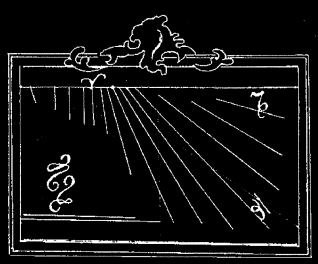
14th EFTF

EUROPEAN FREQUENCY AND TIME FORUM

Omne tempus impensum



quomodo fluit expensum ?

PROCEEDINGS

14-16 March 2000
Torino, Italy

14th EFTF

EUROPEAN FREQUENCY AND TIME FORUM

14-16 March 2000, Torino (Italy)

Organized by

Istituto Elettrotecnico Nazionale Galileo Ferraris
Politecnico di Torino

Under the High Patronage of
the President of Italian Republic

With the sponsorship of

Regione Piemonte

Provincia di Torino

Comune di Torino

FONDAZIONE CRT
Cassa di Risparmio di Torino
Fondazione CRT

AEM
AEM

14th EUROPEAN FREQUENCY AND TIME FORUM

2000 Conference Chairman

Sigfrido Leschiutta	IEN Galileo Ferraris / Politecnico di Torino, Italy

2000 Technical Program Chairman

Andrea De Marchi	Politecnico di Torino, Italy

2000 Conference Secretariat

Antonella Alotto	Politecnico di Torino, Italy
Elisabetta Melli	IEN Galileo Ferraris, Italy

Technical Program Committee

Andrea De Marchi	Politecnico di Torino, Italy - Chairman 2000 EFTF
Jean-Pierre Aubry	Oscilloquartz SA, Switzerland
Claude Audoin	CNRS/LHA, France
Andreas Bauch	Physikalishe-Technische Bundesanstalt, Germany
Raymond J. Besson	LCEP/ENSMM, France
Jean-Simon Boulanger	National Research Council of Canada, Canada
Michel Brunet	CNES, France
Giovanni Busca	Observatoire Cantonal, Switzerland
Leonard Cutler	Hewlett Packard Laboratories, USA
Gerrit De Jong	NMi Van Swinden Laboratory, The Netherlands
Philippe Defranould	Thomson-Microsonics, France
François Deyzac	ONERA, France
Klaus Dorenwendt	Physikalishe-Technische Bundesanstalt, Germany
Pekka Eskelinen	Lappeenranta University of Technology, Finland
Steven Feltham	ESA/ESTEC, The Netherlands
Michael Garvey	Frequency and Time Systems Inc., USA
Aldo Godone	IEN Galileo Ferraris, Italy
Michel Granveaud	BNM/LPTF, France
Gulayaev Yuri	Russian Academy of Sciences, Russia
Jörg Hahn	DLR, Germany
Daniel Hauden	CNRS/LPMO, France
Kurt Hilty	Swiss Federal Office of Metrology, Switzerland
Carl K. Hruska	York University, Canada
Kalevi Kalliomaki	University Oulu, Finland
Masami Kihara	NTT Optical Network Systems Lab., Japan
Dieter Kirchner	Technische Universität Graz, Austria
Peter W. Krempl	AVL List GmbH, Austria

Sigfrido Leschiutta	IEN Galileo Ferraris / Politecnico di Torino, Italy
Lute Maleki	NASA Jet Propulsion Laboratory, USA
Bernd W. Neubig	Tele Quarz GmbH, Germany
Léon Prost	Office Fédéral de Métrologie, Switzerland
Wofgang Schlüter	Institut für Angewandte Geodäsie, Germany
James Mc A. Steele	National Physical Laboratory, United Kingdom
Pierre Thomann	Observatoire Cantonal, Switzerland
Adolph A. Uljanov	Kvarz Inst. of Electronic Measurements, Russia
Unverdross Rainer	Unverdross Technik, Germany
Jacques Vanier	Canada
Fred L. Walls	NIST, USA
Gernot M.R. Winkler	Innovative Solutions International, USA
Walter Zingg	Micronas Semiconductor SA, Switzerland

Scientific and Organizing Committees of the 14th European Frequency and Time Forum are not responsible for the opinions expressed in the papers published in these Proceedings. The contents of the papers express entirely the views of the authors.

ISBN 88 - 900352 - 2 - 6

TECHNICAL SESSION CHAIRMEN

PLENARY PAPERS
Andrea De Marchi, Politecnico di Torino, Torino, Italy

PIEZOELECTRIC MATERIALS
Daniel Hauden, LPMO, Besançon, France

MICROWAVE FREQUENCY STANDARDS
Andreas Bauch, PTB, Brauschweig, Germany

GEODETIC GPS
Sigfrido Leschiutta, Politecnico di Torino and IEN, Torino, Italy

ATOMIC FOUNTAINS
Christoff Salomon, ENS, Paris, France

CLOCKS FOR SERVICES
Pekka Eskelinen, Lappeenranta University of Technology, Finland

HIGHLY STABLE MICROWAVE OSCILLATORS
Elio Bava, Politecnico di Milano, Milano, Italy

RESONATORS AND FILTERS
Raymond Besson, ENSMM, Besançon, France

PHASE NOISE MEASUREMENTS
Aldo Godone, IEN, Torino,Italy

PIEZOELECTRIC SENSORS
Clemens. Ruppel, Siemens AG, Munich, Germany

INFRARED LASER STABILIZATION
Fabrizio Bertinetto, IMGC-CNR, Torino, Italy

PIEZOELECTRIC OSCILLATORS
Berndt Neubig, Tele Quarz GmbH, Germany

OPTICAL FREQUENCY SYNTHESIS
Giovanni Daniele Rovera, LPTF-BNM, Paris, France

TIME AND FREQUENCY DISSEMINATION
Michael J. Underhill, University of Surrey, Guilford, UK

