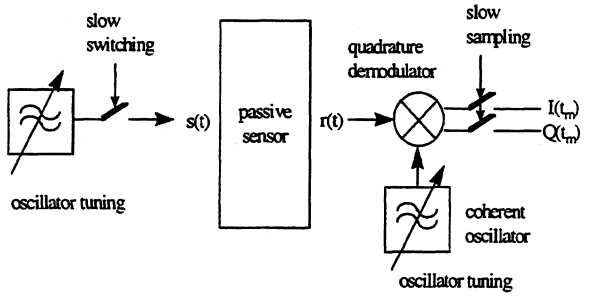


Figure 2: FDS transmitter and receiver principle

For TDS and FDS, the analog signal received is sampled and digitized by an analog to digital converter. Then, a two column vector is gained. One column contains the detector voltage of the component in phase (I) and the other orthogonal (Q) to the reference. The row number is the number of the sample in time and frequency, respectively, depending if TDS or FDS was applied. Utilizing a discrete Fourier transform (FFT) algorithm, time samples can be converted to frequency and vice versa. At sensor readout, time delay ΔT, frequency f_c and magnitude A of RF signals have to be measured or estimated, respectively, in the receiver.

Apart of counters, time to digital converter, etc., the widely utilized method of **time measurement** is phase measurement in coherent systems by projection into an orthogonal base system. The relative phase of the received signal is calculated from arctan (Q/I). The time delay is gained from phase shift by

$$\tau_{tot} = \frac{\varphi_{tot}}{2\pi \cdot f_c}. \qquad (1)$$

Frequency measurement, e.g. of the decaying resonator's response, is done by TDS and FFT, TDS and parameter fitting (MSE, mean square error minimization) and by FDS and gated phase locked loops.

Since it is strongly affected by the radio transmission, in passive radio sensor applications, **amplitude information** is used rarely. TDS or FDS with inverse FFT (IFFT) and an interpolation process are employed.

2.2 Parameter estimation

	Transmission	Receiver	Detection	Others
Wideband DL	burst TDS 100 ns up to 2 µs 10 ... 40 dBm E_0 = 1 nWs to 20 µWs	AGC or limiting receiver	coherent TDS. I / Q sampling. coherent integration for e.g. 128 radio request cycles	for high speed sensor applications with high resolution. data reduction by non periodic sampling
			incoherent detection envelope sampling	for low range. low cost ID purposes. data reduction by non periodic sampling
	stepped FDS e.g. 32x 20 µs. 0 ... 30 dBm E_0 = 0.3 ... 600 µWs	linear receiver	coherent FDS	for sensor and ID applications with low speed requirements but high distance range
Dispersive DL	burst $\tau << 1/B$	matched filter	incoherent detection envelope sampling	*experimental only*
Transponder DL (impedance loaded SAW two port device)	burst TDS 100 ns up to 2 µs 10 ... 40 dBm E_0 = 1 nWs to 20 µWs	AGC or limiting receiver	coherent TDS. I / Q sampling. coherent integration for e.g. 128 radio request cycles	for high speed sensor applications with high resolution.
			incoherent detection envelope sampling	for low range. high SNR applications. only less signal processing effort
Resonators (high-Q)	burst 10 ... 100 µs 0 ... 30 dBm	linear receiver	TDS. high rate sampling & digital MSE criteria processing	high resolution radio request of high-Q resonators with a relatively long off line processing time
		logarithmic receiver	Gated Phase locked loop (GPLL)	real time signal processing. SNR threshold. no multisensor capability

Table 2: Actually employed signal processing for passive radio sensors

2.3 Enhancement of signal's energy

For the low energy TDS methods, the signal energy for detection is enhanced by a factor N by coherent integration of N radio request cycles. By a linear operation, samples of signal amplitude $s_k(t)$ of different interrogation cycles are summed up with correct phase. Due to the linear process, no crosscorrelation between signal and noise occurs. Assuming additive white Gaussian noise, the crosscorrelation between different noise samples vanishes. While the mean signal energy is constant, the normalized noise power P_n after averaging with $\overline{n_k(t)^2}$ the mean power of the stochastic noise is

$$P_n = \frac{1}{N^2} \cdot \sum_{k=1}^{N} \overline{n_k(t)^2} = \frac{\overline{n_k(t)^2}}{N} \cdot \qquad (2)$$

If no correlation between signal and interference signal exists, coherent integration is also applicable for SIR enhancement. Feed through effects of local oscillator within the radio request system cannot be mitigated.

Nonlinearity causes an irreversible correlation between signal and noise terms. Above a threshold SNR, less effective non linear integration measures, e.g. post detection integration, can be applied.

2.4 Multiple access

Today, response signals of different SAW sensors mainly are distinguished in

- **space**, by use of narrow antenna beam widths. The whole code family can be exploited, with exception of a start and a stop bit.
- **time**, by time interleaving of reflectors of the DLs, a high receiver's dynamic range can be achieved, log. amplifiers can be employed.
- **frequency**, for SAWR, dual to time division.

3. Radio transmission

The propagation via the radio channel mainly is affected by:

- The signal attenuation due to the distance d, free space attenuation $a_{path} = (4\pi d/\lambda)^k$, with $\lambda = c/f$, actual exponents k are 2 (up to 4)
- Slow fading due to obstacled propagation, shadowing, etc.
- Rayleigh fading due to multipath propagation and interfering signal components
- Thermal and man made noise
- Interference caused by other RF systems

The sensor design has to consider the statistical channel parameters, typical delay spreads σ_τ are some µs for outdoor and up to 200 ns for indoor channels.

4. Measurement parameters and errors

The **sensitivity** of a sensor usually is defined as the change of the sensor readout due to the effect of certain parameters, e.g. 92 ppm/K for 128° $LiNbO_3$.

Resolution is the minimum change of measurement readout due to a changing measurand normalized to the full span, e.g. time delay. Employing DL and phase instead of envelope detection, it is enhanced from 10^{-3} to 10^{-6}. For amplitude detection using DLT, it is better than 1 % in worst case.

Accuracy, the fidelity of the system's readout to the true measurand and affected by errors, often is mixed up with resolution.

origin	effect	result
sensor	manufacturing	deviation of delay or frequency
	aging	
	bandwidth	deviation of amplitude and phase
system	tuning -gain (includes S/H hold drift, ...) -phase orthogonality	phase and amplitude error
	phase noise	
	bandwidth	assumed to be properly
radio channel	noise	phase and amplitude error

Table 3: Deterministic and stochastic errors for radio request of passive SAW sensors

Manufacturing and aging: The manufacturing tolerances are given to be smaller than 100 ppm = 10^{-4} for low cost SAW mass products [5], representing the accuracy limit for the SAW sensors without calibration. For resonant devices, the total aging remains smaller than ±50 ppm. Aging of delay lines is not specified, since it only depends on the almost neglectible aging of the crystal substrate.

Bandwidth: The bandwidth of measurement means the maximum bandwidth of the dynamic measurand allowed for admissible errors. It depends on the sensor, its environment, converting the measurand into an effect to the sensor, the transmission and the signal processing. For thermal measurements, the sensor's bandwidth is only about 1 Hz, for mechanical effects 10 kHz are achieved. In [9] the integration errors due to the signal processing, evaluating a delay, are estimated. A 3 µs delay line with a resolution of up to 10^{-6} can be operated up to a frequency of $f_{meas} \approx 3$ kHz to keep the additional bandwidth error smaller than 1 percent. Utilizing DLT, with amplitude detection, the measurand's bandwidth is up to $BW_{RF}/10$ (several MHz), limited to a much narrower range by the external impedance sensors from stock.

Receiver tuning: In the receiver coherent detection is performed. For actual systems, the base signals, the

received signal is multiplied by, have a deviation ε from exact orthogonality. The amplification and A/D conversion of the baseband signals is not exactly equal, but differ by a factor b/a (I'=I·a, Q'=Q·b).

Both, ε and b/a ≠ 1, yield phase and amplitude errors.

$$\Delta\varphi_{b/a} = \mathrm{atn}\left\{\frac{b\cdot Q}{a\cdot I}\right\} - \mathrm{atn}\left\{\frac{Q}{I}\right\} \qquad (3)$$

$$V'/V\big|_{b/a} = \sqrt{\left(\frac{b}{a}\right)^2\cdot\sin^2\varphi + \cos\varphi} \qquad (4)$$

$$\Delta\varphi_{\varepsilon} = \arctan\left\{\frac{\cos(\pi/2-\varphi-\varepsilon)}{\cos(\varphi)}\right\} - \varphi \qquad (5)$$

$$V'/V\big|_{\varepsilon} = \sqrt{\cos^2(\varphi)+\cos^2(\pi/2-\varphi-\varepsilon)} \qquad (6)$$

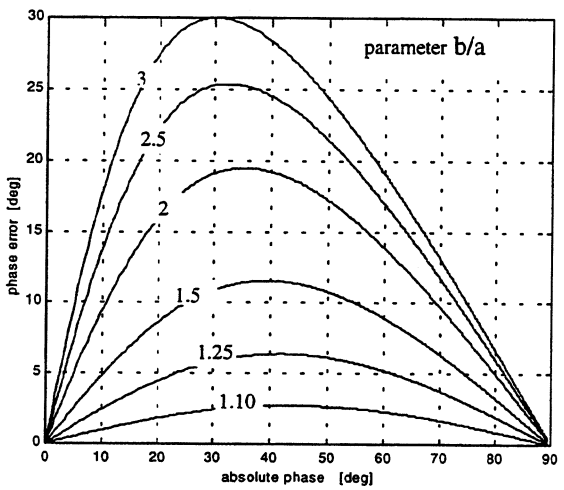


Figure 3: Phase error $\Delta\varphi_{b/a}$ due to different base band gain vs. absolute phase, parameter b/a

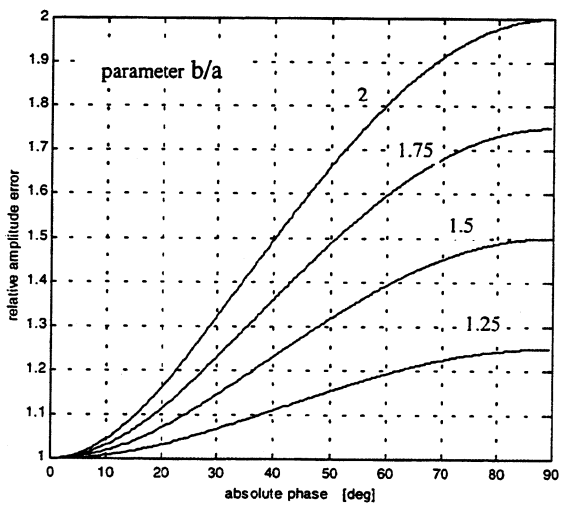


Figure 4: Amplitude error $V'/V|_{b/a}$ due to different base band gain vs. absolute phase, parameter b/a

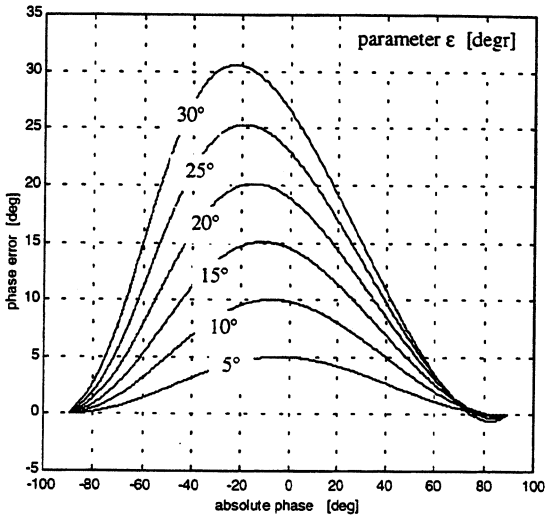


Figure 5: Phase error $\Delta\varphi_{\varepsilon}$ due the non orthogonality of the reference, parameter ε.

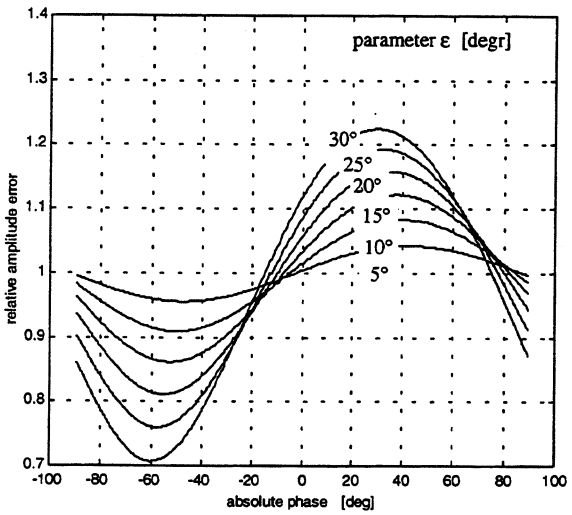


Figure 6: Amplitude error V'/V_{ε} due the non orthogonality of the reference, parameter ε.

Phase noise: The oscillators used for signal generation and coherent detection are afflicted with phase noise. Since multiplication is performed, the phase errors are considered only. We find the rms phase deviation due to the phase noise in a bandwidth 2B

$$\sigma_{\theta}^2 = 2\cdot\int_0^B S_{\theta}(f)\mathrm{d}f \cdot \qquad (7)$$

Dividing the spectrum into subbands, a mask for the admissible phase noise of the oscillators is gained. For a DL with a ΔT up to 10 µs, a flat course of noise density of approx. -90 dBc between 10 kHz and 1 MHz and increasing with 20 dB/decade towards lower frequencies is required to keep the total error smaller than 1 degree.

Integration error: Coherent integration yields an error, if the measurand is not stationary, but has a phase shift γ between each pair of the N samples. The total amplitude error $E_{a,N}$ is

$$E_{a,N} = \frac{N - 2 \cdot \sin\left[\left(\frac{1}{2} + N\right) \cdot \frac{\gamma}{2}\right]}{N \cdot \sin\left(\frac{\gamma}{4}\right)} . \qquad (8)$$

Additive noise and interference because of radio transmission: During radio transmission, noise (assumed to be a white stochastic process with Gaussian probability density function of amplitude, AWGN) and interference from other RF radio systems are added. The probability, that the measurement error exceeds $k \cdot A_m$, is calculated with the Q function from

$$P_{amplitude\_error} = 2 \cdot Q\left\{k \cdot \sqrt{SNR}\right\} \text{ for } k \geq 0. \qquad (9)$$

For small phase deviations due to a large SNR, the probability to exceed a phase error Θ_r is

$$P_{\Delta\varphi \geq \Theta} \approx 2 \cdot Q\left\{\sqrt{2SNR} \cdot \sin\Theta\right\}, \qquad (10)$$

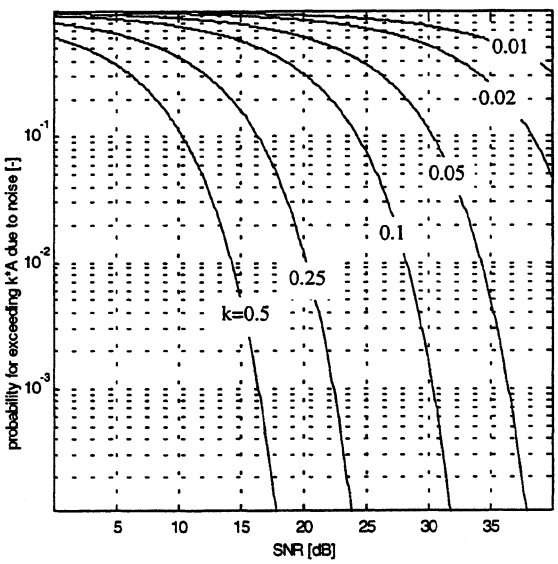


Figure 7: Probability of measurement error versus SNR for errors exceeding $k \cdot A_m$ with parameter k.

Investigating interference, coherent jamming, usually caused by feed through of local oscillators, etc., and incoherent jamming is distinguished. The amplitude error is $\pm (I_c+R_0)/R_0$. A ratio of -40 dB yields a maximum error of 1% in amplitude. The maximum phase error is $\pm\arcsin\{I_c/R_0\}$. Incoherent jammers have to be considered by $I_c \rightarrow (I_i/\sqrt{2})^2$.

Discussion

In literature, sensitivity, accuracy and resolution usually are given simultaneously with a maximum distance range for radio request. No errors due to a

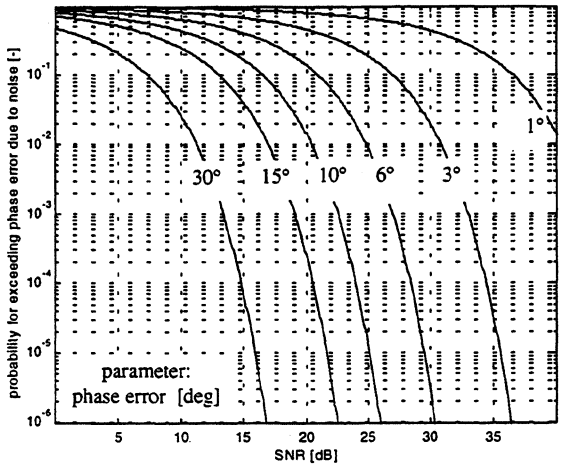


Figure 8: Probability of phase error ($\Delta\varphi > \Theta_r$) versus the SNR, with parameter Θ_r

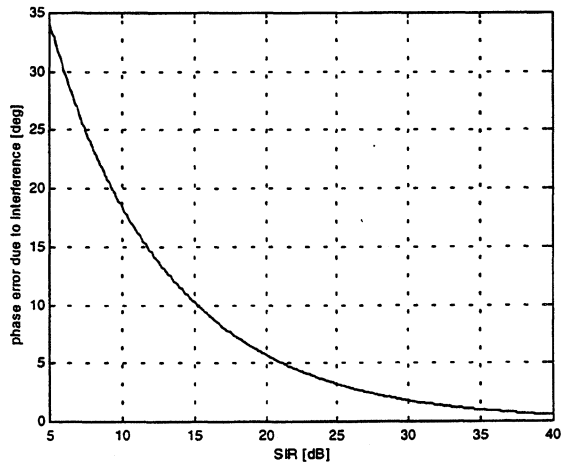


Figure 9: Phase error due to interference versus SIR

signal to noise ratio decreasing with distance and other noise contributions are considered. For an actual system design, the errors have to be considered, of course. For example, a $LiNbO_3$ DL temperature sensor is employed with ΔT = 5 µs, operating at 869 MHz. A resolution of 10^{-5} in >99.9 % of samples should be achieved. A phase of 2π corresponds to $2.30 \cdot 10^{-4} \cdot 5$ µs, a maximum phase error of 15 degree is admissible.

Standard adjustment of low cost systems, e.g. 10% I/Q unbalance b/a and a 5 degree phase misadjustment ε, cause up to approx. 6 degree phase error, the maximum resolution without further effects is approx. $4 \cdot 10^{-6}$, only to be exceeded by exact tuning of the system. For 10^{-5}, a margin of 9 degree for the phase error can be used up. The signal to noise ratio, necessary to keep the error limit in 99.9%, can be found to be higher than 21 dB. The maximum distance range is calculated from the link budget. Assuming a gain of the antennas G_{sens} [dB] at the sensor and G_{sys} [dB] at the request system, a transmitted power P_t [dBm], a sensor attenuation a_{sensor}

[dB] and a coherent integration of N request cycles, the maximal path attenuation a_{path} [dB] is calculated from

$$a_{path} \le P_t + 2 \cdot G_{sys} + 2 \cdot G_{sens} - a_{sensor} - SNR + 10 \cdot \log(N) - \\ - M - F_{mm} - N_0 - 10 \cdot \log(B / 1Hz) - NF$$

The term $N_0 = k \cdot T$ [dBm/Hz] means the thermal noise on the radio channel, F_{mm} [dB] the additional noise figure due to the collected man made noise. The term NF [dB] is the noise figure of the receiver, B [Hz] its bandwidth. M [dB] represents a margin for slow fading due to shadowing, etc. For our example, we assume P_t = 20 dBm, $G_{...}$ = 0 dB, a_{sens} = 20 dB, N_0 = -174 dBm/Hz, B = 1.7 MHz, NF = 10 dB, M = 10 dB and N = 1. Then, the sum a_{path}+SNR is calculated to be 90. For one direction, a path loss of 34.5 dB is allowed. A path length of 1.7 m without any obstacle and without destructive interference due to multipath propagation can be bridged. Enhancing the energy or reducing the bandwidth and therefore reducing the noise power, the distance range is enhanced. A gain of 12 dB in SNR yields a doubling of range.

Employing a **measurement of amplitude**, e.g. for radio request of DLT, a very similar calculation yields a further degradation due to noise and interference and therefore a much higher required SNR.

For **frequency measurement**, the relative frequency error $\Delta f / f$ can be calculated from the phase error $\Delta\varphi$: Measurement of one period length is afflicted with an error in frequency of $\Delta\varphi / 2\pi$.. If the phase is unambiguous, the number of periods is counted for a time interval T_m, the relative frequency error is calculated from

$$\frac{\Delta f}{f} = \frac{\Delta\varphi}{2\pi} \cdot \frac{1}{T_m \cdot f} . \qquad (11)$$

For f = 433 MHz, T_m = 5 µs and $\Delta\varphi$ = 15 degree, the frequency error is approx. 20 ppm.

Conclusion

In the work, the state of the art for the radio request of passive sensors have been discussed. A division into frequency domain sampling and time domain sampling was done. For high resolution, high speed, short range applications wideband delay line sensors advantageously are combined with a burst TDS. To enhance the energy, spread spectrum methods and signal integration are feasible. Applying the proper radio request with comparable total signal's energy and signal processing method, TDS and FDS yield the same performance of measurement for applications without limitations of measurement time. For multiple access to a number of sensors space and time division are actually proven. Investigations about the deterministic and stochastic errors have been presented. The effects of the error sources to amplitude, phase and frequency measurements have been calculated. Finally, an estimation of the distance range for a given reliability was done. It was found, that phase detection is much more robust against misadjustment of the receiver and against noise errors. Therefore it is suggested to be used preferably.

Passive SAW radio sensors have been presented first only a few years ago. Since then, the field of applications is growing. Today, the first industrial setups are in development. In a world of increasing demand for sensors, it is foreseeable, that they will find a fixed position for special applications, other technologies cannot be employed for.

Acknowledgment

The author thanks Prof. Franz Seifert for discussions and scientific advice. Gratefully acknowledged are the support and stimulating discussions with Dr. L Reindl, Dr. C.C.W. Ruppel and W.E. Bulst, form Siemens ZT KM 1, Munich, Germany. Last but not least, I thank all colleagues, accompanying the activities.

References

[1] P.A. Nysen, H. Skeie and D. Armstrong, "System for interrogating a passive transponder carrying phase-encoded information", *US Patent Nos. 4725841; 4 625 207; 4 625 208* (1983-1986).

[2] W. Buff, SAW Sensors, IEEE Transactions on Sensors and Actuators A42 (1992), pp. 117-121.

[3] F. Seifert, W.E. Bulst and C.C.W. Ruppel, Mechanical sensors based on surface acoustic waves, IEEE Transactions on Sensors and Actuators, A44 (1994), pp. 231-239.

[4] A. Pohl, F. Seifert, 'Wirelessly Interrogable SAW Sensors for Vehicular Applications", IEEE Transactions on Instrumentation and Measurement, Vol. 46 (1997), No. 4, pp. 1031-1038.

[5] Siemens Matsushita, SAW components, 1996.

[6] A. Pohl, G. Ostermayer, F. Seifert, "Wireless Sensing Using Oscillator Circuits Locked to Remote High-Q SAW Resonators", IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control, Special issue on sensors and actuators, Vol. 45, No. 5, September 1998, pp. 1161-1168.

[7] A. Pohl, "A Low Cost High Definition Wireless Sensor System Utilizing Intersymbol Interference", IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control, Special issue on sensors and actuators, Vol. 45, No. 5, Sep. 1998, pp. 1355-1362.

[8] G. Ostermayer, A. Pohl, R. Steindl, F. Seifert, "SAW sensors and correlative signal processing – a method providing multiple access capability", Proceedings IEEE ISSSTA 98, South Africa, *in print.*

[9] A. Pohl, F. Seifert, "New Applications of Wirelessly Interrogable Passive SAW Sensors", IEEE Transactions on Microwave Theory and Techniques, Vol. 46, part II, No. 12, Dec. 1998, pp. 2208-2212.

THE INTERFEROMETRIC FREQUENCY MEASUREMENT WITH VARIABLE ACOUSTOOPTIC DELAY LINE

Kludzin V.V., Kulakov S.V., Molotok V.V., Preslenev L.N.

St.Petersburg State University of Aerospace Instrumentation,
67, Bolshaya Morskaya, St.Petersburg, Russia, 190000

ABSTRACT

The conditions for operation of a two-beam interferometer based on an acousto-optic variable delay line are studied. The accuracy of RF signal frequency measurement is estimated, the main physical factors affecting the real measurement accuracy are analyzed. The requirements to the parameters of the acousto-optic delay line are formulated. The possibility to extract and measure the clock frequency of binary pulse trails for synchronizing a receiver in digital communication systems is shown theoretically and experimentally

1. INTRODUCTION

Efficient measurement methods based on optical interferometers of different kinds are widely used for processing optical signals. A tunable interferometer can be used to accurately measure the frequency of monochromatic optical signal [1]. The output result of the interference summation of the direct and lagged optical beams can be easily obtained by optics as the optical wavelength are short. This idea can be also used in the interference systems measuring the RF signal frequency where the variable RF signal delay line with a controllable variation law is the key element.

The pulse-phase method for accurate measurement of delay time is known in the ultrasonic measurement technology [2]. This method is implemented when the direct and lagged signals are summarized when the input signal frequency is varied within some range $\Delta\omega$. In this case, the interference component of the output signal has N extreme values and the measurement conditions has a simple form $\Delta f\tau=N$. Thus by measuring the frequency range Δf and counting the number of the extreme values, the delay time τ can be found rather precisely. Obviously, the inverse problem can also be solved. If a variable delay line is used, and the delay time is varied within the range $\Delta\tau$, the frequency f can also be found as $f=N/\Delta\tau$. So, the condition of operation for an interferometer can be generalized in the form of $\Delta(f\tau)=N$ where the variation symbol Δ can be applied to any parameter in the parenthesis.

2. THE THEORY OF INTERFEROMETER

Let us look at the scheme of a simplest two-beam interferometer (Fig.1) where VDL is a variable delay line.

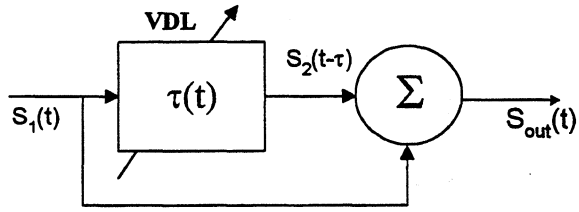


Fig.1. Two-beam interferometer.

Let us assume that the input signal $S_1(t)=S_1\cos(\omega t)$ is a harmonic with an unknown frequency $\omega=2\pi f$. The signal at the output of the variable delay line is $S_2(t-\tau)=S_2\cos[\omega(t-\tau)]$, and correspondingly,

$$S_{out}(t)=A(\tau)\cos[\omega t-\varphi(\tau)]$$

$$A(\tau)=\sqrt{\left(S_1^2+S_2^2+2S_1S_2\cos(\omega\tau)\right)}$$

$$tg[\varphi(\tau)]=\frac{S_2\sin(\omega\tau)}{S_1+S_2\cos(\omega\tau)}$$

Thus, the interference output signal of the summator obtains the envelope $A(\tau)$ and phase shift $\varphi(\tau)$, these parameters containing the entire information. In order to count the number of extremums N, it is better to use the minimal values of the envelope $A(\tau)$ when $S_1=S_2$. At that, $S_{out}=2S\left|\cos(\frac{\omega\tau}{2})\right|\cos\left[\omega(t-\frac{\tau}{2})\right]$, and the envelope derivative $\frac{dA(\tau)}{d\tau}$ has the discontinuities of the first kind

The potentially possible measurement precision δf is determined by the entire interval of delay time variation τ_m in a simple form $\delta f=\frac{1}{\tau_m}$, and the relative error $\Delta=\frac{1}{N}$. At practice, the temperature instability of acoustic velocity $\left(\partial V(T)/V\right)$ and

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

attenuation and dissipation effects of elastic energy can considerably affect the measurements.

3. EXPERIMENTAL SET-UP

The optical scheme of the system for measuring the RF signal frequency that uses an acousto-optic delay line is shown in Fig.2. The optical light from a He-Ne laser (λ=0.63 μm) is formed by Lens L1 as a collimated beam of D1 size and is directed onto the optical scanner (a rotating mirror with rotation speed of Ω). Then Lens L2 focuses the scanned optical bean onto the axis of Acousto-Optic Cell (AOC). After the interaction with an acoustic wave into the AOC, the overlapped diffracted and non-diffracted optical beams are directed with Lens L3 to Photo Receiver (PR), where the lagged RF signal $S_2(t-\tau)$ is formed and obtained. After the input and lagged signals are summarized (Σ), the envelope (Dt) is obtained and the number of extremums N in one measure cycle is counted.

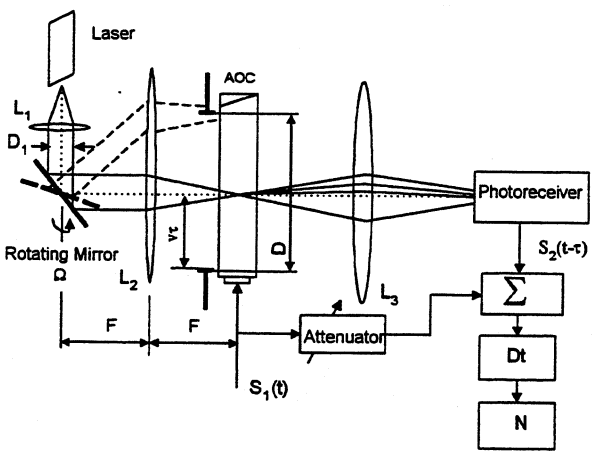


Fig.2 The scheme of a system measuring the RF signal frequency.

Note that the time scale in this scheme can be arbitrary depending on the mirror's rotation speed in the optical scanner, scanner parameters not affecting the measurement precision of RF signal frequency.

The delay time $\tau(t)$ in this system varies linearly in the interval $[0...\tau_m]$ in the time T, and $\tau(t)=\frac{\tau_m}{T}t, \quad 0\le t\le T, \ \tau_m=\frac{D}{v}$, where D is the optical aperture of AOC given by the input diaphragm, v is the acoustic velocity. The scale coefficient $p=\frac{\tau_m}{T}$ is determined as $p=\frac{\Omega D}{2v\cdot arctg(\frac{D}{2F})}$, F is the focal length of Lens L2, and it determines the envelope of the output signal $A(\tau)$ in the form

$$\frac{A(t)}{A_0}=\sqrt{(1+\exp(-2\alpha vpt)+\exp(-\alpha vpt)\cdot\cos(p\omega t))},$$

$$0<t<T$$

Here, the condition S_1=S_2=A_0 is used, and the exponential component reflects the acoustic attenuation along the AOC aperture, α is the attenuation coefficient depending on the properties of the medium of acousto-optic interaction, acoustic anisotropy of the elastic mode and frequency of input signal.

Fig.3 shows the envelope fragments calculated for different attenuations and delays (αvpt = 0;0.5;1.0).

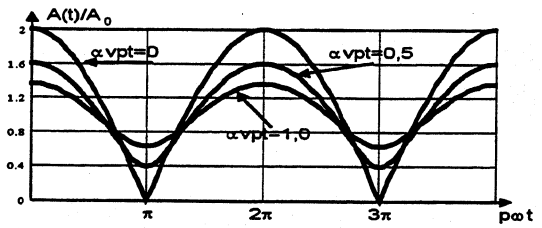


Fig.3. The influence of the elastic wave absorption on the interference signal envelope shape

The acousto-optic cell was made of the TeO_2 crystal with slow shear mode travelling along the [110] direction. This acoustic mode has low velocity (v=0.62·10^6 mm/sec) and thus provides the maximum time delay per crystal length unit. The piezoelectric transducer generating the shear wave was the platelet of X-cut $LiNbO_3$ crystal bonding to the medium of interaction by a thermocompression metallic bonding. The optical beam propagated near the [001] direction and obtained the circular polarization in the medium. The main parameters of the acousto-optic cell are presented in the Table.

Frequency band, MHz	Maximum delay, μsec	Attenuation at the center frequency, dB/μsec	Interaction efficiency, P=0.1 W	Velocity thermal coefficient $\frac{1}{v}\frac{\partial v}{\partial T}$, 1/°C
25-45	16	0,2	0,25	2,5*10^{-4}

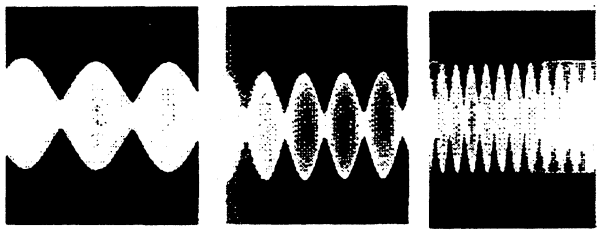
The attenuation in the AOC is caused generally by the effects of the abnormal spread of the slow shear wave due to a large acoustic anisotropy in the TeO_2 crystal.

Optical heterodyning is made by forming the optical beam with a wide angular spectrum [3]. At that, the following condition should be fulfilled.

$$\frac{\lambda f}{v}<\frac{D_1}{F}, \quad if \quad Q<2\pi$$

where $Q = \frac{2\pi\lambda l f^2}{nv^2}$ - is the diffraction parameter [4].

Fig.4 shows the fragments of the interferometer's output signal (without detecting) obtained under different angular velocities of the scanner's mirror (Ω=600, 1000, 2460 rpm and p=0,004; 0,0068 and 0,016 respectively). The measured frequency of the harmonic signal is f=34,01±0,03 MHz, the reference being 34,026 MHz. It is worth noticing that the basic frequency of the output signal has the Doppler shift $\pm\frac{pf}{2}$, caused by the optical beam moving along the AOC aperture.



Ω=600 rpm Ω=1000 rpm Ω=2460 rpm

Fig.4. Oscillograms of the interferometer output signal for different mirror rotation speeds.

The acousto-optic interferometer can be used to estimate and measure the instant rotation velocity of different systems.

4. MEASUREMENT THE CLOCK FREQUENCY OF DIGITAL RF SIGNAL

In the modern communication systems, i.g. in the satellite systems, digital signals like a pulse trails which amplitude, or phase, or both, are changed according to the transmitted message are used. For example, multilevel phase shift keying signals are widely used. The receivers in such systems are synchronized by yielding the repeating frequency (clock frequency) of the elementary information pulses and triggering the receiver's resolver. It is important that when the pulse trails is random and the probabilities of the states (0 or 1) are equal approximately 0.5, there is no any discreet frequency components in the signal spectrum [5, 6]. That is why a special signal processing is required that yields the clock frequency and synchronizes the receiver.

Apart from the communications tasks, the problem of clock frequency measurement can also appear in the electromagnetic surveillance systems, when communication equipment is being tuned, etc.

The problem of yielding the clock frequency of a digital signal can be efficiently solved by multiplying the information stream with its copy delayed by a certain time generally less then the half clock period. At that, the multiplication product contains very short components at all signal inversion points (transitions from 0→1 or 1→0) and such signal has discreet frequency components at clock frequency harmonics in its spectrum regardless of the signal structure.

Fig.5 shows the functional diagram of the algorithm that measures the clock frequency of continuous pulse sequences (BPSK). The universality of this algorithm is determined by the possibility to adjust delay time according to the signal with different parameters. The accuracy and stability of the delay line are determining in this case. Those limitations in practical implementation of acousto-optic delay line lead to using of acousto-optic materials with relatively high acoustic velocities, such as GaP, $LiNbO_3$, quartz, and different glasses.

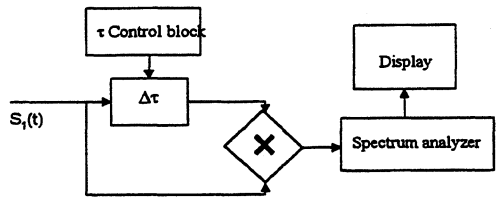


Fig.5. The functional scheme of clock frequency measurement

An acousto-optic delay line made of tellurium glass operating at 60-80 MHz is used in the experimental set-up. Fig.6 shows the spectrograms from the output of the spectrum analyzer for 2 delay times, the input signal being a BPSK signal with the clock frequency of 2 MHz. For the comparison, this figure also shows the spectrum of the input signal that contains no discrete spectral components.

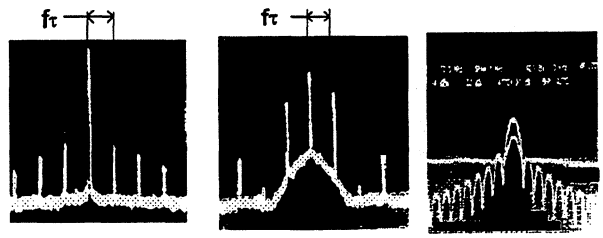


Fig.6. The spectrograms of the output signal from the acousto-optic delay line-multiplier.

The conducted research practically prove that digital signals with the clock frequencies up to 50 MHz can be processed with low signal-to-noise ratio. The limitations in such measurements are determined generally by the finite time of acoustic signal traveling

through the focused optical beam and finite bandwidth of the AOC transducer.

5. CONCLUSION

The general condition of operation of the two-beam interferometer $\Delta(f\tau)=N$ where the variation symbol Δ can be applied to any parameter in the parenthesis is presented.

The possibility to use the interference method for measuring RF signal frequency based on a variable acousto-optic delay line is theoretically and experimentally proved. The measurement precision is determined by obvious physical limitations - maximally possible delay time.

This method and implementation scheme can be used for measuring the instant changing position parameters (linear or angular).

The possibility to use the variable delay line for obtaining and measuring the clock frequency of complex binary code sequences is discussed. This technique can be used for receiver synchronization in communication systems

6. REFERENCES

[1] Born M., Wolf E. Principles of Optics. Pergamon Press, 1964

[2] Physical Acoustics. Vol.1, part A, Edited by W.P.Mason, Acad.Press, 1964.

[3] Kludzin V.V., Preslenev L.N., "RF pulse coherent store in acousto-optic delay lines". In:Acousto-Optic Methods and Systems for Information Processing, Leningrad, 1980, pp.54-60, (in Russian).

[4] Design and Fabrication of Acoustooptic Devices. Ed. by Goutzoulis A., Pape R.D., Marcell Dekker Inc., 1994

[5] Franks L.E. Signal Theory. New York: Prentice Hall Inc., 1969.

[6] Kulakov V.S., Nikitin Yu.I., Nikiforova E.Yu., Preslenev L.N., "Yielding the PSK Signal Clock Frequency by an Acousto-Optic Delay Line Multiplier", Radiotekhnika, #6, 1990, pp.83-86, (in Russian).

1999 Joint Meeting EFTF - IEEE IFCS

THE VIA VIBRATING BEAM ACCELEROMETER : A NEW QUARTZ MICROMACHINED SENSOR

O. Le Traon, F. Deyzac, D. Janiaud, S. Muller
ONERA
BP 72 - 29 avenue de la Division Leclerc
F-92322 CHATILLON Cedex
Tel. : 33 1 46 73 48 35 - Fax : 33 1 46 73 48 24
E-mail : letraon@onera.fr - http://www.onera.fr

INTRODUCTION

For a few years, ONERA, the French National Establishment for Aerospace Research, under the financial support of the DGA (The French MoD Procurement Agency), has been developing a micro-machined quartz vibrating beam accelerometer called VIA (Vibrating Inertial Accelerometer). This sensor exploits the frequency change of a quartz resonator when submitted to an acceleration.

The main targeted applications are the guidance and the attitude control of tactical missiles, as well as the inertial navigation of aircraft, helicopters and ground vehicles (coupled with radio-electric positioning systems such as the GPS receiver).

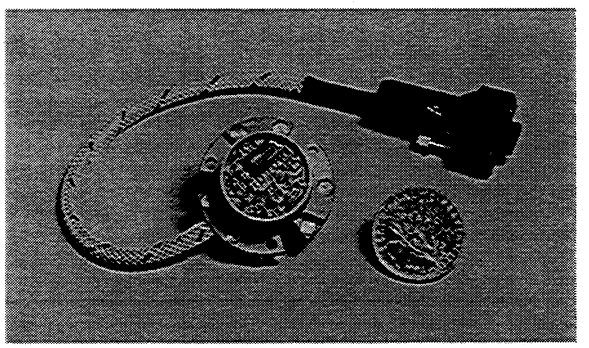
In those fields, industry requires accurate miniature accelerometers (typically 10^{-5} to 10^{-6} of the measurement range), with low cost manufacturing and thus, important research and development activities are carried out by many laboratories[1-4].

The VIA concept has been selected to meet these objectives. First, the monolithic configuration of the transducers is well suited to chemical etching process and miniaturization. Secondly, the choice of vibrating beams as sensitive elements, combined with excellent quartz mechanical properties, gives a very good stability of the scale factor. Third, the performed optimisation of the transducer leads to a very good insulation of the active part with respect to the mounting areas of the transducer, suitable to bias stability.

In order to reduce the thermal sensitivity, the accelerometer is composed of two transducers operating in differential mode.

ACCELEROMETER CONFIGURATION

The VIA accelerometer, shown on picture 1, is made of a sensor and two electronic oscillator circuits which maintain the two vibrating beams at their resonance frequencies by piezoelectric effect. The opposite variations of these two frequencies provide a direct measurement of the applied acceleration.

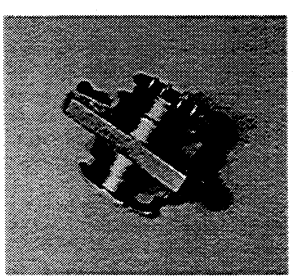


Picture 1 :
VIA accelerometer.

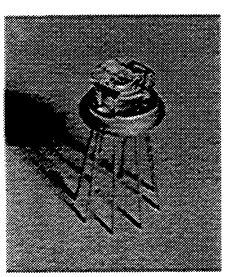
The output signal of the transducer being a frequency, the data is almost completely digitized and thus an analog-digital converter is not required.

The sensor is composed of two quartz transducers mounted into a cooper cylindrical case (picture 2.1). Each transducer is fixed on a standard TO5 base (picture 2.2), and the two bases are mounted at the opposite ends of the cooper case. The transducers operate under vacuum. Their diameter is about 6 mm and the sensor size is about 2 cm^3.

Thanks to the use of Surface Mount Technology, the overall accelerometer dimensions lead to a 10 cm^3 volume only.



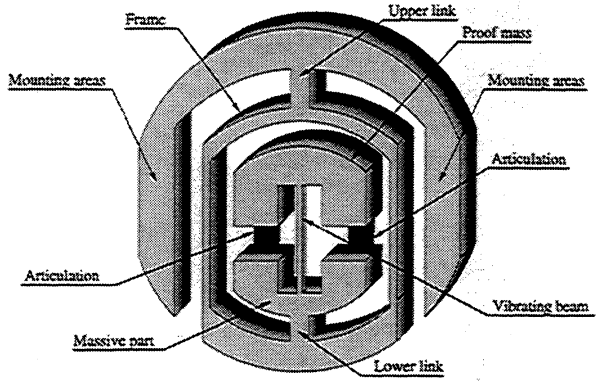
Picture 2.1 :
Sensor composed of 2 transducers



Picture 2.2 :
Transducer fixed on a TO5 base.

TRANSDUCER DESIGN

Picture 3 shows the configuration of the transducer.



Picture 3 :
Transducer design

The active part of the transducer is located in the center of this picture and consists of a proof mass, another massive part, a vibrating beam and two articulations. The beam vibrates in the plane of the chip and the sensitive axis is quasi perpendicular to the chip.

When the transducer is submitted to an acceleration (along its sensitive axis), the proof mass applies tensile or compressive stress to the beam which modifies its resonance frequency. Thus, the measurement of the beam frequency allows to know at any time the applied acceleration.

The resonant frequency of the vibrating beam is about 60 kHz and its sensivity to acceleration is about 12 Hz/g.

The main originality of this design lies in the insulating system (i.e the frame and the upper and lower links), which connects the active part to the mounting areas [5].

This insulating system has two qualities :

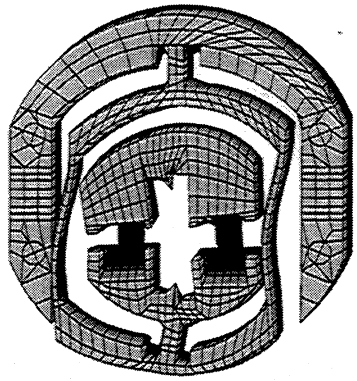
- it *insulates the beam vibrations* from the mounting areas of the transducer,
- it *protects the active part of the transducer from thermal stresses* due to the difference in thermal expansion of quartz and of the case material.

Insulation of the beam vibrations :

When the beam vibrates at its resonance frequency (in the plane of the chip), all the structure is submitted to forced vibrations.

In order to avoid any energy loss and to preserve the high quality factor of the resonator, the mounting areas of the transducer must stay motionless as far as possible.

The insulating system, thanks to its flexibility for the main dynamic excitations, fulfils these requirements. The efficiency of this system is illustrated in picture 4 which represents with much amplification the vibrating displacement of the free - free transducer (i.e. free - free boundary conditions) when the beam vibrates at its resonance frequency. Because of its vibrating amplitude, the beam cannot be represented on the picture. Despite the large amplification, the vibrating movements of the mounting areas appear negligible.



Picture 4 :
Vibrating displacement in free - free boundary conditions

So, when the mounting areas are fixed on a base, the energy loss through the fixation will be very small, and the beam frequency will not be modified by the fixation.

Precisely, a way to appreciate the quality of the vibrating beam insulation is to compare the theoritical frequency values **F** and **F'** of the vibrating beam for two different transducer boundary conditions : first, free-free boundary conditions (picture 4) and secondly transducer fixed on its mounting areas. A very small ratio **(F - F') / F** means a very good vibrating beam insulation. In our case this ratio is about 10^{-7}, which is an excellent result.

The insulation of the beam vibrations is an important point with regards to performance, and especially for the repeatability of the sensor parameters.

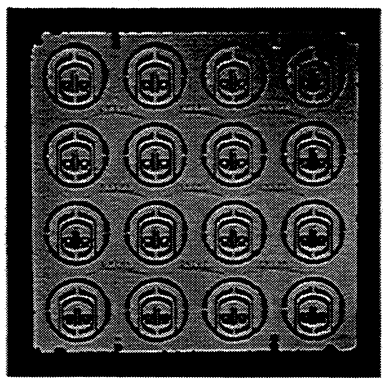
Insulating against thermal stresses :

The transducer is designed to be mounted on a metallic base ; so in the presence of temperature variations, some stresses appear in the whole structure due to differences in the expansion coefficients between quartz and the base material. The active part of the transducer and in particular the vibrating beam must be preserved from those stresses.

As shown in reference [6], the addition of the two links in the insulating system very efficiently preserve the active part from thermal stresses. The residual compressive or tensile stresses in the beam are negligible, so its resonance frequency is not modified.

Manufacturing Process

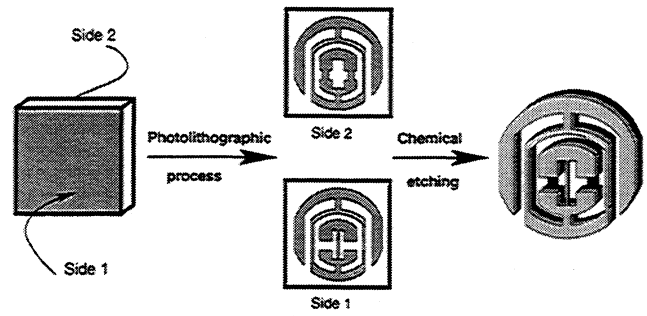
The picture 5 shows a quartz wafer including sixteen chemically-machined transducers. In spite of the three dimensional conformation of the transducer, the wafer is engineered in one step. As a matter of fact, the technique developed by ONERA takes avantage of the etching depth to adjust the machining about the third dimension (perpendicular to the faces of the wafer).



Picture 5 :
Quartz wafer with 16 transducers.
(dimensions : 38,1 mm x 38,1 mm)

Pictures 6.1 to 6.3 illustrate the transducers manu–facturing process :

Picture 6.1 shows a part of a wafer where one transducer will be etched. The two main sides of this wafer are metallized (chromium-gold metal layers).



Picture 6.1 :
Metallized wafer.
(chromium-gold layers)

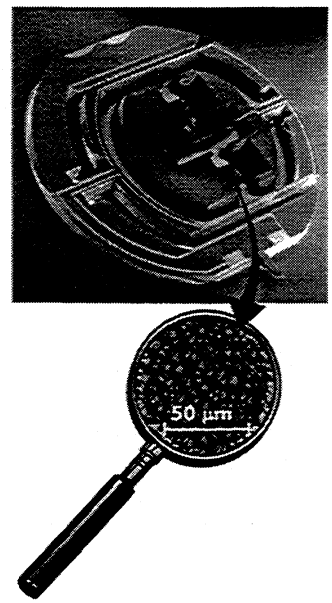
Picture 6.2 :
Photolithographic process
Etching masks of the sides

Pictures 6.3 :
Obtained structure after only one step of chemical etching

Photolithographic technics permit the etching masks to be obtained from the initial metallization of the wafer (picture 6.2). Note that in our case, the etching masks are different for the 2 sides : the vibrating beam mask is on side 1 and the articulations mask on side 2.

The wafer is etched using a fluoride-based solution and the etching is stopped before its depth attains the thickness of the wafer (picture 6.3). The difference between the etching depth and wafer thickness is equal to the thickness of the vibrating beam and articulations.

An adapted chemical etching procedure has been defined in order to answer the specificity of the VIA structure, in particular concerning the roughness of the etched surface. The quality of this surface (picture 7), which defines one of the faces of the vibrating beam, is very important with regard to the vibrating beam behavior (i.e. thermal and acceleration sensitivity).



Picture 7 :
Transducer obtained by chemical etching and roughness after an etching depth of 370 µ

In addition, the chromium-gold mask used for the chemical etching engineering allows also to provide very easily the excitation electrodes of the vibrating beams.

This collective micro-machining technology, developed for the VIA accelerometer, is a very simple process, well suited to low cost manufacturing.

This technique is very close to the one used in the watch-making industry, which guarantees a very low manufacturing cost.

PERFORMANCES

The precise evaluation of the VIA accelerometers is performed at LRBA (DGA Lead Technical Centre for Navigation Systems).

Its main performances are as described in table I.

Measurement range	$\pm$100 g
Bandwidth	> 1000 Hz
Acceleration behavior ($\pm$100 g) :	
Scale factor	$\simeq$ 24 Hz/g
Non-linearity K_2	2 $\mu g/g^2$
Non-linearity K_3	0.02 $\mu g/g^3$
Residual error	150μg
Thermal behavior (-45 ° C to 90 ° C) :	
Bias sensitivity	$\simeq$ 100μg/ ° C
Bias hysteresis	100 μg
Scale factor sensitivity	1 ppm / ° C
Scale factor hysteresis	10 ppm

Table I :
VIA performance.

Full results concerning the evaluation of a VIA accelerometer are provided in reference [7]. Moreover, seventeen months of evaluation have proved the scale factor to be very stable (8 ppm standard deviation) and the bias to be very repeatable (90 μg standard deviation).

As a whole, the VIA accelerometer accuracy is around 300 μg, all errors combined, for a measurement range of $\pm$100 g and in harsh conditions representative of tactical missile environment.

CONCLUSION

The VIA accelerometer concept allows to combine the objectives of good performance and low cost. Indeed, the original design of the transducer takes into account the main insulation requirements for the vibrating beam accelerometers, and the chemically-machined manufacturing, being a very straightforward process, should keep very competitive costs.

This accelerometer has now reached maturity and transfer to industry for civil or military applications is currently under way.

However, the potential of performance has not been fully exploited yet. It is the reason why research activities are still going on, with an emphasis on increase accuracy and miniaturization.

REFERENCES

[1] N. Barbour, J. Connelly, J. Gilmore, P. Greiff, A. Kourepenis, M. Weinberg (Draper Laboratory), "Micro Electromechanical Instrument and Systems Development at Draper Laboratory", in *Proceedings of the 3rd St Petersburg International Conference on Integrated Navigation System*, mai 1996.

[2] M. Helsel, G. Gassner, M. Robinson, J, Woodruff (AlliedSignal Inc.), "A Navigation Grade Micro-Machined Silicon Accelerometer" in *Proceedings of the Position Location and Navigation Symposium*, april 1994.

[3] D.W Burns, R.D Horning, W.R Herb, J.D Zook, H. Guckel (Honeywell Technology Center),"Sealed Cavity Resonant Micro-Beam Accelerometer" in *Proceedings of Transducers'95*, Stockholm, june 1995.

[4] T.V Roszhart, H. Jerman, J. Drake, De Courtis, (Kearfott),"An Inertial Grade Micro-machined Vibrating beam Accelerometer", in *Proceedings of Transducers'95*, Stockholm, june 1995.

[5] O. Le Traon, D. Janiaud, S. Muller (ONERA), "Transducteur monolithique d'accélération" french patent n° 2739190 (published 28 march 97), international patent n° PCT/FR96/01481, deposed by ONERA).

[6] O. Le Traon, D. Janiaud, S. Muller (ONERA), P. Bouniol (LRBA), "The VIA Vibrating Beam Accelerometer : Concept and Performances" in *Proceedings of the Position, Location and Navigation Symposium*, Palm Springs, april 1998.

[7] P. Bouniol (LRBA), D. Janiaud, O. Le Traon, S. Muller (ONERA), "Performance Evaluation of the VIA Accelerometer" in *Proceedings of ION National Technical Meeting*, january 1998.

1999 Joint Meeting EFTF - IEEE IFCS

QUARTZ RESONANT MICRO-STRUCTURES AS SENSING ELEMENTS FOR TEMPERATURE SENSORS

T.G. Leblois and C. R. Tellier
Laboratoire de Chronométrie, Electronique et Piézoélectricité

Ecole Nationale Supérieure de Mécanique et des Microtechniques
26 Chemin de l'Epitaphe – 25030 BESANCON – France

ABSTRACT

This paper describes the development of a temperature sensor based on a bulk acoustic waves micro-resonator vibrating in thickness shear mode. Two different cuts are chosen taking into account the thermal sensitivity, the electromechanical coupling factor, the trapped energy conditions. Care is also taken of technical problems caused by anisotropic chemical etching fabrication process. The active element of the structure is a circular thin plate suspended by four bridges micromachined in a $NH_4F.HF$ solution. The performances of resonators are tested in the range 20-100 C. Experimental results on Q factor and thermal sensitivity for our suspended resonators appear to be encouraging when we compare with experimental values obtained with a bridgeless resonator and with theoretical results.

1. INTRODUCTION

Some temperature sensors using quartz resonators have been already developped [1] because this type of sensors has a good precision, repeatibility and the output signal is easy to process digitally. Besides, it can be used in a large range of temperature. Yet, only three cuts are usually used for this application : the doubly rotated cut LC and two singly rotated cuts $(\theta_0 = 40°\text{and}\ \theta_0 = 0°)$ [2]. The associated resonators are based on a tuning-fork structure fabricated by the so-called micromachined process. The aim of our research is to fabricate by micromachining " suspended resonators " whose advantages are firstly to reduce considerably spurious vibrations due to the fixation of the vibrating plate and secondly to increase the metrological performances.
So, this paper can be divided into three parts :
(i) firstly, we present theoretical results concerning thermal sensitivity of BAW resonators, piezoelectric parameters like electromechanical coupling coefficients, Q factor ...using the equation of propagation of waves in the case of endless planar plate.
(ii) secondly, we present and comment the etching properties in order to determine interesting orientations of plates.
(iii) finally, we propose a fabrication process and analyse experimental results, which are compared with theoretical predictions and with results obtained with a bridgeless resonator.

2. THEORETICAL FORMULATION

2.1 – Thermal sensitivity :
We follow a framework proposed by B. DULMET and R. BOURQUIN [3, 4] which uses Piola-Kirchhoff tensors to definie new coefficients called effective constants. The formulation of the thermal sensitivity in the classical method is more complicated and less precise because all the parameters of the resonators vary with temperature including the cristalline orientation of the plates. In order to obtain frequency-temperature dependance, we follow three stages :

a To express the effective constants, we neglect the thermal dependance of the piezoelectric R and the dielectric N coefficients. On the contrary, the effective elastic coefficients G are determined using elastic constants and temperature T. We use Taylor developments to express the matrice G in the vicinity of the reference temperature T_0.

$$G_{cyme}(T) = G_{cyme}(T_0)\left[1 + \sum_{1}^{N} T^nG_{cyme}(T - T_0)^n\right]$$

$$\text{with}\quad T^nG_{cyme} = \frac{1}{G_{cyme}(T_0)n!}\frac{d^nG_{cyme}}{dT^n}$$

T^nG_{cyme} is the n order effective elastic coefficient which is known up to the third order.

b In order to study any orientation of plate, we have to express R, N and G matrix in the coordinates system associated with the plate $(0, x_1', x_2', x_3')$

$$[G'] = [M][G]^T[M]$$
$$[R'] = [\Omega][R]^T[M]$$
$$[N'] = [\Omega][N]^T[\Omega]$$

where [M] is the Bond matrix and [Ω] the rotation matrix.

c The frequency resonance is then obtained by resolving the following Christoffel linear system given by the propagation and the poisson equations. We suppose that the propagation vector is oriented along the thickness of the resonator that's to say along the x_2' direction.

$$\begin{bmatrix} G'_{99} + \frac{R'^2_{29}}{N'_{22}} & G'_{29} + \frac{R'_{29}R'_{22}}{N'_{22}} & G'_{49} + \frac{R'_{27}R'_{29}}{N'_{22}} \\ G'_{29} + \frac{R'_{22}R'_{29}}{N'_{22}} & G'_{22} + \frac{R'^2_{22}}{N'_{22}} & G'_{24} + \frac{R'_{22}R'_{27}}{N'_{22}} \\ G'_{49} + \frac{R'_{27}R'_{29}}{N'_{22}} & G'_{24} + \frac{R'_{22}R'_{27}}{N'_{22}} & G'_{44} + \frac{R'^2_{27}}{N'_{22}} \end{bmatrix} \begin{bmatrix} u_1^\mu \\ u_2^\mu \\ u_3^\mu \end{bmatrix} = \rho_0 V_\mu^2 \begin{bmatrix} u_1^\mu \\ u_2^\mu \\ u_3^\mu \end{bmatrix}$$

where $\mathbf{u}^\mu$ is the eigenvector in the mode μ and V_μ the associated propagation velocity.
In the general case of anisotropic medium, the upper value of V_μ is associated to the quasi-longitudinal

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

mode, the other correspond to quasi-transverse waves. The resonance frequency f_R is then deduced from the dispersion equation and the boundary conditions :

$$2\pi f_R = \frac{n\pi}{2h_0} V_\mu \left(1 - 4\frac{k_\mu^2}{(n\pi)^2}\right)$$

where $2h_0$ is the thickness of resonator at T_0.
n is this overtone number
k_μ is the electromechanical coupling factor for the mode μ

k_μ is expressed in terms of effective coefficients, eigenvalues and eigenvectors :

$$k_\mu = \frac{R_{2\alpha 2} u_{n\alpha}^\mu}{\left[(\rho_0 V^\mu)^2 N_{22}\right]^{1/2}}$$ where u_n^μ is the normalized eigenvector of mode μ

We remark that k_μ like f_R depends on temperature T by eigenvalues and eigenvectors.
In the vicinity of the reference temperature T_0, we can express the resonant frequency with T by a polynomial function up to the third order.

$$\frac{f_R(T) - f_R(T_0)}{f_R(T_0)} = a_1^\mu (T - T_0) + a_2^\mu (T - T_0)^2 + a_3^\mu (T - T_0)^3$$

a_1^μ, a_2^μ, a_3^μ are called the first, second and third order temperature coefficients for mode μ.
The determination of the orientation of plates is governed essentially by a_1, which must be chosen as high as possible to increase thermal sensitivity. It is reasonnable to neglect a_3 for temperature sensor application and a_2 must be weak if we desire to obtain a linear variation of f_R with temperature.
In this paper, we decide to limit our invistigations on singly rotated cuts. So, figures 1a and 1b give the evolution of the first and second order coefficients with θ_0 in the range $[-90°; 90°]$ for $\varphi_0 = 0$ for the transverse mode (B or C) which is piezoelectrically excitable for singly rotated cuts. Figure 1a gives evidence of high coefficient a_1 for $-20° \le \theta_0 \le 20°$, $\theta_0 \ge 50°$ and $\theta_0 \le -80°$ for which $|a_1| > 60$ppm / C.

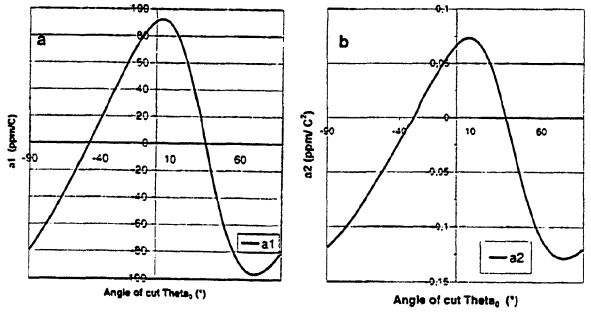


Fig.1 : Evolution of the temperature coefficients with the angle of cut a) First order coefficient
b) Second order coefficient

According to figure 1b, these orientations correspond to high second order coefficients. Indeed, for $|\theta_0| \ge 50°$ a_2 increases with $|\theta_0|$ from 50.10^{-3} to 130.10^{-3}ppm/C². For $-20° \le \theta_0 \le 20°$, $|a_2|$ belongs to the range [30,70] 10^{-3}ppm/C².We can conclude that for the selected cuts, frequency versus temperature characteristics will not be perfectly linear.

2.2 – Electromechanical coupling :
In the last section we establish the expression of k_μ. For singly rotated cuts, we recall that k_A is equal to zero. Only the pure transverse mode along x_1 is piezoelectrically excitable. The associated electromechanical factor is k_C for $\theta_0 \ge -23°$ and k_B for $\theta_0 < -23°$. It takes the following expression.

$$k_{C \text{ or } B} = \frac{R_{26}^2}{\left(G_{66} + \frac{R_{26}^2}{N_{22}}\right) N_{22}}$$

Figure 2 gives the evolution of $k_{C \text{ or } B}$ with θ_0. For $\theta_0 > 0$, we note a decrease of k_C with increasing θ_0. For $\theta_0 > 60°$, k_C becomes weak $(k_c < 2\%)$. For $\theta_0 < 0$, the coupling coefficient varies from $k_c = 13\%$ to $k_c = 0\%$. It is close to 0 for $\theta_0 \le -80°$.

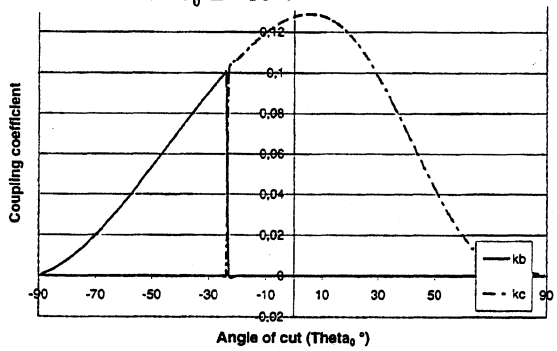


Fig. 2 : Evolution of coupling coefficient with θ_0

According to the results of thermal sensitivity and electromechanical coupling, we can conclude that the orientation to retain for thermal applications are in the ranges $-20° \le \theta_0 \le 20°$ and $50° \le \theta_0 \le 60°$.

2.3 – The quality factor :
The quality factor is one of parameter that characterizes the resonator performance. Q^{-1} is proportionnal to the power dissipated in the resonator due to viscosity, surface defects, loss caused by suspension...
For singly rotated cut, it is expressed as

$$Q = \frac{2h_0 \sqrt{\rho_0 \left(G_{66} + \frac{R_{26}^2}{N_{26}}\right)} \pi n}{\eta_{66} \left(\pi^2 n^2 - 4k_c^2\right)}$$

were η_{66} is the viscosity coefficient expressed in the coordinate axes of the plate.

We will compare this theoretical prediction with experimental results at the end of the paper.

3. CHEMICAL ETCHING BEHAVIOUR

In order to select cuts for our application, we have to take into account properties connected with the anisotropic etching of quartz crystal in a concentrated solution of ammonium bifluoride. For the micromachining of bridges in non-conventionnal cuts, we take care to technical considerations as revealed by previous works which concern surface textures [5] as well as lateral underetching, concave and convex undercuttings [6, 7, 8] :

(i) we must avoid the formation of slightly inclined limiting planes or curved regions bounding the vibration area with respect to the reference orientation.

(ii) the formation of sharp underetch with acute angles under the mask near bridges is not favorable because it can induce mechanical weakness of the suspension.

(iii) the geometrical features of the surface are of prime importance because the performance of high frequency resonators is markedly affected by the roughness of quartz surface [9].

(iv) in order to keep reasonable duration of etching (20 to 30 hours), the etch rate of the reference surface must not be too small.

Analysing results on singly rotated cuts, we can conclude that plates whose angle θ_0 is close to $\theta_0 = 60°$ are more suitable. This orientation coincides with a minimum of the dissolution slowness which implies that the surface texture after etching will be quasi "polished". The plates whose orientation is close to the Y plate are not convenient for micromachining because of their large dissolution slowness. According to these remarks, we retain the two orientations $\theta_0 = 55°$ and $\theta_0 = 60°$.

4. FABRICATION PROCESS AND EXPERIMENTAL RESULTS

4.1 – Fabrication process :

The resonant structure we have to fabricate by micromachining is composed of a circular thin plate (diameter ~ 13.2 mm, thickness ~ 200 µm) with four bridges and a squared mesa (thickness ~ 5 to 10 µm) for the energy trapping. Two successive photolithographic processes are necessary to micromachine mesa and bridges in a concentrated solution of $NH_4F.HF$ maintened at 65C. The duration of etching for bridges and mesa are approximatively 30 hours and 20 mn respectively. The "masks" are Chronium-Gold films deposited by thermal evaporation. They are also used as electrodes for electrical excitation.

4.2 – Experimental results and discussion :

The experimental tests concern the quality factor, the motional resistance and the variations of resonant frequency (at fundamental C mode) with temperature.

a – Q factor and motional resistance :

To determine experimentally the Q factor, the resonator connected to a π network is placed in vacuum.

We determine the motional resistance R_m at the resonant frequency f_R when the phase displacement is zero. The Q factor of quartz is then obtained using the following relation :

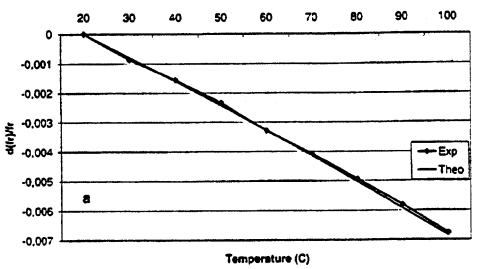
$$Q_Q = Q_{elec}\left(1+\frac{5}{R_m}\right) \text{ where } Q_{elec} = \frac{f_R}{\Delta f}$$

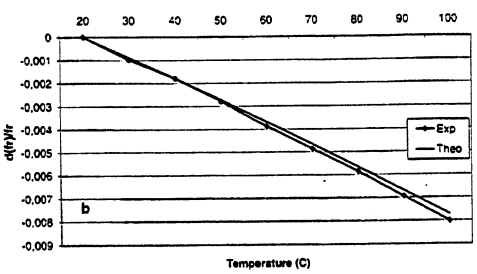
and Δf is the bandwith at – 3dB

Table 1 gives the theoretical and experimental values for the Q factor. We firstly observe that the values of R_m are too high. Such a behavior can result from the viscosity of the quartz plate or from the attenuation of waves at a rough surface. The second assumption is probably good because a prolonged etching causes a degradation of the unmasked surface (especially for the $\theta_0 = 55°$ plate) with the formation of numerus etched pits. The high values of R_m induce a decrease in the Q_Q factor. In order to be sure that the fall of Q is due to the roughness, we measure the same parameters of resonator without bridge machined in a similar cut $\theta_0 = 55°$. Since the thickness for the mesa is close to 20 µm the duration of etching is considerably reduced. In this case, the surface texture is not damaged in the vibrating region. Results in table 1 are not really satisfying because Q_Q factor is also low due to firstly a bad parallelism of the two faces of the plate, to secondly a thin mesa (thickness of mesa 2 µm instead of 20 µm) and to thirsly the stiffness of structure. A good dimensionnement of structure will give better performances.

b – Frequency versus temperature characteristics :

Measurement of the variations in resonance frequency induced by an increase in temperature are performed in ambiant using a frequency synthetiser and a double π network. Figure 3a and 3b give theoretical and experimental evolutions of f_R with T for the two cuts.





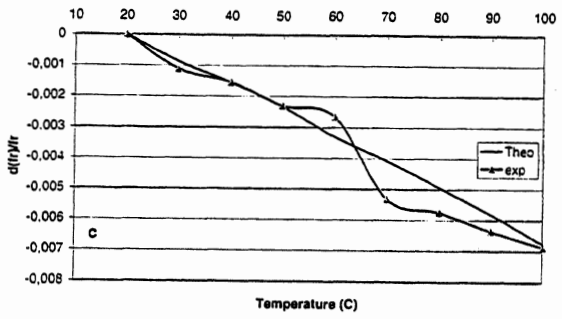


Fig. 3 : Evolution of the resonant frequency with temperature
a) $\theta_0 = 55°$ b) $\theta_0 = 60°$ c) $\theta_0 = 55°$ without bridge

The experimental curves are quasi-linear and coincide with theoretical ones. Figure 3c is related to the bridgeless resonator. The adequation between theory and experiments is not so good because some anharmonic modes disturb the measures essentially between 50 to 80C the temperature range where the frequencies cross each other. If we exclude results in this temperature range, we obtain the expected first order coefficients [Table 2]. Departures between experimental and theoretical second order coefficients are more marked but remain in the same order.

5. CONCLUSION

The feasability of resonators for thermal applications by micromachining is proved. This fabrication process has practically no influence on the frequency versus temperature results. Nethertheless, the fall of Q factor gives evidence of a necessity to overcome some technological difficulties due to chemical etching and specially the degradation of unmasked surfaces after deep etching. But owing to undamagement of annealed Cr-Au films during deep etching, we can try to protect the two faces of vibrating area during micromachining even if the duration of etching is doubled. To optimize the thickness of mesa, etching in less concentrated $NH_4F.HF$ is recommended.

angle of cut θ_0	R_m (Ω)	Q_{elec}	Q_Q	Q_{Th}
55° with bridge $2h_0 = 200$ µm	410	55 000	56 000	$1.8\ 10^6$
60° with bridge $2h_0 = 195$ µm	230	71 600	73 150	$1.81\ 10^6$
55° without bridge $2h_0 = 465$ µm	1 300	80 000	80 000	$4.2\ 10^6$

Table 1 : R_m and Q factors for $\theta_0 = 55°$ and $\theta_0 = 60°$ cuts.

angle of cut θ_0	a_1 exp. ppm/C	a_2 exp. ppm/C^2	a_1 Th. ppm/C	a_2 Th ppm/C^2
55° with bridge	- 78	$-3.\ 10^{-2}$	- 79	$-9.\ 10^{-2}$
60° with bridge	- 86	- 0,13	- 89	- 0,11
55° without bridge	- 70	$-0.2\ 10^{-5}$	- 79	$-9\ 10^{-2}$

Table 2 : Experimental and theoretical thermal coefficients for $\theta_0 = 55°$ and $\theta_0 = 60°$ plates.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge E. Dufromentel and P.Walburger for providing the micromachined structures in section 4.

REFERENCES

[1] D.L. HAMMOND, C.A. ADAMS, P. SCHMIDT, "A linear quartz crystal temperature sensing element", Inst. Society of America, ISA conference, 1984.

[2] T. UEDA, F. KOHSAKA, T. IINO and D. YAMAZAKI, "Temperature sensor using quartz tuning fork resonator", in Proc. 40th AFCS, 1986, PP 224-228.

[3] B. DULMET, "Contribution à l'étude des vibrations à énergie piégée des résonateurs piézoélectriques à ondes de volume. Thèse d'Etat, Université de Franche-Comté, 1991, n° 243.

[4] B. DULMET and R. BOURQUIN, "Influence of the definition of material constants in the study of crystal plates thermal behavior", Proc. 3rd EFTF, 1989, pp 305-312.

[5] C.R. TELLIER, "Micro-usinage chimique du quartz : Modélisation et contrôle par microscopie électronique à balayage", rapport final de contrat DRET N° 851099, 1988, 160 p.

[6] T. LEBLOIS, C.R. TELLIER and T. MESSAOUDI, "Chemical etching of Y-rotated quartz plates : experiments and theoretical approach", Sensors and Actuators A61, 1997, pp 405-414.

[7] T. LEBLOIS and C.R. TELLIER, "Micromachining of quartz resonant structures", proc. 9th CIMTEC Forum on New materials, 1998, to be published.

[8] C. TELLIER, T. MESSAOUDI, T. LEBLOIS and S. DURAND, "Convex and concave corner undercuttings in the micromachining of quartz and silicon mechanical structures ", Proc. 11th EFTF, 1997, pp 386-390.

[9] T. LEBLOIS, C. TELLIER and R. BOURQUIN, "The quality factor of deeply etched quartz resonators : theory and experiments", Revue Phys. Appl. 24, 1989, pp 877-892.

Quartz Microresonator Temperature Sensors Using Lamé-Mode

H.KAWASHIMA*1, H.KANIE*2 and S.YAMAGATA*3

*1 Seiko Instruments Inc.(Presently,Piedek Technical Laboratory)
1-44-1 Kamitakada Nakano-ku Tokyo , Japan 164-0002
*2 Science University of Tokyo , Department of Applied Electronics
Noda, Chiba Japan 278
*3 Hokkaido University of Education , Department of Electrical Engineering
Asahikawa Campus , Hokkaido Japan 070

Abstract

Lamé-mode is very available for realization of a miniaturized quartz crystal resonator because its resonant frequency is principally dependent upon the contour dimensions.Since heat capacity for the quartz crystal resonator gets small, and frequency response versus temprature is very farst, it is useful for use in temperature sensors. In addition,because a Lamé-mode quartz crystal resonator has zero temperature coefficients, designated LQ_1 cut and LQ_2 cut, and particularly, the resonator for LQ_1 cut has a comparatively large value of the second order temperature coefficient β, a Lamé mode quartz crystal resonator can be obtained with the large first order temperature coefficient α when β=0. In this papaer, when cut angles ϕ=45° and θ=45°, α has a value of 44.6×10^{-6}/°C in the calculation and 39.9×10^{-6}/°C in the experiments with β=0 approximately, while when ϕ=51.5° and θ=45°,α=68.1×10^{-6}/°C in the calculation and 62.0×10^{-6}/°C in the experiments with a value of β large than that of ϕ=45° and θ=45°. For both cut angles, frequency change versus temperature (where in after, call temperature sensitivity α) is found to be sufficiently lage, and also the measured data is found to be slightly smaller than the calculated data.

1. Introduction

Quarty crystal is widely used as a material for piezoelectric devices such as resonators, filters and so forth because it is very stable physically and chemically, and also quartz crystal devices can be achieved with a zero temperature coefficient, a samll series resistance R_1 and a high quality factor Q. these quartz crystal devices are actually supported and mounted on two lead wires or a pedestal. However, it is not easy to support and mount them without incresing R_1 and decreasing Q because energy losses of the vibrational portion occur by supporting and mounting the quartz crystal devices. Much effort is, therefore, needed to solve its subject.

When a ratio of width $2z_0$ and length $2x_0$ for a rectangular plate of isotropy in plane becomes integer, a Lamé-mode resonater[1),2)] with nodal points at the four corners can be realized. For an anisotropic material of quartz crystal and lithium niobate, it was reported[3),4)] that Lamé-mode quartz crystal and lithium niobate resonaters are achieved at an aspect ratio R_{xz}(length $2x_0$ / width $2z_0$)=1 for a GT cut resonators coupled between width extensional mode and length extensional mode and at an aspect ratio R_{xz}(=integer) for a lithium niobate resonator, respectively.

In the previous report[5),6)], one of the authors and co-woker introduced that a Lamé-mode resonater is successfully obtained for quartz crystal over a wide cut angle range, and there are cut angles with zero temperature coefficients, designated LQ_1 cut and LQ_2 cut. In particular, a quartz crystal resonator for LQ_1 cut has a comparatively large value of β. Therefore, a Lamé-mode quartz crystal resonator with high temperature sensitivity α can be predicted.

In this paper, temperature sensitivity α, electrical charcterstics and vibrations modes which are analyzed by the finite element simulation are clarified for Lamé-mode quartz crystal resonators. In addition, a res-

0-7803-5400-1/99/ $ 10.00

onator consisting of a vibrational portion and two supporting portions, fabricated by an etching process is employed in the experiments.

First, the frequency equation is derived using an energy method. The calculated values of frequency temperature coefficient α,β where β=0 especially and frequency temperature behavior for Lamé–mode quartz crystal resonators, derived from the frequency equation are then compared with the measured ones.

Second, a theoretical capacitance ratio r which is defined by a ratio of shunt capacitance C_0 and motional capacitance C_1 is shown with the measured data. Lamé–mode quartz crystal resonators are found to be successfully obtained with high temperature sencivity and a small capacitance ratio.

2. Theoretical Approach

2.1 Frequancy equation

Fig.1 shows a Lamé–mode quartz crystal resonator of Y-plate and its coordinate system. The resonator has a dimension of width $2z_0$, length $2x_0$ and thickness $2y_0$, and is expressed by (XYlt)ϕ / θ in the IEEE notation. In addition, excitation electrodes are deposited obverse and reverse of the resonator and their dimensions are width $2z_e$ and length $2x_e$.

Now, when width $2z_0$ and length $2x_0$ are much larger than thickness $2y_0$ and displacements in the x and z axes direction are taken as u and w, lagrangian L is given by

$$L=\frac{1}{2}\left[\rho\int_{-z_0}^{z_0}\int_{-y_0}^{y_0}\int_{-x_0}^{x_0}\left\{\left(\frac{\partial u}{\partial t}\right)^2+\left(\frac{\partial W}{\partial t}\right)^2\right\}dxdydz\right.$$
$$\left.-\int_{-z_0}^{z_0}\int_{-y_0}^{y_0}\int_{-x_0}^{x_0}(T_1S_1+T_3S_3+T_5S_5)dxdydz\right] \quad (1)$$

because stresses T_2, T_4 and T_6 vanish. In addition, ρ is the density and t is the time. On the other hand, the boundary conditions must be satisfied with

$$T_1=T_5=0 \quad on \quad x=\pm x_0$$
$$T_3=T_5=0 \quad on \quad z=\pm z_0 \quad (2)$$

From the relation of eqs, (1) and (2), and ∂L / ∂A=0(A:amplitude), the frequency equation is given as

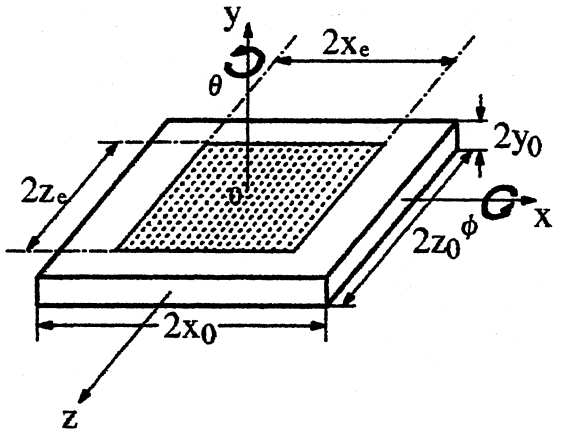


Fig.1 Lamé-mode piezoelectric resonator of Y-plate and its coordinate system

$$f=\frac{m}{2(2x_0)}\sqrt{\frac{c'_{11}-c'_{13}}{\rho}}=\frac{n}{2(2z_0)}\sqrt{\frac{c'_{33}-c'_{13}}{\rho}} \quad (3)$$

because c'_{11}=c'_{33}, where m and n are odd. In addition, c'_{pq} is given by

$$c'_{pq}=\hat{c}_{pq}-\hat{c}_{6q}\hat{c}_{p6}/\hat{c}_{66} \quad (p,q=1,3,5)$$
$$\hat{c}_{pq}=\bar{c}_{pq}-\bar{c}_{4q}\bar{c}_{p4}/\bar{c}_{44} \quad (p,q=1,3,5,6)$$
$$\bar{c}_{pq}=c_{pq}-c_{2q}c_{p2}/c_{22} \quad (p,q=1,3,4,5,6)$$
c_{pq}: elastic stiffness constant (4)

2.2 Motional inductance L_1

Motional inductance L_1 can be calculated by putting kinetic energy equal to electro-magnetic energy in electrical equivalent circuit. Namely, when electric displacement D_2 is taken in the y axis direction, L_1 is given as

$$L_1=\frac{\rho(2y_0)\int_S\left\{\left(\frac{\partial u}{\partial t}\right)^2+\left(\frac{\partial W}{\partial t}\right)^2\right\}dxdz}{\left\{\frac{\partial}{\partial t}\int_{S_e}D_2dxdz\right\}^2}$$

where S: Entire area of vibrational portion
S_e: Electrode area (5)

In addition, when piezoelectric constant e_{2k} is taken as k=1~6, electric displacement D_2 has a form:

$$D_2=e'_{21}S_1+e'_{23}S_3+e'_{25}S_5 \quad (6)$$

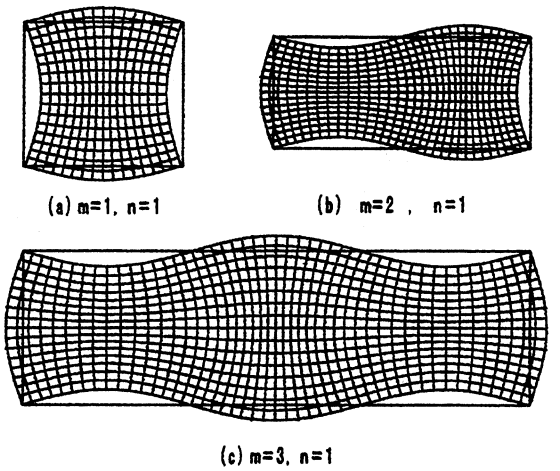


Fig.2 Mode shapes analyzed by the finit element method.

where

$$e'_{2k} = \hat{e}_{2k} - \hat{e}_{26}\hat{e}_{6k}/\hat{c}_{66}(k = 1,3,5)$$
$$\hat{e}_{2k} = \bar{e}_{2k} - \bar{e}_{24}\bar{e}_{4k}/\bar{c}_{44}(k = 1,3,5,6)$$
$$\bar{e}_{2k} = e_{2k} - e_{22}e_{2k}/c_{22}(k = 1,3,4,5,6)$$

As a result, we obtain that

$$L_1 = \frac{\rho\,\xi^2}{32}\,\frac{(2x_0)\,(2y_0)\,(2z_0)}{(e'_{21} - e'_{23})^2}\,\frac{1}{E_p} \qquad (7)$$

where E_p: Constant which is determined by the order of vibration and the area of excitation electrodes

$$\xi = m\pi/2x_0 = n\pi/2z_0$$

2.3 Capacitance ratio r

In designing a resonator, it is of great importance what value the capacitance ratio r has. capacitance ratio r is defined by a ratio of shunt capacitance C_0 and motinal capacitance C_1. First, shunt capacitance C_0 is given as

$$C_0 = \frac{\varepsilon'^S_{22}\,(2x_e)\,(2z_e)}{2y_0} \qquad (8)$$

where $\varepsilon'^S_{22} = \hat{\varepsilon}^S_{22} + \hat{e}^2_{26}/\hat{c}_{66}$

$$\hat{\varepsilon}^S_{22} = \bar{\varepsilon}^S_{22} + \bar{e}^2_{24}/\bar{c}_{44}$$

$$\bar{\varepsilon}^S_{22} = \varepsilon^S_{22} + e^2_{22}/c_{22}$$

ε^S_{22}: Dielectric constant, in the thickness direction

From eqs.(3), (7) and (8), capacitance ratio r is calculated:

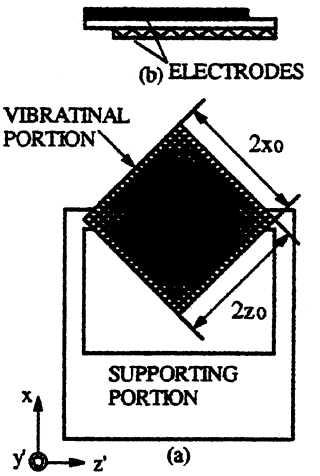


Fig.3 Fundamental Lame-mode quartz crystal resonator m = 1, n = 1(a), formed by etching process and its electrodes construction (b)

$$r = \frac{\varepsilon'^S_{22}\,\xi^4}{32}\,\frac{c'_{11} - c'_{13}}{(e'_{21} - e'_{23})^2}\,\frac{(2x_e)(2z_e)(2x_0)(2z_0)}{Ep} \qquad (9)$$

From the derived equations, frequency temperature coefficients α, β, frequency temperature behavior, motional inductance L_1 and capacitance ratio r are calculated for Lamé–mode quartz crystal resonators.

3. Mode Analysis by the Finite Element Method

Fig.2 shows mode shapes of Lamé–Mode quartz crystal resonators with a cut angle of ϕ=45° and θ=45°, analyzed by the finite element method. As is apparent from Fig 2, it is easily understood that resonators have zero displacement at four coners of the resonators.

4. Resonator Shape and Electrode Construction

Fig.3 shows a fundamental Lamé–mode quartz crystal resonator (m=1, n=1) of a square plate(a), and its sectional view of electrodes construction(b). The resonator consists of a vibrational portion and two supporting portions, and is integratedly fabricated by a chemical etching process. The vibrational portion and the supporting portions are connected at the two corners because a Lamé–mode resonator, as is shown in Fig 2, has zero displacements at four corners of the resonators.

5. Results and Discussion

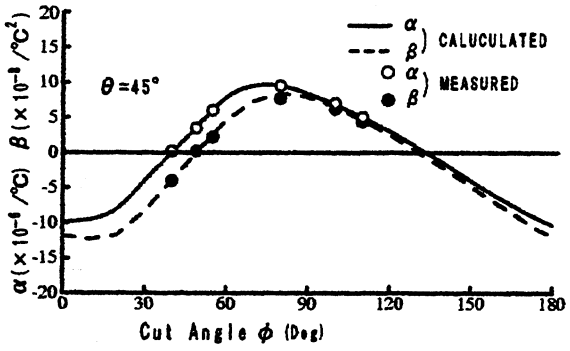


Fig.4 Relationship between the second and first order temperature coefficients β, α and cut angle ϕ of fundamental quartz crystal resonator.

5.1 Frequency temperature coeffcients

Fig.4 shows the relationship of the second and first order temperature coeffcients β, α versus a cut angle ϕ for a Lamé-mode quartz crystal resonator (m=1, n=1) of a square plate with length $2x_0$=1.2mm, aspect ratio $R_{xz}(2x_0/2z_0) = 1$, thickness $2y_0 = 55\mu m$ and cut angle $\theta = 45°$. The solid line is the calculated values and circles are the measured ones. Both results show good agreement. As is apparent from Fig 4, the second order temperature coeffcients β reaches zero approximately at $\phi = 45°$ and 126° in the calculation and in the experiments. In addition, α for $\beta = 0$ has a value of 44.6×10^{-6}/°C in the calculation and 39.9×10^{-6}/°C in the experiments respectively. Lamé-mode quartz crystal resonators for $\phi = 45°$ and 126° approximately are called LQ_3 cut and LQ_4 cut.

On the other hand, the frequency constant ($f \cdot 2z_0$) has a value of 360kHz·cm. Therefore, a miniaturized Lamé-mode quartz crystal resonator with a frequency over approximately 3MHz can be achieved when it is housed in a tubular unit of 2mm diameter and 6mm length.

5.2 Frequency temperature behavior

Fig.5 shows a frequency temperature behavior of a Lamé-mode quartz crystal resonator (m=1, n=1) with length $2x_0$=1.2mm, aspect ratio $R_{xz} = 1$, cut angle $\phi = 45°$ and $\theta = 45°$. In this condition, temperature sensitivity α has 40×10^{-6}/°C approximately.

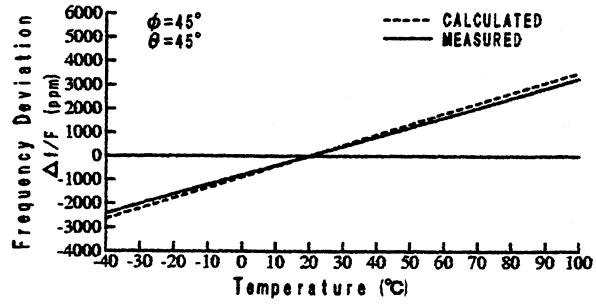


Fig.5 Frequency temperature behavior of a Lamé-mode quartz crystal resonator. (m=1, n=1)

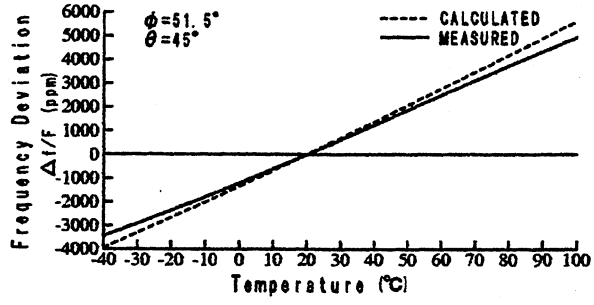


Fig.6 Another Frequency temperature behavior.(m=1, n=1)

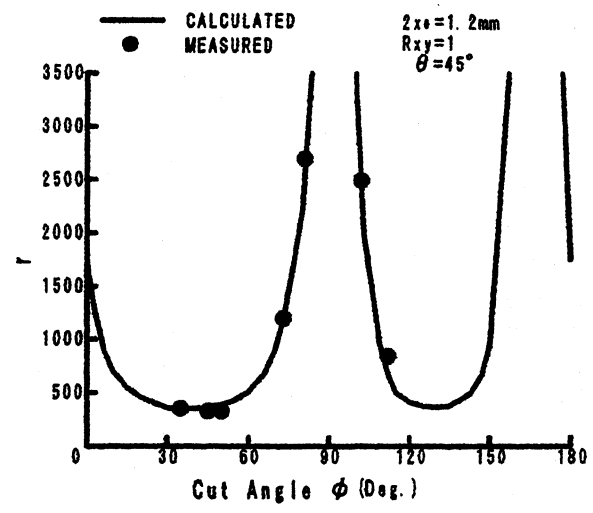


Fig.7 Capacitance ratio r to cut angle ϕ of fundamental quartz crystal resonator. (m=1 , n=1)

Fig.6 shows another example of frequency temperature behavior with cut angles of $\phi = 51.5°$ and $\theta = 45°$. In this case, although β becomes slightly larger,α has a value larger than that of $\phi = 45°$ and $\theta = 45°$.

5.3 Capacitance ratio r

Fig.7 shows the relationship between a cut angle ϕ and capacitance ratio r for a

resonator with the same dimension as that in Fig 4. The excitation electrodes are deposited on the entire area of obverse and reverse of the vibrational portion. As can be easily understood from Fig 7, the capacitance ratio r is very dependent upon the cut angle ϕ, and has a value of $r = 321$ in the calculation and 315 in the experiments at the cut angle $\phi = 45°$ where gives $\beta = 0$.

Table **I** shows typical values of electrical equivalent circuit parameters for Lamé–mode quartz crystal resonators vibrating in fundamental mode(m=1,n=1). Although this Lamé–mode quartz crystal resonator is a very small–sized resonator, it has sufficiently satisfactory characteristics such as series resistance $R_1 = 551\Omega$, motional inductance $L_1 = 0.75H$ and a quality factor $Q = 25\times10^3$ with a frequency of 3MHz.

6. Conclusions

In tHis paper, frequency temperature coefficients, frequency temperature behavior, electrical equivalent circuit parameters, and vibration modes using the finite element method were clarified for Lamé–mode quartz crystal resonators which consist of a vibrational portion and supporting portions, fabricated by a chemical etching process.

First, the frequency equation was derived using an energy method. The calculated values of frequency temperature coefficient β, α and the frequency temperature behavior for Lamé–mode quartz crystal resonators vibrating in fundamental mode, derived using the frequency equation, were then compared with the measuerd ones, so that both result showed good agreement.

Second, capacitance ratio r was shown from motional inductance L_1 and shunt capacitance C_0, and compared with the measured data. Lamé–mode quartz crystal resonators were found to be successfully abtained with a small capacitance ratio r at the cut angle $\phi = 45°$.

Finally, the typical values of electrical eqiuvalent circuit parameters for small–sized Lamé–mode quartz crystal resonators with a frequency of 3MHz were shown, and which were successfully obtained with a small series resistance R_1 and a high quality facter Q.

Table I Typical electrical equivalent circuit parameters for Lamé–mode quartz crystal resonators vibrating in fundamental mode(m=1,n=1).

ϕ (deg)	f (MHz)	R_1 (Ω)	L_1 (H)	Q ($\times10^3$)	r
45	3	551	0.75	25	315

Acknowlegements

We would like to experss our thanks to Mr.T.Kubota and Y.Yokoyama for their help on the typing and the drawings in this paper.

References

[1] R.Bechmann : "Contour modes of square plates excited piezoelectrically and determination of elastic and piezoelectric coefficients", Proc. Phys. Soc.B., Vol. 64, pp.323–336,(1951)

[2] R.Holland : "Contour Extensional Resonant Properties of Rectangular Piezoelectric Plates", IEEE Trans.on Sonics and Ultrasonics., Vol.SU–15 No.2,(1968)

[3] K.Nakamura and T.Tsujimoto : "Lamé-mode piezoelectric resonators using $LiNbO_3$ crystals", Trans. IEICE(C–1), J78–C–I, No.11, pp.494–499, 1995.(in Japanese)

[4] P.C.Lee,M.Nakazawa and J.P.Hou : "Extensional Vibrations of Rectangular crrystal Plates", Proc. 35th Ann. Freq. Control Symposium, 1981.

[5] H.Kawashima and K.Sunaga : "An Analysis of Lamé–Mode Piezoelectric Resonators with Energy Method", Calculation Engineering Vol.1 1996.(in Japanese).

[6] H.Kawashima, M.Matsuyama and K.Sunaga : "Lamé–Mode Quartz Crystal Resonators", Proc. IEEE International Frequency Control Symposium, 1996.

THE PRINCIPLE OF PRESSURE-TEMPERATURE TRANSDUCER OF QUARTZ CRYSTAL RESONATOR

RUIGAO JIA
Department of Applied Physics, University of Petroleum
257062, Dongying Shandong, P. R. China
Telephone: 86-546-8393153
FAX: 86-546-8224374
E-mail: appphy@hdpu.edu.cn

ABSTRACT

This paper expounds that the frequency and the force sensitivity of a quartz crystal resonator are in direct proportion to the order of overtone frequency if the resonator can be made to vibrate at different overtone frequencies in different circuits. This conclusion is tested and verified by experiments. This paper reports the success in experiment that makes the force transducer of circular plate crystal resonator, which has been designed with 3rd order overtone, 5.0 MHz, thickness shear, vibrate at fundamental frequency, 3rd order overtone frequency, 5th order overtone frequency and 7th order overtone frequency separately in different circuits. An investigation has been made in the changing characteristics of both the frequency and sensitivity with the change of the order of overtone. The experimental results show that both the frequency and the force sensitivity are in direct proportion to the order of overtone. This paper analyzes the feasibility and the limitativity of rising force sensitivity method, which is to make the designed resonator with smaller order of overtone frequency vibrate at higher order of overtone frequency. The principle of pressure-temperature transducer is found based on the fact that a crystal resonator can be made to vibrate at differential order of overtone frequency in differential circuits, and on the fact that the sensitivity is a linear function of temperature when the resonator vibrates at an overtone frequency. The distinguished feature of this principle is that transducer be made up only one resonator and is able to measure both the pressure and temperature in fluid. This transducer uses the peculiarity feature of temperature coefficient of sensitivity for avoiding the measurement error caused by temperature coefficient of sensitivity.

1. INTRODUCTION

Radically directed diametrical force applied to the periphery of quartz crystal discs vibrating in thickness shear were found by Bottom [1] to produce frequency shifts in 1947. In the ensuing twenty years the effect has been received considerable attention while other cuts and modes of motion have been investigated. These studies were initially prompted by the need to find mounting points yielding reduced vibration sensitivity, it was also realized that the effect could lead to cost-effective sensors with high accuracy and precision.

Ratajski [1] made more extensive investigations on rotated Y-cut quartz crystal plates. He introduced the coefficient K_f that has subsequently become the norm for quantifying the effect. Lee et al.[3,4] provided a theoretical explanation, showing that the effect arising from a combination of lattice deformation and nonlinear elasticity.

The stress-frequency effect was utilized for the designing QXR pressure transducers. The applied force may be classified as compressive force and extensive force. There were three modes to applied compressive force. The first is periphery force mode. The typical QXR pressure transducer is invented by H. Edward Karrer and Jerry Leach [5], the section of the transducer is shown in figure 1. The second mode to applied force is radically compression force. The typical QXR pressure transducer is designed by E. Karrer and R. Ward [6]. The transducer is shown in figure 3. The third mode to applied force is triplets edge force mode. The frequency changes of circular plano-convex AT and IT-cut plates due to radially applied three forces were investigated by Oura et al. [7] in 1983. The typical pressure transducer applied extensive force was invented by Michael Valdios et al. [8] in 1985. The texture of the pressure transducer is shown in figure 4.

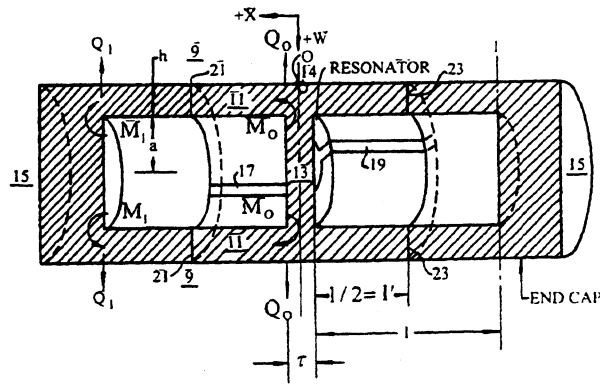


Figure 1

These quartz crystal pressure transducers have very good performance characteristics, but they have a problem in common. The pressure-frequency sensitivity is dependence of temperature. The temperature coefficient of sensitivity causes measurement error. The pressure transducer invented by Henry E. Karrer et al. [5] in 1971 is assembled in a constant temperature case (as shown in figure 2) in order to eliminate the measurement error caused by temperature coefficient of sensitivity, which makes the texture of hole sensor complicated. The

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

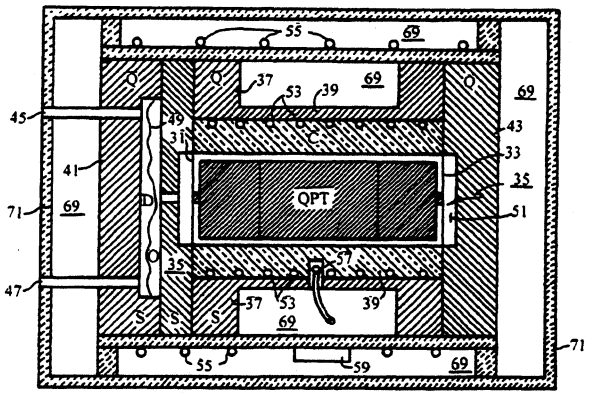


Figure 2

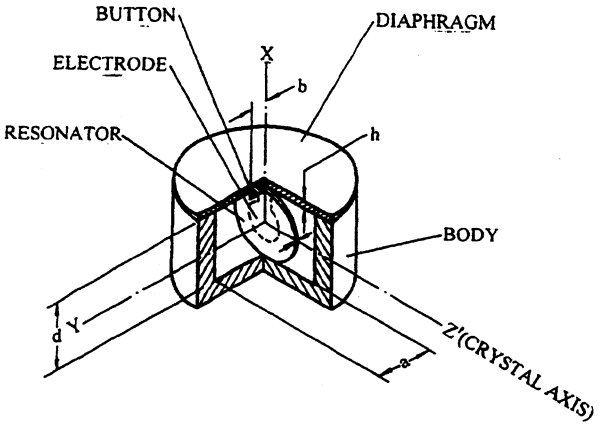


Figure 3

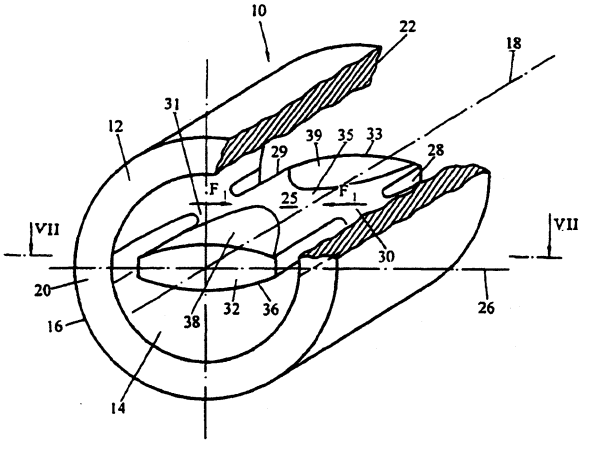


Figure 4

pressure transducer invented by EerNisse [9] in 1985 overcomes the difficulty. The texture of the transducer is shown in figure 5. This transducer has very good performance characteristics but it is harsh working.

C. R. Dauwalter [10] measured the frequency change of circular unconoured 15 MHz fundamental frequency AT-cut resonators due to the application of diametrically opposed forces over the temperature range from 30℃ to 85℃ in 1972. The experimental results show that the relationship of the compressive stress sensitivity of frequency with temperature appears to be linear. In 1979

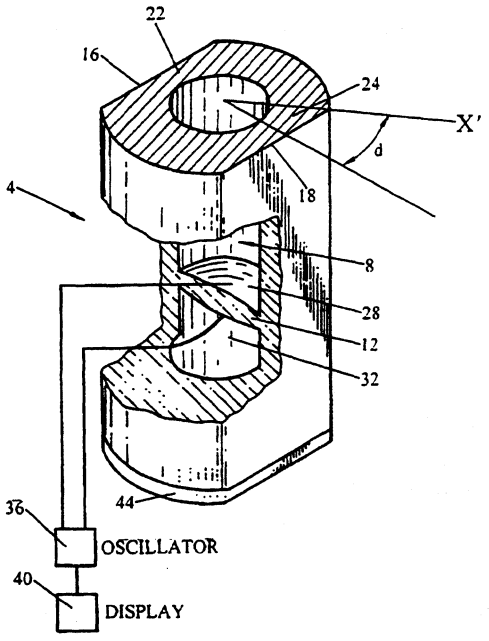


Figure 5

E. P. EerNisse [11] measured the result that the relationship of the force-frequency effect for the rotated X-cut vs. temperature is linear. Ruigao Jia et al. [12] deduced the measurement error formula in 1997 that gives the quantitative relationship of the measurement error with the temperature coefficient of force sensitivity and found a simple method for eliminating the measurement error caused by temperature coefficient of sensitivity for diametrically opposed force mode. This method is selecting the special azimuth angle of applied force, at this angle the temperature coefficient of force sensitivity being zero.

2. CHANGING CHARACTERISTICS OF FREQUENCY AND SENSITIVITY WITH ORDER CHANGING OF OVERTONE FREQUENCY

The resonant frequency of thickness shear mode is determined by

$$f = \frac{n}{2e}\sqrt{\frac{C_{66}'}{\rho}} \qquad (1)$$

Where n is the order of overtone frequency, e is resonator thickness, ρ is quartz density and C_{66}' is elastic constant. The formula (1) may be rewritten as

$$f_n = nf_1 \qquad (2)$$

Where $f_1 = \frac{1}{2e}\sqrt{\frac{C_{66}'}{\rho}}$ is the fundamental frequency.

Formula (2) shows: The overtone frequency f_n is direct proportion to the order n of overtone frequency if the resonator can be made vibrate at several differential overtone frequencies when it is set separately in differential circuits.

The sensitivity for single rotated cut circular resonator is determined by the following formula:

$$S_F^n = \frac{K_f f_n^2 \eta}{nD} \quad (3)$$

Where K_f is Ratajski edge force coefficient, f_n is resonant frequency, D is diameter of crystal plate, η is a constant.

After substituting formula (2) into formula (3), the sensitivity of the resonator vibrating at n order overtone frequency is

$$S_F^n = n\frac{K_f f_1^2 \eta}{D} \quad (4)$$

Let $S_F^1 = \frac{K_f f_1^2 \eta}{D}$, formula (4) can be rewritten as

$$S_F^n = nS_F^1 \quad (5)$$

Where S_F^1 is the sensitivity of resonator vibrating at fundamental frequency.

The formula (5) shows: The force sensitivity of resonator is direct proportion to the order n of overtone frequency if the resonator can be made vibrate at differential overtone frequency in differential circuits. Hence, the method to raise the sensitivity without reducing the load-bearing capacity is found. The method is to make the resonator with smaller order of overtone frequency vibrate at the larger order of overtone frequency.

The reference resistance, reference capacitance and reference inductance must be changed when the resonator is made to vibrate at differential overtone frequency. The change of these parameters must interfere with the stability of the frequency and sensitivity, and other performances. Hence, the raising force sensitivity method is inevitably limited.

3. THE EXPERIMENTAL RESULTS OF FORCE SENSITIVITY CHARACTERISTICS OF CRYSTAL RESONATORS VIBRATING AT DIFFERENTIAL OVERTONE FREQUENCY

The extensive experimental investigations were made in order to verify the formula (2) and the formula (5) and research on the feasibility and the limitation of the method to raise sensitivity. The experiments to test the frequency change and the sensitivity change vs. the order of overtone frequency using 5.0 MHz third overtone AT-cut resonators due to the application of diametrically opposed force (the azimuth angle ψ =0) have been investigated by the author of this paper. Each crystal resonator used in these investigations has two small flats on its edge, nominally perpendicular to the X crystallographic axis. The resonators set in four differential circuits vibrate at fundamental frequency, 3rd order overtone frequency, 5th order overtone frequency and 7th order overtone frequency separately.

Table 1 shows the four vibrating frequencies, f, sensitivities, S_F^0 and the four determined reference frequencies, Ff, determined reference resistance, FR, determined reference capacitance, FC, determined reference inductance, FL with the Crystal Test Meter.

Table 1

n	1	3	5	7
f(Hz)	1705974	5002541	8304447	11723310
$S_F^{(n)}$ (Hz/kg)	-42	46	79	99
Ff(Hz)	1706049	5002556	8304454	11723326
FR(Ω)	2831.8	76.6	157.4	374.1
FC(pF)	4.5	4.4	3.9	3.9
FL(mH)	1983.56	1402.33	2398.48	2133.39

The experimental results show: The vibrating frequencies conform with formula (2) and the sensitivities with the formula (5) excepting when the resonators vibrate at fundamental frequency. The inconformity with formula (5) results from the difference of vibrating mode between fundamental frequency and higher order of overtone frequency.

The experiments results show: Both the stability (short term stability and long term stability) and the linearity of frequency vs. the order of overtone are good when the resonators vibrate at designed order (3rd order) and 5th order of overtone, and that are worse at fundamental frequency and 7th order of overtone.

The experimental results give a method to raise the force sensitivity also, The method is to make the designed resonator with smaller order of overtone frequency vibrate at higher order of overtone frequency. The limitation of the method is that the higher order must be close to the designed order.

4. THE PRINCIPLE OF PRESSURE-TEMPERATURE TRANSDUCER

Assuming a transducer scaled at temperature T_1, operates in some temperature range. Its sensitivity is following linear function of temperature T separately if it vibrates at n and n' order of overtone frequency [12].

$$S_F^{(n)} = \alpha T + S_F^{(0)} \quad (6)$$

$$S_F^{(n')} = \alpha' T + S_F^{(0')} \quad (7)$$

Where $S_F^{(n)}$, α, $S_F^{(0)}$ is the sensitivity, temperature coefficient of sensitivity, the sensitivity at T=0°C when it vibrates at n order of overtone frequency separately, and $S_F^{(n')}$, α', $S_F^{(0)}$ is the sensitivity, temperature coefficient of sensitivity, the sensitivity at T=0°C when it vibrates at n' order of overtone frequency separately.

Assuming the applied force to the transducer is F and the environment temperature is T, the measured frequency change is Δf and $\Delta f'$ separately when transducer vibrates at n and n' order overtone frequency separately.

$$\Delta f = F(\alpha T + S_F^{(0)}) \quad (8)$$

$$\Delta f' = F(\alpha' T + S_F^{(0')}) \quad (9)$$

The applied force F and environment temperature T can be deduced from above two formulae

$$F = \frac{\Delta f}{\alpha T + S_F^{(0)}} \quad (10)$$

$$T = \frac{\Delta f' S_F^{(0)} - \Delta f S_F^{(0)}}{\alpha' \Delta f - \alpha \Delta f'} \quad (11)$$

All the $\alpha, S_F^{(0)}$, α' and $S_F^{(0')}$ could be measured before the transducer is operated.

Both the formula (10) and (11) show that applied force F and environment T can be measured with one transducer. If the force is induced by pressure in liquid or in gases, the pressure P could be determined by following formula:

$$P = kF = k\frac{\Delta f}{\alpha T + S_F^{(0)}} \quad (12)$$

k is the transformation parameter between pressure and force. It depends on the structure of the transducer and it can be measured.

Both the formulae (11) and (12) give the principle of pressure-temperature transducer. Make one transducer vibrates at two orders of overtone frequency in two circuits separately, pressure and temperature in fluid can be measured on the basis of formula (11) and (12). The transducer based on the principle could avoid the measurement error caused by temperature coefficient of sensitivity. This principle uses the peculiarity feature of temperature coefficient of sensitivity for avoiding the error caused by temperature coefficient of sensitivity. All the α, $S_F^{(0)}$, α' and $S_F^{(0')}$ could be measured beforehand, the method to avoid the error is reliable.

The temperature coefficient of sensitivity is zero if the azimuth angle of applied force is the special angle. The measurement error caused by temperature coefficient of sensitivity can be avoided as well as [12]. The special azimuth angle is not the angle that the sensitivity is the maximum. The azimuth angle of pressure-temperature transducer may be the angle that the sensitivity is the maximum.

5. CONCLUSIONS

Both the frequency and sensitivity of quartz crystal resonator are in direct proportion to the order of overtone frequency. Making resonator that has been designed with smaller order of overtone frequency to vibrate at higher order of overtone frequency is a feasible method to raise sensitivity. Based on the experimental results that sensitivity is linear function of temperature, and a resonator can be made to vibrate at differential order of overtone frequency in differential circuits, the pressure-temperature transducer could be made. The pressure-temperature transducer with one resonator could measure pressure and temperature in fluid. The pressure-temperature transducer uses the peculiarity feature of temperature coefficient of sensitivity for avoiding the measurement error caused by temperature coefficient of sensitivity.

REFERENCES

[1] Virgil E. Bottom, Note on the Anomalous Thermal Effect in Quartz Oscillator Plates, Amer. Mineralogist, vol. 32, pp. 590-591, Sept.-Oct. 1947

[2] J. M. ratajski, Force-Frequency Coefficient of Singly Rotated Vibrating Quuartz Crystal, IBM J. Res.Dev.,vol. 12, pp.92-99.

[3] Peter C. Y. Lee, Y. S. Wang, and X. Markenscoff, Elastic Waves and Vibrations in deformed Crystal Plates, in Proc. 27th Ann.Freq. Control Symposium, pp. 1-6, June 1973.

[4] Peter C. Y. Lee, Y. S. Wang, and X. Markenscoff, High Frequency Vibrations of Crystal Plates Under Initial Stresses, J. Acoust. Sco. Amer., pp. 95-105,June 1975.

[5] Henry E. Karrer, Jerry G.Leach, U.S. Patent, 3,561,832, Feb. 9, 1971.

[6] E. Karrer and R. Ward, A Low-range Quartz Resonator Pressure Transducer, ISA Transaction Vol. 16, No. 2, pp. 90-98, 1977.

[7] N. Oura, N. Kuramochi, J. Nakamura, Y. Miyazaki and M. Yamashina, Frequency changes in fifth-Overtone 5-MHz Circular Plano-convex AT and IT-cut Plates Due to Radially Applied Three Force, IEEE Transavtions on Ultrasonics, Vol. SU-30, PP.111-114, Mar. 1983.

[8] Michel Valdois,at el., U.S. Patent 4,547,691, Oct.15 1985.

[9] Errol P. EerNisse, U. S. Patent, 4,550,610, Nov. 5, 1985.

[10] harles R. Dauwalter, The Temperature Dependence of the Force Sensitivity of AT-cut Quartz Crystals, Proceedings of the 27th Ann. Frwq. Control Symposium, pp. 108-112, June 1972.

[11] Errol E. EerNisse, Temperature Dependence of the Force Frequency Effect For the Rotated X-cut, Proceedings of the 33rd Annual Frequency Control Symposium, pp. 300-305, 1979.

[12] Ruigao Jia, Jinsong Wang, Guanping Feng, Temperature Coefficient of sensitivity Dependence of the Measurement Error of Quartz Crystal Resonators, IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, Vol. 44, NO. 6, pp. 1376-1381, November 1997.

ALC CRYSTAL OSCILLATORS BASED PRESSURE AND TEMPERATURE INTEGRATED MEASUREMENT SYSTEM FOR HIGH TEMPERATURE OIL WELL APPLICATIONS.

R. A. Bianchi, J. M. Karam, B. Courtois.

TIMA Laboratory. 46 av. Félix Viallet. 38031 Grenoble Cedex, France.
Tel : +33 4 76 57 48 36. Fax : +33 4 76 47 38 14.
e-mail : Raul-Andres.Bianchi@imag.fr

Abstract

An application specific integrated circuit for the pressure and temperature measurement at the high temperature oil well bottom conditions, is presented in this paper. The system is mainly composed of three high performance Automatic Level Controlled (ALC) oscillator circuits which drive three external crystals (the sensitive elements), and of some mixing and filtering stages. The integrated system was successfully tested up to 220°C, showing a frequency resolution of 0.0003 ppm (0.0007 psi, for the pressure measurement) and a drift of 1.5 Hz/month (~0.5 psi/month), concerning respectively the short and long term measurement stability. A drastic reduction of the whole measurement tool size and cost will be allowed by means of this application specific BiCMOS integrated circuit.

1. INTRODUCTION

This integrated measurement system will provide a permanent measurement of the pressure and temperature at the well bottom conditions during the whole life of the well. Consequently the stability of the system (i.e. the measurement resolution and particularly the long-term stability) is specially concerned.

Looking for new oil resources, oil wells are increasing in depth, and reducing in diameter. As a consequence of this, the measurement tools need to be reduced in size, mainly to reduce the cost of the expensive hermetically-sealed housing, and they must support higher temperatures (175°C). Normally the electronics is implemented with discrete components, which fulfil military specifications (125°C), assembled in a PCB board.

The objectives of drastic size reduction, higher temperature robustness, and better interconnections reliability will be achieved by means of the design of the following BiCMOS Application Specific Integrated Circuit, with an additional increase of the system stability.

2. MEASUREMENT SYSTEM STRUCTURE

The measurement system structure in *Figure 1* consists of three integrated oscillator's circuits which drive three different crystals [3]. The Reference one has a reduced temperature sensitivity and is used to generate a stable reference frequency. The Temperature crystal has a quite larger temperature sensitivity which is used for temperature measurement. External pressure is mechanically transmitted to the Pressure crystal, which is sensitive to pressure and it has also a small temperature sensitivity that must be compensated. The used Reference and Temperature crystals are fundamental mode quartz, but the Pressure crystal is a third harmonic mode one.

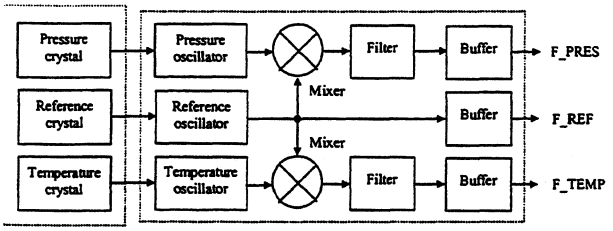


Figure 1. Measurement system structure.

The generated pressure and temperature signals are mixed with the reference one, filtered and amplified to obtain lower frequency digital signals (F_PRES and F_TEMP) for further digital signal processing.

2.1 Oscillators structure

The three oscillators are sinusoidal *Automatic Level Controlled (ALC)* crystal oscillators and they have the same general structure. It consists of a *Current Controlled Inverter Gain stage*, a *Phase Shift Generator Circuit* and a *Level Detector & Bias Controller*, as shown in *Figure 2*. This structure makes the generation of purely sinusoidal signals with a controlled level and a very good frequency stability performance possible.

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

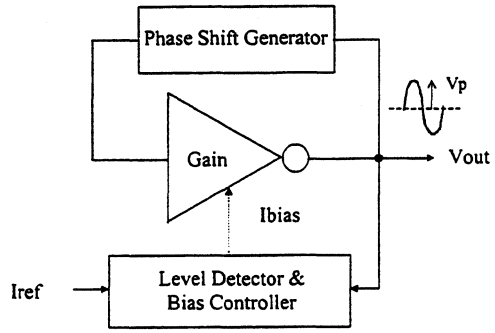


Figure 2. ALC oscillator structure.

2.1.1 Fundamental mode oscillators

In *Figure 3* the structure of the fundamental mode oscillators is described.

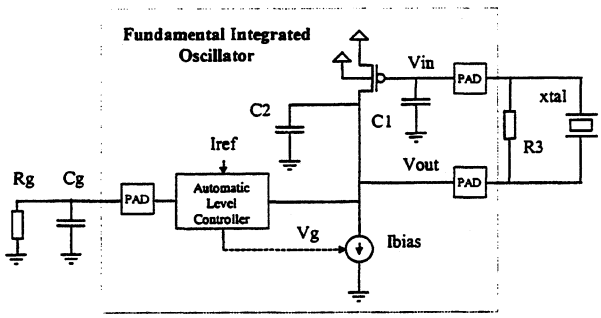


Figure 3. Fundamental mode oscillator circuit.

Oscillations will not occur at the third, fifth and other oscillation modes of the crystal, which are weaker than the fundamental one, because the gain of the circuit was accurately calculated and limited to avoid exciting theses resonance modes.

2.1.2 Third harmonic mode oscillator

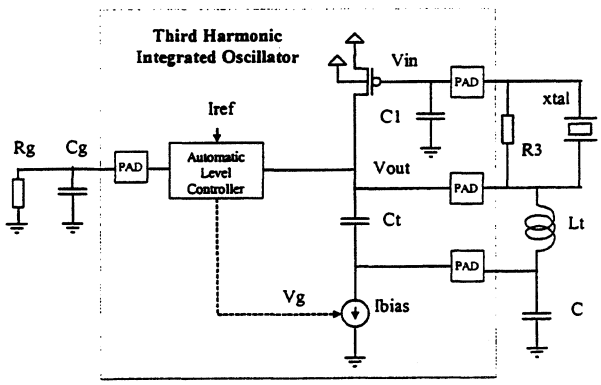


Figure 4. Third harmonic mode oscillator circuit.

For the third harmonic oscillator case, the structure was modified in order to select the desired resonance mode. Thus, another resonant circuit was added (the *Lt-Ct circuit*) which acts as the appropriated capacitor (*C2*) only at frequencies higher than the mid-value between the first and the third harmonic frequencies. Since at the first harmonic frequency the *Lt-Ct* circuit does not allow the generation of the necessary phase shift to produce oscillations, the first oscillation mode is then eliminated. In *Figure 4* the structure of the third harmonic oscillator is described.

2.2 Mixer structure

To meet the low-noise and high-linearity requirements for the mixer stage, a *Fully-Balanced Current Mode Switching* mixer structure (see *Figure 5*) was chosen. It acts as a polarity-switching function of the Temperature and Pressure Oscillator signals in response to the Reference Oscillator signal polarity.

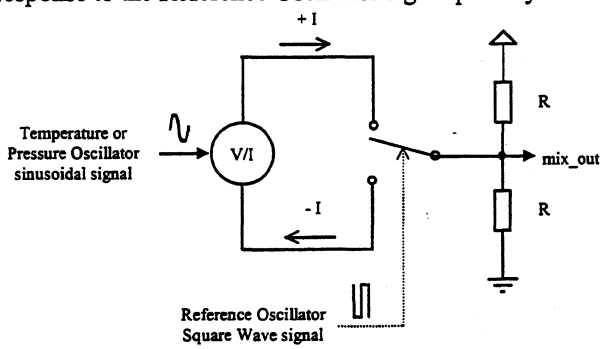


Figure 5. Mixer structure.

Because of the used asymmetric power supply, the polarity switching of signals can be better performed in current mode. So, a high input impedance *Voltage to Current Converter (V/I)*, some *Switches*, and a *Current to Voltage Converter* were implemented for this mixer. To drive those switches a square wave version of the Reference Oscillator signal is then needed.

The spectrum of the output signal of this kind of mixers is quite complex, because the product between a sinusoidal signal (**TP** = Temperature or Pressure oscillator output.) and a unity (+1 ↔ -1) square wave signal (**REF**) is obtained at the output. The latter can be expressed by the following Fourier series,

$$S_{REF} = \frac{4}{\Pi}\cdot\left[\sin(\omega_{REF}\cdot t) - \frac{1}{3}\sin(3\cdot\omega_{REF}\cdot t) + \frac{1}{5}\sin(5\cdot\omega_{REF}\cdot t) - \cdots\right]$$

Thus, the output of the switching mixer will be the product between its S_{TP} input signal,

$$S_{TP} = V_{DC} + V_{AC}\cdot\sin(\omega_{TP}\cdot t)$$

and the above series expansion S_{REF}, giving at the output of the mixer,

$$S_{OUT} = \frac{2}{\Pi}\cdot\{2\cdot V_{DC}\cdot\sin(\omega_{REF}\cdot t)+V_{AC}\cdot\cos[(\omega_{TP}+\omega_{REF})\cdot t]+V_{AC}\cdot\cos[|\omega_{TP}-\omega_{REF}|\cdot t]\}$$
$$-\frac{2}{3\cdot\Pi}\cdot\{2\cdot V_{DC}\cdot\sin(3\cdot\omega_{REF}\cdot t)+V_{AC}\cdot\cos[(\omega_{TP}+3\cdot\omega_{REF})\cdot t]+V_{AC}\cdot\cos[|\omega_{TP}-3\cdot\omega_{REF}|\cdot t]\}$$
$$+\frac{2}{5\cdot\Pi}\cdot\{2\cdot V_{DC}\cdot\sin(5\cdot\omega_{REF}\cdot t)+V_{AC}\cdot\cos[(\omega_{TP}+5\cdot\omega_{REF})\cdot t]+V_{AC}\cdot\cos[|\omega_{TP}-5\cdot\omega_{REF}|\cdot t]\}$$

If both mixed frequencies are relatively close, the mixer output signal can be described with the spectrum diagram of *Figure 6*,

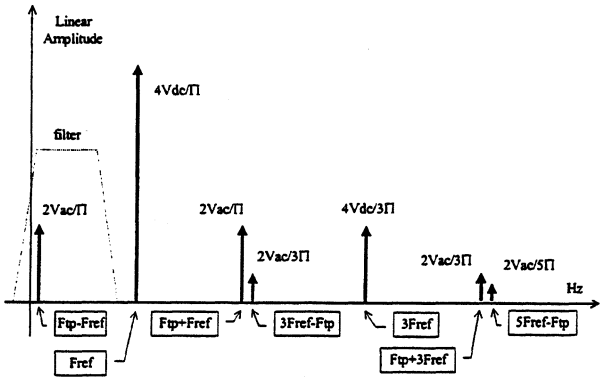


Figure 6. Mixer output frequency spectrum.

where the desired output signal is,

$$S_{Filtered} = \frac{2}{\Pi}\cdot\{V_{AC}\cdot\cos[|\omega_{TP}-\omega_{REF}|\cdot t]\}$$

Therefore, the higher frequency components must be filtered. If the Temperature or Pressure Oscillator signals have some amount of distortion, specially because of some amount of undesired cross-talk with other signals, low inter-modulation frequencies can be produced in the mixer. Which could generate slow variations in the DC level of the mixer output signal and further jitter noise in the comparator output. Special filtering is thus needed also for low frequencies.

2.3 Filters and Buffers

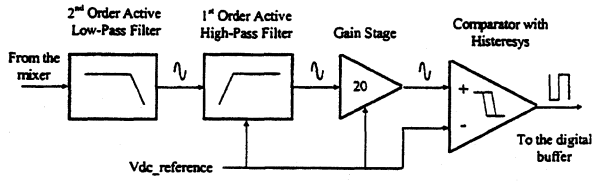


Figure 7. Filter structure

In order to solve the above mentioned problems, produced by the high and low intermodulation frequencies, and by the inaccurate DC level present at the mixer output signal, the three stages Band Pass filter (*10 to 70KHz*) with DC level correction of *Figure 7*, was designed to obtain,

- Gain at 10KHz ≅ 20 dB
- Gain at 70KHz ≅ 20 dB
- Attenuation at ω_{REF} ≅ 100 dB
- Attenuation at 10Hz ≅ 40 dB

3. FABRICATED CIRCUIT

The whole integrated system, was fabricated in a standard 0.8μm BiCMOS technology. The *Figure 8* shows a Photomicrograph of the fabricated and tested final version (3815 x 3815 μm).

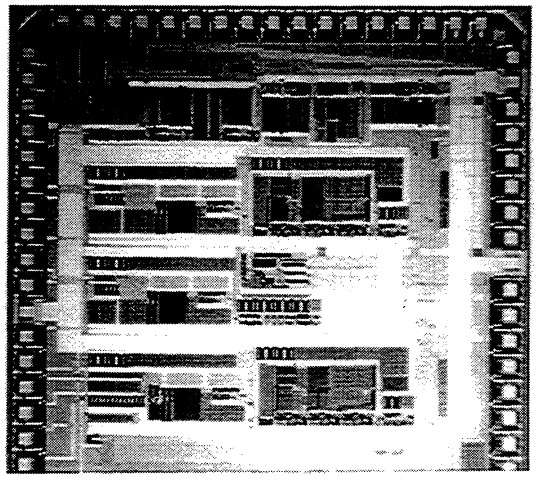


Figure 8. Photomicrograph of the Implemented Circuit.

4. MEASUREMENT RESULTS

Concerning frequency stability, the short-term (about one second) frequency stability of the whole system was measured by means of the frequency resolution observed at 1 count/sec rate (see *Table 1*). The temperature oscillator signal was mainly used for the compensation of the mentioned temperature sensitivity of the pressure crystal.

	Pressure Measurement System
Power Supply Voltage	5 Volts
Maximal Temperature	220°C
Resolution (1/Sec)	0.0003 ppm, 0.0007 psi
Long Term Drift	1.5 Hz/month, 0.5 psi/month

Table 1. Pressure measurement system performance.

Concerning the long-term frequency stability the output signal *F_PRES* frequency was measured every 10 minutes during almost two months, at 185°C and ambient pressure conditions (see *Figure 9*).

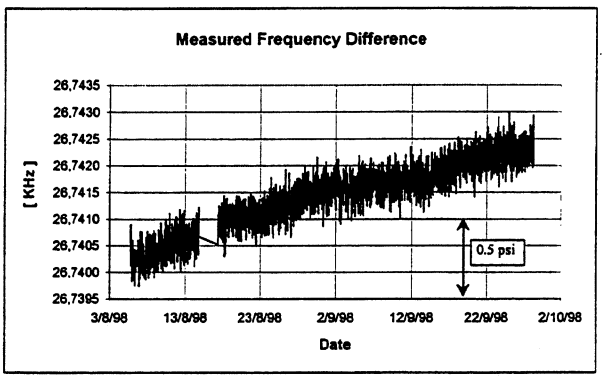


Figure 9. Long term frequency stability

Thus, with stable temperature and pressure conditions, the observed long-term drift is only due to the long-term drift of the integrated devices, their operation points, and the characteristics of the crystals. The *Figure 9* shows the frequency data measured at the *F_PRES* output, where a small long-term frequency drift of about *1.5 Hz/month* (*0.5 psi/month*) was observed.

5. CONCLUSIONS

The design and implementation of this high-performance pressure and temperature measurement integrated system was presented, showing a very good behaviour in terms of measurement resolution, long-term stability, and high temperature robustness, which are compatible with the requirements imposed by the oil well bottom conditions.

The characteristics of the presented application specific BiCMOS integrated circuit will make possible a drastic reduction of the whole measurement tool size, cost, and failure rates; with an additional performance increase.

6. REFERENCES

[1] B. Parzen and A. Ballato, *Design of Crystal and Other Harmonic Oscillators*. New York: Wiley-Interscience, 1983. ISBN: 0-471-088196

[2] R. J. Matthys, *Crystal Oscillators Circuits*. New York: Wiley-Interscience, 1983. ISBN: 0-471-87401-9

[3] M.Willander and H.L. Hartagel, *High Temperature Electronics*. London: Chapman & Hall, 1997. ISBN: 0-412-62510-5

[4] E. A. Vittoz et al., "High-Performance Crystal Oscillator Circuits: Theory and Application", *IEEE Journal of Solid-State Circuits*, Vol 23, No 3, June 1988.

[5] R. G. Meyer and D. Soo, "MOS Crystal Oscillator Design", *IEEE Journal of Solid-State Circuits*, Vol SC-15, No 2, pp 222-228, April 1980.

[6] M. Soyuer, "Design Considerations for High-Frequency Crystal Oscillators", *IEEE Journal of Solid-State Circuits*, Vol 26, No 6, pp 889-893, June 1991.

[7] A. Rusznyak, "Start-Up Time of CMOS Oscillators", *IEEE Transactions on Circuits and Systems*, Vol CAS-34, No 3, pp 259-268, March 1987.

[8] J. T. Santos and R. G. Meyer, "A One-Pin Crystal Oscillator for VLSI Circuits", *IEEE Journal of Solid-State Circuits*, Vol SC-19, No 2, pp 228-235, April 1984.

[9] F. S. Shoucair, "High Temperature Latchup Characteristics in VLSI CMOS Circuits", *IEEE Transactions on Electron Devices*, Vol 35, No 12, pp 2424-2426, December 1988.

[10] F. S. Shoucair et al., "Electrical Characteristics of LSI Silicon MOSFET's at Very High Temperatures, Part III: Modeling and Circuit Behavior", *IEEE Transactions on Components, Hybrids, and Manufacturing Technology*, Vol CHMT-7, No 1, pp 146-153, March 1984.

[11] F. S. Shoucair, "Design Considerations in High Temperature Analog CMOS Integrated Circuits", *IEEE Transactions on Components, Hybrids, and Manufacturing Technology*, Vol CHMT-9, No 3, pp 242-251, September 1986.

Biography

Raúl Andrés Bianchi was born in Córdoba, Argentina, in 1972. He received his Electronic Engineering degree in 1995 from the Catholic University of Córdoba, and his *Diplôme D'Etudes Approfondis en Microélectronique* in 1996 from the Joseph Fourier University in Grenoble, France. Since then he has been working towards his Ph.D. degree, in the *Institut National Polytechnique de Grenoble,* on the subject of integrated microsystems for high temperature applications in standard technologies.

1999 Joint Meeting EFTF - IEEE IFCS

MICROWAVE OSCILLATOR AS A SENSOR OF ELECTROPHYSICS FLOW PARAMETERS

V.V. Boloznev and E.V. Safonova
Kazan State Technical University named by A.N.Tupolev, tel.+7(8432)631741, E-mail: kate@cg.ru

1.ABSTRACT

New results of the authors in development of microwave self-excited oscillating sensors of high-temperature technological processes, mainly processes of burning, are presented. As the majority of microwave-diagnostic devices, sensor reacts to flame's operating body permittivity $\varepsilon_x = 1 - \omega_{cr}^2 / \omega^2$, where ω_{cr} - is a critical frequency, which depends on free electron density, ω - generation frequency, $\omega > \omega_{cr}$. ε_x into ω transformation goes through antenna's resistance.
Main functional characteristics of such sensors have been researched heretofore [1,2,3]. Main disturbance's actions, being created by control process and causing measurement error, have been described in this work. Temperature error has been researched in detail. This research in a physics-mathematical relation is similar to self-excited oscillator's frequency instability analysis, and the means of error reduction coincide with stabilization methods.
Nonuniform sensor heating forms thermal flows and complicated temperature field. Quantitative dependence of frequency instability from temperature and electrophysics, thermophysics and constructive self-excited oscillator parameters has been established by electric and thermal processes comparative analysis. Analysis has been carried out in two modes: quasi-static (the start of the process being controlled) and dynamic (nominal process run under the perturbation actions). In the first one – relational frequency drift is notably high (to 10^{-3}), in the second – two-four orders lower in the perturbation frequencies range $\nu > 10^{-2}$ Hz.
Two original technical solutions have been suggested along with a traditional stabilizing measures: the electric scheme of sensor with a multi-link resonator and the construction, in which all main details are made of porous radio-transparent material.

2. INTRODUCTION

Self-excited oscillating sensors with frequency exit are the most compact devices for high-temperature technological process SHF-control [1]. Sensor with coaxial resonator and ringed slit antenna [2,3] has been researched better and received the most wide usage. Susceptibility to controlling process aggressive influences: heating, vibration, antenna erosion, technological process product deposition, for example, soot deposition on the surface, is a serious drawback of these sensors. These influences are disturbances from the measuring electronics standpoint. The protection from heating, which creates a threat to sensor's efficiency (Gunn-effect diode and self-excited oscillating diode failure) and causes severe measurement error, is the most difficult. Error quantities and sensor threshold sensitivity are proportional to self-excited oscillator's frequency instability. Experimental research of the mentioned above destabilizing factors and their influence run into technical and economical difficulties. This task is particularly difficult in the process of burning control application.

3.THE PURPOSE

Our researches include the analysis of the sensor's measurement error, subjected to destabilizing influence. Such task is closer to self-excited oscillator instability analysis in mathematical standpoint. Purposes of this work are as follows:

- detection of the heating influence mechanisms on the generation frequency, their comparative quantitative estimation;
- Technical suggestions elaboration for stability improvement.

4.ELECTRICAL AND THERMOPHYSICAL MODELS

The construction of the sensor, located on the technological apparatus wall 1 (thereafter referred to as a combustion chamber or chamber) has been shown below.

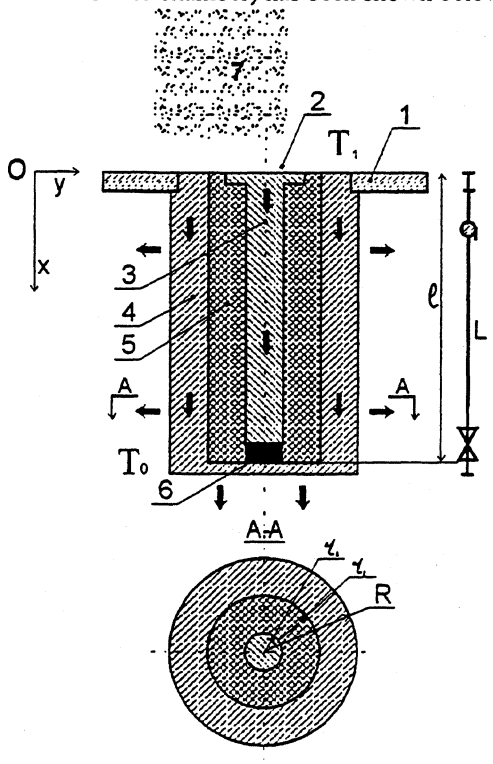


Figure 1

Sensor's circuit (fig.1 right) coincides with [2], antenna is short-circuited and its influence is not considered here. Fire surface 2 coincides with antenna aperture; coaxial resonator includes a rod 3, outer cylinder 4 and a dielectric bush 5. Diode 6 is located near the face. Heating is realized by the operating body 7, which temperature near the wall area is T1, outside temperature is T0. Arrows show thermal flow movement.

From the thermophysics standpoint sensor is equivalent to the rod, heating at the one face. Temperature spreading in the rod is essential only along the coordinate X (fig.1). Diode's temperature is close to the temperature on the lower face.

5. BASIC EQUATIONS

Frequency instability analysis is based on the sensor's electrical and thermophysical models matching. The first model can be studied using the second Kurokava condition, which is equivalent to equation

$$ctg\left[\omega\ell\sqrt{\varepsilon}/C\right]+\rho\omega C_D=0, \quad (1)$$

Where ρ - is coaxial wave resistance, ω - angular frequency, ε - dielectric bush permeability, C_D - open-frame diode capacity, C - light speed.

In practically interesting cases these dependencies could be presented as $q(T)=q+\delta q(T)$, where

$q=\omega,\ell,\varepsilon,\rho,C_D$, $\delta q=\frac{\partial q}{\partial T}\Delta_q T << q$; (this last condition is not being kept for the diode capacity C_D), for the other quantities- ℓ,ε,ρ, - derivatives $\partial q/\partial T$ are temperature coefficients.

Heat transfer equation for the accepted model in a quasi-static mode can be found in the references, for example in [4]. We set a condition, that temperature on the face is:

$$T(L)=T_0+(T_1-T_0)/ch(mL)= \\ =T_0+\Delta T_{01}/ch(mL)\le T_{\max}, \quad (2)$$

- a permissible diode operating temperature. It becomes practical when $ch(mL)>(4..8)$. Then temperature distribution can be brought to the following form:

$$T(x)=T_0+\frac{\Delta T_{01}\{A_1\exp(mx)+A_2\exp(-mx)\}}{(1-n)\exp(mL)}, \quad (3)$$

Where $A_{1,2}=(1\mp n)\exp(\mp mL)$, m, n - heat return coefficients from the face and the sidewalls. This quantities dependence from the thermophysical constants and sensor sizes has been explained in [4,page 52] and reference manuals. It should be noted here, that m is proportional to the heat conduction coefficient n to the cross section area.

6. QUASI-STATIC MODE

To solve the equitation (1) lets expand it (as it is usually done) to the Taylor row relatively to the initial values of q parameters and by δq exponents. Let's equate a sum of first-degree terms to zero:

$$\Delta\omega=\frac{\delta\omega}{\omega}=\frac{-\Delta_\ell-\Delta_\varepsilon/2-M(\Delta_\rho+\Delta_D)}{1-M}= \\ =-\Delta_1-\Delta_2-\Delta_3-\Delta_4 \quad (4)$$

Where $M=(C\rho C_D\sin^2 F)/\ell\sqrt{\varepsilon}$, $F=\omega\ell\sqrt{\varepsilon}/C$, $\Delta=\delta q/q$ - relational increments.

Let's call two first items in numerator geometrical and "electrical" elongation.

Geometrical elongation adds up from the elementary segments elongation ΔX_i, heated to different temperatures. The last are obeyed to the law (2).

Then

$$\Delta_l=\frac{K_T}{(1-M)L}\int_0^l T(x)dx\approx \\ \approx\frac{K_T}{mL}\cdot\frac{n[1-\exp(-mL)]}{(1-n)}\cdot\Delta T_{01}, \quad (5)$$

Where K_T - linear expansion coefficient.

Expression for the "electrical" elongation can be written in the same way:

$$\Delta_\varepsilon=\frac{\beta}{(1-M)2L}\int_0^L T(x)dx\approx \\ \approx\frac{\beta}{2}\cdot\frac{1}{mL}\cdot\frac{n}{(1-n)}\cdot[1-\exp(-mL)]\cdot\Delta T_{01}, \quad (6)$$

Where β - temperature coefficient of permittivity.

Coaxial line wave resistance does not depend on proportional diameters increase and is related with $\varepsilon(T)$ growth only:

$$\Delta\rho[T(x)]\approx\frac{1}{2}\rho\sqrt{\varepsilon}\Delta_\varepsilon[T(x)] \quad (7)$$

Line segment, forming the resonator, becomes heterogeneous, that leads to the distributed reflection appearance in it.

Line segments of such type have been well researched in electrodynamics [5]. So, for the exponential ρ change (this is close to our case).

$$\Delta\Gamma_3=-\frac{Nsh(\alpha L)/\alpha}{ch(\alpha L)-(\gamma_3/\alpha)sh(\alpha L)}, \quad (8)$$

Where γ_3 - is a propagation constant;

$$N=\frac{1}{4}\ln\frac{1+\beta T_1}{1+\beta T_L},\ \alpha=\sqrt{\gamma^2+N^2}$$

Therefore:

$$\Delta_3=\left(\frac{1}{1-M}\right)\left(\frac{C}{2\omega L\sqrt{\varepsilon}}\right)\Delta\Gamma_3$$

Temperature coefficient of the diode capacity is notably high [6]. That is why, quantity Δ_4 essentially exceeds all other items, despite of the comparatively low diode heating. Its estimation by the first item of Taylor row gives a substantial error.

Besides, dependence $C_D(T)$ has been specified graphically. Thus, let's calculate Δ_4, numerically solving the equation (1), and presenting the result graphically (fig.2).

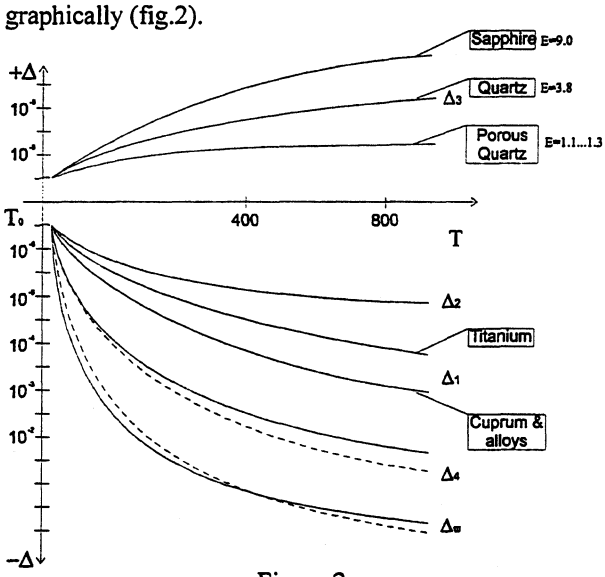


Figure 2

7. DYNAMIC TEMPERATURE MODE

Fast changes are typical for the process of burning spatio-temporal state, especially in a supersonic flow.

Temperature fluctuations on the fire surface run to hundreds degrees, and there spectrum $S_T(\nu)$ spreads to a few kHz. This spectrum envelope shape is weakly researched, it is highly probable that spectrum contains quasideterministic components. One of the cases of there appearance - is cyclic character mechanical processes: turbine and other technological device unites rotatory and alternating motions. Temperature propagation in the sensor body obeys the following law:

$$\partial T/\partial t = a\left[\partial^2 T/\partial x^2\right],$$

And excited thermal waves attenuate fast.

$$\Delta T_m(x) = T_{1m}\exp\left[-x\sqrt{\pi\nu/a}\right] = T_{1m}K(x,\nu)$$

Here T_{1m}, ν - fluctuation amplitude and frequency,

$a = \eta/C_g\gamma$ - diffusivity, C_g - specific heat, η - heat conduction, γ - density.

Normalized fluctuation spectrum:

$$S'_\omega(x,\nu) = S_\omega(x,\nu)/H_q = S_T(\nu)k(x,\nu),$$

Where $k(x,\nu) = K(x,\nu)/H_q$, Hq - normalizing divisor, by way of which we could use the estimation of quasi-static changes power.

Attenuation degree of fluctuation with frequency is notably high.

So for titanium sensor under $x = 10^{-2}m$ at frequencies 1Hz and 100Hz the quantity $K(x,\nu) = 4{,}5\cdot10^{-3}$ and $\sim10^{-5}$ respectively. Thus, temperature disturbance is essential only in the most low frequency range ($\nu < 10^{-2}$ Hz) of useful information spectrum.

8. DIFFERENT INSTABILITY FACTORS COMPARATIVE ESTIMATION.

Temperature dependencies of the frequency, average over the set of electrical, thermophysics and constructional factors, have been shown on fig.2.

Symbols have been matched with equation (4). Curves matching let us enunciate the request to materials: minor values of linear expansion coefficient, heat conduction, and minor permittivity for the dielectrics. The last fact affects the sensor's physical length, which is close to $\lambda/4\sqrt{\varepsilon}$. Considering m and n functions dependencies from the heat conduction η their ratio is $n/m \sim 1/\sqrt{\eta}$ [4]. That is why from the standpoint of instability decrease titanium is $\sqrt{36} = 6$ times better than cuprum.

One more condition plays a role for dielectrics: as permittivity decreases the geometrical length of sensor, i.e. mL product, increases. Here the porous quartz benefits from policor $\sqrt{900} = 30$ times. Diode capacity C_D contributes dominantly only during the technological process start-up.

In dynamic mode $\partial C_D/\partial T$ quantity is relatively small and $\Delta_4(t) \sim \Delta_{1,2}(t)$. Among the other factors, let's note the sensor's cross sectional area, which should also be minimized.

Diode heating and other instability components are increasing as the sensor is shortening. But for the condition $\omega < \omega_{cr}$ observation quarter-wave sensor has to be rather short $L < \pi C/2\omega_{cr}\sqrt{\varepsilon}$

This contradicts to the restriction $T(L) < T_{\max}$. Diode temperature $T(L)$ and instability $\Delta_\omega(L)$ dependencies have been shown on figures 3a, 3b.

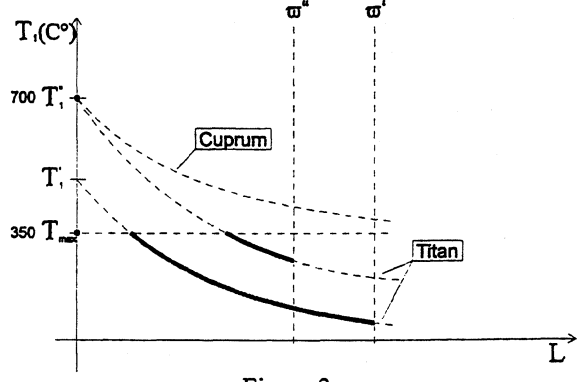


Figure 3a

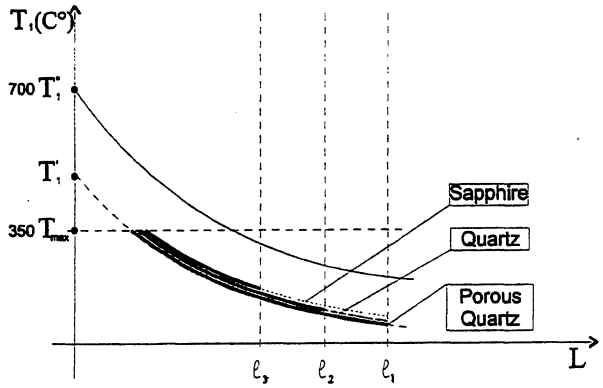


Figure 3b

Stroking lines - vertical (here $\omega''_{cr} > \omega'_{cr}$) and horizontal respectively show two specified restrictions. Curves parts, where both of the restrictions have been kept up, are shown thicker. The difficulty is that T_1 growth in the combustion chambers is attending by ω_{cr} growth, which means that T'_1 corresponds to ω'_{cr}, T''_1 - ω''_{cr}, and operating sector is shortening very fast. That is why quarter-wave sensors usage is limited. Multi-link resonator's sensor removes this problem.

$\Delta_\omega(L)$ curve is different from $T(L)$ by its scale, which is $(10^{-3}..10^{-4})$ 1/degree for the quasi-static case and $< 10^{-5}$ 1/degree for the dynamic fluctuations. Both quantities are essentially smaller than destabilizing antenna influence, which is examined in the separate study.

9.CONCLUSIONS

Retention of the functional characteristics and thermostability produce conflicting requirements to the self-excited sensors of hightemperature technological processes. Two conditions compatibility within the framework of traditional schematic and engineering solutions is possible only for the processes with rather low free electron density in the operating body.

The most critical unit - is diode, especially open-frame diode.

Traditional radiomaterials are of little use.

Titanium and radio-transparent dielectrics with $\varepsilon \to 1$ (spumes melted quartz) have got advantages.

Considerable frequency change can be seen during initial heating, i.e. at the process start-up. In the operating mode temperature fluctuations of the frequency are rather attenuated.

10. TECHNICAL SUGGESTIONS

Two suggestions have a technical novelty. The first one is the normal resonator's circuit replacement with the multi-link one [ ]. It let us carry frequency-specifying elements of circuit out of the intensive heating area. Disadvantage: sensor's sensitivity $\partial\omega / \partial\varepsilon_x$ decrease.

While realizing that proposition we have to study phase thermoinstability influence of transmission line, which connects the resonator with antenna, at the frequency and sensitivity.

The second one has an engineering nature and assumes that all main sensor details are made of dielectric, which requirements are obvious from foresaid. Only current-driving surfaces are subject to plating. Here we expect diode's heating decrease in 2...3 times and frequency instability - in 3...6 times. Disadvantage: lows mechanical durability.

Other suggestions are trivial and resolve into higher-frequency diodes usage, material and different heat removal actions choice.

REFERENCES

[1] V.V. Boloznev, A.N. Pikulev, A.D. Romanychev, "Auto-oscillation Systems and Amplifiers in Transmitters" [in Russian], Proceedings of Moscow Energy Institute, № 180, Moscow (1988), p.35.

[2] V.V. Boloznev, A.Sh.Chabdarov, "Permittivity Sensor for an Ionized Gas Flow", Kazan' Aviation Institute. Translated from Pribory i Tekhnika Eksperimenta, № 5 pp.149-152, September-October, 1992.Original article submitted October20, 1991; revision submitted January 8, 1992.

[3] V.V. Boloznev, E.V. Safonova, "Diode SHF-active oscillator with multi-link resonator in autodyne mode", preliminary program and abstracts of the IEEE 1988 international Symposium on Acoustoelectronics, Frequency Control and Signal Generation, St.Peterburg, June 7-12, 1998, p.146

[4] S.S.Kutateladze," Bases of heat exchange theory"[in Russian], Moscow: Atomizdat, 1979, pp. 52-54.

[5] A.L. Fel'dshtein, L.R. Yavich, "Four-therminal device and eight-therminal synthesis on SHF"[in Russian], Moscow: Svyaz' 1971, p.388.

[6] D.P. Tsarapkin "SHF-oscillators on the Gunn's diodes"[in Russian], Moscow: Soviet radio, 1972

[7] A.V. Luikov " Heat conduction theory"[in Russian], Moscow: High school, 1967, p.600

PROPAGATION OF LAMB WAVES IN 1-3 PIEZOCOMPOSITE BORDERED BY LIQUIDS

F. Teston, G. Feuillard and M. Lethiecq, *Senior Member, IEEE*

GIP Ultrasons/LUSSI, EIVL, 4 rue de la chocolaterie, 41000 Blois
E-mail : teston@univ-tours.fr

ABSTRACT

This paper deals with the effect of liquid loading on the propagation of Lamb waves in a piezocomposite plate. In these materials, the first symmetric mode exhibits a very low phase velocity and a strong electromechanical coupling factor compared to conventional materials (quartz, lithium niobate or ceramics). First, the theoretical dispersion curves are calculated and experimentally verified. Then, the influence of the viscosity and the adiabatic compressibility modulus of the liquid on the frequency and the attenuation of the S_0 mode is studied. The frequency and attenuation variations have been measured when the plate is loaded by a glycerol solution, and are found to be in fair agreement with theoretical calculations. The results indicate that the S_0 mode is suitable for liquid phase sensor applications.

INTRODUCTION

Numerous papers have been devoted to the study of the propagation of Lamb waves in a plate bordered with a liquid and to the description of liquid Lamb wave sensors [1-2]. Lamb waves propagating in thin plates are highly sensitive to boundary conditions when a surrounding medium loads the plate, Lamb wave phase velocity and attenuation are modified due to mass loading and viscoelastic effect. In such devices, the phase velocity of the A_0 mode, which is widely used in sensing, can be smaller than the sound velocity of a liquid. Thus, the radiation of the normal displacement component into the liquid can be minimized. Moreover, most of published papers concern the viscosity and density effects on the Lamb wave propagation [3-5]. They neglect the influence of the adiabatic compressibility modulus of the liquid. However, in opposition to other sensors which use shear waves, Lamb waves have a longitudinal particle displacement which interacts with the liquid and so depends to the adiabatic compressibility of liquid.

In this paper, we propose to use the S_0 mode in a 1-3 piezocomposite to develop liquid phase sensors. The first part is devoted to the validation of the model when the piezocomposite is assumed to be a homogeneous material, the properties of which are deduced from homogeneization models [6]. In the next part, we theoretically investigate the coupling between acoustic waves and the liquid where all the three parameters previously called to mind (ρ, η and κ) are taken into account. Finally, preliminary experiments with water and glycerol solutions are presented.

1. THEORETICAL ANALYSIS

The method used to calculate the vibrating modes in a piezoelectric plate is based on the effective permittivity concept [7] which requires the knowledge of all the mechanical and dielectric constants. The solutions of this problem are the plane waves propagating along the x_1 direction bounded by planes $x_3 = h/2$ and $x_3 = -h/2$.

In the first step, in order to consider the piezocomposite as a homogeneous medium, all the useful effective parameters such as the elastic, dielectric and piezoelectric constants were calculated. Thus, using the effective permittivity concept, the velocity of the Lamb waves in the piezoelectric plate for any electrical boundary conditions of both surfaces was calculated [8].

Then, the method was extended to include a Newtonian liquid overlaying the plate taking into account the mechanical and electrical particle displacement in the liquid and using appropriate boundary conditions.

The complex elastic constants of liquids, c_{ijkl}^{Liquid}, are given by

$$c_{ijkl}^{Liquid} = \lambda\delta_{ij}\delta_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})$$

where δ is the Kronecker symbol, $\lambda = \kappa - j\omega\frac{2}{3}\eta$ and $\mu = j\omega\eta$. κ and η are respectively the adiabatic compressibility modulus and the viscosity.

Here, the introduction of complex constants in the surrounding medium leads to complex solutions. The phase velocity and attenuation of the Lamb modes are then determined in the complex plane using a dichotomy method.

2. RESULTS AND DISCUSSION

For all the computations, the ceramic and the polymer used are PZT-5H (Navy type IV) and epoxy resin (Araldite D - Ciba Geigy).

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

The dispersion curves have been calculated for the unloaded plate according to three electrical boundary conditions on both surfaces. All these results are not given here, but are necessary to calculate the electromechanical coupling factor.

2.1 Electromechanical coupling factor.

The electromechanical coupling factor, k, has also been calculated using the approximation $k^2 = 2(V_0 - V_s)/V_s$, where V_0 (resp. V_s) is the phase velocity corresponding to the plate with both surfaces opened (resp. shorted).

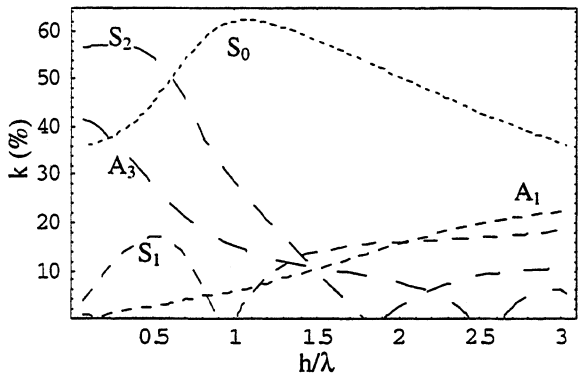


FIGURE 1 : Electromechanical coupling factors for a piezocomposite plate with 66 % of ceramic volume fraction. Label A refers to the anti-symmetric modes and S to the symmetric modes.

Figure 1 shows the electromechanical coupling factor for the eight lowest order modes versus the normalized plate thickness (h/λ). It is interesting to note that the coupling coefficient of the A_0 mode is very low (not represented), due to the low value of the piezoelectric constant e_{15}. That of the S_0 mode shows a rather unusual behavior : k^2 has a maximum value of 36 % at $h/\lambda \cong 1$ and is always higher than 9 %.

2.2 Dispersion curves.

In order to experimentally verify the dispersion curves, two Lamb wave devices were fabricated using the composite with 66 % of ceramic volume fraction. Each device consists of a delay line made of two single transducers : the first has a spatial periodicity, λ, of 940 µm and is 290 µm thick ; the second has a 570 µm spatial periodicity and is 270 µm thick. The propagation path on both surfaces are metallized. In order to consider the substrate as a homogeneous medium, the direction of the electrode fingers is chosen parallel to the diagonal of the ceramic rods and a finger covers at least two adjacent ceramic rods.

The phase velocity was determined from electrical impedance measurements. Figure 2 shows the theoretical phase velocities in a short/open electrical boundary condition and experimental phase velocities for the two devices.

The experimental results are consistent with the calculated results. Moreover, only modes which have a large coupling factor in Figure 1 are observed on Lamb wave devices. These results allowed to confirm that the homogenization model can be used for all the acoustic modes (not only the thickness mode) and extend the validity domain in terms of h/λ.

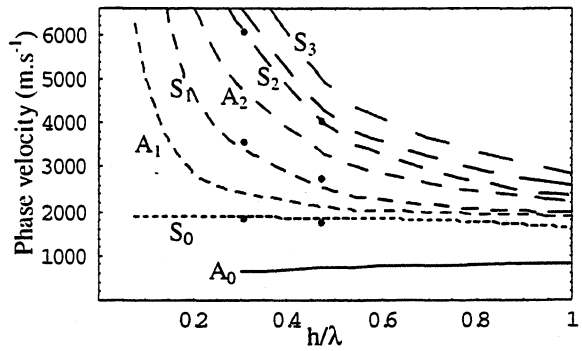


FIGURE 2 : Calculated dispersion curves of Lamb modes and experimental data (points).

All these results endow the S_0 mode with suitable characteristics for liquid phase applications.

2.3 Influence of the liquid on phase velocity.

Now, we focus our attention on the effect of the presence of a liquid on the phase velocity of the S_0 mode. To carry out this study, the effects of the density, ρ, the viscosity, η, and the adiabatic compressibility modulus, κ, of aqueous glycerol mixtures have been separately examined on Figure 3. The following values were used in calculations : for water $\rho = 998.23$ kg.m^{-3}, $\eta = 1.0019\times10^{-3}$ Pa.s and $\kappa = 2.24\times10^{9}$ N.m^{-2} and for glycerol $\rho = 1261.03$ kg.m^{-3}, $\eta = 1.0019\times10^{-3}$ Pa.s and $\kappa = 4.74\times10^{9}$ N.m^{-2} for 20°C temperature.

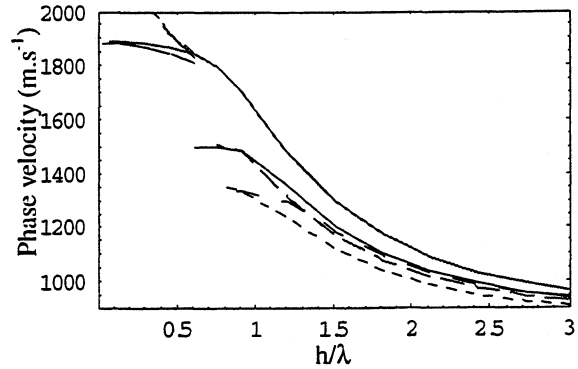


FIGURE 3 : Phase velocities of the unloaded plate (gray line), loaded by water (dark solid line) and loaded by glycerol. In this case, only some parameters are taken into account : density (grey dashed line), viscosity (large dashed line) and viscosity, density and adiabatic compressibility modulus (dashed line).

The behavior of the S_0 mode can be divided in two regions depending on the h/λ ratio and liquid loading.

When the plate is loaded by water and if the phase velocity is lower than the longitudinal wave velocity in water, the difference between the loaded and unloaded plate is significant, elsewhere the phase velocity is lower and the curve tends to merge with that of the free plate. When the plate is loaded by glycerol, the same behavior is observed when the phase velocity is lower than the water sound velocity. However, the velocity shift is much larger and the discontinuity occurs at a lower phase velocity. For the lower h/λ values, the behavior of the phase velocity is not usual. The phase velocity of the S_0 mode is much higher than that of a free plate and seems to increase as h/λ tends to zero.

In order to understand the relationship between the three parameters (ρ, η and κ) of glycerol and the phase velocity shift of the S_0 mode, we have successively changed one of these parameters. The calculation shows that slight changes occur on the phase velocity when the viscosity or the adiabatic compressibility modulus are taken into account. We have only represented the viscosity effect curve. For h/λ>0.7, this parameter slightly modifies the phase velocity, whereas for h/λ<0.7, this effect is more significant. When we examined the mass loading (liquid density effect), we found that the phase velocity is mostly perturbed in the normalized thickness range of 0.8 to 1.2.

2.4 Frequency and attenuation variations versus viscosity.

We have studied the influence of the adiabatic compressibility modulus versus the viscosity-density product when the normalized thickness (h/λ) is equal to 0.47 corresponding to the second device described in section 2.2.

Figures 4 and 5 represent the relative frequency variations and the attenuation variations versus the square root of the viscosity-density product with the water loading as reference. In order to describe the influence of the adiabatic compressibility modulus of aqueous glycerol mixtures, the two limit cases have been considered (pure water or pure glycerol).

For the lower value of $\sqrt{\eta\rho}$ on Figure 5 , the attenuation is negligible and does not depend on the κ value. It is important to note that a phase velocity condition (i.e., lower than sound velocity of liquid) is not required for the S_0 mode. Indeed, contrary to the A_0 mode, the particle displacement of the S_0 mode is mainly parallel to the surface (transverse component, u_1) and the radiation of the longitudinal component into the liquid is not significant even for $V_{S_0} > V_{Liquid}$.

On the other hand, the frequency variations depend on the κ value and are negative (Figure 4). The effect is larger for low values of $\sqrt{\eta\rho}$. Indeed, as the adiabatic compressibility modulus behaves linearly as a function of weight concentration (90 % variation in glycerol weight mixture leads to 90 % variation in the κ value while viscosity only varies of 10%), it has a value close to that of glycerol.

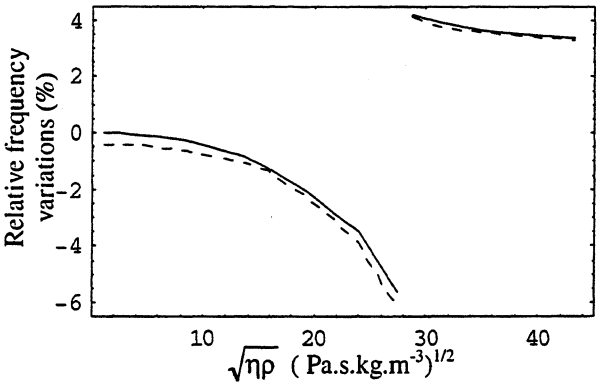


FIGURE 4 : Relative frequency variations (%) versus the square root of the viscosity-density product of aqueous glycerol solutions in the case where the adiabatic compressibility modulus, κ, is equal to 2.24×10^9 N.m^{-2} (water) in solid line or 4.74×10^9 N.m^{-2} (glycerol) in dashed line.

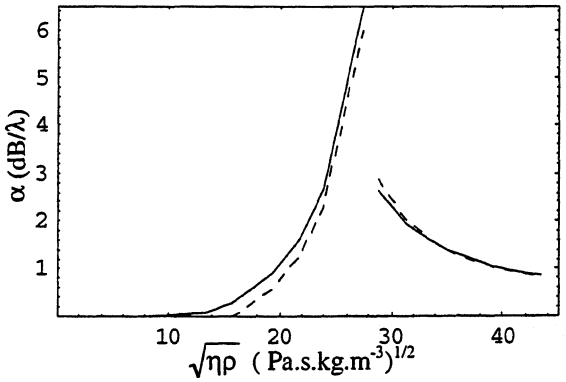


FIGURE 5 : Attenuation variations (dB/λ) versus the square root of the viscosity-density product of aqueous glycerol solutions in the case where the adiabatic compressibility modulus, κ, is equal to 2.24×10^9 N.m^{-2} (water) in solid line or 4.76×10^9 N.m^{-2} (glycerol) in dashed line.

Now, for increasing $\sqrt{\eta\rho}$ values, the frequency and attenuation variations vary rapidly. For $\sqrt{\eta\rho}$ >10, a slight difference is observed between two κ values. In this region, the phase velocity of the S_0 mode is lower than the longitudinal phase velocity in the solution, which means that no radiation loss occurs. Thus, the S_0 mode behavior is mainly due to the viscous effect.

A sharp attenuation peak is observed for a value of $\sqrt{\eta\rho}$ = 28 Pa.s.kg.m^{-3} which corresponds to a discontinuity in the relative frequency variation. At this point, the phase velocity of the S_0 mode (or resonant frequency) is greater than that of the plate loaded by water and takes the same value as that of the aqueous glycerol solution. The sign of the relative frequency variations changes from negative to positive and attenuation variations become lower and slowly decrease as the viscosity-density product increases. Another observation is that the adiabatic compressibility modulus effect is low and tends to become negligible .

2.5 Transfer function measurements.

The frequency responses of the second device previously described are presented in Figure 6 when the back side of transducers is in contact with air, water or pure glycerol.

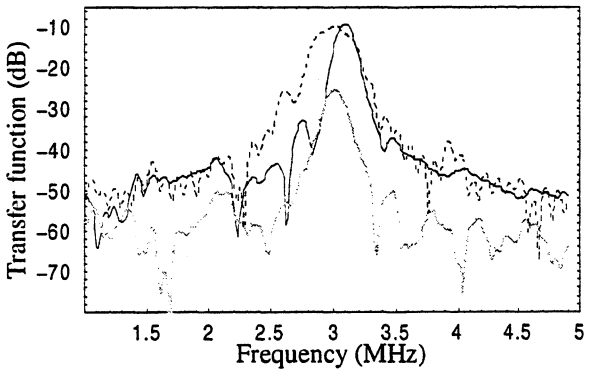


FIGURE 6 : Frequency response of the device loaded with air (solid line), water (dashed line) and glycerol (gray line).

The acoustic signal of the S_0 mode does not suffer attenuation (0.4 dB) when the surface is loaded by water and the resonant frequency is shifted to lower frequencies (73 kHz). These measured values are consistent with the theoretical results shown in Figure 3. One can observe that, the bandpass of the device increases.

When the plate is loaded by glycerol, the S_0 mode is highly attenuated (15.7 dB) due to the viscosity effect. As the center-to-center IDT separation is 23×λ, this result is in agreement with Figure 5. However, the resonant frequency does not seem to be strongly perturbed as expected (Figure 4).

3. CONCLUSION

Numerical analysis of the Lamb wave propagation in piezocomposite plates, with and without liquid loading, was carried out. The dispersion curves have been experimentally validated, which confirms that homogenization models can be used for all the acoustic modes and extends their validity domain.

The influence of three properties of liquids, i.e. the density, the viscosity and the adiabatic compressibility modulus has been studied for the S_0 mode. We can conclude that the density is not an important parameter compared to the two others for our experimental device. The measured results for water and glycerol solutions are partially consistent with the theoretical predictions. Other experiments must be carried out in order to confirm these results and, in particular, to explain the low experimental frequency shift compared to theoretical predictions. This study shows that S_0 mode in 1-3 piezocomposites are suitable for liquid phase sensor applications.

Acknoledgments

The authors are grateful to VERMON S.A. (Tours, France) for the piezocomposite manufacturing.

4. REFERENCES

[1] Zhu Z. and Wu J., "The propagation of Lamb waves in a plate bordered with a viscous liquid", *J. Acoust. Soc. Am.*, 98(2), pp. 1057-1064, (1995).

[2] Zaitsev B. D., Kuznetsova I. E. and Joshi S. J., "Propagation of acoustic waves in plates bordered with viscous liquid", *IEEE Ultrasonics Symp.*, (1997).

[3] White R. M., Wicher P. J., Wenzel S. W. and Zellers E. T., "Plate-mode ultrasonic oscillator sensors", *IEEE Trans. Ultrason. Ferroelect. Freq. Contrl.*, vol. 34, n° 2, pp. 162-171, (1987).

[4] Martin B. A., Wenzel W. S. and White R. M., "Viscosity and density sensing with ultrasonic plate waves", *Sensors and Actuators*, A21-A23, pp. 704-708, (1990).

[5] Vellekoop M. J., Lubking G. W., Sarro P. M. nad Venema A.., "Evaluation of liquid properties using a silicon Lamb wave sensor", *Sensors and Actuators*, A43, pp. 175-180, (1994).

[6] Levassort F., Lethiecq M., Certon D. and Patat F., "A matrix method for modeling electroacoustic moduli of 0-3 piezocomposite", *IEEE Trans. Trans. Ultrason. Ferroelect. Freq. Contrl.*, vol. 44 n°2, pp. 445-452; (1997).

[7] Milsom R. F., Reilly N. H. C. and Redwood M., "Analysis of generation and detection of surface and bulk acoustic waves by interdigital transducers", *IEEE Trans. Trans. Ultrason. Ferroelect. Freq. Contrl.*, vol. SU-24 n°3, pp. 147-166; (1997).

[8] F. Teston, G. Feuillard, D. Certon, F. Levassort and M. Lethiecq, "Propagation of lamb waves in 1-3 piezocomposites and their application to liquid sensors", in press to *Ferrolectrics*, (1999).

PIEZOELECTRIC HYDROGEN SENSOR ON THE BASE OF WO_3-Pd FILMS

I.A. Ges, B. A. Budkevich

Institute of Electronics, Belarus National Academy of Sciences;
22 Logoiski Tract, 220841, Minsk, Belarus;
Fax: (37517)2652541; E-mail: i-ges@inel.bas-net.by

ABSTRACT

The questions of hydrogen piezoelectric sensor making are considered in this work. The method of metallic palladium precipitation, based on colour centres transformation is offered. The palladium film possessed a developed surface and served the basis for hydrogen sensor of piezoelectric type on the base of thin-film structure WO_3 / Pd – quartz crystal. Influence of sensor initial loading with sensitive covering on its steady-state characteristics is explored. Hydrogen concentration limits of sensor using are determined. The obtained results confirm the efficiency of an offered method of palladium precipitation when making sensitive layer of hydrogen sensor.

1. INTRODUCTION

In our wok we propose new method of sensitive layer creation for piezoelectric hydrogen sensor.

The various types of hydrogen sensors (semiconductor, fibre-optic, electrochemical, piezoelectric and etc.) exist now [1-2]. The principle of their operation is based on different physical effects in solid body. The presence of selective sensitive material to the hydrogen is general for all sensors.

Mass-sensitive piezoelectric sensor can be successfully used as a detector element in devices of monitoring of an environment gas composition [3-4]. The principle of piezoelectric sensors operation is based on changes of oscillation frequency of quartz crystals covered with special sensitive material, which interacts with researched gas. Different chemical materials, which selectively absorb corresponding gases, are used as sensitive coating of quartz crystal.

2. METODS AND RESULTS

The properties almost all-solid state chemical sensors based on metal–oxide layers may be essentially improved by doping of their surfaces catalytic metal. All advantages and disadvantages of such structure widely discussed in literature, mainly for the chemical sensors having the general configuration: catalytic metal / insertion compound / insulator / semiconductor [5]. As a rule, catalytic metal is a solid thin film of Pt or Pd, while insertion compound is WO_3 (or MoO_3) because of its specific property. Thin amorphous films of WO_3 and MoO_3 are well known photo – and electrochromic materials for wide variety of application. The essence of electrochromic effect is concluded in appearance of an absorption band in electrochromic material under the action of the electrical or electromagnetic fields. As a result on its surface and in the bulk are formed coloration centres [6]. The concentration of coloration centres is easily varied under changes of external influence and technological ways and conditions of films producing.

In this paper we represent result on the fabrication of modified layers WO_3 and MoO_3 allowed to change the concentration and structure catalytic metal grains on its surface.

For vacuum evaporated thin films WO_3 and MoO_3 there are three simple methods of their preliminary coloration, which we used in work. It is action of:

– UV light [6],
– electric fields [7],
– chemical active metal in the acid solution [8],

which give rise to different levels of colour centres concentration in it's from 10^{10} to 10^{22} cm^{-3}. As we have found, after such operations electrochromic film can react with the ions of noble metals in aqueous solutions with formation of metallic phase on their faces. Such electron – ion reactions have been investigated for Pt^{4+}, Pt^{2+}, Pd^{2+}, Ag^{+}, Au^{3+} ions [9].

Here we discuss the process of coloration centres transformation in more detail. In our experiments the tungsten oxide films were coloured at the contact of the film surface with the active metal in the acid ambience. The films were coloured to optical density D ~ 2. The quantity of tungsten atoms, participating in electron capture determines the concentration of coloration centres.

$$m\ W^{6+} + ne^{-} \Leftrightarrow n\ W^{5+} + (m-n)W^{6+}$$

Fig.1 shows the succession of appearance Pd metallic phase on the surface of WO_3 films. The velocity of such process depends only both the initial concentration of colour centres and ions of corresponding noble metal in aqueous solutions.

According to results of SEM-analysis amorphous coloured film is characterised by developed surface microrelief with the granularity size 0,05- 0,2 мкм [10]. At the contact of coloured film with water solution of a palladium chloride a bleaching of electrochromic film with precipitation of metallic palladium particles on surface and bulk centres of coloration occurs. Reaction of metallic phase extraction can be presented in the manner of:

$$ne^{-} + mA^{n+} \Leftrightarrow mA^{0} + nH^{+}$$

0-7803-5400-1/99/ $ 10.00

where n - an amount of electrons, participating in reactions of forming one metal atom; m- a minimum number of metal atoms, giving development centre; A - noble metal atoms.

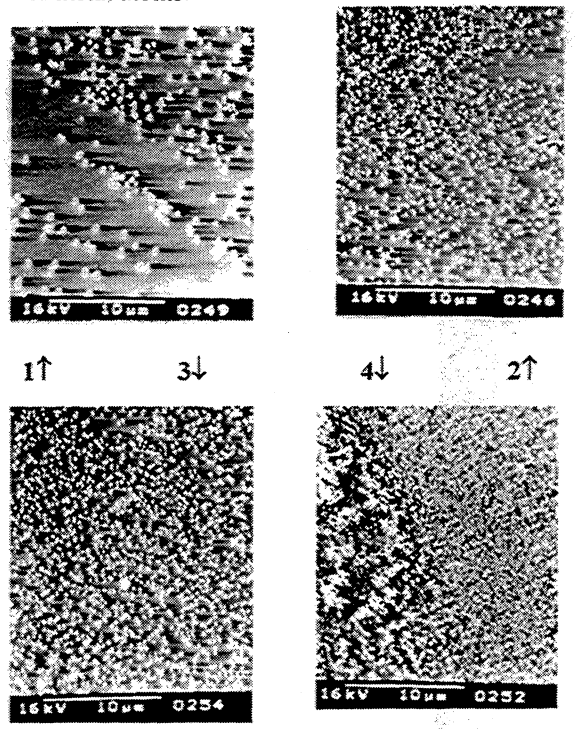


Fig.1. The appearance Pd metallic phase on the surface of WO_3 films after 20 min (1), 40 min (2), 90 min (3) and 120 min (4) in palladium chloride solution

Further contact of WO_3 film with a palladium chloride solution leads to the growing of metallic palladium particles and after 12 hours of interaction with solution a surface of a palladium polyatomic particles possesses developed relief, where size of particles reaches several microns. An investigation of metal particles growing process points to presence of two-stage process connected with the coloration centre transformation on the surfaces and in the bulk of film. Observed experimentally dependence of quick stage velocity from noble metal ions concentration testifies about diffusion nature of the metallic centres precipitating process. Fig.2 shows the modified surface of WO_3 films with well forming catalytic metal grains for its using as gate layer of sensitive structure.

Selectively sensitive elements are a base of all sensors. A metallic palladium can be used as such element for hydrogen detection. The palladium more actively absorbs hydrogen. One atom of palladium can absorb one atom of hydrogen. Such absorption leads to increasing of covering mass that causes of reducing fluctuation frequency of quartz crystal [4,11]. Usually palladium is used as film, which produced by the method of vacuum evaporation or electrochemical precipitating in the manner of black palladium. However such films are characterised by sufficiently high density, that only expects the application of surface layer as absorb surfaces. The sensitivity and dynamic range of piezoelectric sensor can be greatly raised when producing a palladium porous film with developed surface.

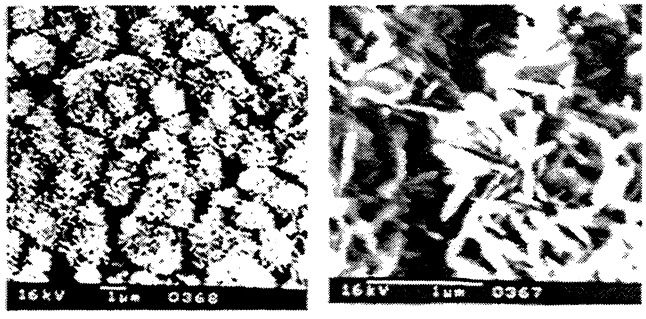


Fig.2 The surface of Pd film after 12 hours of interaction with palladium chloride solution (different resolution).

Our method of formation of catalytic active metal layer on surfaces of tungsten oxide film has a number of advantages in comparison with a method of vacuum deposition of metal films. The coloration centres are direct initiators of metallic phase emanation on surfaces of WO_3 film, forming under their photochemical or electrical activation. The coloration centres are formed both on the surface and in the bulk of tungsten oxide film and conditions of external influence and technological method of preparation can change their amount. In this case the forming of metallic phase has a chaotic nature and the film can have high catalytic activity owing to developed microrelief.

The offered method of precipitation of noble metal (Pd) layer on the tungsten oxide film was used for making of a hydrogen sensors (fig.3), formed on base of thin-film structure WO_3/ Pd.

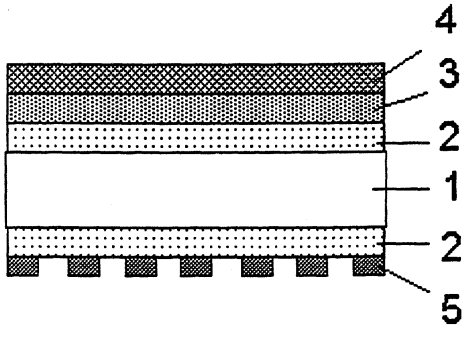


Fig. 3. The structure of piezoelectric sensor: 1 – quartz crystal; 2 – Ag–electrode; 3 – WO_3 film; 4 – Pd–film; 5 – heater.

Amorphous tungsten oxide thin films 0,2-0,5 μm thick were prepared by the thermal evaporation technique, in vacuum of $1{,}3 \times 10^{-3}$ Pa, 99,9% WO_3 powder. The WO_3 electrochromic film was coloured at the contact of its surface with the active metal (Zn) in

the acid ambience (1N H_2SO_4). The precipitation of Pd layer on surfaces of coloured WO_3 film was effected in water solution of palladium chloride acidifying by HCl. Four sensors were prepared for study of the sensor initial loading influence by sensitive material (WO_3 modified films) on its steady-state characteristics. Two pairs of sensors with different loading, equivalent reducing of a base sensor frequency from 7 kHz to 36 kHz for two quartz resonators with different pedestal frequencies (2 181 kHz and 3 299 kHz) were prepared. Steady-state characteristics of sensors are brought on the fig. 4 and fig.5. They have practically linear nature in the field of small concentrations of hydrogen (before 1,5 mg/l). Further increasing of a hydrogen concentration is characterised by the exponential dependency. The sensor enters to the saturation mode at the concentration more than 3 mg/l, when increasing of a hydrogen concentration does not lead to a frequency changing.

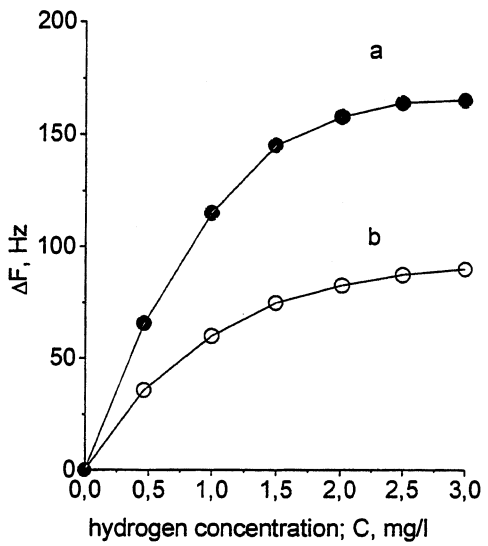


Fig.4.Steady-statec characteristic of piezoelectric sensor (basic frequency is 3 299 kHz) with different loading (a - 36 kHz; b – 18 kHz)

From steady-state characteristics seen that sensitivity (curve steepness on the linear area) sensor that more, than more preliminary sensor loading. Increasing of the quartz resonator pedestal frequency also leads to increasing of the sensor sensitivity. However, the reliable work is observed for the sensors with the pedestal frequency before 10 MHz. From our results it is possible to make a conclusion that piezoelectric sensor on the base of WO_3 modified films is usable for hydrogen detection within the range of its concentrations from 0,1 to 4 mg/l.

The influence of temperature on the static characteristic of sensor has brightly expressed character and is one of main destructive factor, which influence on sensor operation. By results of researches were

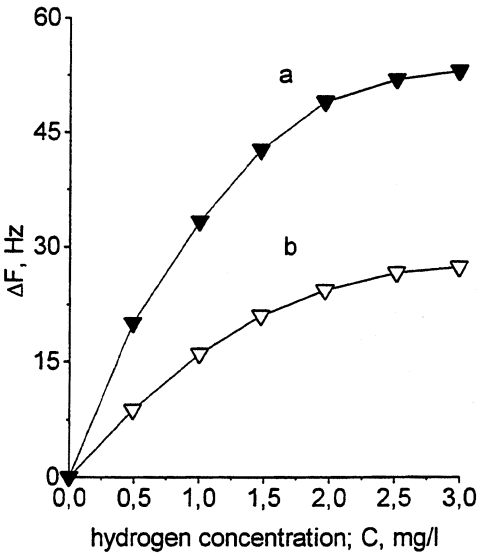


Fig. 5. Steady-state characteristic of piezoelectric sensor (basic frequency is 2 181 kHz) with different loading (a - 14 kHz; b – 7 kHz).

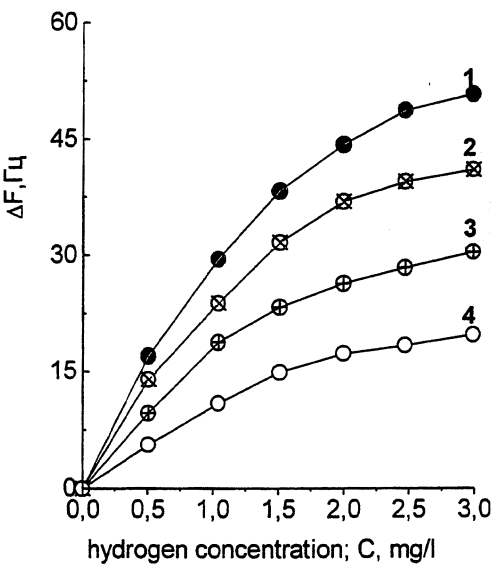


Fig. 6. The temperature influence on steady– state characteristic of piezoelectric sensor with loading of 9 kHz: 1–15; 2–20; 3–30; 4 – 40 °C.

obtained the dependence of frequency changes ΔF on hydrogen concentration for four significance of temperature T = 15, 20, 30, 40 °C. Fig.6 shows the steady-state characteristics for sensor with preliminary loading 9 kHz for called temperatures. The decrease of sensitivity with growth of temperature is a general tendency for all piezoelectric sensors. The sensitivity to temperature changing, besides the only chemical reasons, is connected also with thermosensitivity of piezoresonator itself. For piezoelectric chemical sensor

one of most temperature–steady piezoshear, namely AT-shear is used. At the changing of frequency at a level 10^{-5} - 10^{-7} from a resonant the oscillation of frequency owing to oscillations of temperature become noticeable and are perceived as noise. It results in decrease of a threshold of sensor sensitivity [11]. For decrease of temperature influence on operation of sensor it is necessary to apply thermostatically controlled regime or temperature compensation, or to use the scheme of matching with similar basic quartz, located in immediate adjacency from measuring quartz. But in the latter case the chemical component of remains without attention [3]. All these methods of thermosensitivity reducing are characterised by advantage and limitation and, according to them, find the applications.

Our result point to a possibility of the sensitive covering using on the base of tungsten oxide modified films (as a result of coloration centres transformation) for making of piezoelectric type hydrogen sensors.

3. REFERENCES

[1] A. Mingmei and K. Yufang , "A new hydrogen sensor with the rare earth complex fluoride", Sensors and Actuators, B, vol.15-16, pp. 179-180, 1992.

[2] M.A. Butler, "Optical fiber hydrogen sensor", Appl. Phys.Lett., vol. 45, No. 10, pp. 1007-1008, 1984.

[3] N.A Burlaenko and A.V. Maslennikov, "Ammonia piezoelectric sensor its steady-state characteristics". Defectoskopia, No. 4, pp. 23-26, 1996 (In Russian).

[4] L. Spassov, D.Y. Yankov, A.N. Mogilevski and A.D. Mayorov, "Piezoelectric sorption sensor for mercury vapors in air using a quartz resonator", Rev. Sci. Instrum, vol. 64, No1, pp. 225-227, 1993

[5] R.D.Rauh." Micronic Insertion Sensors". Solid State Ionics, vol. 28-30, pp. 1479-1485, 1988

[6] S.K. Deb, "Optical and photoelectric properties and color centres in thin films of WO_3", Phil. Mag., vol.27, pp. 801-821, 1973

[7] B.W. Faughnan and R.S. Crandal, "Electrochromic display". In book edited by J.I. Pankov: Display devices. Topics in Applied Physics, vol. 40. Springer-Verlag, New-York, 1980.

[8] C.G. Granqvist, Handbook of inorganic electrochromic materials. Elsevier, Amsterdam, 1995.

[9] B.A. Budkevich, I.A. Ges, V.A. Pilipovich, I.M. Romanov and L.I. Romanova, "Colour centres transformation to development centres in amorphous WO_3 thin films", Vesti Akad. Nayk BSSR, seria fiz.-math. Nauk,. No. 1, pp. 56-60, 1988 (In Russian).

[10] B.A. Budkevich, I.A. Ges, S.P. Zhvavyi, G.D. Ivlev and V.A. Pilipovich. "Modification of electrochromic WO_3 films by pulsed ultraviolet laser irradiation", Phys.stat.sol.(a). vol.110,pp.155-164, 1988

[11] V.V. Malov, Piezoelectric devices. Energo-atomizdat, Moscow, 1989 (In Russian).

Optimization of Sensing Film for Quartz Crystal Microbalance Odor Sensor Using Self-Assembly Deposition Method

A.SAITOH, T.NOMURA, S.MUNOZ[1] and T.MORIIZUMI[1]
Dept. of Communication Eng., Shibaura Inst. of Tech.,
[1]Dept. of Electrical and Electronics Eng., Tokyo Inst. of Tech.,
3-9-14 Shibaura, Minato-ku, Tokyo 108-8548, Japan.

ABSTRACT

We demonstrated a new self-assembly method using thiol compounds for making a sensing film with a desired specific response property, using the interaction between odor molecules and sensing film molecules due to their affinity. The sensing film was deposited self-assembly on a QCM with gold electrode in the mixed solution of sensing film materials and odor molecules. This method realizes the easily making the sensor with a desired response property.

In addition, we investigated response property of the sensor having a surface function group controlled sensing film by self-assembly method. In this experiment, amino group, hydroxyl group, and $-CH_3$ group were used for the function group. The results of a series of the sensor response measurement indicated the importance of the function groups on the sensing film surface and the necessity of controlling the surface.

1. INTRODUCTION

Chemical nanogravimetry sensing using quartz crystal microbalances (QCM)[1] or surface acoustic wave (SAW) devices[2] in conjunction with organic sensing film is a powerful tool in the fields of odor sensing and biosensing. Because these sensors are coated with selective film composed of organic compounds, the selectivity of the sensor is determined by the properties of the sensing film, specifically, the material and structure of the film. Previous studies have concentrated on the role of film material in the determination of selectivity, thus overlooking the importance of the relationship between selectivity and film structure. Recently, host-guest interaction[3] and the molecular imprinting method[4] have been applied in the area of molecular recognition. The host-guest interaction is a result of structural interaction between a host molecule, such as cyclodextrin, crown ether, or calixarene, and a guest molecule. The molecular imprinting method forms specific ab/adsorption sites using molecularly imprinted polymers.

In this paper, to compose a sensing film with a specific and a desired response property, we developed a new self-assembly method utilizing molecular imprinting method. Actual film deposition is performed using thiol compounds

In addition, the influence of the function group of a sensing film surface for sensor response properties was investigated.

2. EXPERIMENTAL

A thiol molecule is deposited via the chemical binding between thiol groups (-SH) and gold surface[5] and forms monomolecular film. This method was used for the deposition method of sensing films. In this experiment, four different thiol compounds were used as the sensing film materials. These included two different alkane thiols, hexadecane thiol:$CH_3(CH_2)_{15}SH$ (HDT) and decane

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

thiol :$CH_3(CH_2)_9SH$ (DT), and a compound that contains both an amino group and a thiol group, aminoethane thiol:$H_2N(CH_2)SH$ (AET), and a compound that contains both a hydroxyl group and a thiol group, p-hydroxy thiophenol:HOC_6H_4SH (HTP). Each thiol compound was prepared as 5mM ethanol solution. Sensing film was deposited on the surface of the device over a period of more than 20 hours immersing in the solution.

For the sensor devices, quartz resonators equipped with gold electrodes were used at a resonant frequency of 20MHz. Sensor responses were measured in the confined chamber under a controlled temperature (18℃). Sensor response was defined as the difference between the oscillation frequencies before and after odor molecule adsorption.

3. RESULTS AND DISCUSSION

3.1 HDT/DT/AET mixed film sensor using additional odor method

A new self-assembly method for the deposition of the sensing film with a desired response property is described in this section. This method makes use of the affinity between odor molecules and sensing film materials in a depositing solution. In the mixture solution of odor molecules and sensing film materials, they interact due to their affinity. That results in the most suitable formation of those molecules in the solution and on the electrode of the QCM.

For the sensing film materials, HDT, DT, and AET were mixed in a 1:1:1 molar ratio in an ethanol/water solvent (7:3). Water was added to the solvent to increase the hydrophobic interaction between film materials and odor molecules. In this experiment, benzene and 1-hexanol were used as additional substances (as the objective odor gas samples) and were introduced into the depositing solution. The quartz resonator was then immersed in the solution.

Figures 1 and 2 show the concentration dependence of the sensors for benzene and 1-hexanol gas, respectively. The result for the sensor coated with the mixed film without the addition of any odor molecules in the deposition process is shown as the dotted line and is labeled "no addition". Others are indicated with the name of the added odor molecule. The results indicate improved selectivity when odor molecules are added during the deposition process. A suitable film structure for the added odor molecule seemed to be formed by the interaction between the film materials and the odor molecules.

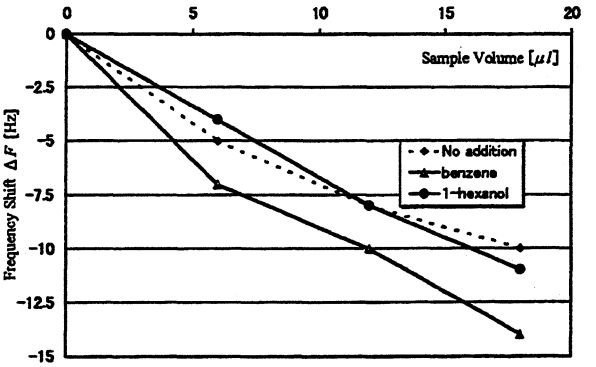


Figure 1: Concentration dependence of sensor response with HDT/DT/AET mixed sensors for benzene gas.

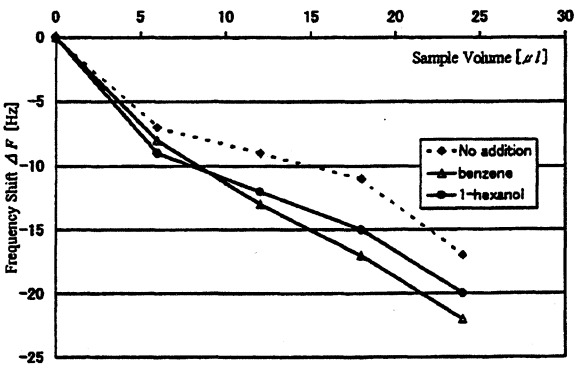
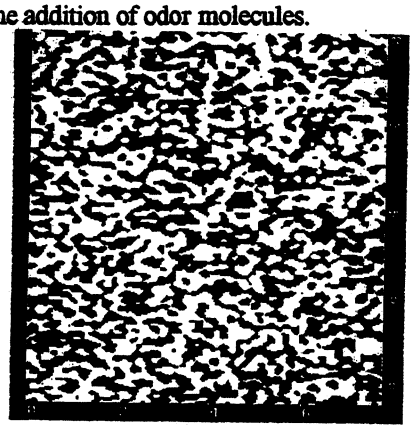
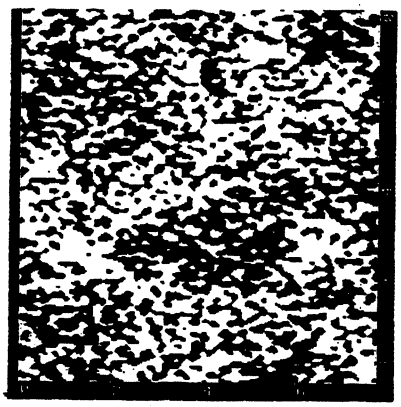


Figure 2: Concentration dependence of sensor response with HDT/DT/AET mixed sensors for 1-hexanol gas.

For investigation of structural changes in the sensing film caused by the addition of the odor molecules, the surface potential distribution in the sensing film is observed using scanning Maxwell stress microscopy (SMM). Figures 3(a) and 3(b) show SMM images of pure HDT/DT/AET mixed film and HDT/DT/AET mixed film with 1-hexanol added during the deposition process, respectively. In both figures, the light color indicates parts consisting of HDT or DT molecules, and the dark color indicates the parts of AET molecules. Differences in the SMM images show that the film structure changes with the addition of odor molecules.



(a)



(b)

Figure 3: SMM images of HDT/DT/AET mixed film.
(a) Mixed film without the addition of any odor molecules. (b) Mixed film with 1-hexanol added in the deposition process. The images are shown with two colors, treating the average value of the distribution of the surface potential as the boundary.

Figure 3(a) shows a state that occurs when the three thiol compounds are mixed and deposited uniformly because of no addition of any odor molecules. Figure 3(b) shows a nonuniform image, which indicates the existence of a large domain of AET molecules in the thiol film with 1-hexanol added. The domain surface has the $-NH_2$ group of the AET molecule. Thus, domains play important roles as adsorption sites for hydrophilic groups of odor molecules, such as the -OH group in the 1-hexanol molecule.

3.2 Influence of sensing film surface for sensor response property

Function groups on sensing film surface are important as well as sensing film material and structure. We investigated the relation between the function group and the sensor response property.

Self-assembly deposition method realizes the controllability of the group of the sensing film surface. In this experiment, amino group (AET), hydroxyl group (HTP), and $-CH_3$ group (DT) were used as the function group. The sensor responses were measured for several volatile organic compounds. Figure 4 shows the time responses of each sensor. The sensor response patterns are shown in figure 5. The patterns are normalized by the maximum response for each gas sample. Two types of response pattern are appeared in figure 5. One of the patterns is the response for non-polar molecules (benzene, n-hexane) and the other is for polar molecules (acetone, 2-hexanone). That indicates the functional groups influence the sensor response properties. From this result, molecular occupied areas were calculated. Figure 6 shows the molecular occupied area for each sample. From the figure, we can obtain that the molecular occupied area in AET

and HTP is almost same and that the molecular occupied area in AET and HTP sensors is smaller than that of DT. AET and HTP have the polar function group on the sensing film surface. The group interacts with polar odor molecules and determines the adsorption phenomena (orientation of adsorbed molecules onto the surface due to dipole – dipole interaction). On the other hands, the molecular occupied area of the polar molecules onto the DT film is larger than that of the non-polar molecules. Because DT film doesn't have the functional group onto the sensing film surface, the interaction force between sensing film and odor molecules is weak.

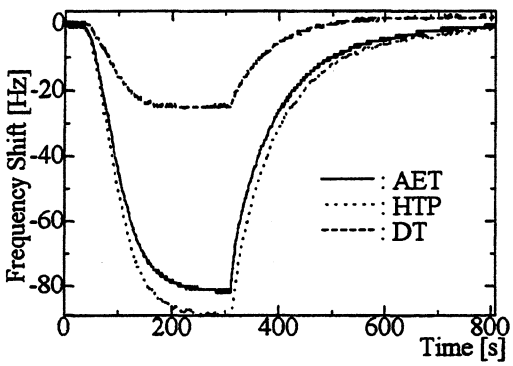


Figure 4: Time dependence of the sensor response for benzene

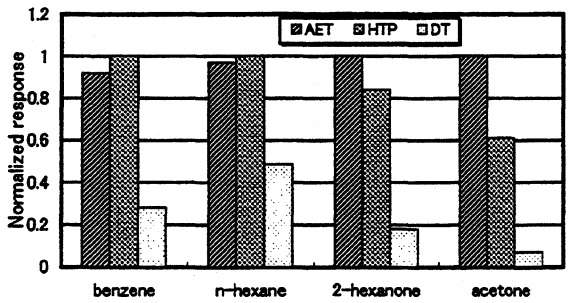


Figure 5: Response patterns for volatile organic compounds

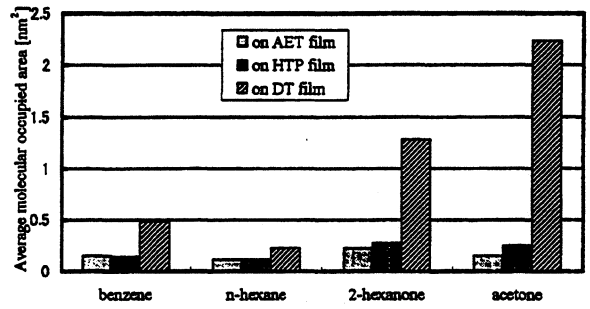


Figure 6: Average molecular occupied area

4. CONCLUSIONS

In this experiment, thiol compounds were used for the sensing film materials.

We demonstrated a new self-assembly method for the deposition of sensing film with a desired response property, using the affinity between odor molecules and sensing film materials in the depositing solution. By using this method, we can easily obtain a sensor with the desired response property.

The sensing film surface is very important in the first stage of the adsorption of odor molecules and determines the response property of the sensor. From the experiment, surface polar function groups on a sensing film are important for determination of adsorption phenomena and the orientation of the adsorption molecules.

5. REFERENCES

[1] G. Sauerbrey: Z. Phys **155** (1959) 206.

[2] H. Wohltjen: Sens. & Acts. **5** (1984) 307.

[3] W. Gopel: Sens. & Acts. **A 56** (1996) 83.

[4] K. Mosbach and O. Ramstom: Biotechnology **14** (1996) 163.

[5] L. Bertilsson and B. Liedberg: Langmuir **9** (1993) 141.

1999 Joint Meeting EFTF - IEEE IFCS

METHODS OF HARMFUL EFFECTS LIMITATION IN SAW GAS SENSORS

Judyta Hechner , Waldemar Soluch, Tadeusz Wróbel

Institute of Electronic Materials Technology
Wólczyńska 133, 01-919 Warsaw, Poland

ABSTRACT

In this paper some problems concerning work stability of SAW NO_2 sensors employing copper phthalocyanine (PcCu) as the chemical interface and 128°YX lithium niobate as substrate are presented. The proposed sensing element consists of two delay lines: one of them is covered with a chemical coating (PcCu), the another is covered with Al and then PcCu. In this case a response of the sensor depends only on the conductivity changes of the PcCu layer during NO_2 exposure. Other effects are identical and therefore are canceled.

1. INTRODUCTION

SAW gas sensors are made by applying a thin chemically sensitive film onto the surface of the SAW delay line. In general the detection principle of the SAW gas sensor relies on the change of the SAW velocity upon adsorption of the reactant by the sensor film. This velocity change are related to the following effects: the change in the mass density of the film, the change in its electrical conductivity or the change in its elastic constants.

It is common to build SAW chemical sensors in a dual-delay-line configuration, operating in a differential frequency mode. As it follows from the studies of other authors [1, 2] and from our own experiments [3, 4], the dual-delay-line concept is not an ideal solution of the stability problem in sensor work. These experiments revealed some phenomenons which destabilize detection process: unidirectional drift of generated frequencies at constant temperature (±0,1°C) and ambiguous sensor response (during the reaction with investigated gas), which did not give clear results.

Considering these factors we have started experiments which should define and eliminate the harmful elements of detection process.

We have been working with nitrogen dioxide SAW sensors. The experiments have been carried out for SAW delay lines on 128°YX lithium niobate. This substrate characterized by high electromechanical coupling constant and low level of bulk waves. The interface material is copper phthalocyanine (PcCu), which is an organic p-type semiconductor. PcCu is not an intrinsic semiconductor [5-7]. However, the chemisorption of NO_2 causes the creation of a charge transfer complexes with PcCu molecules by accepting an electron from PcCu, which results in increasing of the hole concentration.

In this paper, the influence of SAW sensor configuration, substrate properties, operation temperature, structure and geometry of the sensitive layer on stability, inversibility and reproducibility of gas SAW sensor response is presented.

2. EXPERIMENTAL PROCEDURE

A dual-delay-line SAW device on 128°YX lithium niobate with two different transducers has been designed. The long one was a ladder type (294 electrodes), the short one was an ordinary type (38 electrodes). We used four electrodes per wavelength to eliminate reflections. Insertion loss at the center frequency (70 MHz) was about 15dB [8].

SAW delay lines were fabricated using standard lift-off photolithography.

Thin films of PcCu were prepared in a two stage process. At the first stage the input material was purified by vacuum sublimation and at the second it was vacuum deposited on piezoelectric substrates.

The standard SAW sensors consist of two identical delay-lines. One of them is covered with a chemical coating, another uncoated is used as a reference. Our sensing delay line is coated with PcCu, while the another one is covered with Al+PcCu (Fig. 1).

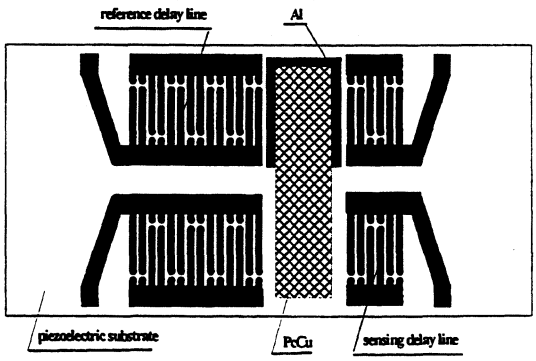


Fig. 1. Sensor element configuration

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

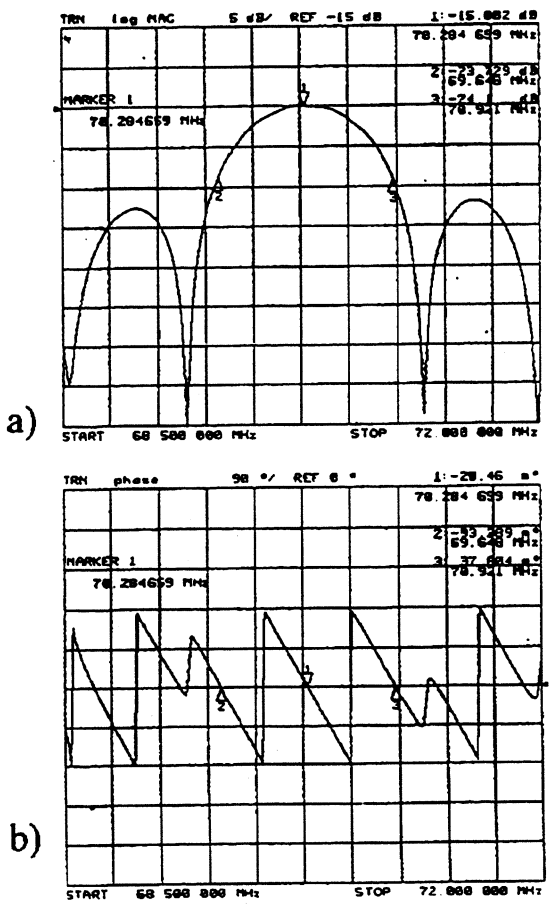


Fig. 2. Delay line responses
(a) amplitude
(b) phase

The experimental measuring system contained the following modules:
- the gas - mixing system.
- SAW device, which is introduced into the exposure chamber, containing the gas inlet and outlet and heating system.
- electronics circuits which include amplifiers, mixer and frequency counter. Amplitude and phase responses of delay-lines have been measured use of Hewlett Packard Network Analyzer 8725 A (Fig. 2).

3. EXPERIMENTAL RESULTS

3.1. Phase shift stability

The drift of generated frequency is noticed by most experimentators, who are working on SAW sensors.

The directions of the phase shift due to absorption of the reactant by chemical coating and of spontaneous drift (hence also the drift of frequency) are illustrated in Fig. 3.

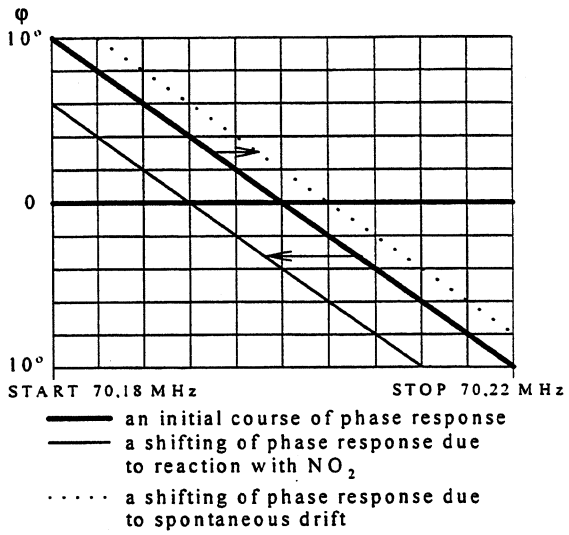


Fig. 3. The direction of phase response shift (φ).

Opposite direction of phase changes due to these phenomenons lead to destabilization of the sensor work, especially in the case of low concentration of investigated gas.

The spontaneous frequency drift is probably connected with pyroelectric effect. The surface charges on $LiNbO_3$ wafer produce unidirectional drift of the phase response at the constant temperature.

The degree of susceptibility to charging depends on the operation temperature, gas atmosphere, as well as on the kind of delay line coating.

The phase stability of a single delay line at a constant temperature was investigated in the temperature range 30°÷170°C. The difference in behavior between the coated and uncoated delay lines has been studied. The phase versus time curves for the uncoated delay line at different temperatures in N_2 atmosphere are shown in Fig. 4.

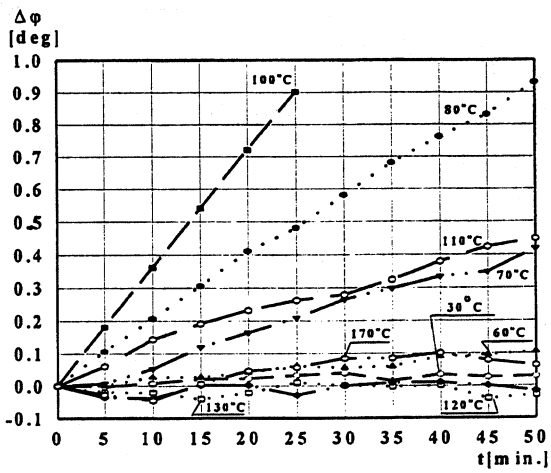


Fig. 4. Phase shift variation as a function of time for different temperature for the uncoated delay line.

The temperatures at which generated frequency drift diminishes for uncoated delay line in different gases (N_2, He, air) are presented in Table 1.

Table 1. The temperatures at which the generated frequency drift diminishes

Kind of gas	Air	Nitrogen	Helium
Temperature	≅110°C	120°C	125°C

Our measurement show that for some investigated temperature range there is a region of strongly unidirectional drift of the phase shift (Table 2).

Table 2. The temperature regions of the unidirectional phase drift.

Uncoated delay-line	PcCu coated delay-line	Al + PcCu coated delay-line
60÷120°C	50÷90°C	45÷80°C

The differential frequency drift of SAW device can be explained by difference in susceptibility to pyroelectric effect of coated and uncoated delay lines. In the case of our device the discrepancy between delay line covered with PcCu film and Al+PcCu is insignificant (Fig. 5).

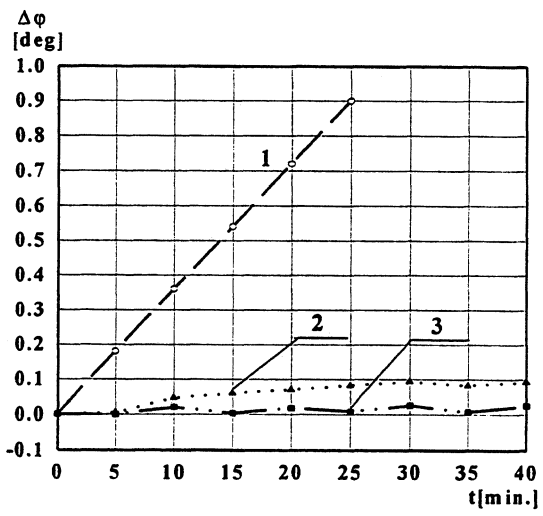


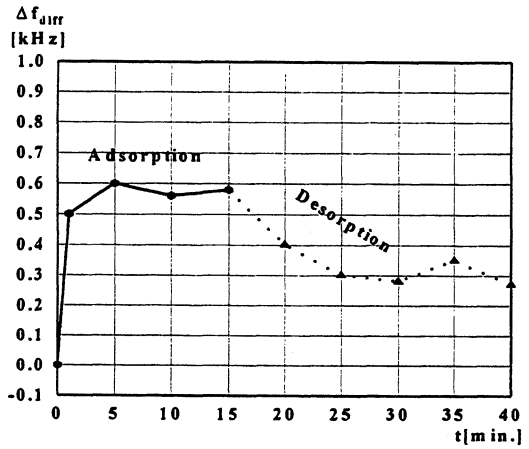
Fig. 5. Phase shift versus time at 100°C.
1 - uncoated delay line
2 - PcCu coated delay line
3 - Al and PcCu - coated delay line

3.2. Sensor responses

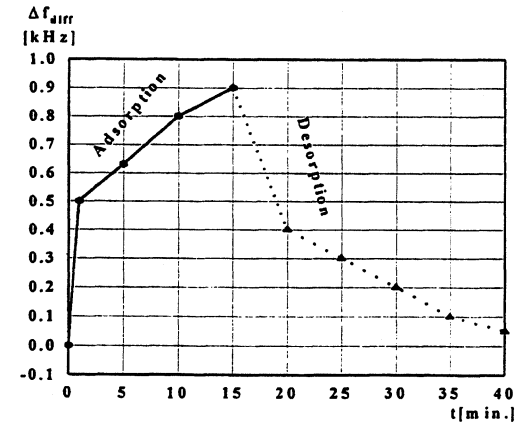
The interaction between PcCu layer and oxidizing gases (e.g. NO_2) is related to such processes as physisorption, chemisorption and desorption.

Both physisorption and chemisorption are exothermic reactions. In contrary to physisorption, chemisorption requires substantial activation energy, however after delivering of energy needed for initiation goes on with heat emission. Desorption is endothermic reaction. Therefore opposite energetistic effects are revealed in sensor responses. In the case of uncovered reference delay line, the above processes and self-drift of the difference frequency (Δf_{diff}) disturb the adsorption and desorption curves (Fig.6.a).

In the case when both delay lines are covered with PcCu layer a response of the sensor depends only on the conductivity changes of PcCu coating during the NO_2 exposure. Other effects are identical and therefore are canceled (Fig. 6.b.). As a result reproducible and univocal response has been obtained.



(a)

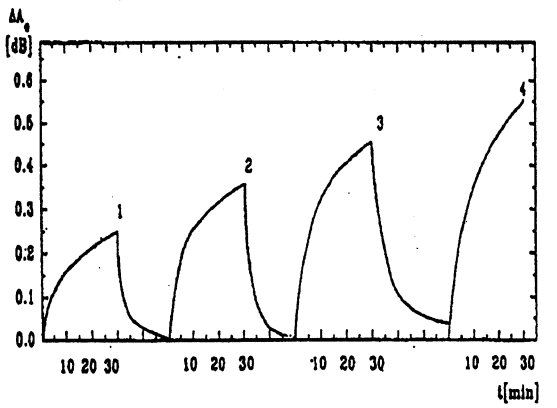


(b)

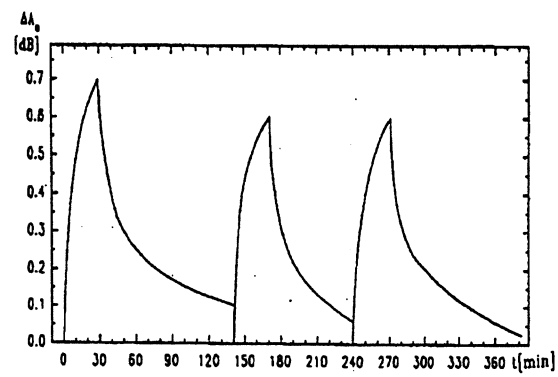
Fig.6. Sensor response
(a) uncoated referens delay line
(b) reference line with PcCu layer

It was found that PcCu layers before the proper analysis process requires a special activation ensuring recurrent results (Fig. 7). Since, surface adsorption centers are closed by parasitic molecules, activation process leads to unblock this centers by intermittent periodic exposing to NO_2.

In the case of α-PcCu formed by small particles (diameter = 100nm) an activation process leads to higher detection sensitivity and for β-PcCu (formed by needles length ≥ 1μm) the sensitivity is the same but it can be obtained in shorter time.



(a)



(b)

Fig. 7. Sensor responses to NO_2
(a) before activation
(b) after activation

4. CONCLUSION

Our measurements show that for the investigated range of temperatures there is a region of unidirectional differential frequency drift. This phenomenon can be explained by the difference in susceptible pyroelectric effect of coated and uncoated delay lines. In the case of the described configuration the discrepancy between delay line covered with PcCu film and Al+PcCu is insignificant.

Preliminary activation of PcCu layer ensure unequivocal and recurrent detection process.

The results of this work confirm that the use of SAW delay line coated with Al and PcCu as a reference element improve stability, inversibility and reproducibility of sensor responses.

5. REFERENCES

[1] Z. Liron et all, „ Temperature effect and chemical response of surface acoustic wave (SAW) single-delay-line chemosensors", Sensors and Actuators B-12, 1993, pp. 115-122.

[2] A.J.Ricco, „SAW chemical sensors" The Electrochemical Society Interface, Winter 1994, p.38.

[3] J. Hechner. T. Wróbel, „The influence of PcCu layer crystalline structure on the parameters of SAW Gas sensors" in Proceedings of Tenth European Frequency and Time Forum, 1996, pp. 370-375.

[4] J. Hechner, W. Soluch, T. Wróbel, „ Using SAW sensors to characterise adsorption and desorption of NO_2 modules on copper phthalocyanine film" in Proceedings of the 11-th European conference on Solid-state Transducers, 1997, pp. 801-804.

[5] M. Passard et al, „Gas sensitivity of phthalocyanine thin films", Sensor and Actuators B.18-19, 1994, pp 489-492.

[6] R. Zhou et al, „ NO_x sensitiwity of mono-meric and polimeric-N-macrocylic compounds", Sensor and Actuators B. 15-16, 1993, pp. 312-316.

[7] Shu Ji Qin and B. Bott, „The sensitivity to NO_2 of sandwich devices based on lead phthalocyanine and copper phthalocyanine", Sensor and Actuators B. 24-25, 1995, pp. 54-57.

[8] W. Soluch, „Design of SAW delay lines for sensors in Proceedings of the 11-th European conference on Solid-state Transducers, 1997, pp. 797-800, and Sensors and Actuators A-67. 1998, pp. 60-64.

ACKNOWLEDGEMENTS

This work was supported by the State Committee for Scientific Research (KBN) under grant No. 8 T10C 013 11.

SURFACE ACOUSTIC WAVE VAPOR SENSOR USING ULTRATHIN MUTILAYER FILMS

T.NOMURA, A.SAITOH and S.FURUKAWA*
Shibaura Institute of Technology, *Kyushu Institute of Technology

ABSTRACT

A surface acoustic wave (SAW) sensor incorporating an ultra thin selective film is presented. The ultra thin films were made by both self-assembled method and by Langmuir-Blodgett method. SAW delay lines coated with SA film and LB film has been used as vapor sensor. Rapid and totally reversible sensor responses were obtained by using the thin monolayer film.

1. INTRODUCTION

Surface acoustic wave devices have the ability to directly respond to the inertial mass and elastic mechanical properties of materials in contact with the device surface. This feature enables them to directly sense mass and mechanical properties of their environment. Therefore a potential application of SAW technique is an environment sensor.

In general, most of the SAW device using as a sensor is coated with a thin film that is capable of interaction with the chemical of interest through physical and chemical absorption. The quality of the selective film, such as life, uniformity, and thickness, influences the sensitivity, response time, and reversibility of the sensor response.

Ultra thin organic films are currently gaining interest in many areas such as integrated optics, sensors, and coatings reducing friction or surface orientation layers. The Langmuir-Blodgett (LB) film technique provides a very useful convenient method to make a uniform and ultra thin film on a substrate [1]. The other film is an organic film by self-assembled (SA) method. This method has been developed by G.Decher et al [2], [3]. These methods are very suitable to make an ultra thin and uniform film.

In this paper, we show a new SAW gases sensor which utilize SA and LB film. Experimental results show that the SAW velocity change was quite linear in proportion to the concentration of organic gases. Moreover we show that the thin film is suitable to obtain good linearity, high sensitivity and rapid response.

2. SAW SENSOR STRUCTURE

In order to check the performance of the ultra thin film as a chemical interface of gas sensor, experiments were conducted on SAW devices shown in Fig.1. The SAW device was consists of dual delay lines on the same piezoelectric substrate. In the structure one delay line is used as a reference channel, and a thin metal film is deposited on the path of the reference channel in order to short the electric fields at the surface. The other delay line is used as a sensing channel and remains a bear surface. For both delay lines, the whole of the propagation path is coated with the thin selective film. Absorption of chemical gas causes an increase in the density, mass, permittivity, and electric conductivity of the film. All these factors tend to decrease the velocity of the SAW. In the present sensor structure, the dual delay lines configuration cancels out any systematic changes which are common to the two delay lines, such as the effects of the changes in temperature, and nonspecific mass loading. However, the nonsystematic effects, such as the electrical change of the SA film, cannot be cancelled by the present sensor configuration. Therefor, only the change in the electrical property due to the moisture and the chemical gas in air is detected as a sensor response.

SAW sensor employing LB and SA film were fabricated on a 128° YX $LiNbO_3$ substrate. The entire surface of the two devices was covered with an ultra thin monolayer. The each delay line had a center frequency 30MHz and a center-to-center spacing between the input and output IDT of 10mm. The insertion losses of these delay lines were approximately 20 dB.

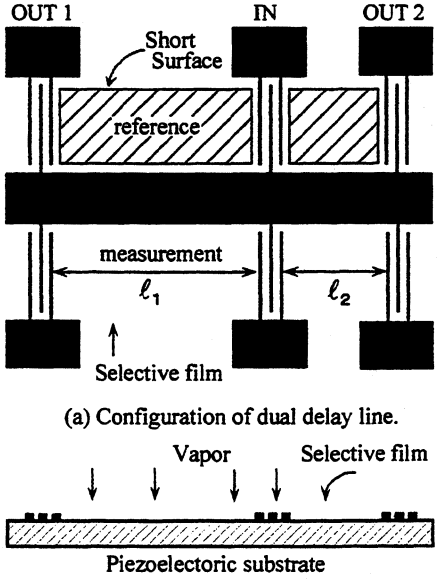


Fig.1 Schematic diagrams of a SAW humidity sensor using LB film and SA film.
(a) Configuration of dual SAW delay line.
(b) Cross-section view of the SAW delay line.

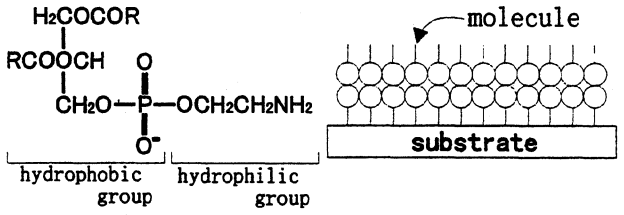
0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

3. SAW SENSOR COATED WITH LB FILM

The SAW device is coated with a thin film that is capable of interacting with chemicals of interest through various physical and chemical sorption mechanisms. These interactions determine the sensitivity, selectivity, and reversibility of the SAW sensor. The LB film method provides a convenient way to obtain precise and highly reproducible mass loading on SAW device [4,5]. The LB method is also suited to application of thin organic films one a time. By precisely controlling the number of deposited LB monolayers on a SAW device, we are able to estimate the mass loading effects of a thin selective coating [6].

In this study LB films were investigated for a selective film of SAW sensor. The LB film transfer technique is described in detail elsewhere [5]; therefore only a short explanation is given here. In the LB transfer technique, a small amount of the solution on the surface of clean deionized water in the LB trough. After the materials were compressed by reduction of water surface area, the dense and one molecule thick film were formed and transferred onto the SAW substrate by a vertical dipping method. Each stroke of the substrate through the air-water interface adsorbed lipids and allowed two monolayers to be transferred onto the SAW device. In this way, the ability to control the thickness with precision permits a highly reproducible mass loading to be applied to the SAW device.

In order to determine the sensitivity of the sensor, fundamental experiments were conducted. The propagation path was coated with a hygroscopic polymer film deposited by the above Langmuir-Blodgett (LB) transfer technique. The phosphatidylethanol-amine was used as a hygroscopic material. Figure 2 shows the chemical structure of phosphatidylethanol-amine and its image of the LB film.



(a) Structure (b) LB film

Fig.2 Chemical structure of the phosphatidylethanol-amine and LB film.

Measurements have performed using a vector voltmeter (HP model 4193A). The variation of the SAW velocity corresponding to the humidity is shown in Fig.3. It is found that the variation of the film due to humidity causes linear changes in the SAW velocity. Between 10 and 70% in RH, the SAW velocity varies almost linearly with humidity, with a slope of 0.33 m/s per 10% change in RH. The above results suggest that a SAW humidity sensor utilizing LB film is promising.

The variation of the phase difference between the reference and measurement channel, as a function of humidity, is shown in Fig.4. The variation of the phase, sensor output, varies with a slope of about 0.025 degree per 1 % change in RH. The results clearly indicate that it is possible to design the LB film a priori so as to achieve a desired linear response in the SAW sensor.

Absorption and desorption of the moisture is also shown in Fig.4. The hysteresis in absorption and desorption cycle is very small, and the two curves show good linearity.

Moreover, the results showed that the fractional velocity change of the SAW delay line depends on the electrical effect due to moisture in air. The SAW humidity sensor offer good linearity, high sensitivity and rapid response.

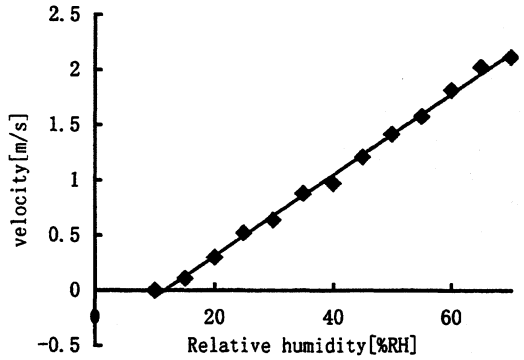


Fig.3 Variations of SAW velocity due to a relative Humidity (%RH).

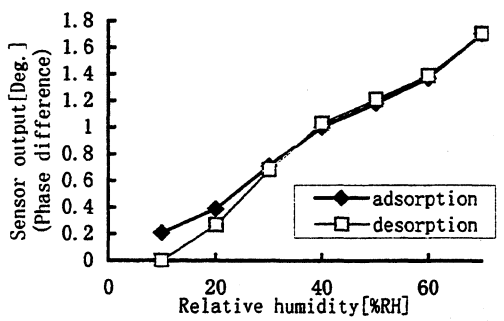


Fig.4 Output of SAW sensor produced by moisture absorption and desorption cycle.

4. SAW SENSOR WITH SA FILM

The buildup of the SA films is shown as follows by G.Decher. (Described in detail in [2] and [3]) The buildup of mono-layer films is shown in Fig.5. A substrate with a positively charged planar surface is immersed in the solution containing the anionic polyelectrolyte and a monolayer of the polyanion is adsorbed. Since the adsorption is carried out at relatively high concentrations of polyelectrolyte, a number of ionic groups remain exposed to the interface with the solution and thus the surface charge is reversed (a similar "non-flat" adsorption has also been observed by others. After rinsing in pure water, monolayers of the SA film are obtained. In the process of the film formation, controlling pH, concentration, ionic strength, etc. make it possible to adjust the thickness at the order of 0.5-2nm.

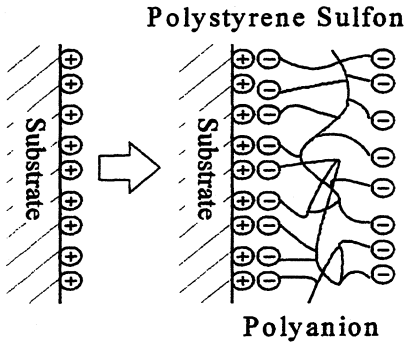


Fig.5 Schematic representation for the buildup of self-assembly mono- and multi-layer.

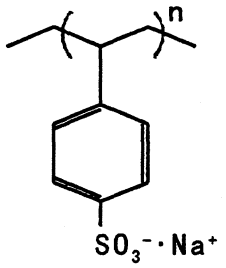
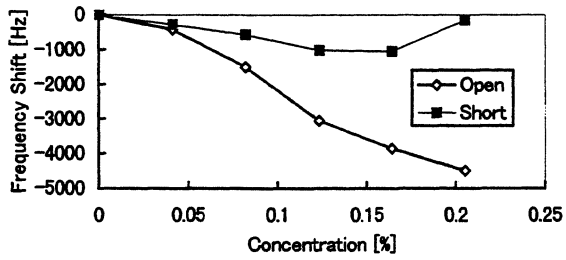


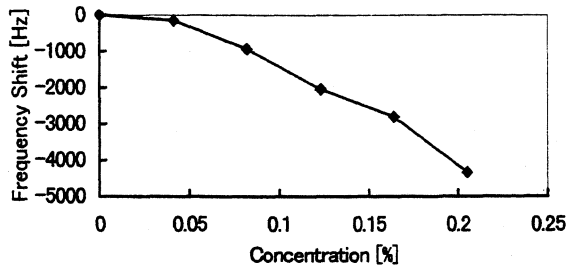
Fig.6 Chemical structure of the monomer unit of the $PSS^- \cdot Na^+$ polyelectrolyte.

Polystyrene sulfonic acid sodium salt ($PSS^- \cdot Na^+$) was used as a film material. The polymer is well known as a material for resistive humidity sensor [7]. The chemical structure is shown in Fig.6. The molecules are dissociated in pure water to form polyanions solution. Meanwhile the surface of SAW device is usually chemically active and is oxidized to be in possession of hydroxyl group. Therefor in an acid solution the surface of the devices is charged positively owing to hydration reaction. The polyanion solution by $PSS^- \cdot Na^+$ is controlled in a weak acidic solution with pH 6.5. From this reason, the SAW devices to be a substrate are charged positively, and PSS^- that became anions is adsorbed. Thus mono-layer film formation is made. Actually, by immersing a substrate in a 0.5mM solution of $PSS^- \cdot Na^+$ for 30 min, adsorption is made. The substrate, which is pulled up at a speed of 2cm/min after that, is immersed in pure water for 2 min to remove excessive film molecules. With dehydration at a room temperature for 24 hours, sensing film is obtained.

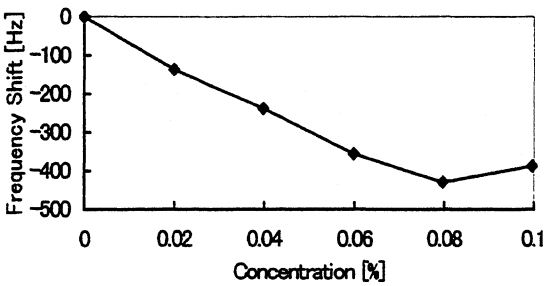
In this experiment, these delay lines are connected in the feedback circuit of an amplifier, resulting in SAW oscillator. The frequency change was measured as the sensor responses. Two organic gases such as methanol and ethanol have been chosen as gas substance. The frequency variation of the each SAW oscillator corresponding to the concentration of the methanol gas is shown in Fig. 7(a). Figure 7(b) shows the difference frequency between the short and open channel. As expected, frequency of the both oscillators decreases with increasing the concentration of the sample gases. The individual oscillators exhibit a unique frequency shifts for ranges 10% to 20%. The large fluctuation generated on the individual channels is believed to be due to the temperature change when the sample is poured. However in the case of Fig. 7(b) the output in proportion to the concentration of the sample can be obtained. It is explained that the SA film function as a selective sensing film. Figure 7(c) shows the response result complying with ethanol, and the output in proportion to the sample concentration can likewise be obtained. When the results complying with the 2 samples are compared, difference is noted with the two output responses. Thus it turns out that the electric characteristics of the film are also dependent on the sample.



(a) Frequency shifts on the reference and measurement channels for the methanol gas.



(b) Output of the sensor verses methanol concentration.



(c) Output of the sensor verses ethanol concentration.

Fig.7 Frequency shifts as a function of organic gases concentration on the reference and measurement channels.

The sensitivity and reproducibility of the sensor are largely affected by the stability of the organic selective film. In order to check the medium stability of the SAW sensor, after 3 months, the response was measured again. Figure 8 shows as an example that the SA film is applied to a quartz oscillation in comparison with the characteristics after 3 months, none of so remarkable deterioration in the characteristics can be seen. It can be deduced from the above observation that sufficient stability is kept with the sensing film in comparison with the fact that the life of a lipid membrane often used for a smelling sensor is approximately 1 month.

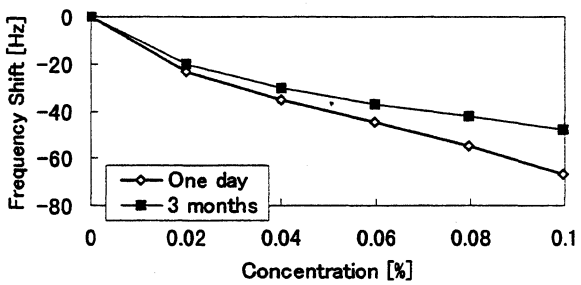


Fig.8 The stability in a time interval of 3 months to the ethanol gases by using a quartz oscillation.

5. CONCLUSION

The SAW sensor incorporating an ultra thin selective film was investigated. The ultra thin films were made by both self-assembled method and by Langmuir-Blodgett method. SAW delay lines were coated with the SA and LB films have been used as vapor sensors. A dual delay line was fabricated on a single lithium niobate (128° YX $LiNbO_3$) substrate and the propagation path was coated with the selective polymer film deposited by these transfer techniques. The phosphatidylethanol-amine was used as an LB film material and the polystyrene sulfonic acid sodium was used as an SA film material. Due to sorption of vapor molecules, the density, permittivity and electrical conductivity of these selective films changed. The effects caused to reduce the velocity of the SAW. Experimental results showed that the SAW velocity change was quite linear in proportion to the concentration of organic gases. Rapid and totally reversible responses were obtained by using the thin monolayer film.

In this study, we found that the ultra thin film is suitable to obtain good linearity, high sensitivity and rapid response. However, the sensitivity of the sensor with LB film decreases a few weeks after the first experiments.

This study was supported by the Proposal-Based New Industry Creative Type Technology R&D Promotion Program from the New Energy and Industrial Technology Development Organization (NEDO) of Japan.

REFERENCES

[1] K.Blodgett, "Films built by depositing successive monomolecular layers on a solid surface," J. Amer. Chem. Soc., vol.57, pp.1007-1022, 1935.

[2] G.Decher, J.D.Hong and J.Schmitt, "Buildup of ultrathin multilayer films by a self-assembly process: III. Consecutively alternating adsorption of anionic and cationic polyelectrolytes on charged surfaces", Thin Solid Films, 210/211(1992) pp.831-835.

[3] G.Decher, Y.Lvov and J.Schmitt, "Proof of multilayer structural organization in self-assembled polycation-polyanion molecular films", Thin Solid Films, 244 (1994) pp.772-777.

[4] B.Holcroft, G.G.Roberts, A.Barraud, and J.Richard, "Surface-acoustic-wave device incorporating conducting Langmuir-Blodgett films," Electronic Lett., vol.23, no.9, 1987.

[5] H.Wohltjen, W.R.Barger, A.W.Snow, and N.L.Jarris, "A vapor-sensitive chemiresistor fabricated with planar microelectrodes and a Langmuir-Blodgett organic semiconductor film," IEEE Trans. Electron Devices, vol.ED-32, pp.1170-1175, 1985.

[6] T.Nomura, M.Takeyabayashi and S.Furukawa, "Chemical sensor based on surface acoustic wave resonatr incorporating Langmuir-Blodgett films", Proc. IEEE Ultrson. Symp., pp.445-555(1995).

[7] Keiryou Handobukku ed. Kougyougijituin (Korona Ltd., Tokyo, 1987) p.1076. [in Japanese]

AN ATTEMPT FOR NH_3 DETECTION BASED ON QUARTZ RESONATOR WITH THIN SnO_2 FILM

V.Georgieva, L. Spassov, V. Georgiev
Institute of Solid State Physics, Bulgarian Academy of Sciences
72 Tzarigradsko Chaussee blvd., 1784 Sofia, Bulgaria

ABSTRACT

Flammable and toxic gases are very common components in the environment. Various types of physical and chemical sensors are created for their detection. A lot of them are based on different kinds of MOS field effect transistors with a combination of high sensitive thin layers.

In this paper the sensitivity of quartz resonators with thin layer of tin oxide (SnO_2) deposed on them is investigated for the registration of ammonia in the air. The measurements are carried out at NH_3 concentration from 10 ppm to 5000 ppm. It is found correlation between relatively change of the resonator's frequency ($\Delta f/f$) and the concentration in the investigated interval. The provided dependence can be described by equation like $f(x)=ax^{\alpha}+b$. To obtain the relationship between the thickness of the thin tin oxide and sensitivity of quartz resonator, the thickness of SnO_2 is changed from 40*nm* to 120*nm*. With the increasing of the thin SnO_2 thickness the resonator's sensitivity considerably increases. The highest values in the alteration of the frequency of the quartz resonators are registered by layer's thickness of 120*nm*.

The changes in the sensitivity of sorption as function of ammonia concentration at different SnO_2 thickness are calculated. It is shown that at small concentration sensitivity of the thinner layers is higher than at the thicker. At concentration 70 ppm this dependence becomes inversely and shows tendency to constant values at concentration over 5000 ppm. The results obtained show that the system quartz resonator-thin SnO_2 film could be used for detection of NH_3 concentration down to 10 ppm. Such system could be used for development of acoustic sensor for monitoring NH_3 contamination in the environment.

1. Introduction

The flamable gases like CH_4, liquefied petroleum gas (LPG) and H_2, as the toxic gases like CO, H_2S, NO_x and NH_3, are very common components in the environment [1]. The measurements and control of these gases are important for human being and also for a broad spectrum of industry and technologies. For this aim various kinds of physical and chemical sensors are created. The main requirements to gas-sensors for their wide applications are: 1) good sensitivity in a wide range of concentrations of the toxic gases; 2) quick response; 3) good reproducibility and no hysteresis; 4) long life and etc [2].

The basic structure as well as the physical existence of the MOS field-effect transistor is no doubt of great importance for the development of a whole series of sensors for the measurement of physical and chemical environmental parameters. The equation for the MOSFET drain current already shows a number of parameters, that can be directly influenced by the external quantity, but small technological variations of the original MOSFET configuration also rise to a large number of sensing properties. FET-based sensors such as the GASFET, OGFET, ADFET, SAFET, CFT, PRESSFET, ISFET, CHEMFET, REFET, ENFET, IMFET, BIDFET etc, are developed up to the present [3].

Semiconductor gas sensors based on the combination of the sensing behaviour of thin SnO_2 films with the advantages of the microelectronic structures are widely investigated too. Tin dioxide is one of the most useful materials for gas sensor [4,5]. The main reason for this is its high surface sensitivity to gas adsorption, simplicity for preparation and fast response time [6]. Usually the layers of thin SnO_2 are deposited by common microelectronic processes such as chemical vapour deposition (CVD), resistively and e-beam evaporation or RF sputtering. In recent years the concentrated efforts have been made to develop gas sensitive acoustic wave sensors, which exploit the various acoustic waves [7]. In this paper are presented results obtained by investigation of the quartz-resonators with thin SnO_2 film for detection of NH_3 concentrations in the air. Investigations are based on the correlation, which exists between mass sensitive resonator's parameters and deposited additional mass.

2. Experiment

The experiments are carried out with quartz resonators, which are formed on polished AT-cut quartz plates - 8 *mm* diameter and 4 *mm* diameter of gold electrodes. The electrodes are prepared from two layers: thin Cr sublayer about 10*nm* thick and upper *Au*-layer - 80*nm* thick. The *Cr*-layers are deposited from *Cr*-target by RF sputtering in *Ar*-ambient at vacuum 3*Pa*. The golden layers are formed by resistivity evaporation at initial vacuum $5.10^{-2}Pa$. The sensitivity thin SnO_2 layers are deposited on the quartz piezoelement by RF sputtering from SnO_2-target in *Ar*-ambient containing 16% O_2 at the total pressure ($P_{Ar}+P_{O2}$) of 0.3*Pa* with deposition rate about 2 *nm/min*. The processes of thin layer formation are carried out without heating of the substrates.

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

Obtained in this way resonator structures with thin SnO_2 layer are kept over tempered water solution of NH_3 with various concentrations. The details about the experiments and the measurements are described in [8]. The alteration in the resonator's frequency (Δf) calculation in the case of AT-cut quartz can be expressed by equation (1) [9].

$$\Delta f = -2{,}26.10^6 . f^2 . \Delta m / S_\perp, \qquad (1)$$

where:

Δf - is a resonant ferquency change, due to the sorbed additional mass (Δm),

f - is resonant frequency and

$S_\perp$ - is covered area of the electrodes (0.125 cm^2).

Here f is in *MHz*, Δf - in *Hz*, Δm - in g and $S_\perp$ - in cm^2.

3. Results and discussion

The dependence of resonator's frequency change versus ammonia concentration in the air over solutions with five different concentrations is presented in Fig.1(a). The measurements are carried out at concentration of NH_3 respectively: 10 ppm, 100 ppm, 500 ppm, 1000 ppm and 5000 ppm. It is obvious, that by increasing of NH_3-concentrations, relative change of the frequency shows tendency to increase too in all investigated range.

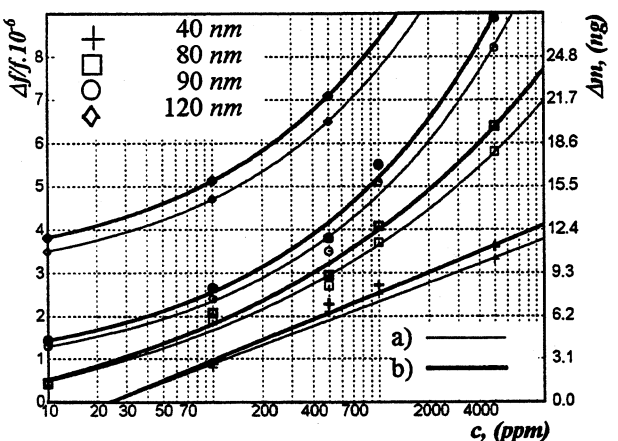


Fig.1. Relative change of resonator's frequency *Δf/f* (a) and sorbed mass *Δm* (b) versus amonia concentration *c (ppm)* at different thickness of SnO_2

The provided dependence can be described by equation like $y = Ax^{\alpha} + B$, where coefficients A, α and B change their values for different curves. Calculated Δm from equation 1 as a function of ammonia concentration is shown in Fig.1(b). The dependence of Δm from NH_3 concentration shows the same tendency with those at Fig.1(a). It can be explained by the equation (1), which gives the relationship between Δf and Δm. It is clear, that by increasing NH_3 concentration the quantity of sorbed mass increases too. It may be supposed, that equilibrium between sorbed and desorbed molecules is reached at higher quantity of the sorbed mass, when NH_3 concentration rises. Figure 2 illustrates the depen-

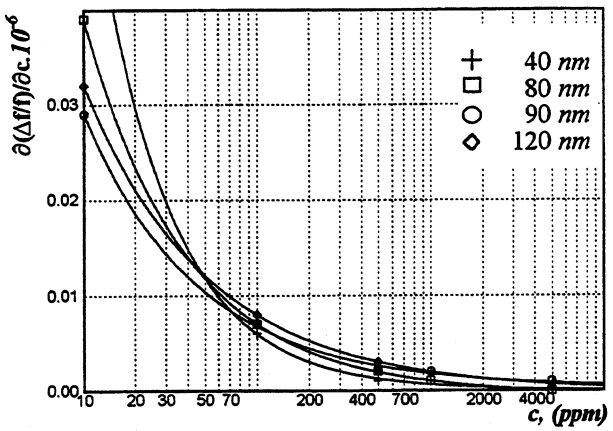


Fig.2. Mass sensitivity *∂(Δf/f)/∂c* versus amonia concentration *c(ppm)*

dence of mass-sensitivity (amonia-sensitivity) ∂(Δf/f)/∂c of resonator's structure at different NH_3 concentrations. All curves show one and the same character not dependent on of SnO_2 thickness. The changing of mass-sensitivity reaches constant values by high concentrations in range of 5000 ppm. The highest values of ∂(Δf/f)/∂c by a constant NH_3 concentration appear at the thinest SnO_2 films. By concentration of NH_3 – 10 ppm mass-sensitivity of resonator's structure with SnO_2 thickness – 120*nm* equals to $0.032.10^{-6}$, whereas for thickness of 80*nm* it is $0.038.10^{-6}$ or almost twice as high.

Figure 3 illustrates the correlation between resonator's frequency relative change and the thickness of the tin oxide by different ammonia concentrations. The thickness of tin oxide films is changed from 40*nm* to 120*nm*. This range of thicknesses is chosen for investigation as a result of experiments published in [10]. It is shown that the relative change of the resonator's frequency increases with the increase of tin oxide thickness in all investigated interval by constant ammonia concentration. For resonators with thickness of SnO_2 - 40*nm* the alteration of Δf/f is $0{,}8.10^{-6}$, whereas by the thickness of SnO_2 - 120*nm* Δf/f reaches $4{,}7.10^{-6}$ by

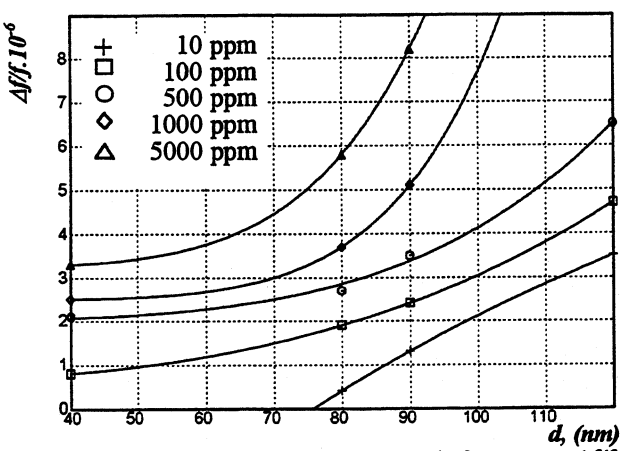


Fig.3. Relative change of resonator's frequency *Δf/f* versus the tin oxide thickness *d* at different concentration g NH_3 c_{NH3}

constant ammonia concentration 100 ppm. The dependence of the relative change of the resonator's frequency versus the thickness of thin tin dioxide film shows one and the same character of gradually increasing at low ammonia concentrations in range from 10 ppm to 500 ppm. At higher ammonia concentrations (1000 ppm and 5000 ppm) the dependence is most strongly expressed. Probably one of the reasons is the penetration of NH_3 molecules in the depth of the layer as a result of higher concentration gradient. These dependence of $\Delta f/f$ as a function of the thickness of SnO_2 film also are decribed as grade function. On the base of the obtained dependence it can be supposed, that simultaneously with the physical adsorption on the surface of the sensitive tin oxide layer, adsorption takes place also in the depth of the layer.

Fugure 4 shows mass-sensitivity of the resonator structures as a function of the sensitive layer's thickness by different ammonia concentrations. Obviously, that dependence by C_{NH3} = 10 ppm is different in character from those in the interval from C_{NH3} - 100 ppm to C_{NH3} - 5000 ppm. In the former case the increase of the thickness does not lead to an increase in sensitivity. It may be supposed, that by those low concentrations, the quantity of adsorbed substance is determined by surface active centers. In this case between the adsorbed NH_3 and the one in the gas phase is established dynamic equilibrium for the determined experimental conditions (temperature, NH_3 - concentrations).

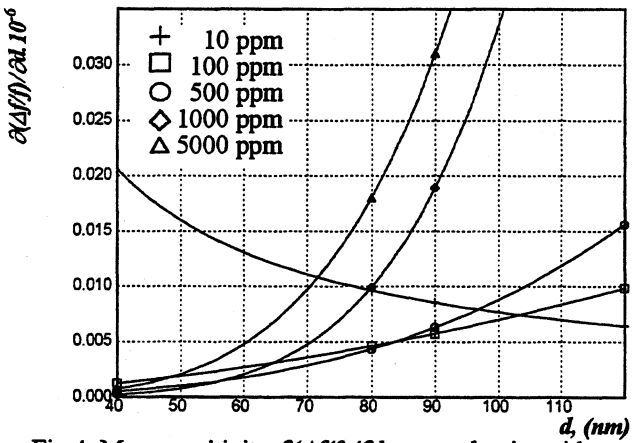


Fig.4. Mass sensitivity $\partial(\Delta f/f)/\partial d$ versus the tin oxide thickness d

Increasing of the SnO_2 thickness results in lowering the mass-sensitivity almost three times from $1{,}92.10^{-8}$ at d_{SnO2} = 40*nm* to $0{,}64.10^{-8}$ at d_{SnO2}=120*nm* by a constant ammonia concentration - 10 ppm. Just the opposite character is this dependence at all other higher concentrations. Most strongly is this dependence expressed at NH_3 concentrations of 1000 ppm and 5000 ppm. These results confirm the supposition about existance of NH_3 diffusion in the depth of the sensitive SnO_2 layers.

4. Conclusion

Relationship is obtained between the thickness of the thin tin dioxide film and ammonia concentration on the one hand and mass sensitivity of quartz resonator at NH_3 detection on the other hand. It is found out, that the structure quartz resonator with thin SnO_2 layer can be used with success for NH_3 detection. This system could serve as a base for development of acoustic sensor for monitoring NH_3 in the environment.

5. References

[1] G. Sberveglieri, "Recent developments in semiconducting thin film gas sensors", Sensors and Actuators, B 23, pp. 103-109, 1995.

[2] N. Yamazoe, Y. Shimizu, "Humidity sensors: principles and applications", Sensors and Actuators, vol. 10, pp. 379-398, 1986.

[3] P. Bergveld, "The impact of MOSFET - based sensors", Sensors and Actuators, vol. 8, pp. 109-127, 1985.

[4] W. Gopel, K. D. Schierbaum, "SnO_2 sensors: current status and future prospects" , Sensors and Actuators, B 26-27, pp. 1-12, 1995.

[5] S. K. Andreev, L. I. Popova, V. K. Georgiev, G. D. Beshkov, "Characteristics and gas-sensing behaviuor of a tin-oxide-gate FET", Sensors and Actuators, B 8, pp. 89-91, 1992.

[6] J. Robertson, "Defect levels of SnO_2", Physical Review B, vol.30, no. 6, pp. 3520-3522, 1984.

[7] L. M. Dorojkine, V. V. Volkov, V. S. Doroshenko, A. A. Lavrenov, D. A. Mourashov, I. A. Rozanov, "Thin film piezoelectric acoustic sensors. Application to the detection of hydrocarbons", Sensors and Actuators, B 44, pp. 488-494, 1997.

[8] V. Lazarova, L. Spassov, V. Georgiev, S. Andreev, E. Manolov, L. Popova, "Quatrz resonator with SnO_2 thin film as acoustic gas-sensor for NH_3", Vacuum, vol.47, no. 12, pp. 1423-1425, 1996.

[9] V. V. Malov, Piezoresonancnie datchiki. Moskva: Energoizdat, 1989, p.68 (in Russian).

[10] L. Spassov, V. Lazarova, E. Manolov, "Mass-sensitive quartz resonator with SnO_2 thin film as ammonia gas-sensor", in Proceedings of 4th International Symposium on Surface Waves in Solid and Layered Structures (ISSWAS-4), June 7-12, 1998, pp. 98-102 (in press).

1999 Joint Meeting EFTF - IEEE IFCS

AUTOMATIC ULTRASONIC MEASUREMENT OF SOUND VELOCITY IN LIQUIDS BY A PHASE-LOCKED LOOP METHOD

Kiyoshi Ikeda

Faculty of Engineering, Shinshu University,
500 Wakasato, Nagano 380-8553, Japan

Abstract

This paper describes the workings of a new ultrasonic technique for measuring sound velocity in liquids which relies on a phase-locked loop method to automatically eliminate dependence on temperature. The sound velocity v is easily found from the Main VCO frequency f multiplied with the coefficient k as the ratio of a known sound velocity v_S and its frequency f_S. Validity and usefulness for the system was verified experimentally using NaCl solutions as test-liquid specimens. The compensated frequency f_C is gained from the output of the compensatory VCO (Comp VCO) which is automatically eliminates the strong temperature dependence of sound velocity in liquids. The Comp VCO is constructed identically to the Main VCO. The sound velocity for eliminating the temperature dependence was easily obtained with $v = v_S f_C / f_S$. The system with compensation circuit was applied to NaCl solutions of varying concentrations and was shown to yield 0.0005% accurate velocity determination, independent of temperature variations of ± 1.3°C at approximately 24.9°C.

1 Introduction

The ultrasonic technique used for measuring sound velocity in liquids has become very important, because automatic measuring systems are currently being used to advance research into physical and chemical properties and in instruments used for industrial measurements [1], [2]. For this purpose, various ultrasonic techniques using continuous waves, pulse waves, an ultrasonic interferometer, a correlation method, ultrasonic resonance and a reverberation method have been conventionally developed [3]. There are a variety of ways to take extremely accurate measurements of sound velocity using the pulse method, but some of these are suitable only for research purposes [4]. Recently, the high-precision pulse-echo method [5] has been reported. This method automatically measures sound velocity in liquids and can be used to obtain detailed structural information about the liquid medium. However, it requires a temperature control of ± 0.5mK with a thermostat system and an extremely complicated digital electronic system using a computer in order to reach a precision of 2×10^{-5} for sound velocity measurements. This system, which is constructed with an extremely accurate thermostat and a computer, requires a complicated circuit and high-cost equipment. In a previous study we proposed an innovative ultrasonic technique for measuring sound velocity in liquids [6] that automatically eliminates temperature dependence without requiring the usual thermostat system, the technique relies on a phase difference method, and is constructed with an elaborate circuit. Here we propose an original ultrasonic technique that relies on a phase-locked loop method. The strong temperature dependence of sound velocity in liquids is automatically eliminated via a temperature-compensation circuit constructed with a Comp VCO measuring system. Therefore, the Comp VCO is very important for ultrasonic velocity measurement systems. We derived an empirical equation for NaCl soutions which is approximated over a concentration range of 0.0 – 1.0% and a temperature range of 15 – 45°C while analyzing various sound velocities by the method of least squares. A temperature compensation factor was derived by analyzing the Comp VCO and using the empirical equation. The proposed technique has a very simple circuit system which exclusively treats continuous sinusoidal waves without relying on modulated waves and is comprised of analog circuits only. The obtained temperature range is 1.3 times wider than that obtained with the phase difference method. We report here on the principle of measurement, experimental results, the frequency of the Comp VCO and the temperature elimination characteristics.

2 Principle of Measurement

Figure 1 shows a block diagram on the experimental measurement system considered in this study. The measurement system consists of an ordinary acoustic cell (AC), enclosed with a dotted line, and an electronic circuit enclosed with solid line. The phase-locked loop is constructed with a Main VCO (Main Voltage-Controlled Oscillator), AC, PD (Phase De-

0-7803-5400-1/99/ $ 10.00

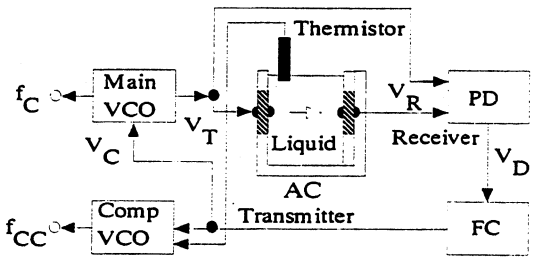


Figure 1: Block diagram of sound velocity measurements in liquids which automatically cancels the dependence on temperature by a phase-locked loop method.

tector) and FC (Feedback Circuit) which consists of an integrator circuit and comparator circuit. The frequency of the Main VCO f_C is adjusted until V_D equals the reference DC voltage V_{ref} on the comparator circuit: the feedback loop locks the phase difference between V_T and V_R at a constant value θ_l. If the ultrasonic velocity of the liquid is a function of concentration s and temperature t and is expressed as $v = F(s,t)$, the phase-locked frequency f_C fixed at a certain point is given by

$$f_C = K_B F(s,t)/2\pi l. \quad (1)$$

where l is the length of the cell, $K_B = \sin^{-1}(V_{Dl}/K_A) - K$, $V_{Dl} = V_{ref} = K_A \sin(\theta_l + K)$, K_A is the overall gain of the phase detection and K is a constant. The term V_{Dl} is the output of the phase detector when θ is locked at θ_l. Replacing $F(s,t)$ with v and solving this equation with respect to v, we obtain

$$v = \frac{2\pi l f_C}{K_B} = K_C f_C \quad (2)$$

where $K_C = 2\pi l/K_B$. Provided that both phase difference θ and l are fixed, K_C is reduced to a constant value and the sound velocity v is easily found from the VCO frequency f_C multiplied by the coefficient K_C. We obtain the velocity of sound in liquids by determining the value of f_C and calculating Eq. (2). If the liquid temperature varies from t_1 to t, v_{t1}, the sound velocity at t_1 changes to v_t at t and is given by

$$v_t = \frac{f_{Ct} v_{t1}}{f_{Ct1}} = K_{Ct} f_{Ct}, \quad (3)$$

where $K_{Ct} = v_{t1}/f_{Ct1}$, f_{Ct1} and f_{Ct} are, respectively, f_C at t_1 and t. The sound velocity v_t is easily found from the VCO frequency f_{Ct} multiplied by the coefficient K_{Ct} as the ratio of a known sound velocity v_{t1} and its frequency f_{Ct1}. Therefore if v_{t1} at t_1 is known, we can obtain v_t at t from Eq. (3).

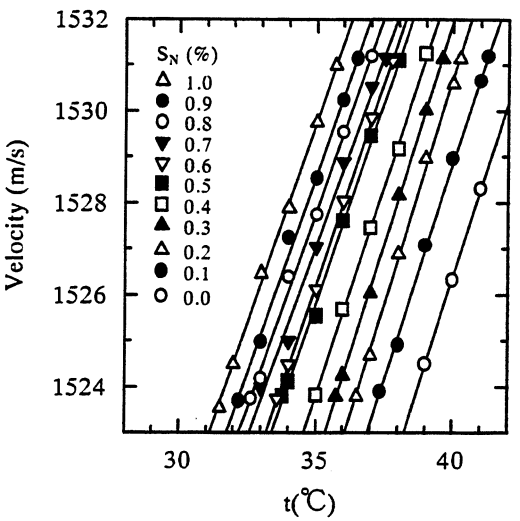


Figure 2: Sound velocity v versus an NaCl solution temperature range of 30.5 – 41.5°C for a concentration range of 0.0 – 1.0%.

3 Experimental Results

The container for the AC was an acrylic tank 50 mm square and 75 mm in height. We determined v_S using an ultrasonic wavelength obtained by varying the distance between the transmitter and receiver of the AC. We calculated a value for v_S of 1.505 $\times 10^3$ m/s at f_S = 2.150 MHz in a 0.5% NaCl solution and liquid temperature of t_S = 25.0°C. Figure 2 shows sound velocity by solution temperature, obtained for l = 42mm, V_{ref} = 3.0V and different concentrations ranging from 0.0 – 1.0%. The results clearly show that observed sound velocity increases in proportion to solution temperature. This figure indicates an increase of 1°C and causes an increase of 1.75 m/s in sound velocity over a temperature range of 31.5 – 35.7 °C at a 1.0% concentration. Therefore, sound velocity cannot be measured accurately either under constant solution temperature conditions or by eliminating temperature dependence via a compensation technique. The compensated frequency f_C is gained from the output of the Comp VCO which is constructed as a part of the temperature-compensation circuit.

4 Empirical Equation of Sound Velocity

Generally the ultrasonic velocity of the solution is a function of concentration s and temperature t and expressed as $v = F(s, t)$. The Taylor expansion of F(s, t) around reference concentration s_0 and refer-

ence temperature t_0 is given by

$$v = F(t_0 + \Delta t,\ s_0 + \Delta s)$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \left(\Delta t \frac{\partial}{\partial t} + \Delta s \frac{\partial}{\partial s} \right)^n F(t_0,\ s_0) \qquad (4)$$

If we express v approximately by the first three terms of series expansion, Eq. (4) becomes

$$v = V_k + V_{t1} t + V_{t2} t^2 + V_{t1s1} ts + V_{s1} s + V_{s2} s^2 \quad (5)$$

where,
$V_k = F(t_0,\ s_0) - \alpha t_0 - \beta s_0 + \frac{\gamma t_0^2}{2} + \delta t_0 s_0 + \frac{\eta s_0^2}{2}$
$V_{t1} = \alpha - \beta t_0 - \delta s_0$, $V_{t2} = \frac{\gamma}{2}$, $V_{t1s1} = \delta$,
$V_{s1} = \beta - \delta t_0 - \eta s_0$, $V_{s2} = \frac{\eta}{2}$,

$$\alpha = \left(\frac{\partial F}{\partial t} \right)_{t_0,s_0}, \beta = \left(\frac{\partial F}{\partial s} \right)_{t_0,s_0}, \gamma = \left(\frac{\partial^2 F}{\partial t^2} \right)_{t_0,s_0},$$

$$\delta = \left(\frac{\partial^2 F}{\partial t \partial s} \right)_{t_0,s_0}, \eta = \left(\frac{\partial^2 F}{\partial s^2} \right)_{t_0,s_0}$$

The proportional coefficient $\frac{\partial F}{\partial s}$ for thinly concentrated solutions is independent of solution temperature, and the sound velocity in solutions of a given temperature and concentration is almost a linear function. Consequently, an empirical equation for NaCl solutions which is approximated by the method of least squares over a concentration range of 0.0 – 1.0% and a temperature range of 15 – 45 °C is given by

$$v = 1418 + 4.05t - 0.03254t^2 + 11.81 s_N \qquad (6)$$

5 Temperature Compensation Factor

We derive a temperature compensation factor for designing the Comp VCO by considering the empirical Eq.(6) as an equation describing the temperature-dependence of sound velocity in NaCl solutions. The compensated frequency f_{CC} is obtained from the output of the Comp VCO which is constructed as part of the temperature compensation circuit. If the solution temperatures studied varies from t_1 to t, f_C undergoes a change Δf_C, and the temperature compensation factor F_A – which is an extremely important specific value for designing the Comp VCO – is defined by

$$F_A = \Delta f_C/(t - t_1) = K_B(V_t - V_{t1})/2\pi l(t - t_1). \quad (7)$$

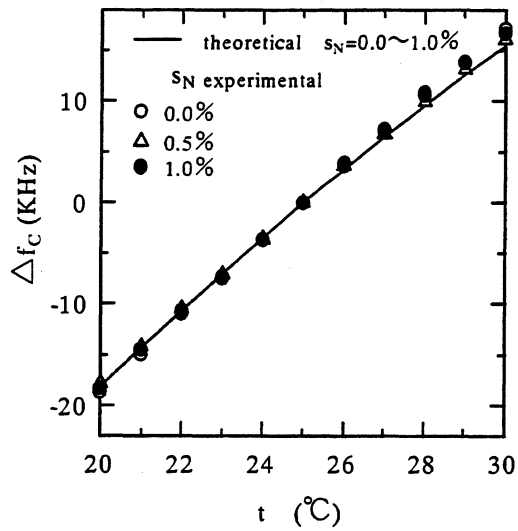


Figure 3: Change in phase-locked frequency observed over a temperature range of 20 – 30°C. Concentration of NaCl solutions is s_N = 0.0, 0.5 and 1.0%.

where $V_t = 4.05t - 0.03254t^2$ and $V_{t1} = 4.05t_1 - 0.03254t_1^2$. Figure 3 shows a graph of Δf_C observed around 25°C for solutions with s_N = 0.0 ,0.5 and 1.0%, together with theoretical values. These results indicate that the same value of F_A can be used over a concentration range of 0.0 – 1.0% and temperature range of 23 – 27°C. The gradient in experimental values for these curves gives a specified value F_A = 3.57 kHz/°C, which is reasonably consistent with F_A = 3.38 kHz/°C calculated with Eq. (7). These experimental and theoretical results strongly suggest the possibility of temperature compensation for F_A.

6 Frequency of the Comp VCO

Provided the Comp VCO is constructed identically to the Main VCO, i.e. from relationship $K_{HV} = K_{HC}$, a frequency with compensation f_{CC} is given by

$$f_{CC} = 11.81 K_{CI} s_N + K_{CI} V_K + \frac{V_{RR} + kV_R + V_{KK}}{K_{HC}} \qquad (8)$$

where V_K = 1418 and K_{HV} represents the relationship between control voltage and oscillation frequency of the Main VCO; K_{HC} represents the relationship between control voltage and oscillation frequency of the Comp VCO; $K_{CI} = \frac{1}{K_C} = K_B/2\pi l$. $V_{Kt} = V_K + V_t$. $V_{KK} = K_{HV} K_{CI} V_t - kR_f V_r \{a(R_{Sf} + b_i R_O) - b_i R_a\}/R_a(R_S + b_i R_O)$; R_{Ti} is the resistance value of the thermistor at liquid temperature $t = t_i$°C;
$a = \frac{R_A}{R_A + R_B}$, $R_a = R_S + R_f + aR_A$, $R_{Sf} = R_S + R_f$;

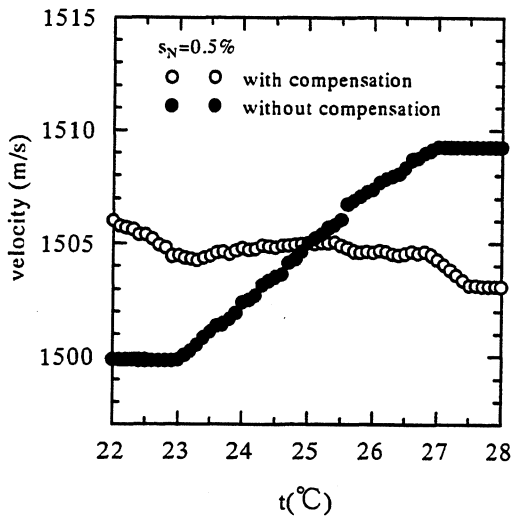


Figure 4: Sound velocity with (open circles) and without (closed circles) compensation circuit observed over a range of 22 – 28°C. Sample is a 0.5% NaCl solution.

$b_i = \frac{R_{Ti}}{R_{Ti}+R_C}$, $k = R_E/R_D$;
R_A, R_B, R_S, R_f, R_D, R_E, R_O, are the resistance values of the resistor which comprises the voltage-linearized circuit of Comp VCO; V_r is the reference voltage in the voltage-linearized circuit; and V_R and V_{RR} are the voltages for the level-converting circuit.

Substituting Eq.(8) for Eq.(2), we obtain

$$v_C = 11.81 s_N + V_K + V_{OO} \qquad (9)$$

where $V_{OO} = K_C \frac{V_{RR}+kV_R+V_{KK}}{K_{HC}}$. Because V_{OO} is independent of solution temperature, v_C is a constant value without the influence of temperature.

7 Characteristics of Temperature Elimination

The temperature compensation circuit previously described satisfies the conditions of the temperature compensation factor for linear compensation of temperature, that is, $\Delta t = 4$°C. Moreover, it was confirmed that the designed circuit is valid and useful as a temperature compensation circuit. We added it to our measuring system for sound velocity and measured f_C in 0.5% NaCl solution while varying the temperature from 22 to 28 °C to test the operation of the compensation circuit. The sound velocity v_t is obtained using Eq.(3) with $f = 2.150$ MHz, $l = 42$ mm, and $v_{t1} = 1.505\times10^3$ m/s at $t_1 = 25$°C. The open and closed circles in Fig. 4 represent the results with and without the compensation circuit, respectively. It is clear from the results with compensation that sound velocity shows the least dependence on temperature over this range. The effect of temperature variation on the oberved sound velocity is safely compensated for over a temperature range of 24.9°C ± 1.3°C.

8 Conclusions

We developed an innovative ultrasonic technique for measuring sound velocity in liquids which automatically eliminates dependence on temperature by a phase-locked loop method. We analyzed the measured values of sound velocity in NaCl solutions as liquid specimens and sought an empirical equation for NaCl solutions approximated over a concentration range of 0.0 – 1.0% and a temperature range of 15 – 45°C. The Comp VCO is designed using the temperature compensation factor derived with the empirical equation of sound velocity in NaCl solutions. A Comp VCO with a specific value of 3.38 KHz/°C over a range of 25°C ± 2.0°C was degined. This value strongly agreed with the experimental value of 3.36 KHz/°C. The system with the Comp VCO was applied to NaCl solutions of varying concentrations and was shown to yield 0.0005% accurate velocity determination, independent of temperature variations of ±1.3°C at approximately 24.9°C.

References

[1] L. C. Lynnworth, Ultrasonic Measurements for Process Control. New York: Academic Press, 1989.

[2] Edited by Erika Kress-Rogers , Instrumentation and Sensors for the Food Industry. Oxford, Butterworth-Heinemann Ltd, 1993.

[3] K. Negishi and K. Takagi,Ultrasonic Technique. Tokyo: Tokyo Daigaku Syuppankai, 1987.

[4] R. N. Thurston and A. D.Pierce, Ultrasonic Measurement Methods. New York: Academic Press, 1990.

[5] G. Tardajos, G. G. Gaitano and F. R. Montero de Espinosa, "Accurate, sensitive, and fully automatic method to measure sound velocity and attenuation," Rev. Sci. Instrum, vol.65, pp.2933-2938, 1994.

[6] K. Ikeda, "Ultrasonic Measurement of Sound Velocity in Liquids which Automatically Eliminates Temperature Dependence," presented at the IEEE Conference on Ultrasonic Symposium, Sendai, Japan, October 5-8,1998.

Total Variance Explained*

D. A. Howe†
National Institute of Standards and Technology, 325 Broadway, Boulder, CO 80303

1. Abstract

I explain the difference between the Total variance and the Allan variance and what is gained for estimating frequency stability especially at long term. I also describe the property that adding up Total variance values in the usual "power-of-2" increments yields twice the sample standard variance.

2. Introduction and Summary

Total variance uses all of the available time-difference data $\{x_n\}$ like the standard variance in a data run of duration T (or length N_x of sample x_n-values), but unlike the Allan variance which uses only three (first, median, and last values). I describe a model of the Allan variance as a measure of asymmetry of first and last points with respect to the median by the misalignment of this triplet over T. The concept of Total variance involves scanning the entire interval for all such asymmetries about the median value and averaging them. We still obtain a convergent τ-domain variance for all typical oscillator FM power-law noise types (white denoted as WHFM, flicker as FLFM, and random-walk as RWFM). Remarkably, however, the equivalent degrees of freedom (or edf) shows an increase from 1 to 1.5, 2.1, and 3 in the presence of RWFM, FLFM, and WHFM respectively for computations at $\tau = T/2$ thus supporting the basic model criteria. Because of its efficient use of available data, Total variance is recommended for long-term measurements of frequency stability [1]. A particularly simple and recommended computation of frequency stability dubbed "Totvar" has demonstrated improved estimation of frequency stability at long-term τ-values while essentially computing the usual max-overlap sample Allan variance at short- and mid-term τ-values.

3. Measuring Frequency Stability

If the time or the time fluctuations between two oscillators can be measured directly, an advantage is obtained over just measuring frequency fluctuations. The reason is that we can readily see time behavior from actual measurements, and frequency can easily be inferred from time. To avoid measurement-system dead time and simultaneously measure the underlying frequency stability of the best oscillators often nearly at the same frequency, we use the dual mixer time difference (DTMD) scheme [2]. Measurement samples of time fluctuations occur at a rate f_s having an interval $\tau_0 = \frac{1}{f_s}$. Given a sequence of time deviates $\{x_n : n = 1, \ldots, N_x\}$ with a sampling period between adjacent observations given by τ_0, we define the $m\tau_0$-average fractional frequency deviate as

$$\overline{y}_n(m) \equiv \tfrac{1}{m} \textstyle\sum_{j=0}^{m-1} y_{n-j},$$

where $y_n = \frac{1}{\tau_0}(x_n - x_{n-1})$ and we regard $\{\overline{y}_n(m) : n = m, \ldots, N_y\}$ as a finite realization of a stochastic process $\{\overline{Y}_n(m) : n = 0, \pm 1, \pm 2, \ldots\}$. Allan in ref. [3] devised a characterization of frequency stability based on an ensemble average of a 2-sample standard deviation. The Allan variance is defined as [4]

$$\sigma_y^2(m) \equiv \frac{1}{2} E\left\{\left[\overline{y}_{n+m}(m) - \overline{y}_n(m)\right]^2\right\},$$

$$\text{and } E\left\{\left[\overline{y}_{n+m}(m) - \overline{y}_n(m)\right]\right\} = \mathrm{Dr}(m),$$

where E throughout this paper means an expected average or infinite mean, and $\mathrm{Dr}(m)$ is a linear trend and assumed to be linear frequency drift which is usually estimated and removed. If the first difference $\{\overline{y}_n(1) - \overline{y}_{n-1}(1)\}$ is stationary, then the stochastic process is such that the expectations above depend on the averaging time index m but not on the time index n. Note that each point estimate of the Allan variance computed at m requires a $2m$ interval. A hat " $\hat{}$ " denotes a sample estimate of the function.

The usual max-overlap sample Allan variance $\hat{\sigma}_y^2(\tau, T)$ involves averaging time $\tau = m\tau_0$ and sample data run T. Called Avar, it is given by [2,5]

$$\hat{\sigma}_y^2(\tau, T) = \mathrm{Avar}\,(m, \tau_0, N_{x,y}) =$$

$$\frac{1}{2(N_y-(2m-1))} \textstyle\sum_{n=m}^{N_y-m} \left(\overline{y}_{n+m}(m) - \overline{y}_n(m)\right)^2 = \quad (1)$$

$$\frac{1}{2(m\tau_0)^2(N_x-2m)} \sum_{n=m+1}^{N_x-m} \left(x_{n+m} - 2x_n + x_{n-m}\right)^2, \quad (2)$$

*Contribution of the U. S. Government, not subject to copyright.

†E-mail: dhowe@nist.gov

for $1 \leq m \leq \frac{N_x-1}{2}$. Known as a central difference, summand terms in (2) involve a second difference of $\{x_n\}$ expressed symmetrically over a $2m\tau_0$ span. Each central difference consists of only a first, middle, and last x_n value, or x_n's taken in triplet, then subsequently squared and averaged. Having removed drift and other deterministic error sources, an oscillator's random FM noise will have triplets which *on average* fall on a straight line. Thus, any particular triplet is more likely to fall nearly on a straight line, hence the central difference is often uncharacteristically low as judged by the rest of the x_n values in the $2m\tau_0$ interval.

4. Motivating Concept of Total Variance

Total variance [6–10], denoted as $\sigma^2_{total}(\tau, T)$, has been developed to exploit the time fluctuations in and around Avar's forementioned triplets on the belief that we are at liberty to choose a range of neighboring quantities which can serve as proper surrogates and average them to obtain a better estimator $\hat{\sigma}^2_y(\tau, T)$. These "proper surrogates" originate in the fact that Avar measures only a *symmetry*, or a lack thereof, in equispaced triplet values of $x_{n\pm m}$. Total applies a *consistency hypothesis* in addition to measures of symmetry. "Consistency" means that *averages* of certain individual estimates can equally serve as any other individual estimate. Since oscillator and clock designers seek persistent frequency, it follows that estimates of frequency and frequency stability can be derived in a variety of ways which take advantage of consistency within statistical variability.

For example, the form of Avar in (1) originated because values of $\{\overline{y}_n(m)\}$ are actually measured asymetrically with respect to $\{x_n\}$ values, that is, *post facto* in which $\overline{y}_n(m) = \frac{1}{m\tau_0}(x_n - x_{n-m})$. For symmetry however, we can substitute

$$\overline{y}^{\circ}_n(m) = \tfrac{1}{m\tau_0}\left(x_{n+\frac{m}{2}} - x_{n-\frac{m}{2}}\right). \qquad (3)$$

By a linear interpolation, $\overline{y}^{\circ}_n(m) =$ average of:

$$\tfrac{1}{m\tau_0}\left(x_{n-m} - x_n\right) \text{ and } \tfrac{1}{m\tau_0}\left(x_n - x_{n+m}\right),$$

whose result is $\frac{1}{2m\tau_0}\left(x_{n+m} - x_{n-m}\right)$, or equivalently $\overline{y}^{\circ}_n(2m)$. In other words, we can use $\overline{y}^{\circ}_n(m)$ defined symmetrically in (3) in place of $\overline{y}^{\circ}_n(2m)$. $\overline{y}^{\circ}_n(2m)$ can in turn be used as a "surrogate value" of $\overline{y}_n(2m)$ in addition to the usual asymmetrically-computed values of $\overline{y}_n(2m)$ in (1) or (2) for the Allan variance estimate $\hat{\sigma}^2_y(2m)$. Likewise, $\overline{y}^{\circ}_n(2m)$ can serve in place of $\overline{y}_n(4m)$. We can extend this idea to computations at $\overline{y}^{\circ}_n(4m)$, $\overline{y}^{\circ}_n(8m)$, and so forth, and find by this pyramid method a greater number of terms that can serve as quantities in estimation (1) while still maintaining its basic properties. Repeating this process reveals a pattern which, when combined with either (1) or (2), serves as a motivating idea for a new statistic called "Total variance."

5. Implementing the Concept

Using a symmetry argument, we can define a useful expectation which is independent of values of averaging time index m and the time index n. Redefine $\overline{y}_n(m)$ as centered at n by

$$\overline{y}_n(m) \equiv \tfrac{1}{m}\left(\textstyle\sum_{j=0}^{\frac{m}{2}-1} y_{n-j} + \sum_{j=0}^{\frac{m}{2}-1} y_{n+j}\right), \qquad (4)$$

or in terms of $\{x_n\}$ values, $\overline{y}_n(m) = \frac{1}{m\tau_0}(x_{n-\frac{m}{2}} - x_{n+\frac{m}{2}})$, for m-even. Random FM noise processes also center before and after n such that,

$$E\left\{\overline{y}_{n+\frac{m}{2}}(m) - \overline{y}_n(m)\right\} = E\left\{\overline{y}_{n-\frac{m}{2}}(m) - \overline{y}_n(m)\right\}.$$

Therefore

$$E\left\{\overline{y}_{n+\frac{m}{2}}(m) - \overline{y}_{n-\frac{m}{2}}(m)\right\} = 0,$$

and we can derive the symmetric form of the Allan variance as

$$\sigma^2_y(m) \equiv \frac{1}{2}E\left\{\left[\overline{y}_{n+\frac{m}{2}}(m) - \overline{y}_{n-\frac{m}{2}}(m)\right]^2\right\}. \qquad (5)$$

Substitute s for m in (4), and define (5) in terms of $\overline{y}_n(s)$ to obtain

$$\sigma^2_{y^\circ}(m, s) \equiv \frac{1}{2}E\left\{\left[\overline{y}_{n+\frac{m}{2}}(s) - \overline{y}_{n-\frac{m}{2}}(s)\right]^2\right\}. \qquad (6)$$

The separation between any two $\overline{y}$ samples in (6) is still m as in (5), but now we are in a position to positively and negatively vary or symmetrically "modulate" the averaging-time s of each $\overline{y}$ in the neighborhood of its usual value m by a small range, say, $\pm\delta$. This has the effect of smoothing $\sigma^2_y(m)$ by using interpolated or "surrogate" values in addition to the usual values of $\overline{y}_{n+\frac{m}{2}}(m) - \overline{y}_{n-\frac{m}{2}}(m)$ as described in the previous section. Finally, we increment n, repeat the process, square differences and average to obtain the value of $\hat{\sigma}^2_y(m)$. Generally speaking,

$$\sigma^2_{total}(m) = \overline{\sigma^2_{y^\circ}(m, s)} = \text{smoothed version of } \sigma^2_y(m).$$

The smoothing operation above picks up additional estimates of $\sigma^2_y(m)$ as m increases, improving the usual max-overlap Allan variance estimate, especially if that estimate is uncharacteristically high or low. Equation (6) and hence $\sigma^2_{total}(m)$ differ from the Allan variance because each term constituting a $2m$ interval yields its result dependent on m and s. This

is because the difference-pair of average frequencies $\overline{y}_{n+\frac{m}{2}}(s) - \overline{y}_{n-\frac{m}{2}}(s)$ may be separated when $s < m$ or overlapped when $s > m$. They are conjoined, or adjacent only when $\delta = 0$ making $s = m$. With WHFM, $\sigma^2_{total}(m)$ is unbiased with respect to $\sigma^2_y(m)$ and is biased negatively with FLFM and RWFM. The bias depends on the depth of modulation δ relative to m but ought to be limited to $\delta = \frac{m}{2}$ or a full range given by no more than m itself. This is so that $\sigma^2_{total}(m)$ is controlled in a reasonable manner over data run of length N_x, an issue discussed next.

Returning to measuring frequency stability from an actual data run, recall that m is a parameter which defines interval $\tau = m\tau_0$. The estimate of $\sigma^2_{total}(\tau)$ for a data duration T involves two issues. The first is that our ability to smooth becomes more and more restrictive as $m \to N_x$ (or equivalently, as $\tau \to T$) because the extent with which we can modulate m and increment n is bounded by the beginning and end points of a fixed-length data run. Second, I have mentioned one way to more fully use the available data. There are other ways from which *ad hoc* manipulation and averaging of $\overline{y}_{n\pm\frac{m}{2}}(s)$ differences can yield improved estimates of the Allan variance. In concept, any could be designated as a "Total" variance if it includes essentially all average frequency differences in a symmetric *totality* over interval 2τ (to be consistent with the Allan variance). One way in particular is simple to implement, has been tested on the expected FM noises and other oscillator error sources, and has an important connection to the classical standard variance explained in the last section. This Total variance evolved from experiments in which the usual max-overlap Avar estimator was applied to periodic extensions of the original data [6].

The most efficient analysis of frequency stability would be to apply forementioned surrogate values where they are needed most, namely at long-term τ-values, beyond $\tau = m\tau_0 = 10\%$ of the data run $T = N_x\tau_0$. For a given τ-value and data run of length T, the number of samples in the standard estimator for the Allan variance is of order $\frac{T}{\tau}$, thus (2) is usually more than adequate for determining noise level at short- to medium-term τ-values. A recommended characterization of frequency stability now includes the use of Total variance for the range at long term $\frac{\tau}{T} \geq 10\%$ [1]. This recommendation resulted from what can be regarded as a "hybrid estimator" which uses the standard Allan estimator for short- and medium-term τ-values and rather conveniently applies more surrogate values in long term until they are all applied at the usual Avar function limit of $\tau = \frac{N_x-1}{2} \cdot \tau_0$, half the data run. Such an estimator, now dubbed "Totvar," was introduced in [7]

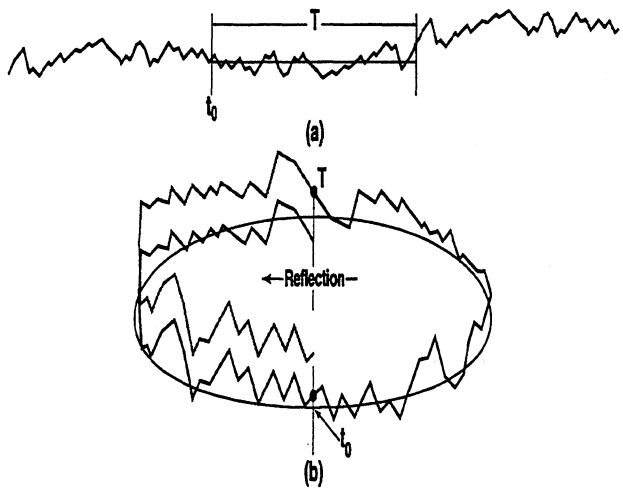


Figure 1: Circular extension of the original $x(t)$ data set for computation of Totvar. (a) Extension of a phase record by "reflection" at both ends; (b) circular representation of extended phase record.

and refined in [9] and [10]. Totvar here is regarded as a composite function of the normal Avar statistic $\hat{\sigma}^2_y(\tau, T)$ and a sample Total statistic $\hat{\sigma}^2_{total}(\tau, T)$ as

$$\text{Totvar}(\tau, T) = f(\hat{\sigma}^2_y(\tau, T), \hat{\sigma}^2_{total}(\tau, T)).$$

Thus Totvar is always a sample variance of a data run of length T. The definition of Totvar is

$$\text{Totvar}(\tau, T) = \text{Totvar}\,(m, \tau_0, N_x) =$$

$$\frac{1}{2(m\tau_0)^2(N_x-2)} \sum_{n=2}^{N_x-1} \left(x^{\#}_{n-m} - 2x^{\#}_n + x^{\#}_{n+m} \right)^2, \quad (7)$$

for $1 \leq m \leq N_x - 1$ where an extended virtual sequence $\{x^{\#}_n\}$ is derived as follows: for $n = 1$ to N_x let $x^{\#}_n = x_n$; for $j = 1$ to $N_x - 2$ let

$$x^{\#}_{1-j} = 2x_1 - x_{1+j}, \quad x^{\#}_{N_x+j} = 2x_{N_x} - x_{N_x-j}. \quad (8)$$

Constructing the extended virtual sequence as in (8) is illustrated in Figure 1 and is called *extension by reflection*. Because of the symmetry of the extended data, the number of summands in (7) does not depend on m. Surrogates values in Totvar emerge from this data extension. Rather than doing extensions of the original vector $\{x_n\}$ and applying the straight second-difference, we alternatively can resample within the original vector. Applied to Totvar, this exercise only points out that the procedure used in the sampling function is intricate and not very intuitive because sampling on $\{x_n\}$ is no longer in terms of equispaced triplets spaced 2τ [9]. For example, note that τ can go to $(N_x - 1)\,\tau_0$ in (7)-(8) instead of the usual limit of $\lfloor (N_x - 1)/2 \rfloor \tau_0$.

Totvar can also be defined in terms of extended normalized frequency averages by

$$\text{Totvar}(\tau, T) = \text{Totvar}\,(m, \tau_0, N_y + 1) =$$

$$\frac{1}{2(N_y-1)} \sum_{n=2}^{N_y} \left[\bar{y}_n^{\#}\,(m) - \bar{y}_{n-m}^{\#}\,(m)\right]^2, \qquad (9)$$

where $\bar{y}_n^{\#}\,(m) = \left(x_{n+m}^{\#} - x_n^{\#}\right) / (m\tau_0)$.

The concept of Total variance was motivated by a simple need for improved long-term estimation of the Allan variance and encouraged by the results using a straight circularization technique on original data $\{x_n\}$, but this technique could not work in the presence of RWFM and/or significant drift [6,11]. Totvar here is classed as a hybrid statistic combining the benefits of sample Allan variance with sample Total variance. Totvar has been tested on the range of FM power-law noise types [4] and other expected oscillator and measurement-system error sources [2,7].

We can illustrate (7)-(8) as a hybrid statistic in the following way. If $m = 1$, the virtual extension $\{x_n^{\#}\}$ only needs to be τ_0 longer than $\{x_n\}$ at both ends to compute (7). Thus (7) is essentially the standard Allan estimator (2). As m increases, the virtual extension needs to be longer until at $m = \frac{N_x-1}{2}$, the extensions at each end are length $\frac{N_x-1}{2}$. Of course, there is no standard Allan estimator in the region $\tau > \frac{T}{2}$, so if the "hybrid" called Totvar in (7) is allowed to compute values for $m > \frac{N_x-1}{2}$, it reverts to a region defined by the Total variance but not the Allan variance. Computations of Totvar should not extend beyond $\tau = \frac{T}{2}$ to be consistent with the limit of the standard Allan estimator, but these higher order terms will be considered in the last section.

For accurately estimating the Allan variance, an adjustment must be made to the hybrid estimator Totvar as defined by (7) and its extension $\{x_n^{\#}\}$ in (8) to remove a normalized bias (denoted as nbias) which depends on the ratio $\frac{\tau}{T}$ and whether the noise type in long-term is FLFM or RWFM rather than WHFM. The most noteable adjustments using Totvar in this manner involve formulae for nbias and increased edf compared to the Allan estimator. These can be summarized as [10]

$$\text{nbias}\,(\tau) = \frac{E\{\text{Totvar}(\tau,T)\}}{\sigma_y^2(\tau)} - 1 = -a\frac{\tau}{T}, \qquad (10)$$

$$\text{edf}\,(\tau) = \text{edf}[\text{Totvar}\,(\tau, T)] \approx b\frac{T}{\tau} - c, \qquad (11)$$

where $0 < \tau \leq \frac{T}{2}$ and a, b, and c are given in Table 1. The values of nbias and edf for the important longest-term case $\tau = T/2$ are tabulated in Table 2. The edf formula (11) is empirical, with an observed error below 1.2% of numerically computed exact values; the tabulated values of edf $(T/2)$ in Table 2 are exact.

Table 1: Coefficients for computing normalized bias and edf of Totvar in the presence of FM noises.

Noise	a	b	c
WHFM	0	3/2	0.000
FLFM	$(3\ln 2)^{-1}$	$24\,(\ln 2)^2\,\pi^{-2}$	0.222
RWFM	3/4	140/151	0.358

Table 2: Tabulated exact quantities for $\tau = T/2$.

Noise	nbias$(T/2)$	edf$(T/2)$
WHFM	0	3.000
FLFM	−0.240	2.097
RWFM	−3/8	1.514

Both Totvar and Avar are invariant to certain manipulations of the vector $\{x_n\}$. The simplest example is that we can reverse and/or invert the sequence $\{x_n\}$ without affecting either's result. Unlike Avar's simple sampling function however, Totvar's many sampling functions between $\tau_0 \leq \tau \leq \frac{T}{2}$ are complicated and can be derived from formulae in ref. [9], but unraveling useful information from them is difficult. It is as informative and easier to look at the frequency-response function associated with Totvar compared to Avar as in Figure 2 for a comparison of the effect of their sampling functions. The dashed curve in Figure 2 is a constant-Q, one-octave passband filter response considered to be ideal for extracting typical power-law noise levels [12–16]. Totvar implements a circular convolution of Avar's frequency response, thus significantly reducing the depth of periodic nulls.

6. Uncertainty of Estimates

Returning to the topic of characterizing noise, the reason for using Totvar is for very efficient extraction of commonly-encountered integer power-law noise types and levels of an oscillator's spectral FM noise. This means greater certainty in the extraction of these parameters and others such as drift and quasi-sinusoidal modulation shown in Figure 3. In retrospect, the formulation of the two-sample Allan ensemble average will contain $N_{\overline{y}(m)}$ average frequencies and only $\frac{N_{\overline{y}(m)}}{m} - 1$ independent intervals with which to do a computation. This represents the actual number of "degrees of freedom". In general the 1σ accuracy of the computation for $N_{\overline{y}(m)}$ sets is simply given by

$$\%\ error,\ Allan\ deviation = \frac{100}{\sqrt{2(\frac{N_{\overline{y}(m)}}{m} - 1)}}. \qquad (12)$$

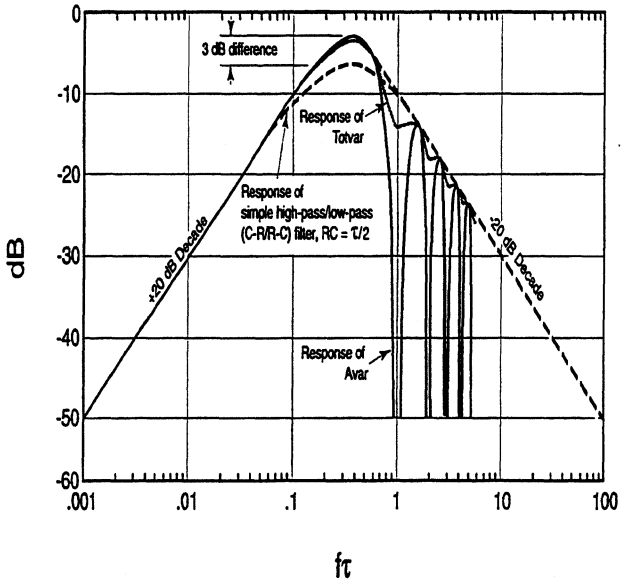


Figure 2: A comparison of frequency responses of Avar (solid curve), Totvar (shaded curve), and a passband variance consisting of a simple cascade of a single-pole high-pass followed by a low-pass filter with identical break points at $RC = \tau/2$ (dashed curve [13]).

This expression is an uncertainty which is adequate for a quick upper-bound approximation for a confidence interval or error bars above and below each value of the Allan deviation *vs.* τ. We assume that the probability distribution is chi-squared, and exact confidence intervals can be determined based on the equivalent degrees of freedom (edf) in overlapping statistical averages for a given noise type, rather than the actual number [5,17].

Sample estimates of the Total variance (given by the Total deviation or root Totvar, convenient plots of interest) have edf's that are greater than even those using the max-overlap estimates of the Allan variance, and significantly greater at long-term τ-values. Chi-squared distribution functions are used for the Allan variance, but it turns out that the distribution functions are slightly narrower using Totvar (which is another of its benefits) at long averaging times. Thus a conservative upper-bound approximation for either the Total deviation or root Totvar 1σ accuracy is

$$\%\ error,\ Total\ deviation = \frac{100}{\sqrt{2(\mathrm{edf})}}, \qquad (13)$$

where edf values are conveniently derived from (11) and Table 1 based on the computation of the Total deviation as a ratio of τ to the length of the data run T, rather than the Allan deviations's $\frac{N_{\overline{y}(m)}}{m}$ independent sets as above.

7. Analysis of Variance

Consider a function of independent variables. In analysis of variance, we explain the total variability of the function in terms of each variable. In the discussion here, we address functionals which depend on a time interval Δt. At this point we can recall a conservation principle regarding the standard sample variance, which states that the mean of the interval variances plus the variance of the interval means equals the standard variance of the entire series. This is true for any process, stationary or not. An analysis of variance in terms of the mean of k interval variances and variance of the k interval means is derived in Appendix I.

The standard variance of finite series $\{X_{ij}\}$ in Appendix I is simply a number, partly due to the variance of k interval means and the remaining part due to the mean of k interval variances. Now consider intervals of duration Δt and a whole data run of length T. The longest possible set of equal-length intervals would be $\Delta t = T/2$, thus there are $k = 2$ consecutive interval means. We recognize that the variance of such two-interval means is the special-case two-sample variance equaling $\frac{1}{2}\hat{\sigma}_y^2(T/2)$, half the sample Allan variance at $\tau = T/2$. But half the sample Allan variance will differ from the standard variance by a remaining portion attributable to the sample variance *within* each of the two intervals by the conservation principle just stated. By double-sampling at $\Delta t = T/4$, we find the two-sample variance ($k = 2$) now must consist of two *nono*verlapped variance estimates whose average, denoted as $\hat{\sigma}^2_{y,\mathrm{nono}}(T/4)$, would be the remaining portion if that were as far as the data could be sampled. Repeating this process until there are no remaining interval variances left unaccounted for, we find that

$$\tfrac{1}{2}\left(\hat{\sigma}^2_{y,\mathrm{nono}}(\tau_0) + \hat{\sigma}^2_{y,\mathrm{nono}}(2\tau_0) + \ldots \right. \qquad (14)$$

$$\left. + \hat{\sigma}^2_{y,\mathrm{nono}}(T/4) + \hat{\sigma}^2_{y,\mathrm{nono}}(T/2)\right) = \hat{\sigma}^2_{std}(T),$$

where $T = m\tau_0, m = 2^j$ for $j = 0, 1, \ldots, J-1$ and the nonoverlapped estimator of the Allan variance is

$$\hat{\sigma}^2_{y,\mathrm{nono}}(2^j) \equiv \frac{2^j}{2N_y} \sum_{k=1}^{\frac{N_y}{2^{j+1}}} \left[\overline{y}_{2k2^j}(2^j) - \overline{y}_{(2k-1)2^j}(2^j)\right]^2.$$

The composite in (14) is a common property of what is called a "multiresolution pyramid" [18]. The nonoverlapped $k = 2$ condition requires that the τ-intervals occur in power-of-two increments. This nonoverlapping sample Allan variance would relate

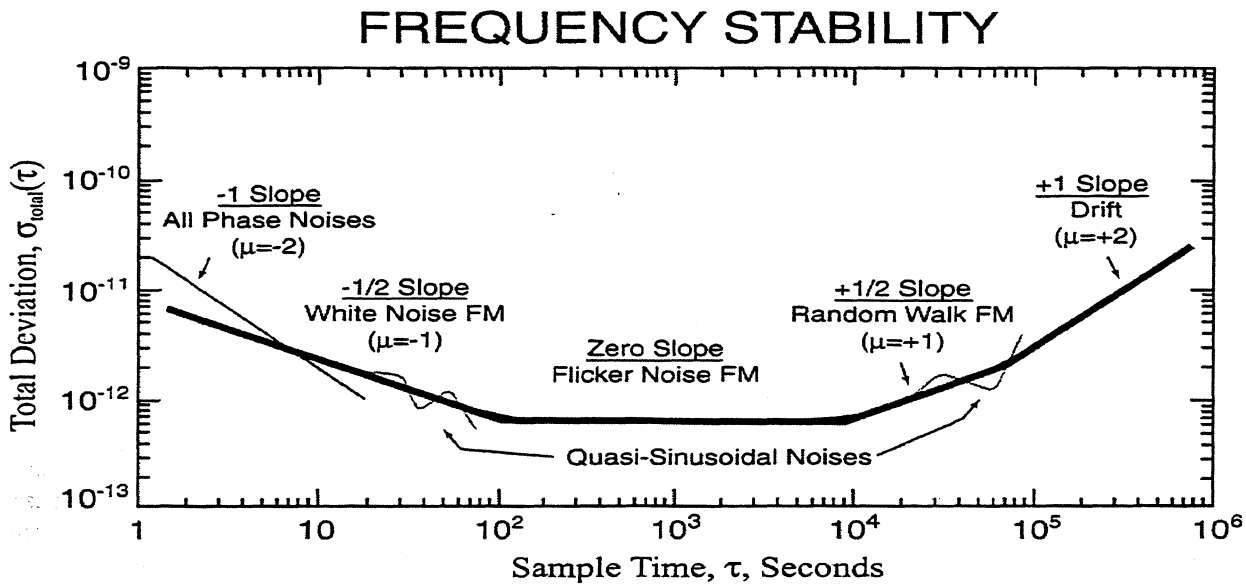


Figure 3: Total deviation plot (or root Totvar) showing power-law noises as straight lines in addition to other error sources. The goal is to identify noise sources and accurately estimate their levels with this kind of frequency stability plot.

directly to the sample standard variance as in (14) but is inefficient as an estimator [15]. Unfortunately the sample Allan variance from its *definition*, for example at $\Delta t = T/4$, calls for *three* variance estimates, not two nonoverlapped, because its definition includes a $\tau = \Delta t$ overlap and the straightforward relationship to the standard variance is lost rather quickly. In otherwords, even for a short series,

$$\frac{1}{2}\left(\hat{\sigma}_y^2(T/4) + \hat{\sigma}_y^2(T/2)\right) \neq \hat{\sigma}_{std}^2(T),$$

in contrast to (14). Since the definition of the Allan variance contains one τ-overlap, we can admit all possible overlaps to obtain an improved estimator in order to minimize its error bars. Known as the standard "max-overlap" Allan estimator [5] given as (2), it also departs from a tractable connection to the standard variance for the same reason as the original τ-overlap estimator.

Functionals which depend on a time interval Δt have such a strong connection to spectral functions that (14) is a "decomposition" of the sample standard variance and seems an appropriate jargon and so is commonly used. In this regard, decomposition of the standard variance is suited to frequency-domain analysis, and Totvar maintains a straightforward relationship with the sample standard variance. It abides by the conservation principle if we consider an *infinite* extension by reflection. This means that the virtual sequence generated by (8) and shown in Figure 1 recurs indefinitely [10]. Percival [19] was the first to point out that for the case in which Totvar is computed in power-of-2 increments above $T/2$ as estimated from data-run T as in (7), a remaining portion, the sum of Totvar terms of all power-of-2 intervals $\tau > T/2$ for $\tau \to \infty$, adds to the usual multiresolution pyramid to precisely equal the standard sample variance. These leftover higher-order components are never actually reported but are an artifact of infinitely extending the orginal sequence. They can be regarded as the sum of 0-frequency aliases, a remaining "D.C." term to make up $\hat{\sigma}_{std}^2(T)$. Greenhall [10] coined the term Remvar($\frac{T}{2}$) to designate this portion. Totvar beyond T soon drops to nearly zero, so the remaining portion above T is generally very small. Nevertheless, Remvar accounts for this portion and was used in the proof of the decomposition of the standard variance. Summing all the familiar "power-of-2" τ-values in a Totvar plot leads to exactly twice the standard sample variance much in the same way that integrating an estimate of a spectrum also yields the sample variance.

Knowledge that we can account for <u>all</u> variations in a data-run by its standard variance as "decomposed" in calculations of the sample Total variance is especially useful. For example at a long-term τ-value of $T/2$, an equal remaining portion (Totvar($\frac{T}{2}$)=Remvar($\frac{T}{2}$)) would indicate that we have summarized completely the variations at $T/2$.

8. References

[1] E.S. Ferre-Pikal, et al., "Draft Revision of IEEE Std 1139-1988: Standard Definitions of Physical Quantities for Fundamental Frequency and Time Metrology—Random Instabilities," *Proc.*

1997 IEEE Int. Freq. Cont. Symp., pp. 338–357 (1998; in revision).

[2] D.B. Sullivan, D.W. Allan, D.A. Howe, and F.L. Walls (Editors), *Characterization of Clocks and Oscillators*, Natl. Inst. Stand. Technol. Technical Note 1337, 1990.

[3] D.W. Allan, "Statistics of Atomic Frequency Standards," *Proc. IEEE* **54**, pp. 221–230, Feb. 1966.

[4] J.A. Barnes, A.R. Chi, L.S. Cutler, D.J. Healy, D.B. Leeson, T.E. McGunigal, J.A. Mullen, Jr., W.L. Smith, R.L. Sydnor, R.F.C Vessot, G.M.R. Winkler, "Characterization of frequency stability," *IEEE Trans. Instrum. Meas.*, **IM-20**, pp. 105–120, 1971.

[5] D. A. Howe, D. W. Allan, and J. A. Barnes, "Properties of Signal Sources and Measurement Methods," *Proc. 35th Annual Symposium on Frequency Control*, pp. 1–47, 1981.

[6] D. A. Howe, "An Extension of the Allan Variance with Increased Confidence at Long Term," *Proc. IEEE International Frequency Control Symposium*, pp. 321–329, 1995.

[7] D. A. Howe and K. J. Lainson, "Effect of Drift on TOTALDEV," *Proc. IEEE International Frequency Control Symposium*, pp. 883–889, 1996.

[8] D. A. Howe, "Methods of Improving the Estimation of Long-term Frequency Variance," *Proc. European Frequency and Time Forum*, pp. 91–99, 1997.

[9] D. A. Howe and C. A. Greenhall, "Total variance: a progress report on a new frequency stability characterization," *Proc. 29th Ann. PTTI Systems and Applications Meeting*, pp. 39–48, 1997.

[10] C. A. Greenhall, D. A. Howe and D. B. Percival, "Total Variance, an Estimator of Long-Term Frequency Stability," *IEEE Trans. Ultrasonics, Ferroelectrics, and Freq. Control*, in process, 1999.

[11] F. Vernotte, 1995. A circularized Allan variance statistic is defined but not invesigated in F. Vernotte, "Stabilité temporelle des oscillateurs: nouvelles variances, leurs propriétés, leurs applications," PhD thesis, Université de Franche-Comte, Besançon, 1991.

[12] J. Rutman, "Characterization of frequency stability: A transfer function approach and its application to measurements via filtering of phase noise," *IEEE Trans. Instrum. Meas.*, vol. IM-23, pp. 40-48, Mar. 1974.

[13] R. G. Wiley, "A Direct Time-Domain Measure of Frequency Stability: The Modified Allan Variance," *IEEE Trans. Instrum. Meas.*, vol. IM-26, pp. 38-41, Mar. 1977.

[14] D.W. Allan, M.A. Weiss, and J.L. Jespersen, "A Frequency-domain View of Time-domain Characterizations of Clocks and Time and Frequency Distribution Systems," *Proc. 45th Freq. Cont. Symp.*, pp. 667–678, 1991.

[15] D. B. Percival, "Characterization of frequency stability: frequency-domain estimation of stability measures," *Proc. IEEE*, vol. 79, pp. 961–972, 1991.

[16] D.A. Howe and D.B. Percival, "Wavelet Variance, Allan Variance, and Leakage," *IEEE Trans. Instrum. Meas.*, **IM-44**, pp. 94–97, 1995.

[17] C.A. Greenhall, "Recipes for Degrees of Freedom of Frequency Stability Estimators," *IEEE Trans. Instrum. Meas.*, vol. **40**, pp. 994–999, 1991.

[18] O. Rioul and M. Vetterli, "Wavelets and Signal Processing," *IEEE SP*, pp. 14–38, October 1991.

[19] D. B. Percival and D. A. Howe, "Total variance as an exact analysis of the sample variance," *Proc. 29th Ann. PTTI Systems and Applications Meeting*, pp. 97–105, 1997.

Appendix I

Consider a series $\{X_{ij}\}$ with k intervals each having n values and means $m_j, j = 1, ..., k$. Assume

$$\sum_1^{nk} X = 0, \sum_{j=1}^{k} m_j = 0 \tag{15}$$

and put

$$X_{ij} = x_{ij} + m_j, \tag{16}$$

so that

$$\sum_{i=1}^{m} x_{ij} = 0. \tag{17}$$

The standard variance of the data run is denoted $V = \frac{SS}{mk}$ where

$$SS = \sum_{j=1}^{k} \sum_{i=1}^{n} (x_{ij} + m_j)^2. \tag{18}$$

$$SS = \sum_{i=1}^{n} x_{i1}^2 + ... + \sum_{i=1}^{n} x_{ik}^2 + nm_1^2 + nm_k^2 \tag{19}$$

plus terms of the form

$$\sum_{i=1}^{n} 2x_{ij}m_j = 2m_j \sum_{i=1}^{n} x_{ij} = 0.$$

$$V = \frac{1}{nk} SS = \frac{1}{k} \left[\sum_{i=1}^{n} \frac{x_{i1}^2}{n} + ... + \sum_{i=1}^{n} \frac{x_{ik}^2}{n} \right]$$

$$+ \frac{m_1^2}{k} + ... + \frac{m_k^2}{k} = \frac{1}{k} \sum_{j=1}^{k} \hat{\varepsilon}_{ss,j}^2(\Delta t) + v_m = \overline{\hat{\varepsilon}_{ss,j}^2(\Delta t)} + v_m \tag{20}$$

where $\overline{\hat{\varepsilon}_{ss,j}^2(\Delta t)}$ is the mean of the interval variance $\hat{\varepsilon}_{ss,j}^2(\Delta t)$, and v_m is the variance of the interval means m_j.

1999 Joint Meeting EFTF - IEEE IFCS

NOISE PREDICTIONS FOR OPTOELECTRONIC OSCILLATORS USING DIFFERENT MODELS.

S. Römisch and A. De Marchi

Dipartimento di Elettronica del Politecnico di Torino
c.so Duca degli Abruzzi, 24 10129 Torino, Italy
e-mail: romisch@boulder.nist.gov

ABSTRACT

In this paper an optoelectronic oscillator (OEO) is analized with two different approaches. The two approaches lead to two models which allow prediction of the oscillator phase noise on the basis of the noise of its components. The first method, called Leeson's method, applies control system theory to an oscillator described by functional blocks. The second, called Van der Pol's method, uses the time domain equation of the oscillator, which is linearized about a equilibrium under the assumption af slow variations of the state variables.

In both cases the predicted noise is compared with the measured noise of a real OEO.

1. INTRODUCTION

The two models derived in this paper allow the analysis of the various noise contributions in the new generation of microwave oscillators called opto-electronic oscillators (OEO) [2], [3].

In 1966 Leeson presented a simple linear model which can be used to predict the effect of amplifier phase noise on oscillator phase noise.[1] This model is based on heuristic considerations of the basic oscillator scheme: a frequency selective element and an amplifier in a closed loop configuration.

The application of control system theory in the frequency domain allows us to derive results that contain those stated by Leeson. Furthermore in this case the number of elements is not limited to an amplifier and a resonator, and the system structure can be more complicated.

The OEO has also been analyzed using a different approach, with the application of the Van der Pol method, which linearizes the time domain equations of the oscillator with the hypothesis of *slow variations* of signal's phase and amplitude about the oscillation condition.

Both models are linear. Otherwise, it would be impossible to describe the system through the transfer function of its functional block (Leeson's method), and it would also be impossible to perform the Fourier transform of the time domain equations (Van der Pol's method).

Therefore these two models are useful to analyze the noise processes associated with the oscillation signal without investigating the non linear dynamics involved in the oscillator.

Noise prediction from both models are compared with the experimental results obtained at NIST on an OEO operating at 10.6 GHz [3].

2. THE NOISELESS OPTOELECTRONIC OSCILLATOR

A key element in all OEOs is the optical fiber that provides the long delay which will determine the oscillation frequency. A second selective element chooses among the large number of possible oscillator modes created by the fiber. The oscillator also contains a laser, a detector, and an electro-optic amplitude modulator (EOM).

After optical detection, the RF signal is amplified and fed back to the modulator, thus closing the loop. The basic scheme is shown in Fig. 1.

In our case the EOM is a Mach-Zehnder type modulator, which has a cosine-shaped transmittance versus drive voltage

$$P_{out}(t) = P_{in}\,\gamma\left[1+\varepsilon\cos\left(\pi\frac{V_{bias}+e_{out}(t-\tau_d)}{V_\pi}\right)\right], \quad (1)$$

where P_{in} and P_{out} are the optical power incident on the modulator and detected at the end of the fiber. The modulator's parameters are γ, a factor related to the insertion loss; ε, a factor related to the extinction ratio; and V_π, the voltage that is required to move from a maximum to a minimum of the optical power transmittance. The bias point of the modulator is chosen to be one half of V_π.

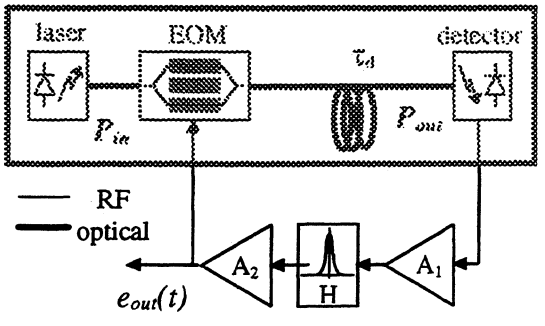


Figure 1. Basic scheme of an opto-electronic oscillator.

We assume that the signal around the loop will have the form

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

$$e_{out}(t) = V_0 \cos(\omega_0 t), \quad (2)$$

where $\omega_0=2\pi\nu_0$ is the oscillator angular frequency.

After expansion of the cosine in Eq. (1), Eq. (1) can be written as

$$P_{out}(t) = P_{in}\gamma\left[1-2\varepsilon\sum_{n=0}^{\infty} J_{2n+1}\left(\frac{V_0\pi}{V_\pi}\right)\cos\left((2n+1)\omega_0(t-\tau_d)\right)\right]. \quad (3)$$

Harmonic distortion in the EOM is neglected because the filter in Fig. 1 selects only the fundamental frequency. In general this element is a bandpass filter; in our case is a microwave resonant cavity with a Lorentzian transfer function

$$H(\omega) = \frac{H_0}{1+j2Q_L\dfrac{\omega-\omega_R}{\omega_R}} \cong H_0 e^{-j2Q_L\frac{\omega-\omega_R}{\omega_R}}, \quad (4)$$

where Q_L is the loaded quality factor and ω_R is the resonance frequency of the cavity. The exponential approximation in Eq. (4) is valid when

$$\omega-\omega_R << \frac{\omega_R}{2Q_L}; \quad (5)$$

that is, the cavity needs to be close to resonance at the oscillator frequency. The bandwidths of all the remaining elements in the oscillator, such as amplifiers, the detector, and the modulator, are considered to be wide enough that they will not affect the dynamics of the system. In particular the amplifiers' transfer functions will be

$$A_1(\omega) = A_{10}\, e^{-j\Phi_1} \quad \text{and} \quad A_2(\omega) = A_{20}\, e^{-j\Phi_2}. \quad (6)$$

The detected signal is simply $\rho P_{out}(t)$, where ρ is the detector responsivity. If we follow the rest of the signal path from the detector back to the modulator, we obtain

$$e_{out}(t) = A_{20}H_0A_{10}\rho P_{out}(t-\tau_{RF}), \quad \text{with} \quad \tau_{RF} = \frac{\Phi_1+\Phi_2}{\omega_0} + 2Q_L\frac{\omega_0-\omega_R}{\omega_0\omega_R}, \quad (7)$$

which, using Eqs.(2) and (3), can be written as

$$V_0\cos(\omega_0 t) = -2A_{20}H_0A_{10}\rho\varepsilon\gamma J_1\left(\frac{V_0\pi}{V_\pi}\right)\cos(\omega_0(t-\tau)), \quad \text{with} \quad \tau = \tau_d + \tau_{RF}. \quad (8)$$

It is now possible to write the oscillation condition as

$$\begin{cases} V_0 = 2A_{20}H_0A_{10}\rho\gamma P_{in}\varepsilon J_1\left(\frac{V_0\pi}{V_\pi}\right), \\ \omega_0\tau_d + (\Phi_1+\Phi_2) + 2Q_L\dfrac{\omega_0-\omega_R}{\omega_R} = (2K+1)\pi. \end{cases} \quad (9)$$

In absence of the resonator, the oscillation frequency is determined by the fiber alone:

$$\omega_F = \frac{2(K+1)\pi-(\Phi_1+\Phi_2)}{\tau_d}. \quad (10)$$

The fiber has a free spectral range of $1/\tau_d$. The effective oscillation frequency ω_0 depends on both cavity and fiber and its expression can be derived from Eq. (9):

$$\omega_0 = \frac{\omega_F\tau_d+2Q_L}{\omega_R\tau_d+2Q_L}\omega_R. \quad (11)$$

The detuning between the effective oscillation frequency and that determined by the fiber depends on both the initial detuning between the two selective elements and the two parameters Q_L and τ_d:

$$\omega_0-\omega_F = \frac{2Q_L}{\omega_R\tau_d+2Q_L}(\omega_R-\omega_F). \quad (12)$$

These are the working conditions for an oscillator with noiseless components; the derivation of the transfer functions for the noise spectra in the oscillator differs in the two approaches that will be addressed as the *Leeson's method* (application of control system theory) and the *Van der Pol's method* (time domain analysis).

3. LEESON'S METHOD

3.1 The noise transfer functions (NTF)

This method applies, in the frequency domain, control system theory to the transfer functions of the elements which describe the oscillator.

In particular, we represent the ensemble of the laser, the EOM, the fiber, and the detector as a whole "black box", called *optical link*. This optical link has $e_{out}(t)$ as input signal and the detected $e_d(t)$ as output signal. Recalling Eq. (3), we write the transfer function of the optical link in the frequency domain as

$$L(\omega) = \frac{E_d(\omega)}{E_{out}(\omega)} = -L_0 e^{-j\omega\tau_d}. \quad (13)$$

The system described by the "black boxes" defined above is shown in Fig. 2, with all the transfer functions involved.

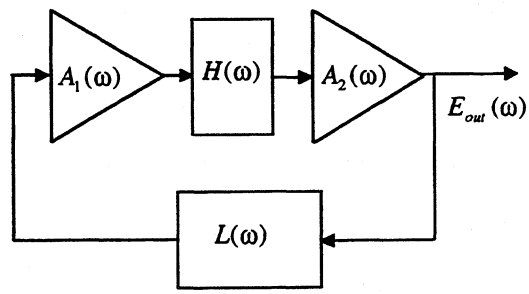


Figure 2. Block scheme of the optoelectronic oscillator.

To investigate the noise in the system, we make two assumptions. First, using the vectorial representation of a noisy signal shown in Fig. 3, we consider separately

the amplitude and phase contribution of the noise component of the signal, and therefore

$$S_\varphi(f) \quad \Leftrightarrow \quad (\Delta\Phi(f))^2 = \left(\frac{V_n(f)}{V_0\sqrt{2}}\right)^2,$$
$$S_\alpha(f) \quad \Leftrightarrow \quad \left(\frac{\Delta V(f)}{V_0}\right)^2 = \left(\frac{V_n(f)}{V_0\sqrt{2}}\right)^2, \tag{14}$$

where both thepower spectral densities of amplitude (α) and phase (φ) are represented. Second, it is possible to represent the noise of a device as an input equivalent noise which is transferred to the output by a noiseless transfer function.

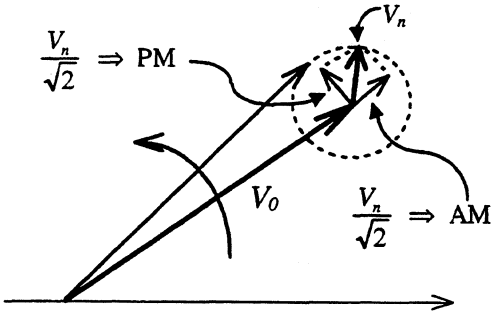


Figure 3. Vectorial representation of a noisy signal.

As a consequence of these two assumptions, the block diagram of the system with noise sources appears as shown in Fig. 4.

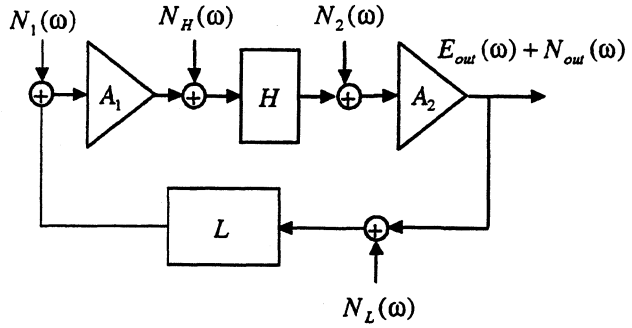


Figure 4. Block diagram of the noisy OEO.

These noise spectra can represent phase noise or amplitude noise as shown in Eq. (15),

$$N_i(\omega) = \begin{cases} V_i \cdot \Delta\Phi(\omega-\omega_0) & \text{(PM)}, \\ \Delta V(\omega-\omega_0) & \text{(AM)}, \end{cases} \tag{15}$$

where $i = 1, H, 2, L,$ and V_i is the signal at the respective summing junction.

The complete set of transfer functions for all noise contributions deduced from Fig. 4 is

$$\begin{aligned} &1^{st}\text{ amplifier} && N_{out} = \frac{A_1HA_2}{1-LA_1HA_2}N_1, \\ &\text{cavity} && N_{out} = \frac{HA_2}{1-LA_1HA_2}N_H, \\ &2^{nd}\text{ amplifier} && N_{out} = \frac{A_2}{1-LA_1HA_2}N_2, \\ &\text{optical link} && N_{out} = \frac{1}{1-LA_1HA_2}N_L. \end{aligned} \tag{16}$$

If we substitute the explicit expression for phase noise or amplitude noise shown in Eq. (15), the transfer functions for all the noise contributions and for both kind of noise spectra appear to be the same and can be expressed as

$$NTF(\omega) = \frac{1}{1-LA_1HA_2} = \left[1 + L_0A_{10}H_0A_{20}e^{-j\left(\Phi_1+\Phi_2+\omega\tau_d+2Q_L\frac{\omega-\omega_R}{\omega_R}\right)}\right]^{-1}. \tag{17}$$

Because the analysis takes place around the oscillation frequency, we can write

$$\omega = \omega_0 + \Omega, \tag{18}$$

where $\Omega/2\pi = f$ is the frequency offset from the carrier (Fourier frequency). Through application of Eq. (9), (10) and (12), we obtain

$$NTF(\Omega) = \frac{1}{1-e^{-jF(\Omega)}},$$
$$\text{with} \quad F(\Omega) = \Omega\left(\tau_d + \frac{2Q_L}{\omega_R}\right) + 2Q_L\frac{\omega_R\tau_d(\omega_0-\omega_F) - 2Q_L(\omega_R-\omega_0)}{\omega_R(\omega_R\tau_d+2Q_L)}, \tag{19}$$

and therefore the transfer function of the power spectral densities appears as

$$|NTF(\Omega)|^2 = \frac{1}{[1-\cos(F(\Omega))]^2 + \sin^2(F(\Omega))}. \tag{20}$$

If there is no detuning between the oscillation frequency set by the fiber and the resonance frequency of the cavity, the Eq. (19) is simplified, and we can add some consideration. In fact if the Fourier frequency is small enough, the noise transfer function can be approximated by

$$|NTF(\Omega)|^2 \approx \frac{1}{F^2(\Omega)} = \frac{\omega_0^2}{\Omega^2(\omega_0\tau_d+2Q_L)^2}, \tag{21}$$
$$\text{with} \quad \Omega \ll \frac{1}{\tau_d} \quad \text{and} \quad \Omega \ll \frac{\omega_0}{2Q_L}.$$

This is a simple result, identical to that found by Leeson in his paper.

3.2 Prediction on the experimental device.

In order to apply this model and calculate the total expected noise spectrum of an OEO, we have to know the equivalent input noise of all components in the system. It is also possible simply to calculate the effect of a single component on the spectral quality of the signal produced by the oscillator.

In our case the phase noise of the amplifiers is known from direct measurement [3], and its effect calculated using this model, is given by

$$S_\varphi(f)\Big|_{osc}^{pred} = \frac{S_\varphi(f)\Big|_{amp}}{\left[1-\cos\left(\Omega\tau_d+\frac{2Q_L\Omega}{\omega_0}\right)\right]^2+\sin^2\left(\Omega\tau_d+\frac{2Q_L\Omega}{\omega_0}\right)} \quad (24)$$

and is compared with the measured oscillator phase noise [3] in Fig. 5.

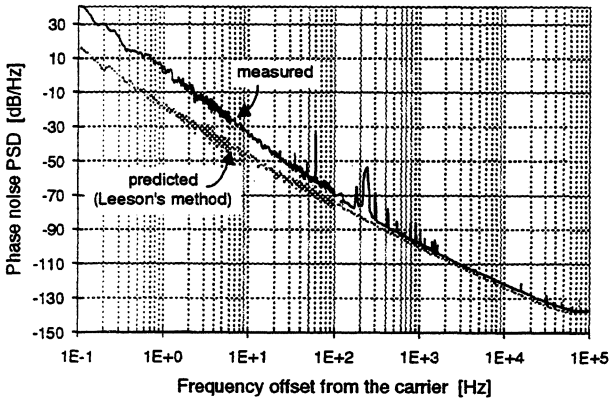


Figure 5. PSD of the measured oscillator noise and predicted amplifier phase noise contribution according to the model achieved with the Leeson's method.

The discrepancy between the two curves for Fourier frequencies below 100 Hz is due to the presence of excess noise on the oscillator signal, caused mainly by environmental factors like thermal fluctuation and vibrations.

4. VAN DER POL'S METHOD

4.1 The time domain analysis

This analysis is done mainly in the time domain and is based on the assumption that the oscillator signal can be written as

$$e_{out}(t) = V(t)\cos(\omega_0 t+\varphi(t)), \quad (22)$$

with both amplitude and phase varying *slowly* in time. The starting point is the oscillation condition expressed by Eq. (8) and rewritten as

$$V(t+\tau)\cos(\omega_0 t+\omega_0\tau+\varphi(t+\tau)) = -2GP_{in}\varepsilon\gamma \cdot J_1\left(\frac{V(t)\pi}{V_\pi}\right)\cos(\omega_0 t+\varphi(t)), \quad (23)$$

where

$$G = A_2 H A_1 \rho,$$
$$\tau = \tau_d + \frac{\phi}{\omega_0} + 2Q_L\frac{\omega_0-\omega_R}{\omega_0\omega_R}, \quad (24)$$
$$\phi = \phi_1+\phi_2.$$

Under the assumption of *slow variations*, we can write

$$V(t+\tau) = V(t)+\dot V(t)\tau+\ddot V(t)\tau^2,$$
$$\varphi(t+\tau) = \varphi(t)+\dot\varphi(t)\tau+\ddot\varphi(t)\tau^2, \quad (25)$$

with $\tau >> 2\pi/\omega_0$.

Using Eqs. (23) and (25), we can write the differential equation which describes the amplitude and phase fluctuations of the oscillation signal:

$$\left(V+\dot V\tau+\ddot V\tau^2\right)\cos\left(\omega_0 t+\omega_0\tau+\varphi+\dot\varphi\tau+\ddot\varphi\tau^2\right) = -2GP_{in}\varepsilon\gamma J_1\left(\frac{V\pi}{V_\pi}\right)\cos(\omega_0 t+\varphi), \quad (26)$$

where are $V(t)$ and $\varphi(t)$ the independent variables.

The stationary condition is represented by the derivatives of the state variables equal to 0. Under these assumptions, we recover the oscillation condition already calculated in Eq. (9):

$$\begin{cases} V_0 = 2G_0P_{in}\gamma\varepsilon J_1\left(\frac{V_0\pi}{V_\pi}\right), \\ \omega_0\tau_0 = (2K+1)\pi \quad \text{with} \quad \tau_0 = \tau(\omega_0). \end{cases} \quad (27)$$

The parameters of the system which affect its equilibrium are: the loop gain G, the optical power P_{in}, the fiber delay τ_d, the phase delay introduced by the amplifiers ϕ, and the resonance frequency of the cavity ω_R.

If Eq. (26) is linearized around the stationary condition with respect to both the independent variables and the parameters, it is possible to calculate its Fourier transform. The linearized equations appears as

$$\begin{cases} \left(\frac{V_0\pi}{V_\pi}\frac{J_0\left(\frac{V_0\pi}{V_\pi}\right)}{J_1\left(\frac{V_0\pi}{V_\pi}\right)}-2\right)\cdot\delta V(t)-\tau_0\cdot\delta\dot V(t)-\tau_0^2\cdot\delta\ddot V(t) \\ \qquad +\frac{V_0}{G_0}\cdot\delta G(t)+\frac{V_0}{P_{in}}\cdot\delta P_{in}(t)=0, \\ \tau_0\cdot\delta\dot\varphi(t)+\tau_0^2\cdot\delta\ddot\varphi(t)+\omega_0\cdot\delta\tau_d(t)+\delta\phi(t) \\ \qquad -2Q_L\frac{\omega_0}{\omega_R^2}\cdot\delta\omega_R(t)=0, \end{cases} \quad (28)$$

where the Greek letter δ indicates difference with respect to the equilibrium value.

The application of the Fourier transform produces

$$\begin{cases} \dfrac{\Delta V(\Omega)}{V_0} = \dfrac{\left(\dfrac{\Delta G(\Omega)}{G_0} + \dfrac{\Delta P_{in}(\Omega)}{P_{in}}\right)}{2 - \dfrac{V_0\pi}{V_\pi}\dfrac{J_0\left(\frac{V_0\pi}{V_\pi}\right)}{J_1\left(\frac{V_0\pi}{V_\pi}\right)} + j\Omega\tau_0 - \Omega^2\tau_0^2}, \\ \Delta\varphi(\Omega) = \dfrac{\left(\omega_0\Delta\tau_d(\Omega) + \Delta\phi(\Omega) - 2Q_L\dfrac{\omega_0}{\omega_R^2}\Delta\omega_R(\Omega)\right)}{\Omega\tau_0(\Omega\tau_0 - j)}. \end{cases} \quad (29)$$

A direct consequence is the relationship between all noise power spectral densities involved in the system.

In particular, the phase noise of the oscillator, is

$$S_\varphi(\Omega) = \frac{\omega_0\tau_d \cdot S_{\Delta\tau_d/\tau_d}(\Omega) + S_\phi(\Omega) - 2Q_L\dfrac{\omega_0}{\omega_R}\cdot S_{\Delta\omega_R/\omega_R}(\Omega)}{\Omega^2\tau_0^2(\Omega^2\tau_0^2+1)}. \quad (30)$$

Recalling Eq. (24), which contains the explicit form of τ, we can calculate its value at equilibrium and find the part of Eq. (30) relative to the amplifier phase noise (that is, $S_\phi(\Omega)$) which matches the results obtained with Leeson's method.

4.2 Prediction on the experimental device

As with previous method, we can use the results produced by this model to predict the effect of amplifier phase noise on the oscillator noise.

The transfer function that we are interested in is

$$S_\varphi(\Omega)\Big|_{osc}^{predicted} = \frac{S_\phi(\Omega)\Big|_{amp}}{\Omega^2\tau_0^2(\Omega^2\tau_0^2+1)} \approx \frac{S_\phi(\Omega)\Big|_{amp}}{\Omega^2\tau_d^2}, \quad (31)$$

where $S_\phi(\Omega)\Big|_{amp}$ has been measured [3]. The measured noise of the OEO and the predicted amplifier noise contribution to it are compared in Fig. 6.

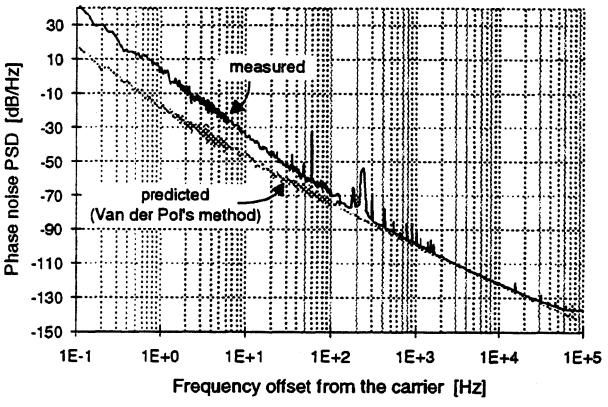


Figure 6. PSD of the measured oscillator noise and predicted amplifier phase noise contribution according to the model achieved with the Van der Pol's method.

The discrepancy between the two curves in the lowest frequency part of the spectrum is due to the presence of excess noise mainly caused by environmental factors as thermal variation and vibrations.

5. CONCLUSIONS

The two model described in this paper have different characteristics. The first model, calculated with control system theory (Leeson's method), is applicable when it is possible to describe the system in terms of black boxes whose associated noise is measurable without knowing the details of their operation.

The model calculated with Van der Pol's method permits us to identify the single contribution to the noise due to the fiber length variations or the cavity resonant frequency fluctuations. On the other hand, it is clearly less flexible and in fact requires the ability to write the system equations in the time domain and to verify all the assumption leading to the final expressions.

Although the two models presented belong to different analyses, their agreement with the experimental data and between them is good, as shown in Fig. 5 and Fig. 6.

We thank S. R. Jefferts for support and many helpful discussions. We are also grateful to M. Young for a careful reading of the manuscript.

6. REFERENCES

[1] D. B. Leeson, "A simple model of feedback oscillator noise spectrum," Proc. IEEE (Lett.), Vol. 54, pp. 329-330, February 1966.

[2] X. S. Yao and L. Maleki, "Optoelectronic oscillator for photonic system," IEEE J. Quant. Electron., Vol. 32, no. 7, pp. 1141-1149, 1996.

[3] S. Römisch, J. Kitching, E. Ferrè-Pikal, L. Hollberg and F. Walls, "Performance evaluation of optoelectronic microwave oscillators," in these proceedings.

[4] B. Van der Pol, "The nonlinear theory of electric oscillations," Proc. IRE, Vol. 22, no. 9, pp. 1051-1086, 1934.

[5] S. Römisch "Tecniche optoelettroniche per la generazione e il trattamento di frequenze stabili in banda X," Ph.D. Thesis, Politecnico di Torino (Italy), February 1998.

A PRACTICAL METHOD TO PROCESS TIME AND FREQUENCY SIGNAL

Wei Zhou, Zhiyong Li

Dept. of Measurement and Instrumentation, Xidian University, Xi'an, 710071, P. R. China

ABSTRACT

A new practical time and frequency signal processing method can be used to generate an accurate time signal, stable phase step, frequency change and synthesis, and in other practical uses. The method can accomplish the time and frequency signal processing by changing the period and phase of the machine cycle of a microcomputer. It is based on pulse deletion of high frequency clock signal, quantified phase delay, and other phase lock and variation in a microcomputer clock signal circuit. The pulse deletion and phase delay can be implemented periodically or singly for different purpose, and the frequency change and synthesis or phase shift of the microcomputer output signal can be obtained respectively. With this method and software control, a simple narrow frequency range regulation can generate an accurate frequency and time signal in a wider frequency range. Many multiple frequency difference, phase and time processing methods, period deletion methods can be used here, and different precision can be obtained. With this method the time signal processing precision can be from several tens ns to less than 1 ns. Combined with other techniques, 100 ps to 10 ps precision can be obtained. Therefore, the practical technique can be used widely.

1. INTRODUCTION

Based on a highly stable frequency standard signal, generation of accurate time signal, variation of stable phase, synthesis and change of frequency signal, and precision measurement of time and frequency can be used widely. Using a microcomputer to do the works is a common used approach. However, the long machine cycle of the microcomputer influences the resolution of its output signal. To enhance the resolution of processing time and frequency with the computer, some techniques have been used. However, some simpler approaches based on new principles are more effective choice.

2. THE FUNCTION OF THE CLOCK CYCLE AND ITS PHASE CHANGE OF COMPUTER

Based on software design, the period T of periodic signal generated by computer can be

$$T=nt \qquad (1)$$

where t is the machine cycle of the computer; n is an integer controlled by software. The machine cycle of the computer is much longer than its clock cycle. For example, their ratio is 12 in many microcomputers. Therefore, the time resolution generated by computer is limited a lot. If t_1 is clock cycle of the computer,

$$T=n(mt_1) \qquad (2)$$

where m is a integer ratio between machine cycle and clock cycle and it often is fixed (for example, 12). To enhance the resolution, one must pay attention to m and t_1. Obviously, with deleting clock pulses one can change m, and with shifting clock phase one can also change t_1.

If only based on the machine cycle, the computer can generate a frequency signal f_0. After deleting the computer clock pulses (processing cycle) and shifting its phase (with quantified time delay approach), the computer will generate a frequency difference Δf. That is $f_1=f_0+\Delta f$

$$\frac{\Delta f}{f_0}=\frac{\Delta T}{\tau} \qquad (3)$$

and

$$f_1=f_0+\Delta f=f_0(1+\frac{\Delta T}{\tau}) \qquad (4)$$

where ΔT represents in τ period time (phase) variation generated by deleting clock pulses and shifting phase of the computer. For convenient, we can choose τ as the period of f_1. We can used any method to delete the clock pulses and shift the clock phase. If equal amount time including clock cycles and phase differences are deleted in the period of f_1, the f_1 frequency must be stable.

The period T of f_1 is

$$T=nt+n_1t_1+n_2t_2 \qquad (5)$$

where t_2 is the unit delay time of the delay chain, and n, n_1, n_2 are controlled by the computer, they are integers. With this method to process time or frequency signal, the basic principle is time and phase processing method. Here $t>n_1t_1$, $t_1>n_2t_2$.

Quantified time delay method can be used to shift the clock phase, and different gate circuits can be used as the delay components [1]. To delete the pulses of a high frequency signal then divide its frequency as the clock signal of microcomputer, higher resolution can also be obtained.

From equation (4) and (5), only the clock pulse deletion and its phase delay are done periodically, the stable output frequency signal of the computer with frequency change and synthesis can be obtained. If they

0-7803-5400-1/99/ $ 10.00

are done singly, the phase shift of out signal can be obtained.

With this method and software control, a simple narrow range frequency (period) and phase regulation can generate an accurate frequency and time signal in a much wider frequency range. A simple frequency synthesizer can also be used as the clock signal of the computer to obtain better time resolution.

Because of limitation of computer command cycle and its software operation time, its output frequency can't be very high. With a simple method, the computer can only generate a square wave signal. In many practical designs one can use the square wave signal with a high resolution as a base of different signals and make different devices. Based on the square wave signal, combined with a phase-locked loop and other circuits, a sine oscillation signal in a very wide frequency range can be obtained.

This method can be combined with a phase locked loop and a variable frequency divider together to generate some higher frequency signals. Its block diagram is shown in Fig. 1.

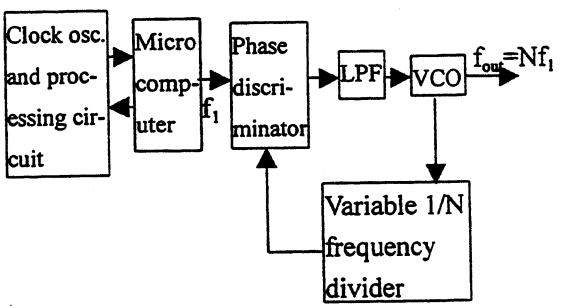


Fig. 1 Higher frequency generator

3. PRACTICAL APPROACHES

To enhance the resolution of the output frequency signal generated by computer with deleting clock pulses approach, one can delete pulses of a stable and high frequency signal, then divide it into a suitable frequency signal as the clock signal of the computer. With this way, in the equation (5) t_1 is the period of the high frequency signal.

A simple deleting clock pulses circuit is shown in Fig. 2, and Fig. 3 is its waveform diagram.

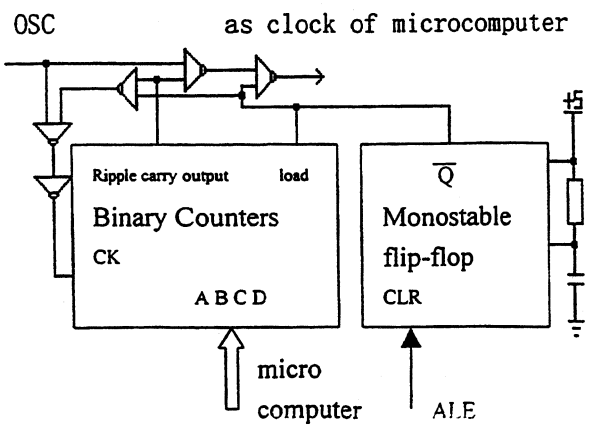
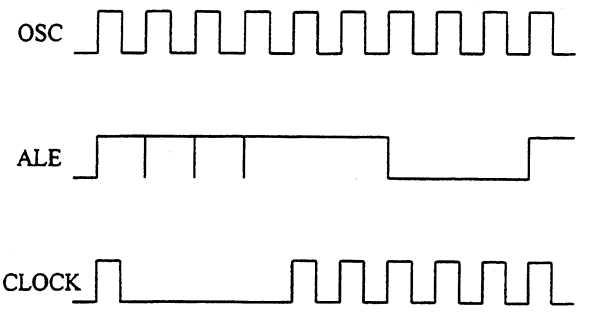


Fig. 2 A simple deleting clock pulses circuit



Deleting 3 paulses

Fig. 3 The waveform diagram of deleting pulse circuit

Using the deleting clock pulses approach, the highest resolution is limited by the period of the clock frequency or the relative high frequency signal. For higher resolution, the phase processing method to the clock signal can be used. Here, the period and phase processing of the clock have the same meaning, all are based on time processing.

The clock phase processing approach used to enhance time and frequency processing resolution is shown in Fig.4.

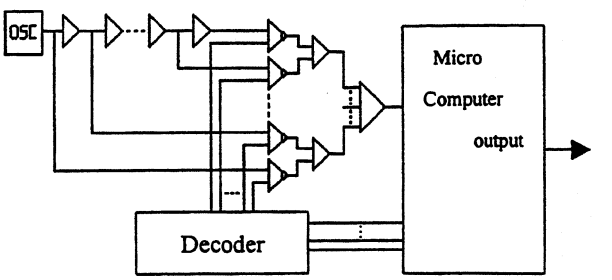


Fig. 4 With shifting clock phase approach to enhance the resolution of the output signal of a microcomputer

Here, the delay circuit is used to shift the clock phase, and it is composed of a gate chain. The delay stability must be stable and the performances of other components ought to be symmetrical. Generally, with more delay components and shorter delay time of the delay unit, the processing precision can be higher. In this case, one must choose high quality delay units. There are many different methods to build the delay chain, in series or series parallel. Sometimes suitable delay feedback can decrease the components number. With periodically shifting the clock phase, higher time processing resolution than deleting clock pulses approach can be obtained. However, in this case the frequency stability will be lowered and some problems will be found, if the delay circuit is too simple. Because of the delay error of delay unit in the delay chain, in the output signal the frequency fluctuation will be generated.

The total delay time of the delay chain must be equal

to the clock cycle exactly, then periodically shifting phase combined with deleting clock pulse can generate a high stable output signal. However, it is difficult, or a "coincidence control circuit" must be used. If the total delay period can not coincide with the period of the clock cycle, the "no coincidence error" will be generated, and the periodic phase modulation too. The unstability and unsymmetry of other circuits used to control the phase shift can also influence the signal stability. Therefore, choosing suitable delay steps is very important. When gate circuit are used as the delay unit, around 20 steps are suitable.

Two clock signals with a suitable frequency difference can also be used to enhance the time signal resolution generated by a microcomputer. In this case some attached circuits (counters and coincidence detection circuits) must be used.

A high frequency standard can combined with the shifting phase circuit, deleting pulse circuit and frequency divider to generate the clock signal of a microcomputer. Its block diagram is shown in Fig. 5.

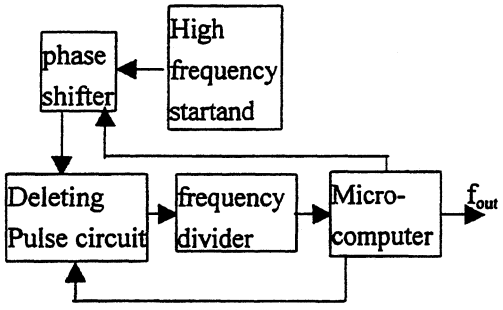


Fig. 5 A clock signal processing circuit

This approach takes advantages of different methods and avoids the problem from too many delay units. The frequency of standard can be from 50 MHz to more than 100 MHz, to obtain high time precision with less delay units. According to the approach, a simple device has been made with 1 ns time resolution from output of a microcomputer. We also combine the approach with a phase-locked VCXO through a conversion circuit with multiple frequency difference function. In this case, the time resolution of the output signal can be from 100 ps to 10 ps according to the different multiple ratio.

4. TESTING AND RESULT

Using the method shown in Fig. 2 we tested the frequency stability of output signal of a microcomputer with deleting clock pulses function. A high stable 5 MHz frequency standard is used as the clock signal of the microcomputer and the standard of a high precision frequency meter. The frequency meter is used to measure the frequency of the output signal. Changing the software several consecutive frequency output signals are measured and shown in table 1. In the experiment the output signals are obtained from ALE terminal of the computer.

Table 1. Output signal test with deleting clock pulses function

the clock phases number to be deleted	1	2	3
output signal frequency(KHz)	624.9999989 625.0000001 624.9999996 624.9999994 624.9999995 625.0000002 624.9999993 625.0000004 625.0000003 624.9999995 625.0000000 625.0000004 624.9999996 624.9999990 625.0000002	555.55555549 555.55555556 555.55555558 555.55555563 555.55555545 555.55555553 555.55555542 555.55555544 555.55555554 555.55555543 555.55555544 555.55555543 555.55555554 555.55555557 555.55555543	500.00000001 499.99999993 499.99999993 499.99999993 499.99999998 500.00000006 500.00000001 499.99999991 500.00000001 500.00000004 499.99999993 499.99999998 500.00000004 500.00000001 499.99999991
maximum error	$\pm 1.2\times 10^{-10}$	$\pm 1.9\times 10^{-10}$	$\pm 1.5\times 10^{-10}$

Using the same method but without deleting the clock pulses of the microcomputer, we tested the frequency signal. The consecutive frequency signals are measured and shown in table 2.

Table 2. Output signal test without deleting clock pulses function

	consecutive output frequencies based on computer software	
output signal frequency (KHz)	104.16666668	83.333333326
	104.16666662	83.333333353
	104.16666669	83.333333337
	104.16666664	83.333333346
	104.16666666	83.333333344
	104.16666666	83.333333351
	104.16666662	83.333333359
	104.16666662	83.333333354
	104.16666663	83.333333353
	104.16666665	83.333333363
maximum error	$\pm 3.3\times 10^{-10}$	$\pm 2.2\times 10^{-10}$

The precision of the frequency meter can be 1×10^{-10}/second. If the frequencies to be measured is near or lower than 100 kHz, the precision is only 3×10^{-10}/second. In the measurement we choose 1 second gate time. Because of the limitations of the stability of the frequency standard and the precision of the frequency meter, the measurement results can not represent the influence limit on which the hardware and software of the microcomputer. However, according to the testing

results, at least the influence of the hardware and software on the signal frequency stability is less than 0.1 ppb. From the test results, with deleting clock pulses function the consecutive frequency interval of the output signal is much finer, and the frequency resolution is enhanced. Using a delay chain with 10 ns delay unit, we tested the output frequency signal, and the result is shown in table 3.

Table 3. output signal test with phase shift function

	consecutive output frequencies based on deleting clock pulses and clock phase processing		
output signal frequency (KHz)	502.51256260	505.05050535	507.61421311
	502.51256320	505.05050499	507.61421297
	502.51256281	505.05050561	507.61421289
	502.51256275	505.05050478	507.61421267
	502.51256295	505.05050496	507.61421355
	502.51256289	505.05050505	507.61421311
	502.51256239	505.05050521	507.61421347
	502.51256242	505.05050555	507.61421319
	502.51256244	505.05050511	507.61421363
	502.51256322	505.05050500	507.61421301
	502.51256298	505.05050440	507.61421289
	502.51256277	505.05050498	507.61421337
maximum error	$\pm 8.2\times 10^{-10}$	$\pm 1.2\times 10^{-9}$	$\pm 9.4\times 10^{-10}$

From table 3 one can find that the consecutive frequency interval of the output signal with shifting clock phase is much finer than that with clock pulse processing, but its frequency stability is lowered.

5. CONCLUSION

With microcomputer to process frequency and time signals is a simple and convenient method. Using deleting clock pulses and shifting its phase approach finer frequency interval and higher resolution of the output signal of the microcomputer can be obtained. Many multiple frequency difference, phase and time processing methods, period deletion methods can also be used in the clock circuit, and different precision can be obtained with different structure. This method have been used in a new MCXO [2] and some frequency synthesizer, and will be used widely.

REFERENCES

[1] Wei Zhou, et al "Some New Methods for Precision Time Interval Measurement", The Proceedings of the 1997 IEEE International Frequency Control Symposium, pp. 418 - 421, 1997

[2] Wei Zhou, et al "An Improvement Method of MCXO", 1999 IEEE International Frequency Control Symposium

1999 Joint Meeting EFTF - IEEE IFCS

ESTIMATION OF THE POWER SPECTRAL DENSITY OF PHASE: COMPARISON OF THREE METHODS

François Vernotte

Observatoire de Besançon

41 bis av. de l'Observatoire, BP 1615, F-25010 Besançon Cedex

Phone: +33-3-81.66.69.22, email: francois@obs-besancon.fr

Abstract—**The power spectral density (PSD) of time error (or phase) is generally modeled as a sum of power laws from f^{-4} (long term instabilities) to f^0 (short term instabilities), but other features may be observed in the PSD. In this paper, we compare three estimation methods of the frequency PSD without assuming a particular model : a classical Fourier analysis, an estimation using discrete n^{th} differences and an analysis based on orthogonal polynomials. After a description of these methods, their advantages and drawbacks are reviewed in terms of computation, of spectral sensitivity and of adaptability to particular cases (small number of samples, unregularly spaced data, ...). This paper also deals with the statistics of the spectral estimates. The sensitivities of these three estimation methods are compared in the case of an assumed power law model.**

1. INTRODUCTION

In many applications, it is necessary to estimate the measurement uncertainty of the mean frequency of an oscillator. Furthermore, it may be helpful to extrapolate the range of variation of the mean frequency from now to a given date. Although this problem is easy for a white process, such an estimation is quite difficult in the case of non-stationary processes, particularly for a random walk frequency noise. The result of such a study depends on the value of the different noise levels [1][2]. In this connection, the estimation of these noise levels is essential.

Thus, the aim of this paper is not the proposal of a n^{th} variance, but the choice of a tool, as simple as possible, which could give a quick survey of the general features of the PSD of time error. In order to achieve this issue, this tool must satisfy several criterions.

On the right hand, it must converge for very low frequency noise. Since the PSD of time error may be modeled as:

$$S_x(f) = \sum_{\alpha=-4}^{0} h_\alpha f^\alpha \qquad (1)$$

the estimator must converge up to f^{-4} PM, without taking into account a low cut-off frequency.

On the other hand, the deterministic drifts which may also alter the time error data, must not be confused with long term random fluctuations. Subsequently, the estimator must be insensitive to the deterministic drifts.

2. ESTIMATION OF THE PSD $S_x(f)$

It is possible to estimate $S_x(f)$ from N time error samples $\{x(t_1), \ldots, x(t_N)\}$ by using a set of M sampling functions $\{h_1(t), \ldots, h_M(t)\}$ such as :

$$\widehat{S_x}(f_j) = \left| \int_{-\infty}^{+\infty} h_j(t) x(t) dt \right|^2 \qquad (2)$$

where $\widehat{S_x}(f_j)$ is an estimate of $S_x(f)$ for a given frequency f_j.

It can be shown that the expectation of this estimate is:

$$\left\langle \widehat{S_x}(f_j) \right\rangle = \int_{-\infty}^{+\infty} |H_j(f)|^2 S_x(f) df \qquad (3)$$

where $H_j(f)$ is the Fourier transform of the sampling function $h_j(t)$.

In order to really estimate the PSD of phase, such a function must verify several general properties. Moreover, the considerations described in introduction implies that the estimators verify a set of special additional properties.

2.1 General properties of the estimators

- $H_j(f)$ must be narrow and centered around f_j
- the relative bandwidth $\Delta f / f_j$ of $H_j(f)$ must be independent on j
- the estimators must be normalized :

$$\int_{-\infty}^{+\infty} |h_j(t)|^2 dt = \int_{-\infty}^{+\infty} |H_j(f)|^2 df = 1 \qquad (4)$$

- the estimators should be independent (orthogonality condition) in order to ensure independent estimates of $S_x(f)$:

$$\int_{-\infty}^{+\infty} h_j(t) \cdot h_k(t) dt = \delta_{jk} \qquad (5)$$

(this is only true for white PM with this condition).

2.2 Special properties of the estimators

The estimators must be insensitive to drifts :

$$\int_{-\infty}^{+\infty} h_j(t) \cdot t^\beta \cdot dt = 0 \qquad (6)$$

with $\beta = 0$ (constant phase term), 1 (linear phase drift) and 2 (quadratic phase drift).

Moreover, the estimates must converge for low frequency noises such as f^{-4} PM :

$$\int_{-\infty}^{+\infty} |H_j(f)|^2 f^{-4} df \quad \text{must be finite} \qquad (7)$$

From the *moment condition* [3], it can be proved that the insensitivity to linear phase drift yields the convergence up to f^{-4} PM. Thus, only (6) has to be verified for ensuring both drift insensitivity and convergence properties.,

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

2.3 Examples of estimators

Obviously, we can use an estimator whose transfer function is a Dirac distribution:

$$\begin{cases} H_j(f) = \delta(f - f_j) \\ h_j(t) = \exp(i2\pi f_j t). \end{cases} \quad (8)$$

This leads to the classical Fourier analysis.

It is well known that n^{th} order differences may be considered as estimators. For instance, the Allan variance, which is a 2^{nd} order difference for $x(t)$ data, is defined by its calculation sequence or its transfer function:

$$\begin{cases} h_j(t) = \frac{1}{\tau\sqrt{2}}[\delta(t+j\tau) - 2\delta(t) + \delta(t-j\tau)] \\ |H_j(f)|^2 = 8\frac{\sin^4(\pi\tau f)}{\tau^2}. \end{cases} \quad (9)$$

Finally, a set of polynomials of ascending degrees (Chebyshev polynomials) may also be used as a set of estimators of the PSD:

$$h_j(t) = \sum_{p=0}^{j} a_p \cdot t^p. \quad (10)$$

The degree j of such a polynomial is directly linked to the analysis frequency f_j.

3. FIRST METHOD: n^{th} ORDER DIFFERENCE

The variances are generally plotted versus the integration time τ. However, since τ is linked to the frequency f_j of the maximum of the transfer function of the variance by the following relationship:

$$f_j = \frac{1}{2\tau}, \quad (11)$$

the variance curve may also be plotted versus f_j. According to a correct normalization of the variance, this plot may be an estimate of the PSD $S_x(f)$.

By using the structure function approach [4][5], it is possible to define a variance verifying the convergence properties needed. In a previous paper [6], we defined the "pulsar variance" (P-var), whose one of its property was the insensitivity to quadratic phase drift.

3.1 The "pulsar variance" (P-var)

P-var is defined by the following calculation sequence (see figure 1):

$$h_j(t) = \sqrt{\frac{\tau_0}{20n}} \sum_{i=1}^{n} [\delta(t+(n+i)\tau_0) - 3\delta(t+i\tau_0) + 3\delta(t-i\tau_0) - \delta(t-(n+i)\tau_0)]. \quad (12)$$

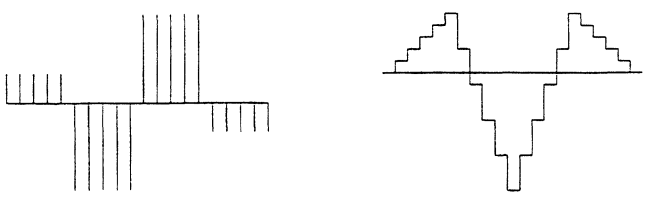


Fig. 1. Calculation sequence of P-var applied to $x(t)$ data (left) and to $\bar{y}_k$ samples (right).

$S_x(f)$	P-var(f)		Bias
k_0	$\frac{k_0}{2}$	$\approx 0.5k_0$	1
$k_{-1}f^{-1}$	$\frac{9}{10}\ln(\frac{16}{9})k_{-1}f^{-1}$	$\approx 0.518k_{-1}f^{-1}$	1.036
$k_{-2}f^{-2}$	$\frac{\pi^2}{15}k_{-2}f^{-2}$	$\approx 0.658k_{-2}f^{-2}$	1.316
$k_{-3}f^{-3}$	$\frac{\pi^2(27\ln 9 - 80\ln 2)k_{-3}}{40f^3}$	$\approx 0.956k_{-3}f^{-3}$	1.912
$k_{-4}f^{-4}$	$\frac{\pi^4}{60}k_{-4}f^{-4}$	$\approx 1.623k_{-4}f^{-4}$	3.246

TABLE I

RESPONSE OF P-VAR FOR THE DIFFERENT TYPES OF NOISES. THE LOW FREQUENCY NOISES ARE A BIT OVERESTIMATED (BIAS > 3).

Its transfer function may be approximated (asymptotic approximation) by:

$$|H_j(f)|^2 \approx \frac{16\sin^8(\pi\tau f)}{5\pi^2 f^2 \tau} \quad (13)$$

with $\tau = n\tau_0$ and $f_j = \frac{1}{2\tau}$.

3.2 Properties of P-var

3.2.1 General properties

Only the last condition of those defined in 2.1 is not verified: the estimators are not independent. This implies that, for different frequencies, the estimates are partially correlated. However, it may be shown that these correlations are weak.

3.2.2 Special properties

Since P-var is a 3^{rd} order difference, it is insensitive to quadratic phase drifts. Subsequently, P-var converges from f^{-6} PM to white PM.

3.3 Distribution of the estimates

If $x(t)$ is a gaussian process, the estimates are χ^2 distributed. The number of degrees of freedom may be calculated as:

$$\nu = 2\left(\frac{\langle \widehat{S_x} \rangle}{\sigma_{S_x}}\right)^2. \quad (14)$$

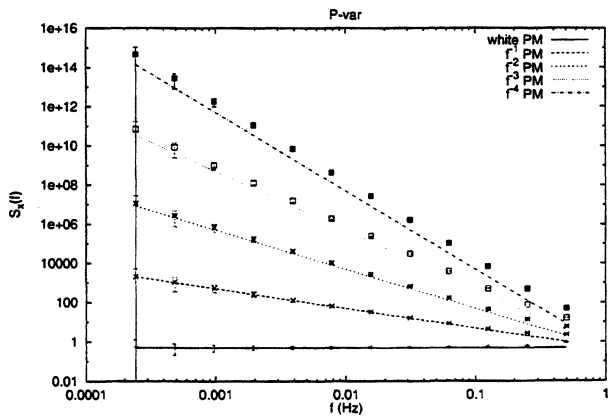


Fig. 2. Graph of P-var versus the frequency for the different types of noise. The continuous lines represent the asymptotic behavior, the points are the mean value obtained by simulations and the error-bars are the dispersions around these mean values.

From a Monte-Carlo method ($N = 8192$ $x(t)$ data), we obtained a number of degrees of freedom between 1 (for the lowest frequency) and 6000 (for the highest frequency).

4. SECOND METHOD: ORTHOGONAL POLYNOMIALS

Polynomials of ascending degrees look like sines of increasing frequencies (see figure 3). Thus, they may be used as estimators [3][7].

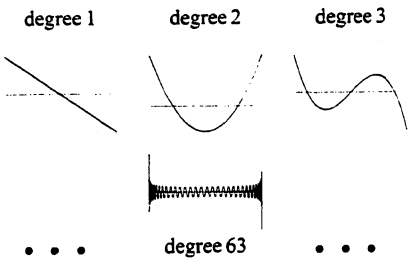


Fig. 3. Graph of several Chebyshev polynomials of ascending degrees.

4.1 Calculation of the polynomials

4.1.1 The Chebyshev polynomials

If $x_i \in \{0, 1, \ldots, N\}$, the Chebyshev polynomials $f_j(x)$ are defined by:

$$\begin{cases} f_0(x) = 1 & f_1(x) = 1 - \frac{2x}{N} \\ & \text{and} \\ (j+1)(N-j)f_{j+1}(x) & = (2j+1)(N-2x)f_j(x) \\ & \quad -j(N+j+1)f_{j-1}(x) \end{cases} \tag{15}$$

The Chebyshev polynomials verify the orthogonality property :

$$\sum_{i=0}^{N} f_j(x_i) f_k(x_i) = 0 \quad \text{if} \quad j \neq k. \tag{16}$$

4.1.2 Sampling polynomials

The Chebyshev polynomials must be normalized. We used the sampling polynomials:

$$p_j(t/\tau_0) = \frac{f_j(t/\tau_0)}{\sum_{i=0}^{N} f_j^2(t_i/\tau_0)} \tag{17}$$

where τ_0 is the sampling period and $t_i = i\tau_0$.

Since it is necessary to use a huge number of significant digits in order to avoid truncature error, we calculated "only" 2048 polynomials for a sequence of 8192 time error data.

4.2 Properties of the polynomials

4.2.1 General properties

The conditions of 2.1 are verified except the constancy of the bandwidth of the transfer functions. The bandwidth of the estimators must be normalized.

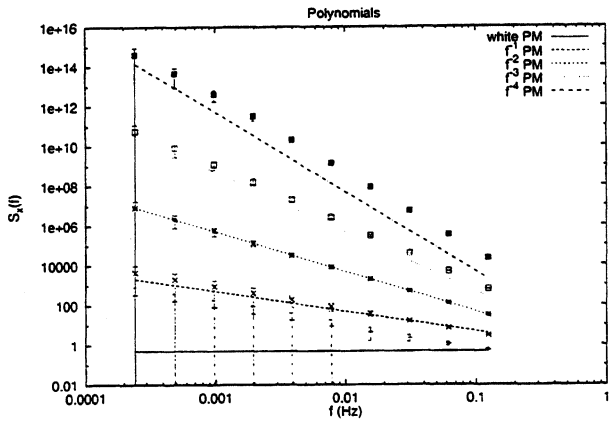


Fig. 4. Graph of the polynomial responses versus the frequency for the different types of noise. This method is not suitable for white PM (slope $\approx f^{-1}$ and large dispersion of the estimates).

4.2.2 Special properties

Thanks to the orthogonality of the polynomials, we just have to remove the first 3 polynomials for ensuring insensitivity to drifts. These estimators converge from f^{-6} PM to white PM.

4.3 Normalization of the bandwidth

Since we only need one estimate per octave, but with a constant bandwidth, we used a weighted average of the "rough estimates" :

$$\left\langle \widehat{S_x}(f_j) \right\rangle = \frac{1}{K} \sum_{k=0}^{2j} \sin^2\left(\frac{\pi}{2}\frac{f_k}{f_j}\right) \widehat{S_x}(f_j) \tag{18}$$

where $K = \sum_{k=0}^{2j} \sin^2\left(\frac{\pi}{2}\frac{f_k}{f_j}\right)$.

However, the "averaged estimates" are no longer independent.

4.4 Distribution of the estimates

If $x(t)$ is a gaussian process, the rough estimates are χ^2 distributed with 1 degree of freedom.

The degrees of freedom of the averaged estimates are about the half of the number of averaged rough estimates.

$S_x(f)$	$\left\langle \widehat{S_x}(f_j) \right\rangle$	Bias	
k_0	-	-	Confusion with f^{-1} PM
$k_{-1}f^{-1}$	$0.379k_{-1}f^{-1}$	0.758	
$k_{-2}f^{-2}$	$0.499k_{-2}f^{-2}$	0.998	
$k_{-3}f^{-3}$	$1.232k_{-3}f^{-3}$	2.464	
$k_{-4}f^{-4}$	$5.453k_{-4}f^{-4}$	10.906	large bias

TABLE II

RESPONSE OF THE POLYNOMIALS FOR THE DIFFERENT TYPES OF NOISES. THE LOW FREQUENCY NOISES ARE QUITE OVERESTIMATED (BIAS > 10).

From a Monte-Carlo method ($N = 8192$), we obtained a number of degrees of freedom between 1 (for the lowest frequency) and 1200 (for the highest frequency).

5. THIRD ESTIMATION: USING A FFT ALGORITHM

The FFT is designed to give a sharp analysis of a spectrum. Due to this sharpness, the PSD of a random process estimated with a FFT exhibits a high dispersion versus the frequency. However, since we only need one estimate per octave, we used the normalization of the bandwidth described in 4.3. This estimation method is called the averaged FFT.

5.1 Properties of the averaged FFT

5.1.1 General properties

Unlike the classical FFT, the relative bandwidth of the averaged FFT is constant, but the estimates are a bit correlated. The other properties of 2.1 are verified.

5.1.2 Special properties

In order to ensure convergence and insensitivity to drifts, we used the second derivative of the data. Then, the estimates must be multiplied by $16\pi^4 f^4$ in order to retrieve estimates of $S_x(f)$.

5.2 Distribution of the estimates

If $x(t)$ is a gaussian process, the rough estimates are χ^2 distributed with 2 degrees of freedom (real part and imaginary part).

The degrees of freedom of the averaged estimates are about the half of the number of averaged rough estimates.

From a Monte-Carlo method ($N = 8192$), we obtained a number of degrees of freedom between 1 (for the lowest frequency) and 1200 (for the highest frequency).

6. COMPARISON OF THE METHODS

Versatility: only P-var is able to distinguish all types of noise.

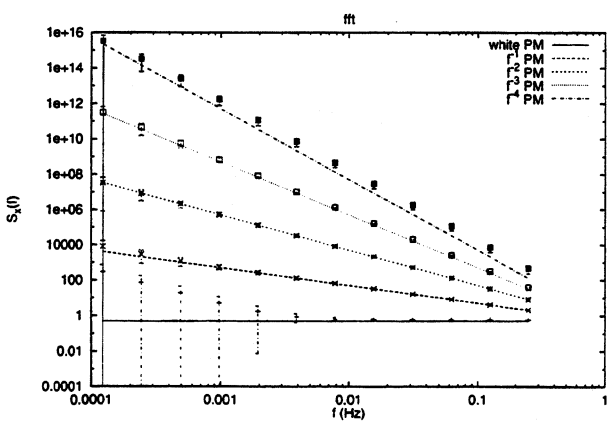


Fig. 5. Graph of the averaged FFT responses versus the frequency for the different types of noise. We may notice a confusion between white PM and f^{-1} PM for the low frequencies and a large dispersion for f^{-4} PM.

$S_x(f)$	$\langle \widehat{S_x}(f_j) \rangle$	Bias	
k_0	$0.567k_0$	1.133	Confusion with f^{-1} PM for low f
$k_{-1}f^{-1}$	$0.500k_{-1}f^{-1}$	1.000	
$k_{-2}f^{-2}$	$0.500k_{-2}f^{-2}$	1.000	very low bias
$k_{-3}f^{-3}$	$0.610k_{-3}f^{-3}$	1.220	
$k_{-4}f^{-4}$	$1.66k_{-4}f^{-4}$	3.32	large dispersion

TABLE III

RESPONSE OF THE AVERAGED FFT FOR THE DIFFERENT TYPES OF NOISES. THE LOW FREQUENCY NOISES ARE A BIT OVERESTIMATED (BIAS > 3).

Accuracy: P-var and the averaged fft are equivalent (except for white PM and f^{-4} PM).

Measurement range: the PSD is estimated over $1 + \ln(N/4)/\ln 2$ frequencies for P-var and the averaged fft (shifted to the low frequencies for the averaged fft) and $1 + \ln(N/16)/\ln 2$ for the polynomial method (2 measures less).

Calculation time: the averaged fft method is the quickest one.

Thus, P-var seems to be the most reliable method of these three ones.

The averaged FFT method, which is almost equivalent, may be used for a quick look to the PSD, thanks to its fast and easy implementation.

But only the polynomial method may be used with unequally spaced data.

REFERENCES

[1] F. Vernotte and M. Vincent, "Estimation of the uncertainty of a mean frequency measurement," in *11th European Frequency and Time Forum*, Neuchâtel, Suisse, Mars 1997, pp. 553–556.

[2] F. Vernotte and M. Vincent, "Estimation of the measurement uncertainty of the drift coefficients versus the noise levels," in *12th European Frequency and Time Forum*, Varsovie, Pologne, Mars 1998, pp. 222–227.

[3] J. E. Deeter and P. E. Boynton, "Techniques for the estimation of red power spectra. i. context and methodology," *The Astrophysical Journal*, vol. 261, pp. 337–350, October 1982.

[4] W. C. Lindsey and C. M. Chie, "Theory of oscillator instability based upon structure function," *Proceedings of the IEEE*, vol. 64, pp. 1652–1666, December 1976.

[5] J. Rutman, "Characterization of phase and frequency instabilities in precision frequency sources: Fifteen years of progress," *Proceedings of the IEEE*, vol. 66, pp. 1048–1075, September 1978.

[6] F. Vernotte, M. McHugh, and G. Zalamansky, "Cut-off frequencies and noise power law model of spectral density: Adaptation of the multivariance method using the structure function approach," in *9th European Frequency and Time Forum*, Besançon, France, Mars 1995, pp. 373–376.

[7] F. Vernotte, G. Zalamansky, M. McHugh, and E. Lantz, "Cut-off frequencies and noise power law model of spectral density : adaptation of the multi-variance method for irregularly spaced timing data using the lowest mode estimator approach.," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 43, no. 3, pp. 403–409, 1996.

Limited Live-time Measurements of Frequency Stability *

D. A. Howe and E. E. Hagn †
National Institute of Standards and Technology, 325 Broadway, Boulder, CO 80303

1. Abstract

For common FM noises in oscillators, discontinuous measurements with dead-time, small sample statistics, and an assumed χ^2 distribution, the RMS frequency fluctuations *vs.* time-interval may be the only reportable measure of frequency stability. We show in a group of simulation trials that in a typical experimental scenario, the RMS frequency stability can seriously underestimate true flicker frequency (FLFM) and random-walk frequency (RWFM) noise and can have larger long-term uncertainty with respect to the zero dead-time Allan deviation and should not be reported as Allan deviation.

2. Introduction and Summary

This paper offers insights into the problem of interpreting the RMS frequency stability with limited live-time frequency measurements, that is, measurements having dead-time of more than 30%. Consider a pair of clocks or oscillators being compared which run continuously, however, measurement live time is a τ averaged frequency difference but the duration between these averages is T and thus no measurement is made during a dead time of $T - \tau$. A result is summarized as the square root of the mean-square of the frequency differences with a corresponding duration T. Thus the root-mean-square frequency error is a measurement of RMS frequency change *vs.* T and is presumed to be a suitable substitute for $\sigma_y(\tau)$, the Allan deviation. We show in a group of simulation trials involving common FM noises, dead-time, small sample statistics, and an assumed χ^2 distribution that this procedure underestimates the FLFM and RWFM noise level.

Measurements of frequency stability with dead-time are biased relative to the zero-dead-time Allan deviation for stochastic noise, and the usual relationship to the fractional-frequency spectral density is often lost. Although dead-time $T-\tau$ bias for given FM noise types can be removed in principle, it is often not removed because this is tedious and is easily susceptible to guesswork regarding the validity of important assumptions. For example, the selection of noise type to use for determining bias is obscured by degraded confidence intervals, a major problem in the presence of FLFM and RWFM.

Results may be reported as an RMS frequency change *vs.* T instead of as an Allan deviation plot. Even though this characterization of frequency stability might have been determined using limited live-time measurements, it may still be useful for some applications. However, it is not the Allan deviation and usually is not a fair estimate of characteristic oscillator FM noise level.

3. RMS Frequency Stability: $\psi_y(\tau, T)$

We define a generic two-sample frequency variance and compare it to the two-sample Allan variance. Frequency instability is generally regarded as an uncertainty on an oscillator's expected or predicted average frequency. At long averaging times, the dominant component of frequency prediction is often the error due to linear frequency drift. In general, a sample estimate of linear frequency drift between two oscillators is $\frac{\Delta^\tau \overline{y}(t)}{T}$ where $\Delta^\tau \overline{y}(t)$ is a *change* in a pair of measured values of frequency offset $\overline{y}_n$ and $\overline{y}_{n+1}$ separated by T, the span of time over which the change occured. τ is the averaging time used to compute each value of frequency offset. The Allan deviation $\sigma_y(\tau)$ is $(\Delta^\tau \overline{y}(t))_{rms}$ and division by a time interval τ is implied because *adjacent* values of $^\tau\overline{y}(t)$ must be used by definition, making $T = \tau$ [1]. Thus the Allan deviation can be interpreted as an uncertainty of a τ-sample estimate of systematic linear frequency drift. Some experimentalists measure quantities such as "rms frequency deviation," labeled as $(\Delta f)_{rms}$ *vs.*T and mean the RMS of measured values of $\left(\frac{\Delta \overline{y}(t)}{T}\right)$. The fact that $\tau < T$ is not an important consideration or is constrained by other measurement factors. In otherwords, some experimentalists often desire the same drift uncertainty measure as the Allan deviation, so compute $\left(\frac{\Delta \overline{y}(t)}{T}\right)_{rms}$. We will define this experimentalist's statistic more carefully as the

†E-mail: dhowe@nist.gov and hagn@nist.gov

square root of $\psi_y^2(\tau,T)$, or psi-variance. Its square root, psi-deviation $\psi_y(\tau,T)$ is often taken to be the root Allan variance $\sigma_y(T=\tau)$ thinking that this function is close enough to serve as a counterpart. At the very least, it is easily confused with the Allan deviation $\sigma_y(\tau)$ which strictly is $(\Delta^{\tau}\overline{y}(t))_{rms}$ *associated* with $T \neq \tau$ and corrected based on ratio $r = \frac{T}{\tau}$. The effect of dead-time is always somewhat obscured by the restriction that the ratio r of dead-to-live time be constant. This restriction is necessary so that integer power-law noise processes have a constant slope (on a log-log scale) corresponding to a constant slope in the frequency domain [1,2]. $\psi_y^2(\tau,T)$ is a convenient experimental characterization in which the live measurements are a constant τ and the dead time is a free parameter $T-\tau$.

We show in simulation that psi-variance $\psi_y^2(\tau,T)$ is a suitable substitute for $\sigma_y^2(T)$ in the presence of white frequency (WHFM) noise but that its level is always too low with FLFM and RWFM noises which are likely to occur in long term. Coupled with the fact that these same long-term estimates may derive from only one or two measurements, they are subject to a negatively skewed probability distribution, hence the reported frequency stability may be significantly low.

4. Mathematical Details and Dilemma

Measurement samples of the time-error function $x(t)$ occur at a rate f_s having an interval $\tau_0 = \frac{1}{f_s}$. Given a sequence of time errors $\{x_n : n = 1,\ldots,N_x\}$ with a sampling period between adjacent observations given by τ_0, we define the $m\tau_0$-average fractional-frequency deviate as

$$\overline{y}_n(m) \equiv \frac{1}{m}\sum_{j=0}^{m-1} y_{n-j},$$

where $y_n = \frac{1}{\tau_0}(x_n - x_{n-1})$. Define psi-variance

$$\psi_y^2(\tau,T) = \frac{\tau}{T} < [{}^{\tau}\overline{y}(t) - {}^{\tau}\overline{y}(t-T)]^2 >, \qquad (1)$$

where $< \cdot >$ denotes an ensemble average and ${}^{\tau}\overline{y}(t)$ is the mean frequency over duration $\tau = m\tau_0$. Thus, in terms of $x(t)$, ${}^{\tau}\overline{y}(t) = \frac{x(t)-x(t-\tau)}{\tau}$. Definition (1) is based on taking sequential mean frequency measurements spaced T apart, differencing them, and computing the mean square. Figure 1 shows the sampling function associated with $\psi_y^2(\tau,T)$ acting on $\{y_n\}$. τ is called the averaging time and $T-\tau$ is the measurement dead time. $\psi_y^2(\tau,T)$ becomes twice the two-point standard (Allan) variance if $\tau = T$ [1]. $\psi_y^2(\tau,T)$ expressed in terms of the time-error function $x(t)$ is

$$\psi_y^2(\tau,T) = \frac{1}{\tau T}\psi_x^2(\tau,T), \qquad (2)$$

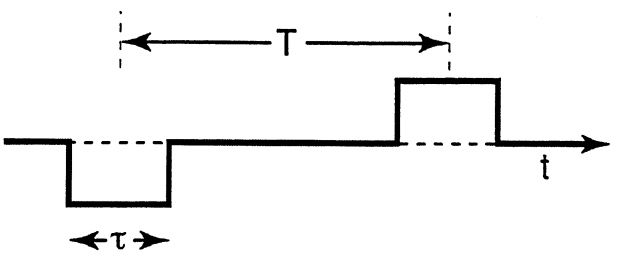


Figure 1: Measurement sequence for mean-square measurements of frequency stability $\psi_y^2(\tau,T)$.

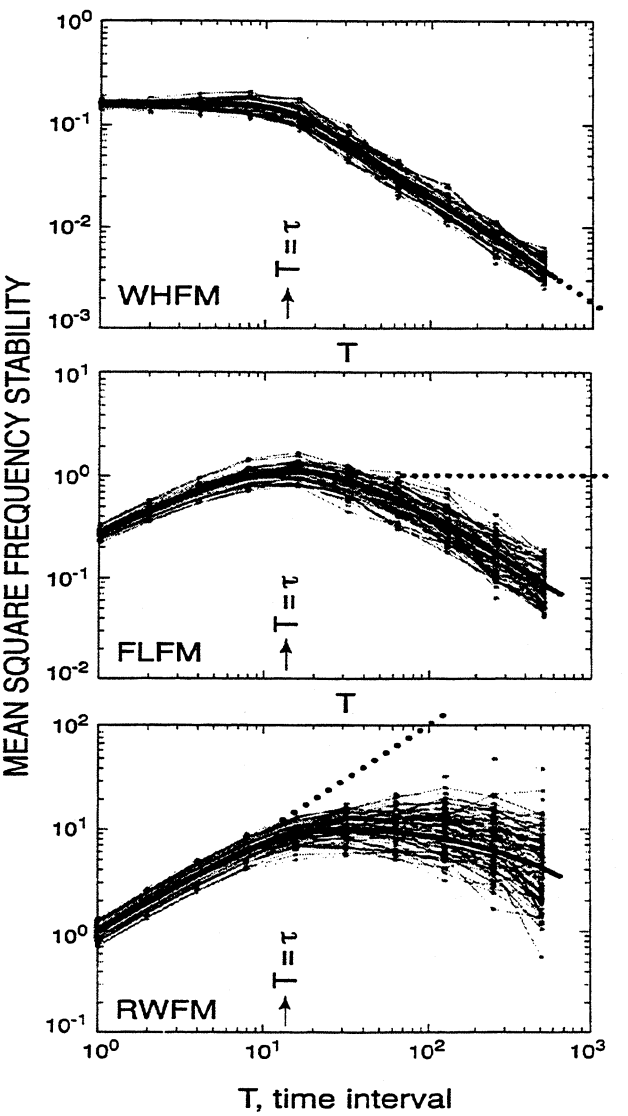


Figure 2: Mean-square frequency fluctuations ψ_y^2 with respect to T of 100 simulations of FM noises whose mean is the solid curve. $T > \tau$ to the right of the arrow are limited live-time measurements, dead time $= T - \tau$. Dashed line is the expected result if the zero dead-time Allan variance were used.

where $\psi_x^2(\tau,T) = < \{[x(t+\frac{\tau}{2}) - x(t-\frac{\tau}{2})] - [x(t-T+\frac{\tau}{2}) - x(t-T-\frac{\tau}{2})]\}^2 >$, the mean-square time error for Doppler radar [1] rewritten as a central difference. (We are not relating this writing to Doppler radar.) Thus,

$$\psi_y^2(\tau,T) = \frac{1}{\tau T} \langle \{[x(t+\tfrac{\tau}{2}) - x(t-\tfrac{\tau}{2})] - [x(t-T+\tfrac{\tau}{2}) - x(t-T-\tfrac{\tau}{2})]\}^2 \rangle . \quad (3)$$

How do we interpret "$\frac{1}{\tau T}$" in (3) when converting from a mean-square time error $\psi_x^2(\tau,T)$ to a mean-square frequency error $\psi_y^2(\tau,T)$? Terms in (3) can be rearranged and written as

$$\psi_y^2(\tau,T) = \frac{1}{\tau T} \langle \{[x(t+\tfrac{\tau}{2}) - x(t-T+\tfrac{\tau}{2})] - [x(t-\tfrac{\tau}{2}) - x(t-T-\tfrac{\tau}{2})]\}^2 \rangle . \quad (4)$$

This can be interpreted to mean that frequency differences that are *overlapped* by $T-\tau$, squared, averaged, and divided by T produce the same value as frequency differences that are *separated* by $T-\tau$, squared, averaged, and divided by τ. In the frequency domain, (3)-(4) act on the same underlying spectrum with two passband filters. The responses of (3) and (4) are identical but interpreted as filters centered at two different Fourier frequencies corresponding to $f_\tau = \frac{1}{2\tau}$ and $f_T = \frac{1}{2T}$.

Averages of frequency error are often made with constant τ and varying T, the result being reported as an RMS frequency error corresponding to T. At this point, there are several intricate issues involved in estimating FM noise level. For example, there is a cutoff frequency f_c between a maximum of f_s (the sampling frequency) and $\frac{1}{2m\tau_0}$ (half the reciprocal of the τ average used in determining $\overline{y}_n(m)$). There is also a measurement-system high-frequency cutoff f_h which must be accounted for. Most of the determinations of FM noise level with dead time depend on an assumption that this high-frequency cutoff f_h is modeled as a rectangular or "brickwall" filter. In theory, the relationship among the parameters f_τ, f_T, f_s, f_c, and f_h is known, but calculating noise levels and simultaneously verifying that the assumptions remain valid for a given experiment can make for an arduous task [2–5].

5. Simulation Study

For more direct insights into statistical errors from dead time, we used simulation trials acting on defined FM noises (WHFM, FLFM, and RWFM). We are in a position with (3) to measure the effect on $\psi_y^2(\tau,T)$ of varying T with respect to a constant τ. Figure 2 shows the result of $\psi_y^2(\tau,T)$ in simulation trials in which T is varied with a fixed value of τ indicated at the arrow, that is, where $T=\tau$. The arrow points to $\psi_y^2(\tau,T) = 2\sigma_y(\tau)$, or twice the Allan variance. Now let $r = T/\tau$. To the right of the arrow are r-ratios corresponding to dead-time measurements $T > \tau$. Notice that if T is reported as if it were "τ" used in the Allan variance, then values will be underneath the expected τ^μ slopes indicated by the dashed lines with FLFM and RWFM. For WHFM, the result is not biased low, but the problem is that it would be an easy mistake to *interpret* the slope as WHFM even though the characteristic noise type is actually FLFM, or FLFM if the noise type is RWFM.

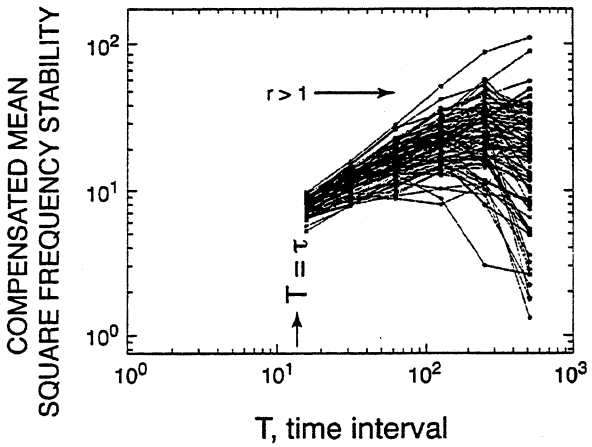


Figure 3: Compensated mean-square frequency fluctuations with respect to T of 100 simulations of RWFM noise (compensated for dead-time $r > 1$ bias). Note that the spread in the distribution is significant and can often indicate two possible FM noise types.

6. Correcting for Dead-time Bias

The proper procedure in experimental design is to fix τ and then adjust the assocated $\psi_y^2(\tau,T)$ result to remove bias in order to estimate $\sigma_y^2(\tau = T)$ due to τ. Then translations to the FM noise coefficients h_α are straightforward [6]. But if τ cannot be held constant, then an adjustment must be made for each estimate and all estimates subsequently averaged, a potentially tedious task.

The bias is determined from the fact that the variance of the process function $y(t)$ increases in proportion to $\ln T - \ln \tau$ and $T-\tau$ for FLFM and RWFM respectively and can be expressed in terms of r [1]. Compensating for this, we have

Flicker-y FLFM: $\psi_y^2(\tau = T) \propto (1 + 3\ln r) \cdot \psi_y^2(\tau,T)$,

for $1 < r < \sim 100$ and

Random-walk-y RWFM: $\psi_y^2(\tau = T) \propto r \cdot \psi_y^2(\tau,T)$.

Applying this compensation to the dead-time region $r > 1$ to simulation trials of RWFM noise yields the results illustrated in figure 3. Now T indeed can be used as "τ" in the Allan variance but there remains the problem of the spread in the distribution. For accurate measurements of long-term FM noise level, we may be unable to determine FM noise type because as many as all three noise types may be consistent with the error bar uncertainty at long term associated with a given measurement.

7. Uncertainty of Estimation

Dead time is likely to be a serious issue when we can make only a limited number of measurements. Lesage and Audoin were the first to theoretically consider the problem of small samples on estimating the Allan variance [8]. Yoshimura showed that the probability distribution in cases with dead time $r > 1$ is governed by χ^2 statistics for the FM noises considered here (but not PM noises) [9]. This means that in the case of limited numbers of measurements, the results are more likely to be below the underlying characteristic oscillator noise level than above it. In the case of one sample as in the very longest T value, the result is more than twice as likely to be too low. Furthermore, its most probable value is 0.

8. "First Results" are Likely to be Excellent

When faced with significant periods of dead time between separated measurements of average frequency difference between two oscillators, it is common to assign a confidence interval to each frequency measurement and to call the absolute value of the overall frequency difference an estimate of the Allan deviation. In this case, should frequency stability still be reported as an Allan deviation? No, because the uncertainty of $\psi_y(\tau, T)$ with significant dead time and only one or two estimates is usually greater than the uncertainty on each frequency measurement. Depending on the assumed power-law noise type, the uncertainty on each frequency measurement will usually underestimate the uncertainty of $\psi_y(\tau, T)$ [10]. Even if the ratio of live to dead time is held constant, long-term uncertainty is a major concern in noise typing. Should frequency stability be reported as an Allan deviation with bias removed? No, because this study shows that with limited live-time measurements, the RMS frequency stability in long-term is likely to fall below the actual noise level, and we can no longer reliably judge the noise type based on slopes of frequency stability *vs.* τ and finally, cannot feasibly correct for this because of increased uncertainty.

9. References

[1] J.A. Barnes, A.R. Chi, L.S. Cutler, D.J. Healy, D.B. Leeson, T.E. McGunigal, J.A. Mullen, Jr., W.L. Smith, R.L. Sydnor, R.F.C Vessot, G.M.R. Winkler, "Characterization of frequency stability," *IEEE Trans. Instrum. Meas.*, **IM-20**, 105–120, 1971.

[2] J.A. Barnes, "Tables of Bias Functions, B_1 and B_2, for Variances Based on Finite Samples of Processes with Power Law Spectral Densities," *Nat. Bur. Stands. (U.S.) Technical Note*, **TN-375**, January 1969.

[3] J. Rutman, "Characterization of frequency stability: A transfer function approach and its application to measurements via filtering of phase noise," *IEEE Trans. Instrum. Meas.*, vol. IM-23, pp. 40-48, Mar. 1974. Also J. Rutman, "Characterization of Phase and Frequency Instabilities in Precision Frequency Sources: Fifteen Years of Progress," *Proc. IEEE* **66**, 1048–1075 (1978).

[4] P. Lesage and C. Audoin, "Effect of Dead-Time on the Estimation of the Two-Sample Variance," *IEEE Trans. Instrum. Meas.*, **IM-28**, 6–10, Mar. 1979.

[5] F. Vernotte, G. Zalamansky, and E. Lantz, "Time stability characterization and spectral aliasing, Parts I and II," *Metrologia*, **35(5)**, 723–738, 1998.

[6] D.B. Sullivan, D.W. Allan, D.A. Howe, and F.L. Walls (Editors), *Characterization of Clocks and Oscillators*, Natl. Inst. Stand. Technol. Technical Note 1337, 1990.

[7] E.S. Ferre-Pikal, J.R. Vig, et al., "Draft Revision of IEEE Std 1139-1988: Standard Definitions of Physical Quantities for Fundamental Frequency and Time Metrology—Random Instabilities," *Proc. 1997 IEEE Int. Freq. Cont. Symp.*, 338–357 (1997; in revision).

[8] P. Lesage and C. Audoin, "Characterization of Frequency Stability: Uncertainty due to the Finite Number of Measurements," *IEEE Trans. Instrum. Meas.*, **IM-22**, 157–161, June, 1973.

[9] K. Yoshimura, "Characterization of Frequency Stability: Uncertainty Due to the Autocorrelation of the Frequency Fluctuations," *IEEE Trans. Instrum. Meas.*, vol. 27, 1–7, Mar. 1978.

[10] F. Vernotte, "Estimation of the uncertainty of a Mean Frequency Measurement," *Proc. 11th European Frequency and Time Forum*, pp. 553-556, March 1997.

PHASE NOISE EVALUATION DIGITAL ALGORITHM FOR PRECISION OSCILLATIONS

D. V. Bogomolov, D. S. Ochkov

JSC "Radiophyzika", Geroev Panfilovtsev 10, Moscow123363, Russia
dvb@atom.ru

Abstract

The accuracy increasing for digital algorithms of noise phase modulation calculations is urgent as for some measurements and especially for processing of signals which have been obtained by means of Computer Aided Design software. Calculation accuracy limits an application of digital algorithms for precision oscillations research. This paper considers accuracy of phase noise calculation common methods when precision oscillation is described by a single sequence (without quadrature components) of digital samples in time domain and presents the author's algorithm for this purpose. The algorithm is based on a signal points tracking by means of the signal interpolation in a vicinity of these points to determine the time moment when oscillation crosses an ordinates' axis. The calculation accuracy can be previously set and can provide an error less then 0,01 rd without large computation expenses. The algorithm output sequence has a minimum dimension for researched phase fluctuation description (two samples within oscillation period). Therefore it allows to obtain phase noise spectral density floor equal to a minus 160 dBc/Hz . The calculated sequence of output phase noise samples allows to monitor both short-term phase fluctuation and slow drift of a phase.

Introduction

This work deals with the design of a numeric algorithm of precision oscillations phase noise calculation when oscillation is presented as a one sequence of digital samples in time domain. Let the digital sequence represent initial oscillation, which can be expressed as:

$$S(t)=A_0[1+\varepsilon(t)]\sin(2\pi f_0 t + \varphi(t)), \quad (1)$$

where $\varepsilon(t)$ - random process describing amplitude A_0 disturbers, $\varphi(t)$ - random process describing phase modulation, f_0 - oscillation frequency (carrier).

This task of function $\varphi(t)$ calculation can be solved with the following methods:

1) Digital simulation of the quadrature signals transformation [1];

2) Construction an additional quadrature sequence for digital signal samples on the basis of an analytical signal model with using of Gilbert transformation [2];

3) Method of phase modulation calculation on the basis of characteristic signal points (for example zero) scanning by means of the signal interpolation in a vicinity of samples points.

The first two ways demand additional calculations because quadrature components are calculated for each input sample even there are more than two samples within one period. Two samples are a necessary minimum for one-value representation of oscillation according to the samples theorem. If such samples frequency was applied, the phase noise is represented in spectrum width equal to double frequency of the researched oscillation. It is enough, because (shown in [3]) the noise with higher spectral components doesn't influence on the phase modulation of a signal and can be described as additive amplitude noise.

Methods description

The first method consists of digital simulation of quadrature transformation of the researched oscillation. The fig. 1. explains a procedure of signal transformations.

0-7803-5400-1/99/ $ 10.00

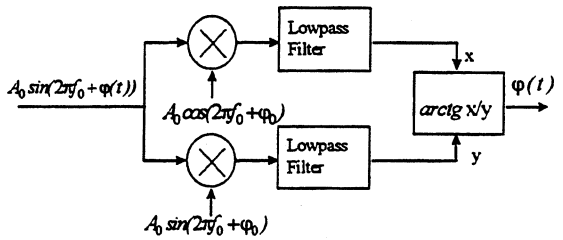


Fig. 1. The phasemeter circuit on base quadrature components obtaining.

The researched signal feeds on two ideal mixers on which reference signals have phase shift 90^0. Output mixers signals pass through lowpass filters. The amplitude constant A_0 of the researched signal doesn't influence on the result of calculations, because the phase modulation function is determined as function *arctg z*, where *z* is equal to the ratio of signals amplitudes in x and y channels.

The accuracy of function $\varphi(t)$ evaluation depends on the out band attenuation of digital lowpass filters. Caused by digital signals multiplication the harmonious components disturb amplitude and phase of investigated oscillation. But in any case the noise spectrum will be deformed by overlapping of noise spectral tails in the considered frequency band. Spectral tails appear from digital signals multiplication (signal's combination and harmonic components can arise at samples frequency band). So, the distortions influence estimation of the noise spectrum is necessary in each particular case and it cannot be carried out in common case. It is necessary to notice, that researches of a sensitivity threshold of such digital phasemeter has been executed in work [1]. Thus, the noise floor of this method realised on a basis Computer Aided Design software has been evaluated of about - 117 dBc/Hz in the best case.

The second method is considered in work [2]. The calculation of function $\varphi(t)$ consists of analytical signal construction with use of Gilbert transformation. If the initial oscillation is a real function $u(t)$, function $v(t)= H[u(t)]$ is constructed as the function coupling with $u(t)$ by Gilbert transformation Entire analytical signal $\chi(t)$ can be written in the following way:

$$\chi(t) = u(t) + iv(t).$$

The integral representation of Gilbert transformation is possible to use for $v(t)$ function calculation it [4]:

$$v(t)=\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{u(s)}{t-s}ds=\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{u(t-s)-u(t+s)}{s}ds \quad (2)$$

where first integral have the sense in the main Koshe value. The numerical integration accuracy of the expression (2) depends on kinds of function $u(t)$ and requires permanent control of calculation errors. So it demands significant computation expenses. The calculation volumes can be reduced, if spectral representation of fluctuation $u(t)$ and Fast Fourier Transformation (FFT) will be used to execute Gilbert transformation [3]. Such way also allows to monitor distortions of a signals spectrum from oscillation discrete representation and from transformation procedure.

This method of noise phase modulation $\varphi(t)$ determination was realised by the authors as the computer program. Small calculation errors can be obtained if whole observation time interval divisible by oscillation period and if there aren't oscillation $u(t)$ amplitude or derivative jumps on the boundaries of samples sequence. When $\varphi(t)$ is random process, it is impossible to eliminate the significant jumps on the sequence boundaries. Even small amplitude modulation of the quadrature component $v(n\Delta t)$ obtained by Gilbert transformation brings significant phase errors. Estimations show, that samples sequence for FFT must have dimension more then 2^{15} - 2^{16} terms to achieve calculation accuracy of function $\varphi(t)$ equal to 0.2 - 0.3 rd.

Algorithm based on interpolation

The authors have investigated a method of phase modulation calculation on the basis of characteristic signal points (for example zero) tracking by means of the signal interpolation in a vicinity of these points. The main question of algorithm construction is an achievement of required accuracy independently of samples number within one oscillation period.

The algorithm includes the following operations:

1) numbers n and $n+1$ of adjacent samples pair with different sine of oscillation amplitude are determined;

2) the signal amplitude corresponding to a middle of a time interval between these samples is calculated according to the formula of the samples theorem

$$u(m\Delta t) = \sum_{k=-\infty}^{\infty} u(k\Delta t)\frac{\sin\omega_S\Delta t(m-k)}{\omega_S\Delta t(m-k)} \quad (3)$$

where ω_S - boundary spectrum frequency of the oscillation $u(t)$, $m=n+1/2$;

3) the kind of phase point (π or 2π) which oscillation had passed is determined by sine of these three samples;

4) linear interpolation oscillation $u(t)$ between the pair of samples which have different sine and more short distance to each other by time;

5) linear interpolation error of time determination when oscillation pass through a zero level is calculated;

6) if an error magnitude exceeds a required value the time interval is being divided in half and all steps from 2 repeat, else the next oscillation crossing through zero well be searching.

We shall consider these operations in more details. We shall denote the oscillation sample which arranges before a point of signal crossing of a zero level as $u_{BF}(n)$, and the first after crossing, as $u_{CN}(n + 1)$. Amplitude of oscillation calculated according to the formula (3), we shall designate as $u_{MD}(n + 1/2)$. It is easy to show, that the calculation error Δ_m not exceeds the magnitude of m + 1 term, if this alternating series will be cut on term number m. As the amplitude of the samples is almost constant the reduction of the series terms (3) with growth of m value is defined by function *sinc (x)*. So we can write following limitation for Δ_m

$$\Delta_m \le \left| u(n) \frac{\sin \pi(m-1/2)}{\pi(m-1/2)} \right| \le \frac{2}{\pi(2m-1)},$$

for $m>1$. (4)

This expression shows, that the error decreases in inverse proportion to amount of the terms in series, and only 7 terms must be summed up to achieve the accuracy of 10 %.

If a discrete step in time domain is less than a quarter of the oscillation period, we can one -valued determine what the kind of phase point (π or 2π) oscillation had passed. In case when the discrete step in time domain is about a half of the oscillation period to calculate the sign of sample $u_{MD}(n + 1/2)$ is necessary. We can one-valued determine the kind of phase point by use of the sign of these three samples. Therefore, in algorithm for any values of discrete step table 1 is used, This table specify demanded conformity.

Table 1

Transition from square	Sign $u_{BF}(n)$	Sign $u_{MD}(n+1/2)$,	Sign $u_{CN}(n+1)$
II → III	+	+	-
II → III	+	-	-
IV→ I	-	-	+
IV→ I	-	+	+

The algorithm is based on linear interpolation between the last pair of calculated signal samples to determine the time moment when oscillation crosses an ordinates' axis. Therefore the error of calculations is determined by a function *sin (x)* deviation from linear function in a vicinity $x=0$. A decomposition of function *sin(x)* in a Tailor series in a vicinity $x=0$ is used for evaluation of calculation error. The deviation δ_n from the linear law is calculated with taking into account the series terms with a degree not exceeding by a third. Value δ_n is limited by a following expression:

$$|\delta_n| < \frac{1}{6}\left(\frac{2\pi\Delta t}{T}\right)^3 < \frac{1}{6}\left(\frac{2\pi}{n_T}\right)^3, \quad (5)$$

where T - researched oscillation period, Δt - discrete step in time domain, n_T - amount of the samples within one oscillation period. Moreover, it is necessary to take into account central symmetry of function *sin(x)* relative to coordinates middle, so calculation errors are being decreased when a pair of samples has almost the same absolute value. The point when oscillation crosses an ordinates' axis is calculated by the following expression:

$$t_{0n} = n\Delta t + \Delta t \frac{u_{BF}(n)}{|u_{BF}(n)| + |u_{MD}(n+1/2)|}. \quad (6)$$

The calculation error equals to zero in case when :

$$|u_{BF}(n)| = |u_{MD}(n+1/2)| \quad (7)$$

Therefore the right side of expression (5) is necessary to be multiplied by following term:

$$\frac{|u_{BF}(n)| - |u_{MD}(n+1/2)|}{|u_{BF}(n)| + |u_{MD}(n+1/2)|}. \quad (8)$$

A simple formula for an error determination of t_{0n} allows to check this value at each iteration of time step reduction and to stop this process when demanded accuracy will be achieved.

The curve of phase modulation function $\varphi(t)$, obtained by this method is represented on fig. 2. The random process which simulates fast phase fluctuations has uniform distribution in a range from -0.05 till 0.05 rd. Bolded line represents calculated function $\varphi(n\Delta t)$ and thin line - initial noise process $\varphi(t)$. Estimations show that about 5 iterations are enough to achieve the accuracy of phase modulation calculation equal to 0.016 - 0.03 rd.

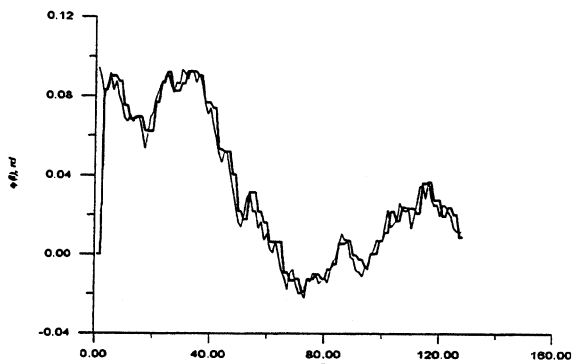


Fig. 2. Initial and calculated function $\varphi(t)$.

Estimation of noise spectral density

We shall estimate spectral density of algorithm noise, arising from finite calculations accuracy of digital algorithm. Unfortunately as it is shown in work [4], the spectral density of noise $S_\varphi(f)$ cannot be analytically expressed through maximum calculations error Δ_c in common case. However we can obtain an estimation of a noise spectral density caused by calculations errors, using model of signal with harmonic phase modulation. Let write the signal in the following kind:

$$u(t) = A\sin(\Omega_0 t + \frac{\beta}{\alpha}\sin\alpha t), \quad (9)$$

where A_{nd} - signal amplitude, Ω_0 - carrier frequency, β - maximum frequency deviation, α - modulating frequency. For such kind of modulation according expression from [5] the spectral density $S_\varphi(f)$ of a signal at α - offset from carrier frequency can be obtained as:

$$S_\varphi(\alpha) = \frac{\eta^2}{2} = \frac{\beta^2}{2\alpha^2}. \quad (10)$$

The value of parameter β on the observation interval T_S can be determined with accuracy equals to $\Delta_c/2\pi T_S$. Substituting at equation (9) parameter $T_S = 10^{-2}$ (it is enough for the control a spectral component in a range 100 Hz - 10 MHz) and $\alpha=10^7$ Hz, we evaluate $S_\varphi(f)$ as:

$$S_\varphi(10^7) < 1.26\ 10^{-16}$$

Such spectral threshold of phase noise calculation of digital algorithm allows to realise sensitivity of phase noise detection at a level of good measuring devices (HP 3048A) and to use it not only for the phase noise analysis of oscillations formed with the help of CAD software, but also for real measurements.

Conclusions

The represented algorithm solves a calculation task of phase modulation caused by noise for precision oscillations which are represented by single (without quadrature components) sequence of digital samples in time domain. The algorithm reflects integrally-derivation dualism of the instantaneous frequency and phase concepts. The additional samples are calculated by summation under the equation (3) (integration). The point when oscillation pass through zero level is determined with help of the increments calculations (derivation). Algorithm has the following advantages: the high calculation accuracy which can be previously set, minimum number of term in an output sequence, small computation and memory expenses.

References

1. S. Gallion, M. Mourey, R. J. Besson, "Comparison of the effects of intermodulation and amplitude to phase convertion in the oscillator phase noise", Proc. of the IEEE International Frequency Control Symposium, Pasadena, U.S. A., May 27-29, 1998.
2. D. E. Vakman, 'Frequency measurement of the analytical signal," Journal of Communications Technology and Electronics, vol. 24, N5, May 1978.
3. Vanshtein L. A., Vakman D. E. Frequency Dividing Method in Oscillation and Waves Theory, "Nayka", Moscow, 1983.
4. J. Rutman, "Oscillator specification: A review of classical and new ideas", Proc. of the 31st Annual Frequency Control Symposium, Jun 1-3, 1977.
5. S.J. Goldman Phase Noise Analysis in Radar Systems. Using Personal Computers, JOHN WILEY and SONS, 1989.

1999 Joint Meeting EFTF - IEEE IFCS

DIRECT APPROACH TO MTIE CALCULATION

Andrzej Dobrogowski, Michal Kasznia
Institute of Electronics and Telecommunications
Poznan University of Technology
60-965 Poznan, Poland
dobrog@et.put.poznan.pl, mkasznia@et.put.poznan.pl
phone ++48 61 8782-293, fax ++48 61 8782-572

ABSTRACT

In the paper the application of time effective algorithms for the Maximum Time Interval Error (MTIE) calculation is considered. In the first section authors introduce the problem of time effective MTIE calculation. In the next section the methods of direct and rough calculation of MTIE estimate are described. Then two time effective methods of MTIE calculation based on the estimator formula are described and used for MTIE calculation. In the experiment several time error sequences obtained from the real measurement process were used.

1. INTRODUCTION

To evaluate the quality of timing signal some parameters are commonly adopted. Among them the maximum time interval error MTIE is considered as very important parameter since it contains information which can be used for dimensioning of the memories (FIFOs) in the telecommunication network circuitry. The limit values of MTIE for the particular types of clocks and hierarchical interfaces in the synchronization network are defined and recommended in the standard [1,2]. Finding the estimate of MTIE within a reasonable period of time is not a simple task. That is the reason we consider it in the paper.

The MTIE calculation directly founded on the MTIE estimator's formula given in international telecommunication standards is dramatically time consuming. Several time effective methods of the MTIE calculation were presented in the literature. Some of them based on statistical nature of time error process enable indirect assessment of the MTIE estimate in much shorter time [3,4]. However, the methods give only the rough estimate of MTIE value and estimation can only be performed for time error samples measured in synchronized clock configuration. Another methods, presented by authors of the paper in [5,6] enable time effective calculation of MTIE estimate based on the parameter's estimator formula. The results of calculation using these methods show the real MTIE value characterizing the analyzed sequence of the time error samples. It was presented [6] that the time effectiveness of the methods depends on the data values. In the extreme situation the MTIE calculation time may be equal or longer than the time of calculation directly based on estimator's formula. The time error sequences considered in [6] were mostly the computer generated series. In the paper we present the MTIE results and their calculation time for several sequences obtained from the real measurement process.

2. MAXIMUM TIME INTERVAL ERROR

The maximum time interval error is defined in international standards as the maximum peak-to-peak time error variation of a given timing signal with respect to an ideal timing signal within a particular time period [1, 2]. If the results of time error function measurements x(t) take the form of N equally spaced samples $\{x_i\}$, MTIE can be estimated from the formula

$$\hat{MTIE}(n\tau_0) = \max_{1 \le k \le N-n} \left(\max_{k \le i \le k+n} x_i - \min_{k \le i \le k+n} x_i \right) \qquad (1)$$

where $\{x_i\}$ is a sequence of N samples of time error function x(t) taken with sampling interval τ_0, $\tau = n\tau_0$ is an observation interval, and n=1, 2, ..., N–1.

Following the formula (1) directly in order to find the estimate of MTIE for the observation interval τ, all intervals with width τ existing in the sequence of N time error samples must be reviewed. The window with width $\tau = n\tau_0$ and embracing n+1 samples is set at the beginning of data sequence $\{x_i\}$ and then it is shifted with the step of τ_0 to the end of the sequence. For each window's location the peak-to-peak value of time error in the window is found. The maximum peak-to-peak value found for all existing locations of the window is the value of MTIE(τ) estimate. The complexity of calculation grows with n and therefore the direct method is time consuming.

3. TIME EFFECTIVE METHODS OF MTIE CALCULATION

3.1. Indirect methods

In the literature [3,4] some methods of MTIE estimate calculation were presented. First method enables the estimation of the percentile of MTIE basing on the standard deviation of the phase noise. It enables the assessment of the MTIE estimate only for the measurement in the synchronized clock configuration. Second method, measurement using disjointed intervals, enables the assessments of the MTIE estimate on line with a course of time error measurement process without the memory consuming storage of the data. This method gives rather rough evaluation of MTIE for the measurement time error sequence. Using this method some errors in the results of MTIE calculation may be present. Some peak-to-peak value of the time error

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

associated with small distance between maximum and minimum values (the MTIE value for small observation interval) may be detected during the measurement for greater observation interval.

3.3. Boundary decision method

In [6] the MTIE calculation method was proposed which reduces the number of window's review operations. The decision on window's review depends on the values of the samples at the window's boundaries at the following locations. Two samples: the earliest value, which leaves out of the window and the new sample, which appears at the window end (Fig. 1) are compared with current maximum and minimum samples. The result of the comparison determines the next operation. The new extreme value should be searched, when the current extreme sample leaves out the window and simultaneously the new value is not the new extreme.

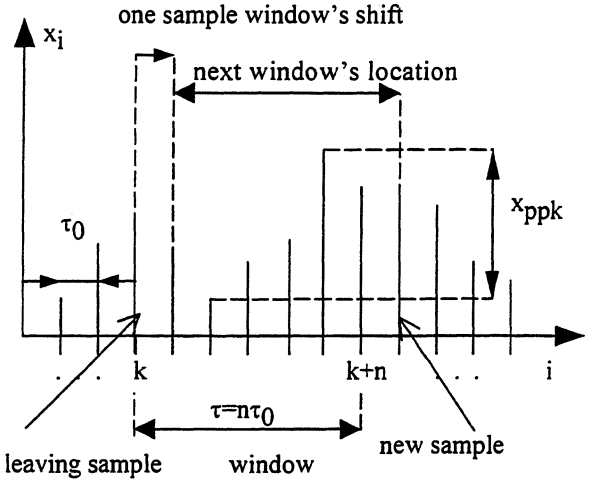


Fig. 1. The window's shift in the boundary decision method

3.3. Extreme fix method

In the another method of MTIE calculation proposed and developed in [6], some window's locations are excluded from inspection, if the peak-to-peak value for these locations is not greater than the value found until now, or if this value may be found for next window's locations. The method called extreme fix method is based on fixing the positions of minimum and maximum samples for a given window's location. After finding the positions of the extremes the window's shift to the position of the first extreme (denoted as p1) is performed (Fig. 2). There are no extreme values in the distance between the starting position of previous window's location and the p1 position. After the shift the peak-to-peak value for the window's location p1 should be found. Because the samples between the position p1 and the last samples in the previous window's location (k+n) were reviewed and the extreme values are known they are excluded from the inspection. All these operations can make the MTIE calculation much faster.

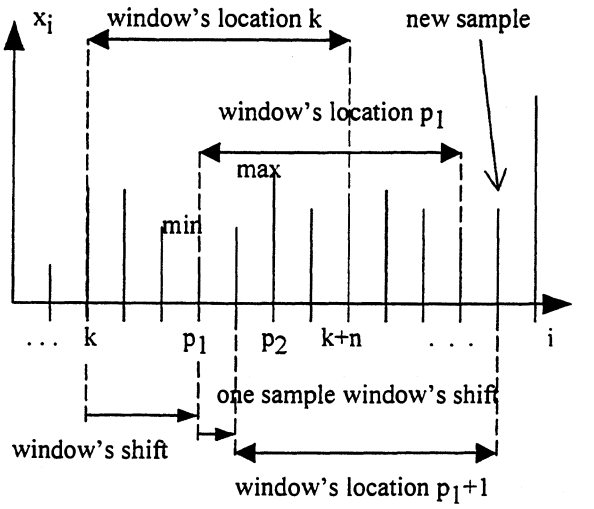


Fig. 2. The window's shift in the extreme fix method

4. EXPERIMENT

It was mentioned that the disadvantage of the methods presented in [6] is the dependence of its effectiveness on the data values. In the experiment described in [6] the computer generated time error sequences were studied. In this paper we will analyze the time effectiveness of the method using the time error sequences obtained from the measurement process.

Two considered sequences (denoted as GPS1 and GPS2) were obtained from the measurement of two different oscillators controlled by the GPS signals (Fig. 3 and 5). The other sequences (denoted as DCF1 and DCF2) were obtained from the measurement of the oscillator controlled by the DCF signal with the GPS disciplined oscillator as a reference (Fig. 7 and 9). The time error samples were taken with the sampling interval τ_0=1/30 s during the period of 4000 s. The fifth sequence was obtained from the measurement of the signal controlled by the DCF (DCF3) with sampling interval τ_0=1/3 s during the period of 40000 s (Fig. 11). Therefore in the every sequence we have the same number of the samples.

In the experiment 81 values of MTIE estimate were computed using the boundary decision method (BD) and the extreme fix method (EF). The observation interval varied from τ=0.1 s to 1000 s for the GPS1, GPS2, DCF1 and DCF2 sequences. For the DCF3 sequence the observation interval took the values from τ=1 s to 10000 s. The results of calculations are presented in Fig. 4, 6, 8, 10 and 12.

In the Table 1 the time of calculation for the whole range of 81 MTIE values and for the decade components of the whole range is presented. In the Table 2 the time of calculation for the chosen observation interval is given. The format of the time is as follows: hours:minutes:seconds (zeros at the beginning are omitted). As the reference the time of calculation using direct method in both tables was included [6]. The PC computer with Pentium II 450 MHz processor was used for all calculation.

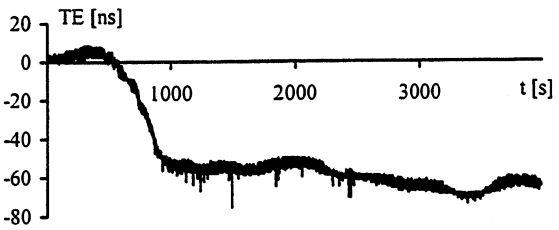


Fig. 3. Time error for the GPS1 signal measurement

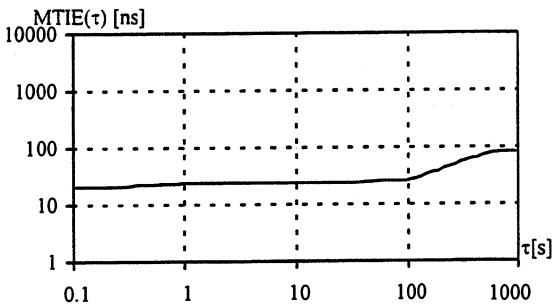


Fig. 4. MTIE for the GPS1 time error sequence

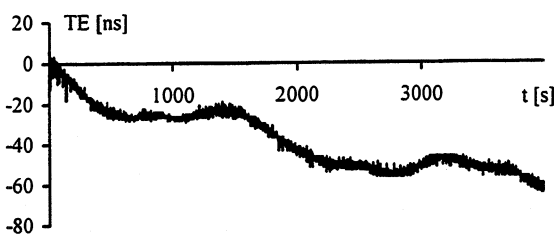


Fig. 5. Time error for the GPS2 signal measurement

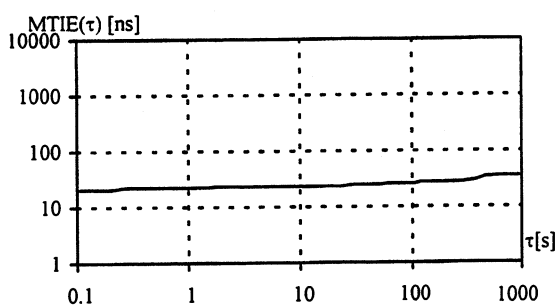


Fig. 6. MTIE for the GPS2 time error sequence

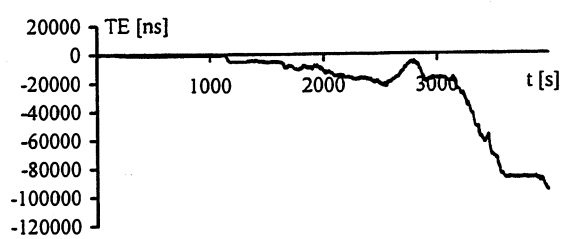


Fig. 7. Time error for the DCF1 signal measurement

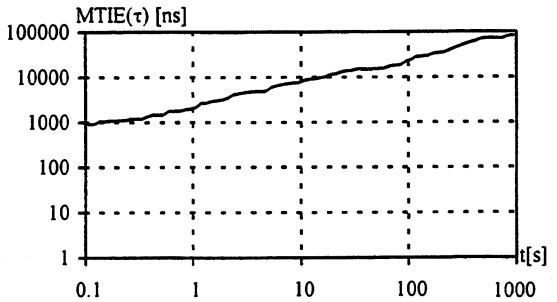


Fig. 8. MTIE for the DCF1 time error sequence

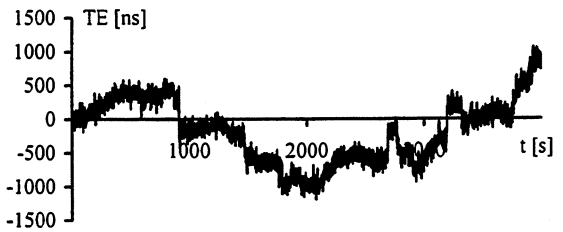


Fig. 9. Time error for the DCF2 signal measurement

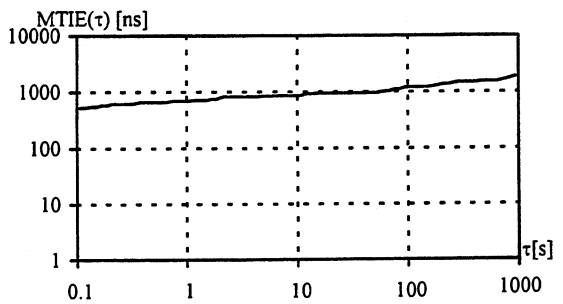


Fig. 10. MTIE for the DCF2 time error sequence

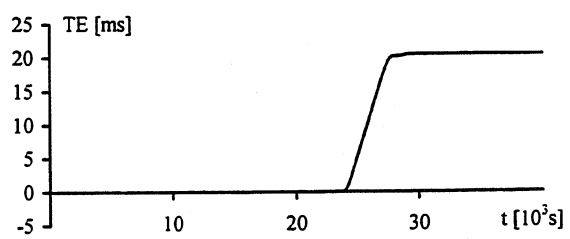


Fig. 11. Time error for the DCF3 signal measurement

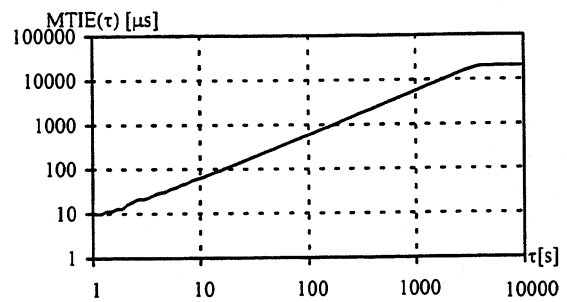


Fig. 12. MTIE for the DCF3 time error sequence

Table 1. Time of calculation for the range of observation intervals

		the range of observation intervals [s]				
TE	method	0.1-1000	0.1-1	1-10	10-100	100-1000
	direct	44:05:45	4:01	34:04	5:25:41	38:01:59
GPS1	BD	1:14:13	3:07	4:30	13:25	53:11
	EF	4:53	1:17	40	41	2:15
GPS2	BD	2:13:36	2:49	4:42	17:11	1:48:54
	EF	4:26	1:15	37	39	1:55
DCF1	BD	1:04.24	2:31	3:25	15:14	43:14
	EF	22:41	1:55	1:43	5:27	13:36
DCF2	BD	11:09	2:23	2:11	2:25	4:10
	EF	7:27	1:54	1:18	1:25	2:50
		1-10000	1-10	10-100	100-1000	1000-10000
DCF3	BD	5:44:22	2:15	4:50	32:37	5:04:40
	EF	3:59:37	1:47	4:05	31.37	3:22:08

Table 2. Time of calculation for chosen observation intervals

		observation intervals [s]				
TE	method	0.1	1	10	100	1000
	direct	4.35	25.32	3:55.31	37:17.11	3:17:35.31
GPS1	BD	7.58	13.68	22.90	1:17.72	4:49.30
	EF	4.45	3.13	1.70	3.02	12.19
GPS2	BD	7.08	9.56	12.19	1:58.73	11:24.18
	EF	5.09	2.31	1.70	2.69	9.45
DCF1	BD	6.37	7.30	4.12	1:29.69	1:12.89
	EF	6.04	4.23	6.86	30.76	32.96
DCF2	BD	6.65	6.74	6.43	8.73	15.22
	EF	6.31	4.45	3.57	5.50	10.27
		1	10	100	1000	10000
DCF3	BD	5.94	7.30	26.64	3:50.22	33:57.16
	EF	5.38	5.39	24.16	3:34.71	16.09

The results presented in Table 1 and 2 show the supremacy of the extreme fix method over the boundary decision method. The time effectiveness of the EF method is visible especially for the long observation intervals. The time of MTIE calculation using direct method is shorter, than the calculation time using both analyzed methods only for observation intervals enclosing small number of samples (four samples in the case presented above). In the other cases the EF and BD methods are better than the direct method. The calculation time extends if in the time error process the frequency offset is present, like in the DCF1 and DCF3 time error sequences.

5. CONCLUSIONS

In the paper the application of two MTIE calculation method was presented. Both considered methods allow to find the MTIE estimate characterizing the analyzed time error sequence. In the experiment the calculations were performed using several time error series. The results of the experiment reveal good time effectiveness of the methods proposed in [6]. Short calculation time was obtained for the time error sequences expressing good quality of the measured clocks (GPS1 and GPS2 time error sequences) or the random behavior of the measured values (DCF2 time error sequence). Some frequency offset of the measured clock (DCF1 and DCF3) makes the MTIE calculation time longer. However the time is still shorter than for the direct method. In such case another method of MTIE calculation may be suitable [6].

The work was performed in the frame of the project TB-44-576.

6. REFERENCES

[1] ETSI Draft ETS DE/TM-3017, "Generic requirements for synchronization networks", Helsinki 1995.

[2] ITU-T Rec. G.810 "Considerations on Timing and Synchronization Issues"

[3] P. Tavella, A. Godone, S. Leschiutta, "The Range Covered by a Random Process and the New Definition of MTIE", Proc. of 28th Precise Time and Time Interval Meeting, Reston, Dec. 1996.

[4] S. Bregni, "Measurement of Maximum Time Interval Error for Telecommunications Clock Stability Characterization", IEEE Trans. on Instrumentation and Measurement, pp. 900-906, vol. 45, No. 5, October 1996.

[5] A. Dobrogowski, M. Kasznia, "The Methods of Calculation of Maximum Time Interval Error", Proc. of Poznan Telecommunications Workshop PWT'97, pp. IV.3-1-IV.3-5, Poznan, Poland, 10-11 December 1997 (in Polish).

[6] A. Dobrogowski, M. Kasznia, "Time effective methods of calculation of Maximum Time Interval Error", submitted for IEEE Transactions on Instrumentation and Measurement

Correlation-Based Noise Measurements Below the Thermal Noise Floor

Enrico Rubiola
Politecnico di Torino, Dip. Elettronica
c.so Duca degli Abruzzi no. 24,
10129 Torino, Italy
e-mail: rubiola@polito.it

Vincent Giordano
LPMO CNRS
32 av. de l'Observatoire
25044 Besançon, France
e-mail: giordano@lpmo.univ-fcomte.fr

Abstract

In the characterisation of the phase noise of a component, it is a common practice to measure the cross spectrum at the output of two phase detectors that simultaneously compare the component output signal to a common reference. This technique, which is based on correlation and averaging, allows the rejection of the phase detector noise. Working in the VHF band with an improved version of the above idea, we realized that, under certain hypotheses, the correlation and averaging mechanism makes the instrument insensitive to thermal noise. This feature makes the instrument suitable to investigate noise phenomena below the thermal floor, but — if not understood — is a potential cause of experimental mistakes. Both theoretical explanations and experimental proof are given.

1 Introduction

Spectral purity of electronic devices is usually described in terms of $S_\varphi(f)$, i.e. the power spectrum density (PSD) of the device phase lag φ. To measure $S_\varphi(f)$, we use the scheme shown in Fig. 1. It is similar to the microwave instrument described in [1], but operates in the VHF band at the carrier frequency $\nu_c = 100$ MHz. Properly setting the detection phases γ'', the voltages $a(t)$ and $b(t)$ are proportional to the phase of the device under test (DUT), and consequently the average cross power spectrum density $S_{ab}(f)$ is proportional to $S_\varphi(f)$. This type of measurement, derived from [2], extracts the shared-path noise and rejects the single-arm noise processes, provided that they are independent. Averaging on m measures, a rejection of the single-arm noise by a factor $\sqrt{m}$ is expected. In our improved scheme we use the interferometric phase detector [3] because it is less noisy than the double balanced mixer. In short, this detector works as a synchronous receiver preceded by a low noise amplifier. For best noise performances, the amplifier works in small signal regime, amplifying only the DUT noise sidebands. This condition is ensured by a carrier suppression mechanism taking place inside the hybrid couplers when the phases γ' are set equal to the DUT phase and the attenuators A_1 and A_2 are set equal to the DUT attenuation ℓ. Details of the interferometric detection technique, together with design guidelines, are given in [4].

The two low noise amplifiers of Fig. 1 are impedance matched, which implies that thermal noise is present at their input. In the absence of the DUT — the latter is replaced by a short cable — all this noise comes from the internal terminations R_1, R_2 and R_3 of the power splitters. These power splitters are 3 dB 90° hybrid couplers with one input connected to a matched resistive load. In fact, as well known, the reactive reciprocal power splitter must be a 3 dB hybrid coupler terminated at one input, otherwise it could not be impedance matched. Because R_1 R_2 and R_3 generate thermal noise shared by the two amplifiers, at first sight one could believe that the thermal noise limits the instrument sensitivity. In reality it is much more complex, as explained in the following sections.

2 Noise Theory of the Double Interferometer

Let us analyse in detail the circuit signals when the instrument measures an attenuator of loss ℓ as the DUT. For this purpose, we define six noise processes indicated with $n_1(t), n_2(t), \ldots, n_6(t)$; each one is the thermal voltage noise available across a resistor of value R_0, which is the characteristic impedance of the whole circuit. n_1, n_2 and n_3 are the available noise voltages across the resistors R_1, R_2 and R_3, respectively; n_4 and n_5 come from the internal resistors of the attenuators A_1 and A_2; n_6 is the thermal noise due to the resistive loss of the DUT. For all these processes, the power spectrum density $N_i(f)$ is equal to $R_0 k_B T_0$, where $k_B = 1.38\times10^{-23}$ J/K is the Boltzmann con-

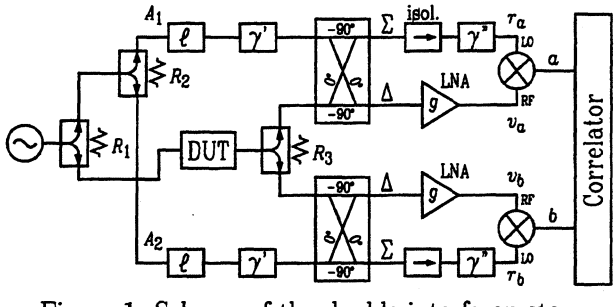


Figure 1: Scheme of the double interferometer.

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

stant and $T_0 = 290$ K the reference temperature. The temperature of the whole instrument is assumed to be close to T_0, uniform and constant. With n_4, n_5 and n_6, the noise is generated by the internal resistors of an attenuator of power loss ℓ. Hence the corresponding voltage available at the attenuator output is $\sqrt{(\ell-1)/\ell}\, n_i$. This occurs because the total output noise at the attenuator output must be $R_0 k_B T_0$ when the attenuator input is terminated to a resistor. Furthermore, we assume that the DUT adds *extra noise* $\check{n}_6$; the latter can be any noise process, including flicker. Under this hypothesis, the DUT output noise is

$$n_{\rm dut}(t) = \sqrt{\frac{\ell-1}{\ell}}\, n_6(t) + \check{n}_6(t) \ . \qquad (1)$$

In the vicinity of the carrier frequency, noise processes can be divided in in-phase and quadrature components as

$$n(t) = n_x(t)\cos(\omega t) - n_y(t)\sin(\omega t) \ . \qquad (2)$$

With thermal noise, the PSDs of the low frequency processes $n_x(t)$ and $n_y(t)$ are $N_x(f) = R_0 k_B T_0$ and $N_y(f) = R_0 k_B T_0$, so that the PSD of the radiofrequency process $n(t)$ is $N(f) = R_0 k_B T_0$. Taking the oscillator signal $\cos(\omega t)$ as the phase reference, the DUT output signal is $\sqrt{2R_0P_c}\sin(\omega t)$, where P_c is the carrier power. This implies that the x noise component is responsible for phase noise; accordingly, $S_\varphi(f) = N_x(f)/R_0P_c$.

For the sake of simplicity, we neglect the loss ℓ_h of the hybrids and power splitters, and the amplifier noise. The former can be introduced later in the equations, and the latter vanish in the correlation function because the two amplifiers are independent.

The reference signals at the mixer LO ports are

$$r_a(t) = -V_p\cos(\omega t) \qquad (3)$$
$$r_b(t) = V_p\sin(\omega t) \ . \qquad (4)$$

Consequently, arm a detects the $\cos(\omega t)$ signal, while arm b detects the $\sin(\omega t)$ component. Due to the circuit phase relationships, the signals at the mixer RF inputs are

$$\begin{aligned} v_a(t) = \sqrt{g}\Bigg[&-\frac{1}{\sqrt{2\ell}}n_{1x} + \frac{1}{2\sqrt{\ell}}n_{2x} + \\ &+\frac{1}{2}n_{3y} + \sqrt{\frac{\ell-1}{2\ell}}\,n_{4x} + \\ &-\frac{1}{2}\sqrt{\frac{\ell-1}{\ell}}\,n_{6x} - \frac{1}{2}\check{n}_{6x}\Bigg]\cos(\omega t) + \\ &+\sqrt{g}\begin{bmatrix}\text{non}\\ \text{detected}\\ \text{terms}\end{bmatrix}\sin(\omega t) \end{aligned} \qquad (5)$$

$$\begin{aligned} v_b(t) = \sqrt{g}\Bigg[&\frac{1}{\sqrt{2\ell}}n_{1x} + \frac{1}{2\sqrt{\ell}}n_{2x} + \\ &+\frac{1}{2}n_{3y} - \sqrt{\frac{\ell-1}{2\ell}}\,n_{5x} + \\ &+\frac{1}{2}\sqrt{\frac{\ell-1}{\ell}}\,n_{6x} + \frac{1}{2}\check{n}_{6x}\Bigg]\sin(\omega t) + \\ &+\begin{bmatrix}\text{non}\\ \text{detected}\\ \text{terms}\end{bmatrix}\cos(\omega t) \ . \end{aligned} \qquad (6)$$

Due to saturation at the LO port of the mixer, the mixer loss ℓ_m must be introduced; in agreement with most databooks, ℓ_m also includes the intrinsic loss due to power splitting into up converted and down converted signals. Hence, after filtering out the 2ω components, the detected signals are

$$\begin{aligned} a(t) = \sqrt{\frac{2g}{\ell_m}}\Bigg[&\frac{1}{\sqrt{2\ell}}n_{1x} - \frac{1}{2\sqrt{\ell}}n_{2x} - \frac{1}{2}n_{3y} + \\ &-\sqrt{\frac{\ell-1}{2\ell}}\,n_{4x} + \frac{1}{2}\sqrt{\frac{\ell-1}{\ell}}\,n_{6x} + \frac{1}{2}\check{n}_{6x}\Bigg] \end{aligned} \qquad (7)$$

$$\begin{aligned} b(t) = \sqrt{\frac{2g}{\ell_m}}\Bigg[&\frac{1}{\sqrt{2\ell}}n_{1x} + \frac{1}{2\sqrt{\ell}}n_{2x} + \frac{1}{2}n_{3y} + \\ &-\sqrt{\frac{\ell-1}{2\ell}}\,n_{5y} + \frac{1}{2}\sqrt{\frac{\ell-1}{\ell}}\,n_{6x} + \frac{1}{2}\check{n}_{6x}\Bigg] \end{aligned} \qquad (8)$$

The correlation function of $a(t)$ and $b(t)$ is

$$\mathcal{R}_{ab}(\tau) = \lim_{\theta\to\infty}\frac{1}{\theta}\int_\theta a(t)\, b^*(t-\tau)\,{\rm d}t \qquad (9)$$

and its Fourier transform is the cross power spectrum density

$$S_{ab}(f) = \mathcal{F}\{\mathcal{R}_{ab}(\tau)\} \ . \qquad (10)$$

Making the noise processes appear explicitly in the latter, all the cross terms vanish. Hence

$$\begin{aligned} S_{ab}(f) = \frac{g}{\ell_m}\Bigg[&\frac{1}{\ell}N_{1x} - \frac{1}{2\ell}N_{2x} - \frac{1}{2}N_{3y} \\ &+\frac{\ell-1}{2\ell}N_{6x} + \frac{1}{2}\check{N}_{6x}\Bigg] \ . \end{aligned} \qquad (11)$$

By definition, $N_i = R_0 k_B T_0$ for all $i = 1\ldots6$. Hence

$$S_{ab}(f) = \frac{g}{2\ell_m}\check{N}_{6x}(f) \qquad (12)$$

Phase noise is related to the radiofrequency spectrum by $S_\varphi(f)=N(f)/R_0P_c$. Combining this with (12), we derive the instrument gain

$$K_\varphi = \frac{S_{ab}(f)}{S_\varphi(f)} = \frac{gR_0P_c}{2\ell_m} \qquad (13)$$

Equations (11) and (12) state that, under the hypothesis of temperature uniformity, the instrument compensates for the thermal noise, and consequently only $\check{N}_{6x}(f)$ contributes to the measured phase noise.

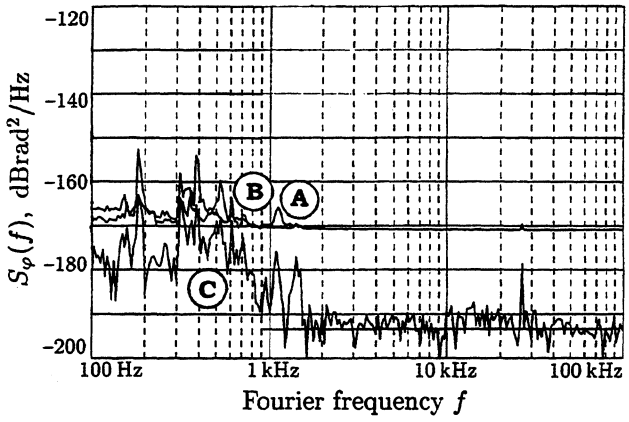


Figure 2: Noise floor of the 100 MHz double interferometer prototype. A and B: single-arm. C: correlation.

3 Experimental Proof

3.1 Noise Floor

We measured the noise floor of a VHF version of the scheme shown in Fig. 1, designed for the carrier of 100 MHz, and described in [5]. In this prototype, the amplifiers show gain $g = 40$ dB and noise figure $F = 2$ dB. The signal power at the mixer LO inputs is 8 dBm, and the DUT power is $P_c = 8$ dBm. The hybrids and the power splitters, that are actually hybrids of the same type terminated at one input, show losses $\ell_h = 0.8$ dB. The mixer loss is $\ell_m = 6$ dB.

Fig. 2 shows the results averaged on $m = 32767$ measures, which is the maximum of the available correlator. The single-arm noise floor (curves A and B), is -172 dBrad^2/Hz, which is close to the expected value $S_{\varphi a}(f) = S_{\varphi b}(f) = 4FkT_0\ell_h^2/P_c \simeq -172.3$ dBrad^2/Hz. The thermal noise, calculated for the same conditions, is $S_{\varphi \mathrm{th}} = k_BT_0/P_c \simeq -182$ dBrad^2/Hz. Yet, the measured floor (curve C) is $S_{\varphi 0} \simeq -194$ dBrad^2/Hz, which is 12 dB lower than the thermal floor. Furthermore, we are still not able to assess the ultimate noise floor of the instrument. In fact, the observed $S_{\varphi 0}$ is 22 dB lower than the single-arm noise; as this difference is close to $\sqrt{m}$, which is the noise reduction due to averaging, there are no reasons to believe that increasing m the measured floor would not further decrease.

3.2 Noise Measurement Below the Thermal Floor

The DUT is now replaced with the circuit shown in Fig. 3. All the settings reported in Section 3.1 are kept except for m, which is reduced when possible. Neglecting the thermal noise N_6 because it is expected to be rubbed out, this circuit injects calibrated noise $\tilde{N}_6(f) = g_aR_0F_akT_0/(\ell_vk_c)$. $\tilde{N}_6$ can be set to the desired value adjusting ℓ_v. The equivalent phase noise thereby injected is $S_{\varphi i}(f) = g_aR_0F_akT_0/(\ell_vk_cP_c)$. Fig. 4 shows the measured S_φ as a function of the injected $S_{\varphi i}$. Going to the left of the figure, ℓ_v increases and the injected noise becomes negligible compared to

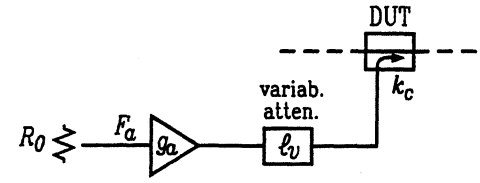


Figure 3: Injection of reference noise $\tilde{N}_6$.

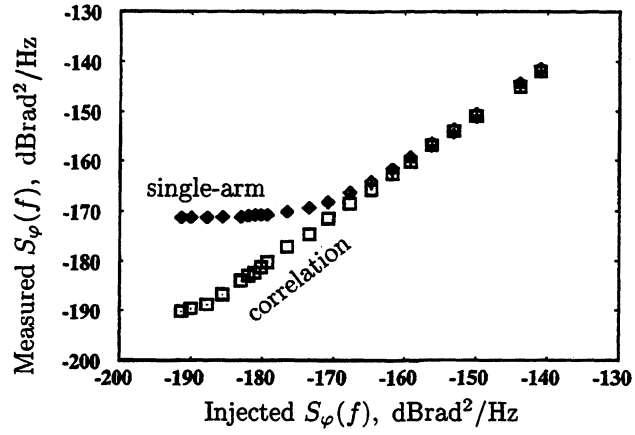


Figure 4: Reference noise measurement.

the equivalent noise at the amplifier inputs. Therefore, the single-arm S_φ approaches the value of -172 dBrad^2/Hz that we measured in the absence of the DUT. By contrast, the correlated noise fits the straight line $S_\varphi = S_{\varphi i}$. Consequently, still under the hypothesis of temperature uniformity, the instrument compensates for the thermal noise and measures the extra noise only.

3.3 Noise of an Attenuator

With the double interferometer, that is a highly sensitive instrument, the common belief that the correlator rejects the single-arm noise and detects the shared noise only is a poor approximation of reality. Accordingly, we analyse in detail the measurement of the noise generated by an attenuator comparing the two configurations shown in Fig. 5.

In Fig. 5A the DUT is replaced by a cable, which is noiseless; two 16 dB attenuators, B_1 and B_2, are inserted at the power splitter outputs. In Fig. 5B these attenuators are replaced by one attenuator of the same value inserted as the DUT. Hence in A the two attenuators are independent noise sources; they also isolate by a factor $\ell = 16$ dB any preceding source of correlated noise. In B the attenuator is a shared noise source. All the experimental conditions are the same for the two configurations, except the detail of Fig. 5. The power splitter driving power, significantly different in the two cases, is not relevant because the reactive power splitters are not power-sensitive; moreover, we choose a device specified for continuous operation up to 36 dBm, while the input power does not exceed 22 dBm. Finally, the oscillator power is +26 dBm and

the residual carrier at the mixer RF input is not higher than -25 dBm.

The experimental results, shown in Fig. 6, are reported in terms of voltage PSD $S_V(f)$ averaged on $m = 1024$ measures. $S_V(f)$ is used instead of $S_\varphi(f)$ because in this case $S_V(f)$ is easier to understand. We focus our attention on the white noise floor, for $f > 1$ kHz. The expected single-arm noise is $S_V(f) = 2gR_oFk_BT_0/\ell_m \simeq -148$ dBV^2/Hz, which is experimentally confirmed with both configurations. With configuration B, the shared noise at the DUT output is $(\ell-1)/\ell N_6 \simeq R_0k_BT_0$. In the absence of the noise compensation, the correlated noise would be $S_{ab}(f) = (gR_ok_BT_0)/(2\ell_h^2\ell_m) \simeq -157.5$ dBV^2/Hz. Yet, the observed noise is significantly lower, -165 dBV^2/Hz, and almost equal for both configurations. Once again, this value is close to $S_{sa}(f)/\sqrt{m}$ that we expect if the correlated noise is negligible. To sum up, the noise compensation mechanism makes the individual arm attenuator indistinguishable from the shared attenuator.

4 Conclusions

It has been theoretically and experimentally shown that the double interferometer scheme, in conjunction with the correlation and averaging mechanism, compensates for the thermal noise. The instrument noise floor is lower than the thermal noise. This feature makes it possible to measure low noise phenomena, well below the thermal floor without the need for cooling. On the other hand, the same fact turns into a potential cause of experimental mistakes in the common laboratory practice. In fact, the thermal noise of the device under test, which vanishes in this measurement process, is still present and reappears when the device is used in a normal application.

The proof given applies to the double interferometer. There are other phase noise measurement systems [2, 6] based on correlation; in these systems, double balanced mixers are used as phase detectors insteasd of the interferometric detectors. Attempts are in progress to show that the high sensitivity of those systems is due to the same thermal noise compensation mechanism here shown.

Acknowledgments

We wish to acknowledge prof. Michele Elia of the Politecnico di Torino for having greatly contributed to the theoretical comprehension of our experiments.

References

[1] E. Rubiola, V. Giordano, J. Groslambert, "Double correlating interferometer scheme to measure PM and AM noise", *Electronics Letters* vol. 34 no. 1 pp. 93–94, January 8th, 1998.

[2] D. Fest, J. Groslambert, J. J. Gagnepain, "Individual characterization of an oscillator by means of cross-correlation or cross covariance method", *IEEE Trans. on IM* vol. 32 no. 3 pp. 447–450, September 1983.

[3] K. H. Sann, "The measurement of near-carrier noise in microwave amplifiers" *IEEE Trans. on MTT* vol. 16 no. 9 pp. 761–766, September 1968.

[4] E. Rubiola, V. Giordano, J. Groslambert, "Very high frequency and microwave interferometric PM and AM noise measurements", *Review of Scientific Instruments* vol. 70 no. 1 pp. 220–225, January 1999.

[5] E. Rubiola, V. Giordano, J. Groslambert, "VHF and microwave interferometric PM and AM noise measurements", Proc. *12th European Frequency and Time Forum* pp. 286–291, Warszawa (Poland), 10–12 March 1998.

[6] F. L. Walls, S. R. Stain, J. E. Gray, D. J. Glaze, "Design considerations in state-of-the-art signal processing and phase noise measurement systems", *Proc. 30th Frequency Control Symposium* pp. 269–274, Atlantic City (NJ, USA) 2–4 Jun 1976.

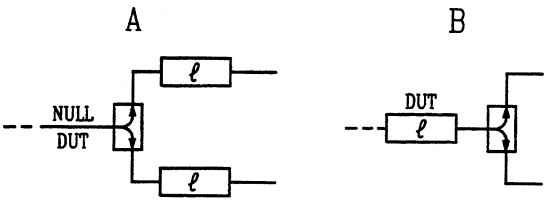


Figure 5: Measurement schemes with the 16 dB attenuator.

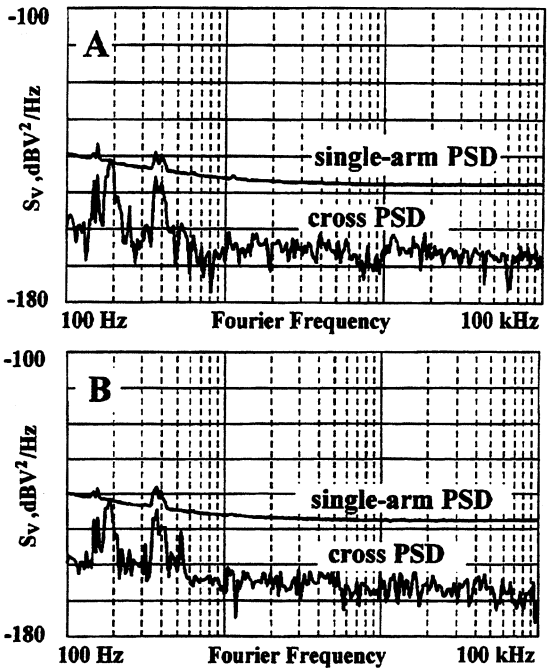


Figure 6: Noise measured with the 16 dB attenuator.

1999 Joint Meeting EFTF - IEEE IFCS

CHARACTERIZATION OF PHASE ERRORS USING GRAY DYNAMIC MODEL

Ching-Haur Chang and Chia-Shu Liao

National Standard Time & Frequency Laboratory
Telecommunication Laboratories of Chunghwa Telecom Co., Ltd.
12, Lane 551, Min-Tsu Road Sec. 5, Yang-Mei, Taoyuan, Taiwan 326, ROC
Tel: +886-3-424-4246 Fax: +886-3-424-5178 Email: changc@ms.chttl.com.tw

ABSTRACT

Characterization of phase errors based on the gray dynamic model is discussed in this paper. The model has been proven useful in the prediction of the response of a system with energy. In this work, we use a model, called GM(1,1), to characterize the phase error between two clocks. Advantage of using the model is that it is capable of predicting future data simply based on a sequence of data samples. Two examples are conducted to justify the value of using the model.

1. INTRODUCTION

The modeling of phase errors is critical in the field of frequency calibration. In this paper, the theory of gray dynamic model, denoted GM(n, h) with n being the order of the system and h being the number of variables in the system, is used to model the phase error between two clocks. It has been shown in [1] that the response of a system with energy is well predictable by the use of the model. According to [1], the response of a system with energy is exponential. In clock modeling, it was shown that the phase error could be described by a second-order systematic model [2]. An exponential model can fit the second-order model from short-term point of view. This is why the gray dynamic model is employed in this study. The advantage of using the model is that it is capable of predicting future data simply based on a set of data samples.

Two examples are conducted in this work. The first one applies the data recorded in our laboratory to see its capability in predicting the phase error between two clocks. The result shows that the maximum error resulted is about 50 % lower than that resulted from the use of least squares fit. The second one applies a set of noisy data to estimate the frequency offset of a clock and, then, to steer the clock based on the estimate. The result shows that a clock with frequency offset on the order of 10^{-13} can be improved to the order of 10^{-14} by steering with the estimate.

2. BASIC THEORY OF GM(1,1) MODEL

Given a finite set of information

$$x^{(0)} = \{x^{(0)}(1),\ x^{(0)}(2),\ \ldots,\ x^{(0)}(N)\}\quad \forall x^{(0)}(k) \geq 0,\ k = 1,2,\ldots,N \qquad (1)$$

with elements spaced at a regular interval, the source model, called GM(1,1) gray dynamic model [1], is described by

$$x^{(0)}(k) + az^{(1)}(k) = b \qquad (2)$$

with $z^{(1)}(k) = (x^{(1)}(k) + x^{(1)}(k-1))/2$ and $x^{(1)}(k) = \sum_{m=1}^{k} x^{(0)}(m)$. To obtain a response equation for (2), a differential equation is used as the

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

replacement. The differential equation is of the form

$$\frac{dx^{(1)}(t)}{dt} + ax^{(1)}(t) = b \tag{3}$$

Solving for (3) yields

$$x^{(1)}(t) = e^{-at}(be^{at}/a + c) \tag{4}$$

where c is a constant. Substituting the initial condition of $x^{(1)}(1) = x^{(0)}(1)$ into (4), we obtain $c = (x^{(0)}(1) - b/a)e^{a}$. As a result

$$x^{(1)}(t) = (x^{(0)}(1) - b/a)e^{-a(t-1)} + b/a \tag{5}$$

By changing the independent variable, we have

$$x^{(1)}(t+1) = (x^{(0)}(1) - b/a)e^{-at} + b/a \tag{6}$$

Note that t is a continuous variable. The discrete version of (6) can be found by sampling the time function at the same interval as that in the source sequence. This results in

$$x^{(1)}(k+1) = (x^{(0)}(1) - b/a)e^{-ak} + b/a \tag{7}$$

The next step is to find the coefficients $[a \quad b]$. From (2), we have

$$x^{(0)}(2) + az^{(1)}(2) = b$$
$$x^{(0)}(3) + az^{(1)}(3) = b$$
$$\cdots$$
$$x^{(0)}(N) + az^{(1)}(N) = b$$

In matrix form, this becomes

$$\begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(N) \end{bmatrix} = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(N) & 1 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \end{bmatrix} \tag{8}$$

According to the method of least squares, the vector $[a \quad b]^T$ can be found as

$$\begin{bmatrix} a \\ b \end{bmatrix} = (\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{X}_N \tag{9}$$

where

$$\mathbf{B} = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(N) & 1 \end{bmatrix} \tag{10}$$

and

$$\mathbf{X}_N = \left[x^{(0)}(2) \quad x^{(0)}(3) \quad \cdots \quad x^{(0)}(N)\right]^T \tag{11}$$

Substituting (9) into (7), we have an estimate form of white response given by

$$\hat{x}^{(1)}(k+1) = (x^{(0)}(1) - b/a)e^{-ak} + b/a \tag{12}$$

From (2), the source sequence can be estimated by

$$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k) \tag{13}$$

or

$$\hat{x}^{(0)}(k+1) = (x^{(0)}(1) - b/a)(1 - e^{a})e^{-ak} \tag{14}$$

This is the equation by which original sequences are constructed.

3. EXAMPLES AND RESULTS

To demonstrate how GM(1,1) model can be used to characterize the phase error, two examples are conducted here. Firstly, we use the GM(1,1) to model the phase error between two clocks. Table 1 gives the data taken from MJD 51015, which are the phase errors measured hourly between a clock called "FTG" and the master clock used to generate the UTC(TL). The data in the table are in sequence of left to right and top to down. For this clock, the coefficients a and b using $N = 7$ are found as

$$[a \quad b] = [5.722 \times 10^{-3} \quad 4.449 \times 10^{-7}] \tag{15}$$

The choice of $N = 7$ is arbitrary here. Substituting this result into (14), we obtain a source sequence of the form

$$\hat{x}^{(0)}(k+1) = 4.437 \times 10^{-7} \times e^{5.722 \times 10^{-3} \times k} \tag{16}$$

Table 2 gives a comparison between the data predicted by (16) and the measuremen

t results recorded in our laboratory. It can be seen that the prediction errors are less than 3.45% for the data shown. This means that the phase error at 62 (i.e., $69 - 7$) hours later could even be predictable with such a small error using only the first seven ones. For the same case, the prediction error resulted by extending the line obtained from the method of least squares fit is as large as -7.52%. Fig. 1 compares the errors using GM(1,1) with those using least squares fit. It can be seen that the maximum error for the former is less than 4%, while that of for the latter may up to 8%.

The former is about 50 % lower than the latter. Note that it is meaningless trying to predict any future data by extending the curves obtained from polynomial fits using this segment of data samples.

Secondly, a sequence of noise corrupted data is used to test the GM(1,1) model. Fig. 2 shows a set of phase errors obtained from comparing two clocks (a master clock and a remote clock) via a short baseline GPS common-view. According to the schedule used in our system, up to 48 data can be collected in one day. It is obviously seen from the figure that the remote clock has a significant frequency offset away from the master clock. To avoid using data with irregular space, only data with equal interval are chosen in the following modeling. Table 3 gives 28 equally spaced data, which are spaced 30-minute apart. Again the data are read in sequence of left to right and top to down. The coefficients resulted from the sequence using $N = 28$ are as follows

$$[a \quad b] = [-8.784 \times 10^{-3} \quad 5.299 \times 10^{-8}] \qquad (17)$$

Together with the first data given in Table 3, $x^{(0)}(1) = 5.41 \times 10^{-8}$ s, the source sequence can then be modeled as

$$\hat{x}^{(0)}(k+1) = 5.323 \times 10^{-8} \times e^{8.784 \times 10^{-3} \times k} \qquad (18)$$

Fig. 3 shows the phase errors predicted by (18) accompanied with the data given in Table 3. From (18) the phase error computed at the middle of the first period, $k = 11$ under the data set chosen, is $\hat{x}^{(0)}(12) = 58.63$ ns (the error measured is 62.90 ns), while that at the middle of the next period, $k = 57$ under the data set chosen, is $\hat{x}^{(0)}(58) = 87.82$ ns (the error measured is 78.60 ns). According to the phase errors collected in this example, the two points computed above exactly cover a span of one day. It can be seen from (18) that the source sequence is a growing process. This tells that the remote clock runs slower than the master clock. Hence, the frequency offset of the remote clock relative to the master clock can be estimated as

$$-(\hat{x}^{(0)}(58) - \hat{x}^{(0)}(12)) \times 10^{-9} / 86400 = -3.38 \times 10^{-13} \qquad (19)$$

If the remote clock is to be made traceable to the master clock, then it needs to steer with a quantity opposite to that given in (19), that is 3.38×10^{-13}. Fig. 4 shows the phase errors measured after steering the clock with that estimate. Note that there is a separate adjustment of phase offset before obtaining this result. The frequency offset estimated from Fig. 4 is on the order of 10^{-14}. This tells that a clock with frequency offset on the order of 10^{-13} is now improved to the order of 10^{-14}.

Table 1 A segment of the phase data (in unit of seconds) recorded in our laboratory.

4.4424E-07	4.4035E-07	4.3687E-07	4.3846E-07
4.3535E-07	4.3111E-07	4.2697E-07	4.2893E-07
4.2531E-07	4.2330E-07	4.2076E-07	4.1812E-07
4.2008E-07	4.1818E-07	4.1434E-07	4.1121E-07
4.0863E-07	4.1012E-07	4.0680E-07	4.0361E-07
4.0021E-07	4.0187E-07	3.9869E-07	3.9582E-07

Table 2 Comparison of prediction values with measurement results.

k	Measured (E-07)	Predicted (E-07)	Error (%)
9	4.233	4.214	-0.44
19	4.036	3.980	-1.39
29	3.817	3.759	-1.53
39	3.556	3.550	-0.18
49	3.370	3.352	-0.53
59	3.157	3.166	0.28
69	2.890	2.990	3.45

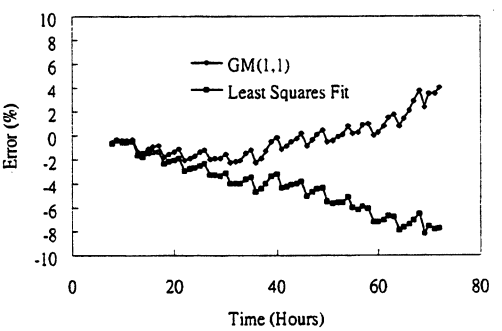


Fig. 1 Prediction errors using GM(1,1) and least squares fit.

Table 3 A set of the phase data (in unit of ns) obtained via GPS common-view.

54.10	51.30	56.30	55.50	52.00	54.50	56.80
55.70	55.70	57.60	56.90	62.90	62.90	66.70
60.90	55.80	61.80	58.40	62.50	64.20	65.40
64.60	62.30	60.20	61.90	66.20	70.90	69.30

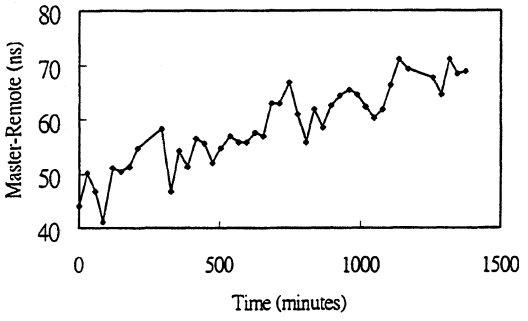


Fig. 2 A set of phase errors obtained via GPS common-view.

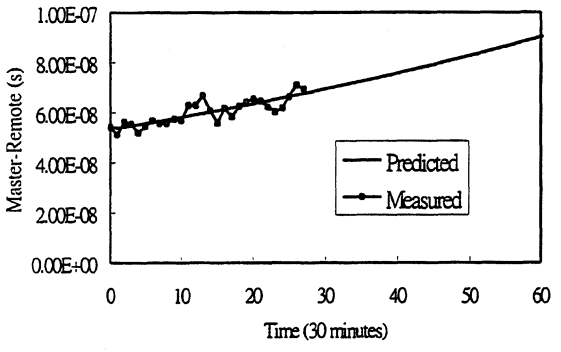


Fig. 3 Phase errors predicted by (20) accompanied with those used for modeling.

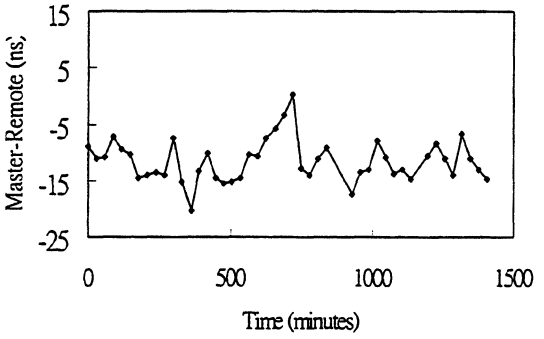


Fig. 4 Phase errors measured after steering the clock.

4. CONCLUSIONS

The use of GM(1,1) to characterize the phase error is presented in this paper. Two examples are conducted to verify the value of using the model. The first one indicates that lower prediction error than that achieved by using the method of least squares fit can be obtained. The second one demonstrates the capability of using the model to steer a clock with prediction determined by it.

5. REFERENCES

[1] J. Deng and H. Guo, Principle of gray prediction and its applications (in Chinese), Chuan Hwa, 1996.

[2] D.W. Allan, "Time and frequency (time-domain) characterization, estimation, and prediction of precision clocks and oscillators," IEEE Trans. On Ultrasonics, Ferroelectrics, and Frequency Control, Vol. 34, No. 6, pp. 647-654, Nov. 1987.

CORRELATION OF FREQUENCY HOPPED VCO PHASE SETTLING TO VARACTOR TRANSIENT CAPACITANCE

L. A. Mallette, M. J. Delaney, S. Killman, T. Folk, B. Wong
Hughes Space and Communications
SC-S10-S325
P. O. Box 92919
Los Angeles CA 90009 USA

1. ABSTRACT

Microwave VCO and system designers should be aware that attoFarad level changes in the dynamic performance of a varactor can correlate to changes in system settling response of fast hopping frequency synthesizers.

It is theorized that a moisture/silicon interaction occurred within a packaged varactor diode due to imperfect wafer processing of the die and a flawed ceramic enclosure. The flawed ceramic allowed ambient moisture to penetrate and come in contact with exposed silicon at the surface of the die junction, causing an increase in interface (or surface) states which in turn increased the settling time constant of diode capacitance on the order of fifty microseconds. The increased time constant was correlated to an observed degradation in VCO phase settling. The increased phase transient was shown to recover with several time/temperature combinations in accordance with Arrhenius predictions.

2. BACKGROUND

Military system architectures increasingly rely on the utilization of frequency agile (hopped) synthesizers to maintain secure communications by minimizing detection. A performance measure of a synthesizer is its ability to quickly settle to within a phase threshold in a specified amount of time.

This paper briefly describes what is thought to be the underlying cause of degradation in the dynamic performance of such a synthesizer. Sensitivities to ambient atmospheric humidity were observed in-situ in a 4.5 GHz VCO through dynamic phase settling tests and confirmed by capacitive transient measurements of its varactor diode. The phase transient of the hopped frequency at a given time was shown to double after humidity exposure on the varactor diode. The increased phase transient was shown to recover with several time/temperature combinations in accordance with Arrhenius predictions.

The degraded characteristic of capacitive transience (Reference 1 and Appendix) and phase settling are illustrated in figures 1 and 2 respectively.

The varactor diode is a silicon, mesa constructed diode (Figure 3) manufactured using diffusion/ion implantation semiconductor wafer processing. The package consists of a cylindrical ceramic body having a kovar pedestal base at one end and an open kovar flange at the other. The diode cathode is eutectically mounted to the pedestal and the anode is electrically connected to the open ended flange (top). A kovar lid is welded in place to hermetically seal the package.

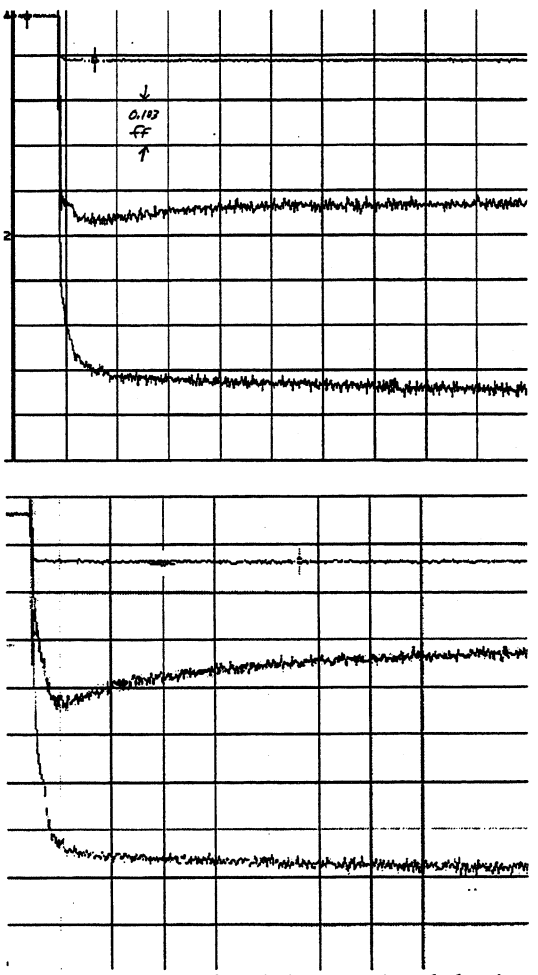


Figure 1: Capacitive and resistive transient behaviors (top and bottom waveforms of each graph respectively). Top graph shows nominal. Bottom graph shows degraded condition. (Scale: 100 aF/div. vs. time).

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

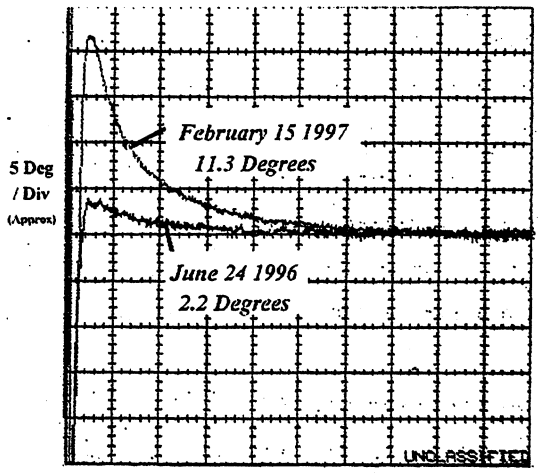


Figure 2: Phase settling degradation - Normal behavior (lower curve) and degraded behavior (upper curve) after being stored, non-operational in air. (5 deg/div. vs. time).

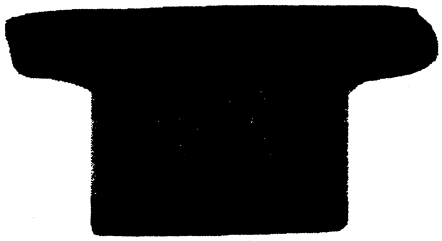


Figure 3: Varactor diode x-ray showing internal construction. Scale: about 20 X.

Data was obtained from three separate wafer lots of varactor diodes. The lots are categorized as shown in the following table:

Wafer Lot	Description
-1	Bare die - No anomalous phase/capacitive transient
-2	Packaged die - Anomalous phase/capacitive transient
-3	Packaged die - No anomalous phase/capacitive transient

3. EXPOSED SILICON

It was observed that only the -2 wafer lot diodes were experiencing phase settling drift while the two other lots exhibited no change. The theory was that exposed silicon was somehow the site of moisture induced electron/ion traps or hydrogen contamination.

The as-deposited semiconductor and contact metal layer structures of the different diode lots were examined using a focused ion beam (FIB) system which produced cross sections to be imaged in a Field Emission low voltage Scanning Electron Microscope (FE-SEM). The FE-SEM image (figure 4) shows a perspective view of one whole die. In the center of this die is the circular mesa diode structure covered by a gold bonding pad. Figure 5 illustrates a FIB cut into the mesa structure.

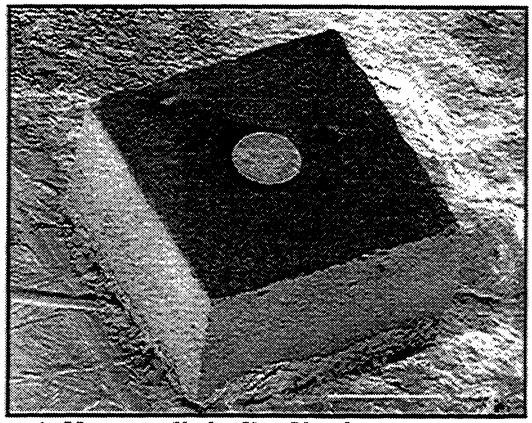


Figure 4: Varactor diode die. Circular structure on top is button of gold for contact. Figure 5 illustrates close-up

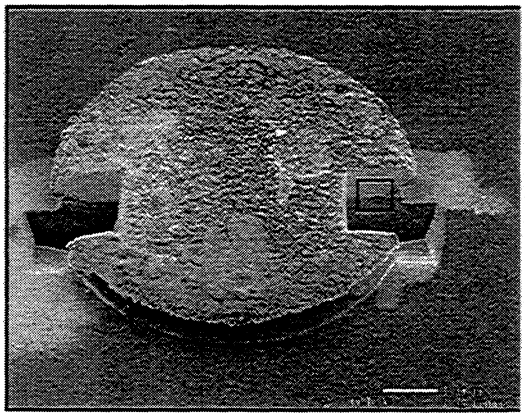


Figure 5: FIB cut of mesa structure. Box outlines close-up shown in Figure 6.

Figure 6 shows the contact metal configuration of a -1 wafer lot varactor diode. The adhesion metals titanium (Ti) and molybdenum (Mo) are used because gold will not adhere well to silicon. These metals are deposited first and the gold last. Etching is performed and the die is passivated with an oxide layer resulting in a structure whose edge is sealed in such a way to protect the underlying silicon from environmental exposure. This die has proper etch characteristics and is properly passivated.

Figure 7 shows a contact metal processing exhibiting exposed silicon surfaces as a result of an over etch condition during processing. Normally the contact metals should cover the silicon out to the oxide as was shown previously. In multiple diode samples of both the -2 and -3 wafer lots, exposed silicon was always found.

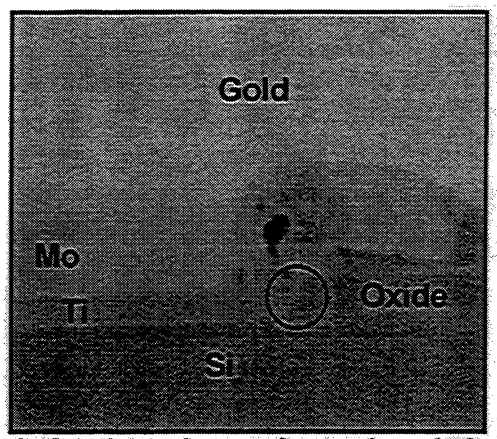


Figure 6: Constituent layers of properly etched -1 lot varactor diode.

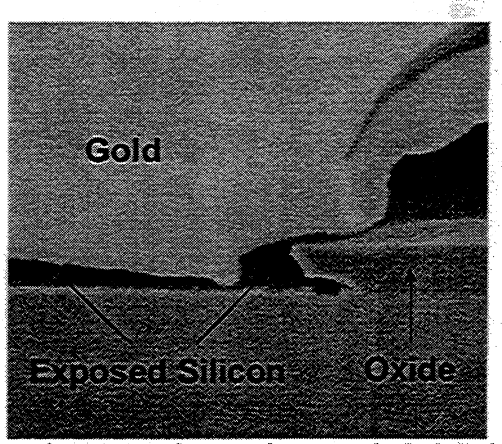


Figure 7: Constituent layers of over etched -2 and -3 lot varactor diodes.

4. EFFECT OF MOISTURE

Several samples of both the -2 and -3 lot diodes were subjected to 85/85 moisture tests in an attempt to induce degradation. Each sample had a small hole made into the package to ensure moisture/silicon contact. In all cases, the part's capacitive transient changed and had a response similar to the original anomaly. A conclusion was made that moisture was indeed interacting with the exposed silicon.

5. HERMETICITY

In an effort to explain why the -3 diode parts did not exhibit degradation under normal environmental conditions but were indeed sensitive to induced moisture, fine and gross leak tests were performed to determine hermeticity of the different packages. It was found that a fine leak test method using helium detection of 5 X 10^{-8} atm cc/s was not sufficient due to the small cavity volume of the diode pill package of $4x10^{-4}$ cm^3. Instead, an ultra sensitive method (utilizing radioactive krypton Kr^{85}) with sensitivity of 1 x 10^{-13} atm cc/s was used. Gross leak tests were made using a red dye penetrant. Several diodes were tested with the following results:

- All -2 diodes failed the fine leak test while all the -3 packaged parts passed.
- Leak rates of the -2 package parts translated to an atmospheric exchange ranging from hours to months while the -3 packages had a translated atmospheric exchange greater than 10 years.
- Red dye penetrant tests of the -2 parts were positive while tests of the -3 parts were negative.
- The leak rates of the -2 diodes were two orders of magnitude better than its specification and normal leak test requirements. A gross leak test (10^{-8}) is the standard acceptance leakage. This part required 10^{-12} or better for acceptance.

6. CERAMIC MICROSTRUCTURE

The ceramic housings of the two packaged diode lots were evaluated using thermomechanical and fracture-mechanics analysis, as well as microstructural characterization.

6.1 Microstructure Review

Scanning electron microscopy (SEM) revealed differences in the microstructures of the two ceramics. The ceramic from the -2 lot consisted of very large particulate agglomerates and a substantial amount of grain-boundary glassy phase while the -3 lot was more uniformly structured with fine and dense (crystalline) grains containing significantly less glassy phase along their boundaries. In addition, major cracks and very large crack-like pores along particulate agglomerate boundaries were observed in the -2 lot while no evidence of major cracking was revealed from the microstructure of the -3 ceramic (figure 8). The microstructural anomalies in the -2 lot ceramics were believed to be all ceramic-processing-initiated, resulting from improper powder treatment, green-body forming process, drying/calcining, and/or sintering.

6.2 Fracture-Mechanics Analysis

Based on microstructural characteristics, linear-elastic-fracture-mechanics (LEFM) was applied to estimate the fracture properties of both diode lot ceramics. Analysis results showed the fracture strength of the -3 ceramic with a finer, denser and more uniform microstructure to be 35 to 43 kpsi while the defective -2 ceramic was 26 to 30 kpsi.

6.3 Thermomechanical Consideration

It was theorized, based on a thermomechanical analysis, that assembly processing of the diodes, dip tinning in particular, may have induced a thermal shock stress

comparable to the fracture strength of the -2 lot ceramic. This stress then rapidly extended and merged the pre-existing processing flaws into a near-through-thickness major crack in the material. The major crack then continued growing through the grain-boundary glassy phase driven by a moderate post-brazing residual stress in the presence of ambient moisture -- a result of moisture-induced-stress-corrosion cracking [2, 3]. Full development of a through-wall-thickness crack in the ceramic housing could have occurred under storage conditions causing a premature hermetic failure.

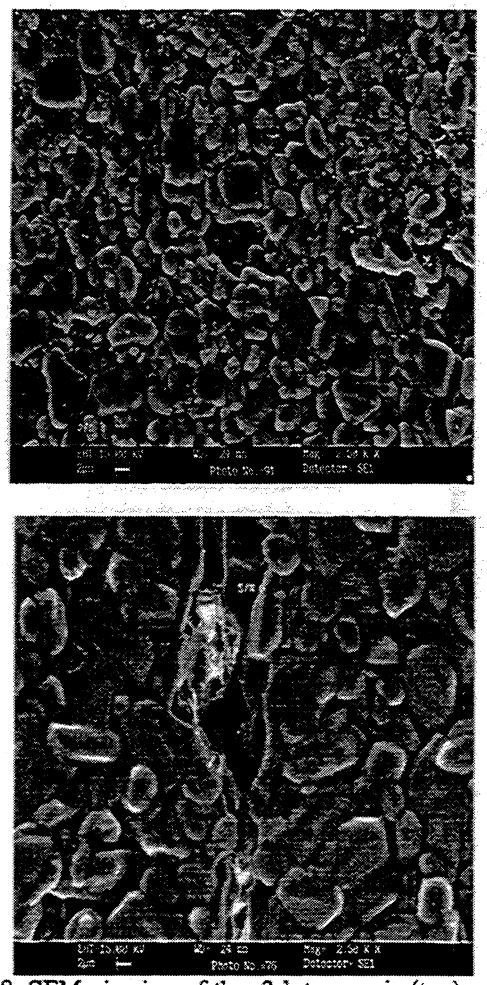


Figure 8: SEM viewing of the -3 lot ceramic (top) and the -2 lot ceramic (bottom).

7. RECOVERY TESTING

The hypothesis suggested that an increase in interface (or surface) state density on the exposed silicon of the varactor diode occurred due to the presence of moisture. It was also hypothesized that this condition should be reversible by thermal activation and follow an Arrhenius relationship. A series of experiments involving vacuum and heat to drive off the moisture demonstrated the theory. An activation energy for the recovery of the phase settling was estimated. From the data it was possible to estimate a recovery time at normal operating temperature in a space environment.

7.1 Vacuum Bake - 110° C (230° F)

Figure 9 shows the phase settling recovery plot of a VCO after being subjected to a vacuum bake at 110° C (230° F). Data points were obtained at only two intervals of 24 hours and 88 hours, but recovery was clearly evident and followed a straight line as expected for an Arrhenius behavior. A 50% recovery required only about 10 hours with complete recovery projected at 1000 hours.

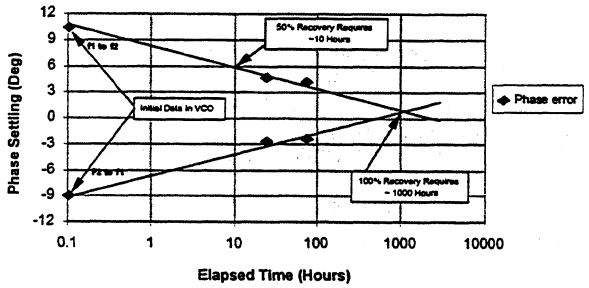


Figure 9: Phase settling recovery. Vacuum bake at 110°C

7.2 Vacuum Bake - 49° C (119° F)

Another VCO was tested at a lower vacuum bake temperature of 49° C (119° F). Here, as shown in figure 10, a 50% recovery is obtained after approximately 500 hours with a full projected recovery occurring on the order of 20,000 hours.

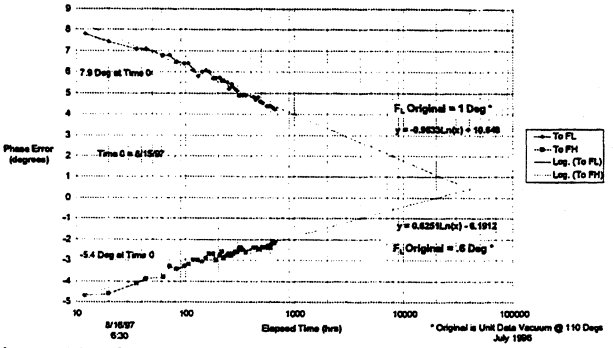


Figure 10: Phase settling recovery. Vacuum bake at 49°C.

7.3 Thermal Vacuum - On-Orbit Temperature

An Arrhenius plot was generated using the 50% and 100% recovery point data from the 110° C and 49° C thermal vacuum experiments. An estimation of the thermal activation energy (0.3 eV) for a recovery process was

obtained from the Arrhenius plot.

A long-term thermal vacuum at a temperature of 29° C (84° F -- the expected on-orbit operating temperature) was performed on yet another VCO to test the ambient temperature recovery prediction. After 2000 hours, an approximate 50% recovery occurred and agreed with the Arrhenius prediction.

The results of the recovery experiments provided adequate evidence to conclude the hypothesis for the cause of phase settling and capacitive transient degradation.

8. CONCLUSIONS

It is theorized that a moisture/silicon interaction occurred within the packaged varactor diode due to imperfect wafer processing of the die and a flawed ceramic enclosure. The flawed ceramic allowed ambient moisture to penetrate and come in contact with exposed silicon at the surface of the die junction, causing an increase in interface (or surface) states which in turn increased the settling time constant of diode capacitance on the order of fifty microseconds. The increased time constant was correlated to an observed degradation in VCO phase settling.

The visual examinations of the varactor die and ceramic package as well as the data obtained from 85/85 tests and recovery experiments provide strong evidence showing that an over etching of the metal contact layers occurred during wafer fabrication, which exposed bare silicon to the package atmosphere. Flaws in the ceramic made the package relatively porous, thus allowing ambient moisture to enter the interior atmosphere of the package. The source of the moisture was normal environmental humidity.

The conceptual understanding of the underlying cause is that interface trap levels were created at the exposed silicon-oxide junction, which changed the transient capacitive response of the diode. This transient capacitance effect is very subtle (transient $\Delta C \sim 100$ attoFarad), extremely difficult to measure, and not easy to model from first principles. The empirical data, which shows that the effect can be induced by moisture exposure and reduced by vacuum baking, is consistent with a surface moisture process and not indicative of a hydrogen poisoning effect.

8. AKNOWLEDGEMENTS

This study could not have been performed without the help of many key individuals, including Robert Galano, Dan Guttierez, Martin Hummel, Ronel Pulmano, and Al Young.

9. REFERENCES

1. Scarpulla, J., Young, A. M. and Chen, J., "Measurement of Capacitance Transients with attoFarad Resolution in a Microwave Varactor Diode after Co60 Irradiation", IEEE Radiation and Nuclear Effects Conference, Tucson, 1994.
2. Freiman, S. W., "Environmentally Enhanced Fracture of Ceramics," Materials. Research Society Symposium Proceedings, Volume 125, pp. 205-215, Materials Research Society, 1988.
3. Engineering Materials Handbook, Volume 4, "Ceramics and Glasses," ASM International, 1991.

APPENDIX
Capacitive Transient

The transient capacitance measurement system measures the dynamic response of the varactor diode capacitance as the voltage applied across its depletion region is stepped from one potential to another. The sensitivity of the system allows measurements to be made within the hundreds of attoFarads. The system consists of a bridge network (reference 1) that separates the real and imaginary components of the response which describe the time dependent behavior of the diode; the imaginary term being the capacitive transient and the real term a resistive transient which serves no useful purpose. The identifying feature of the capacitive transient plot that relates to phase settling degradation is the overshoot of the capacitive transient curve. It is hypothesized that the overshoot transient capacitance consists of high frequency components that the PLL has difficulty tracking out and becomes the dominant delay.

1999 Joint Meeting EFTF - IEEE IFCS

Spurious Signals in Direct Digital Frequency Synthesizers Due to the Phase Truncation

Věnceslav F. Kroupa, *member IEEE*, Václav Čížek, Jarmil Štursa, and Hana Švandová

The Institute of Radio Engineering and Electronics
Academy of Sciences of the Czech Republic, 182 51 Praha, Czech Republic

Abstract

One of the sources of spurious signals in DDFS is the use of smaller number, **W**, of the most significant bits (MSB) applied for the output sine wave reconstruction from all **R** bits stored in the accumulator. The result is a phase modulation of the output signal.

The problem was first solved by Nicholas et al. [1,2] in a rather complicated way with the result that the level of the largest spurious signal is about -6W dB below the carrier with an increase of 3.9 dB in some instances.

In the present paper shall investigate the validity and correctness of these generally cited results [3, 4].

1. Introduction

The block diagram in Fig. 1. illustrates the set-up of DDFS with a Modulo-N accumulator ($N=2^R$).

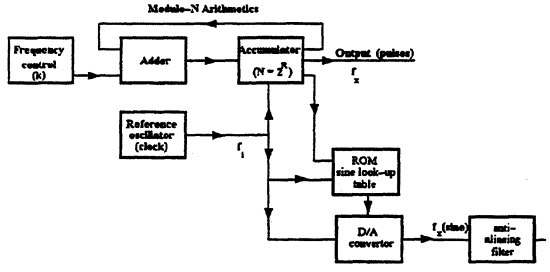


Fig. 1 Block diagram of the Modulo-N DDS.

In this case the output frequency f_x is

$$f_x = f_i\frac{k}{2^R} = f_i\frac{X}{Y} = f_i\frac{X}{2^L} \quad (L \le R) \qquad (1)$$

where X and Y are relatively prime integers and L the effective bit length of the accumulator (in the case that k contains factors of 2's).

More then ten years ago Nicholas and Samueli [1,2] tried to solve the problem of spurious signals in DDFS generated by the phase truncation, caused by neglecting some low significant bits (LSB) in the accumulator -see Fig. 2. However, a major difficulty with the paper [1] is a rather complicated derivation and presentation of the results including a correction table. In addition we encounter another problem, namely, that these results are repeated, without any doubts, even in recent books on DDFS [3, 4].

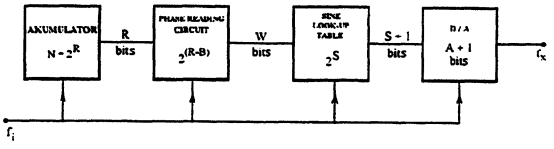


Fig. 2 The most important "generators"of spurious signals in DDS.

The first author of the present contribution found a simpler solution of the problem of spurious signal level due to the phase truncation in DDS [5, 6]. At this occasion he was also faced with the validity of the corrections suggested by Nicholas [1].

In the present paper we shall first recall some important results from the paper [5] and with their assistance reconsider results published in papers [1 and 2] and books [3,4].

2. Analysis of the phase truncated sine waves

2.1 Basic principles

The high frequency resolution of DDFS requires very large accumulator capacity (R=32, 48, or even 64 bits). It is evident that the would be complexity of the hard ware would prevent to refer to all these minute phase increments. Generally, we use only a small number of the most significant bits (**MSB**), ***W***, and disregard all the remaining least significant bits (**LSB**), ***B***, i.e.,

$$B = R - W \quad or \quad B = L - W \qquad (2)$$

Thus from the original number stored Modulo-Y in the accumulator after the *m-th* clock pulse

$$x(m) = mX - sY \le Y = 2^L \le 2^R \qquad (3)$$

and $s=0,1,2,\ldots$ and *s-th* overflowing we refer to

$$X'(m) = integer\left(\frac{mX - sY}{2^B}\right) \qquad (4)$$

Since the information, passed to the "sine look-up table",

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

than indicated in Tab. 1, particularly, for $\xi_x > 0.1$.

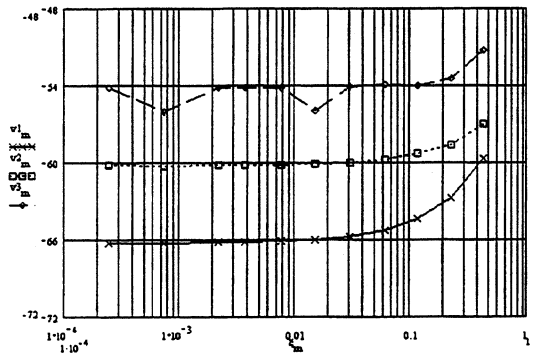


Fig. 5. The largest spurs for B=1 ($v1_m$), B=2 ($v2_m$), and B=3 ($v3_m$) as function of ξ_m for R=12.

Another lesson learned from Fig. 5 is that for B=3 the respective characteristic, $v3_m$, is oscillating about the expected level (11). After arriving at this result we started with a more exact investigation of the problem.

2.3 Computer simulations - large number of accumulator bits neglected.

An example is shown in Fig. 6 for the normalized frequency $\xi_x = 1625/4096 \approx 0.4$ and **4** bits neglected. s_k and s_{kk} (o) indicate sorted spurious signals due to phase modulation, computed via FFT, whereas rectangles the idealized slope of harmonics $20*log(2^{-W}/p)$. Note that for small harmonic number p both slopes coincide, but for p nearing 2^B we notice an increasing difference which for $p = 2^{B-1}$ reaches $+3.922\ dB$; the preceding ones are $+0.913\ dB$, $+0.235\ dB$ etc. which recall Tab. 1.

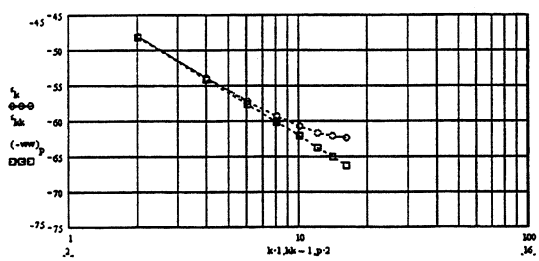


Fig. 6 Level of the spurious signals computed with DFT for $\xi_m = 1625/4096 \approx 0.4$ and $B = 4$.

On the other hand application of the Fourier transform on the analog stair case output sine wave revealed for small normalized frequencies (1) nearly the ideal slope $20*log(r)$ - see Fig. 7a - with increasing deviations for larger and larger normalized frequencies $\xi_x = f_x/f_c$ - see Figs. 7b through 7d.

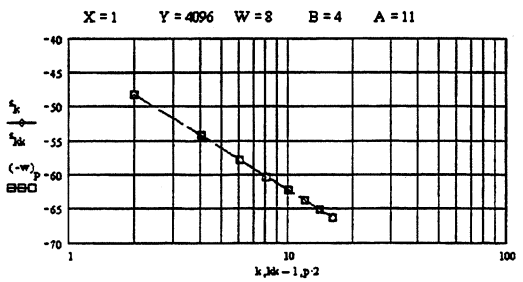


Fig. 7a

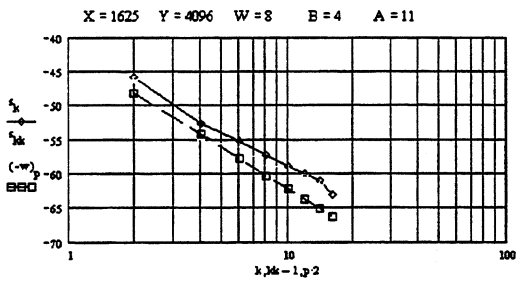


Fig. 7b

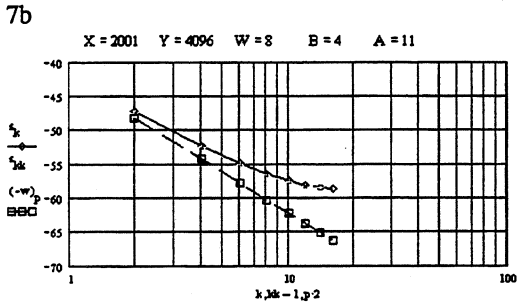


Fig. 7c

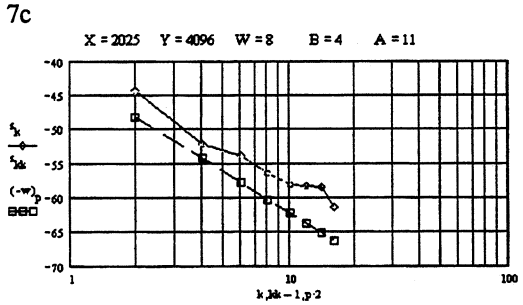


Fig. 7d

3. Mathematical investigations

By considering the validity of (7) we can simplify relation (5) into

$$\sin\frac{2\pi mX}{Y} - 2\pi\frac{2^B}{Y}s(m)\cos\frac{2\pi mX}{Y} \qquad (12)$$

After simplifying modulation function $s(m)$ with the approximation (8) we face a nearly ideal sine wave with amplitude 1 and a disturbing cosine wave with a small sinusoidal modulation. After applying relations between

must be of the same order as the original ***X(m)*** we have to multiply ***X'(m)*** by 2^B and the sine readings are ideally given by (cf. Fig. 2.)

$$\sin[2\pi\frac{2^B}{Y}integer(\frac{mX}{2^B})] = \sin\{2\pi\frac{2^B}{Y}[\frac{mX}{2^B} - s(m)]\} = \sin\{2\pi[\frac{X}{Y}m - 2^{-W}s(m)]\} \quad (5)$$

where the sampled phase modulation function

$$0 < s(m) = \frac{mX}{2^B} - integer(\frac{mX}{2^B}) \le 1 \quad (6)$$

is a periodic one. Furthermore we find that the phase modulation index in (5) is generally small compared with "1"

$$2\pi\frac{2^B}{Y}s(m) = 2\pi 2^{-W}s(m) \ll 1 \quad (7)$$

Investigation of the properties of the function ***s(m)*** reveals that for one discarded bit, i.e. **B = 1**, we get a rectangular wave, with the space-mark ratio 1:1 and the amplitude $\mathbf{a_s}$ = **.5**. However, for larger **B** we usually face a superposition of sawtooth waves - Fig.3.

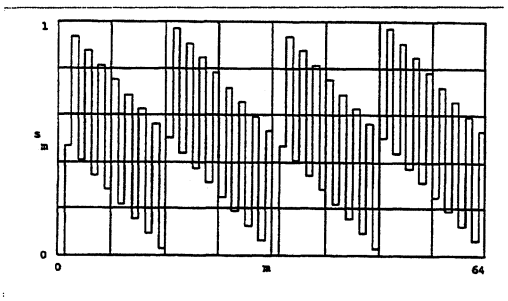


Fig. 3 Plot of *s(m)* for $\xi_x=207/2^{10}$ for *B=5*: Superposition of rectangular and sawtooth waves.

After Fourier series expansion of **s(m)** we get the following approximation [6 p. 163] of *s(m)*

$$s(m) \doteq \frac{1}{\pi}\sum_{r=1,2,\ldots}^{2^{(B-1)}}\frac{1}{r}\sin(2\pi rmR_b) \quad (8)$$

where $\mathbf{R_b}$ is defined as

$$R_b = \frac{X}{Y}*2^W - integer(\frac{X}{Y}*2^W) = \frac{P}{2^B} \quad (9)$$

where **P** is an integer from **1 to 2^{B-1}**. Note that we can always hold the remainder in the range from **-.5 to +.5**. This restriction simplifies further calculations.

In an idealized case we get for sine wave modulated with the largest spurious signal from (5), (7), and (8)

$$\sin[\omega_x t + 2*2^{-W}\sin(vt)] \approx J_o(2*2^{-W})\sin(\omega_x t) \pm J_1(2*2^{-W})\sin(\omega_x \pm v)t \approx \sin(\omega_x t) \pm 2^{-W}\sin(\omega_x \pm v)t \pm \ldots \quad (10)$$

Thus in the ideal case we find for the level of the largest spurious signal in respect to the carrier

$$20\log|\frac{J_1(2\cdot 2^{-W})}{J_o(2\cdot 2^{-W})}| \approx 20\log(2^{-W}) \approx -6*W \quad (11)$$

2.2 Computer simulations - small number of accumulator bits neglected

Nicholas [1] pointed out that in some instances where **B** is a small number (1, 2, or 3) the level of the largest spur is larger than the ideal value given by (11). He solved the problem with a correction factor in [1]. The results are recapitulated in the Table 1:

B = 1	corr. -3.922 [dB]	
B = 2	-0.912 [dB]	
B = 3	-0.224 [dB]	Table 1
B ≥ 4	negligible	

However, our computer simulations for one LSB phase bit neglected, i.e. B =1, reproduced in Fig. 4, learned us that the correction was, first of all, function of the normalized frequency $\xi_m = f_x/f_i = X/Y$

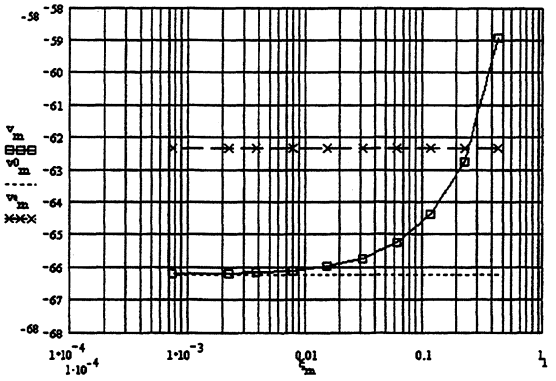


Fig. 4 The largest spur levels, v_m, at the output of a DDFS with R=12, W=11 B=1; (□) with the use of the Fourier transform of eq.(5); (x) with DFT of eq.(5); the dotted line is the ideal value in accordance with (11).

To explain this discrepancy we have applied on eq. (5) a simple discrete Fourier transform (DFT) and arrived at the line with crosses in Fig. 4 which is in agreement with the Nicholas results. In Fig. 5 we have plotted levels of the largest spurs, below the carrier, computed with the assistance of the Fourier transforms of eq.(5) for three different B's. Note that for the greater B the differences, with respect to relation (11), are smaller but still larger

trigonometric functions we get for the spurious side bands caused by the ***r-th*** harmonics

$$\frac{2^B}{rY}\{\sin[2\pi m(\frac{rP}{2^B}+\frac{X}{Y})] + \sin[2\pi m(\frac{rP}{2^B}- \quad (13)$$

Evidently the above expressions represent staircase values of spurious sine waves around the carrier. After computing their amplitudes with the assistance of the Fourier transform we get

$$c_{sr} = \frac{2^B}{rY}\frac{\sin[\pi(\frac{X}{Y} \pm \frac{rP}{Q})]}{\pi(\frac{X}{Y} \pm \frac{rP}{Q})} \quad (14)$$

and after dividing the above relation by the amplitude of the carrier, c_X, we arrive at the ***r-th*** order spurious signal level

$$\frac{c_{sr}}{c_X} \approx \frac{2^B}{rY}\cdot\frac{X/Y}{X/Y \pm rP/Q}\cdot\frac{\sin[\pi(X/Y \pm}{\sin(\pi\frac{X}{Y}} \quad (15)$$

However, we must also take into account that some ***rP/Q*** exceed the Nyquist pass band ***Y/2*** but are returned as alias. As a consequence the spurious spectral line numbers (bins) must meet condition

$$n_r = |X + (r\frac{PY}{2^B} + sY)| < Y/2 \quad (16)$$
$$(s,r = \ldots,-2,-1,0,1,2,\ldots)$$

After introducing r=1 into relation (16) we can estimate the level and the position of the largest spurious signal due to the first harmonic of the spurious phase modulation signal. Consequently we can simplify the above relation to

$$n_1 = X \mp 1*YR_b = X \pm 2^W P \quad (P=1, \quad (17)$$

4. Conclusions

Evidently we cannot accept the statement, e. g. by Crawford [3] and others, that application of the simple DFT on (1), with corrections given in Table 1, provide an actual DDFS output spectrum, more exactly, the level of the largest spurious signal caused by the truncation of some small number of accumulator bits.

There is another needless complication with introducing the largest common divisor between F_r and 2^B, since for each DDFS system the given constants are: the clock frequency f_c, the number of accumulator bits R and often the number of effective phase bits W. Consequently the input is either the desired frequency f_x or the number of phase steps F_r, but the ratio $F_r/2^R$ is automatically changed into relatively prime X/Y, it is to the normalized frequency ξ_x.

5. Appendix

To examine validity of the approximation of eq.(6) with the expansion (8) we have computed the respective spectral lines. One result is shown in Fig. 8.

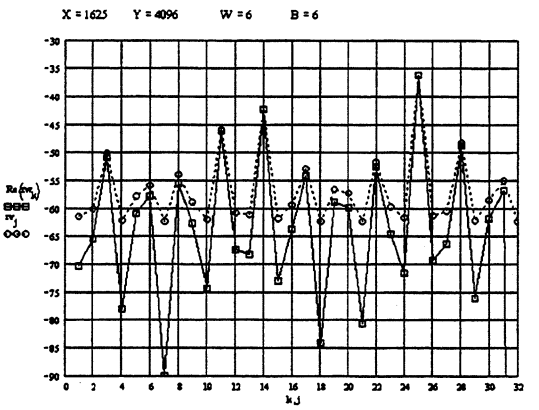


Fig. 8 DFT spectra of *s(m)* according to (7), (o), compared with spectra of the approximation (8) (□).

6. Acknowledgement

This work has been supported by the Grant Agency of the Czech republic under the contract No. 102/96/1240.

7. References

[1] H.T. Nicholas, III, and H. Samueli, "An analysis of the Output Spectrum of Direct Digital Frequency Synthesizers in the Presence of Phase-Accumulator Truncation," Proceedings of the 41st Annual Frequency Control Symposium 1987, Proc. pp. 495-502 (reprinted in [6]).

[2] H.T. Nicholas, III, H. Samueli, and B. Kim, The Optimization of Direct Digital Frequency Synthesizer Performance in the Presence of Finite Word Length Effects, Proceedings of the 42nd Annual Frequency Control Symposium 1988 (reprinted in [6]).

[3] J.A. Crawford, Frequency Synthesizer Design Handbook. Boston/London: Artech House, 1994.

[4] Bar-Giora Goldberg, Digital Techniques in Frequency Synthesis. New York: MacGraw-Hill, 1996.

[5] V.F. Kroupa, "Discrete spurious signals and background noise in Direct Digital Frequency Synthesizers, IEEE 1993 International Frequency Control Symposium, pp. 242-250.

[6] V.F. Kroupa,ed.:"Direct Digital Frequency Synthesizers" IEEE Press, 1999.

1999 Joint Meeting EFTF - IEEE IFCS

On the frequency and amplitude spectrum and the fluctuations at the output of a communication receiver

M. Planat

(planat@lpmo.univ-fcomte.fr)
Laboratoire de Physique et Métrologie des Oscillateurs du CNRS,
Associé à l'Université de Franche-Comté
32 Av. de l'Observatoire, 25044 Besançon Cédex, France

Abstract

A multiplier (or mixer) cascaded with a low pass filter is the basic piece of any communication receiver. It is also used to register minute frequency fluctuations of an external oscillator (RF) under test versus the frequency of a local oscillator (LO). Such a scheme may also be viewed as the basic model of an electronic oscillator, with the amplifier noise at the RF input and the resonator signal at the LO input. We have investigated experimentally the whole spectrum of frequencies and amplitudes of beat signals and their frequency fluctuations at the IF output of the multiplier + filter. We have found evidence that all the dynamics follows arithmetical rules. As for the frequency of the beat signal it is defined from a diophantine approximation of the frequency ratio of input oscillators. As for the amplitude it is defined globally from the position of resolved fractions with respect to the uniform graduation. As for the frequency fluctuations a transition from white frequency noise to 1/f frequency noise is observed close to resonance. It is explained on the basis of number theory in relation to the Riemann problem concerning the distribution of prime numbers.

1. The frequency spectrum

The experiments were performed using a wideband Schottky diode mixer, a LO with constant frequency $f_l \approx 10$ MHz, a RF synthesizer with variable frequency f_0 and a third order low frequency filter with cut-off frequency 375 kHz [1]. Fig. 1 shows the frequency f_i and the amplitude a_i of the beat note. It is observed that they are many subharmonic modes $f_i = | p_i f_o - q_i f_l |$ in addition to the fundamental one at 1/1, and that the ones with p_i or q_i even are strongly rejected due to the doubly balanced structure of the mixer.
In [1] we gave a nice interpretation of the frequency spectrum of output beat signals $\mu_i = f_i / f_o$ as a function of the frequency ratio $\nu = f_l / f_o$. Since the three frequencies operation may be interpreted as:

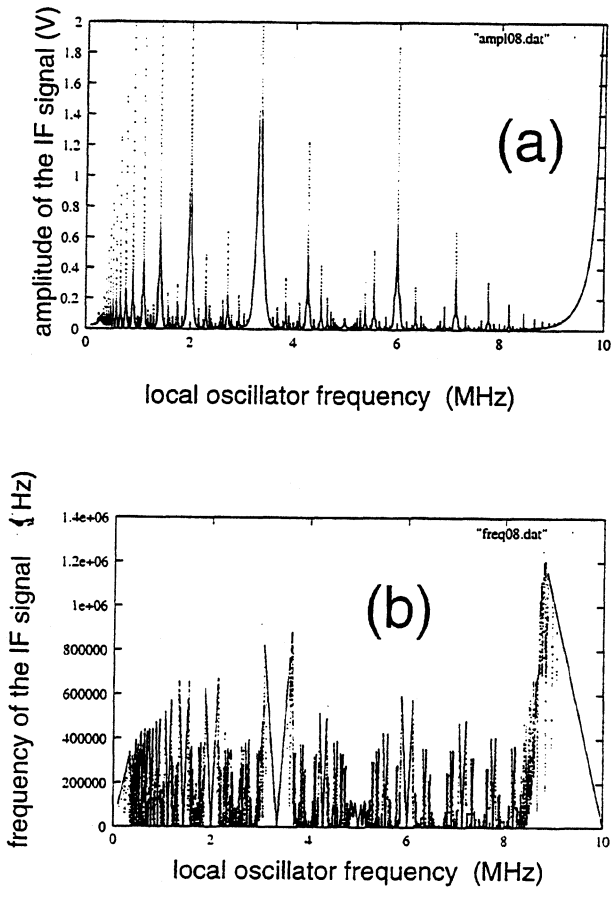


Fig.1
The frequency spectrum μ_i (b) and the amplitude spectrum $v_i = A\,\delta_i$ (a) at the output of a communication receiver versus the frequency ratio ν between the input oscillators

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

$$\mu_i = q_i \, | \nu - p_i/q_i \, | \qquad (1)$$

it is expected that the device produces a rational approximation p_i/q_i of the real number ν allowed by the physical constraints, in particular the type of synchronization that takes place in the mixer, the properties of the filter, the finite resolution of countings and the unavoidable thermal noise which is present in the set-up.

The observed approximation is of the diophantine type, not of the decimal one. It is given from continued fraction expansions (CFE):

$$\nu = a_0+1/\{a_1+1/\{a_2+\ldots 1/\{a_i+\ldots\}\}\} \qquad (2)$$

with $a_0=[\nu]$, $\alpha_0=\{\nu\}$, $a_i=[1/\alpha_{i-1}]$, $\alpha_i=\{1/\alpha_{i-1}\}$, where $[\nu]$ and $\{\nu\}$ denote the integer and fractional part of ν. Successive best approximants of ν, that is convergents p_i/q_i are obtained by truncating the expansion at some stage i. The plot in Fig. 1b may be very well fitted by using truncated CFE whenever a low partial quotient $a_i = a_{max}$ is reached [1]. This implies that each basin circumventing a given resonance corresponds to te same convergent p_i/q_i and has width of order $\delta\nu \approx 1/(q_iq_{i+1}) \leq 1/(a_{max}q_i^2)$.

2. The amplitude spectrum

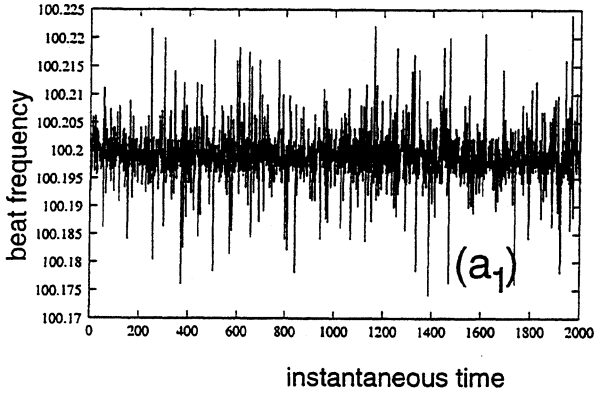
We also found an arithmetical interpretation of the resonant amplitude at mode p_i/q_i. Knowing the position of all the n fractions up to the denominator q_{max} ($n \approx 3\, q_{max}^2/\pi^2$), it is possible to compute the shift between the position of fraction p_i/q_i on the unit interval and the point i/n of the corresponding uniform scale, that is

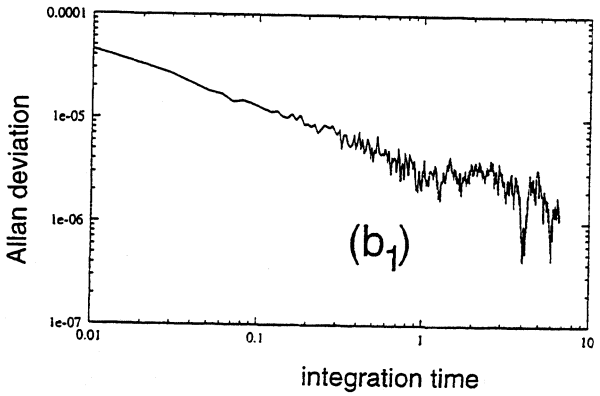
$$\delta_i = | \, i/n - p_i/q_i \, | \qquad (3)$$

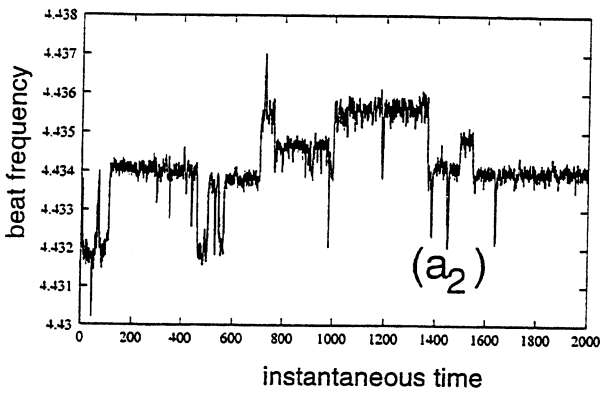
The amplitude of the beat signal at p_i/q_i follows from $v_i = A\, \delta_i$ where A is a reference voltage depending on the sensitivity of the mixer.
The shift δ_i was introduced in 1924 by Franel and Landau (FL) in the context of a conjecture equivalent to Riemann hypothesis [7]. The hypothesis relies on the position of zeroes of the function $\zeta(s)$ which was introduced by Riemann in 1859 in his celebrated paper on prime numbers. According to FL the sum $S = \sum_{i=1}^{n} \delta_i \approx 0(n^{1/4+\varepsilon})$ whatever ε is equivalent to RH, so that the mean deviation should be of order $\sigma \approx S/n \approx n^{-3/4}$.
A drawing of the FL shift δ_i closely follows Fig. 1a if we account for the rejection of even intermodulation products [1].

3. The frequency fluctuations

We performed experiments of the time fluctuations in the instantaneous period close to the resonance 1/1. Allan deviation $\sigma_y(\tau)$ of the relative frequency $y^{(1)}(\tau) = f_i^{(1)}(\tau) / f_i(\tau)$ (with $f_i(\tau)$ the mean frequency) is represented for the three l time series in Fig. 2. We remind that Allan variance $\sigma_y^2(\tau)$ is defined as the mean squared value of the relative frequency deviation between adjacent samples of length τ. A transition from a $1/\sqrt{\tau}$ dependence for $f_i \approx 100$ Hz (which is white frequency noise) to a constant law for $f_i \approx 0.5$ Hz (which is 1/f noise) is observed when the resonance is approached. The transition occurs through an intermediate region with a random telegraph type signal (RTS) with slow divergence $\tau^{1/4}$ for $f_i \approx 4.4$ Hz.







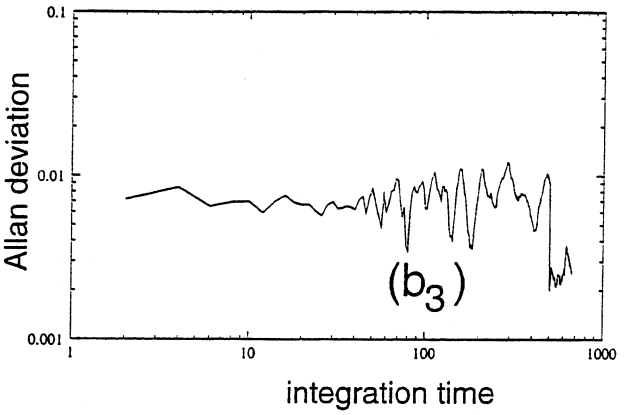
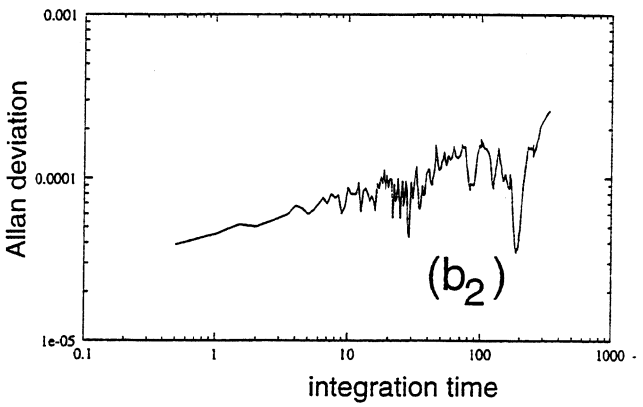
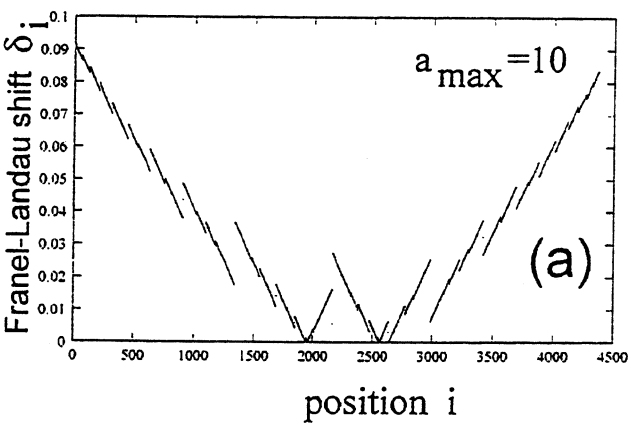
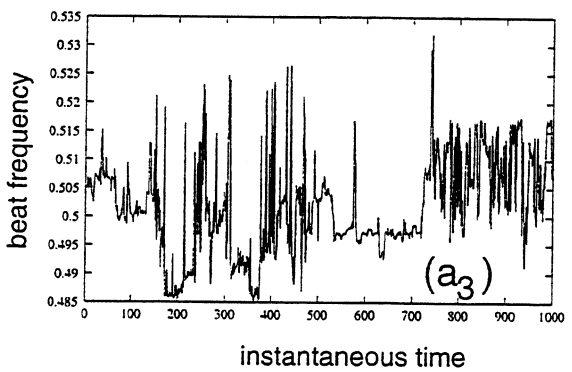


Fig. 2

Frequency fluctuations μ_i of beat signals versus time 1 (a) and the corresponding Allan deviation $\sigma_y(\tau)$ (β) in the transition from white frequency noise (1) to 1/f frequency noise (3). The intermediate regime (2) shows a RTS type noise.







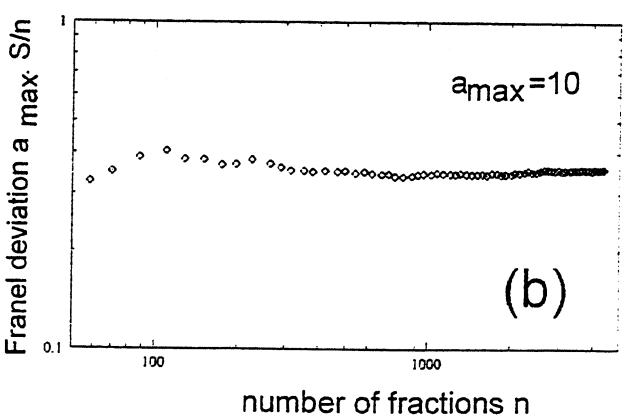


Fig.3

The Franel-Landau shift δ_i for fractions with CFE truncated at low level a_{max}=10 (q_{max}=150) (a) and the corresponding deviation $a_{max}\sigma$ versus n (b).

We could find an arithmetical interpretation of the 1/f noise observed close to resonance. We used (2) as the definition of the frequency deviation, i.e. in such a region the device works as a discriminator with an output amplitude proportional to the frequency deviation.

In addition resonance induces a lack of resolution in the CFE. Accounting for this by setting a prescribed low length i_{max}=1, 2 or 3 for the CFE leads to a mean shift $\sigma \approx S/n \approx n^{1/4}$ as calculated in [1]. The approach may explain results in Fig. 2b associated to the RTS signal. Finally by setting a prescribed low value $a_i \leq a_{max}$ for the partial quotient in the continued fraction expansion of the input frequency ratio ν leads to a Franel-Landau shift as shown in Fig. 3a. The plot looks similar to a fractal attractor. This is confirmed by plotting the corresponding mean deviation a_{max} $S/n \approx 0.35$. Also it has been found (C. Eckert, private communication) that the attractor has a fractal dimension close to unity.

4. Conclusion

This work on resonances at the output of a mixer extends similar results obtained from carrier-envelope resonance in a GaAs field effect transistor [3]. It confirms the idea that 1/f frequency noise is of deterministic origin [4,5]. It arises from the lack of resolution of synchronization states close to resonance. This could be observed close to the phase locked loop region in [6]. Since the visible effect of nonlinearity near the mode-locking zone is the increased slope of the effective frequency shift between the input oscillators versus the bare one, this should be converted into an increased 1/f noise versus this bare shift and the open loop gain. The increased 1/f noise close to resonance in agreement with the predicted law was announced in [6] and the present paper relies on the fundamental explanation of its origin. Detailed results concerning the above dynamics and related mathematical and physical effects are given in our forthcoming book [2].

5. References

[1] M. Planat, S. Dos Santos, N. Ratier, J. Cresson and S. Perrine, "Close to resonance interaction of radio frequency waves in a Schottky diode mixer: 1/f noise and number theory", in "Quantum 1/f noise and other low frequency fluctuations in electronic devices", edited by A. Chung and P. Handel, AIP Press (1999).

[2] M. Planat, "Noise of frequencies and the dynamics of algebraic numbers", Lecture Notes in Physics, to be published, end 1999.

[3] M. Planat, F. Lardet-Vieudrin, G. Martin, S. Dos Santos and G. Marianneau, "Frequency stability enhancement from carrier-envelope resonance in a surface acoustic wave delay line oscillator", J. Appl. Phys., 80, 2509 (1996); see also IEEE IFCS Proc., Honolulu (1996), p. 815

[4] M. Planat, V. Giordano, G. Marianneau, F. Vernotte, M. Mourey, C. Eckert and J.A. Miehe, "Is the frequency noise of an oscillator of deterministic origin?", IEEE Trans. On UFFC, 43, 326 (1996); see also EFTF Proc., Besançon (1995), p. 377

[5] M. Planat, "Annals of Telecommunications", Special Issue on Oscillators, Vol 51, 7-8 and 8-9 (1996)

[6] S. Dos Santos and M. Planat, "Arithmetical fractals in an electronic loop", in "Fractals and Beyond" edited by M.M. Novak, World Scientific, Singapore (1998), p. 297; see also EFTF Proc., Warszawa (1998), p. 402

[7] M.H. Edwards, "Riemann's Zeta Function", Acad. Press, New York, (1974)

EVALUATION OF PASSIVE COMPONENT SHORT-TERM STABILITY VIA USE IN LOW LOOP DELAY OSCILLATORS

M. M. Driscoll, Northrop Grumman Corporation

Electronic Sensors and Systems Sector, P.O. Box 746, MS 134, Baltimore, MD 21203

ABSTRACT

Measurement of the PM noise generated by passive components operated at HF carrier frequencies was reported by NIST at the 1998 Frequency Control Symposium [1]. The results were obtained using a two channel, cross-correlation measurement system. The results indicated that the flicker portion of short-term reactance fluctuations of inductors, capacitors, and varactor diodes could result in carrier signal spectral degradation, especially when the components were used in moderate and narrow bandwidth tuned circuits (i.e., bandpass filters.

In this paper, the results of similar measurements made in the VHF range are reported. Determination of the short-term reactance stability of several types of inductors, capacitors, varactor and PIN diode-based attenuators and helical resonators has been made. The measurements were made by using these components in simple tuned circuits and multipole filters that are used as the frequency-determining element in low loop delay oscillators. Using this method, the near-carrier phase noise of the oscillator signal is poor enough as to be easily measured. Further, if the oscillator loop delay is known, the equivalent open loop phase noise of the oscillator components can be accurately determined from the oscillator signal measurements [2]. Using modular amplifiers having very low 1/f PM noise as the oscillator sustaining stage and DUT circuit bandwidths of 10% or less, it has been possible to measure reactive element short-term stability down to a measurement limit at or slightly below $PSD(\Delta X/X)^2 = 10^{-16}/f$ from f = 1Hz to 1KHz. In virtually every case, the measured levels of oscillator flicker-of-frequency and white frequency noise were almost entirely traceable to the sustaining stage amplifier itself. The conclusions drawn from this work are that "active" components such as amplifiers, mixers, diode and transistor frequency multipliers and dividers, as opposed to the types of passive components measured, constitute the primary contributor to VHF and UHF carrier signal 1/f PM noise.

Low Loop Delay Oscillator Measurements

Figure 1 shows a block diagram of the type of oscillator circuit used. Figure 2 shows the relationship between oscillator signal closed and open loop phase noise.

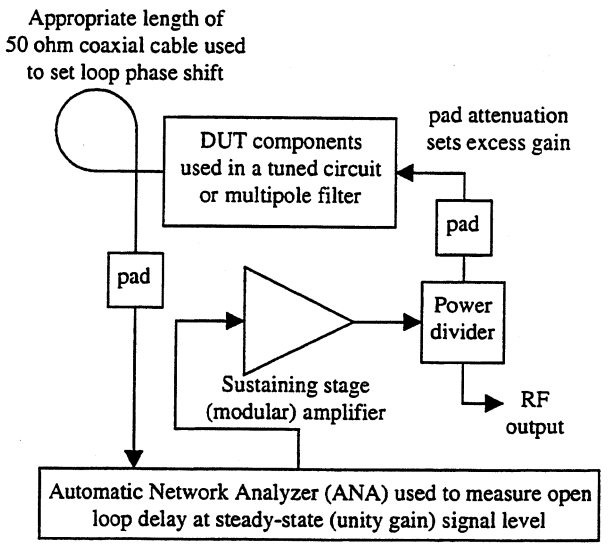


Figure 1. Oscillator Block Diagram, Showing ANA Measurement of Loop Delay

As shown in the figure, there is a conversion of open loop signal phase noise to closed loop frequency noise that carrier signal offset frequencies below $1/2\pi\tau$, where τ is the total loop delay in the oscillator circuit [3]. For low values of loop delay, the closed loop phase noise levels are quite high and easily measurable. Open loop white phase noise is converted to white (20dB/decade) frequency noise, and 1/f PM open loop noise converted to 1/f FM (30dB/decade) noise. If the loop delay is known, the open loop phase noise levels are accurately determined from measurement of the closed loop (oscillator output signal) phase noise spectrum.

The oscillator phase noise measurements were made using a HP3048 Noise Measurement System and phase-locking a HP8662A synthesized signal generator to the oscillator under test. In order to maintain phase-lock to the low loop delay oscillators, the tuning sensitivity of the HP8662A was increased by using it in the DCFM mode. Although DCFM operation results in increased phase noise levels in the HP8662 (the internal crystal oscillator is bypassed), the levels were at least 10dB below those of the oscillator under test for carrier offset frequencies below several KHz.

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

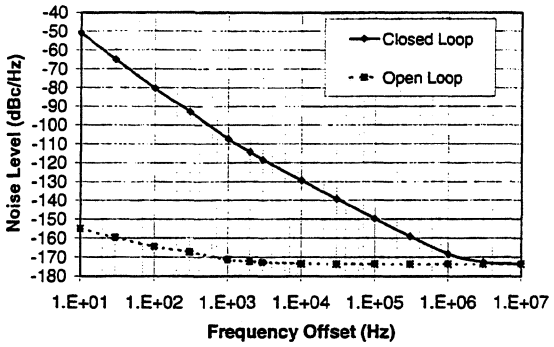


Figure 2. Typical Signal Phase Noise Relationships in a Low Loop (100nsec) Delay Oscillator

Measurement Results: Inductors and Capacitors

Figure 3 shows a measurement result for a 32MHz oscillator using a single pole, tuned circuit comprised of powdered iron (MS21422) inductors and ceramic dielectric (M39014/01) capacitors. The measured oscillator open loop delay was 88nsec, and the tuned circuit loaded Q was approximately 8.5. The closed loop phase noise at f=100Hz and 10KHz is –79dBc/Hz and –121dBc/Hz, respectively. Noting that $1/2\pi\tau$ = 1.8MHz, the values of oscillator open loop noise are (-79-85) = -164dBc/Hz at f=100Hz and (-121-45) = -166dBc/Hz at f=10KHz. The 100Hz noise level is approximately 2-3dB higher and the 10KHz noise level 5dB higher than that attributable to the (gain=20dB, P_{1dB}=18dBm, NF=4dB) sustaining stage amplifier, power divider, and pad alone for the case of amplifier operation at 1.5dB gain compression.

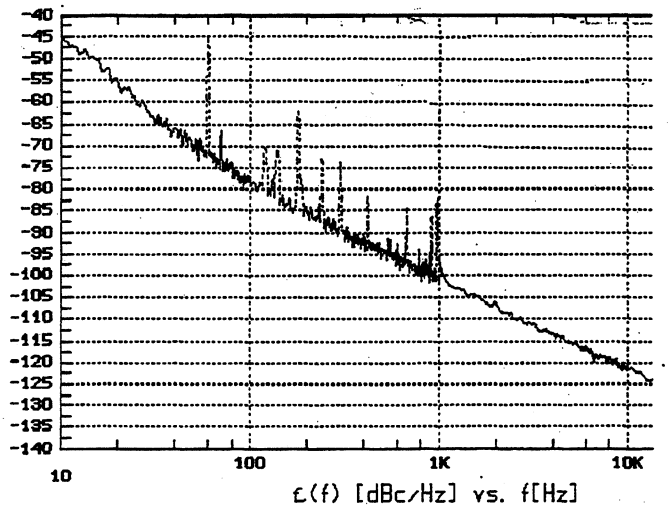


Figure 3. Phase Noise Measurement Results for a 88nsec Loop Delay, 32MHz Oscillator

However, even if the measured phase noise was considered entirely due to the tuned circuit components, the short-term reactance fluctuations attributable to the inductors and capacitors would correspond to a power spectral density given approximately by:

$$PSD(\Delta L/L)^2 \text{ or } PSD(\Delta C/C)^2 = 4X10^{-18}/f + 7X10^{-19} \quad (1)$$

It should be noted that the degree of carrier signal phase modulation induced by the short-term reactance fluctuations of an inductor or capacitor in a tuned circuit are "magnified". The relationship is given by:

$$PSD(\Delta L/L)^2 \text{ or } PSD(\Delta C/C)^2 = S\varphi(f)/QL^2 \quad (2)$$

Obviously, one drawback associated with the in-oscillator measurement scheme is that one can only measure down to the limit established by the sustaining stage amplifier/ power divider/pad combination in figure 1. However, these limits correspond to very low levels of component instability when the components are used in tuned circuits and/or filters in order to intentionally amplify the effect. Furthermore, component use in tuned circuits and/or band-pass filters represents an actual "real life" use. Similar results were obtained at 320MHz and 640MHz using smaller inductance and capacitance value, MS21423 type inductors and M39014/02 type ceramic and M23269/10 type glass capacitors.

Measurement Results: Multi-pole Band-pass Filters

Multi-pole, band-pass filters were also evaluated via use as oscillator frequency control elements. Figure 4 shows the phase noise measured for a 320MHz oscillator using a 32MHz bandwidth, 4 pole, band-pass filter. The oscillator open loop delay was 28nsec. The filter resonators were formed using air-wound inductors and ceramic chip capacitors. As shown in the figure, the closed loop phase noise level at 100Hz and 10KHz is –68dBc/Hz and –118dBc/Hz at 10KHz. This corresponds to open loop phase noise sideband levels of –163dBc/Hz at 100Hz and –173dBc/Hz at 10KHz. The 100Hz noise level is almost entirely attributable to the oscillator sustaining stage amplifier (gain=12dB, P_{1dB}=22dBm, NF=6dB) used. On the other hand, the oscillator white FM noise level at fm=10KHz (which corresponds to an open loop white PM noise) is 4-5dB higher than that attributable to the amplifier, even when the effects of noise floor degradation due to in-compression amplifier operation are taken into account.

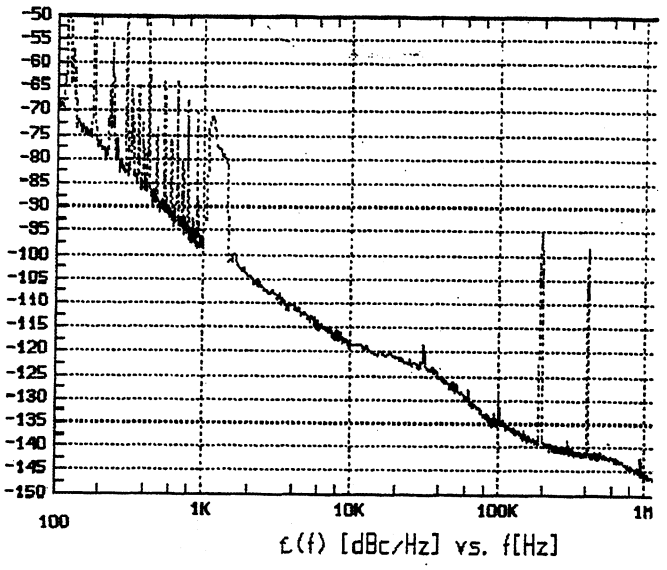


Figure 4. Phase Noise Measurement Results for Oscillator Incorporating a 4 pole 320MHz, L-C Band-pass Filter

In-oscillator phase noise measurements were also made using a higher (3usec.) delay, 30MHz, 4 pole helical resonator filter. In the case of the helical filter, the delay was large, and resulted in relatively lower levels of oscillator 1/f and white FM noise. Therefore, the HP8662A signal generator had to be operated using an externally supplied, 10MHz VCXO. Oscillator phase noise levels above carrier signal offset frequencies of several hundred Hz could not be measured because they fell below those of the HP8662A. The measured oscillator signal phase noise level of –110dBc/Hz at fm=100Hz corresponds to an equivalent open loop phase noise level of –165dBc/Hz and is also almost entirely attributable to the sustaining stage amplifier used.

Measurement Results: AGC Attenuators

A reverse-biased, varactor diode-based attenuator was designed and evaluated in a low loop delay oscillator in order to compare the PM noise with that of a (forward-biased) PIN diode-based attenuator [4]. The attenuator was configured as a tuned circuit and also served as the oscillator frequency control element. A schematic diagram for the attenuator is shown in figure 5. A plot of the measured oscillator (with attenuator) open loop response at various control voltage (attenuation) settings is shown in figure 6. The attenuator was designed to exhibit negligible center frequency phase shift vs. attenuation. Figure 7 shows the measured signal phase noise spectrum for a 29MHz, 66nsec loop delay oscillator using the varactor diode-based attenuator to accomplish an AGC. As shown in the figure, the oscillator signal phase noise sideband levels were –80dBc/Hz at fm=100Hz and –123dBc/Hz at fm=10KHz using the same amplifier used to obtain the figure 3 data. For 66nsec loop delay, this corresponds to oscillator open loop phase noise sideband levels of (-80-88)=-168dBc/Hz at fm=100Hz and (-123-48)=-171dBc/Hz at fm=10KHz, respectively. These levels are almost entirely attributable to the sustaining stage amplifier open loop phase noise and indicate that the short-term reactance fluctuations of the inductors, capacitors, and the two, hyper-abrupt varactor diodes in the AGC attenuator are well below $PSD(\Delta L/L)^2$, $PSD(\Delta Cv/Cv)^2 = 10^{-17}$ levels at 1Hz. Changing the attenuation values of the pads used on either side of the attenuator in the oscillator loop allowed a comparison in the phase noise levels to be made for different levels of attenuator drive (varactor diode RF voltage swing). No differences in oscillator signal phase noise level were measured for attenuator drive levels in the range 1mW to 30mW, corresponding to varactor diode RF voltage swings of 130mV peak to 700mV peak on varactor diodes nominally biased at/near 7 volts.

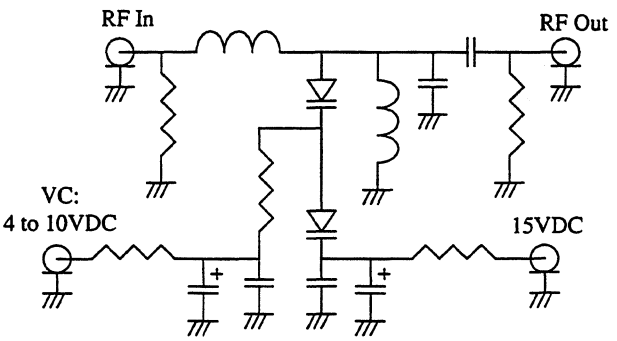


Figure 5. Schematic Diagram of Varactor Diode-Based, Voltage-Controlled Attenuator

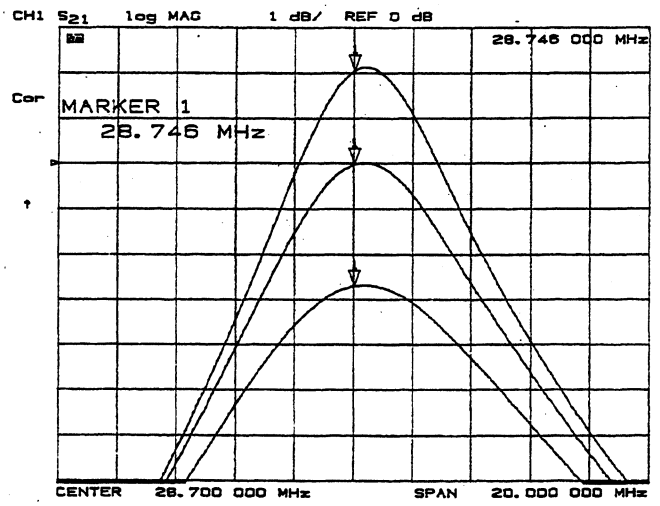


Figure 6. Measured Open Loop Gain of Low Loop Delay Oscillator with Varactor Diode-Based AGC Attenuator

In order to compare the PM noise of that varactor diode-based attenuator with that of a PIN diode-based device, a commercially available, connectorized, PIN diode attenuator was obtained. The PIN attenuator was used to perform AGC in the oscillator circuit in a manner similar to that used with the varactor attenuator. The PIN attenuator was operated at a relatively low attenuation setting where the published value of third order inter-modulation distortion levels were low (approx. -70dBc). In the case of the PIN attenuator, a LC tuned circuit similar to that used to obtain the figure 3 data was used as the frequency control element. The overall oscillator open loop delay was 88nsec, and the frequency of operation was 31MHz. The measured oscillator signal phase noise levels at fm=100Hz and 10KHz were –78dBc/Hz and –125 dBc/Hz, respectively, and they correspond to open loop phase noise levels of –163dBc/Hz and –170dBc/Hz.

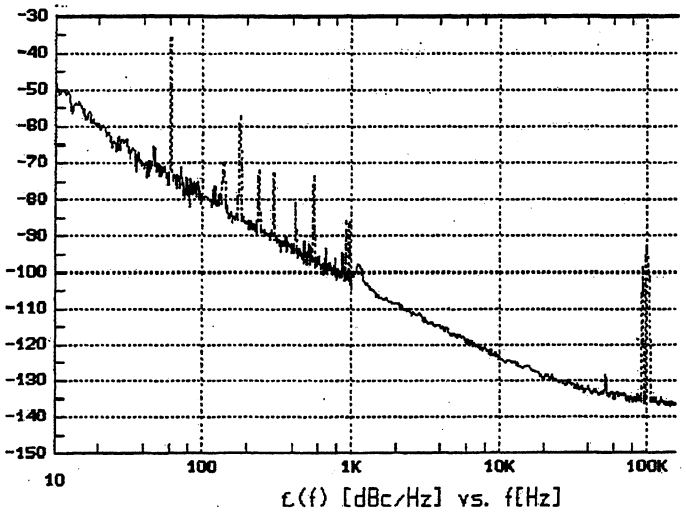


Figure 7. Measured Phase Noise of Oscillator with Varactor Diode-Based AGC Attenuator

It cannot be assumed from this data that the 1/f PM noise generated by the PIN attenuator was necessarily higher than that generated by the varactor diode-based attenuator. This is because the 100Hz open loop phase noise level, while 5dB higher, is consistent with that measured for a (figure 3 data) oscillator containing a tuned circuit similar to that used in the PIN AGC oscillator.

Conclusions

Figure 8 shows a generalized comparison between measurements of component short-term stability at VHF reported here and earlier referenced results obtained by NIST at HF. It is important to note that, because of the in-oscillator measurement method employed, the VHF results in the figure represent highest possible instability levels due, in most cases, to the sustaining stage amplifier and not the DUT. In spite of this, the trend is that the component short-term stability at VHF appears to be significantly superior to that measured for similar type devices at HF. This may be due to differences in device material properties or design in that per volume inductance and capacitance values tend to be much larger at HF.

The results indicate that evaluation of component short-term stability can be made down to levels as low as $PSD(\Delta X/X)^2 = 10^{-16}/f$ if/when the components are used in low loop delay oscillators. In this regard, the components need to be used in moderate Q factor, tuned circuits in order to intentionally enhance the carrier signal phase modulation sensitivity to component reactance fluctuation.

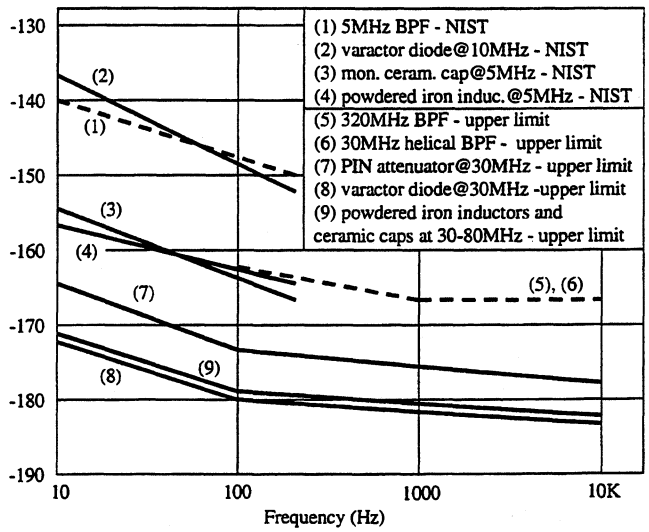


Figure 8. Comparison of Component Short-Term Stability Measured at HF (NIST) and VHF. Results for Filters and PIN Attenuator are on a $\mathcal{L}(f)=1/2[S\phi(f)]$ basis; all other results on a $PSD(\Delta X/X)^2$ Basis.

References

1. Ascarrunz, Zang, Ferre-Pikal, and Walls, "PM Noise Generated by Noisy Components", Proc. 1998 IEEE Freq. Contr. Symp., June 1998, pp. 210-217.
2. M.M. Driscoll, "Prediction of Acoustic Resonator Stabilized Oscillator Sustaining Stage Phase Noise Characteristics from Low Q Resonator Oscillator Noise Measurements", Proc. 1993 IEEE Freq. Contr. Symp., June 1993, pp. 202-208.
3. D.B. Leeson, "A Simple Model of Feedback Oscillator Noise Spectrum", Proc. IEEE, Vol. 54, No. 2, Feb. 1966, pp. 329-330.
4. E.S. Ferre-Pikal and F.W. Walls, "Voltage-Adjustable Attenuation with Low 1/f Noise", Proc. 1998 IEEE Freq. Contr. Saymp., June 1998, pp. 186-191.

1999 Joint Meeting EFTF - IEEE IFCS

A STUDY OF THE FREQUENCY STABILITY LIMITS OF OSCILLATORS BASED ON SAPPHIRE DIELECTRIC RESONATORS

J.G. Hartnett, E.N. Ivanov, M.E. Tobar
Department of Physics, University of Western Australia, Nedlands, WA, 6907, Australia

Email: john@physics.uwa.edu.au Tel:+61 8 9380 3443 Fax:+61 8 9380 1014

ABSTRACT

A novel approach to the design of ultra-stable microwave oscillators is presented. The method allows the short to medium term frequency stability to be accurately predicted *a priori* even before the beginning of the construction stage. This is achieved by characterising the fluctuations in various components and control systems of the oscillator in terms of their Allan variances. Stability limits are then determined by combining these results with the known Q-factors and frequency-temperature characteristics for both room temperature and cryogenic temperature compensated sapphire dielectric resonators.

1. FREQUENCY DISCRIMINATOR

The square root Allan variance (SRAV) of an oscillator is ideally determined by all contributing noise sources of the resonator and the frequency and temperature control systems. The most sensitive frequency control schemes that reduce the SRAV over short time scales usually comprise a Pound frequency discriminator[1]. A schematic is shown in figure 1. The SRAV of voltage fluctuations at the output of the Pound frequency discriminator, normalised to 1 volt, σ_u, is related to the SRAV of the fractional frequency fluctuations, σ_y^{FD} by

$$\sigma_y^{FD} = \frac{1}{S_{FD}} \frac{1\ Volt}{f_{RES}} \sigma_u \qquad (1)$$

where S_{FD} is the frequency discriminator conversion ratio and f_{RES} the resonator operational mode frequency. The conversion ratio may be expressed as

$$S_{FD} = \eta \gamma_P [P_{inc}^{res} K_{amp}^{mw} 4 J_0(\varphi) J_1(\varphi) \frac{2\beta}{1+\beta}] \frac{2}{bw} \qquad (2)$$

where η is the RF mixer conversion ratio, γ_P is the diode detector conversion ratio (in Volts/Watt), P_{inc}^{res} is the power incident on the resonator, K_{amp}^{mw} the gain of a microwave amplifier placed in front of the detector. J_0 and J_1 are Bessel functions, φ the modulation index, bw the bandwidth of the operational mode and β the resonator coupling coefficient. If the detector is operating in the small signal regime, γ_P is independent of power and can be assigned a constant value. However, this is not true at high input power. In order to have an idea of the voltage fluctuations of the discriminator, independent of other resonator noise sources, the resonator was replaced with an equivalent load.

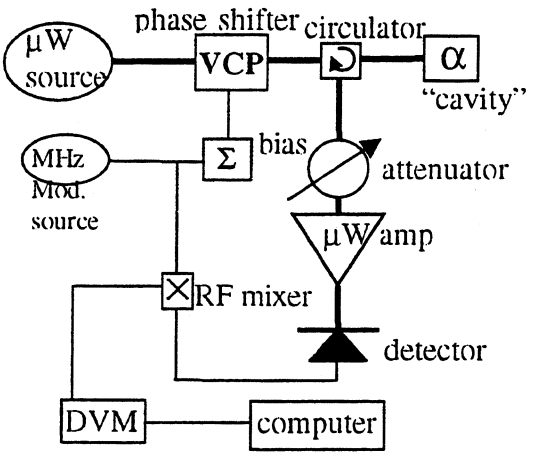


Figure 1: Schematic of Pound frequency discriminator circuit. Thick lines are microwave circuit and the thin lines are the RF circuit. The microwave source was supplied at 9 GHz. The "cavity" was an attenuator set to an equivalent coupling of 0.8. A reference voltage was used for the bias and results were read out by a digital volt meter (DVM) and recorded by a computer.

Three voltage controlled phase shifters (VCP) were investigated and all had similar voltage to phase characteristics (curve 1 of figure 2) but quite different insertion loss behaviour (curves 2-4 of figure 2). From equation (2) it is clear that in order to maximise S_{FD}, coupling, β, should be as close to unity as possible, the bandwidth, bw, should be as small as possible (ie. the Q factor as high as possible) and $J_0(\varphi)J_1(\varphi)$ should be maximised (this occurs when the modulation index φ = 1.08). Another consideration is that spurious AM modulation is produced by the VCP and this is minimised when the gradient of the insertion loss-bias voltage characteristic goes to zero, ie. when a turning point exists. This is true for curves 2 & 3 in figure 2 but not for curve 4. Voltages noise measurements were performed for various modulation bias voltages and amplitudes of the modulating supply (which in turn translates to different modulation index values).

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

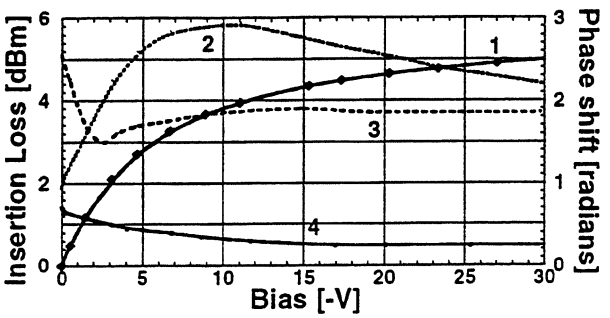


Figure 2: Phase shift as a function of bias voltage for a VCP (curve 1) and insertion loss for three different VCPs (curves 2-4).

The frequency discriminator (figure 1) was configured with and without a microwave amplifier in front of the detector. The net gain introduced by the amplifier was varied by a variable attenuator, placed in front of the microwave amplifier. Two different detectors were tried and the modulation frequency was initially chosen at 1 MHz. The SRAV of the relative voltage fluctuations are plotted in figure 3 along with the noise floor. Adding the microwave amplifier further lowered the noise floor by the introduced net gain.

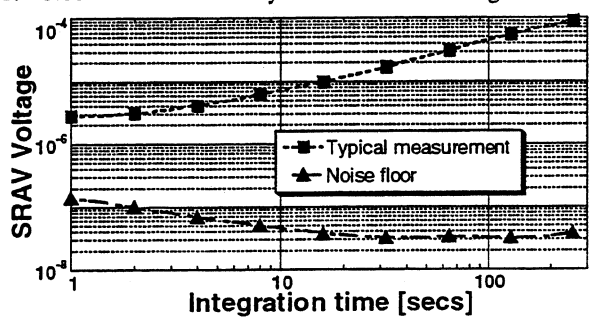


Figure 3: A typical SRAV voltage measurement without the microwave amplifier in the Pound circuit. The noise floor was determined by measuring voltage fluctuations into the DVM from a 50 Ohm termination.

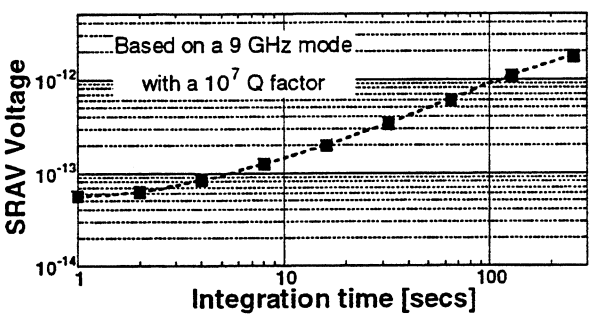


Figure 4: Predicted fractional frequency SRAV extrapolated from the voltage SRAV of VCP#1 shown in figure 3.

The SRAV voltage data can be converted to fractional frequency fluctuations using equations (1) and (2). For an operational mode at 9 GHz with a Q factor of 10^7 (typical for a frequency temperature compensated 77 K sapphire resonator[2]) and P_{inc} of 1 dBm the data of figure 3 has been converted and is plotted in figure 4.

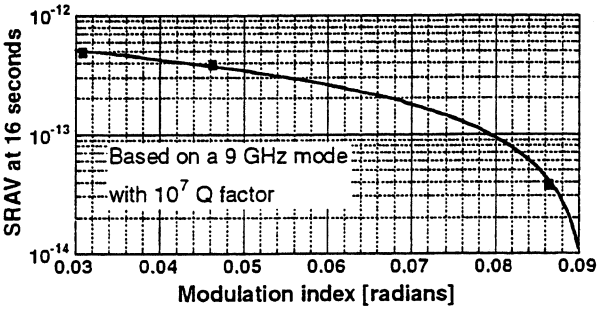


Figure 5: The SRAV frequency fluctuations for VCP#3 as a function of the modulation index, at 16 seconds integration time and constant power on the detector, without the microwave amplifier. This VCP has almost constant insertion loss vs bias (curve 4 of figure 2).

For all three VCPs, the modulation index was varied at specific bias voltages, and in all cases the SRAV improved with increasing φ. A linear dependence of the SRAV at small modulation index was observed (figure 5) as expected from equation (2) (when φ is small $J_0(\varphi)J_1(\varphi) = \varphi/2$). To test the effect of the insertion loss-bias turning point, the SRAV was compared with constant φ at different bias voltages. A minimum in SRAV was always observed at the turning point.

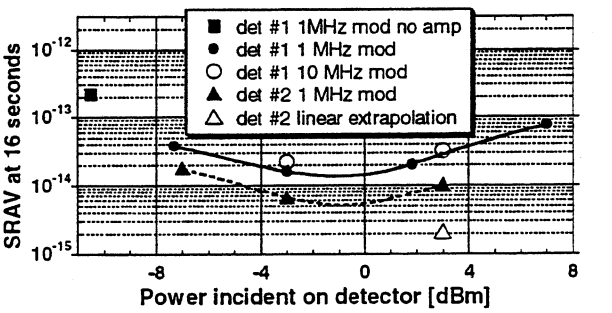


Figure 6: The SRAV frequency fluctuations of VCP#1 as a function of incident detector power, at 16 seconds integration time and a constant modulation index of 0.07.

In order to further optimise the discriminator the microwave amplifier was added to the circuit. An attenuator was included to keep the input power to the amplifier between –40 to –26 dBm. The SRAV of VCP#1 was measured at the same modulation index and integration time as previous and is shown in figure 6. The data point at –10.5 dBm (solid square) shows the SRAV without the amplifier and the solid circles show the SRAV with the amplifier as a function of power into the detector.

A reduction in the SRAV was observed, which can be understood in terms of the noise temperature, T_{RS} of the readout system. Without the amplifier it maybe written as $T_{RS} = T_0 + T_{det}$ but if we add an amplifier it becomes

$$T_{RS} = T_0 + T_{amp} + \frac{T_{det}}{K_{amp}^{\mu w}} \quad (3)$$

where T_0 is ambient temperature, T_{amp} is the amplifier noise temperature, T_{det} is the detector noise temperature and $K_{amp}^{\mu w}$ the gain of the amplifier. T_{det} typically is of the order of millions of degrees K and by adding the amplifier gain, T_{RS} was significantly reduced.

A minimum is reached in the SRAV as a function of input power to the detector at about −1 dBm (figure 6). Detector #1 (solid circles) was replaced with detector #2 (solid triangles) which saturates at a higher input power. Ideally a detector which has a linear power dependence at high input power is desired. The open triangle at 3 dBm is the predicted value if detector #2 remained linear. It was suspected that this minimum was a result of the flicker corner of the detector becoming greater than the modulation frequency. Thus, the measurements with detector #1 were repeated with a 10 MHz modulation frequency. However, degradation of SRAV was recorded (open circles). Measurements of the flicker corner showed that at 3 dBm input to the detector the flicker corner was still within the 1 MHz bandwidth. The minimum was a result of the saturation of the detector.

2. TEMPERATURE STABILITY

A copper cavity loaded with a 3 cm diameter sapphire resonator was inserted into a bench top vacuum chamber with a thermoelectric (Peltier) module. The cavity was cooled to 10 °C and stabilised through a lock-in circuit (patented) supplied by PSI Pty Ltd. The circuit employed a thermistor bonded to the base of the cavity close to the Peltier element. The thermistor was incorporated in an ac bridge to control the temperature around a set point. To monitor the temperature fluctuations in the Cu cavity another nominally identical thermistor was attached close to the other. Voltages were read out by a DVM and recorded on a computer. An additional thermistor was attached to the surface at the centre of the sapphire cylinder, to measure the temperature fluctuations at the sapphire element.

The SRAV of voltage fluctuations normalised to 1 volt was measured (figure 7). The SRAV of fractional temperature fluctuations was calculated from the voltage fluctuations and the thermistor characteristics. To reduce the effects of the ambient temperature fluctuations the apparatus was placed in an insulated box (curve 2 in figure 7).

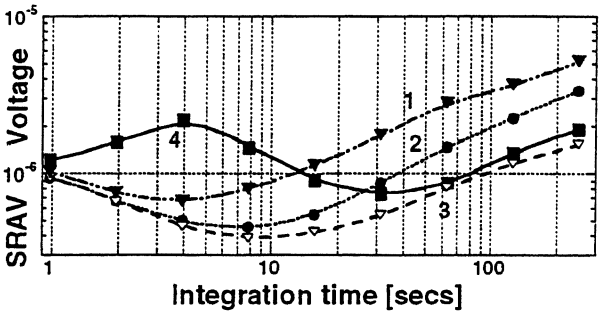


Figure 7: The SRAV of relative voltage fluctuations for the thermistor attached to the surface of the sapphire with and without an insulated box (curves 2 & 1 respectively). The SRAV voltage for the thermistor (curve 4) and the metal resistor (curve 3) attached to the base of the cavity. The metal resistor (curve 3) determines the noise floor of our measurement system.

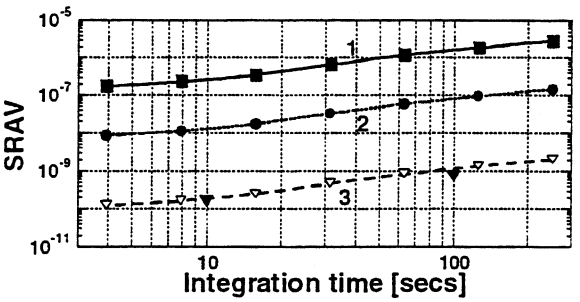


Figure 8: Curve 1 shows the SRAV of voltage fluctuations at the sapphire surface. Curve 2 shows the SRAV of temperature fluctuations calculated from curve 1. Curve 3 shows the inferred SRAV of fractional frequency fluctuations from curve 2 and the known temperature coefficient of sapphire[3]. The solid black triangles are measured values[4] of Allan variance for two nominally identical 9 GHz oscillators using the same temperature controller.

3. DISCUSSION

Very good agreement is observed between the predicted SRAV from temperature measurements and the measured SRAV from beat frequency measurements. This allows us to then predict *a priori* the SRAV for other oscillators at both room temperature and at liquid nitrogen temperature.

Based on a rutile/sapphire compensated resonator[2] a curvature of 3×10^{-7} can be expected at 77 K. The above ac lock-in has a temperature resolution of 3 - 30 μK so to remain within 1mK of the compensation point should not be difficult. Therefore, assuming the known coefficients for a platinum

thermometer the SRAV of fractional frequency due to temperature control is estimated and plotted in figure 9. Curve 4 is the calculated SRAV for a 77 K resonator-oscillator with temperature-frequency compensation and curve 3 is without. Curves 2 and 1 are room temperature calculations for thermistor controlled cavities with and without turning points respectively. The curvature for curve 2 is estimated from the 5 times larger gradient of df/dT of sapphire at room temperature compared to 77 K.

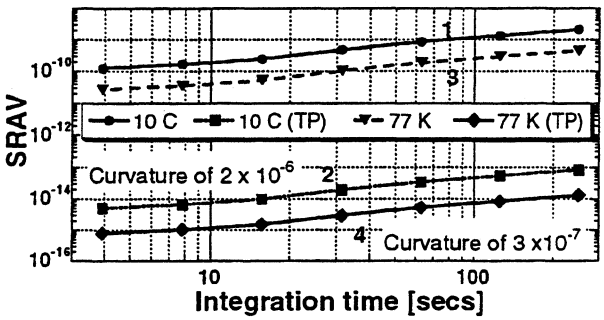


Figure 9:The projected SRAV of fractional frequency fluctuations based of the temperature stability data of figure 8. TP indicates a turning (compensation) point in the frequency temperature dependence of the sapphire resonator.

The use of an amplifier in the Pound frequency discriminator scheme presents an improvement, particularly if we consider the situation where the cavity is near critically coupled and the option of high modulation index is not available. The incident power on the resonator may be increased in such a case with further improvement in S_{FD}. At the "normal" optimum value of the modulation index (that is, φ = 1.08 radians) nearly 50% of the power is in the side bands. Even if the carrier is suppressed, the sidebands are reflected and would saturate the amplifier and detector if the power incident on the cavity was high. But the same discriminator conversion ratio can be maintained at a very low value of φ where little power is in the sidebands by increasing the power incident on the cavity, together with good carrier suppression. In addition the improvement in the discriminator noise temperature by adding an amplifier results in further gains.

Therefore, *a priori* we may predict the Pound frequency discriminator noise limits based on our assessments and voltage noise measurements. Further improvements over those plotted in figure 4 may be expected by increasing the power on the cavity, near critically coupling the cavity, using an amplifier when the modulation index is chosen to be small, and choosing a detector which saturates at a higher input power.

4. CONCLUSION

In an optimal frequency stabilised oscillator the electronic noise floor of the discriminator determines the SRAV. Our results infer a SRAV of 5.6×10^{-15} due to the discriminator at 10 seconds integration time at room temperature. Also, a lockin circuit employing an ac bridge and a thermistor to control the temperature of a sapphire loaded cavity has been implemented. The relative temperature fluctuations of a thermistor on the sapphire determined the limit due to the temperature control system. At room temperature, and at 10 seconds integration time, a SRAV of 2×10^{-10} is predicted where no frequency temperature compensation point exists, and a SRAV of 5×10^{-15} if it does. At 77 K, 1.2×10^{-15} and 4×10^{-11} are the predicted limits for an oscillator operating on a mode with and without a turning point respectively. This assessment has been based on known thermistor characteristics at room temperature and platinum at 77 K, however further investigation is needed to find a temperature sensor with a higher temperature coefficient than platinum at 77 K with low noise. These results suggest that it is possible to construct ultra-stable fly-wheel oscillators at 77K suitable for use with the new generation of frequency standards.

5. ACKNOWLEDGEMENTS

The authors wish to acknowledge the contribution of the ARC in funding this research through the ARC's Collaborative Research program SPIRT Scheme in conjunction with Poseidon Scientific Instruments Pty Ltd. Also we acknowledge Paul Stockwell and Cameron McNeilage of PSI Pty Ltd for their valuable assistance with the temperature controller.

6. REFERENCES

[1] R. V. Pound, "Electronic Frequency Stabilization of Microwave Oscillators," *Rev Sci Intruments*, vol. 17, pp. 490-505, 1946.

[2] M. E. Tobar, J. Krupka, J. G. Hartnett, E. N. Ivanov, and R. A. Woode, "High-Q sapphire-rutile frequency-temperature compensated microwave dielectric resonators," *IEEE Trans. on UFFC*, vol. 45, pp. 830-836, 1998.

[3] M. Tobar, J. Krupka, R. Woode, and E. Ivanov, "Dielectric frequency-temperature compensation of high quality sapphire dielectric resonators," *Proc. IEEE Int. Freq. Cont. Symp.*, pp. 799-806, 1996.

[4] C. McNeilage, "private communication," 1999.

INVESTIGATION OF DUAL-MODE EXCITATION OF CRYSTAL OSCILLATOR

Anatoly V. Kosykh, Alexander N. Lepetaev, Sergey A. Zavjalov.

Omsk State Engineering University, 644050, 11 Mira avenue, Omsk, Russia
e-mail: info@most.omsk.ru

ABSTRACT

The concepts of dual-mode oscillator design have been discussed in this work. Considerable attention has been given to the constant B-mode excitation level keeping under external factors effects and resonator motional resistance variation. The goal of the work submitted here is development of suitable for production dual-mode oscillators of small size intended for DTCXO or OCXO.

Keywords: Dual-mode oscillator, temperature-compensated oscillator.

INTRODUCTION

One of the main problems in thermocompensation and thermostabilization technique is piezocrystal temperature measuring. An external thermosensor application leads to the appearance of static and dynamic temperature measurement error. As a result is temperature stability worsening especially under non-uniform temperature conditions [1]. The most effective method of piezocrystal temperature measurement accuracy increasing is the use of additional mode of the same crystal as a temperature sensor [2]. The main requirement is high abruptness of temperature vs. frequency curve of the "thermosensive" mode. Anharmonics of fundamental oscillation mode [3], frequency difference between the 1-st (multiplied by 3) and 3-rd harmonics [4,5,6], or C-mode (in resonator of double rotated cut, for example SC-cut) are used as "temperature" modes. The problem of resonator type choice is considered in detail in [7]. From our point of view the use of C-mode in SC-cut resonators is the most preferable.

However the simultaneous dual-mode excitation makes serious problems in providing stability of two oscillations. Technical methods allowing to overcome these difficulties are considered in [7]. The scheme and technical parameters of dual-mode oscillator designated for synchronizing frequency synthesizer with thermodepended (on mode B frequency) frequency control are given in the article. The dual-mode oscillator itself in that scheme appears to be uncontrolled. This simplifies stabilization of stationary mode oscillation. In some cases, however one needs dual-mode oscillator to be fundamental mode frequency controlled (e. g. in thermocompensated oscillators with open regulation loop). The complexity of the problem increases because the requirements are imposed on the oscillator as well as on the band and linearity of change.

1. SUBSTANTIATION OF REQUIREMENTS TO OSCILLATOR

Dual-mode crystal oscillator which is designated for the use in oscillator with digital thermocompensation must meet the following requirements:

- the fundamental mode must have a flat F-T curve with minimal swing in the whole band of operation frequencies;
- thermosensitive mode must have monotonous F-T curve (it's advisable to be close to linear) with relatively stable motional resistance in the temperature range (2-4 times smooth Rq changes are possible);
- the amplitude of reference and thermosensitive oscillator output signal must be sufficient for formation of logical levels (these signals are used for further processing);
- it is desirable to use low- frequency difference signal (Fb-Fc) as thermosensor signal;
- the range of fundamental mode frequency tuning must exceed the swing of temperature vs. frequency curve in the operating temperature interval (not less than $8x10^{-5}$ at 0-5 V voltage control;
- non-linearity of control law shouldn't exceed 10%;
- energy consumption must be minimal;
- the level of side-line components (B-mode and summary-difference B and C-mode combinations) must be minimal, no more that -50 dB;
- oscillator must have low frequency sensitivity to the changes of supply voltage.

Besides, the oscillator must have minimal number of tuned coils and inductances (as it is done in [8]).

2. BASIC OSCILLATOR CIRCUIT

SC-cut resonator with the first harmonic C-mode frequency of 10 MHz are used in the work. The choice of resonator cut is determined by next reasons:

- SC-cut has minimal dynamic temperature frequency coefficient;
- in SC- cut C-mode frequency vs. resonator drive current is low;
- these resonators have low aging rate;
- thermosensitive B-mode has abrupt F-T curve(330 ppm in the temperature range).

0-7803-5400-1/99/ $ 10.00

For simultaneous dual-mode resonator excitation (B and C modes) the modified Kolpitts circuit was used. The difference in single-mode and dual-mode circuit is shown in Fig. 1A and 1B.

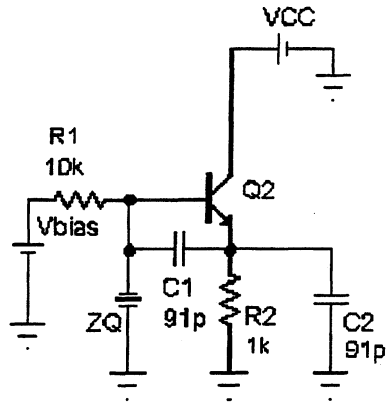


Fig. 1A.

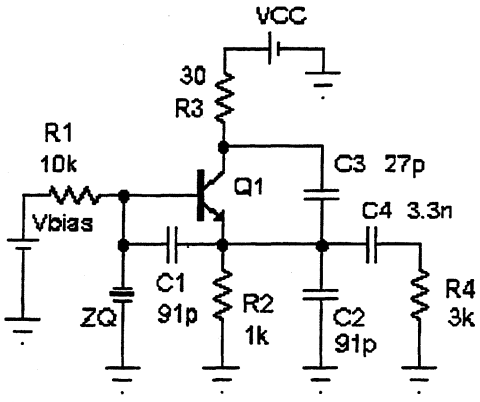


Fig 1B.

3. MODELING AND CALCULATION

For estimation of dual- mode excitation conditions the following method of analysis is suggested. The essence of this method is in calculation of two-dimensional chart of the oscillator's input impedance. This chart allows determining dual-mode excitation regime for the given frequencies under known ratios of active resistance's and Q-factors of both resonator modes. Besides this method means the estimation of transient process of resonator amplitude and current's frequency setting for each mode.

In the analyzed dual-mode oscillator model the crystal resonator is presented as two ideal sources of sinusoidal current (Fig 2), connected in parallel to resonator static capacitance (this capacitance will be taken into account below as an element of oscillator circuit). When input currents of each mode pass through input oscillator clamps the alternate voltage will appear on them (which in general case will be unsinusoidal. This speaks for the circuit non-linearity).

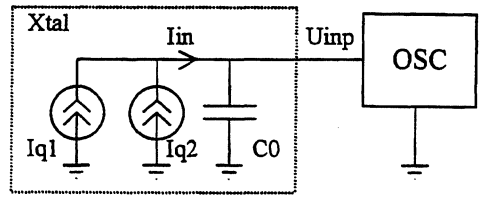


Fig 2

Taking into account energetic considerations, non-linear circuit of self-excited oscillator can be changed into linear one (as complex effective resistance). In this case the balance of active and reactive parts of energy (which is being given by each source of current modes and being dispersed on equivalent input resistance of oscillator) must be observed. Total energy in the steady state regime for each mode must be equal to zero.

It's easy to prove that effective resistance of oscillator for each mode can be calculated from the expression

$$Zg_k = \frac{2}{Iq_k^2 \cdot T}(Wr_k + j \cdot Wi_k) \qquad (1)$$

where Zg_k - effective input resistance of oscillator at the k-th mode frequency (k = 1...2),

Iq_k - resonator k-th mode current amplitude,

T - integration period (integral number of periods of both mode oscillations must be kept within this period).

Wr, Wi - real and imaginary part of energy at input of oscillator during time T.

The component parts of energy can be expressed by formulas

$$Wr_k = Iq_k \cdot \int_0^T U_{inp}(t) \cdot \sin(\omega_k t)dt \qquad (2)$$

$$Wi_k = Iq_k \cdot \int_0^T U_{inp}(t) \cdot \cos(\omega_k t)dt \qquad (3)$$

Input voltage of oscillator $U_{inp}(t)$ can be determined as circuit reaction on input current $I_{in}(t)$ at the set regime: $I_{in}(t) = Iq1 \cdot \sin(\omega_1 \cdot t) + Iq2 \cdot \sin(\omega_2 \cdot t)$, where ω_1, ω_2 - frequencies of the corresponding modes.

Calculations of the expressions (1)...(3) were made by using program packet MICROCAP-5 at discrete points of input current meanings Iq1, Iq2 in the current's band 0...5mA for each mode. The received massif of discrete meanings of Zg_k is used as basic data for making continuos functions $Zg_k = f_k(Iq1,Iq2)$ by means of MATHCAD-7.

Fig.3A and Fig.3B show real parts of effective input resistance of the oscillator (Rg1 and Rg2) under investigation for modes C and B correspondingly.

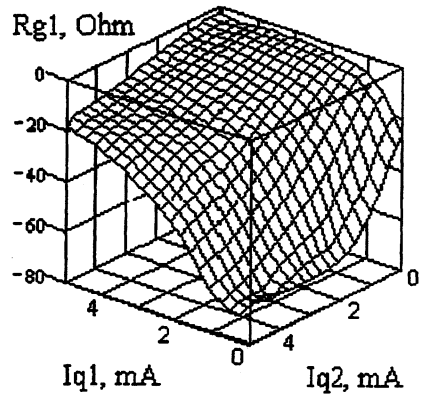


Fig 3A.

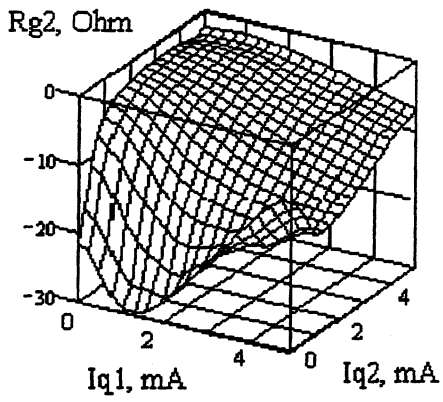


Fig. 3B.

The steady-state regime of oscillator describes by expression:

$$Zq_k+Zg_k=0, \qquad (4)$$

which is equivalent to the conditions of energy balance (1)...(3), where $Zq_k=Rq_k+j\cdot Xq_k$. Here Rq_k is the resistance of crystal oscillator losses (in the given model it is considered to be inde pendent of mode current's value), Xq_k - reactive resistance of crystal resonator. Besides the solution of the equation system (4) must possess stability. Fig. 4A and Fig. 4B presents functions

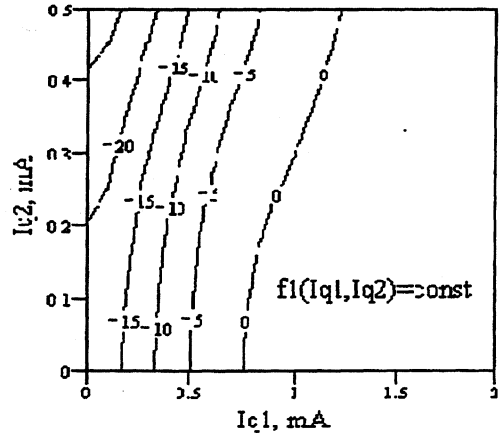


Fig. 4A.

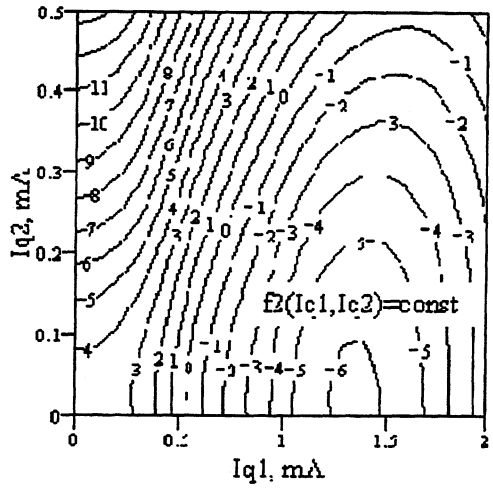


Fig. 4B.

$f1=Rq_1+Re(Zg_1)$ and $f2=Rq_2+Re(Zg_2)$ as level lines.

In this case it was considered that $Im(Zq_k)$ is definitely determined by the value $Im(Zg_k)$, that is the classical condition of phase balance. Initial data for calculation of oscillator circuit (Fig. 1B): Rq_1 = 8 Ohm, Rq_2 = 25 Ohm, Q_1 = 377000, Q_2 = 55000, F_1 = 10MHz, F_2 = 11MHz.

Amplitude balance for each mode corresponds to the lines of zero level f1=0, f2=0 (Fig 4). Dual-mode regime corresponds to points of intersection of zero level lines. It is seen from the figures that the point of intersection of zero level line really exists. From the practical point of view the process of dual- mode oscillator self-excitation and estimation of its stability at the point of zero level line is of great interest. The following assumptions were used:

1. Current amplitudes Iq_k are considered to be constant within the limits of one period of oscillations.
2. $Iq_k >> \Delta Iq_k$, where ΔIq_k - current increment for the period of oscillation $T_k=1/F_k$.

Is is easy to prove that the transition process under such assumptions is expressed in the following way:

$$\frac{dIq_k}{dt} = -\frac{Iq_k \cdot f_k(Iq_1, Iq_2)}{Rq_k \cdot \tau_k},$$

where $\tau_k=2\cdot Q_k/\omega_k$ (resonator's k-th mode time delay).

As the initial conditions the following currents are chosen: $Iq_k(0)=(Rq_k\cdot Q_k)^{-1}$ (currents initiated by voltage jumps up to 1V in the oscillator equivalent circuit).

The transient process of oscillations setting is represented on Fig. 5 as hodographs. The dashed lines on this Figure are the lines of a zero level for functions $f_1(Iq_1,Iq_2)$ and $f_2(Iq_1,Iq_2)$.

The graphs of transients contain obtained at the different initial conditions hodographs . The line 1 is

obtained at the initial conditions for calculated from expression $Iq_k(0) = (Rq_k \cdot Q_k)^{-1}$ (currents of modes are initiated by a 1V voltage jump).

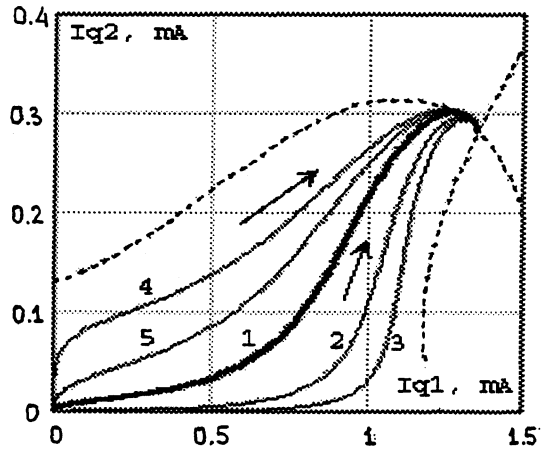


Fig. 5.

For resonators parameters mentioned above, $Iq_1(0) = 2.7 \cdot 10^{-4}$ mA, $Iq_2(0) = 6.3 \cdot 10^{-4}$ mA. For remaining lines the initial conditions are chosen arbitrary and their values (in 10^{-4} mA) are represented in a Table

	Curve 2	Curve 3	Curve 4	Curve 5
$Iq_1(0)$	3	100	1	1
$Iq_2(0)$	1	1	100	20

The similar hodographs of oscillations setting process (for other values of modes motional resistances a resonator (Rq_1 =10 Ohm, Rq_2=29 Ohm) but with the same initial conditions) are shown on Fig. 6.

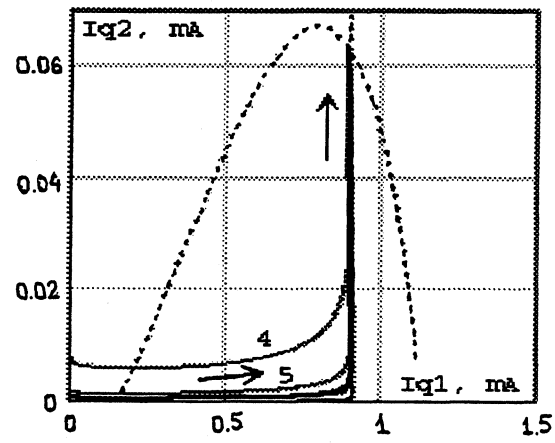


Fig. 6.

Fig. 7. shows the single-mode oscillation regime, having place when Rq_1 =8 Ohm, Rq_2=29 Ohm.
The comparison of Figures shows: the small Rq modification can essentially change a self-excitation condition of the second mode - from soft excitation up to rigid one.

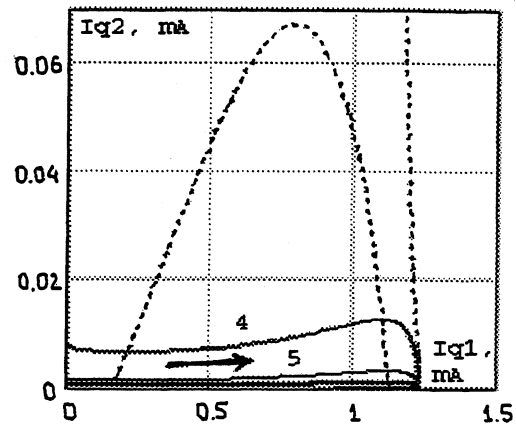


Fig. 7.

4. CONCLUSION

Presented method allows calculate both static and dynamic regimes of dual-mode oscillation. Except mentioned above, transient process of spectral components setting may be determined. This method makes it possible to optimize dual-mode oscillator circuit under

REFERENCES

1. Kosykh A., Ionov B., Vasiliev A. Temperature dynamic characteristics and temperature-dynamic compensation of crystal oscillators. Proc. 46-th A.F.C.S.., 1992. p. 425-433.
2. Kusters J., Leach J., Fisher M. Quartz resonator cut to compensate for static and dynamic thermal transient. US Patent N 4079280, 1978.
3. Shmaly Y. The modulation method of the precision quartz-crystal oscillators and standards frequency stabilization. Proc. 49-th A.F.C.S., 1995, p. 579 -589.
4. Marianovsky L., Vasecky G. Crystal oscillator. USSR Patent N 758472. 1980.
5. Shodovsky S. Resonator self-temperature sensing using a dual-harmonic mode crystal oscillator. Proc. 43-rd A.F.C.S., 1989, p. 2-7.
6. Benjaminson A., Stalling S. A microcomputwr compensated crystal oscillator using a dual-mode resonator. Proc. 43-rd A.F.C.S. 1989, p.20-26.
7. Kosykh A., Abramson I., Bagaev V. Dual-mode crystal oscillators with resonators excited on B and C modes. Proc. 48-th A.F.C.S., 1994, p. 578-586.
8. Research work report N01830010869. VINITI, USSR,. Bagaev V, Samoilenko V., Kosykh and other.

1999 Joint Meeting EFTF - IEEE IFCS

NOISE IN OSCILLATORS WITH TWO ASYNCHRONOUS OSCILLATIONS

Dmitry P. Tsarapkin*and Fred L. Walls**

* MPEI, Krasnokazarmennaya 14, Moscow, 111250, Russia
** NIST, Time and Frequency Division, 325 Broadway, Boulder, CO 80303, USA

ABSTRACT

This paper describes noise phenomena in oscillators with two degrees of freedom, sustaining two-frequency asynchronous oscillations. The oscillator is comprised of two parallel high-Q-resonant, RLC circuits connected in series to an active voltage-controlled one-port device with a symmetric volt-ampere characteristic having a nonlinearly described by a function arctg(x). The oscillator under analysis is isochronous, as it uses a purely resistive active device and has no additional phase shifts in the positive feedback loops. Phase noise at the two frequencies due to white noise sources is uncorrelated, while amplitude noise shows some mutual correlation. The main features of the noise characteristics arise due to the interaction of the two asynchronous oscillations via the common bias.

1. INTRODUCTION

Asynchronous oscillation in dynamic systems with two degrees of freedom is a classical problem of nonlinear oscillation theory [1]. Aside from pure scientific interest, such oscillations are worthy of our attention, since they can be used to improve frequency stability of precision quartz oscillators [2-7] and can cause distortions in microwave and other sources [8].

A typical arrangement of the oscillator under analysis (Fig. 1) is comprised of two parallel high-Q-resonant RLC circuits connected in series to an active voltage-controlled non-linear one-port device. If the tank circuits resonant frequencies ν_1 and ν_2 are incommensurable and sufficiently separated, the total voltage waveform across the active device (AD) in a steady-state regime u(t) consists of a bias voltage u_o and one or two fundamental voltage components $u_1(t)$ and $u_2(t)$:

$$u(t) = u_1 + u_2 + u_o = U_1 \cos(2\pi\nu_1 t + \phi_1) + U_2 \cos(2\pi\nu_2 t + \phi_2) + u_o. \quad (1)$$

In general, the dc term u_o is not necessarily identical to the applied bias voltage U_{00} since there may be some rectification of the rf voltage components.

This work was sponsored by the US National Institute of Standards and Technology under a grant #RE0-609 administered through CRDF and partly by the Russian Foundation for Basic Investigations under a grant #98-02-17210.

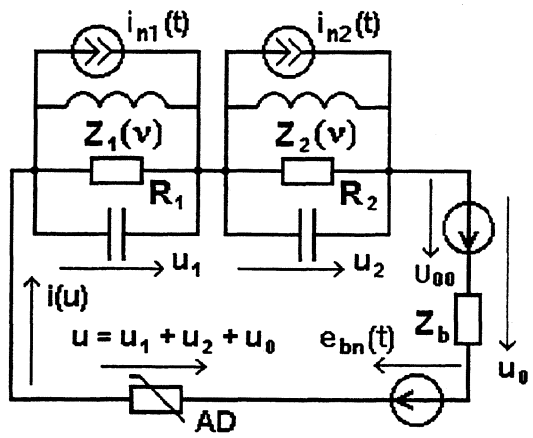


Figure 1. An equivalent network of an oscillator sustaining two-frequency asynchronous oscillations with added noise sources.

The kind of oscillation which actually occurs in the network depends, first of all, on the particular AD non-linearity. Van der Pol came to the conclusion that stable oscillation at two asynchronous frequencies simultaneously could not occur if the AD volt-ampere characteristic (VAC), i(u), was represented by a cubic relation [1]. Instead, the device oscillates at either one or the other frequency, depending on initial conditions. It was soon was discovered that this result is only a consequence of the specific form of VAC and that for an appropriate VAC, simultaneous oscillation at two frequencies is possible. Apparently, the first time this was studied was by Chikhachev [9]. Twelve years later Skinner [10] showed that for stable asynchronous oscillations to occur, there must be at least a fifth-order term in the power-series expansion of VAC, that is

$$i(u) = a_o + a_1 u + a_2 u^2 + a_3 u^3 + a_4 u^4 + a_5 u^5. \quad (2)$$

where the parallel resonant (antiresonant) impedances of the two tank circuits, R_1 and R_2, be not too unequal. However, as pointed out by Schaffner [11] and Anisimov [9], these conditions for oscillations are not self-starting because the effective resistance of the system is positive for small oscillations. A 5th-order polynomial VAC, which lead to sharply excited asynchronous oscillations, can only have about the same values for the stationary amplitudes U_1 and U_2. This restriction, however, does not apply if VAC

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

is represented by a 7th-order polynomial [8].

The common peculiarity of analysis in [8, 9, 11] is a suggestion that the AD output current responds to an instantaneous value of the applied voltage while parameters of the VAC and an operating point stay unchanged. Disman and Edson [10] checked the last condition and found that replacement of a constant bias by an automatic one, when the bias is derived by rectification of the rf voltage applied to AD, changes the oscillator properties drastically. In particular, mild excitation of stable asynchronous oscillations are now possible. Additional publications in this area revealed interesting details about asynchronous oscillations in different cases but did not change the understanding.

2. STEADY STATE REGIMES

In contrast to other work, the AD described here has a symmetric VAC, which in our opinion, is a better match to many practical cases. A specific form of the VAC nonlinearity is represented by the function

$$i(x) = A[(2/\pi)\,\mathrm{arctg}(x) + 1], \qquad (3)$$

where $x = C_n \times u$ is the normalized total instantaneous voltage across the one-port AD, and C_n is the normalizing coefficient. The VAC chosen has odd symmetry relative to the argument value $x = 0$.

The analysis is based on the method of symbolic shortened equations (SSE method) developed by Evtyanov [12]. The partial oscillations without noise components are written as

$$u_k(t) = U_k\cos\Psi_k(t) = U_k\cos[2\pi\nu_k t + \phi_k(t)]. \qquad (4)$$

Bearing in mind that $Q_{1,2} \gg 1$ for resonators of interest, the amplitudes $U_{1,2}(t)$ and the phases $\phi_{1,2}(t)$ can be considered slowly varying functions of time. This permits us to replace instantaneous values of variables by their envelopes and thus to lower the order of the initial differential equations. Using the SSE method to derive simplified differential equations, requires us to approximate the circuit impedance in the vicinity of all operating frequencies and use the expression obtained as differential operators acting on the complex amplitudes $\mathbf{U}_k(t) = U_k(t)\exp[\phi_k(t)]$.

The main equations describing behavior of the oscillator can be obtained from its equivalent circuit in Fig. 1. For single tank circuits, the approximated impedances, $\mathbf{Z}_1$ and $\mathbf{Z}_2(\mathbf{p})$, have the form

$$\mathbf{Z}_i = R_i/(1 + T_i\mathbf{p}), \qquad (5)$$

where $R_k = c^2_k\,\rho_k\,Q_k$ is the parallel resonant impedance, c_k is the coupling coefficient, $\rho_k = (L_k/C_k)^{1/2}$ is the characteristic impedance, Q_k is the loaded Q-factor, $T_k = 2Q_k/\omega_{0k}$ is the resonator time constant, $\mathbf{p} = d/dt$ is the differential operator with respect to time, $k = 1, 2$. In a steady-state regime, $\mathbf{p} = j\Omega$, where $\Omega = 2\pi f = \omega - \omega_{0k}$ is the angular offset frequency.

$\omega_{0k} = 2\pi\nu_{0k} = (L_kC_k)^{-1/2}$ is a self-resonant angular frequency, and $f = \nu - \nu_{0k}$ is the frequency shift in hertz with respect to the current frequency ν and also a Fourier frequency in the noise analysis.

The full set of the equations with the noise sources consists of two equations describing the asynchronous carriers, two equations for the frequency shifts $\dot{\phi}_{fk}$ due to noise and the relationship for the bias circuit

$$\begin{aligned} T_k\dot{X}_k &= (G_kR_{kn} - 1)X_k + I_{\|k}; \\ \dot{\phi}_{fk} &= I_{\perp k}/I_{1k}T_k = I_{\perp k}/G_k(X_k)X_kT_k; \qquad (6) \\ T_b\dot{X}_0 &= X_{00} - X_0 - R_{bn}I_0 + X_{bn}; \quad (k = 1,2) \end{aligned}$$

where $X_k = C_n \times U_k$, $X_{0(0)} = C_n \times U_{0(0)}$ is the normalized values of partial oscillations and bias, X_{00} is the constant part of the resulting bias voltage X_0, $G_k = I_{1k}/X_k$ is the normalized averaged partial transconductances, $I_{1k}(X_1, X_2, X_0)$ is the amplitudes of partial fundamental rf currents, and $k = 1, 2$.

The operating frequencies ω_k have no regular shifts regarding ω_{0k} ($\dot{\phi}_k = \omega_k - \omega_{0k} = 0$) since both partial currents $\mathbf{I}_{1k}$ are in-phase with the voltages $\mathbf{U}_k$.

In Eq. (6) $R_{bn} = C_n \times R_b$ and T_b arise from the expression for bias impedance $\mathbf{Z}_b(\mathbf{p}) = R_b/(1+\mathbf{p}T_b)$ which describes the inertial properties of the bias network.

Noise terms in (6) are represented by slowly varying noise currents in vicinity of the operation frequencies

$$\mathbf{I}_{nk}(\omega) = (I_{\|k} + jI_{\perp k}) \times \exp(j\phi_k), \qquad (7)$$

where X_{bn} is the averaged low frequency noise voltage arising from bias self-noise e_{bn} and AD noise current.

It follows from (6) that in the steady-state regime

$$G_{1(2)}(X_1, X_2, X_0)/G_M = G_{1(2)n} = 1/FR_{1(2)}, \qquad (8)$$

where $FR_k = C_n \times G_M \times R_k$ are the partial regeneration factors representing maximal small signal gain of the partial positive feedback loops, $G_M = 2A/\pi$ is the maximal small signal value of G_k which is reached when $X_0 = X_{1,2} = 0$.

The nonlinear functions $G_{1(2)n}(X_1, X_2, X_0)$ define all the main properties of the oscillator. The analysis is based on computer simulation. We find that the oscillator with a constant bias has in general, five stationary points (Fig. 2). Points 2 and 3 describe stable

one-frequency regimes while the three points 1, 4, and 5 corresponded to two-frequency solutions. The central point 1 is stable but the biharmonic oscillation is not self-starting (so called *sharp* excitation) since the positive feedback loop

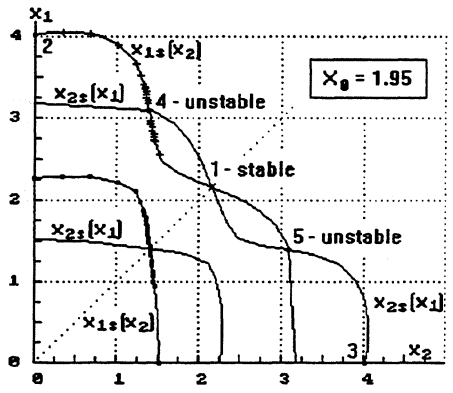


Figure 2. An example of symmetric joint steady-state solutions.

gain is less than unity for small X_1 and X_2. If exited, the two-frequency oscillation can only have small stationary amplitudes X_{1s}, X_{2s}, because the stable point disappears when FR_1 and FR_2 differ too much. If $FR_1 = FR_2 = 3$, a symmetric two-frequency regime can be observed for $|X_0| \in (1.7, 2.03)$ as in Fig. 3.

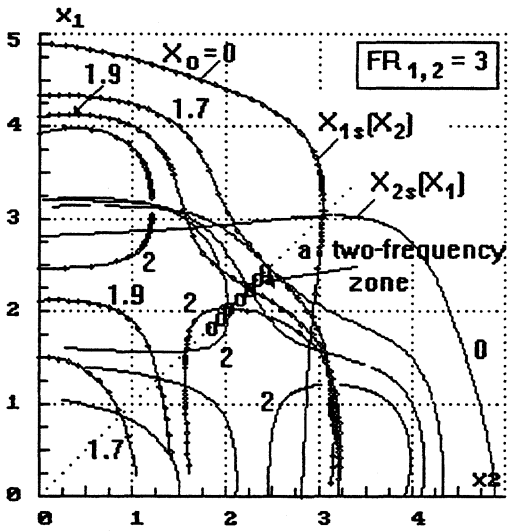


Figure 3. Evolution of the two-frequency oscillator symmetric steady-state regimes with bias change.

In an oscillator with automatic biasing, U_0 depends on a direct current I_0, which in turn, is a function of U_0 and alternating voltages amplitudes. In a steady-state regime the normalized bias is equal to

$$X_0 = X_{00} - R_{bn} \times I_0 (X_1, X_2, X_0). \qquad (9)$$

Fig. 4 illustrates a gradual transition from an unstable to stable joint solution with R_{bn} growth. A symmetric solution is unstable when $R_{bn} = 0$, still unstable for $R_{bn} \le 2.8$, and becomes stable if $2.8 < R_{bn} < 4.46$. The range of mild excitation is narrower and takes place within $2.8 < R_{bn} < 3.6$. If $R_{bn} > 3.6$, in vicinity of the origin there is a zone where regeneration is not sufficient to sustain oscillations. The result is that the excitation is getting sharp. For $R_{bn} > 4.46$ the joint solution completely vanishes.

For the chosen X_{00}, R_{bn}, and FR_1 the joint solution exists in some range of FR_2. If $FR_1 = 3$ and $R_{bn} = 10$, the stable simultaneous asynchronous oscillations exist for $FR_2 \in (2.5, 3.36)$. For even larger FR_2 there is only oscillation at frequency ν_2. If $FR_2 < 2.5$, the only oscillation is at frequency ν_1.

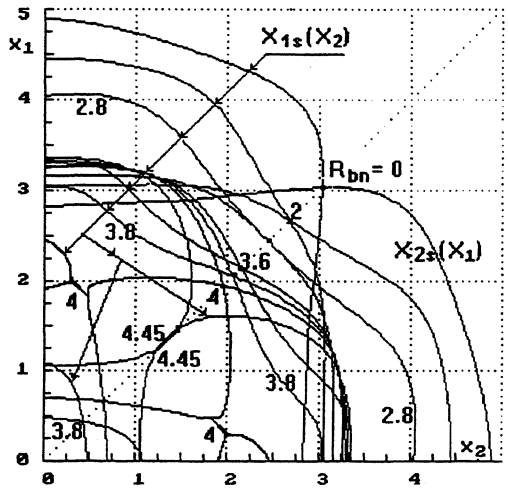


Figure 4. Symmetric steady-state regimes in the case of $FR_1 = FR_2 = 3$, $X_{00} = 0$, and different bias resistors.

3. NOISE CHARACTERISTICS

The noise components $U_f(t)$ and $\phi_f(t)$ arise in the output signal as the oscillator reaction to primary noise sources (Fig. 1). For simplicity consider the case of fundamental δ-correlated noise. Taking each particular variable X_k as the sum of a steady value X_{sk}, and a noise variation X_{fk}, using the SSE method, we obtain a system of linear equations for the first approximation to the disturbances

$$\begin{pmatrix} a_{11} - pT_1 & a_{12} & a_{13} \\ a_{21} & a_{22} - pT_1 & a_{23} \\ a_{31} & a_{32} & a_{33} - pT_b \end{pmatrix} \times \begin{pmatrix} X_{f1} \\ X_{f2} \\ X_{f0} \end{pmatrix} = \begin{pmatrix} -I_{\|1} \\ -I_{\|2} \\ -X_{bn} \end{pmatrix}. \qquad (10)$$

The elements a_{ij} of the matrix (**A**) are first-order partial derivatives of right hand side regular parts of equations

(6) with respect to the variables X_1, X_2, X_0 evaluated at the stationary point. Thus, $a_{33} = -(1 + R_{bn}\frac{\partial I_0}{\partial X_0})$.

Oscillator noise depends on the specific form of VAC. The graphs in Fig. 5-10, obtained for a mono-frequency oscillator with a constant bias, allow us to check our «arctg»-type VAC in this sense.

From (10) it follows that the power spectral density of AM noise takes a form

$$S_a(f) = \frac{S_{\parallel}(f)}{(G_1 - \sigma_1)^2 + (2\pi T G_1 f)^2},$$

where $\sigma_1(U) = \partial I_1/\partial U$ is the local first harmonic AD transconductance unlike the averaged one $G_1 = I_1/U$. Random amplitude perturbations decay with a time constant $T_a = T(1 - \sigma/G_1)$ which is regime dependent (Fig. 5). For regime stability $T_a > 0$, that is, $G_1 > \sigma_1$. Under this condition, σ_1 can have any sign and value.

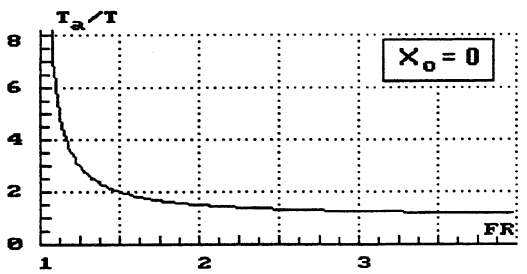


Figure 5. Dependence of the time constant for amplitude noise on a regeneration factor.

The plots of AM noise mean square value $\overline{U_f^2}(U) = \frac{\pi}{2TG_1(G_1 - \sigma)} S_{\parallel}$ and the fractional AM noise $\overline{U_f^2} / U_{ss}^2$ in Fig. 6 demonstrate quite different behavior with respect to FR.

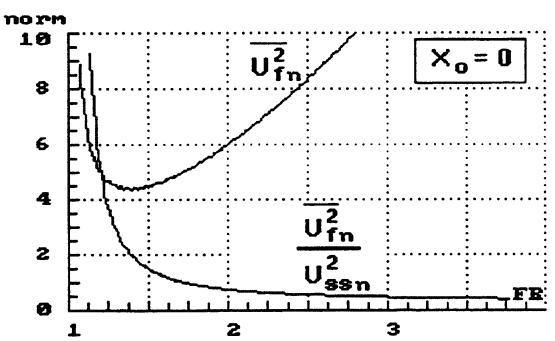


Figure 6. Mean squared AM noise and fractional AM noise vs. regeneration factor.

PM noise is described by $\pounds(f) = S_{\perp}(f) / 2G_1^2 U_{ss}^2 \xi^2$. Here $S_{\perp}(f)$ is the power spectral density of the quadratic noise current component $I_{\perp}$, $\xi = \Omega T = 2\Omega Q/\omega_0 = 2fQ/\nu_0$ is the extended tank circuit detuning at the Fourier frequency f. This formula gives the same results as Leeson's formula for the amplifier input [13].

To the first approximation, $S_{\perp} = S_{\parallel} \propto I_0$. The dependence of PM noise on the regime calculated on this basis is shown in Fig. 7. The larger FR the lower the PM noise.

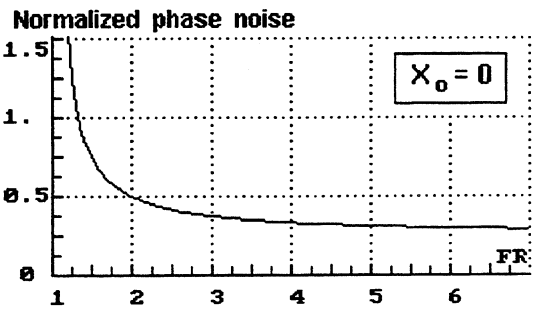


Figure 7. Influence of the regeneration factor on oscillator output PM noise spectral density.

So far our oscillator noise model is based on the assumption of a linear time-invariant system. In reality, any oscillator is a periodically time-varying system and its time-varying nature must be taken into account to permit accurate modeling of noise [14,15]. Periodic nonstationarity of the noise process leads to unequal correlation functions, and so power spectral densities of the in-phase and orthogonal noise current components. In general, there is some mutual correlation. The values of possible AM and PM noise changes due to noise cyclostationarity are illustrated in Fig.8 where we defined [15]

$$\Delta S_a = (F_0 + F_{2c}) / F_0, \quad \Delta\pounds = (F_0 - F_{2c}) / F_0. \quad (11)$$

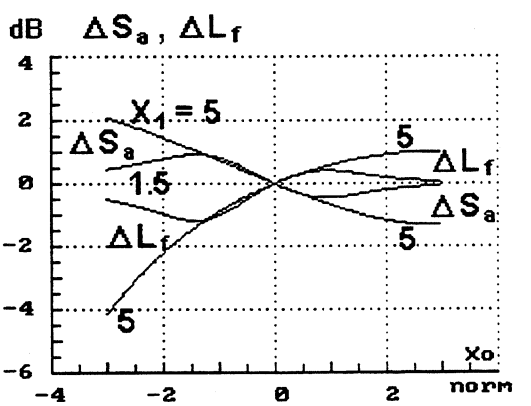


Figure 8. Influence of the regeneration factor on oscillator output phase noise spectral density.

The variations in Fig. 8 do not exceed ±2 to 4 dB. This means, noise cyclostationarity can be neglected when we calculate oscillator noise for the chosen VAC form. In the biharmonic regime the second oscillation influences the average current thus changing the initial noise. However, the effects stipulated by noise cyclostationarity remain the same order of magnitude as in a mono-frequency case.

Consider now PM noise in the symmetric two-frequency regime with $FR_1 = FR_2 = FR$. This leads to $X_{s1(2)} = X_{(s)}$, $I_{11(2)} = I_1$. The information on PM noise regime dependence is represented in Fig. 9 where we drew in logarithmic scale the direct to first harmonic current ratio vs. FR. This ratio reproduces PM noise behavior since $I_{\perp i} \propto I_0$. The solid line characterizes the two-frequency regime, and the marked one "-" the mono- frequency regime.

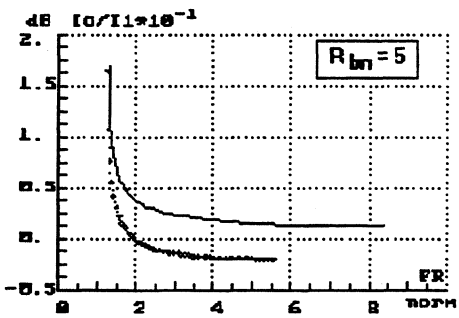


Figure 9. PM noise dependence on the regeneration factor for an oscillator with auto-biasing.

It follows that the symmetric two-frequency regime has 3.to.4 dB higher PM noise than the mono-frequency one. This is because the second oscillation suppresses the rf current more than the direct one.

AM noise in the two-frequency regime, as well in a mono-frequency one, is influenced strongly with bias circuitry inertiality described by the T_b/T_1, T_b/T_2 ratios. According to Eq. (10), AM noise at each frequency depends on the two partial in-phase rf noise components and the noise associated with the bias. Thus, AM noise at the two frequencies is partially correlated.

4. CONCLUSION

Our investigation confirmed that the chosen symmetric VAC allows us to reproduce all the basic effects known earlier for two-frequency asynchronous oscillations. Our results reveal the features of steady-state regimes and noise in such oscillators. Unfortunately, our analysis also predicts that the PM noise is higher in a two-frequency regime than in the mono-frequency regime when the oscillator uses a single AD operating in a symmetric mode.

5. REFERENCES

[1] B. van der Pol, "An Oscillation Hysteresis in a Triode Generator with Two Degrees of Freedom," Phil. Magazine, 1922, vol. 43, ser. 6, April, pp. 700-719.

[2] V. Ya. Bargin, A. A. Zelensky, F. F. Kolpakov et al., "A Multi-Mode Quartz Resonator-Thermosensor," Electronnaya Technika, ser. X, Radiocomponenty, 1972, n. 2, pp. 54-57 (in Russian).

[3] L. S. Marjanovsky, G. V. Vasetsky, "Quartz Oscillator," USSR Pat. # 758,472; Bul. Izobr., 1980, n. 31, p. 322 (in Russian).

[4] V. P. Bagaev, A. V. Kosykh, A. N. Lepetaev, and V. F. Samoylenko, "Dual-Mode Digitally Compensated Crystal Oscillator," Electrosvvaz, 1986, n. 3, pp. 48-51 (in Russian).

[5] S. Schodowski, "Resonator self-temperature-sensing using a dual-harmonic-mode crystal oscillator," Proc. 43rd Ann. Symp. Freq. Control, 1989, pp.2-7.

[6] A. V. Kosykh, S. A. Zavjalov, "Modulation Type Dual-Mode Oscillator Intended for Micro-Chip Realization," Proc. 1995 IEEE Int. Freq. Contr. Symp., 31 May-2 June 1995, San Francisco, pp.542-545.

[7] D. E. Pierce, Y. Kim, and J. R. Vig, "A Temperature Insensitive Quartz Microbalance," Proc. 1997 IEEE Int. Freq. Contr. Symp., 28-30 May 1997, Orlando, pp.41-48.

[8] D. P. Tsarapkin, E. P. Stroganova, "Asynchronous Oscillations in a Two-Tank-Circuit Oscillator with a Volt-Ampere Characteristic Approximated by the 7th Order Polynomial," Radiotechnika i electronika, 1981, vol. 26, n. 11, pp. 2315-2320 (in Russian).

[9] V. V. Anisimov, "On biharmonic oscillations excitation in an oscillator with two degrees of freedom," Vestnik MGU, ser. physics, astronomy, 1956, n. 1, pp. 137-146 (in Russian).

[10] M. I. Disman, W. A. Edson, "Simultaneous Asynchronous Oscillations in Class-C Oscillators." Proc. IRE, 1958, vol. 46, n. 5, pp. 895-903.

[11] J. S. Shaffner, "Simultaneous Oscillations in Oscillators," IRE Trans., 1954, v. CT-1, n.2, pp.2-8.

[12] S. I. Evtyanov, "On conformity between symbolic and "shortened" equations," Radiotechnika, 1946, vol. 1, n. 1, pp. 68-79 (in Russian).

[13] D. B. Leeson, "A Simple Model of Feedback Oscillator Noise Spectrum," Proc. IEEE, 1966, vol. 54, n. 2, February, pp. 329-330.

[14] A. Hajimiri, T. L. Lee, "A General Theory of Phase Noise in Electrical Oscillators," IEEE Journal of Solid-State Circuits, 1998, vol. 33, n. 2, pp. 179-194.

[15] A. N. Bruevich, "Fluctuations in Oscillators at Periodically Nonstationary Random Noise," Radiotechnika, 1968, vol. 23, n. 5, pp. 35-42 (in Russian).

EXPERIMENTAL STUDIES OF NOISE IN A DUAL MODE OSCILLATOR

H. Ascarrunz*, F.L. Walls+, E.S. Ferre-Pikal#, D. Tsarapkin•, and J. Vig×
*SpectraDynamics, Inc., Louisville, CO; +NIST Boulder, CO; #University of Wyoming, Laramie, WY; •MPEI (TU), Moscow, Russia; ×US Army CECOM, Ft. Monmouth, NJ

ABSTRACT

We present preliminary results of an experimental study of phase modulation and amplitude modulation noise in a specific type of dual mode oven controlled crystal oscillator. Our experiments indicate some correlation between the amplitude modulated and phase modulated noise in two anharmonically resonant oscillations. No correlation was found between the PM noise in both oscillations.

INTRODUCTION

The purpose of this paper is to discuss preliminary results of an ongoing investigation of amplitude modulation (AM) and phase modulation (PM) noise in dual mode oven controlled oscillators (DMOCXO). A DMOCXO has two anharmonic resonant modes. Since both modes share the same environmental effects, DMOCXOs have a wide range of applications in precision measurement and temperature sensing applications. [1] Of particular interest is the potential to greatly reduce long term frequency drift attributed to oven aging and frequency changes due to thermal transients. The problem is that the effects of an additional mode on the noise performance of the mode of interest are not well documented. Our objective is to characterize the correlation between the AM and PM noise in each resonant mode of a DMOCXO and document the effects of dual mode (DM) operation on the noise performance of the fifth overtone (5OT) in 100 MHz SC cut quartz resonators of medium electrode size (3.05 mm). We have approximately 40 resonators made with different polishing and masking processes and small medium and large electrode areas that have been previously characterized for PM noise at the 5OT in single mode (SM) operation. Since the medium electrode resonators with no masking more closely fit the standard design of current resonators, we chose two resonators with different SM operation PM noise performance, for our measurements. R1 had the median SM PM noise performance of its batch, and R14 had the optimum SM PM noise performance of the same batch. [2]

MEASUREMENT EQUIPMENT

The DMOCXO used in our experiments, consists of a dual mode test bed (DMTB) and removable DM resonators. The fundamental mode (f_1), is centered usually about 20.002 500 MHz ±1 kHz, and the 5OT is usually centered around 99.999 850 MHz ± 100 Hz. The DMTB consists of separate excitation and separate output stages for the two resonant modes, with only the oven, the input and output diplexers, and the resonator, in common. Tested low noise isolation amplifiers were added at both outputs of the DM oscillator to prevent feedback and provide additional channels.

The AM noise was measured using a conventional diode detector AM measurement system. A three cornered hat cross correlation PM measurement system was used to measure the single sideband PM noise in the 5OT. [3,4]

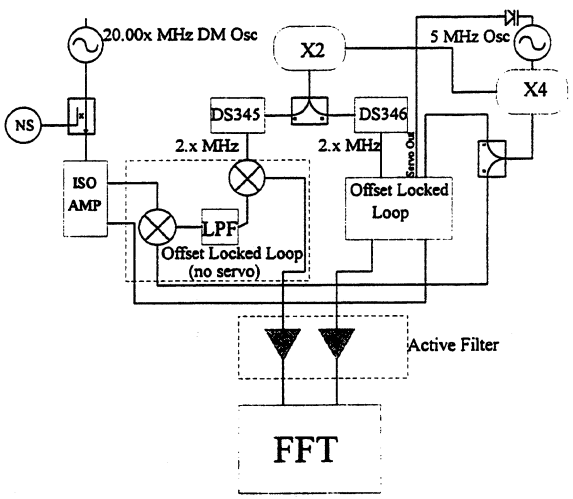


Figure 1. Modified 20.00x MHz PM measurement system

The single-sideband PM noise in the fundamental (f_1) was measured using a modified PM measurement system shown in Figure 1. In this approach, the fundamental is locked to a 5 MHz low noise OCXO using an offset-locked loop. A multiplier chain provides a 20 MHz signal that is mixed with the fundamental. The output is low pass filtered and mixed with a signal from a direct digital synthesizer (DDS). The mixer output drives the phase locked loop and an IF amplifier to a signal analyzer. To ensure phase coherence and low noise, the DDS was locked to the low noise 5 MHz source. A (1/1000) divider at the output of the DDS reduced the digital noise. To verify that the PM noise added by our modified PM noise measurement system (at the f_1) was negligible, we constructed a second channel and verified the single channel results of the modified PM noise measurement system.

Cross correlation measurements on the PM noise of two different carrier frequencies were made using single channel combinations of the measurement systems described above. Cross correlation measurements were also made on the single channel AM and PM measurements of one frequency, i.e. the cross correlation of the 100MHz PM with the 100 MHz AM, and of both frequencies, i.e. the cross correlation of the 20 MHz AM with the 100 MHz PM. Typically the channels with the lowest noise were used in the cross correlation measurements. The gain of each channel was measured using the noise source method. [5] The cross correlation gain was calculated as the logarithmic average of the single channel gains.

MEASUREMENT RESULTS

Previous measurements on the 5OT by Ferre-Pikal et al indicated that the 1/f PM noise in some resonators varied with drive current. [2] We tried to determine if the drive current of one mode had an effect on the PM noise of the other mode by measuring the PM noise in SM and DM operation. However, turning an additional mode on caused a frequency shift and made it unclear if the PM noise degradation was attributable to changes in drive current. Figures 2, 3, and 4 are normalized SM and DM PM noise measurements of the f_1 and 5OT oscillations for Resonators 1 and 14. They are normalized by adding 30log(f) to the measured PM noise.

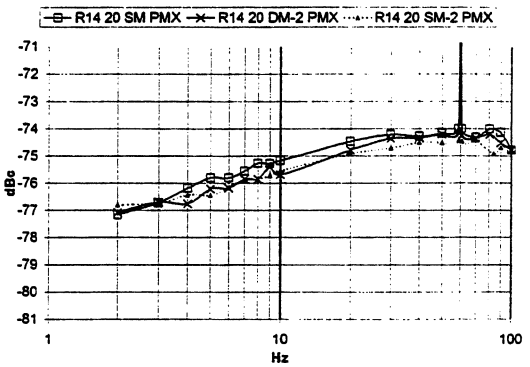


Fig. 2. Normalized 20 MHz (f_1) PM noise measurements of Resonator 14 in SM and DM operation.

The f_1 (20 MHz) PM noise shown in Fig. 2 for Res. (Resonator) 14 indicates little sensitivity of the PM noise to DM or SM operation, even though the frequency shifted by about 150 Hz. Similar results were obtained for Res. 1. The 5OT (100 MHz) PM noise appeared to be more sensitive to the presence of the f_1. The traces labeled SF and FS in Figures 3 and 4, respectively, represent the PM noise of the 5OT in DM operation with the f_1 operating at different mechanically tuned frequencies. Table 1 shows the different operating frequencies for each measurement. The PM noise performance of the 5OT seems to degrade as the f_1 is tuned in frequency even though the drive level in all our measurements remained unchanged.

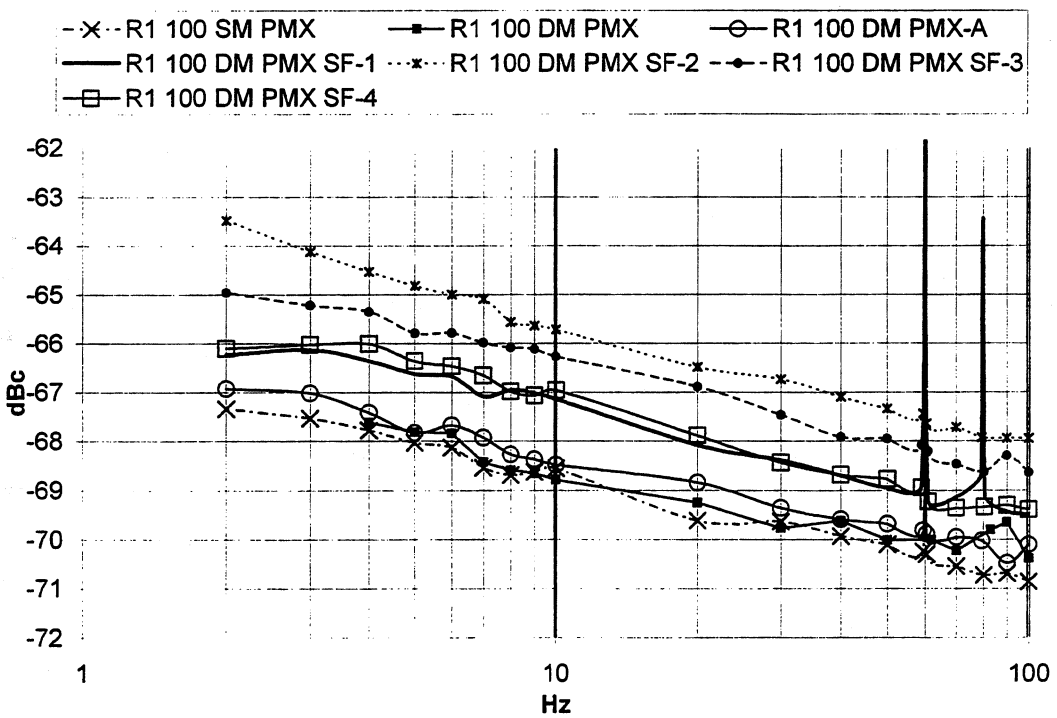


Figure 3. 100 MHz (5OT) PM noise measurements of Res. 1 in DM and SM operation.

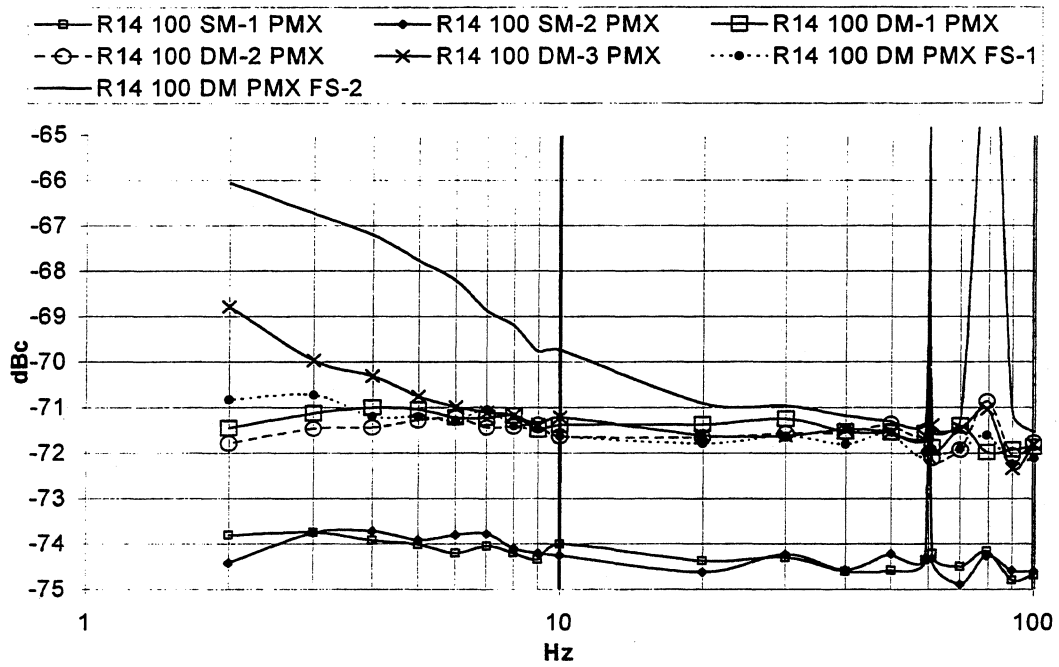


Figure 4. Normalized 100 MHz (5OT) PM noise measurement in Res. 14 for DM and SM operation

Table 1. 5OT frequency at different f_1 frequencies for Resonator 1 (SF traces in Fig. 3) and resonator14 (FS traces in Fig. 4).

Label	Frequency of f_1	Frequency of 5OT
Resonator 1		
SF-1	20.002 708 MHz	99.999 962 MHz
SF-2	20.002 043 MHz	99.999 942 MHz
SF-3	20.003 023 MHz	99.999 957 MHz
SF-4	20.001 605 MHz	99.999 914 MHz
Resonator 14		
FS-1	21.907 274 MHz	99.999 841 MHz
FS-2	20.003 165 MHz	99.999 850 MHz

The observed frequency pulling suggested some interaction between modes, however, cross correlation measurements of the 20 MHz (f_1) PM noise and the 100 MHz (5OT) PM noise, shown in Fig. 5 and 6 suggests that the PM noise in the two oscillations is independent.

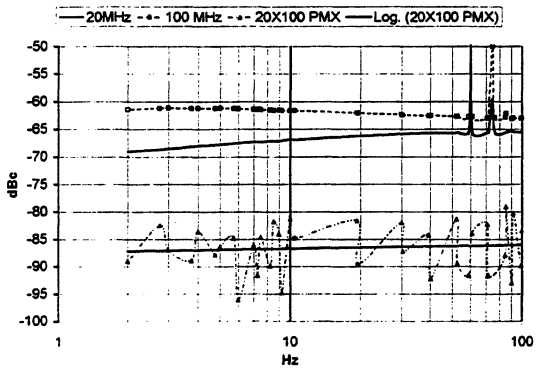


Figure 5. Normalized 20MHz and 100 MHz PM noise cross correlation measurement for Res. 1.

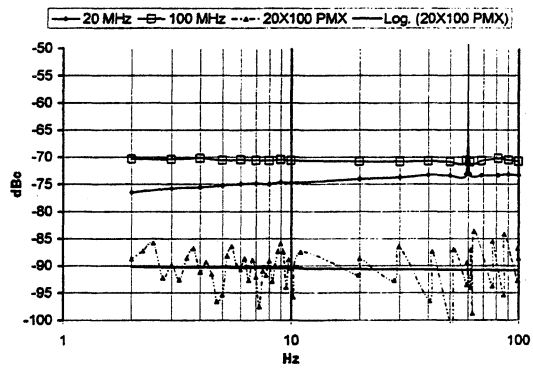


Figure 6. Normalized 20 and 100 MHz PM noise cross correlation measurement for Res. 14.

At 100 percent correlation the cross correlation trace would be a geometric average of the single channel traces. If there is no correlation between the single channel measurements, we expect the cross to average away as (5) log (N), where N is the number of averages. [3,4] At 2000 averages we would expect at least 16.5 dB of rejection, roughly the difference between our measured cross and the geometric mean of the single channel traces. In addition, we would expect that the cross correlation trace would become smoother with more averaging. In all our cross correlation measurements of the 20 MHz PM noise and 100 MHz PM noise, we observed that the cross correlation trace averaged lower but did not become smoother.

Cross correlation measurements of AM noise and PM noise were taken at both resonant frequencies (Figs. 7 and 8). Fig. 7 depicts cross correlation of the 100 MHz AM noise and 100 MHz PM noise. The 20

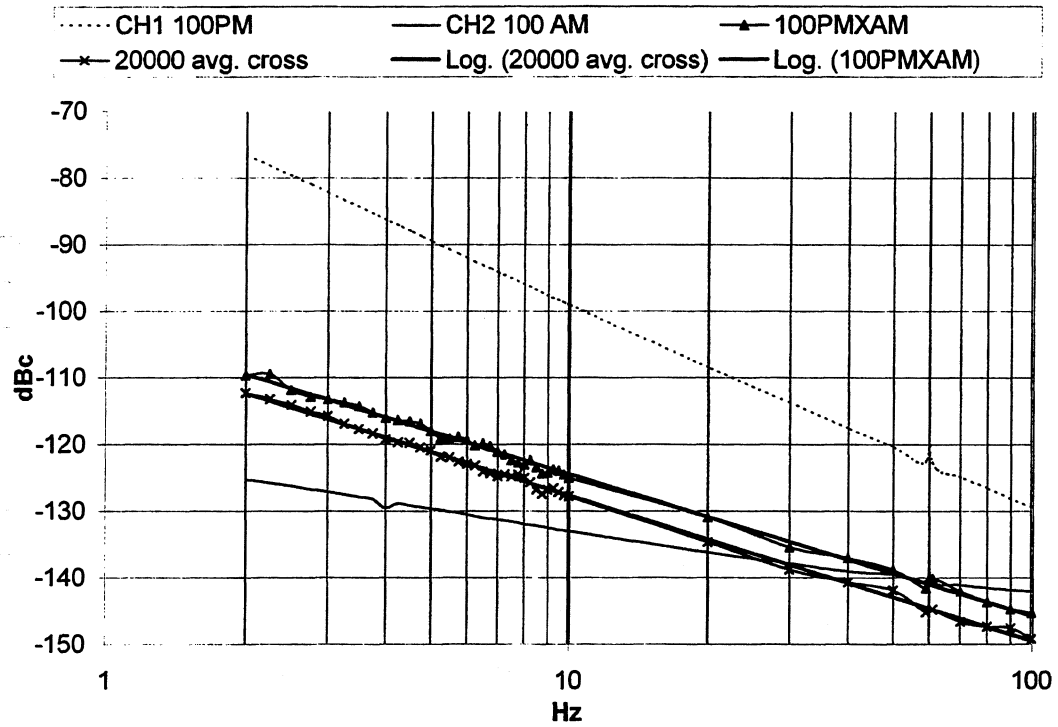


Figure 7. Res. 14 correlation between the 100 MHz PM noise and 100 MHz AM noise in DM operation

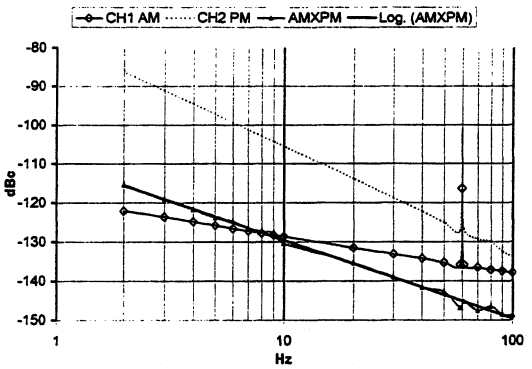


Figure 8. Cross correlation measurement of the AM noise with the PM noise at 20 MHz for Res. 14.

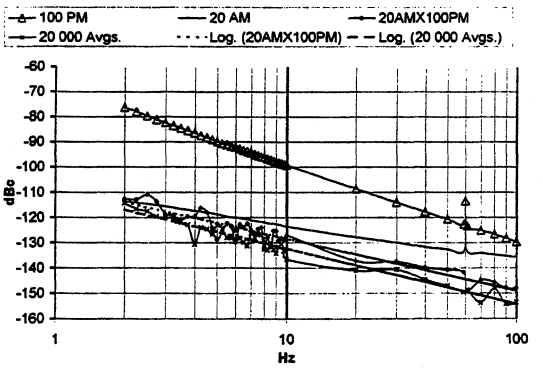


Figure 9. Correlation between the 20 MHz AM noise and 100 MHz PM noise for Res. 14.

MHz PM noise and 20 MHz AM noise cross correlation measurement is shown in Fig. 8. In both measurements the PM and AM noise seemed to be slightly correlated.

The 100PMXAM trace in Figure 7 represents the cross correlation between the AM and PM noise at 2000 averages, the trace labeled 20000 avg. cross is the same measurement at 20000 averages. If they were uncorrelated, we would expect the 20000 average trace to be about 5 dB lower than the 2000 average trace. In Fig. 7 we observe that at 20000 averages the cross correlation is close to 3 dB lower than at 2000 averages, this would indicate some correlation between the AM and PM noise of the individual frequencies. In addition, the cross correlation traces in Figs. 7 and 8 are smoother and their slopes ($\approx f^{-2}$) correspond to the slope of the geometric mean of the single channel PM noise and AM noise measurements.

Figures 9 and 10 show cross correlation measurements between the PM noise in one resonant frequency and the AM noise in the other resonant frequency. Although the cross correlation averages below the single channel traces in Figure 9, the averaging rate is roughly half as large as we would expect if there was no correlation. The f^{-2} slope observed in all the PM noise and AM noise cross correlation measurements indicate that there a small amount of correlation between the AM noise in one resonant frequency and the PM noise of the other resonant frequency.

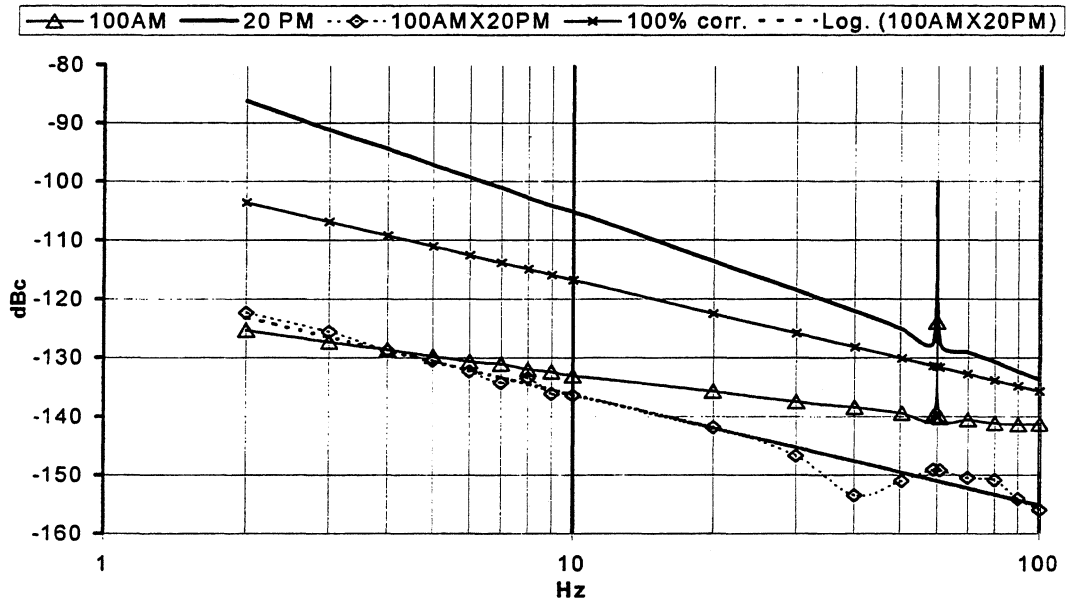


Figure 10. Correlation between the 20 MHz PM noise and 100 MHz AM noise in Res. 14.

DISCUSSION OF RESULTS

PM noise measurements in DM and SM operation indicate some interaction between the modes in our DM oscillator. It is unclear that the frequency pulling observed when tuning either mode is attributable to mode to mode interaction in the crystal. The use of mechanical jumpers to toggle the modes, and small impedance changes in the diplexers probably account for most of the observed cross mode tuning interactions. Experiments done on the 5OT in DM operation (Figure 4) indicate that the PM noise performance degrades when we move the f_1. The measured 5OT PM noise increased when the 5OT was pulled farther from its optimized frequency by the f_1. The cross mode PM noise correlation analysis in both resonators suggests no measurable frequency or phase modulation between the modes.

The PM and AM correlation analysis performed on the individual tones and cross both tones indicates some PM to AM correlation in our oscillator. If the AM noise was 100% correlated to the PM noise, you would expect the correlation curve to be the geometric mean of the both, as shown in Figure 12 by the calculated trace labeled "100% corr", and to intersect with the AM and PM noise curves at some Fourier frequency above 100 Hz. The 100AM trace in Figure 12 intersects the 100AMX20PM trace close to 3 Hz. This indicates that the component of the PM noise attributable to AM to PM conversion is at least 32 dB below the total PM noise, or, less than .07 % of the total noise. The same applies to AM noise attributable to PM to AM conversion. The 20AMX100PM at 2000 averages, intersects the 100 AM about 36 dB from the PM. In the same tone AMXPM measurements, Figs. 8, 9, and 10, we also see very small correlation.

From our measurements, we can infer that the mechanisms in our DMOCXO that produce AM and PM noise in the separate modes are largely independent. We observed that the largest contributing factor to PM noise degradation in the 5OT is the frequency "pulling" caused by tuning the f_1.

REFERENCES

[1] D.E. Pierce, Y. Kim, and J.R. Vig, "A Temperature Insensitive Quartz Microbalance," Proc. 1997 IEEE Intl. Freq. Control Symp., 1997, pp. 41-48.

[2] E.S. Ferre-Pikal, F.L. Walls, "Experimental Studies on Flicker Noise in Quartz Crystal Resonators as a Function of Electrode Volume, Drive Current, Type of Quartz, and Fabrication Process," Proc. 1996 IEEE Intl. Freq. Control Symp., 1996, pp. 844-851.

[3] W.F. Walls, "Cross-correlation Phase Noise Measurements," Proc. 46th Ann. Freq. Control Symp., 1992, pp. 257-261.

[4] F. L. Walls, "Secondary Standard for PM and AM Noise at 5, 10 and 100 MHz," IEEE Trans. Instrum. Meas., 42, 1992, pp. 136-143.

[5] F.L. Walls, A.J.D. Clements, C.M. Felton, M.A. Lombardi, and M.D. Vanek, "Extending the Range and Accuracy of Phase Noise Measurements," Proc. 42nd Ann. Symp. Freq. Control, 1988, pp. 432-441.

PHASE-NOISE MEASUREMENT IN DUAL-MODE SC-CUT CRYSTAL OSCILLATORS

Yasuaki WATANABE, Toshihito OKABAYASHI*, Shigeyoshi GOKA and Hitoshi SEKIMOTO
Graduate School of Engineering, Tokyo Metropolitan University
1-1, Minami-Osawa, Hachioji, Tokyo 192-0397, Japan

*Hitachi Electronics Services Co. Ltd.
504-2 Shinano-cho, Totsuka-ku, Yokohama 244-0801, Japan

Abstract - This paper reports the phase-noise characteristics and their analysis model of an SC-cut dual-mode oscillator. The C-mode phase-noise sideband levels of -124 dBc at 10 Hz and -154 dBc at 10 kHz have been demonstrated using a dual-mode oscillator with a 10 MHz 3rd overtone SC-cut crystal resonator. These levels are not especially low but they are only 3 to 6 dB higher than the corresponding levels of the signal obtained by the single-C-mode oscillation using one active circuit. It is therefore evident that the proposed dual-mode oscillator can be used as a reference oscillator in communication systems. Moreover, the phase-noise levels of the C-mode in the dual-mode oscillation correspond well to those obtained at stopping the B-mode oscillation. These results demonstrate that the B-mode oscillation does not disturb the phase-noise sideband levels of the C-mode signal.

I. INTRODUCTION

Quartz oscillators are widely used as stable frequency sources in communication systems. Recent advancements in mobile communications technology requires that quartz oscillators be highly stable and have good spectrum purity. We have already developed an SC-cut quartz oscillator that uses an AT-cut quartz resonator as a gain-band filter, and applied it to the ultra-stable OCXOs [1,2]. Because this circuit has a very narrow gain-band width, we also applied this circuit to a dual mode oscillator [3] that is used in a temperature compensated quartz crystal microbalance system [4].

For the purpose of applying this dual mode oscillator to the communication field, this paper describes that the measurements and calculations of the phase noise characteristics of this oscillator.

The configurations of a dual-mode oscillator that uses an SC-cut resonator is described in Section II, and the phase-noise measurement method is described in Section III. The phase-noise analysis in dual-mode oscillation based on the Leeson's model [5,6] is discussed in Section IV. From the experimental results and analyses, it is shown that the proposed dual-mode oscillator can be used as a reference oscillator in communication systems and the proposed analysis model can predict the phase-noise characteristics in the dual-mode.

II. DUAL-MODE OSCILLATOR USING NARROW-BAND OSCILLATORS

Fig. 1 shows the dual-mode oscillator that we presented at the 1997 FCS [7]. This circuit employs two narrow-band active circuits that have different negative-resistance bands which are connected in parallel at a common resonator terminal. We demonstrated that this oscillator could excite any modes in a common resonator even if their resonant frequencies are very close to each other. Fig. 2(a) shows an example of a frequency spectrum of the dual-mode oscillator corresponding to the resonator frequency response shown in Fig. 2(b).

We rearranged this circuit in order to excite the B and C modes in the 10 MHz 3rd overtone SC-cut resonator. Table 1 shows the electrical equivalent parameters of the SC-cut resonator. In a practical circuit the buffer amplifiers are used to obtain the signal strength for measuring their phase noise characteristics. A relay has been also employed to detach the oscillation circuits in order to compare the phase noise characteristics in the C-mode oscillation and the dual-mode oscillation.

Fig. 3 shows the frequency response of the SC-cut resonator for dual-mode oscillation. The oscillator shown in Fig. 1 was tuned to excite the C and B modes.

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

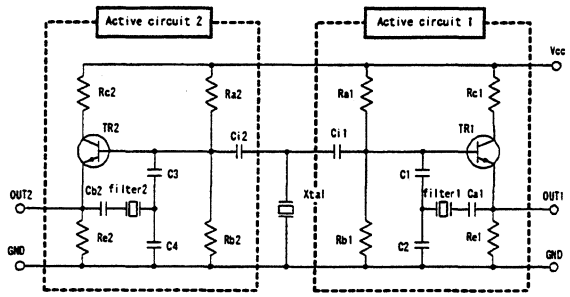


Fig.1 A dual-mode Colpitts oscillator.

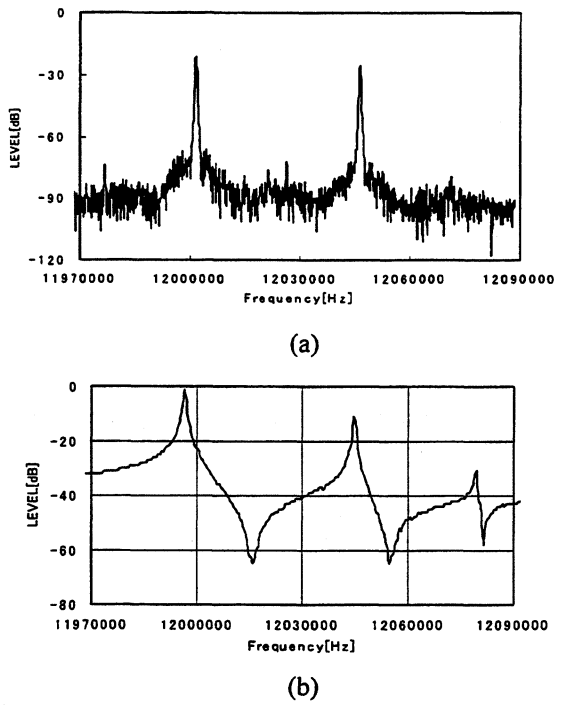


Fig.2 (a) Frequency spectrum of the dual-mode oscillator, (b) frequency response of the AT-cut resonator.

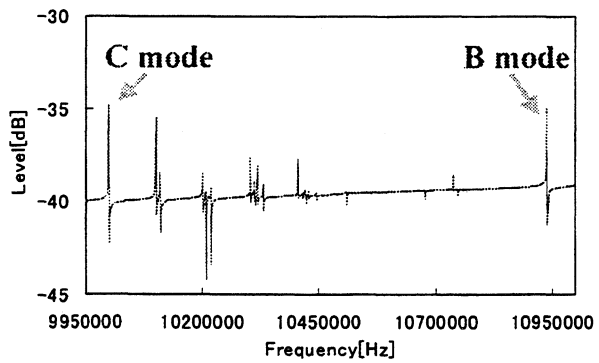


Fig.3 Frequency response of the 10 MHz SC-cut crystal resonator.

Table 1. Equivalent parameters of the SC-cut resonator.

Parameters	C mode	B mode
$R_1[\Omega]$	56.36	36.81
$L_1[mH]$	1170.91	1239.98
$C_1[fF]$	0.22	0.17
$C_o[pF]$	5.8	
$f_s[Hz]$	9999740.74	10937669.15
Q	1350467	2315140

III. PHASE NOISE MEASUREMENT SYSTEM

Fig. 4 shows the block diagram of the phase noise measurement system. The oven was kept at 70 °C. A voltage controlled SC-cut OCXO that had the similar phase noise characteristics to the dual-mode oscillator was used as the reference oscillator. To reject the B-mode signal that influences the phase-locked-loop condition, a ceramic resonator filter was inserted between the oscillator output and the phase discriminator.

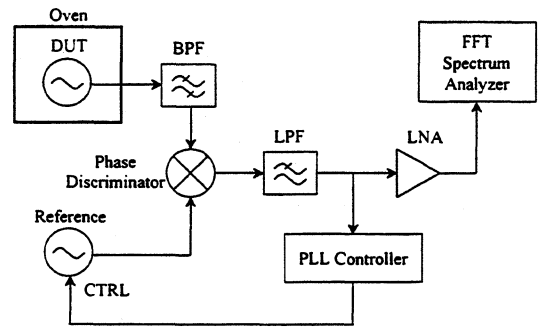


Fig.4 Block diagram of the measurement system.

IV. PHASE NOISE ANALYSIS

Fig.5 illustrates Leeson's model of phase noise for a closed-loop oscillator circuit. Leeson's model has some application limitations for negative resistance oscillators, such as the Colpitts oscillator. One significant limitation is that it does not take into account that the noise observed at the amplifier output is affected by the impedance-frequency characteristics of the resonator and other filter elements. However, this model is easy to understand and can be very useful in identifying sources and characteristics of oscillator noise. In this regard we use this model to identify the phase noise in dual-mode oscillators in the areas where the impedance characteristics of resonator and filter devices are negligible.

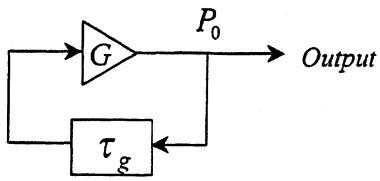


Fig.5 Leeson's model of phase noise for the closed-loop oscillator circuit.

The approximate phase noise of an oscillator, based on the Leeson's model, can be expressed in equation form as shown below,

$$S_\phi(f)=\alpha\cdot(\upsilon_0/2Q_L)^2/f^3+\beta\cdot(\upsilon_0/2Q_L)^2/f^2+\alpha/f+\beta \quad \ldots(1)$$

where α is *1/f* noise factor in the open loop circuit, ν_0 is nominal frequency, Q_L is a resonator load Q, and β is noise floor level.

The coefficient β is obtained by,

$$\beta = 2GFKT/P_0 \quad (2)$$

where G is the compressed amplifier gain, F is the noise factor of the amplifier, K is Boltzmann's constant, T is the absolute temperature, and Po is the carrier power at the amplifier output.

Assuming that the B-mode oscillation does not directly influence the C-mode spectrum, the active circuit that drives the B mode behaves as a load Q depressor and a white PM noise source. If the same bias parameters are used in each circuit, the load Q for the C mode will be half as much as a single-mode operation and the floor noise will be double.

Following those assumptions, the phase noise spectrum $S_{\phi d}(f)$ of the C mode signal during dual-mode oscillation is expressed in next equation,

$$S_{\phi d}(f)=4\alpha\cdot(\upsilon_0/2Q_L)^2/f^3+8\beta\cdot(\upsilon_0/2Q_L)^2/f^2+\alpha/f+2\beta \quad \ldots(3)$$

It is understood from Eq. 3 that the flicker FM noise in the $1/f^3$ term is four times (6 dB) as large as a single-mode oscillation, the white FM noise in the $1/f^2$ term is eight times as large and the white PM noise in the β term becomes twice as large as a single mode oscillation.

In order to calculate the coefficients that are required in Eq.3 we measured the phase noise for the single-mode oscillator and then fit Eq. 1 to the measured data, since the parameters such as loaded Q, amplifier gain G, or noise factor F are not easy to quantify. We estimated the spectral purity for the dual-mode oscillation by substituting the coefficients into Eq.3.

V. RESULTS

The experiments were carried out in the following three patterns [a]~[c] and the results were compared with the analyses described in Section IV.

[a] Single-mode oscillation: only the C-mode was excited, that is, the circuit that excites the B-mode was detached from the circuit. The results obtained indicate intrinsic phase noise in the oscillator.
[b] Dual-mode oscillation: B and C modes were simultaneously excited.
[c] Single-mode oscillation: only the C-mode was excited but the circuit that excites the B mode remained connected to the oscillation circuit.

Fig.6 shows the results for pattern [a]. The line shows the measurements and the circles are the fitted values based on Eq. 1. The results indicate that this oscillator has a high-stable profile that is characterized by $1/f^3 + 1/f + 1/f^0$ spectrum. The phase-noise sideband levels were -130 dBc at 10 Hz and -157 dBc at 10 kHz. Table 2 shows the coefficients and determinable constants for Eq. 1. By putting those values into Eq. 3, the dual-mode oscillation phase noise was estimated.

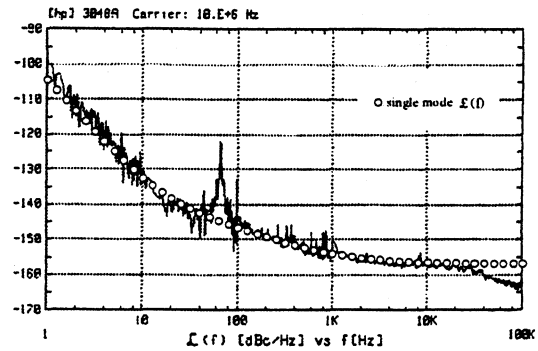


Fig.6 Phase noise of C-mode signal: the oscillator excites only C mode and the circuit that excites the B mode is detached from the oscillator.

Table 2. Power low coefficients for the single mode oscillation

Coefficients	α	β	ν_0	Q_L
Value	1.84E-13	1.98E-16	1.00E+7	3.63E+5

Modes	$1/f^3$	$1/f$	$1/f^0$
Coefficients	3.50E-11	1.84E-13	1.98E-16

Fig.7 shows the results for [b]. The calculated data is indicated with circles. The phase-noise sideband levels were -124 dBc at 10 Hz and -154 dBc at 10 kHz. By comparing Fig. 6 and 7 we observe that the phase noise in dual-mode oscillation is 6 dB higher in the flicker ($1/f^3$) region than the noise in the single-mode oscillation and that the phase noise floor is 3 dB higher than that of the single-mode oscillation. The calculated results corresponded well to the measurements, that is, Eq.3 can predict the phase noise in those regions.

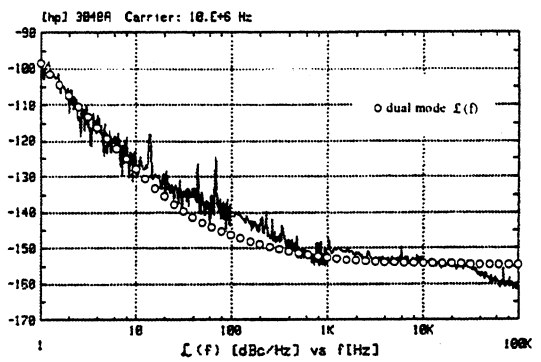


Fig.7 Phase noise of C-mode signal: the oscillator excites C and B mode simultaneously.

Fig. 8 shows the results for [c]. The phase-noise sideband levels corresponded well to Fig. 7. This agreement indicates that the C-mode oscillation signal is not directly affected by the B-mode signal, and that the only additional-circuit-elements increase the phase noise levels.

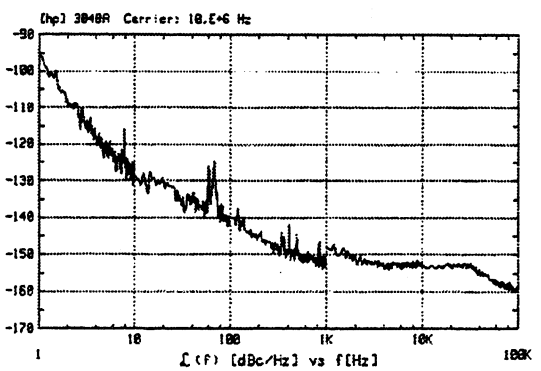


Fig.8 Phase noise of C-mode signal: the oscillator excites only C mode and the circuit that excites the B mode is connected to the main oscillator circuit.

Unexpectedly, the experimental noise levels in the regions of 20 Hz to 500 Hz were increased by 3 to 6 dB and disagree with the calculations in Figs. 7 and 8. To account for this discrepancy we assume that Leeson's model can not represent the $1/f^2$ noise, resulted from impedance characteristics of crystal filters, and that the $1/f^2$ noise is latent in the other noise components at the single-mode oscillation. This noise is emphasized by the dual-mode oscillation since the $1/f^2$ noise in dual-mode is eight times (+9 dB) as large as that in single mode operation.

VI. CONCLUSIONS

C-mode phase-noise sideband levels of -124 dBc at 10 Hz and -154 dBc at 10 kHz were achieved using a dual-mode oscillator with a 10 MHz 3rd overtone SC-cut crystal resonator. In these regions the analyzed sideband levels corresponded well to the experiments. These levels are not especially low, but if the intrinsic noise levels are improved then the proposed dual-mode oscillator can be used as a reference oscillator in communication systems. We will report on the application of the dual- mode signals and the noise level increment at the $1/f^2$ region in the future.

References

[1] Y.Watanabe, H.Sekimoto and Y.Oomura "A Low Phase-Noise Oscillator Design for High Stability OCXOS" ,Proc. IEEE Freq.Cont.Symp., pp.749-751,June.1996

[2] T.Uchida, K.Kasahara, Y.Watanabe and H.Sekimot "Long-Term Stability of High-Precision OCXO with Quartz Resonator filters" ,IEE Japan ECT-98-76,1998(in Japanese)

[3] I.Niimi, Y.Watanabe, H.Sekimoto and S.Goka "Active-Impedance Analysis of Narrow-Band Crystal Oscillators with Resonator Filters and Its Application to Dual-Mode Crystal Oscillators" IEICE Trans,vol.E81-C,no.2, pp.284-289, Feb.1998

[4] S.Goka, Y.Watanabe and H.sekimoto "Stress Sensitivity Distribution Measurement Using a Low-Frequency Sound Wave and Its Application to Quartz Crystal Microbalances" ,Jpn. J. Appl. Phys. Vol.36(1997) pp.3178-3179 Part 1,no.5b,May 1997

[5] D.B.Leeson, "A Simple Model of Feedback Oscillator Noise Spectrum" ,Proc. of the IEEE, vol.54, no.2,pp.329-330,Feb.1966

[6] Thomas E.Parker and Gary K.Montress "Spectral Purity of Acoustic Resonator Oscillators" ,Proc.IEEE Freq.Cont.Symp., pp.340-348,May.1992

[7] Y.Watanabe, H.Sekimoto, S.Goka and I.Niimi "A Dual Mode Oscillator Based on Narrow-Band Crystal Oscillators with Resonator Filters" ,Proc.IEEE Freq.Cont.Symp., pp.932-937,May.1997

Flicker Noise Measurement of HF Quartz Resonators

J. Groslambert*, V. Giordano*, M. Brunet◇, E. Rubiola△

* LPMO CNRS, 32 av. de l'Observatoire 25044 Besançon, France
◇ CNES, 18 av. Edouard Belin, Toulouse, France
△ Politecnico di Torino, Dip. Elettronica c.so Duca degli Abruzzi no. 24, 10129 Torino, Italy
e-mail: rubiola@polito.it or giordano@lpmo.univ-fcomte.fr

Abstract

The frequency flicker of quartz resonators can be derived from the measurement of $S_\varphi(f)$, i.e. the power spectrum density of phase fluctuations φ. Besides, the interferometric method turns out to be the best choice to measure the quartz resonator phase fluctuations because it high sensitivity even in the low power conditions required for these devices. Combining these two ideas, we built an instrument suitable to measure the resonator frequency flicker floor and we measured the stability of some 10 MHz high performances resonators as a function of the dissipated power. The stability limit of our instrument, described in terms of Allan deviation $\sigma_y(\tau)$, is of some 10^{-14}.

1 Introduction

The frequency stability of commercially available state-of-the-art quartz oscillators, specified in terms of Allan deviation $\sigma_y(\tau)$, can be as good as 7×10^{-14} for the measurement time τ in the 3–100 s range. In that τ range, where $\sigma_y(\tau)$ is nearly constant, frequency flicker is the most relevant noise process. Whether this floor is due to the frequency fluctuation of the quartz resonator, or it comes from the phase flicker of the amplifier converted into frequency flicker by the Leeson effect, is still a matter of discussion. To answer this question a noise measurement system is needed, one that is suitable to low power, 100 μW or less, and that exhibits a stability $\sigma_{y0}(\tau)$ of some 10^{-14}, which is out of reach for current instrumentation. A method is being proposed, based on frequency domain measurement of phase fluctuations, that shows improved sensitivity. A measurement system has been implemented and successfully used to measure a few high stability 10 MHz quartz resonators. The flicker floor of the described prototype is close to the 10^{-14} target, depending on the quartz driving power.

This work has partially been supported by the Centre National d'Etudes Spatiales, the French space agency.

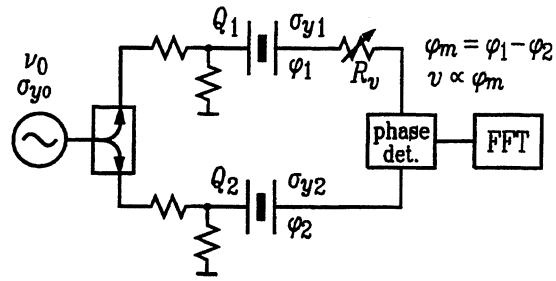


Figure 1: Bridge quartz measurement scheme.

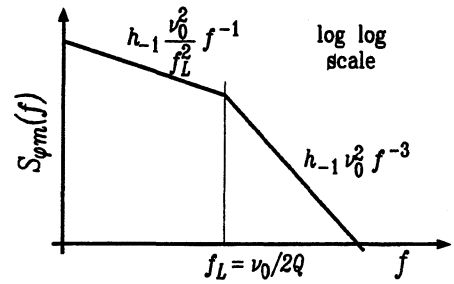


Figure 2: Expected $S_{\varphi m}(f)$. Frequency flicker only is considered, characterized by $S_y(f) = h_{-1}f^{-1}$.

2 Basics of the proposed method

With reference to Fig. 1, let us assume that two equal quartz resonators are inserted in a bridge [1, 2] driven by a noiseless oscillator of frequency ν_0, tuned at their exact resonance frequency. Provisionally, we also assume that one resonator flickers, while the other one is perfectly stable. In this condition, phase shift φ_m fluctuates because the resonance frequency of the quartz flickers. Hence, the ultimate frequency flicker floor of an oscillator based on that quartz can be derived from measurement of the power spectrum density $S_{\varphi m}(f)$.

Regarding the relative frequency shift $y_q = \frac{\nu-\nu_0}{\nu_0}$ as the input signal and φ_m as the output, the quartz is equivalent to a low pass filter characterised by the Leeson cutoff frequency $f_L = \frac{\nu_0}{2Q}$. With the 10 MHz resonators we measured on, f_L is of the order of 5 Hz, limited by the loaded merit factor Q; for reference, the unloaded Q never exceeds 1.5×10^6 at that frequency.

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

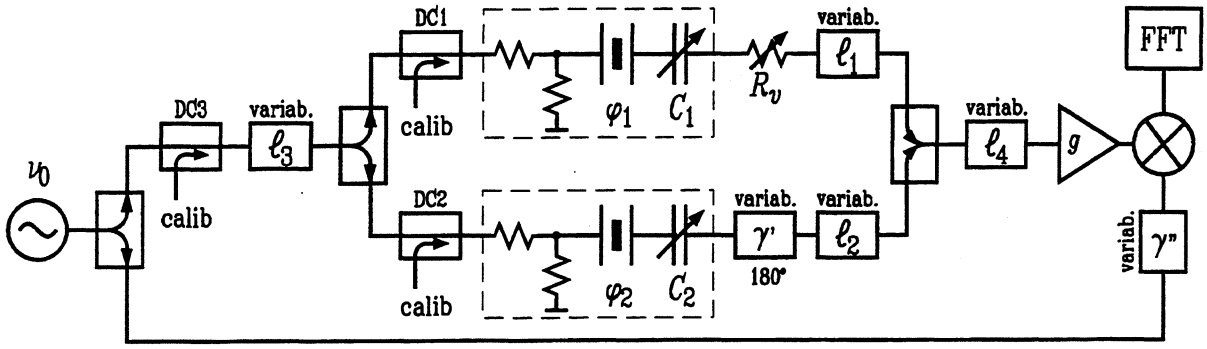


Figure 3: Interferometric measurement system.

This frequency to phase conversion, written in terms of power spectrum densities, is

$$S_{\varphi m}(f) = \frac{1/f_L^2}{1+f^2/f_L^2}\,\nu_0^2\,S_{yq}(f) \qquad (1)$$

Obviously, $S_{\varphi m}(f)$ must be taken in the frequency range where frequency flicker is the dominant process. Fig. 2 reports the $S_{\varphi m}(f)$ plot thereby expected.

For frequency flicker, usually described as $S_y(f) = h_{-1}f^{-1}$, it holds $\sigma_y^2(\tau) = 2\ln 2\; h_{-1}$. Combining the above relationships we get $S_y(f) = \sigma_y^2(\tau)/(2\ln 2\, f)$. The latter, inserted in (1) yields

$$\sigma_{yq}^2(\tau) = \frac{2\ln 2}{4Q^2}\left(1+\frac{f^2}{f_L^2}\right) f S_{\varphi m}(f)\,. \qquad (2)$$

To derive $\sigma_{yq}^2(\tau)$, we measure $S_{\varphi m}(f)$ for $f < f_L$; consequently

$$\sigma_{yq}^2(\tau) = \frac{2\ln 2}{4Q^2}\, f\, S_{\varphi m}(f)\,. \qquad (3)$$

There are two reasons to measure $S_{\varphi m}(f)$ at $f < f_L$ instead of $f > f_L$: i) higher $S_{\varphi m}(f)$ relaxes noise specification for the phase detector, and ii) slope f^{-1}, instead of f^{-3} yields lower spectrum analyser uncertainty. Nevertheless, the approximation of equation (2) for $f > f_L$, together with the corner at $f = f_L$, is a useful diagnostic tool.

Finally, letting both resonators flicker, the variance of each resonator is half of that stated by equations (2) and (3) for a given $S_{\varphi m}(f)$.

Resonators being tested also convert the oscillator frequency fluctuations $\sigma_{yo}(\tau)$ into phase fluctuations that are undistinguishable from these due to the resonators themselves. That oscillator, based on a quartz, can not be more stable than the resonators we are interested in. Yet, due to circuit symmetry, only a fraction of $\sigma_{yo}(\tau)$ is taken in. Accordingly, the instrument noise floor is

$$\sigma_{y0}^2(\tau) \simeq \frac{1}{2}\left(\frac{Q_1-Q_2}{Q}\right)^2 \sigma_{yo}^2(\tau) \qquad (4)$$

where Q_1 and Q_2 refer to the individual resonators and $Q = (Q_1+Q_2)/2$. After adjusting the circuit by means of a variable resistor that damps the higher Q resonator, a Q mismatch within 10% is easily obtained, which warrants an oscillator noise rejection of 20 dB.

Phase detection turns out to be a critical point because high sensitivity must be achieved with low power, which are mutually exclusive constraints. In fact, although details of state-of-the-art oscillators are not published, we expect that with a 10 MHz quartz the loaded Q is in the 7×10^5 to 10^6 range and the dissipated power is of some 10 μW. Taking $Q = 7\times10^5$ as a conservative value, a noise floor $\sigma_{y0}(\tau) = 10^{-14}$ implies that the instrument noise specified in terms of $S_{\varphi m0}(1\,\mathrm{Hz})$ should not exceed -155.5 dBrad^2/Hz.

A double balanced mixer used as the phase detector could offer the desired low noise, provided it is driven with sufficient power, of the order of 10 dBm, but the signal at the quartz output must be amplified. Yet, according to our experience, commercially available radiofrequency amplifiers do not meet the phase flicker requirement when they deliver some 10 dBm. For comparison, the best prototype built in our laboratory shows a phase noise $S_{\varphi m0}(1\,\mathrm{Hz}) = -140$ dBrad^2/Hz when the quartz dissipated power is $P_d = 50$ μW. Still under the assumption of $Q = 7\times10^5$ and $\nu_c = 10$ MHz, the reported noise is equivalent to $\sigma_{yq}(\tau) = 4\times10^{-14}$ for each quartz, which is suitable to most resonators but is not sufficient for our purposes. Whereas the amplifier noise can be rejected by means of a correlation scheme, in which 4 amplifiers and two mixers are used [3], we opted for quite a different solution.

3 Interferometric Measurement System

The proposed measurement scheme, shown in Fig. 3, is basically an interferometric phase detector modified for quartz resonators. This kind of detector, first proposed as a microwave instrument [4], has been ameliorated and adapted to lower frequencies [5]. In short, phase noise is regarded as a sideband pair that carries information. Hence, after adjusting the phase γ' and the attenuations ℓ_1 and ℓ_2 for best circuit symmetry, the carrier is suppressed at the input of the amplifier;

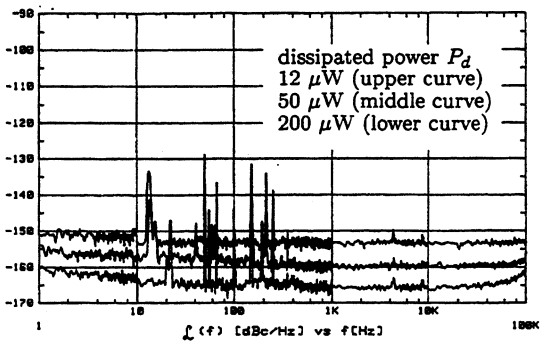


Figure 4: Phase detector noise floor.

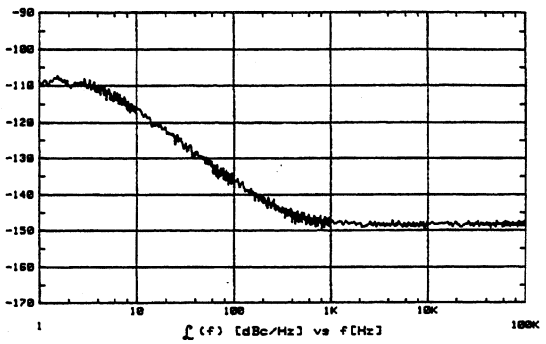


Figure 5: Transfer function $|H(f)|$ of a BVA resonator.

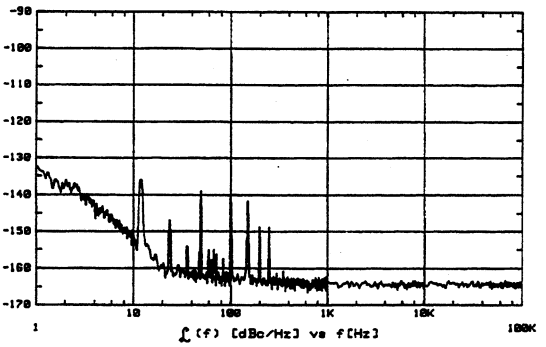


Figure 6: Phase noise of a BVA resonator pair.

the latter amplifies the noise sidebands only. Properly setting the phase shift γ'', the mixer down converts to baseband the phase noise sidebands, rejecting amplitude noise. Finally, at the fast Fourier transform (FFT) analyser input a voltage $v(t)$ is present, proportional to the instant value of $\varphi_m(t)$.

Setting ℓ_1 and ℓ_2 to the lowest possible values, the overall phase detector gain is

$$K_\varphi = \frac{S_v(f)}{S_{\varphi m}(f)} = \frac{gP_cR_0}{\ell_h\ell_m} \qquad (5)$$

for one quartz; g is the amplifier gain, P_c is the carrier power at the quartz output, R_0 is the mixer output impedance, ℓ_h is the loss of the 180° hybrid used as the power combiner in which the carrier is suppressed, and ℓ_m is the mixer loss; ℓ_m includes the 3 dB intrinsic loss due to conversion into upper and lower bands. With the described prototype, actual values are $g = 43$ dB, $\ell_h \simeq 0.3$ dB and $\ell_m \simeq 6$ dB.

The interferometric scheme is suitable to low power measurements because the quartz outputs drive an amplifier that rises the noise sidebands only. Moreover, this scheme is a good choice for low Fourier frequency measurements, close to $f = 1$ Hz, because it shows particularly low flicker phase noise. In fact, the latter comes from the dc-bias flicker up converted by the amplifier nonlinearity. Nevertheless, in this case the amplifier works in small signal regime, warranted by the carrier suppression mechanism, whereat the effect of nonlinearity becomes negligible. With the described implementation and after proper adjustment, residual carrier does not exceed -25 dBm at the amplifier output, which is some 40 dB lower than the maximum deliverable power of that device.

White noise is limited by the amplifier input noise $FkBT_0$, where $F = 1.7$ dB is the amplifier noise figure, $k_B = 1.38\times10^{-23}$ J/K is the Boltzmann constant and $T_0 = 290$ K is the reference temperature, close to the room temperature. In order to account for flicker, we replace F with $F(f)$. Fig. 4 shows the measured noise floor of the phase detector, reported in terms of $\mathcal{L}(f) = \frac{1}{2}S_\varphi(f)$; this choice is due to the available software package. The phase noise floor is $S_{\varphi m0} = 2\ell_hFk_BT_0/P_c$, ascribed to a single quartz, or $S_{\varphi m0} = \ell_hFk_BT_0/P_c$ for each quartz of a pair. Inspecting Fig. 4, we observe that 1 Hz noise is 5 dB higher than the floor level; combinig this piece of information with the amplifier noise figure $F = 1.7$ dB, far from the carrier, we get $F(1\,\text{Hz}) = 6.7$ dB. The noise floor thereby calculated is $\sigma_{y0}(\tau) = 6.6\times10^{-17}/\sqrt{P_c}$. Yet, P_d is dissipated by the motional resistance R_s of the resonator, which turns out to be in the 50–120 Ω range for 10 MHz high performance resonators (see Table I). Because output and dissipated power are related by $P_d = \frac{R_s}{R_0}P_c$, for the sake of simplicity we take a 3 dB margin instead of including R_s in the sensitivity model. Accordingly, we take $\sigma_{y0}(\tau) = 10^{-16}/\sqrt{P_d}$ as a conservative estimate of the measurement stability limit.

Some technical problems had to be solved prior to make the instrument function properly, the main of which are vibration and electromagnetic pollution by the ethernet, which also works at 10 MHz. For these reasons, the prototype is screwed on a 4 mm copper plate put on a 12 cm thick sand layer. All the circuits, plate and sand are enclosed in an iron box that serves as a magnetic and electric shield. In addition, the resonators are temperature stabilized, each one at its turning point.

4 Experimental Results

The instrument first needs adjustment and calibration. As the procedure requires some iterations, we provide

TABLE I: QUARTZ RESONATORS.

device	no.	$R_s\,\Omega$	L H	unl. Q	T °C
AT premium	1	57.5	1.25	1.37×10^6	48
	2	58.2	1.25	1.35×10^6	51
AT swept	1	59.0	1.21	1.28×10^6	52.5
	2	57.8	1.20	1.31×10^6	48
QAS	1	70.0	1.37	1.23×10^6	67
	2	71.9	1.44	1.26×10^6	72
BVA	1	105	2.22	1.33×10^6	80
	2	106	2.24	1.33×10^6	80

general guidelines only.

Both resonators are initially replaced with their equivalent resistances R_s. The detection phase γ'' is set for zero dc level at the mixer output when $\ell_1 = 0$ dB and $\ell_2 = \ell_{\max}$, so that the arm 2 is isolated. Then γ' is set for zero dc level at the mixer output with $\ell_2 = 0$ dB and $\ell_1 = \ell_{\max}$. Afterwards, the resonators are reinserted in the circuit. Small differences in the resonance frequency are corrected acting on C_1 and C_2, the effects of which are separated by setting to a high value the attenuator in series to the quartz we are not working on. All the above adjustments require the machine to work in non-suppressed carrier conditions, whereat ℓ_4 must be set to a high value, sufficient to prevent the amplifier to saturate; in all other cases, ℓ_4 is 0 dB. At this point, the carrier can be suppressed adjusting ℓ_1 and ℓ_2, and slightly retuning C_1 or C_2 if needed. The residual carrier is monitored by means of a directional coupler (not shown) inserted between the first and the second stage of the amplifier.

The gain K_φ and the source rejection are measured injecting through the appropriate directional coupler a known sideband at $f_s = 1$ Hz apart from ν_c; f_s must be lower than f_L, so that the quartz filter action is negligible.

The transfer function of one quartz is measured injecting white noise through the directional coupler in series to it. As an example, Fig. 5 shows the tranfer function of a 10 MHz BVA resonator. The cutoff can be easily individuated at $f_L = 4.5$ Hz, from which we calculate $Q = \frac{\nu_0}{2f_L} = 1.1\times10^6$.

Fig. 6 shows the phase noise of a pair of the same BVA, each one dissipating $P_d = 200\ \mu$W. The phase noise $S_{\varphi m}(1\,\text{Hz}) = -131$ dBrad^2/Hz taken from that figure, inserted in equation (3) yields a flicker floor $\sigma_y(\tau) = 10^{-13}$ for each quartz.

The flicker floor of the resonators reported in Tab. I has been measured, in some cases as a function of the dissipated power P_d. Results are shown in Fig. 7, together with the instrument stability limit. The QAS pair comes from oscillators of known parameters, flicker floor $\sigma_y(\tau) \simeq 1.7\times10^{-13}$ and dissipated power $P_d \simeq -13$ dBm [6]. The stability of those oscillators, reported in that figure, is close to the stability of the resonators, with a discrepancy of 1.5 dB.

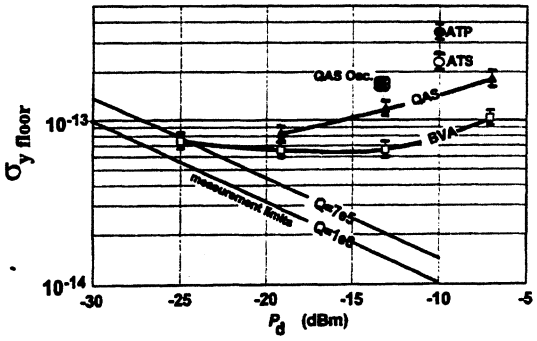


Figure 7: Allan deviation of some quartz resonators, deduced from phase noise measurements.

Acknowledgments

We would like to thank prof. M. Olivier from the LPMO for many scientific discussions, Mr. V. Candelier from CEPE and G. Robichon from BVA Industries for providing some quartz resonators, and Mr. E. Girardet of AR Electronique for supplying some high performance miniature ovens.

References

[1] F. L. Walls, A. E. Wanwright, "Measurement of short term stability of quartz crystal resonators and the implications for oscillator design and applications", *IEEE Trans. on IM* vol. 24 no. 1 pp. 15–20, March 1975.

[2] J. J. Gagnepain, "Fundamental noise studies of quartz crystal resonators" Proc. *30th Frequency Control Symposium* pp. 84–91, Atlantic City (USA), 2–4 Jun 1976.

[3] G. S. Curtis, The relationship between resonator and oscillator noise and resonator noise measurement technique", Proc. *41st Frequency Control Symposium* pp. 218–225, Philadelphia (USA), 1–3 Jun 1983.

[4] K. H. Sann, "The measurement of near-carrier noise in microwave amplifiers" *IEEE Trans. on MTT* vol. 16 no. 9 pp. 761–766, September 1968.

[5] E. Rubiola, V. Giordano, J. Groslambert, "Very high frequency and microwave interferometric PM and AM noise measurements", *Review of Scientific Instruments* vol. 70 no. 1 pp. 220–225, January 1999.

[6] V. Candelier, unpublished.

PM NOISE MEASUREMENTS OF 10 MHZ BVA QUARTZ CRYSTAL RESONATORS USING A PHASE BRIDGE SYSTEM WITH CARRIER SUPPRESSION

F. STHAL *, M. MOUREY *, F. MARIONNET *, W.F. WALLS ‡

* Laboratoire de Chronométrie, Electronique et Piézoélectricité,
Ecole Nationale Supérieure de Mécanique et des Microtechniques
26, Chemin de l'Epitaphe - 25030 BESANÇON CEDEX – France
Tel: (33) 3 81 40 28 31, Fax: (33) 3 81 88 57 14, email: fsthal@ens2m.fr

‡ Femtosecond Systems, Inc.
690 Arbutus Street; Golden, CO 80401 – USA
Tel: 303-462-0799, Fax: 303-462-0766, email: femtosecond@uswest.net

ABSTRACT

In this paper we review a new piece of equipment that allows one to characterize the phase noise of crystal resonators using a phase bridge system with carrier suppression. This equipment allows one to measure the inherent phase stability of quartz crystal resonators in a passive circuit without the noise usually associated with an active oscillator.

We achieved a system noise floor of approximately -150 dBc/Hz at 1 Hz and -160 dBc/Hz at 10 Hz. A PSPICE characterization of the carrier suppression system is given. An investigation of the phase modulation (PM) noise in 10 MHz BVA, SC cut, quartz crystal resonator pairs is presented.

1. INTRODUCTION

The first carrier suppression techniques to measure PM and AM noise were demonstrated by K.H. Sann[1] in 1968. A similar technique was proposed by C.H. Horn[2] in 1969. Recently several teams, working particularly in the microwave band, have used these techniques along with some modifications and improvements[3-5].

In this paper we evaluate a crystal resonator tester that was designed to assist in the PM noise characterization of quartz crystal resonators in the 1 to 200 MHz region*. This unit uses carrier suppression based on the bridge technique introduced by K.H. Sann[6]. This system allows one to measure the inherent phase stability of quartz crystal resonators in a passive circuit without the noise usually associated with an active oscillator.

We achieved a system noise floor of approximately -150 dBc/Hz at 1 Hz and –160 dBc/Hz at 10 Hz. A PSPICE characterization of the carrier suppression system is also given. Attenuation of the input signal introduced by the carrier suppression is shown. An investigation of the phase modulation (PM) noise in several pairs of 10 MHz BVA, SC cut, quartz crystal resonator is presented.

* The system used is a CR200A crystal resonator tester, Femtosecond systems, Inc.

2. MEASUREMENT SYSTEM

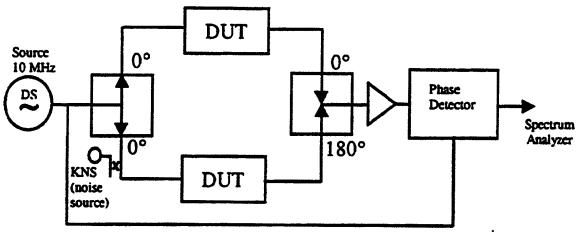


Fig. 1: Carrier suppression principle.

Fig. 1 shows the principle of the carrier suppression technique. The carrier signal of the driving source (DS) is split into two equal parts to drive both devices under test (DUT). Devices under test can be resistors (to measure the noise floor of the system) or crystal resonator pairs. The frequency of each arm of the bridge is tuned to the DS frequency with a series capacitor. The carrier signal is partially canceled when the two signals are combined 180° out of phase. Since phase noise is defined relative to the carrier power, reducing the carrier has the effect of amplifying the phase noise of the DUT. The output signal is amplified and then detected by the phase noise detector.

Calibration of the measurement system is obtained by injecting a known amount of phase noise on one of the arms of the bridge. This known noise source (KNS) is calibrated independently and relative to the carrier level on one arm of the bridge. The noise source is turned on to measure the sensitivity of the measurement system. The noise source is then turned off in order to measure the noise of the DUT. The noise of the DUT, as seen on the FFT analyzer, is corrected using the calibration factor determined when the noise was on. This calibration is done for each frequency measured and gives a very precise calibration of the system.

A carrier suppression of 10-60 dB can be obtained, but a sufficient level at the input of the phase noise detector is necessary to maintain a low noise floor (4 dBm is the minimum level). The struggle in using this system was to get enough carrier suppression to get a good measurement of the noise of the resonators

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

without saturating the output amplifiers with the noise of the KNS. Since the KNS is only on one arm of the bridge, it doesn't get any suppression before it sees the full gain of the output amplifier string.

3. PSPICE MODEL

The carrier suppression system was simulated using PSPICE. The equivalent circuit of the system is shown in Fig. 2.

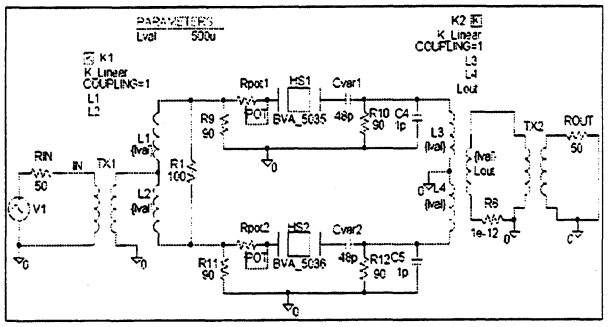


Fig. 2: PSPICE simulation of the carrier suppression.

The "IN" point is the input of the device. A perfect 0° power splitter allows the source to be split into two identical signals[7]. The Quartz crystals were simulated with their classical equivalent circuit around their resonant frequency and matched by several resistors. The resonant frequency of each arm of the bridge is tuned to the driving source frequency with a series capacitor $Cvar_i$ (i = 1,2). The carrier suppression is obtained using a perfect power combiner 0°-180°. The power combiner subtracts the output signals of the quartz crystals. Calibration of the simulation was obtained by measuring the power in the system. Fig. 3 shows the variation of the measured power after the power combiner.

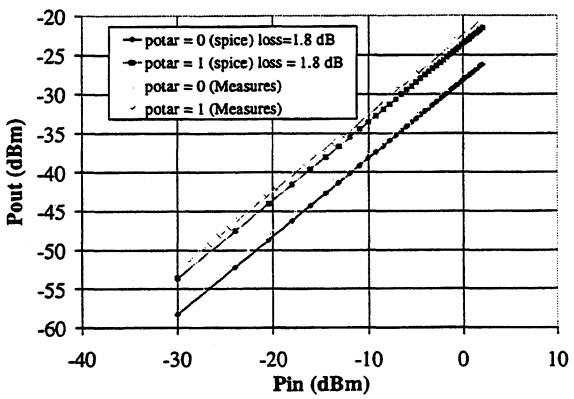


Fig. 3: Power at the output of the system.

The loss was introduced as a correction factor that is the sum of the loss in the real combiners and real coupler. The power introduced to the quartz crystal, according to the input power, is given in Fig. 4. This simulation is obtained with a resonator on only one side of the bridge. The "Potar" values correspond to the minimum and the maximum power that is possible to drive the quartz crystal. We can see that the power goes to 300 μW.

In practice, we have observed an amplitude frequency effect of the quartz crystal of about 2 Hz between the two extreme positions of the potentiometer when the input power was at a maximum. That corresponds to a 200 μW variation of the power applied to the crystal since the amplitude frequency effect is around $\Delta f/f = 10^{-9}/\mu W$ for this cut.

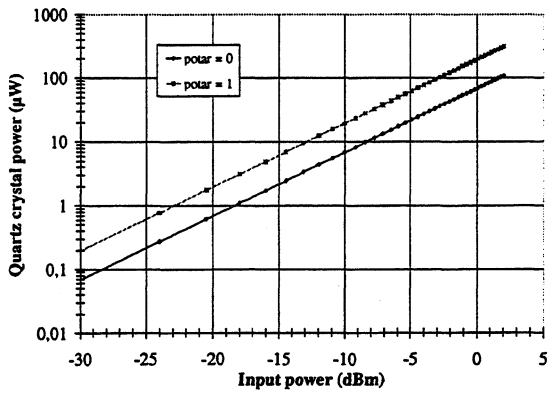


Fig. 4: Quartz crystal power versus input power.

4. NUMERICAL SIMULATION

4.1 Effect of the unbalanced power

The simulation is given using a linear approach. The transfer function of the carrier suppression system gives the attenuation of the noise of DS. This transfer function is shown in Fig. 5.

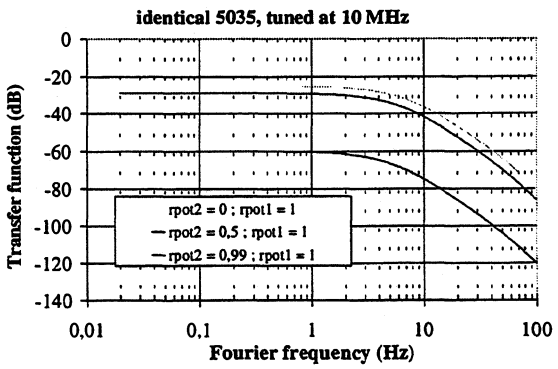


Fig. 5: Transfer function of the carrier suppression

In this simulation two identical quartz crystals are used. We can see that the difference of the power between both arms allows us to regulate the carrier suppression. At 1 Hz, the source noise can be suppressed between 20 to 60 dB. The source noise is not negligible if we have to unbalance the power in both arms to keep the phase noise detector calibrated. For instance, a noise of -100 dBc/Hz at 1 Hz from the

carrier can be included between −120 dBc/Hz and -160 dBc/Hz. Therefore, the noise of the driving source will be above the noise of the DUT. The noise of the driving source must be as low as possible.

4.2 Effect of the inherent frequencies of the resonators

The transfer function of the carrier suppression device is represented according to the resonant frequency of one resonator in Fig. 6. Nu_1 is the inherent resonant frequency of one resonator. The resonant frequency of the second resonator is chosen to equal 9,999,986 Hz. The resonant frequency of each arm of the bridge is tuned to the frequency of DS with a series capacitor. Adjusting the potentiometer allows one to set the signal level to about 4 dBm at the input to the phase noise detector.

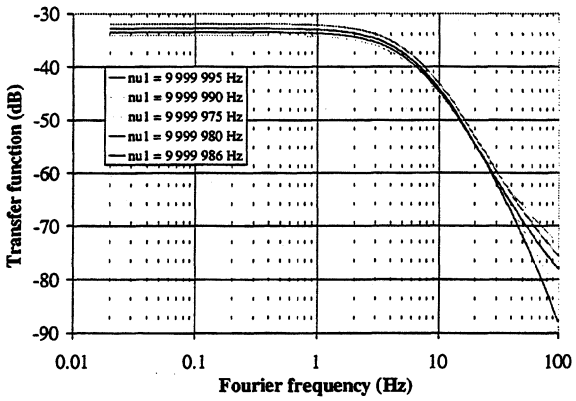


Fig. 6 : Effect of the frequencies of the resonators.

Following the previous conditions, one can see that the influence of the frequencies of the resonators seems to be very low in the transfer function if we adjust the series capacitor to obtain the resonant frequency of each arm of the bridge at the frequency of DS.

5. MEASUREMENTS

The noise floor of the system is given in Fig. 7.

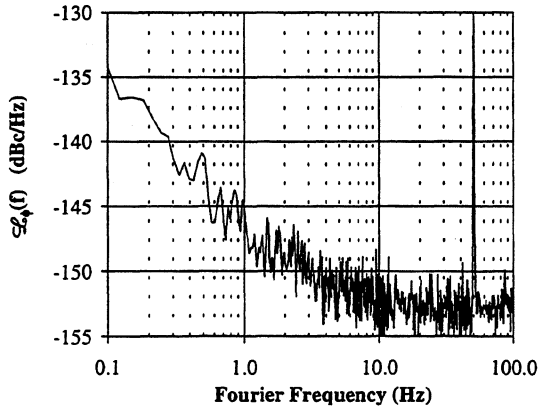


Fig. 7: Noise floor of the system.

This measurement was obtained using two 86 Ω resistors. The driving source, DS, used in these measurements has a noise floor of −110 dBc/Hz at 1 Hz, -138 dBc/Hz at 10 Hz and a noise floor around -150 dBc/Hz. The measurement results are given in terms of the single side band power spectral density of the phase fluctuations, $\mathcal{L}_\phi$(f). We achieved a system noise floor of approximately -145 dBc/Hz at 1 Hz and -153 dBc/Hz at 10 Hz.

Noise of a couple of 10 MHz, SC cut, BVA Quartz crystal is given in Fig. 8. The crystal measurement and the resistor measurement was obtained with a similar sensitivity and drive. The noise floor appears to be about 10 dB below the crystal signal level close to the carrier. We can observe the 1/f attenuation of the quartz crystal noise.

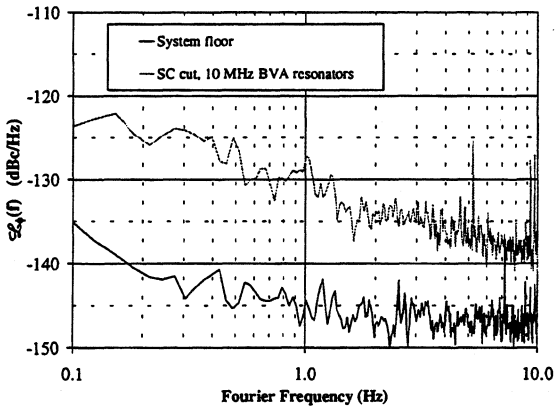


Fig. 8: Phase noise of the 10 MHz, SC cut, BVA quartz crystal resonators.

Short-term stability usually is given in the time domain with an Allan variance. The Allan variance is computed from the Fourier frequency and from the power spectral density of the phase fluctuations [8]:

$$\sigma_y^2(\tau) = \int_0^\infty S_\phi(f) \frac{2\sin^4 \pi f\tau}{(\pi \upsilon_o \tau)^2} df$$

with f = Fourier frequency, ν_o = source frequency.

Standard deviation is computed at $2 \cdot 10^{-14}$ at 1 s following the previous results (Fig. 8). The numerical integral calculus of the σ_y(1s) is obtained with integral limits equal to 0.1 Hz and 10 Hz. The power spectral density of the phase fluctuations S_ϕ(f) which is used in this computation is an 1/f approximation of the previous results. It is an asymptotic curve including eight hundred points and these values: S_ϕ(0.1Hz) = -120 dBc/Hz, S_ϕ(1Hz) = -130 dBc/Hz and S_ϕ(10 Hz) = -140 dBc/Hz.

Fig. 9 shows a comparison between several quartz crystals. The following curves show that quartz crystal "2311" is the crystal that has the most phase noise.

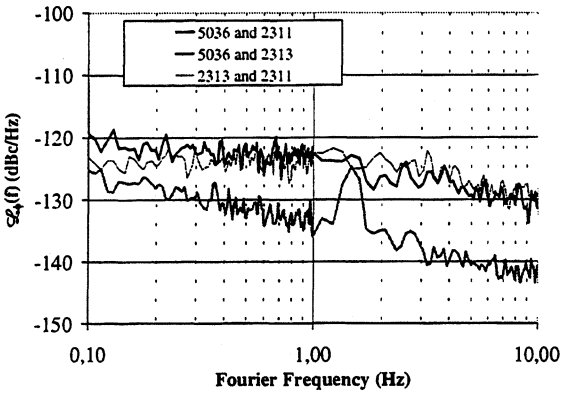


Fig. 9: Comparison between three 10 MHz, SC cut BVA quartz crystal resonator.

DUT noise is shown clearly in the range 0.1 Hz-10 Hz for the "5036" and "2313" resonators. If the noise of the DUT is above the calibration noise source, calibration is obtained with an unknown noise. This phenomenon appears below 1 Hz for the "2311" resonator. The results are therefore not calibrated.

6. CONCLUSION

We achieve the measurement of the phase noise of 10 MHz, SC cut, BVA quartz crystal resonators in their frequency bandwidth. Standard deviation is computed at $2 \cdot 10^{-14}$ at 1 s for the best measured resonators.

Calibration of the system is difficult. The noise level of KNS is very important and must be above the noise of the DUT. Measurement of the noise outside the bandwidth of the quartz crystal is impossible with the proposed calibration. The calibration is obtained with the quartz crystals and, when the KNS is turned off, the results are independent of the resonators as soon as the noise floor of the system is measured.

In reviewing these results and the system overall, it may be helpful to have an even better 10 MHz DS so that less carrier suppression is required. This would allow one to use the KNS with fewer problems. Another approach would be to work from a $S_y(f)$ viewpoint and calibrate the system with using frequency shifts in the reference source (DS) to view different sensitivities at the output of the phase noise detector. For a known step in frequency you would get a change at the output of the phase noise detector.

ACKNOWLEDGEMENT

We thank F.L. Walls and S. Galliou for their helpful discussions.

7. REFERENCES

[1] K.H. SANN, "The measurement of near-carrier noise in microwave amplifiers", IEEE Trans. On Microwave Theory and Tech., vol. MTT-16, No. 9, pp.761-766, Sept., 1968.

[2] C.H. HORN, "A carrier suppression technique for measuring S/N and carrier/sideband ratios greater thean 120 dB", Proc 23rd Ann. Freq. Cont. Symp., Atlantic City, NJ, pp. 6-8, May, 1969.

[3] E.N. IVANOV, M.E. TOBAR, R.A. WOODE, "Ultra-low noise microwave oscillator with advanced phase noise suppression system", IEEE Trans on Microwave and Guided Wave letters, vol. 6, No 9, pp. 312-315, 1996.

[4] E.N. IVANOV, M.E. TOBAR, R.A. WOODE, "A study of noise phenomena in microwave components using an advanced noise measurement system", IEEE Trans. Ultrason. Ferroelectr. Freq. Cont. Vol. 44, No 1, pp. 161-163, Jan., 1997.

[5] E. RUBIOLA, V. GIORDANO, J. GROSLAMBERT, "VHF and microwave interferomatric PM and AM noise measurements", Proc. 12th EFTF, Warsaw, Poland, pp. 286-291, March, 1998.

[6] F.L. WALLS, "Suppressed carrier based PM and AM noise measurement techniques", Proc. 51st Ann. Freq. Cont. Symp., Orlando, FL, pp. 485-492, May, 1997.

[7] F. STHAL, M. MOUREY, "PSPICE simulation of a phase modulation measuring system of quartz crystal resonators", Proc. 12th EFTF, Warsaw, Poland, pp. 155-160, March, 1998.

[8] J. RUTMAN, "Characterization of phase and frequency instabilities in precision frequency source: Fifteen years of progress", Proc. IEEE, vol. 66, pp. 1048-1074, Sep., 1978.

1999 Joint Meeting EFTF - IEEE IFCS

PM AND AM NOISE IN NONLINEAR BJT AMPLIFIER

V.N.Kuleshov, Moscow Power Engineering Institute (Technical University),
14 Krasnokasarmennaya, Moscow 111250, Russia, e-mail: kuleshov@srv-vmss.mpei.ac.ru

ABSTRACT

An approach to nonlinear BJT amplifier PM and AM noise calculation is presented. Polyharmonic presentation of voltages and currents is used and periodical nonstationarity of BJT noise sources due to their modulation by periodic signal is taken into consideration. Due to these effects one has to introduce different "noise factors" for PM and AM noise calculation (if one expands the formulae obtained for linear mode of operation to nonlinear one), and these noise factors depend on the input signal level.

Common emitter (CE) amplifier with both bypassed and unbypassed resistors in emitter circuit is considered in detail. Power spectral densities of AM and PM noise are calculated as functions of signal amplitude. An influence of bypassed and unbypassed emitter resistor on PM and AM noise is investigated.

1. INTRODUCTION

Most of the papers that were devoted to theoretical analysis of PM and AM noise in bipolar junction transistor (BJT) amplifiers were based on using linear model of BJT. Using of such model gives an opportunity to obtain simple formulae for wideband PM and AM noises [2], based on noise factor calculation. This approach was also used for oscillators PM noise calculations based on Leeson's formula [3].

In many cases amplifiers with low level of PM noise are used in nonlinear mode of operation. They are almost always nonlinear in oscillators. But for nonlinear amplifier it is not correct to use simple formulae that contain such characteristics as noise factor, because in nonlinear amplifier input oscillation and noises are transformed to the output in different ways, and power spectral densities of PM and AM noises introduced by nonlinear amplifier are different. It was discussed for amplifier with low frequency nonlinear model of BJT in the book [4].

In this paper we consider PM and AM noises introduced by wideband noise sources of nonlinear amplifier using high frequency charge control model of BJT for active region of its operation. Results of this consideration are usefull to find limitations of linear model and present opportunity to calculate PM and AM noise in nonlinear BJT amplifiers.

2. AMPLIFIER CIRCUIT AND MODEL

The circuit under consideration is shown in Fig.1. It is CE amplifier with feedback impedance Z_E in emitter circuit. Nonlinear circuit model of the amplifier is shown in Fig.2. The transistor is replaced by its charge control model with noise sources. It is supposed that

$$i_c(\vartheta_{b'e}) = I_S\left[\exp\left(\frac{\vartheta_{b'e}}{V_T}\right) - 1\right] \qquad (1)$$

and

$$i_b(\vartheta_{b'e}) = \left(\frac{1}{\beta} + \tau_F p\right) i_c(\vartheta_{b'e}) \qquad (2),$$

where I_S - saturation current, V_T - thermal voltage, β - CE configuration current gain, τ_F - forward injection charge control parameter, $p=d/dt$ - operator of differentiation. We don't take into consideration dependence of junction capacitances C_{je} and C_{jc} on the voltages.

In the circuit diagram (Fig.1) $v_{SS}(t) = \mathrm{Re}[V_{SS}\exp(j\omega_s t)]$ - signal source voltage, and $Z_{SS}(p)$ - signal source impedance.

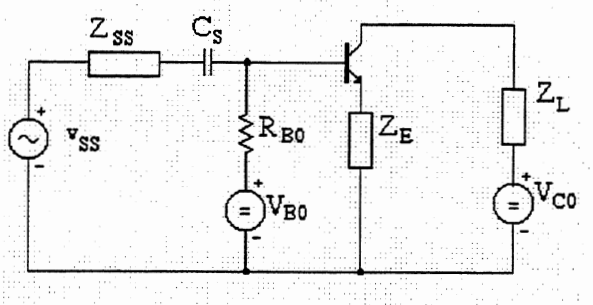


Fig.1. Circuit diagram of CE amplifier.

In the circuit model (Fig.2) $v_S(t)=K_S(p)v_{SS}(t)$ where $K_S(p)=R_{B0}/(R_{B0}+Z_{SS}(p)+(1/pC_S))$ and $Z_S(p)=(Z_{SS}(p)+(1/pC_S))/(1+(Z_{SS}(p)/R_{B0})+(1/pC_SR_{B0}))$.

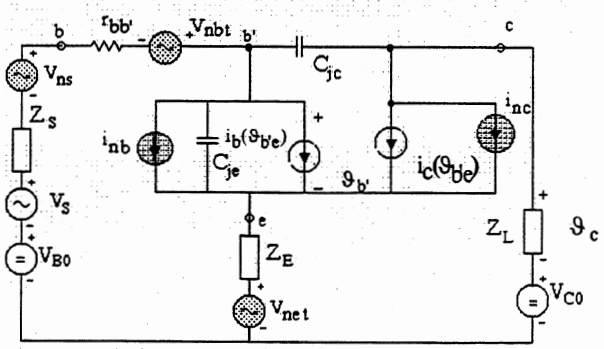


Fig. 2. Circuit model of the CE amplifier.

Although an approach to PM and AM noise calculation is applicable to the circuits with any linear impedances $Z_{SS}(p)$, $Z_L(p)$, $Z_E(p)$, in this paper we consider

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

quantitatively the case. when $Z_{SS}(p)=R_{SS}$. $Z_L(p)=R_L$ and $Z_E(p)=R_E+ R_{E0}/(1+C_{E0}R_{E0})$. Here R_{E0}. C_{E0} - elements of emitter self biasing circuit. and R_E - a resistor in feedback circuit.

Five wideband noise sources are taken into consideration in the model shown in Fig.2: thermal noise voltages $v_{ns}(t)$. $v_{nbt}(t)$ and $v_{net}(t)$ of impedances $Z_S(p)$. $r_{bb'}$. $Z_E(p)$. and noise components $i_{nb}(t)$ and $i_{nc}(t)$ of base current and collector current. These two noise currents become periodically nonstationary when input signal of the amplifier becomes large.

3. BASIC EQUATIONS

Nonlinear model of the BJT amplifier shown in Fig.2 can be described by two symbolic equations

$$\vartheta_{b'e} = -Z(p)i_c(\vartheta_{b'e}) + \bar{\vartheta}_{s.b'e} + \vartheta_{n.b'e}, \quad (3)$$

$$\vartheta_c = -Z_c(p)i_c(\vartheta_{b'e}) + \bar{\vartheta}_{s.c} + \vartheta_{n.c}, \quad (4)$$

where $Z(p)$ and $Z_c(p)$ - symbolic impedances that describe an influence of $i_c(\vartheta_{b'e})$ on $\vartheta_{b'e}$ and ϑ_c through liner part of the circuit model shown in Fig.2. $\bar{\vartheta}_{s.b'e}$ and $\bar{\vartheta}_{s.c}$ - input voltages transformed to emitter junction and to the amplifier output by linear part of the circuit model. $\vartheta_{n,b'e}$ and $\vartheta_{n.c}$ - noise voltages that appear at the emitter junction and the amplifier output due to influence of all wideband noise sources through linear part of the circuit model.

In steady state mode of the amplifier operation $\vartheta_{b'e}(t)$ and $\vartheta_c(t)$ can be presented as sums of periodical functions of time $\bar{\vartheta}_{s.b'e}$ and $\bar{\vartheta}_{s.c}$ with period $T_S = 2\pi/\omega_S$, and small fluctuations $\vartheta_{f,b'e}$ and $\vartheta_{f,c}$, so that

$$\vartheta_{b'e} = \bar{\vartheta}_{b'e} + \vartheta_{f,b'e}, \qquad \vartheta_c = \bar{\vartheta}_c + \vartheta_{f,c}. \quad (5)$$

Periodic voltage $\bar{\vartheta}_{b'e}$ is a solution of the equation

$$\bar{\vartheta}_{b'e} = -Z(p)i_c(\bar{\vartheta}_{b'e}) + \bar{\vartheta}_{s.b'e}. \quad (6)$$

After its solution $\bar{\vartheta}_c$ can be found as follows

$$\bar{\vartheta}_c = -Z_c(p)i_c(\bar{\vartheta}_{b'e}) + \bar{\vartheta}_{s.c}. \quad (7)$$

If one substitutes (5) in (3). (4). linearize $i_c(\bar{\vartheta}_{b'e} + \vartheta_{f.b'e})$ with respect to fluctuations. and take into consideration (1), (6) and (7), one obtains the next equations for fluctuations:

$$\left(1 + Z(p)\frac{i_c(\bar{\vartheta}_{b'e})}{V_T}\right)\vartheta_{f.b'e} = \vartheta_{n.b'e}, \quad (8)$$

$$\vartheta_{f.c} = -Z_c(p)\frac{i_c(\bar{\vartheta}_{b'e})}{V_T}\vartheta_{f.b'e} + \vartheta_{n.c}. \quad (9)$$

As $i_c(\bar{\vartheta}_{b'e})$ is a periodic function of time. fluctuations $\vartheta_{f,b'e}$ and $\vartheta_{f.c}$ are periodically nonstationary. To find them one has to use an approach that is different from the one used for calculation of stationary fluctuations [5].

4. CALCULATION OF THE PM AND AM NOISE PSD

It was shown in [5] that any periodically nonstationary processes v(t) with period of nonstationarity $T_S = 2\pi/\omega_S$ can be presented in the form

$$v(t) = \sum_{k=-\infty}^{\infty} U_k(t)\exp(jk\omega_S t) \quad (10)$$

where $U_k(t)$ - stationary complex random processe having PSD located in Fourier frequency band $-\frac{\omega_S}{2} \le \omega < \frac{\omega_S}{2}$. So one can say that vector $\bar{U} = (\ldots U_{-k}(t), \ldots U_{-1}(t), U_0(t), U_1(t), \ldots, U_k(t)\ldots)^T$. where «T» is a symbol of matrix transposition. corresponds to the process v(t). A matrix of PSD's of the vector $\bar{U}$ gives full description of spectral characteristics of v(t).

It was shown also that using such approach one can obtain from (8). (9) systems of linear equations

$$\left[\hat{E} + \frac{1}{V_T}\hat{Z}(j\omega)\hat{I}_C\right]\bar{U}_{f.b'e}[j\omega] = \bar{U}_{n.b'e}[j\omega], \quad (11)$$

$$\bar{U}_{f.c}[j\omega] = -\frac{1}{V_T}\hat{Z}_c(j\omega)\hat{I}_C\bar{U}_{f.b'e}[j\omega] + \bar{U}_{n.c}[j\omega], \quad (12)$$

where matrixes $\hat{Z}(j\omega)$, $\hat{Z}_c(j\omega)$, $\hat{I}_c$ are defined as follows

$$\hat{Z}(j\omega) = \|Z_{mk}\|, \text{ where } \begin{vmatrix} Z_{mk} = Z(j\omega + jm\omega_S), \text{if } m = k, \\ Z_{mk} = 0, \text{if } m \ne k, \end{vmatrix} \quad (13)$$

$$\hat{Z}_c(j\omega) = \|Z^c_{mk}\|, \text{ where } \begin{vmatrix} Z^c_{mk} = Z_c(j\omega + jm\omega_S), \text{if } m = k, \\ Z^c_{mk} = 0, \text{if } m \ne k, \end{vmatrix} \quad (14)$$

$$\hat{I}_C = \|I^C_{mk}\| = \|I^C_{m-k}\|, \quad (15)$$

where I^C_k - complex semi amplitudes of Fourier series

$$i_c(\bar{\vartheta}_{b'e}) = \sum_{k=-\infty}^{\infty} I^C_k \exp(jk\omega_S t).$$

If we introduce matrixes

$$\hat{D}(j\omega) = \hat{E} + \frac{1}{V_T}\hat{Z}(j\omega)\hat{I}_C, \quad \hat{C}(j\omega) = \frac{1}{V_T}\hat{Z}_c(j\omega)\hat{I}_C, \quad (16)$$

we can write formal solution of equations (11). (12) in the form

$$\bar{U}_{f.c}[j\omega] = -\hat{C}(j\omega)\hat{D}^{-1}(j\omega)\bar{U}_{n.b'e}[j\omega] + \bar{U}_{n.c}[j\omega] \quad (17)$$

Noise voltages $\vartheta_{n,b'e}$ and $\vartheta_{n.c}$ can be presented as linear combinations of primary noise sources $v_{ns}(t)$. $v_{nbt}(t)$, $v_{net}(t)$, $i_{nb}(t)$ and $i_{nc}(t)$.

$$\vartheta_{n.b'e} = K_{s.b'e}(p)[v_{ns} + v_{nbt}] + K_{e.b'e}(p)v_{net} + \\ + Z_{b.b'e}(p)i_{nb} + Z_{c.b'e}(p)i_{n.c} \quad (18)$$

$$\vartheta_{n.c} = K_{s.c}(p)[v_{ns} + v_{nbt}] + K_{e.c}(p)v_{net} + \\ + Z_{b.c}(p)i_{nb} + Z_{c.c}(p)i_{n.c} \quad (19)$$

where symbolic transfer functions and transfer impedances depend only on linear passive components of the model shown in Fig.2.

If we introduce

- matrixes, corresponding to the transfer functions and impedances of eq. (18),(19), using the same algorithm as in (13), (14).
- vectors of random spectra $\bar{U}_{ns}[j\omega], \bar{U}_{nbt}[j\omega], \bar{U}_{net}[j\omega], \bar{I}_{nb}[j\omega], \bar{I}_{nc}[j\omega]$ corresponding to v_{ns}, v_{nbt}, v_{net}, i_{nb}, i_{nc}.
- matrixes

$$\left.\begin{aligned} \hat{K}_{nsb}(j\omega) &= -\hat{C}(j\omega)\hat{D}^{-1}(j\omega)\hat{K}_{s.b'e}(j\omega) + \hat{K}_{s.c}(j\omega), \\ \hat{K}_{ne}(j\omega) &= -\hat{C}(j\omega)\hat{D}^{-1}(j\omega)\hat{K}_{e.b'e}(j\omega) + \hat{K}_{e.c}(j\omega), \\ \hat{Z}_{nb}(j\omega) &= -\hat{C}(j\omega)\hat{D}^{-1}(j\omega)\hat{Z}_{b.b'e}(j\omega) + \hat{Z}_{b.c}(j\omega), \\ \hat{Z}_{nc}(j\omega) &= -\hat{C}(j\omega)\hat{D}^{-1}(j\omega)\hat{Z}_{c.b'e}(j\omega) + \hat{Z}_{c.c}(j\omega), \end{aligned}\right\} \quad (20)$$

then from (17),(18),(19),(20) we obtain

$$\bar{U}_{f.c}[j\omega] = \hat{K}_{nsb}(j\omega)\left[\bar{U}_{ns}[j\omega] + \bar{U}_{nbt}[j\omega]\right] + \\ + \hat{K}_{ne}(j\omega)\bar{U}_{net}[j\omega] + \hat{Z}_{nb}(j\omega)\bar{I}_{nb}[j\omega] + \hat{Z}_{nc}\bar{I}_{nc}[j\omega]. \quad (21)$$

It is known that matrix of PSD of periodically nonstationary random process (for example, $\vartheta_{fc}(t)$) is connected with vector of random spectra of this process ($\bar{U}_{f.c}[j\omega]$) by equation

$$\left(\overline{\bar{U}^{*}_{f.c}[j\omega]\bar{U}^{T}_{f.c}[j\omega]}\right) = \delta(\omega - \omega')\hat{S}^{fc}_{U}(\omega) \quad (22)$$

where «*» denotes complex conjugate value, $\delta(\omega - \omega')$ - delta function and matrix $\hat{S}^{fc}_{U}(\omega)$ has the next form

$$\hat{S}^{fc}_{U}(\omega) = \begin{pmatrix} \dots & \dots & \dots & \dots & \dots \\ \dots & S_{-1-1}(\omega) & S_{-10}(\omega) & S_{-11}(\omega) & \dots \\ \dots & S_{0-1}(\omega) & S_{00}(\omega) & S_{01}(\omega) & \dots \\ \dots & S_{1-1}(\omega) & S_{10}(\omega) & S_{11}(\omega) & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix} \quad (23)$$

It corresponds to vector complex stationary random process

$$\bar{U}_{f.c} = (\dots\ U_{fc.-2}\ \ U_{fc.-1}\ \ U_{fc.0}\ U_{fc.1}\ \ U_{fc.2}\ \dots). \quad (24)$$

For example $S_{11}(\omega)$ is PSD of $U_{fc.1}(t)$, and $S_{-11}(\omega)$ is mutual PSD of $U_{fc.-1}(t)$ and $U_{fc.1}(t)$

As $$\vartheta_{f.c}(t) = \sum_{k=-\infty}^{\infty} U_{fc.k}\exp(jk\omega_s t), \quad (25)$$

the PSD of the process $\vartheta_{fc}(t)$) can be presented in the form

$$S^{fc}_{\vartheta}(\omega) = \begin{cases} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ S_{-1-1}(\omega + \omega_s), & \frac{-3\omega_s}{2} \le \omega < -\frac{\omega_s}{2}, \\ S_{00}(\omega), & \frac{-\omega_s}{2} \le \omega < \frac{\omega_s}{2}, \\ S_{11}(\omega - \omega_s), & \frac{\omega_s}{2} \le \omega < \frac{3\omega_s}{2}, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \end{cases} \quad (26)$$

Specific feature of periodicaly nonstationary random processes is nonzero mutual correlation of components of vector (24).

From (21) and (22) we obtain

$$\hat{S}^{fc}_{U}(\omega) = \hat{K}^{*}_{nsb}(j\omega)\left[\hat{S}^{ns}_{U}(\omega) + \hat{S}^{nbt}_{U}(\omega)\right]\hat{K}^{T}_{nsb}(j\omega) + \\ + \hat{K}^{*}_{ne}(j\omega)\hat{S}^{net}_{U}(\omega)\hat{K}^{T}_{ne}(j\omega) + \\ + \hat{Z}^{*}_{nb}(j\omega)\hat{S}^{nb}_{i}(\omega)\hat{Z}^{T}_{nb}(j\omega) + \hat{Z}^{*}_{nc}(j\omega)\hat{S}^{nc}_{i}(\omega)Z^{T}_{nc}(j\omega) \quad (27)$$

Matrix of primary thermal noise of complex impedance $Z(j\omega)$ is

$$\hat{S}^{n}_{U}(\omega) = 2kT\begin{pmatrix} \dots & \dots & \dots & \dots & \dots \\ \dots & \mathrm{Re}(Z(j\omega - j\omega_s)) & 0 & 0 & \dots \\ \dots & 0 & \mathrm{Re}(Z(j\omega)) & 0 & \dots \\ \dots & 0 & 0 & \mathrm{Re}(Z(j\omega + j\omega_s)) & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix} \quad (28)$$

where k - Boltzmann constant and T - absolute temperature. Matrixes $\hat{S}^{ns}_{U}(\omega), \hat{S}^{nbt}_{U}(\omega)$ and $\hat{S}^{net}_{U}(\omega)$ can be obtained from $\hat{S}^{n}_{U}(\omega)$ if one replaces $Z(j\omega)$ by $|K_S(j\omega)|^2 Z_{SS}(j\omega)$, $[Z_S(j\omega) - |K_S(j\omega)|Z_{SS}(j\omega) + r_{bb'}]$ and $Z_E(j\omega)$. Matrix of shot noise of collector current $i_{nc}(t)$ and base current $i_{nb}(t)$ are

$$\hat{S}^{nc}_{i}(\omega) = e\hat{I}^{T}_{C}, \qquad \hat{S}^{nc}_{i}(\omega) = \frac{1}{\beta}e\hat{I}^{T}_{C}, \quad (29)$$

where e -electron charge and $\hat{I}_C$ -defined by (15).

Using (27), (28), (29) one can calculate matrix of PSD's that gives full description of vector $\bar{U}_{f.c}$ (24) and periodically nonstationary noise voltage at the output of the amplifier.

To calculate PM $\phi(t)$ and AM a(t) noises introduced in first harmonic of the output voltage (5) we can use a definition of these noises

$$a(t) + j\,\phi(t) = \frac{U_{fc.1}}{U_{c.1}}, \quad (30)$$

where $U_{c.1}$ - a complex semi amplitude of the first harmonic of $\bar{\vartheta}_c(t)$. From (30) it follows that

$$\begin{pmatrix} a(t) \\ \phi(t) \end{pmatrix} = \frac{1}{2|U_{c1}|}\hat{X}\begin{pmatrix} U_{fc.-1} \\ U_{fc.1} \end{pmatrix}, \quad (31)$$

where $\hat{X} = \begin{pmatrix} e^{j\varphi} & e^{-j\varphi} \\ je^{j\varphi} & -je^{j\varphi} \end{pmatrix}$, $e^{j\varphi} = U_{c1} / |U_{c1}|$,

$$\begin{pmatrix} S_a(\omega) & S_{a\phi}(\omega) \\ S_{\phi a}(\omega) & S_\phi(\omega) \end{pmatrix} = \frac{1}{4|U_{c1}|^2} \hat{X}^* \begin{pmatrix} S_{-1-1}(\omega) & S_{-11}(\omega) \\ S_{1-1}(\omega) & S_{11}(\omega) \end{pmatrix} \hat{X}^T, \quad (32)$$

where elements $S_{-1-1}(\omega)$, $S_{-11}(\omega)$, $S_{1-1}(\omega)$ and $S_{11}(\omega)$ are taken from matrix (23), calculated in (27).

To find one sided (physical) PSD's of $\phi(t)$ and a(t) one has to multiply $S_\phi(\omega)$ and $S_a(\omega)$ by 2.

Formula (32) gives full solution of the problem PM and AM noise calculation in nonlinear amplifier. Basing on this approach a program for PM and AM noise calculation was developed.

5. RESULTS OF PM AND AM NOISE CALCULATION

Some results of PM and AM noise calculation in the amplifier Fig.1 are shown in Fig.3.

To show a difference between linear amplifier and nonlinear one, we presented $S_\phi(\omega)$ and $S_a(\omega)$ using the same form as in linear amplifier [2]. But because of the difference of PM and AM noises we have to introduce here «phase» F_ϕ and «amplitude» F_a noise factors (NF):

$$S_\phi(\omega) = F_\phi \frac{kT}{P_{av}}, \qquad S_a(\omega) = F_a \frac{kT}{P_{av}},$$

where $P_{av} = V_{SS}^2 / 8 \operatorname{Re} Z_{SS}$.

Calculations were fulfilled for CE amplifier considered in [2]. Parameters of the circuit model were: $r_{bb'}$=30 Ohm, β=50, $(\omega_T/2\pi)$=500 MHz, C_{je}=5 pF, C_{jc}=4pF, $Z_{SS}=R_{SS}$=250 Ohm, C_S=10 nF, R_{BO}=2,5kOhm, C_{E0}=1000 nF, $Z_L=R_L$=500 Ohm. Collector current in operating point without signal at the input I_{c0}=5mA. Signal frequency - $(\omega_s/2\pi)$=2,5 MHz.

Fig.3,a shows dependencies of first harmonic average gain $G=U_{c1}/V_{SS}$ and second harmonic «gain» $H=U_{c2}/V_{SS}$. Fig.3,b shows dependencies of PM and AM NF's on V_{SS}.

Three cases were investigated. The first case (R_{E0}=0, R_E=0) shows significant growth of NF's with increasing V_{SS} in nonlinear region. The second case (R_{E0}=250 Ohm, R_E=0) shows that bypassed resistor R_{E0} decreased AM NF $F_a^{(2)}$ significantly, and lowered PM NF $F_\phi^{(2)}$ a little bit. The third case (R_{E0}=250 Ohm, R_E=25 Ohm) shows an influence of unbypassed emitter resistor on NF's $F_\phi^{(3)}$ and $F_a^{(3)}$.

6. CONCLUSIONS

An approach and a program presented in this paper give an opportunity to investigate an influence of signal frequency, R_{BO}, R_{EO}, Z_{SS}, R_E and Z_L on PM and AM noise in nonlinear BJT amplifiers due to wideband electrical noises sources.

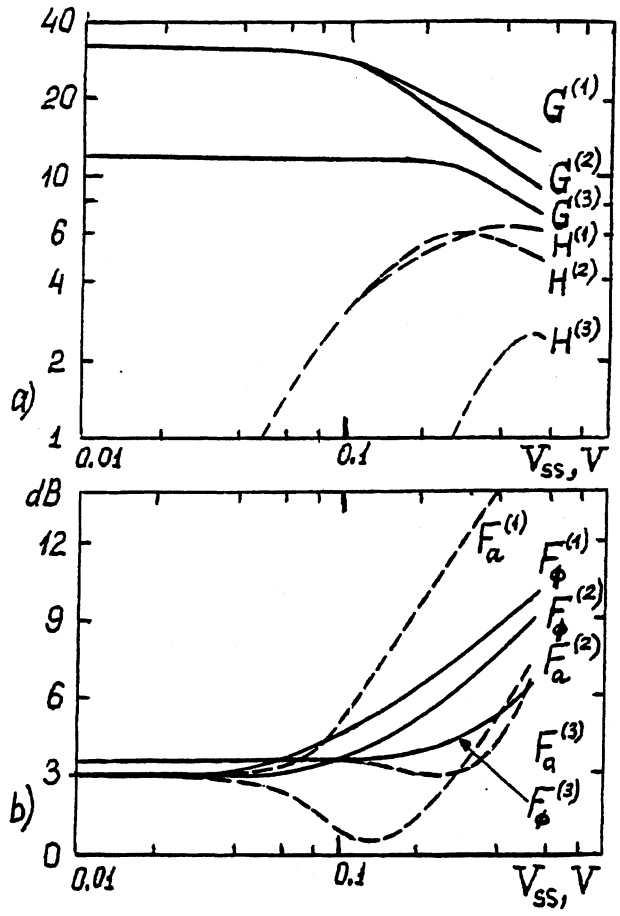


Fig.3. Average «gains» and PM and AM «noise factors» as functions of input voltage V_{SS}.

REFERENCES

1. F.L.Walls, E.S.Ferre-Pikal, and S.R.Jefferts. «The Origin of 1/f PM and AM Noise in Bipolar Junction Transistor Amplifiers».- presented at the IEEE Int. Freq. Contr. Symp., San Francisco, California, U.S.A., May - June, 31 - 2, 1995.
2. T.I.Boldyreva. «Comparative Study of Phase and Amplitude Noise of CE, CE-CB and CC-CC Bipolar Transistor Amplifiers», presented at the IEEE Int. Freq. Contr. Symp, Pasadena, California, U.S.A.,May 27-29, 1998.
3. D.B.Leeson. «A Simple Model of Feedback Oscillator Noise Spectrum». Proc. IEEE, vol. 54, N2, pp.329-330,1966.
4. V.Zalud, V.N.Kuleshov. Noise in Semiconductor Devices, Soviet Radio, Moscow, 1977, (In Russian).
5. V.N.Kuleshov. «Polyharmonic Analysis of Fluctuations in Oscillators». Radioengineering, N. 12, 1989.

1999 Joint Meeting EFTF - IEEE IFCS

PM AND AM NOISE IN COMMON BASE AMPLIFIER

T.I.Boldyreva and V.N.Kuleshov, Moscow Power Engineering Institute (Technical University),
14 Krasnokasarmennaya, Moscow 111250, Russia, e-mail: kuleshov@srv-vmss.mpei.ac.ru

ABSTRACT

Theoretical analysis of PM and AM noise introduced by common base (CB) bipolar transistor amplifier and comparison of CB and common emitter (CE) amplifiers are presented. The analysis is based on the approach and the models developed and used in the papers [2,4]. Power spectral densities (PSD) of PM and AM noises of CB amplifier (both wideband and 1/f type) are found as functions of a harmonic signal frequency, available power of the signal source and parameters of biasing circuits.

Comparison of PM and AM noises, introduced by CB amplifier and CE amplifier with unbypassed emitter resistor is also presented and discussed. It is offered to compare PM noises of the amplifiers taking into consideration available power of signal source that is necessary to provide the same output voltage across the same load impedance.

1. AMPLIFIER CIRCUIT AND MODEL

A circuit diagram of the amplifier under consideration is shown in Fig.1. It contains two sources of harmonic voltages with frequency ω_S and complex amplitudes U_{se} and U_{sb}. They have output impedances Z_{se} and Z_{sb}.

Supposing that $U_{se}\neq 0$ and $U_{sb}=0$ we obtain CB amplifier with feedback impedance Z_{sb} in base circuit. In this case we consider Z_{se} as a signal source impedance.

If we suppose that $U_{sb}\neq 0$ and $U_{se}=0$ we have CE amplifier with feedback impedance Z_{se} in emitter circuit. Here we consider Z_{sb} as output impedance of the signal source.

Bypassed resistors R_{B0} and R_{E0} in the circuit Fig.1 and voltage sources V_{E0} and $V_{B0}$are used to provide the necessary operating point. Load impedance Z_L and collector dc voltage source V_{C0} provide BJT operation in active region and choosing of the amplifier gain.

We use here the same BJT circuit model as in the paper [4] and the same approach to PM and AM noise calculation.

2. BASIC FORMULAE

Let us consider first CB amplifier. In this case we suppose that in the circuit shown in Fig.1 $U_{sb}=0$. Basing on the paper [4] we present PM and AM noises power spectral densities (PSD) of CB amplifier $S_{\phi,cb}(\omega;\omega_s)$ and $S_{a,cb}(\omega;\omega_s)$ by formulae

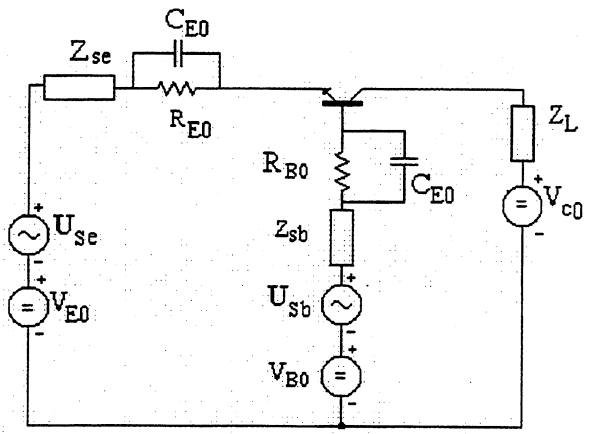


Fig.1. Circuit diagram of CB (or CE) amplifier.

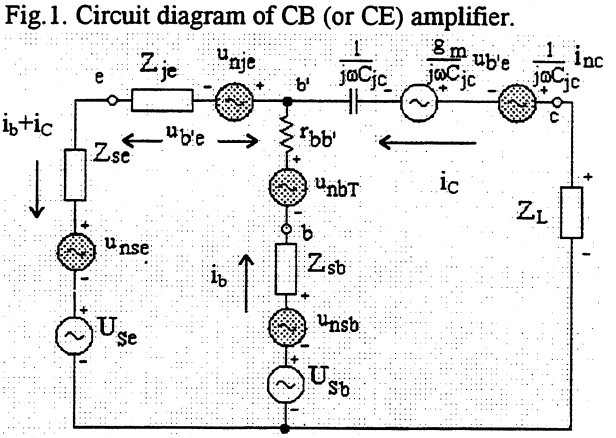


Fig. 2. Circuit model of BJT amplifier.

$$S_{\phi,cb}(\omega;\omega_s)=F_{cb}\frac{kT}{P_{av,cb}}+T^2_{\phi,cb}(\omega_s)S_\mu(\omega), \quad (1)$$

$$S_{a,cb}(\omega;\omega_s)=F_{cb}\frac{kT}{P_{av,cb}}+T^2_{a,cb}(\omega_s)S_\mu(\omega), \quad (2)$$

where F_{cb} - noise factor of the amplifier, calculated if only wideband noise sources are taken into consideration, k - Boltzman constant, T - absolute temperature, $P_{av,cb}$ - available power of the signal source in emitter circuit

$$P_{av,cb}=|U_{se}|^2/8\,\mathrm{Re}\,Z_{se}(j\omega_s), \quad (3)$$

$T_{\phi,cb}(\omega_s)$ and $T_{a,cb}(\omega_s)$ - functions of sensitivities of phase shift $\phi(t)$ and relative amplitude variations a(t) to relative 1/f fluctuations of recombination conductivity $\mu(t)$, $S_\mu(\omega)$ - PSD of $\mu(t)$ [4].

For the case of CE-amplifier, when $U_{sb}\neq 0$ but $U_{se}=0$ we have expressions similar to (1-3)

$$S_{\phi.ce}(\omega;\omega_s) = F_{ce}\frac{kT}{P_{av.ce}} + T^2_{\phi.ce}(\omega_s)S_\mu(\omega), \quad (4)$$

$$S_{a.ce}(\omega;\omega_s) = F_{ce}\frac{kT}{P_{av.ce}} + T^2_{a.ce}(\omega_s)S_\mu(\omega), \quad (5)$$

$$P_{av.ce} = \left|U_{sb}\right|^2 / 8\,\mathrm{Re}\,Z_{sb}(j\omega_s). \quad (6)$$

So the problem of PM and AM noise calculation is reduced to the one that can be solved exactly the same way as it was done for CE amplifier in the paper [4].

The problem of comparison of CB amplifier and CE amplifier is reduced to comparison of the values of noise factors, available powers and functions of sensitivities of these amplifiers under definite conditions of comparison.

3. PM AND AM FLUCTUATIONS DUE TO WIDEBAND NOISE SOURCES

Incremental model of the amplifier Fig.1 is shown in Fig.2. Equations of this model that were used to calculate voltage gains $G_{cb}(j\omega_s)$, $G_{ce}(j\omega_s)$ and $P_{av.cb}$, $P_{av.ce}$, F_{cb}, F_{ce} are given in Appendix.

But some features that are important for comparison of CB and CE amplifiers can be obtained directly from the circuit Fig.2:

- output noise of the circuit Fig.1 has the same PSD both in the case CB amplifier ($U_{se}\neq 0$; $U_{sb}=0$) and in the case CE amplifier ($U_{sb}\neq 0$; $U_{se}=0$);
- noise factors F_{cb} and F_{ce} of CB and CE amplifiers are equal if output impedances of signal sources in both circuits are equal, and feedback impedances are also equal. (Dependencies of these noise factors on ω_s for some particular cases are shown in Fig.3);
- if $Z_{sb}=Z_{se}$ the CB and CE amplifiers with equal values of U_{se} and U_{sb} have the same values of noise factors, signal sources available powers and PM and AM noises.

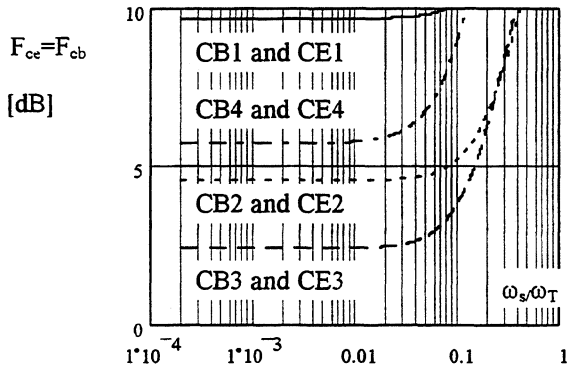


Fig.3. Dependencies of noise factors on ω_s/ω_T for CE and CB amplifiers

Bearing this in mind we consider a set of examples of CB and CE amplifiers The values of Z_{se} and Z_{sb} of these amplifiers are given in Table 1.

Table 1

Example number	$Z_{sb}=R_{sb}$ Ohm	$Z_{se}=R_{se}$, Ohm	Example number	$Z_{sb}=R_{sb}$ Ohm	$Z_{se}=R_{se}$ Ohm
CB1	5	5	CE1	5	5
CB2	5	25	CE2	25	5
CB3	5	125	CE3	125	5
CB4	5	1250	CE4	1250	5

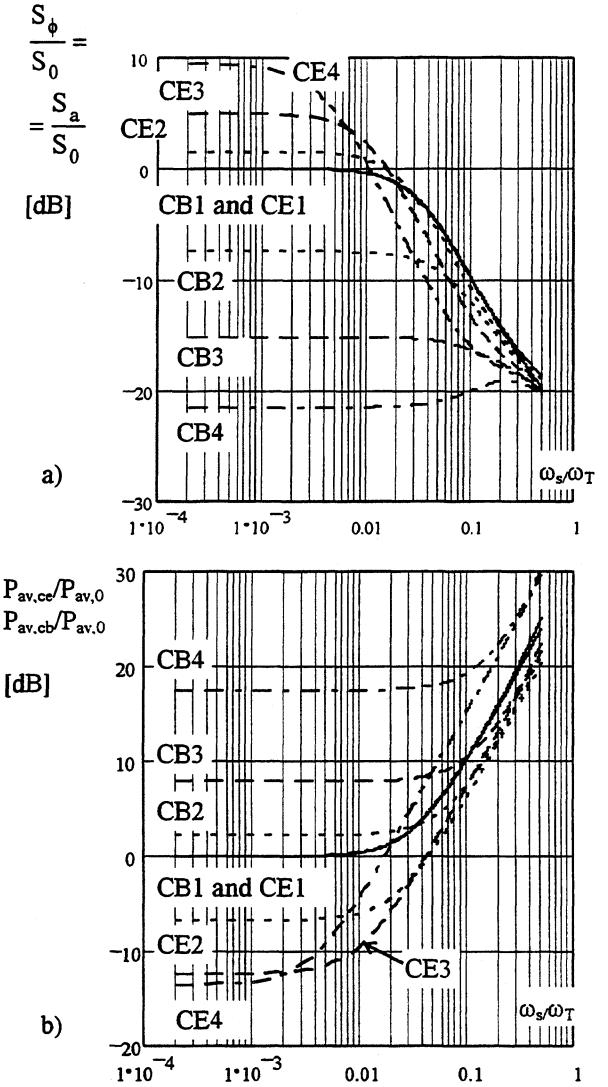


Fig.4. Normalized PSD's $(S_{\phi.cb}/S_0)=(S_{a.cb}/S_0)$ and $(S_{\phi.ce}/S_0)=(S_{a.ce}/S_0)$ - (a), and normalized signal source available powers $P_{av.cb}/P_{av.0}$ and $P_{av.ce}/P_{av.0}$ - (b) as function of ω_s/ω_T for the examples of amplifiers described in the Table 1.

(Parameters of BJT and its current in operating point are taken the same as in the paper [4]: $r_{bb'}$ = 30 Ohm, $(\omega_T/2\pi)$ = 500 MHz , β=50, C_{jc} = 4 pF, I_c=5 mA, g_m = 0.2 A/V, g_π = 0.0004 A/V, C_b = 64 pF, $Z_L = R_L$=500 Ohm). The dependencies of $S_{\phi.cb.ce} = S_{a.cb.ce}$

on ω_s/ω_T for several values of R_{se} or R_{sb} are shown in Fig.4.a. They were calculated for the case when U_L=const for all values of ω_s and R_{se} or R_{sb}, and normalized to PSD $S_0 = S_{\phi,cb,0} = S_{a,cb,0}$ of PM noise in both amplifiers with $R_{se} = R_{sb} = 5$ Ohm and $f_s/f_T = 10^{-4}$. It is easy to find, that, if we take U_L=1 V, $G_{cb}(0)$=50, U_{se0}=20mV, $P_{av,0} = U^2_{se,0}/8R_{se} = 10^{-5}$ W, F≈9.7dB and S_0 = -144.5 dB/Hz.

The curves in Fig.4,a show that one can decrease wideband PM and AM noise introduced by CB amplifier increasing R_{se}, and curves in Fig. 4,b show how signal source available power has to be changed to preserve the same value of output signal with new value of R_{se}. The available powers are normalized to $P_{av,0}$ corresponding to the case when $R_{se} = R_{sb} = 5$ Ohm and $f_s/f_T = 10^{-4}$.

To compare CE amplifier with CB amplifier we considered a set of CE amplifiers CE1...CE4 (Table1). Parameters and operating point of BJT were the same. Results of their consideration are also shown in Fig.4,a,b. One can see that at low signal frequencies increasing of R_{sb} leads to increasing of PM and AM wideband noise. At high frequencies dependencies of $S_{\phi,ce}$ on R_{sb} with ω_s=const are not monotonous. Curves in Fig.4,b show how available power has to be changed to provide a constant value of U_L when R_{sb} and ω_s/ω_T are changed.

4. PM AND AM FLUCTUATIONS DUE TO 1/F NOISE SOURCE

To investigate an influence of signal frequency and signal source impedance on 1/f PM and AM noise we calculated dependencies of functions $T^2_{\phi,cb}$ and $T^2_{a,cb}$ in equations (1), (2) on ω_s/ω_T for the same CB BJT amplifiers that were considered in the previous part of this paper. These dependencies for the cases CB1...CB4 (Table 1) with R_{B0}=0, R_{E0}=0 are shown in Fig.5 a,b. As in the case of wideband noise increasing of R_{se} leads to decreasing of 1/f noise and the rate of this decreasing is approximately the same as in the case of wideband noise sources.

Results of calculations of dependencies $T^2_{\phi,ce}$ and $T^2_{a,ce}$ on ω_s/ω_T for CE amplifiers (examples CE1...CE4 with R_{B0}=0, R_{E0}=0) are also given in Fig.5 a,b. One can see that in these cases increasing R_{sb} leads to significant increasing of 1/f PM and AM noise. It corresponds to the results obtained and discussed earlier [2,4].

The curves, similar to the ones shown in Fig.5 were also calculated for the case when R_{B0}=5 kOhm (high resistance in the base biasing circuit). The values of $T^2_{\phi,ce}$, $T^2_{a,ce}$, $T^2_{\phi,cb}$ and $T^2_{a,cb}$ corresponding to $R_{se}=R_{sb}=5$ Ohm in this case are approximately 12 dB higher then the ones shown in Fig.5. A difference between corresponding curves becomes less when R_{se} grows in CB amplifier, and R_{sb} grows in CE amplifier.

The calculations were made using the same approach as in [4]. Basic formulae are presented in the Appendix. To estimate contribution of 1/f noise sources using (1), (2), (4), (5) one can take $S_\mu(\omega) \cong 4.8 \cdot 10^{-12} I_c^{-1}/\omega$ [4].

5. CONCLUSIONS

CB amplifier built in accordance with the circuit shown in Fig.1 with U_{sb} =0, $R_{B0}=R_{E0}=0$, and signal source (SS) voltage U_{se}, SS impedance $Z_{se} = R_{se}$ equal to $Z_{sb} = R_{sb}$ has the same PM and AM noise as

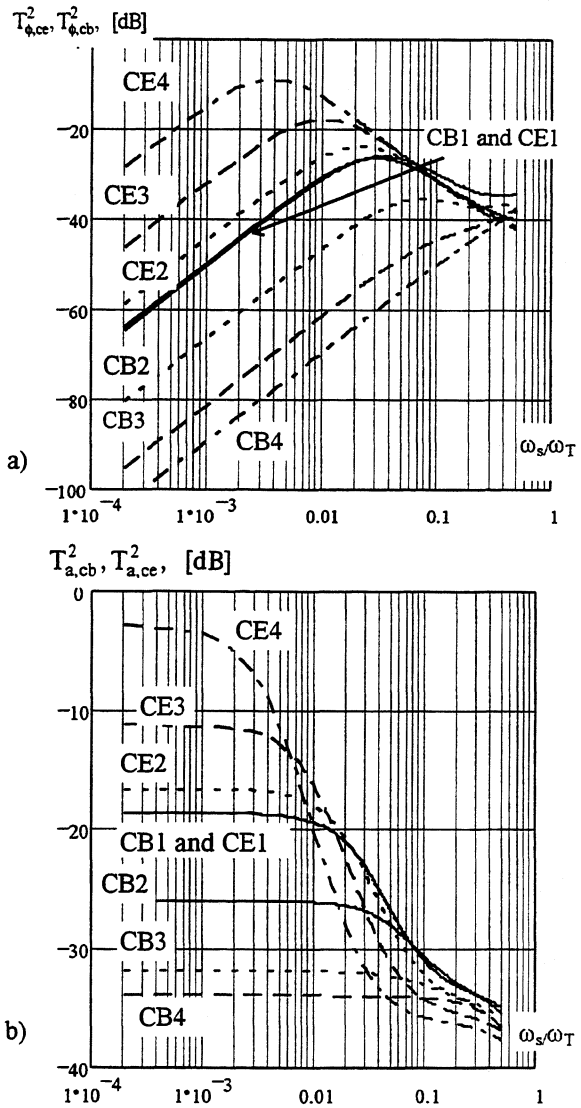


Fig.5. Dependencies of $T^2_{\phi,cb}$, $T^2_{a,ce}$ - (a) and $T^2_{a,cb}$, $T^2_{\phi,ce}$ - (b) on ω_s/ω_T for the amplifiers described in the Table1.

CE amplifier (circuit Fig.1 with U_{se} =0, R_{B0}=R_{E0}=0) and the same values of R_{se}= R_{sb}.

If one increases R_{se} in CB amplifier supporting R_{sb}=const and output voltage U_L=const, PM and AM noise at the output decreases , but one has to increase SS voltage U_{se} and available power $P_{av.cb}$. Decreasing of PM and AM noise when R_{se} grows becomes less when the signal frequency becomes higher (Fig.4,5). PM and AM noise becomes higher if one increases R_{sb}.

Using large bypassed resistor R_{BO} in base biasing circuit increases 1/f PM and AM noise. Bypassed resistor R_{E0} in emitter circuit decreases 1/f PM and AM noise but its influence is significant if R_{E0}>R_{se}.

It is interesting to notice that contributions of CB stages in CE-CB and CC-CB amplifiers that were analyzed in [4] are very different. In CE-CB amplifier the CB stage has SS with high output impedance and its contribution is small by comparison to CE stage of this circuit. In CC-CB amplifier the CB stage has SS with rather low output impedance and its contribution is almost the same as a contribution of CC stage.

CB amplifier with lowest PM and AM noise has to have Z_{sb}=0, R_{BO}=0, R_{se} close to the value providing minimal noise factor in chosen operating point of BJT, and $R_{EO}\approx(3\div5)R_{se}$.

6. APPENDIX

The equations of the incremental model shown in Fig.2 look as follows

$$(Z_{sb}+r_{bb'})i_b+u_{b'e}+Z_{se}(i_b+i_c)= \\ =U_{sb}+U_{se}+u_{nsb}+u_{nse}+u_{nbT}, \quad (A.1)$$

$$(Z_L+\frac{1}{j\omega_s C_{jc}})i_c+(-\frac{g_m}{j\omega_s C_{jc}}+1)u_{b'e}+Z_{se}(i_b+i_c)= \\ =\frac{1}{j\omega_s C_{jc}}i_{nc}+u_{nse}+U_{se}, \quad (A.2)$$

$$u_{b'e}=Z_{je}(i_b+i_c)-Z_{je}(i_{n\pi}+i_{nc}), \quad (A.3)$$

$$Z_{je}=(g_m+g_\pi+j\omega_s C_b)^{-1}, \quad (A.4)$$

$$u_L=-Z_L i_c. \quad (A.5)$$

In these equations:

$u_{nsb}, u_{nse}, u_{nbT}$ - sources of thermal noises of Z_{sb}, Z_{se} and $r_{bb'}$,

$i_{n\pi}, i_{nb}$ – transistor noise currents, that are described the same way as in [4],

g_m, g_π, C_b - parameters of the BJT hybrid pi incremental model [4].

Using them, we calculated voltage gains $G_{cb}(j\omega_s)$, $G_{ce}(j\omega_s)$, noise factors F_{cb}, F_{ce} and available powers of the SS $P_{av.ce}$, $P_{av.cb}$.

To find functions of sensitivities $T^2_{\phi,cb}(\omega_s)$ $T^2_{a,cb}(\omega_s)$ we put all wideband noises in (A.1-A.3) equal to zero, suppose that U_{sb}=0, do in equations (A.2), (A.4) replacements:

$g_\pi \to g_\pi+\mu g_\pi+\Delta g_\pi$, $g_m \to g_m+\Delta g_m$, $C_b \to C_b+\Delta C_b$, where

$\Delta g_\pi=-g_\pi A\mu$, $\Delta g_m=-g_m A\mu$, $\Delta C_b=-C_b A\mu$,

$$A=\frac{(R_{B0}+R_{E0}+r_{bb'}+R_{sb}+R_{se})(g_m/\beta)}{1+(R_{B0}+R_{E0}+r_{bb'}+R_{sb}+R_{se})(g_m/\beta)+(R_{E0}+R_{se})g_m},$$

and express voltage gain as a complex function of μ: $G_{cb}(j\omega_s, \mu)$. After that we find [4]

$$T_{cb}(\omega_s)=\left[\frac{\left(\frac{dG_{cb}(j\omega_s,\mu}{d\mu}\right)}{G_{cb}(j\omega_s,\mu)}\right]_{\mu=0}$$

and

$$T_{\phi,cb}(\omega_s)=\mathrm{Im}(T_{cb}(j\omega_s));$$

$$T_{a,cb}(\omega_s)=\mathrm{Re}(T_{cb}(j\omega_s)).$$

Function $T^2_{\phi,ce}$ and $T^2_{a,ce}$ were calculated the same way.

REFERENCES

1. F.L.Walls, E.S.Ferre-Pikal, and S.R.Jefferts. «The Origin of 1/f PM and AM Noise in Bipolar Junction Transistor Amplifiers».- presented at the IEEE Int. Freq. Contr. Symp., San Francisco, California, U.S.A., May - June, 31 - 2, 1995.
2. V.N.Kuleshov and T.I.Boldyreva. «1/f AM and FM Noise in Bipolar Transistor Amplifiers: Sources, Ways, Technique of Reduction».- presented at the IEEE Int. Freq. Contr. Symp., Orlando, Florida, U.S.A.,May 28-30, 1997.
3. F.L.Walls, E.S.Ferre-Pikal, and C.W.Nelson «Design criteria for BJT amplifiers with low 1/f AM and FM noise».- presented at the IEEE Int. Freq. Contr. Symp., San Francisco, California, U.S.A., May - June, 31 - 2, 1995.
4. T.I.Boldyreva. «Comparative Study of Phase and Amplitude Noise of CE, CE-CB and CC-CC Bipolar Transistor Amplifiers», presented at the IEEE Int. Freq. Contr. Symp, Pasadena, California, U.S.A.,May 27-29, 1998.

1999 Joint Meeting EFTF - IEEE IFCS

PM AND AM NOISE OF BJT AMPLIFIERS WITH QUARTZ CRYSTAL RESONATOR IN EMITTER CIRCUIT

T.I.Boldyreva, Moscow Power Engineering Institute (Technical University),
14 Krasnokasarmennaya, Moscow 111250, Russia, e-mail: fks@srv-vmss.mpei.ac.ru

ABSTRACT

An algorithm of bipolar junction transistor (BJT) amplifier PM and AM noise calculation, developed in the paper [1], is applied to BJT amplifiers, having common emitter (CE) and common emitter - common base (CE-CB) circuit configurations with quartz crystal resonator (QR) inserted in emitter circuits. Power spectral densities (PSD) of PM and AM noise introduced by these amplifiers when signal frequency is close to QR resonance frequency are calculated. Calculations were made for PM and AM noises caused both by wideband noise sources and by sources of 1/f noise.

A comparison of PM and AM noises of CE and CE-CB amplifiers is fulfilled and discussed.

1. INTRODUCTION

Bipolar junction transistor (BJT) amplifiers with quartz crystal resonators (QR) in emitter circuits are widely used to build quartz crystal oscillators [2,3]. Sometimes such narrow band amplifiers are used also to clean up a spectrum of harmonic oscillations. As such amplifiers introduce their own AM and PM noises it is interesting to calculate their power spectral densities and investigate dependence of these noises on circuits parameters. Some results of such calculations and investigation are presented in this paper.

2. AMPLIFIERS CIRCUITS AND PROBLEM FORMULATION

The circuits under consideration are CE and CE-CB BJT amplifiers with QR inserted in emitter circuits. They are particular cases of the circuits shown in Fig.1,a,b, if one replaces an emitter impedance Z_E by QR with inductance L_{EO} and resistance R_{EO}, inserted in parallel to QR (Fig 1,c). These circuits are selected for analysis because they were used to build low noise oscillators [2,3]. General circuits, presented in Fig. 1.a,b coincide with the circuits that were analysed in the paper [1]. But the formulae, presented in the paper [1], were derived for wideband amplifiers. An influence of the circuits elements on dependence of power spectral densities (PSD) of amplitude $S_a(\omega)$ and phase $S_\phi(\omega)$ fluctuations on Fourier frequency ω was negligible and it was not taken into consideration.

In this case we deal with highly selective feedback circuits and the formulae derived in [1] have to be generalized to use them for these kinds of circuits. As algorithms of wideband and 1/f PM and AM noise calculation are different one should obtain two kind of formulae.

To show how to use them for $S_a(\omega)$ and $S_\phi(\omega)$ calculation we'll consider two particular BJT amplifiers with typical parameters of BJT and QR and use these examples to analyze an influence of circuit parameters on PM and AM noise of these amplifiers.

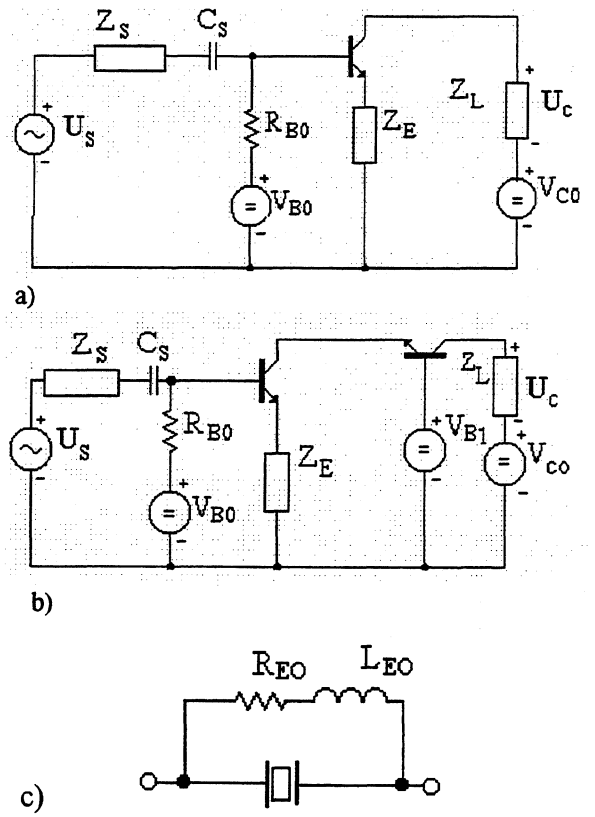


Fig.1. Circuit diagrams of CE (a), CE-CB (b) amplifiers and QR (c)

3. BJT AND QR MODELS AND THEIR PARAMETERS

To calculate gain and noise of the amplifiers we'll use hybrid π model of BJT with noise sources shown in Fig. 2. The model is similar to the one used in [1].

A circuit model of Z_E consists of a QR circuit model, an inductance L_{EO} and a resistance R_{EO}. In the QR circuit model only one series resonance branch with

elements L_q, C_q and R_q and a static capacitance C_0 are taken into consideration (Fig.3).

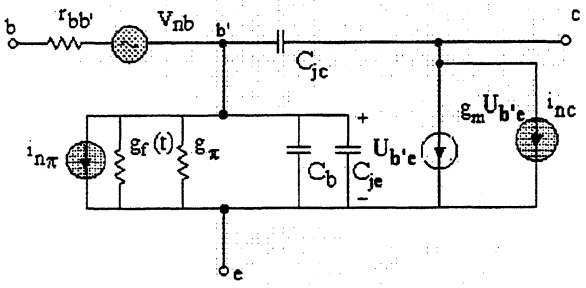


Fig.2. BJT circuit model with noise sources.

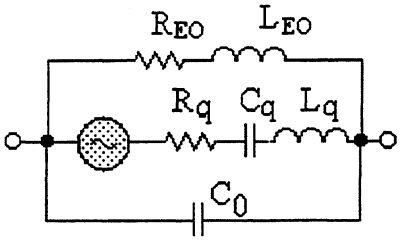


Fig.3. Model of QR with noise source.

In particular circuits that will be considered we take the next sets of QR and BJT parameters.

a) QR: $(\omega_q/2\pi) = f_q = 1/2\pi\sqrt{L_q C_q} = 60$ MHz, quality factor QR $Q_q = \sqrt{\frac{L_q}{C_q}}/R_q = 10^5$, $R_q = 25$ Ohm, $C_0 = 5$ pF. We suppose also that L_{EO} is selected so that $1/\sqrt{L_{EO}C_0} = \omega_q$, and in the basic example $R_{EO} = 25$ Ohm.

b) BJT: $(\omega_T/2\pi) = f_T = 3*10^9$ Hz, CE configuration current gain β=50, $r_{bb'}$=30 Ohm, $C_{jc} = 1$ pF, $C_{je} = 1$ pF.

Basic operating point is selected in active region of BJT operation with average collector current $I_C = 5$ mA. That is why $g_m = 0.2$ A/V, $g_\pi = 0.0004$ A/V, $C_b = 10{,}6$ pF.

4. ALGORITHMS OF PM AND AM NOISE CALCULATIONS

Let us suppose that small harmonic signal

$$v_s(t) = \mathrm{Re}[\dot{U}_s \exp(j\omega_s t)] \qquad (1)$$

is applied to the input of the BJT amplifier (BTA).

Power spectral densities (PSD) of phase $S_\phi(\omega)$ and relative amplitude $S_a(\omega)$ fluctuations, that are introduced by BTA, consist of two components

$$S_\phi(\omega) = S_{\phi 1}(\omega) + S_{\phi 2}(\omega) \qquad (2)$$

$$S_a(\omega) = S_{a1}(\omega) + S_{a2}(\omega) \qquad (3)$$

where $S_{\phi 1}(\omega)$, $S_{a1}(\omega)$ arise due to wideband noise sources (thermal, recombination and shot noises), and $S_{\phi 2}(\omega)$, $S_{a2}(\omega)$ arise due to 1/f fluctuations of BJT parameters.

To derive formulae for PM and AM noise of narrow band amplifier we write at first a formula for PSD of the output noise due to wideband noise sources in the BJT model

$$S_n(\omega) = |G(j\omega)|^2 F(\omega) 4kT(\mathrm{Re}\, Z_s(j\omega)), \qquad (4)$$

where G(jω) - transfer function of the amplifier from signal source voltage to the output voltage, F(ω) - noise factor of the BTA, $Z_s(j\omega)$ - impedance of the signal source.

Taking into consideration that

$$\dot{U}_C = G(j\omega)\dot{U}_s$$

and using results, presented in the book [4] one obtains

$$S_{\phi 1}(\omega) = S_{a1}(\omega) = \frac{Z_{\phi n}(j\omega_s + j\omega) + Z_{\phi n}(j\omega_s - j\omega)}{2|G(j\omega_s)|^2 \mathrm{Re} Z_s(j\omega_s)}\left(\frac{kT}{P_{av}}\right), \qquad (5)$$

where

$$Z_{\phi n}(j\omega_s + j\omega) = |G(j\omega_s + j\omega)|^2 F(\omega_s + \omega)\mathrm{Re} Z_s(j\omega_s + j\omega),$$

$$Z_{\phi n}(j\omega_s - j\omega) = |G(j\omega_s - j\omega)|^2 F(\omega_s - \omega)\mathrm{Re}\, Z_s(j\omega_s - j\omega),$$

$$P_{av} = U_s^2/8\,\mathrm{Re}\, Z_s(j\omega_s).$$

To calculate $S_{\phi 2}(\omega)$ and $S_{a2}(\omega)$ supposing that one knows PSD $S_\mu(\omega)$ of parameter g_π relative fluctuations $\mu_f(t)$ one has to start from symbolic transfer function (TF), $G(p,\mu_f)$ that defines

$$v_C = G(p;\mu_f)v_s,$$

where p=d/dt. After linearization of TF with respect to μ_f one obtains

$$\phi_2(t) = T_{\phi\mu}(p,\omega_s)\mu_f(t),$$

$$a_2(t) = T_{a\mu}(p,\omega_s)\mu_f(t)$$

$$S_{\phi 2}(\omega) = |T_{\phi\mu}(j\omega,\omega_s)|^2 S_\mu(\omega) \qquad (6)$$

$$S_{a2}(\omega) = |T_{a\mu}(j\omega,\omega_s)|^2 S_\mu(\omega) \qquad (7)$$

where

$$T_{\phi\mu}(j\omega,\omega_s) = \frac{1}{2j}\left[\frac{G'_\mu(j\omega + j\omega_s)}{G(j\omega_s)} - \frac{G'_\mu(j\omega - j\omega_s)}{G(-j\omega_s)}\right], \qquad (8)$$

$$T_{a\mu}(j\omega,\omega_s) = \frac{1}{2}\left[\frac{G'_\mu(j\omega + j\omega_s)}{G(j\omega_s)} + \frac{G'_\mu(j\omega - j\omega_s)}{G(-j\omega_s)}\right], \qquad (9)$$

$$G'_\mu(j\omega) = \left.\frac{\partial G(p;\mu)}{\partial\mu}\right|_{\substack{p=j\omega,\\ \mu=0}}, \qquad (10)$$

$$G(j\omega) = \left.G(p;\mu)\right|_{\substack{p=j\omega,\\ \mu=0}} \qquad (11)$$

Specific feature of these formulae is an opportunity to calculate an influence of narrow band amplifiers TF on dependencies of PSD's on Fourier frequency.

In the case of wideband amplifiers these formulae coincide with the ones presented in [1].

5. PM AND AM DUE TO WIDEBAND NOISE SOURCES

Let us consider at first results of PM and AM noise calculation, supposing that only wideband noise sources are present.

Supposing that the QR and BJT have parameters listed in part 3 of this paper, $Z_S=R_S=50$ Ohm, $R_{B0}=2,5$ kOhm and $Z_L=R_L=500$ Ohm, we calculated dependencies of voltage gain $|G(j\omega_s)|$ and the phase shift arg(G(jω$_s$)) on relative frequency offset $(\omega_s-\omega_q)/\omega_q$. They are shown in Fig.4, a,b both for CE and CE-CB amplifiers. One can notice that maximum gain frequency in CE amplifier is shifted with respect to ω_q on the value of about 1/4 of the QR bandwidth, and there is almost no such shift in CE-CB amplifier.

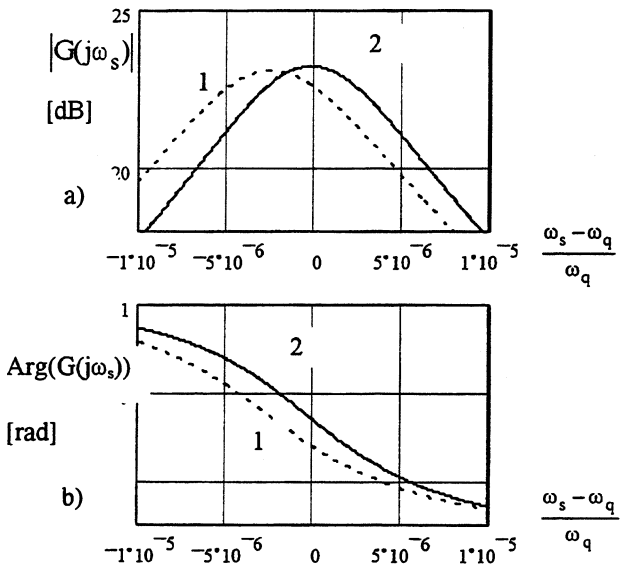


Fig.4. Dependencies of voltage gain $|G(j\omega_s)|$ (a) and the phase shift arg(G(jω$_s$)) (b) on relative frequency offset $(\omega_s-\omega_q)/\omega_q$ (1 - CE, 2 - CE-CB)

Using (5) and algorithms of the paper [1] we have calculated dependencies of AM and PM noise on Fourier frequency $f=\omega/2\pi$ for three values of offset of signal frequency with respect to maximum gain frequencies of each amplifier ($f_{max,1}$ for CE amplifier, $f_{max,2}$ for CE-CB amplifier). Results of this calculation are presented in Fig.5. One can notice that the curves

$\bar{S}(\omega)$ (where $\bar{S}(\omega)=S_\phi/\left(\frac{kT}{P_{av}}\right)=S_a/\left(\frac{kT}{P_{av}}\right)$)

corresponding to the same values $(f_s-f_{max,i})/f_{max,i}$ (i=1,2) in CE and CE-CB amplifiers are almost similar. When the offset increases the band of AM and PM noise becomes wider, and when $(f_s-f_{max,i})/f_{max,i}$ becomes more than $\Delta f_q/2Q_q$, a raising in curves $\bar{S}(\omega)$ appears and increases.

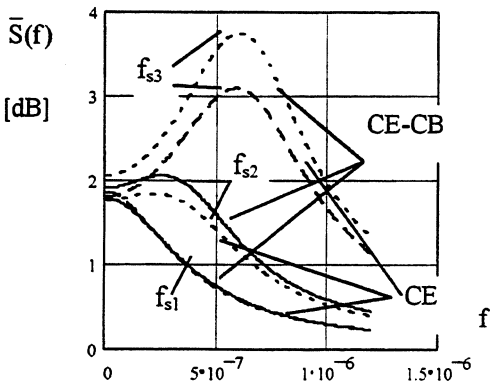


Fig.5. Dependencies of AM and PM wideband noise PSD's on Fourier frequencies for $f_{s1}=f_{max,i}$ $f_{s2}=f_{max,i}(1+0.5\Delta f_q)$, $f_{s3}=f_{max,i}(1+\Delta f_q)$.

6. PM AND AM DUE TO 1/F NOISE SOURCE

We consider here an influence of only 1/f fluctuations of recombination conductivity $g_f(t)$ (Fig.2) on amplitude and phase of output oscillations. In Fig.6 dependencies of $|T_{a\mu}(0,\omega_s)|^2$ and $|T_{\phi\mu}(0,\omega_s)|^2$ on relative signal frequency offset $(\omega_s-\omega_q)/\omega_q$ are shown. They describe an influence of very slow (quasistatic) fluctuations $\mu_f(t)$ on amplitude and phase shift of output oscillations.

One can see from Fig.6 that CE amplifier introduces higher AM noise in the point of maximum gain. But the difference is not very big (about 1,5 dB).

As to the curves shown in Fig.6,b, one can notice deep and narrow minima in dependencies of $|T_{\phi\mu}(0,\omega_s)|^2$ on ω_s. They arise due to the same compensation effects that were noticed in wideband amplifiers and discussed in [1].

It is interesting to find dependencies of transform functions $|T_{\phi\mu}(\omega,\omega_s)|^2$ and $|T_{a\mu}(\omega,\omega_s)|^2$ on Fourier frequencies when ω_s is in the vicinities of the compensation minimum. These dependencies the shown in Fig.7. In this figure curves «1» correspond to signal frequencies $\omega_{s,min1}$ and $\omega_{s,min2}$ of minimal phase shift sensitivities of CE and CE-CB amplifiers to slow $\mu(t)$ variations.

From Fig.6,b one can find that $(\omega_{s,min1}-\omega_q)/\omega_q=-3\cdot10^{-6}$ (CE case) and $(\omega_{s,min2}-\omega_q)/\omega_q=-0{,}22\cdot10^{-6}$ (CE-CB case). Curves «2» correspond to the offsets from these points $(\omega_{sl}^{(2)}-\omega_{s,min1})/\omega_{s,min1}=(\omega_{s2}^{(2)}-\omega_{s,min2})/\omega_{s,min2}=0{,}41\cdot10^{-6}$

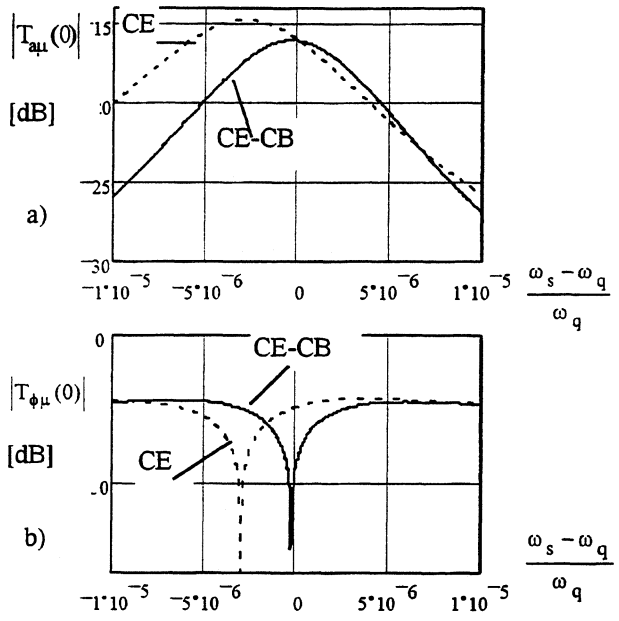


Fig.6. Functions $|T_{a\mu}(0,\omega_s)|^2$ - (a), $|T_{\phi\mu}(0,\omega_s)|^2$ - (b).

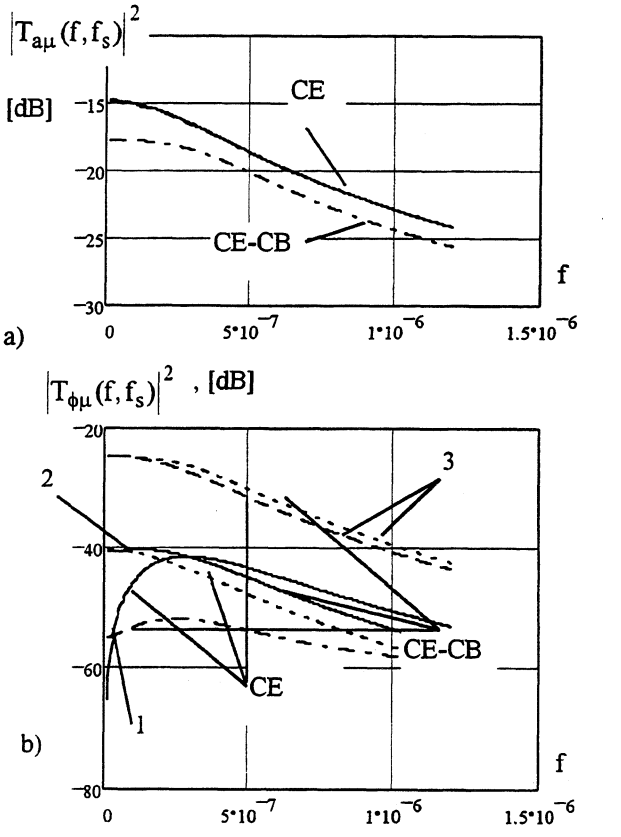


Fig. 7. Dependencies of $|T_{a\mu}(\omega,\omega_s)|^2$ (a), and $|T_{\phi\mu}(\omega,\omega_s)|^2$ (b) on Fourier frequency f with signal frequencies ω_s having small offsets from the point of minimal phase shift sensitivity to $\mu(t)$.

and curves «3» correspond to offsets $(\omega_{s1}^{(3)}-\omega_{s.min1})/\omega_{s.min1}=(\omega_{s2}^{(3)}-\omega_{s.min2})/\omega_{s.min2}=2{,}8\cdot10^{-6}$

Functions $|T_{a\mu}(\omega,\omega_s)|^2$ are not sensitive to small offsets of ω_s. But functions $|T_{\phi\mu}(\omega,\omega_s)|^2$ are very sensitive to these offsets.

One can notice that if a signal frequency is equal to ω_q, an influence of $\mu(t)$ on phase shift fluctuations in CE-CB amplifier is significantly less than in CE amplifier. So if contribution of this source of 1/f fluctuations is significant, it is possible to use compensation effect for 1/f PM noise reduction.

7. CONCLUSIONS

An approach to calculation of PM and AM noise in linear BJT amplifiers, developed in [1], is generalized and applied to the amplifier with QR in emitter circuit. Only one source of 1/f noise is considered, but using the same approach one can calculate and investigate an influence of other sources of such noise (BJT capacitances fluctuations, fluctuations of QR parameters).

8. ACKNOWLEDGEMENTS

The author would like to thank prof. V.Kuleshov for valuable discussions of problems connected with this paper.

REFERENCES

1. T.I.Boldyreva. «Comparative Study of Phase and Amplitude Noise of CE, CE-CB and CC-CB Bipolar Transistor Amplifiers», presented at the IEEE Int. Freq. Contr. Symp, Pasadena, California, U.S.A.,May 27-29, 1998.
2. M.M.Driscoll, «Two Stage Self-Limiting Series Mode Type Quartz-Crystal Oscillator Exhibiting Improved Short-Term Frequency Stability», IEEE Trans. on Instr. and Measurement, Vol. IM-22, N.2, pp.130-138, 1973.
3. D.J.Healey III, «Low Noise UHF Frequency Source». In Proceedings of 27-th Annual Symposium on Frequency Control, U.S.A., 1973, pp.170-179,
4. V.Zalud, V.N.Kuleshov. Noise in Semiconductor Devices, Soviet Radio, Moscow, 1977, (In Russian).

1999 Joint Meeting EFTF - IEEE IFCS

QUANTUM 1/F QUARTZ RESONATOR THEORY VERSUS EXPERIMENT

Peter H. Handel

Department of Physics and Astronomy, Univ. of Missouri St. Louis, St. Louis MO 63121

ABSTRACT

The quantum 1/f theory, as applied to bulk acoustic wave (BAW) and surface acoustic wave (SAW) resonators has recently been generalized to the case of an arbitrary coherence length of the phonons, which may be large or small compared with the size of the resonator. The theory has further been extended to include defect scattering along with the phonon scattering case exclusively considered earlier. The present paper compares the new theoretical results with experimental results obtained by F.L. Walls et al., T. Parker et al., J.R. Vig et al., as well as by other research groups.

1. INTRODUCTION

The quantum 1/f theory, as applied to bulk acoustic wave (BAW) and surface acoustic wave (SAW) resonators, is compared with experimental results obtained by F.L. Walls et al., T. Parker et al., J.R. Vig et al., as well as other research groups. The theory has recently been generalized to the case of an arbitrary coherence length of the phonons, which may be large or small compared with the size of the resonator. The theory has further been extended to include defect scattering along with the phonon scattering case exclusively considered earlier.

Our theoretical results [1], [2] can be written in the form

$$S(f) = \beta' V/fQ^4, \text{ for } V \le \varepsilon, \qquad (1)$$

and

$$S(f) = \beta' \varepsilon^6/fVQ^4, \text{ for } V \ge \varepsilon, \qquad (2)$$

where, with $<\omega>=10^8$/s, with $n=kT/\hbar <\omega>$, T=300K and $kT = 4 \cdot 10^{-21}$ J,

$$\beta' = (N/V)\alpha\hbar <\omega>/12n\pi g^2 mc^2$$

$$=10^{22}(1/137)(10^{-27}10^8)^2/12kT\pi 10^{-27}\, 9\, 10^{20} = 1. \quad (3)$$

For the case of defect scattering, a two-phonon process takes place. A phonon from the main resonator mode scatters on a defect and a phonon of comparable frequency emerges into another mode with much smaller phonon occupation number $n_\omega = kT/\hbar\ \omega$. In this case we have to replace $<\omega>$ by ω and $n_{<\omega>}$ with n_ω, which gives a β-value which is $(<\omega>/\omega)^2$ smaller, i.e. 10^4-10^6 times smaller. In general, therefore, writing $\Gamma = \Gamma' + \Gamma''$, we obtain for the combined phonon and defect scattering case, in general,

$$\beta = \beta'[\Gamma'^2 + (<\omega>/\omega)^2\Gamma''^2]/\Gamma^2. \qquad (4)$$

Although the defect scattering term is small at room temperature, it may become dominant at low temperatures, when the phonon scattering rate Γ' becomes much smaller than the defect scattering rate Γ''.

2. COMPARISON WITH THE EXPERIMENTAL DATA

The form of Eqs. (1)-(4) shows that the level of 1/f frequency noise depends not only as Q^{-4} as previously proposed for quartz [3], but is a non-monotonous function of the volume of the active region. The noise first increases with volume, then after reaching a maximum of the order of the phonon-coherence volume, it decreases with size. Fig. 1 illustrates his dependence on the active resonator volume V.

For quartz, this theory explains the data of Gagnepain who varied the Q-factor with temperature in the same quartz resonator, and of Gagnepain and Uebersfeld who varied Q. It quantitatively fits the data of Fere Pikal, Walls, et.al. (Fig. 2) who considered several quartz resonators which differ in volume and frequency, and the data of Parker for SAW resonators (Table I and Fig. 2), with their relatively low Q values. Indeed, according to [4], the median value of the PM noise L(10 Hz) in dBc/Hz for 12 unswept quartz resonators is -103.1, -101.6 and -97.7 for small, medium and large electrodes respectively, in reasonable agreement with the proportionality with V, which requires a 3 dB difference between the groups with large and small electrodes. The electrode diameters of the 3 size groups were 2.16 mm, 3.05 mm, and 4.32 mm.

0-7803-5400-1/99/ $ 10.00

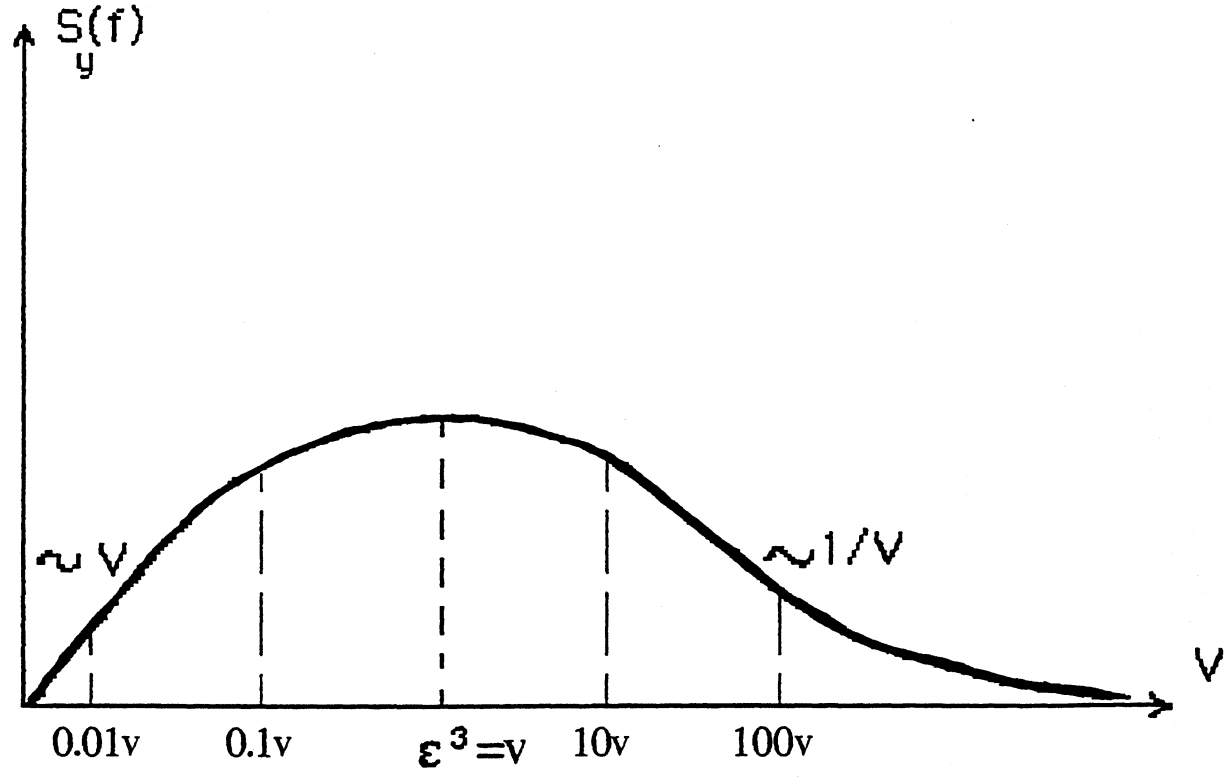


Fig. 1: Dependence of the spectral density of fractional frequency fluctuations on the the active quartz volume V.

TABLE 1
SAW
$\beta=120$, $\varepsilon=316\mu$, $\varepsilon^2=10^{-3}cm^2$

Design Identity	Quality factor Q_u	Eff. averaging area A_a cm^2	Eff. averaging volume V	Coher. βVQ^{-4}	Incoher. $(\varepsilon^2/A)^2\beta VQ^{-4}$	Experim. S(1Hz) Parker/Andres	n
A	21,800	$4.09\cdot10^{-3}$	$2.8\cdot10^{-6}$	$1.44\cdot10^{-21}$	$8.62\cdot10^{-23}$	$10\cdot10^{-23}$	4.09
B	21,800	$6.93\cdot10^{-3}$	$4.9\cdot10^{-6}$	$2.60\cdot10^{-21}$	$5.41\cdot10^{-23}$	$7.0\cdot10^{-23}$	6.93
C	14,500	$1.39\cdot10^{-2}$	$10\cdot10^{-6}$	$7.44\cdot10^{-21}$	$3.85\cdot10^{-23}$	$5.0\cdot10^{-23}$	13.9
D	22,850	$1.26\cdot10^{-2}$	$8.82\cdot10^{-6}$	$6.6\cdot10^{-21}$	$4.16\cdot10^{-23}$	$4.0\cdot10^{-23}$	12.6
E	20,675	$2.53\cdot10^{-2}$	$17.7\cdot10^{-6}$	$1.16\cdot10^{-20}$	$1.81\cdot10^{-23}$	$2.6\cdot10^{-23}$	25.3
F	20,900	$6.23\cdot10^{-2}$	$44.3\cdot10^{-6}$	$2.78\cdot10^{-20}$	$7.17\cdot10^{-24}$	$2.0\cdot10^{-23}$	62.3

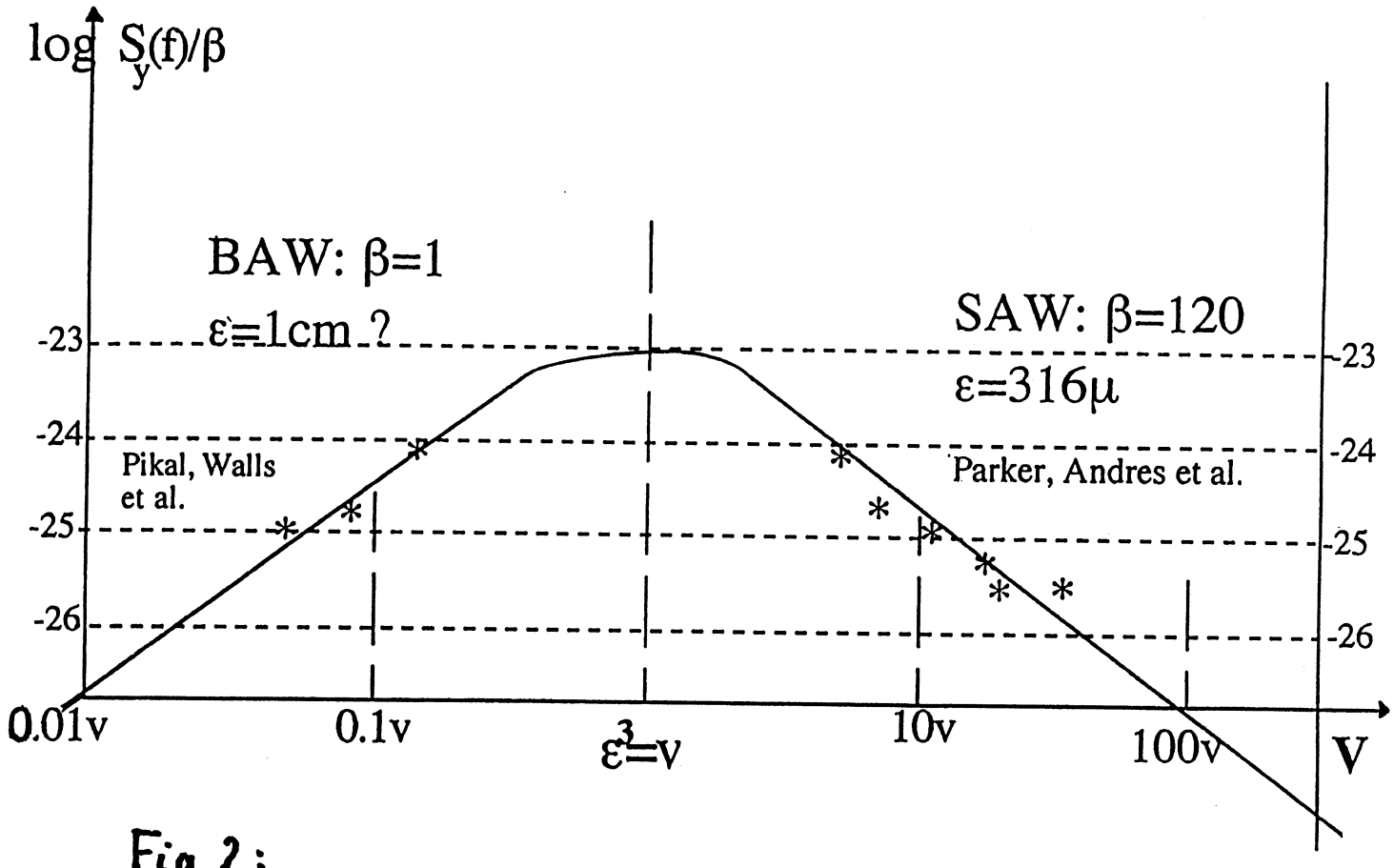


Fig. 2 :

Quantum 1/f noise in the frequency of BAW and SAW quartz resonators as a function of volume at 1Hz, in terms of the phonon-interaction coherence volume $v=\varepsilon^3$ and of the quantum 1/f coefficient β.

The volumes were therefore approximately proportional with the numbers 4.67, 9.30 and 18.66, and therefore with the measured median values. However, the remaining scatter present in the data [4], [5], is analyzed in terms of the defect contributions and coherence corrections given by Eqs. (2)-(4). The theory also provides the basis for predicting from first principles, without adjustable parameters, how to improve the 1/f level of resonators, beyond just improving the Q-factor, which has been known for many years, and which has been related [2] to fluctuations in the dissipation. Since the 1/f noise level depends on the active volume, in the coherent regime one should use the lowest overtone and smallest diameter consistent with other circuit parameters. In the incoherent (low Q) case the opposite should be considered.

3. DISCUSSION

As we see from Table I, there is a number n of between 4 and 62 phonon-interaction coherence regions in the SAW devices studied by Parker and Andres [5]. The quantum 1/f constant β is about 120 times larger in SAW devices. In Fig. 2 we are able to superpose BAW and SAW device data on the same graph by dividing the spectral density of frequency fluctuations by β. In addition, we had to express the volume in units of the coherence volume ε^3 which is much larger in BAW resonators.

REFERENCES

[1] F.L. Walls et al., Proc. 46th Frequency Control Symposium, pp. 327-333, 1992.

[2] P.H. Handel, Proc. 51th Frequency Control Sympos., Orlando, FL, 1997.

[3] P.H. Handel, Solid State Electronics 22, p. 875 (1979).

[4] E.S. Ferre-Pikal et al., Proc. 50th Freq. Control Sympos., 1996, pp. 844-851.

[5] T.E. Parker and D. Andres, Proc. 48th Freq. Contr. Symp., pp. 530-538, 1994.

DEFINITION OF PHASE SPECTRAL DENSITY OF A CRYSTAL RESONATOR THROUGH ITS EQUIVALENT PARAMETER 1/f NOISES

Yuriy S. Shmaliy

Kharkiv Military University and
"Sichron" Center, 4 Skripnika Street, Kharkiv, 310057, UKRAINE, sich@kharkov.ua

ABSTRACT

This paper addresses to the crystal resonator phase and amplitude spectral density definition through its parameter 1/f flicker noises. The approach is based on the Leeson's oscillator open loop model allowing detailed consideration of the both Butler and Colpitts modes of operation with assumption of a resonator inter noise correlation. Examples are given as an illustration of a good agreement of the data with theoretical curves.

1. INTRODUCTION

It has been shown [1] that instability transformation coefficients of crystal resonator equivalent parameters into phase and module of its complex impedance may be effectively used for resonator phase and amplitude spectral density evaluations in a wide range of offset from the carrier if to use the Lesson open loop model of oscillator [2].

In the early works, Wainright *et. al.* [3] and Gagnepain [4] showed that resonance frequency of a quartz crystal resonator within its half bandwidth exhibits 1/f flicker fluctuations. Beyond this range, a resonator noise has $1/f^2$ [3] or $1/f^3$ [4] character due to filtering effect of a resonator. But few years later, Gagnepain et al. [5] noted that the mechanisms contributing to the generation of 1/f noise in crystal resonator were not yet well understood.

In this report, we consider basic equivalent circuit of a crystal resonator with noisy equivalent parameters. We present relatively simple formulas for resonator amplitude and phase spectral densities within and beyond an inter resonance gap. All the coefficients are found out as the functions of a resonator inter resonance operation frequency, Q-factor, and the factor $\kappa = R_1 C_0 / \sqrt{L_1 C_1}$.

2. FLUCTUATION MODEL OF A RESONATOR

Following by [1], we consider equivalent circuit of a crystal resonator as consecutive connection of a motional inductance, losses and capacity with their deterministic parts $\overline{L_1}$, $\overline{R_1}$, and $\overline{C_1}$, and small added fluctuations $\Delta L_1(t)$, $\Delta R_1(t)$, and $\Delta C_1(t)$ respectively [6]. In parallel to this circuit we connect a static capacity with deterministic and random parts $\overline{C_0}$ and $\Delta C_0(t)$ respectively. Further, we transform this circuit into connection of real and imaginary parts $r(t)=\bar{r}+\Delta r(t)$ and $x(t)=\bar{x}+\Delta x(t)$, and bring them to noisy modulus $z(t)=\bar{z}+\Delta z(t)$ and phase $\varphi(t)=\bar{\varphi}+\Delta\varphi(t)$ of a resonator. It yields

$$r = \overline{R_1}\frac{1}{(1-2q\kappa)^2+\kappa^2}, \qquad (1)$$

$$x = \overline{R_1}\frac{2q(1-2q\kappa)-\kappa}{(1-2q\kappa)^2+\kappa^2}, \qquad (2)$$

where $q(t)=\Delta\omega(t)/\Pi(t)$, $\Delta\omega(t)=\omega(t)-\bar{\omega}_1$, $\bar{\omega}_1 = 1/\sqrt{\overline{L_1}\overline{C_1}}$, $\overline{\Pi}=\overline{R_1}/\overline{L_1}=\Delta f_{BW}/2\pi$, and $\kappa = R_1\omega_1 C_0$ are a resonator bandwidth and κ-factor. Relationships (1) and (2) have a common character allowing using them for definition of fluctuation transformation coefficients for different types of crystal units, and which investigation leads to understanding of noise conversion effects in a crystal resonator.

3. TRANSFORMATION COEFFICIENTS OF FLUCTUATIONS

Based on above functions (1) and (2), we come to transformation coefficients of fluctuations of motional and static parameters of resonator into its real and imaginary parts and, further, to module and phase of the resonator complex resistance. Let us write relative departures of resonator parameters caused by fluctuations $\delta_L(t)$, $\delta_C(t)$, $\delta_R(t)$, and $\delta_{C0}(t)$ as follows

$$\delta_{Lr,Cr,Rr,0r} = \frac{r_q(\delta_L,\delta_C,\delta_R,\delta_{C0})-\bar{r}_q}{\overline{R_1}}, \qquad (3)$$

$$\delta_{Lx,Cx,Rx,0x} = \frac{x_q(\delta_L,\delta_C,\delta_R,\delta_{C0})-\bar{x}_q}{\overline{R_1}}, \qquad (4)$$

Relationships (3) and (4) allow definition of transformation coefficients of $\Delta L_1(t)$, $\Delta R_1(t)$, $\Delta C_1(t)$, and $\Delta C_0(t)$ into $\Delta r(t)$ and $\Delta x(t)$ by $K_{yr,yx}(q)=\delta_{yr,yx}/\delta_y$, where in subscripts $Y = L_1, C_1, R_1, C_0$. Correspondingly, a resonator impedance modulus Z may be written in the form $Z(t)=\overline{Z}+\Delta Z(t)=\overline{Z}(1+\delta_z\overline{R_1}/\overline{Z})$, where

$$\delta_z \equiv \frac{\bar{x}\delta_x+\bar{r}\delta_r}{\sqrt{\bar{x}^2+\bar{r}^2}} = \frac{[2q(1-2q\bar{\kappa})-\bar{\kappa}]\delta_x+\delta_r}{\sqrt{1+[2q(1-2q\bar{\kappa})-\bar{\kappa}]^2}}, \qquad (5)$$

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

where $\delta_x = \Delta x/\bar{x}$ and $\delta_r = \Delta r/\bar{r}$. It follows from (5) that at frequencies q_s and q_p where $x = 0$, we get equal instabilities δ_z and δ_r, and at frequencies corresponding to $x = r$ there is $\delta_z = 0{,}5\ (\delta_x + \delta_r)$. Defining phase drift of a resonator complex resistance due to noises as $\Delta\varphi$ = arctg$X$ – arctg$Y$, where $X = \dfrac{\bar{x} + \Delta x}{\bar{r} + \Delta r}$ and $Y = \dfrac{\bar{x}}{\bar{r}}$, in view of $XY > -1$ we write

$$\Delta\varphi \cong \left[(1 - 2q\bar{\kappa})^2 + \bar{\kappa}^2\right]\frac{\delta_x - \left[2q(1 - 2q\bar{\kappa}) - \bar{\kappa}\right]\delta_r}{1 + \left[2q(1 - 2q\bar{\kappa}) - \bar{\kappa}\right]^2}, \quad (6)$$

An analysis of (6) shows that the phase shift with $q = q_s$ and $q = q_p$ is $\Delta\varphi = \delta_x \bar{R}_1/\bar{r}$, since $\bar{x}(q_s, q_p) = 0$. It follows also that certain frequency q_c is possible in a gap of $f_s < f_c < f_p$ at which $r\delta_x - x\delta_r = 0$, and phase shifts caused by parameter drifts of a resonator equivalent circuit are theoretically absent. Coefficients $K_{R\bar{z},L\bar{z},C\bar{z},O\bar{z}}(f)$ and $K_{R\varphi,L\varphi,C\varphi,O\varphi}(f)$ of $\Delta L_1(t)$, $\Delta R_1(t)$, $\Delta C_1(t)$, and $\Delta C_0(t)$ transformation into ΔZ and $\Delta\varphi$ are found out and presented in [7].

4. "LOW-NOISY" OPERATION FREQUENCY

We pointed out above that there is the frequency within an inter resonance gap at which an influence of both ΔL_1 and ΔC_1 to a resonator phase noise tends to zero. This frequency is found out from the identity $K_{L\varphi,C\varphi} = 0$, and is

$$f_C = f_1\left(1 + \frac{1}{4\bar{Q}\bar{\kappa}}\right). \quad (7)$$

As it had been expected, the received value (7) corresponds to maximum of a resonator phase response $\varphi(q)$. Hence, it is naturally to consider it as "low-noisy" and preferable for oscillator mode of operation.

Let us make a comparative theoretical analysis of resonator noises, based on above obtained coefficients and keeping in mind that vibrations in a resonator in the Butler mode are kept around frequency f_s, and that in the Colpitts one are in an inter resonance gap $f_s < f_C < f_p$ respectively. In slightly loaded Butler mode provided in a resonator with small R_1 and C_0 that usually take place in practice ($f = f_s$ and $\bar{\kappa} << 0.5$), and with small offset from the carrier ($F < \bar{\Pi}/4\pi$) we may take that all the coefficients are constant and get $K_{RZ} \cong 1$, $K_{0Z} \cong 0$, $K_{LZ,CZ} \cong -2\bar{\kappa}\bar{Q}$, $K_{R\varphi} \cong 2\bar{\kappa}$, $K_{L\varphi,C\varphi} \cong \bar{Q}$, and $K_{0\varphi} \cong -\bar{\kappa}$. Here, the noise of R_1 is transferred into Z and φ_q with corresponding coefficients of 1 and $2\bar{\kappa}$ at frequency f_s, and of 0 and $4\bar{\kappa}$ at frequency f_c. It speaks in favor of the Colpitts mode of operation. Noises of L_1 and C_1 are transformed respectively in accordance with coefficients $-2\bar{\kappa}\bar{Q}$, $\bar{Q}$ and $-4\bar{\kappa}\bar{Q}$, 0. This also bears witness to preference of the Colpitts mode, as theoretically $K_{L\varphi,Lc}(q_c) = 0$. Double increase of amplitude fluctuations in this case does not give any appreciable negative effect. Just to opposite, it is the other way with transformation of C_0 noise. Here correspondent values are $0, -\bar{\kappa}$ and $1/\bar{\kappa}, -4\bar{\kappa}$. As follows, noise of C_0 is transferred into modulus Z with a transformation coefficient greater by several orders, and into phase φ_z – greater by four times than in the Butler mode. This fact is a comparative disadvantage of the Colpitts mode though it is mostly of a theoretical significance according the real values of $\bar{\kappa}$. At the bifurcation point $\bar{\kappa} = 0{,}5$ the Colpitts mode gives rather better results because all another coefficients, excepting $K_{0\varphi} = -1$, are equal to zero. On the contrary, in the Butler mode only $K_{R\bar{z}} = K_{0\bar{z}} = K_{L\varphi,C\varphi} = 0$ and another coefficients take the values $K_{L\bar{z},C\bar{z}} = -2Q$ and $K_{R\varphi} = 1$, $K_{0\varphi} = -1$.

5. SPECTRAL DENSITIES

For provision of the comparative evaluations we consider a resonator spectral density of phase fluctuations $S_\varphi(F,f)$ at frequencies F tuning-out from the carrier f. Supposing only a linear mode of a resonator operation, we find out an amplitude $\delta_Z(f, t)$ and phase $\delta_\varphi(f, t)$ noises as superposition of that of resonator parameters through the obtained transformation coefficients.

Using covariances $R_{\delta Z}(\tau) = E\{\delta_Z(f,t)\delta_Z(f,t+\tau)\}$ and $R_{\delta\varphi}(\tau) = E\{\delta_\varphi(f,t)\delta_\varphi(f,t+\tau)\}$, where τ is a correlation time we come through the Wiener-Khintchine transform to the spectral densities. Further, since interesting the maximal possible spectrum values, we use the inequality $\left|S_{\xi\eta}(F)\right|^2 \le S_\xi(F)S_\eta(F)$ and write

$$S_\varphi(F) \le K_{R\varphi}^2 S_R + K_{C\varphi,L\varphi}^2\left[S_C + S_L + 2\left(S_C S_L\right)^{0.5}\right] + K_{0\varphi}^2 S_0 +$$

$$2\left[K_{L\varphi,C\varphi}\left(K_{R\varphi}S_R^{0.5} + K_{0\varphi}S_0^{0.5}\right)\left(S_L^{0.5} + S_C^{0.5}\right) + K_{R\varphi}K_{0\varphi}\left(S_0 S_R\right)^{0.5}\right], \quad (8)$$

where $S_R(F)$, $S_L(F)$, $S_C(F)$, and $S_0(F)$ are spectral densities of R_1, L_1, C_1, and C_0 noises, $S_{LC}(F)$, $S_{RL}(F)$, $S_{RC}(F)$, $S_{0L}(F)$, $S_{0C}(F)$, and $S_{0R}(F)$ are complex joint spectral densities of resonator parameters fluctuations.

In view of (8), we may obtain full and zero inter noises correlation multiplying all the joint spectral densities in formulas by coefficient $k = 1$ and $k = 0$ respectively. Based on this, we consider the Butler and Colpitts modes within a half bandwidth of resonator, substituting transformation coefficients into (8).

6. COMPARISON OF MEASURED DATA WITH PREDICTED CURVES

Here we apply (8) for evaluation of power spectral densities of resonator amplitude and phase noises. As the basic practical results, we will take the measured data, obtained in [8,9]. Some another results are given in [7]. Based upon Leeson's model [2], we write a spectral density in a form of oscillator open loop $S'_{Z,\varphi}(F) = S_{Z,\varphi}(-F) + S_{Z,\varphi}(+F)$ using the formula

$$\pounds_\varphi(F) = 10 \cdot \log\left(\frac{S'_\varphi(F)}{2}\right) \text{[dBc/Hz]}, \qquad (9)$$

that is a single-sideband noise-to-carrier ratio. It follows from (10) that $S'_\varphi(F) = 2 \cdot 10^{\pounds(F)/10}$. Since we consider conversion only 1/f noises, we account a white phase noise impact as an additional additive floor $S_{\varphi w}$ or $\pounds_{\varphi w}$ and use instead (9) the relationship

$$\pounds_\varphi(F) = 10 \cdot \log\left(\frac{S'_\varphi(F) + S_{\varphi w}}{2}\right), \qquad (10)$$

Let us compare above obtained curves with the measured data given in literature. The methodology is following. We take that all the resonator deterministic parameters are known. Then we take the measured values of phase and amplitude spectral densities only at the offset frequency 10 Hz or 100 Hz within a half bandwidth of a resonator and the noise floors $S_{\varphi w}$ and S_{Zw} from [8,9]. Further we put dawn $S_L(F) \cong S_C(F) \cong S'_y(F)$ and $S_R(F) \cong S'_z(F)$, and calculate spectral densities obtained by (8) comparing them with correspondent measured data. In the case of unknown resonator parameter, we use its typical value that holds true for the correspondent frequency range.

6.1 VHF crystal resonator of 100 MHz

Driscoll had carried out the measurements while investigation of very high frequency crystal resonators (Fig.3a, [8]). We get from the data the following magnitudes $f_S = 100 \cdot 10^6 Hz$, $Q = f_S / \Delta f_{BW} \cong 1.4 \cdot 10^5$, and $\Delta f_{BW} \cong 700 Hz$. Also, at $F = 100 Hz$ we get $S'_y(10) \cong 3.96 \cdot 10^{-24}$ and being based on above methodology put dawn $S'_y(F) = S_C(F) = S_L(F) \cong 4.0 \cdot 10^{-22} F^{-1}$. As well as in the case of [9], we take $S_{C0} \cong 5 \cdot 10^{-14} F^{-1}$ and R_1 = 36 Ohm. In addition, we initially take S_R = 0 and $\kappa = 0.025$ that corresponds to $C_0 \cong 1.1 \cdot 10^{-12} F$, and $\pounds_{\varphi w} \cong$ -163dBc/Hz . With S_R = 0 we have no good correspondence of the data and accurate prediction curve, and only with $S_R \cong 17 \cdot 10^{-12} F^{-1}$ those fit one to another (Fig.1, curve c). Let us note that Driscoll considered this resonator as noisy. Fig.1 shows also the influence of the resonator static capacity C_0 value and its flicker noise spectral density S_{c0} to the prediction curves for the Butler mode (b,c,d). As may see, we received quite good behavior of the model, excepting offsets with the frequencies less than 30 Hz. Here, probably, a temperature is a main source of fluctuations.

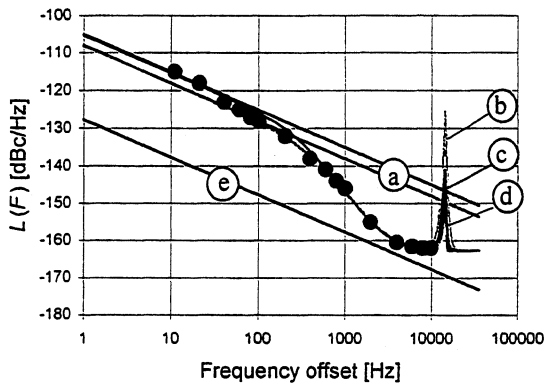


Figure 1.

6.2 VHF crystal resonator of 125 MHz

Curtis *et. al.* in [9, Fig.12] brought the data that show some ranges of spectral density. With F < 50 Hz there is $1/f^2$ noise due to a temperature, with 50 < F < 800 Hz - a flicker noise of 1/f type, with 800 < F < 2000 we see the floor of the noise $1/f^0$, with 2000 < F < 20000 more than half bandwidth of resonator there is a flicker noise of $1/f^3$ type, and with 20000 < F - a white noise floor $1/f^0$. We get from the data $f_S = 125 \cdot 10^6 Hz$, $\Delta f_{BW} \cong 4000 Hz$, and $Q \cong 3.13 \cdot 10^4$. At $F = 100 Hz$ we get $\pounds_\varphi(100) \cong$ -140dBc/Hz and take $S'_y(F) = S_C(F) = S_L(F) \cong 5.1 \cdot 10^{-22} F^{-1}$. As well as in another cases, we take $S_{C0} \cong 5 \cdot 10^{-14} F^{-1}$, R_1 = 36 Ohm, and $\kappa = 0.03$ that corresponds to $C_0 \cong 1.06 \cdot 10^{-12} F$, $\pounds_{\varphi w} \cong$ -163dBc/Hz, and, initially, S_R = 0. With zero level of noise S_R = 0 we get big disparity between the data and prediction curve (Fig.2, curve d). On the contrary, with $S_R \cong 5 \cdot 10^{-12} F^{-1}$ the prediction fit the data quite good in all ranges of offset frequencies, excepting that of the temperature influence (Fig.2, curve c). Further increase of S_R leads to partial rise of the curve.

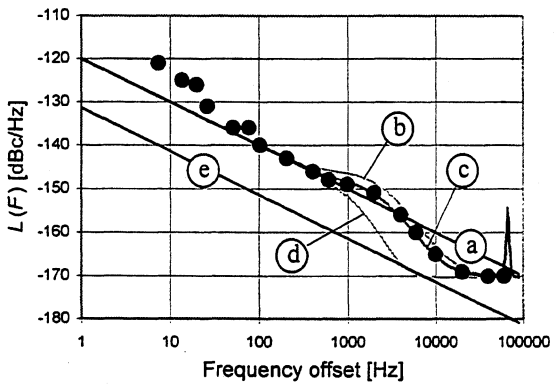


Figure 2.

In addition, Fig.2 displays the influence of S_R changes to predictions for the Colpitts mode. Here we may watch for the small influence of such changes to the spectral density splashes at two special offset frequencies $F = f_p - f_s$ and $F = f_p - f_c$.

6.3 Discussions

Accurate formula (8) allows the influence evaluation of parameters values and their noises into character and magnitude of amplitude and phase spectral densities. It is illustrated by the figures 1 and 2 where expected behaviors of spectral densities at the operation frequency of the Colpitts mode are also presented through the measured data obtained at that of the Butler mode. As the matter of fact, we mark that neglect the cross correlation between resonator parameters noises for the all examples gave rather small error near 3 dBc/Hz that quite appropriate in practice. The comparison showed that phase fluctuations in the Butler mode are reduced noticeably with the growth of $\overline{\Delta}_q$ - factor, though amplitude noises are considerably grown. Just to opposite, increase of $\overline{\Delta}_q$ in the Colpitts mode leads to growth of the phase fluctuations and to theoretically reduced to zero amplitude ones.

Owing to the practical difficulties, experimental estimations of fluctuations in the resonator are mainly based upon measurements made around the frequency f_s. Thus, obtained results are valid for the Butler mode of operation and, as it can be seen from Figure 1 and 2, far from those for the Colpitts one that is most widely used in highly stable and precision oscillators.

The Colpitts mode is characterized by desirable but not easy satisfied "low-noisy" operating frequency f_c, at which theoretically we may provide extremely small influence of a resonator motional inductance and capacity noises into phase instability. Discussion of this mode is given in [7].

7. CONCLUSIONS

We showed in the report that noise transformation coefficients of crystal resonator equivalent parameters into phase and modulus of its complex impedance may be effectively used for resonator phase and amplitude spectral density evaluations over a wide range of offsets from the carrier. It allows definition the "low-noisy" operation frequency of a crystal resonator in oscillator and study the resonator noises at an arbitrary operation frequency.

As follows from Figures 1 and 2, the approximate spectral functions for the Butler and Colpitts modes of operation fit the data good with offsets from the carrier less than half bandwidth of resonator. With this, the accurate formula (8) holds true not only within this range, but also with offsets far apart from the carrier. Here, we may study the splashes of spectral densities at the local frequencies $F = f_p - f_s$, $F = f_p - f_c$, and $F = f_c - f_s$ for the Butler and Colpitts modes correspondingly. Theoretically we get $f_p - f_c = f_c - f_s$ but, in practice, because inaccuracy of f_c setting, we may expect two splashes while operating in the Colpitts mode. In the considered examples, there had not been presented data (Figures 1 and 2) containing the points fallen within the range of theoretical splashes. This is the reason why not to give most complete explanation for the splashes in this paper.

8. REFERENCES

[1] V.N.Kuleshov, G.D.Janushevsky, "Fluctuation Characteristics of Quartz Crystal Resonators," in Proc. of Conf. on Freq. Control, (Moscow), VIMI, vol.2, pp.71-76, 1983 (in Russian).

[2] D.B.Leeson, "A simple Model of Feedback Oscillator Noise Spectrum," Proc. of the IEEE, vol.54, No.2, pp.329-330, 1966.

[3] A.E.Wainright, F.L.Walls, and W.D.McCaa, "Direct Measurements of the Inherent Frequency Stability of Quartz Crystal Resonators," in Proc. of the IEEE AFCS, pp.177-180, 1974.

[4] J.J.Gagnepain, "Fundamental Noise Studies of Quartz Crystal Resonators," in Proc. of the IEEE AFCS, pp.84-91, 1976.

[5] J.J.Gagnepain, M.Oliver, F.L.Walls, "Excess Noise in Quartz Crystal Resonators," in Proc. of the IEEE AFCS, pp.218-225, 1983.

[6] Yu.S.Shmaliy, "Conversion of fluctuations in quartz crystal resonators of thickness-shear vibrations" in Proc. of the Int. Symp. on AFCSG, (Moscow), pp.337-342, 1996.

[7] Yu.S.Shmaliy, "Conversion of 1/f fluctuations in crystal resonator within an inter resonance gap," IEEE Trans. on UFFC, Vol.46, No.1, pp.61-71, 1999.

[8] M.M.Driscoll, W.P.Hanson, "Measured vs. Volume Model-Predicted Flicker-of-Frequency Instability in VHF Quartz Crystal Resonators," in Proc. of the IEEE IFCS, pp.186-192, 1993.

[9] G.S.Curtis, "The Relationship Between Resonator and Oscillator Noise, and Resonator Noise Measurement Techniques," in Proc. of the IEEE AFCS, pp.420-428, 1987.

1999 Joint Meeting EFTF - IEEE IFCS

QUANTUM 1/F QUARTZ RESONATOR THEORY VERSUS EXPERIMENT

Peter H. Handel

Department of Physics and Astronomy, Univ. of Missouri St. Louis, St. Louis MO 63121

ABSTRACT

The quantum 1/f theory, as applied to bulk acoustic wave (BAW) and surface acoustic wave (SAW) resonators has recently been generalized to the case of an arbitrary coherence length of the phonons, which may be large or small compared with the size of the resonator. The theory has further been extended to include defect scattering along with the phonon scattering case exclusively considered earlier. The present paper compares the new theoretical results with experimental results obtained by F.L. Walls et al., T. Parker et al., J.R. Vig et al., as well as by other research groups.

1. INTRODUCTION

The quantum 1/f theory, as applied to bulk acoustic wave (BAW) and surface acoustic wave (SAW) resonators, is compared with experimental results obtained by F.L. Walls et al., T. Parker et al., J.R. Vig et al., as well as other research groups. The theory has recently been generalized to the case of an arbitrary coherence length of the phonons, which may be large or small compared with the size of the resonator. The theory has further been extended to include defect scattering along with the phonon scattering case exclusively considered earlier.

Our theoretical results [1], [2] can be written in the form

$$S(f) = \beta' V/fQ^4, \text{ for } V \le \varepsilon, \quad (1)$$

and

$$S(f) = \beta' \varepsilon^6/fVQ^4, \text{ for } V \ge \varepsilon, \quad (2)$$

where, with $<\omega>=10^8$/s, with $n=kT/\hbar <\omega>$, T=300K and $kT = 4 \cdot 10^{-21}$ J,

$$\beta' = (N/V)\alpha\hbar <\omega>/12n\pi g^2 mc^2$$
$$= 10^{22}(1/137)(10^{-27}10^8)^2/12kT\pi 10^{-27}\ 9\ 10^{20} = 1. \quad (3)$$

For the case of defect scattering, a two-phonon process takes place. A phonon from the main resonator mode scatters on a defect and a phonon of comparable frequency emerges into another mode with much smaller phonon occupation number $n_\omega = kT/\hbar\ \omega$. In this case we have to replace $<\omega>$ by ω and $n_{<\omega>}$ with n_ω, which gives a β-value which is $(<\omega>/\omega)^2$ smaller, i.e. 10^4-10^6 times smaller. In general, therefore, writing $\Gamma = \Gamma' + \Gamma''$, we obtain for the combined phonon and defect scattering case, in general,

$$\beta = \beta'[\Gamma'^2 + (<\omega>/\omega)^2\Gamma''^2]/\Gamma^2. \quad (4)$$

Although the defect scattering term is small at room temperature, it may become dominant at low temperatures, when the phonon scattering rate Γ' becomes much smaller than the defect scattering rate Γ''.

1. COMPARISON WITH THE EXPERIMENTAL DATA

The form of Eqs. (1)-(4) shows that the level of 1/f frequency noise depends not only as Q^{-4} as previously proposed for quartz [3], but is a non-monotonous function of the volume of the active region. The noise first increases with volume, then after reaching a maximum of the order of the phonon-coherence volume, it decreases with size. For quartz, this theory qualitatively fits the data of Gagnepain who varied the Q-factor with temperature in the same quartz resonator (but not frequency or volume), the data of Walls who considered several quartz resonators which differ in volume and frequency, and the data of Parker for SAW resonators, with their relatively low Q values. Indeed, according to [4], the median value of the PM noise L(10 Hz) in dBc/Hz for 12 unswept quartz resonators is -103.1, -101.6 and -97.7 for small, medium and large electrodes respectively, in reasonable agreement with the proportionality with V, which requires a 3 dB difference between the groups with large and small electrodes. The electrode diameters of the 3 size groups were 2.16 mm, 3.05 mm, and 4.32 mm.

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

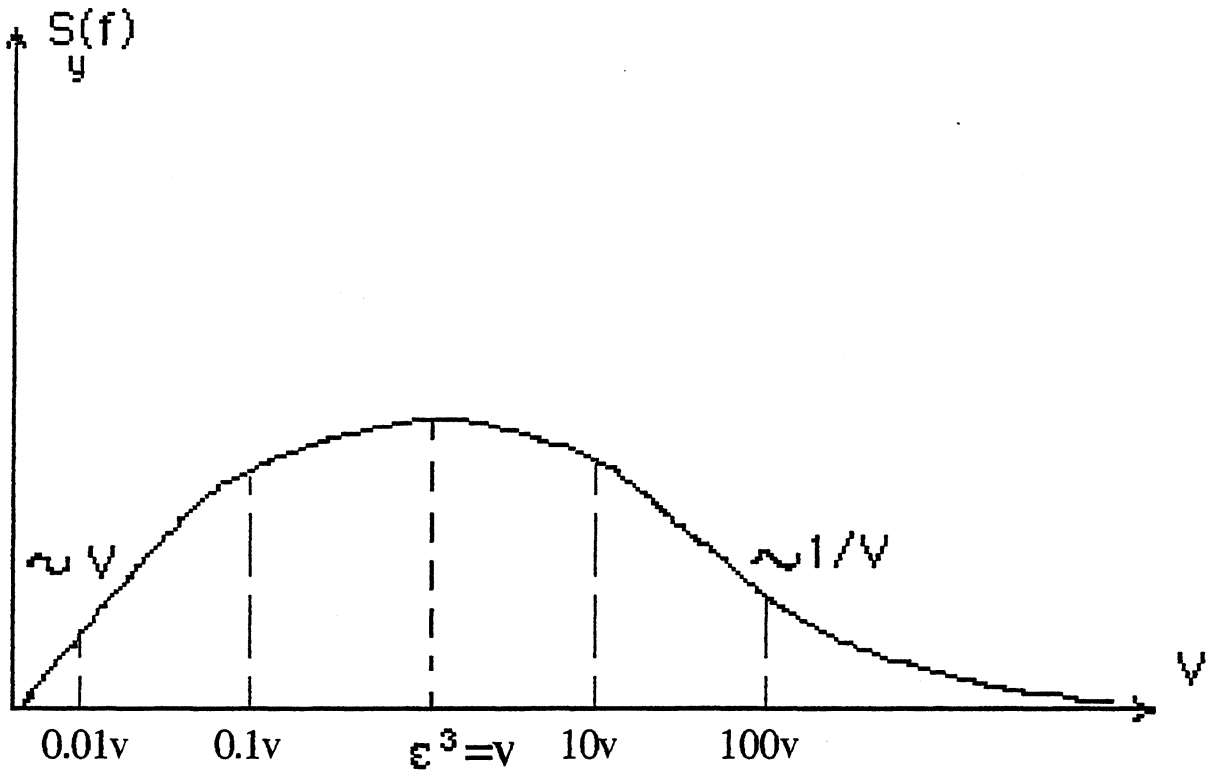
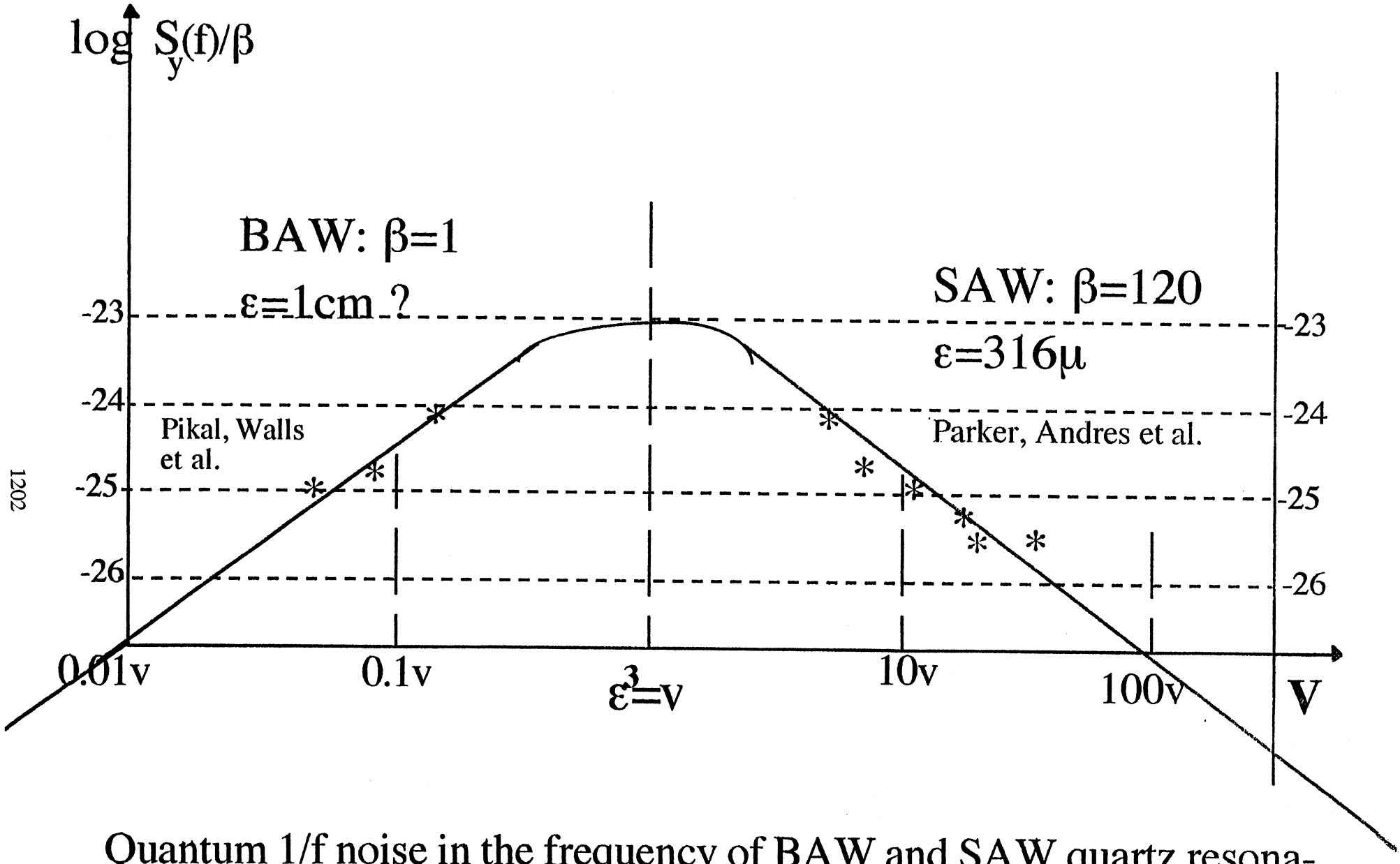


Fig. 1: Dependence of the spectral density of fractional frequency fluctuations on the the active quartz volume V.

TABLE 1
SAW
$\beta=120, \varepsilon=316\mu, \varepsilon^2=10^{-3}cm^2$

Design Identity	Quality factor Q_u	Eff. averaging area A_a cm^2	Eff. averaging volume V	Coher. βVQ^{-4}	Incoher. $(\varepsilon^2/A)^2\beta VQ^{-4}$	Experim. S(1Hz) Parker/Andres	n
A	21,800	$4.09\cdot10^{-3}$	$2.8\cdot10^{-6}$	$1.44\cdot10^{-21}$	$8.62\cdot10^{-23}$	$10\cdot10^{-23}$	4.09
B	21,800	$6.93\cdot10^{-3}$	$4.9\cdot10^{-6}$	$2.60\cdot10^{-21}$	$5.41\cdot10^{-23}$	$7.0\cdot10^{-23}$	6.93
C	14,500	$1.39\cdot10^{-2}$	$10\cdot10^{-6}$	$7.44\cdot10^{-21}$	$3.85\cdot10^{-23}$	$5.0\cdot10^{-23}$	13.9
D	22,850	$1.26\cdot10^{-2}$	$8.82\cdot10^{-6}$	$6.6\cdot10^{-21}$	$4.16\cdot10^{-23}$	$4.0\cdot10^{-23}$	12.6
E	20,675	$2.53\cdot10^{-2}$	$17.7\cdot10^{-6}$	$1.16\cdot10^{-20}$	$1.81\cdot10^{-23}$	$2.6\cdot10^{-23}$	25.3
F	20,900	$6.23\cdot10^{-2}$	$44.3\cdot10^{-6}$	$2.78\cdot10^{-20}$	$7.17\cdot10^{-24}$	$2.0\cdot10^{-23}$	62.3



1202

Quantum 1/f noise in the frequency of BAW and SAW quartz resonators as a function of volume at 1Hz, in terms of the phonon-interaction coherence volume $v=\varepsilon^3$ and of the quantum 1/f coefficient β.

The volumes were therefore approximately proportional with the numbers 4.67, 9.30 and 18.66, and therefore with the measured median values. However, the remaining scatter present in the data [4], [5], is analyzed in terms of the defect contributions and coherence corrections given by Eqs. (2)-(4). The theory also provides the basis for predicting from first principles, without adjustable parameters, how to improve the 1/f level of resonators, beyond just improving the Q-factor, which has been known for many years, and which has been related [2] to fluctuations in the dissipation. Since the 1/f noise level depends on the active volume, in the coherent regime one should use the lowest overtone and smallest diameter consistent with other circuit parameters. In the incoherent (low Q) case the opposite should be considered.

REFERENCES

[1] F.L. Walls et al., Proc. 46th Frequency Control Symposium, pp. 327-333, 1992.

[2] P.H. Handel, Proc. 51th Frequency Control Sympos., Orlando, FL, 1997.

[3] P.H. Handel, Solid State Electronics 22, p. 875 (1979).

[4] E.S. Ferre-Pikal et al., Proc. 50th Freq. Control Sympos., 1996, pp. 844-851.

[5] T.E. Parker and D. Andres, Proc. 48th Freq. Contr. Symp., pp. 530-538, 1994.

A PHENOMENOLOGICAL MODEL OF 1/f NOISE

Dmitry P. Tsarapkin

MPEI, Krasnokazarmennaya 14, Moscow, 111250, Russia

ABSTRACT

This paper deals with a problem of 1/f noise origin and proposes a phenomenological model which describes a possible mechanism of thermal noise to 1/f noise transformation.

1. INTRODUCTION

The 1/f noise displays itself as a universal phenomenon [1-2]. Nevertheless its origin stays far from full understanding despite many years of intensive investigations. In such situation it has a sense to develop one more noise model which could be useful in some cases.

According [3] 1/f noise can be presented by a train of hypothetical random pulses which depend on time t as a constant times t to the power -1/2. Thus, to solve the problem of 1/f noise origin in this approach means to find out a natural mechanism providing such pulses. It is just a goal of this paper.

2. NOISE TRANSFORMATION

To solve the problem, consider an arbitrary nonlinear equilibrium system affected by white noise. Put the equation coupling input, **X**, and output, **Y**, variables having a form

$$\frac{d\mathbf{Y}}{dt} = \mathbf{F}(\mathbf{X}, \mathbf{Y}) . \qquad (1)$$

In general, **X** and **Y** represent multi-dimensional column vectors. Thus, eq. (1) is equivalent to a set of first order differential equations with respect to time t. In the following analysis we consider one-dimensional case for simplicity since it will be enough to illustrate the main idea of this paper.

White noise can be represented by a train of random delta-pulses. Each individual delta-pulse gives rise to a transition process in the system output which form depends on particular non-linearity in vicinity of equilibrium and can be calculated as a solution of a corresponding equation for variations. Put both, X and Y, as a sum of a stationary value and a small variation from it:

$$X = X^0 + \xi, \quad Y = Y^0 + \delta . \qquad (2)$$

After substitution (2) to (1), slightly simplifying, we arrive to the equation for variations:

$$\dot{\delta} - (F'\delta + \frac{F''}{2!}\delta^2 + \frac{F'''}{3!}\delta^3 + \ldots) = \frac{\partial F}{\partial X}\xi \qquad (3)$$

where $F^{(i)} = \partial^i F/\partial Y^i$ — the partial derivatives F with respect to Y calculated in the point of equilibrium.

Consider now possible system responses to noise excitation. In most cases $Y' \neq 0$ .Then to the first order approximation the system response is a trivial relaxation process

$$\delta(t) = \delta_0 \exp[-t/\tau_e] \qquad (4)$$

with the effective time constant $\tau_e = -1/F'(Y^0)$.

The more complicated situation arises when the equilibrium corresponds to a point of extremum. Here the first derivative is equal zero and eq. (3) transforms to

$$\dot{\delta} - \frac{F''}{2!}\delta^2 = \frac{\partial F}{\partial X}\xi . \qquad (5)$$

Its solution has a form

$$\delta(t) = \delta_0 (-\frac{F''(Y^0)}{2!}t)^{-1} , \qquad (6)$$

i.e. δ-correlated input noise pulses lead to relatively slow decaying fluctuations in output.

The most interesting case takes place if both, the first and second derivatives, are zero in an operating point which, thus, accumulates an extremum and a flex point at once. Then the nearest non-zero term in the function F series expansion is F''' . It follows, eq. (3) is modified in the equation

$$\dot{\delta} - \frac{F'''}{3!}\delta^3 = \frac{\partial F}{\partial X}\xi \qquad (7)$$

having the solution

$$\delta(t) = \delta_0 \frac{\sqrt{-3/F'''(Y^0)}}{\sqrt{t}} . \qquad (8)$$

This work was sponsored by the Russian Foundation for Basic Investigations under a grant #98-02-17210.

0-7803-5400-1/99/ $ 10.00 © 1999 IEEE

The response is just proportional t to the desirable power -1/2. Thus, it is proved that in the some specific conditions a nonlinear system gets ability to transform a fundamental initial δ-correlated noise to 1/f one.

With physical point of view this fact arises due to accumulation of individual random shifts by a flat site on a system function plot.

3. EXAMPLE

As a particular example of the situation embodies the above results consider thermal distribution in a doped semiconductor sample.

Let the sample be a cylindrical stub with unity cross-section placed on a plate having a uniform ambient temperature T_A (Fig. 1). Investigate a pointwise sample model described with the inertial thermal impedance [4]

$$Z_T(p) = R_T/(1 + \tau_T p) . \qquad (9)$$

Here $p = d/dt$ - a differential operator with respect to time; in a steady-state regime $p = j\Omega$, Ω being an angular Fourier frequency.

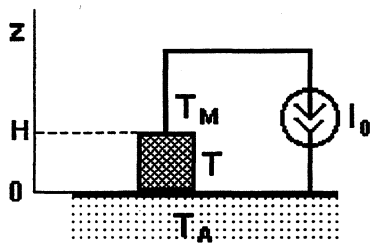


Figure 1. A sketch of the examined sample.

Let, further, the sample be heated with a direct current I_o. In this case the thermal balance equation has a form

$$T = T_A + Z_T(p)P_T \qquad (10)$$

where T — current temperature; $P_T = I^2_0 r$ — dissipation power; $r = H/\sigma_e$ — sample resistance; σ_e — specific electrical conductivity.

Semiconductor parameters are strongly temperature dependent. In doped semiconductor its specific thermal conductivity, σ_T, for $T \geq 300$ K follows the formula

$$\sigma_T \approx \alpha / T \qquad (11)$$

where a value of the constant α defined by particular material [5]. In GaAs, for instance, α = 150 W/cm for moderate doping and goes down up to 120 W/cm in n^+ zones. It follows, $R_T \propto 1/\sigma_T \propto T$. This stimulate sample overheating with temperature growth.

The second nonlinear phenomenon arises due to σ_e dependence on T. In accordance with [5] the graph σ_e vs. 1/T looks like in Fig. 2.

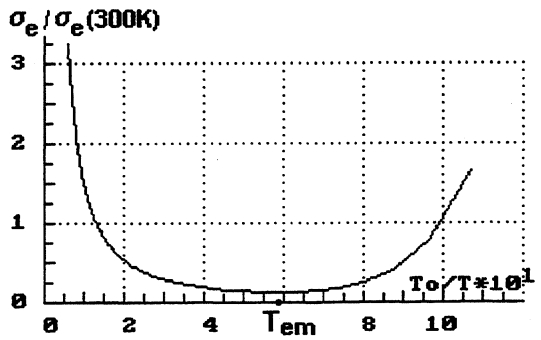


Figure 2. Typical temperature dependence of specific electrical conductivity.

At some particular temperature T_{em}, depending on material, σ_e reaches its minimal value σ_{em}. Usually $T_{em} \in (500, 1000)$ K. In temperature range $T \in (T_0, T_{em})$, where T_0 = 300K — the standard temperature, $\sigma_e \propto T^{-3/2}$. Thus, the form of a curve in Fig. 2 also stimulates sample overheating when $T < T_{em}$ if one uses power supply from a current source.

It is convenient to approximate $\sigma_e(T)$ as

$$\sigma_e(T) = \sigma_e(T_0)\, F_\sigma(T/T_0) \qquad (12)$$

where

$$F_\sigma(T_n) = (1+a)T^{3/2}_n(1+aT^{-(M+3)/2}_n) ; \quad a = (3/M)T^{-(M+3)/2}_n;$$

$M \sim 10...30$; $T_n = T/T_0$ —normalized temperature.

Substitution (9, 11, 12) into eq. (10) leads to a differential equation (13) describing the transient processes with changing working conditions:

$$\tau_T T_n(dT_n/dt) + F_h(T_n) - T_A/T_0 = 0. \qquad (13)$$

Here $F_h(T_n) = T_n(1 - \rho F^{-1}_\sigma(T_n)$ — the «overheating function» taking into account both nonlinear effects discussed; $\rho \propto P_T$ — the «overheating parameter».

In a stationary condition the averaged sample temperature $T = T_s$, being a solution of the equation

$$F_h(T_{ns}) = T_{An} . \qquad (14)$$

Without the nonlinear effects eq. (14) arrives to

$$T_s = T_A/(1-\rho) , \qquad (15)$$

what explains a sense of parameter ρ name.

Depending on ρ value, the function $F_h(T_{ns})$ demonstrates all three basic kinds of plots used in the previous part of this paper. In particular, if T_e = 600K the complex extremum-flex stationary point is born when T_A = 298.7 K, T_s = 526.5 K. It is very impressive that this remarkable state is observed in rather usual thermal regimes.

At next step we solved the heat conduction equation and found that this more exact model supports all previous basic results.

4. CONCLUSION

As revealed by this investigation there is a nonlinear mechanism providing fundamental to 1/f noise transform. The conditions needed this phenomenon takes place arise every time when any system function has an operating point where an extremum coincides with a flux point. Especially often they can be observed in some parts of spread nonlinear medium.

5. REFERENCES

[1] A. Van der Ziel, Noise: Sources, Characterization, Measurement, New York: Prentice-Hall. Ink., 1970.

[2] P.H. Handel, "1/f Noise Universality in High-Technology Applications," Proc. 1994 IEEE Int. Freq. Contr. Symp., 1-3 June 1994, Boston, USA, pp. 8-21.

[3] M.J. Buckingham, Noise in Electronic Devices and Systems, New York: John Wiley and sons, 1983, ch. 6.

[4] O. Muller, and J. Pest, "Thermal Feedback in Power Semiconductor Devices," IEEE Trans., 1970, vol. ED-17, n. 9, pp. 770-782.

[5] G.I. Epiphanov, Physical Basis of Microelectronics, Moscow: Soviet radio, 1971, 376 p. (in Russian).

1999 Joint Meeting EFTF - IEEE IFCS

A GENERIC GENERATION OF $1/F$ NOISE IN LOCKED SYSTEMS WORKING IN NONLINEAR MODE

Serge DOS SANTOS
LUSSI-GIP Ultrasons
École d'Ingénieurs du Val de Loire (EIVL)
Rue de la Chocolaterie BP 3140, F-41034 BLOIS CEDEX, FRANCE

Michel PLANAT
Laboratoire de Physique et Métrologie des Oscillateurs du CNRS associé à l'Université de Franche Comté
32, Av. de l'Observatoire, 25044 BESANÇON CEDEX, FRANCE

Properties of the phase fluctuations of a nonlinear oscillator is investigated using experimental setups which simulate the same behavior. The multiscale aspect of synchronization is experimentally confirmed with the observation of lockings between rational ratios. This global property of synchronization is associated to a local one by the study of the Phase Locked Loop (PLL) under nonlinear behaviors. We show that the discontinuity between the locking and the free-run behavior is characterized by the presence of slow temporal variables which induce a low frequency spectrum. An experimental verification of the increase in the spectrum confirms the classical results which associate $1/f$ noise to nonlinearity, autosimilarity and regularity represented here by the synchronization.

1. INTRODUCTION

Despite a lot of efforts, the ubiquitous $1/f$ noise is still not well understood [1]. In the field of frequency standards such a noise limits the long term stability and as such it is very desirable to improve our knowledge of its generation. In fact, the inability of linear theories to predict it has led several researchers to postulate the general nonlinear mechanism as its origin [2,3]. For example, Putterman observed a $1/f$ spectrum as a result of a nonlinear coupling of high frequency fluctuations [4]. Some years later, in his well-known paper on self-organized criticality, Bak [5] proposed a generic root of $1/f$ spectrum where nonlinearity is present in the dynamical system. Recently, a paper seems to confirm this introducing the concept of hyper-universality which tends to show the independence from the dimension of the system [6]. Likewise, P. Handel claimed that nonlinearity is a general cause if homogeneity is present in the equations of motion [1].

The relation between the indubitable property of nonlinearity in physics and the $1/f$ noise is not sufficient to explain it because chaos, for example, is not $1/f$ noise. If we refer to works given in references, another "ingredient" is needed which seems to be homogeneous, regular or hyper-universal. In this context, Arecchi has shown in an optical experiment that $1/f$ noise appears in nonlinear multistable systems having many attractors with fractal properties [7].

One of the most well-known "stable" states of nonlinear systems is the notion of synchronization which has been constantly studied since the discovery of its many properties, like phase synchronization and spatio-temporal synchronization [8,9]. They have been theoretically extracted thanks to the use of new methods like the Arnold tongues showing devil staircases and phase lockings [8,10,11] and involving studies on arrays of oscillators.

In frequency standards, this work began theoretically and numerically a few years before [12] by showing that this property of the synchronization is present in the global behavior of the oscillator. These results were at the beginning of the hypothesis that $1/f$ noise could hide a correlation [13], and the experimental proof of this complexity in oscillators was given with the carrier-envelope resonance experiment [14]. It has been experimentally shown that the intrinsic behavior of an oscillator results from a complex interaction between high and low frequencies having multiscale properties. With the use of a classical PLL under nonlinear conditions, this result has been confirmed [15] and improved showing that the multiscale properties are in fact directly related to arithmetical frequencies properties of the system [16]. These studies have confirmed that the multiscale property which has a structural coherence is present in nonlinear locked systems.

0-7803-5400-1/99/ $ 10.00

2. THE MULTISCALE PROPERTY OF SYNCHRONIZATION

1. Classical study

The synchronization of nonlinear oscillators has been studied since the original work of Neyfey [17]. If a Van der Pol oscillator, where the output signal $x = A\cos(\omega_s t + \Phi)$ verifies $\ddot{x} + \omega_0^2 x = \lambda(1-4x^2)\omega_0 \dot{x}$, is perturbed by a small component $y = \varepsilon \sin \omega_s t$, the variables A and Φ are the solutions to the nonlinear system:

$$\dot{A} = \frac{\lambda\omega_0 A}{2}(1-4A^2) + \frac{\varepsilon}{2\omega_s}\cos\Phi, \qquad (1a)$$

$$\dot{\Phi} = \Delta + \frac{\varepsilon}{2\omega_s A}\sin\Phi, \qquad (1b)$$

where $\Delta \sim \omega_0 - \omega_s$. The result is obtained using the multiscale method including slow variables. The non integrability of this system is responsible for the fact that we cannot give solutions to Eqs.(1). Nevertheless we can notice a synchronization of the oscillator ($\Phi = 0$) on the perturbation, if $A, \Delta, \varepsilon, \omega_s$ follow $-1 \leq -\frac{2\omega_s A_0 \Delta}{\varepsilon} \leq 1$, which is a synchronization condition. The higher ε is and the smaller Δ is, the better the synchronization can occur.

With the same study, we can show that synchronization can occur if $\omega_s = \omega_0/3$, or $\omega_s = 3\omega_0$, and more generally $\frac{\omega_s}{\omega_0} = \frac{p}{q}$. This synchronization depends on the solutions to Eq.(1b) which is the equation governing the phase of the oscillator. If we suppose the amplitude fluctuations to be negligible, A can be considered as a constant and Eq.(1b) is the well known Adler equation which appears in a nonlinear model of the PLL or in injection-locking systems [18].

2. New methods

A discretization of Eq.(1b) can be performed according to Poincaré sections of nonlinear systems. Defining $\dot{\Phi} = \frac{\Phi_{n+1} - \Phi_n}{h}$, where $h = \frac{2\pi}{\omega_s}$, we obtain

$$\Phi_{n+1} = \Phi_n + 2\pi\Omega + c\sin\Phi_n, \qquad (2)$$

where $\Omega = \frac{\omega_0}{\omega_s}$ and $c = \frac{\pi\varepsilon}{\omega_s^2 A}$. Eq.(2) is the well-known Arnold mapping which is a generic mapping of nonlinear driven systems. If we define a normalized mean frequency $\nu = \lim_{n\to\infty} \frac{\Phi_n - \Phi_0}{2\pi n}$, we obtain the devil staircase (Fig.1) which show the synchronization at all ratios $\Omega = p/q$.

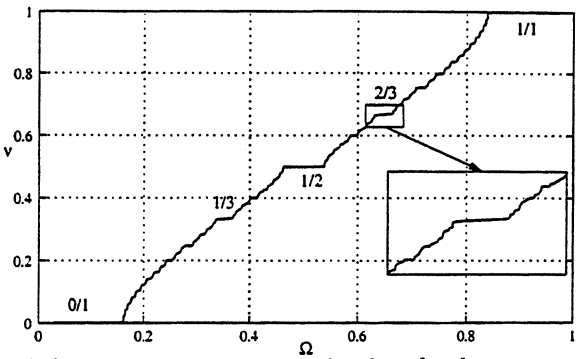


FIG. 1. The devil staircase for $\Omega \in [0;1]$, $n = 1000$, and $c = 0.16$. The plateaus located at $\nu = p/q$ are a signature of the lockings. The detail shows the multiscale invariance of the synchronization

3. Experimental verification

The devil staircase is a generic property of synchronized systems. It has been noted in many experiments [5,11]. We have verified this property in oscillators by studying a nonlinear MLSO (Mode Locked SAW Oscillator) in which hard coupling between high and low frequencies was introduced [14]. The set-up is shown in FIG.2, and measurements of frequency lockings are given in FIG.3. We observe the devil staircase showing the multiscale property of the synchronization in nonlinear oscillators.

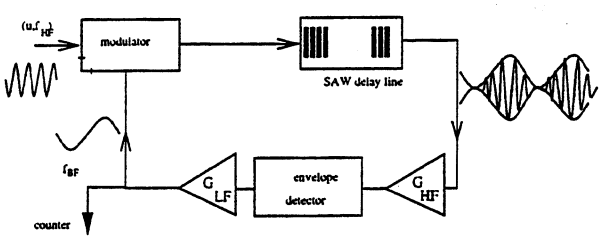


FIG. 2. The carrier-envelope experimental set-up. The frequency resonance occurs between the low frequency f_{LF} elaborated from the SAW delay line and the high frequency $f_{HF} = 91$ MHz of the injection

FIG. 3. Experimental devil staircase showing the synchronization of $\Omega = f_{FH}/f_{LF}$ versus the amplitude u

3. LOCAL PROPERTY OF SYNCHRONIZATION

1. The nonlinear PLL model

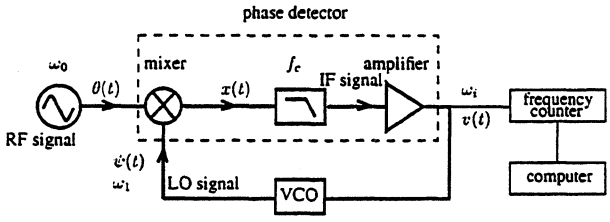


FIG. 4. The PLL used in the experiment. A high accuracy quartz oscillator with a relative stability of 10^{-11} is the RF stable reference. A low phase noise frequency synthesizer is used as the VCO.

In order to study experimentally the local property of the phase synchronization modeled by the Adler equation Eq.(1b), we have considered the nonlinear PLL (FIG. 4) where the phase Φ of the error signal $v(t)$ is given by

$$\dot{\Phi}(t) + K \sin \Phi(t) = \Delta\omega, \tag{3}$$

where $\Delta\omega = \mid \omega_0 - \omega_1 \mid$ is the angular frequency deviation and $K = \mu k A$ is the coupling coefficient (also called open loop gain). We can verify the two different PLL behaviors [19]: the synchronized state ($\Delta\omega < K$) and the free-run state ($\Delta\omega > K$)

2. The Stratanovich approach

An interesting study of such systems has been completed by Stratanovich [20] with a fruitful analogy to a mechanical overdamped system in a periodic potential $V(\Phi)$. Introducing an inertial term $\ddot{\Phi}(t)$ (experimentally possible if we decrease the cut-off frequency f_c in the loop filter), the phase Φ of the PLL follows

$$\frac{1}{f_c}\ddot{\Phi}(t) + \dot{\Phi}(t) + K \sin \Phi(t) = \Delta\omega, \tag{4}$$

which can be seen as the behavior of an overdamped mass along a corrugated inclined plane (FIG. 5) in a periodic potential given by $V(\Phi) = -\Delta\omega\Phi - K\cos\Phi$. In fact, the two different PLL behaviors are identified as follows. If $\Delta\omega < K$, the mass does not have enough energy to jump into the next potential well: the phase is constant and the PLL is synchronized. If $\Delta\omega > K$, the mass can step over each periodic barrier and have an increasing evolution: the associated phase Φ yields a periodic oscillation corresponding to the free run of the PLL.

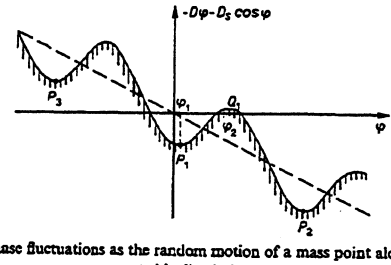


Fig. 26. Phase fluctuations as the random motion of a mass point along a corrugated inclined plane.

FIG. 5. The Stratanovich approach of the nonlinear PLL [20] where $D \sim \Delta\omega$ and $D_s \sim K$

The two different behaviors are separated by an unstable equilibrium position called separatrix which is well-known to exhibit high chaotic and unstable dynamics. When the PLL operates under these conditions, slow variables appear with associated low frequency components (like for example the free-run frequency which tends to 0 when $\Delta\omega \to K$ with a positive value).

3. Cycle skipping and $1/f$ noise

If we add a white frequency noise to the oscillator ω_0, the PLL error signal (under linear conditions) will be composed of the product of a white frequency spectrum and the transfer function of the PLL. With the Stratanovich approach, the mass will fluctuate (after a transient) in the first well (if fluctuations are small). If the fluctuations are greater or if we are near the separatrix, the mass can step over the separatrix by doing a complete phase rotation usually called a cycle-skipping. Wichert *et al.* observed experimentally a $1/f$ spectral distribution in a first order PLL with delay [21]. These results agree with the fact that $1/f$ noise appears near separatrix as observed ten years ago by Yamaguchi [22]. Our observation of the $1/f$ noise near the separatrix was made by applying an additive white frequency noise to the input oscillator. The Allan deviation $\sigma_{\omega_{LF}}$ of the beat frequency $\omega_{LF} = K\sqrt{(\Delta\omega/K)^2 - 1}$ is plotted on FIG.6 versus $u = \Delta\omega/K$. The Allan deviation is then given by [19]:

$$\sigma_{\omega_{BF}} = \sigma_{\omega_0}\frac{\omega_0}{\Delta\omega}\left(1 + \frac{K^2}{\omega_{BF}^2}\right). \tag{5}$$

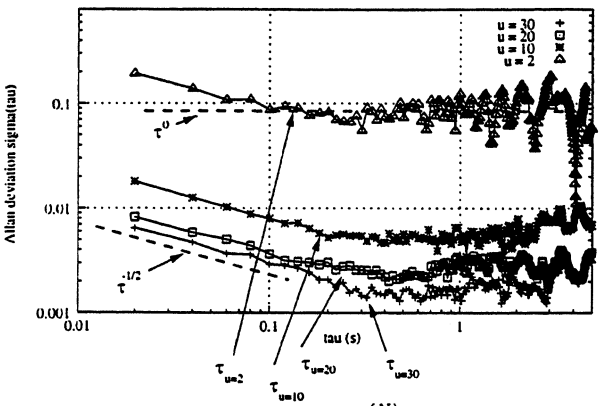


FIG. 6. The Allan deviation $\sigma_{\Phi}^{(N)}(\tau)$ ($N = 2000$) of frequency beat ω_{LF} for different values of u. A high open loop gain $K = 1.5$ kHz was chosen in order to amplify the nonlinear effects. The transformation of the noise is detected in the range $\tau \in [0.1; 0.5]$. The white frequency noise ($\tau^{-0.5}$) is changed into 1/f noise (τ^0)

When $\Delta\omega \to K$ (or $u \to 1$), the relative fluctuations increase as shown in [23]. Furthermore we see a modification of the initial spectrum from the initial white frequency noise (represented as a $\tau^{-1/2}$ slope) to $1/f$ noise (τ^0 slope). This effect is confirmed in FIG.7, where we perform the FFT on relative frequencies. The increase and the modification of fluctuations are clearly visible.

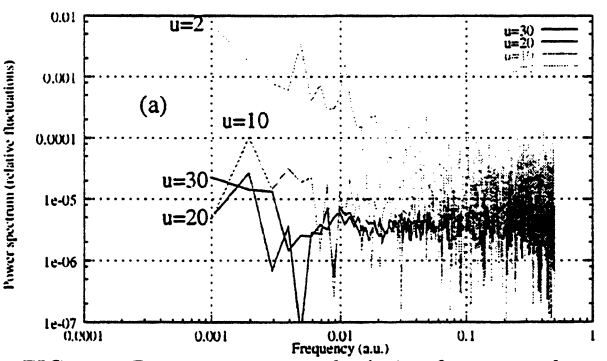


FIG. 7. Power spectrum of relative frequency data (counted with $\tau = 0.1$ s). The modification of the spectrum at low frequency is higher when we approach the separatrix $u = 1$

4. CONCLUSION

We have observed a $1/f$ noise spectrum in a highly nonlinear system, *i.e.* the PLL under particular conditions. The system is then characterized by two different behaviors (the free-run and the synchronization) and the fluctuations between them are $1/f$ fluctuations.

A local analysis of the synchronization process confirms that the $1/f$ noise is a feature of the presence of a discontinuity, a highly nonlinear behavior characterized by separatrix and slow temporal variables. Oscillators exhibit a multiscale characteristic of the synchronization involving arithmetical rules. This arithmetical or fractal network in which we put the synchronization pattern explains the scale invariance of the $1/f$ frequency noise fluctuations.

[1] P. H. Handel, in *Noise in Physical Systems and* $1/f$ *fluctuations* (Peter H. Handel and Alma L. Chung, Saint-Louis, 1993), p. 162.

[2] S. Teitler and M. F. M. Osborne, Phys. Rev. Lett. **27**, 912 (1971).

[3] R. F. Voss, in *Proc. 33rd Ann. Symp. Frequency Contr.* (Atlantic City, NJ, 1979), p. 40.

[4] S. Putterman, Phys. Rev. Lett. **39**, 585 (1977).

[5] P. Bak, C. Tang, and K. Wiesenfeld, Phys. Rev. A **38**, 364 (1988).

[6] P. De Los Rios and Y.-C. Zhang, Phys. Rev. Lett. **82**, 472 (1999).

[7] F. T. Arecchi, R. Badii, and A. Politi, Noise in Physical systems and 1/f noise 127 (1983).

[8] A. Pikovsky, M. Rosenblum, and J. Kurths, Europhysics Letters **34**, 165 (1996).

[9] L. Kocarev and U. Parlitz, Phys. Rev. Lett. **77**, 2206 (1996).

[10] A. S. Pikovsky and J. Kurths, Phys. Rev. Lett. **78**, 775 (1997).

[11] C. Reichhardt and Franco Nori, Phys. Rev. Lett. **82**, 415 (1999).

[12] M. Planat, Proceedings of the 6th European Frequency and Time Forum, Noordwijk 333 (1992).

[13] M. Planat *et al.*, IEEE Trans. on Ultras. Ferroelect. and Freq. Control **43**, 326 (1996).

[14] M. Planat *et al.*, J. Appl. Phys. **80**, 2509 (1996).

[15] S. Dos Santos and Michel Planat, Fractals and Beyond, Ed. by M. Novak, World Scientific, p. 299 (1998).

[16] M. Planat *et al.*, in *Proc. of VII Van der Ziel Symposium on Quantum 1/f noise ,St Louis, august 1998* (1998).

[17] A. H. Neyfey and D. T. Mook, *Nonlinear oscillators* (Wiley-interscience, 1979).

[18] R. Adler, Proc. IRE **34**, 351 (1946), reprinted in Proc IEEE, 61, 1380 (1973).

[19] S. Dos Santos, Ph.D. thesis, Besançon, 1998.

[20] R. L. Stratanovich, *Theory of random noise* (Gordon and Breach, New-York, 1967).

[21] W. Wischert, M. Olivier, and J. Groslambert, in *Noise in Physical Systems and* $1/f$ *fluctuations* (Peter H. Handel and Alma L. Chung, New York, 1993), p. 333.

[22] Y. Yamaguchi and K. Sakai, Physics Letters A **117**, 387 (1986).

[23] S. Dos Santos and M. Planat, Proc. 12^{th} Ann. European Time and Frequency Forum EFTF, Warsaw, Poland (1998).

PROCEEDINGS ORDERING INFORMATION

NO.	YEAR	DOCUMENT #	FROM**
10	1956	AD-298322	NTIS[1]
11	1957	AD-298323	NTIS
12	1958	AD-298324	NTIS
13	1959	AD-298325	NTIS
14	1960	AD-246500	NTIS
15	1961	AD-265455	NTIS
16	1962	PB-162343	NTIS
17	1963	AD-423381	NTIS
18	1964	AD-450341	NTIS
19	1965	AD-471229	NTIS
20	1966	AD800523	NTIS
21	1967	AD-659792	NTIS
22	1968	AD-844911	NTIS
23	1969	AD-746209	NTIS
24	1970	AD-746210	NTIS
25	1971	AD-746211	NTIS
26	1972	AD-771043	NTIS
27	1973	AD-771042	NTIS
28	1974	AD-A011113	NTIS
29	1975	AD-A017466	NTIS
30	1976	AD-A046089	NTIS
31	1977	AD-A088221	NTIS
32	1978	AD-A955718	NTIS
33	1979	AD-A213544	NTIS
34	1980	AD-A213670	NTIS
35	1981	AD-A110870	NTIS
36	1982	AD-A130811	NTIS
37	1983	AD-A136673	NTIS
38	1984	AD-A217381	NTIS
39	1985	AD-A217404	NTIS
40	1986	AD-A235435	NTIS
41	1987	AD-A216858	NTIS
42	1988	AD-A217275	NTIS
43	1989	AD-A235629	NTIS
44	1990	AD-A272017	NTIS
45	1991	AD-A272274	NTIS
46	1992	92CH30833	IEEE*
47	1993	93CH32441	IEEE*
48	1994	94CH34462	IEEE*
49	1995	95CH35752	IEEE*
50	1996	96CH35935	IEEE*
51	1997	97CH36016	IEEE*
52	1998	98CH36165	IEEE*
53	1999	99CH36313	IEEE*[2]

*IEEE members may order Proceedings at half price.
[1] Please check with NTIS and/or IEEE for current pricing.
[2] 1999 Joint Meeting with the EFTF

** NTIS – National Technical Info Services
5285 Port Royal Road
Sills Bldg.
Springfield, VA 22161
Tel: 703-487-4650

http://www.fedworld.gov/ntis/search.htm

** IEEE – Institute of Electrical & Electronics Engineers
445 Hoes Lane
Piscataway, NJ 08854
Tel: 1-800-678-4333

customer.services@ieee.org
http://www.ieee.org/ieeestore/ordinfo.html

EUROPEAN FREQUENCY AND TIME FORUM

PAST PROCEEDINGS AND OTHER INFORMATION

For information about the European Frequency and Time Forum (EFTF), including past proceedings, please contact the Permanent EFTF Secretariat at either of the following addresses:

In Switzerland at:

FSRM
Swiss Foundation for Research in Microtechnology
Rue Jaquet-Draz 1/case Postale 20
CH-2007 Neuchatel, Switzerland

Tel: +41 32 7 200 900

Fax: +41 32 7 200 990

e-mail: eftf@fsrm.ch
http://www.fsrm.ch

In France at:

Société Française des Microtechniques et de Chronométrie
c/o LCEP/ENSMM
26, Chemin de l'Epitaphe
F-25030 Besançon Cedex, France

Tel: +33 3 81 40 28 21

Fax: +33 3 81 88 57 14

e-mail: isabelle.bourgon@ens2m.fr

SPECIFICATIONS AND STANDARDS RELATING TO FREQUENCY CONTROL

Institute Of Electrical & Electronic Engineers (IEEE)

Order from: IEEE Service Center
445 Hoes Lane
Piscataway, NJ 08854
Telephone: (732) 981-0060

176-1987 (ANSI/IEEE)Standard on Piezoelectricity (SH 11270)

177-1966 Standard Definitions & Methods of Measurements of Piezoelectric Vibrators

180-1986 (ANSI/IEEE) Definitions of Primary Ferroelectric Crystal Terms (SH10553)

319-1971 (Reaff 1978) Piezomagnetic Nomenclature (SH02360)

1139-1988 Standard Definitions of Physical Quantities for Fundamental Frequency & Time Metrology (SH12526)

Department of Defense (DOD)

Order from: Naval Pubs & Form Center
5801 Tabor Avenue
Philadelphia, PA 19120
Telephone: (215) 697-2000

General Specs for:

MIL-C-3098 Crystal Unit, Quartz

MIL-C-24523 (SHIPS) Chronometer, Quartz Crystal

MIL-F-15733 Filters & Capacitors, Radio Interference

MIL-F-18327 Filters, High Pass, Band Pass Suppression and Dual Processing

MIL-F-28861 Filters and Capacitors, Radio Frequency Electro-magnetic Interference Suppression

MIL-F-28811 Frequency Standard, Cesium Beam Tube

MIL-H-10056 Holders (Encl), Crystal

MIL-O-55310 Oscillators, Crystal
MIL-O-39021 Oven

MIL-S-4933(ER) Surface Acoustic Wave Devices

MIL-STD-683 Crystal Units, Quartz/Holders, Crystal

MIL-STD-188-115 Interoperability & Performance Standards for Communications, Timing & Synchronization Subsystems

MIL-STD-1395 Filters & Networks, Selection & Use

MIL-T-28816(EC) Time Frequency Standard, Disciplined AN/URQ-23

MIL-W-46374D Watch, wrist: General purpose

MIL-W-87967 Watch, wrist: Digital

General Services Administration (GSA)

Order from: Naval Pubs & Form Center **OR** General Services Administration Business Service Centers in major U.S. cities.

FED-STD-1002 Time & Frequency Reference Information in Telecommunication Systems

Electronic Industries Association (EIA)

Order from: Electronic Industries Assoc.
2001 Eye Street, NW
Washington, DC 20006
Telephone: (202) 457-4900

(a) Holders and Sockets

EIA-192-A, Holder Outlines and Pin Connections for Quartz Crystal Units (standard dimensions for holder types)

EIA-367, Dimensional & Electrical Characteristics Defining Receiver Type Sockets (including crystal sockets)

EIA-417, Crystal Outlines (standard dimensions and pin connections for current quartz crystal units, 1974)

(b) Production Tests

EIA-186-E, (All Sections) Standard Test Methods for Electronic Component Parts
EIA-512, Standard Methods for Measurement of Equivalent Electrical Parameters of Quartz Crystal Units, 1 kHz to 1 GHz, 1985

EIA-IS-17-A, Assessment of Outgoing Non-conforming Levels in Parts per Million (PPM)

EIA-IS-18, Lot Acceptance Procedure for Verifying Compliance with Specified Quality Level in PPM

(c) Application Information

EIA Components Bulletin No. CB6-A, Guide for the Use of Quartz Crystal Units for Frequency Control, Oct. 1987

(d) EIA-477, Cultured Quartz (Apr. 81)

EIA-477-1, Quartz Crystal Test Methods (May 1985)

International Electro-Technical Commission (IEC)

Order from: American Nat'l. Standard Inst. (ANSI)
1430 Broadway
New York NY 1001
Telephone: (212) 354-3300

IEC Publications Prepared by TC 49:

122: Quartz crystal units for frequency control and selection

122-2 (1983) Part 2: Guide to the use of quartz crystal units for frequency control and selection

122-2-1 (1991) Section One: Quartz crystal units for microprocessor clock supply (Amendment 1 - 1993)

122-3 (1977) Part 3: Standard outlines and pin connection (Amendment 2 -1991, Amendment 3 - 1992, Amendment 4 - 1993)

283 (1986) Methods for the measurement of frequency and equivalent resistance of unwanted resonances of filter crystal units

302 (1969) Standard definitions and methods of measurement for piezoelectric vibrators operating over the frequency range up to 30 MHz

314 (1970) Temperature control devices for quartz crystal units (Amendment 1 - 1979)

314A (1971) First supplement

368: Piezoelectric Filters

368-1 (1992) Part 1: General information, standard values and test conditions

368-2 (1973) Part 2: Guide to the use of piezoelectric filters

368-2-1 (1988) Section One - Quartz crystal filters

368B (1975) Second supplement

368-3 (1991) Part 3: Standard Outlines

444: Measurement of quartz crystal unit parameters

444-1 (1986) Part 1: Basic method for the measurement of resonance frequency and resonance resistance of quartz crystal units by zero phase technique in a π - network with compensation of the parallel capacitance C_o

444-4 (1988) Part 4: Method for the measurement of the load resonance frequency f_L, load resonance R_L and the calculation of other derived values of quartz crystal units up 30 MHz

483 (1976) Guide to dynamic measurements of piezoelectric ceramics with high electromechanical coupling

642 (1979) Piezoelectric ceramic resonators and resonator units for frequency control and selection.
Chapter I: Standard Values and Conditions
Chapter II: Measuring and test conditions

642-2 (1994) Part 2: Guide to the use of piezoelectric ceramic resonator units

642-3 (1992) Part 3: Standard outlines

679: Quartz Crystal Controlled Oscillators

679-1 (1980) Part 1: General information, test conditions and methods (Amendment 1 - 1985)

679-2 (1981) Part 2: Guide to the use of quartz crystal controlled oscillators

679-3 (1989) Part 3: Standard outlines and lead connections (First supplement - 1991) (Amendment 1 - 1994)
689 (1980) Measurements and test methods for 32 kHz quartz crystal units for wrist watches and standard values

758 (1993) Synthetic quartz crystal; specifications and guide for use

862: Surface Acoustic Wave (SAW) Filters:

862-1 (1989) Part 1: General Information, standard values and test conditions, Chapter I: General information and standard values, Chapter II: Test conditions

862-2 (1991) Part 2: Guide to the use of surface acoustic wave filters (Chapter III)

862-3 (1986) Part 3: Standard outlines (Chapter IV)

1019: Surface Acoustic Wave (SAW) Resonators

1019-1-1 (1990) Part 1: General information, standard values and test conditions, Section 1 - General information and standard values

1019-1-2 (1993) Section 2: Test conditions

1019-1-3 (1991) Part 3: Standout outlines and lead connections

1080 (1991) Guide to the measurement of equivalent electrical parameters of quartz crystal units

1178-1 (1993) Quartz crystal units - a specification in the IEC Quality Assessment System for Electronic Components (IECQ) Part 1: General Specification

1178-2 (1993) Part 2: Sectional specification - Capability approval

1178-2-1 (1993) Part 2: Sectional specification - Capability approval, Section 1: Blank detail specification

1178-3 (1993) Part 3: Sectional specification - Qualification approval

1178-3-1 (1993) Part 3: Sectional specification - Qualification approval, Section 1: Blank detail specification

1240 (1994) Piezoelectric devices - preparation of outline drawings of surface-mounted devices (MSD) for frequency control and selection, general rules

1253: Piezoelectric ceramic resonators - a specification in the IEC quality assessment system for electronic components (IECQ)

1253-1 (1993) Part 1: Generic specification - qualification approval

1253-2 (1993) Part 2: Sectional specification - qualification approval

1253-2-1 (1993) Section 1 - Blank detail specification - Assessment Level E

1261: Piezoelectric Ceramic Filters for use in Electronic Equipment, a specification in the IEC quality assessment system for electronic components (IECQ)

1261-1 (1994) Part 1: General specifications, qualification approval

1261-2 (1994) Part 2: Sectional specifications, qualification approval

1261-2-1 (1994) Part 2: Section 1, Blank detail specification, Assessment Level E

AUTHOR INDEX

Ces actes ont été imprimés avec l'aide du CONSEIL GENERAL DU DOUBS
et par les soins de son imprimerie.

These proceedings have been printed by courtesy of "CONSEIL GENERAL
DU DOUBS"in its printing house.