POSTER SESSION "A": PIEZOELECTRICITY AND MEASUREMENTS
J.P. Aubry, Oscilloquartz SA, Switzerland

POSTER SESSION "B": TIME AND FREQUENCY
Andrea De Marchi, Politecnico di Torino, Torino, Italy

TABLE OF CONTENTS

GEODETIC GPS

ATOMIC FOUNTAINS

CLOCKS FOR SERVICES

HIGHLY STABLE MICROWAVE OSCILLATORS

RESONATORS AND FILTERS

PHASE NOISE MEASUREMENTS

PIEZOELECTRIC SENSORS

INFRARED LASER STABILIZATION

PIEZOELECTRIC OSCILLATORS

OPTICAL FREQUENCY SYNTHESIS

TIME AND FREQUENCY DISSEMINATION

POSTER SESSION "A": PIEZOELECTRICITY AND MEASUREMENTS

POSTER SESSION "B": TIME AND FREQUENCY

Welcome of Professor Rodolfo Zich
Rettore of the Politecnico di Torino
Corso Duca degli Abruzzi, 24
10129 Torino, Italy

It is indeed for me a pleasure to express on behalf of the Politecnico di Torino, the very warm welcome of our school of Engineering to this distinguished gathering of experts in the realm of Time and Frequency, for the first time in Italy.

Politecnico di Torino, is one of the largest Italian Universities devoted to the technical formation. The enrolment is of about 20 000 students with a faculty of some 800 professors. We are covering near all the branches of the engineering. The School is, at the moment unique in Italy, in the sense that is also formally divided in three levels, leading respectively to a diploma, a laurea and finally a doctorate.

But I am here not only because the Politecnico is the Technical University of Torino, but also because, in the past and also at the present, the links with your community were always very strong.

Let me quote just a score of these accomplishments.

Some of these links go back to 1926, when a researcher that later became professor of our Faculty, Giancarlo Vallaury, organised with Dr Morrison of the National Bureau of Standards, the first international intercomparison of crystal standards between America and Europe. The comparison, with devices made at Bell Laboratories, was repeated two years later, reaching an accuracy, in the comparison of some units of 10^{-5}. The measurements were performed also in Leghorn, at the Italian Navy Research laboratory. It is refreshing, at three-quarter of century after the facts, to read the papers and to remain admired from the ingenuity of those pioneers.

Professor Mario Boella, another of our Professors, who left us just about 10 years ago, performed a number of similar accomplishments in Time and Frequency Metrology.

He constructed in 1935 a direct frequency synthesiser, with a decadic formation of the output frequency, adopting the principle "addition, addition, division" that was industrially adopted after 1950. A more complete and flexible device was made by him in 1946.

But possibly the more lasting activity of Boella was on the piezoelectric resonators and on piezoelectric oscillators made at the IEN. At that period the IEN was also the department of Electrotechnics and Electronics of Politecnico.

In the period 1950-1960 Boella made Y-bar resonators and 100 kHz, and constructed a series of oscillators, some of witch have survived for more of 180000 hours of service. This activity is reflected in about 25 papers, in the period 1930-1960. Boella started also in 1958 the construction of the IEN caesium beam frequency standards, just three years after the Essen device at National Physical Laboratory.

Research was not limited to the frequency standards, because also clocks and experiments in dissemination were performed. Only one is here recalled, the measurement of the HF propagation time between IEN and the National Bureau of Standards, then in Beltsville, using a two-way method. Second pulses were emitted in

Torino by IBF and across the ocean by WWV. With a very elegant method, the propagation delay and its differences and vagaries at the frequencies of 5, 10 and 15 MHz were investigated.

Research activities in the Frequency and Time domain are still performed in our school of Engineering; as you will see two caesium frequency standards are being mounted, a "strong field" and a fountain.

The fountain that is inside ad agreement between Politecnico, NIST and IEN; it is interesting to note that this time-proven cooperation is still active, more that seventy years after the two trips by Morrison to Italy.

One of our third level schools, that leading to a title of doctor, is active since 1985; a fair number of the young Italian authors presenting papers at this Forum, took their doctorate in Time and Frequency Metrology at our Politecnico.

The links between your community and Politecnico di Torino were and are indeed stron and I hope they will remain so or stronger in the future.

Clocks and Tests of Fundamental Physics in Space

Robert F.C. Vessot,
Smithsonian Astrophysical Observatory, Cambridge MA 02138

ABSTRACT

The capability of new clock systems operating in space to test theories of relativity and gravitation is discussed. Possible future tests with new space borne high stability oscillators and frequency discriminators, now under development, are described.

1. INTRODUCTION

A major goal of theoretical physicists is to include gravitation in an overall theory that now includes the electomagnetic, strong and weak forces. To proceed, it is vital to test the underlying assumptions of the principle of equivalence that is the foundation of Einstein's General Theory and to conduct experiments to verify these assumptions with higher and ever higher precision so as to verify new ideas and possibilities for extending the General Theory. New time transfer technology for conducting new tests along with new clocks for space operation provide substantially improved timing accuracy and frequency stability for these tests.

2. PRESENT STATE OF CLOCK TECHNOLOGY AND SYSTEMS FOR SPACE EXPERIMENTS

Operation in the near-zero "g" of a low earth orbiting spacecraft such as the International Space Station (ISS) enables operation of extremely high stability oscillators and frequency discriminators. At present, timing stability of 8.6 picosec per day (and frequency stability of 1 part in 10^{16}) is possible. In the case of cesium clocks a comparable level of frequency accuracy is now expected from a space borne cesium "fountain" frequency discriminator.

Figure 1 shows the expected $\sigma_{y(\tau)}$ versus τ plot for a space borne H-maser controlled by a space fountain system[1,2] The approximation for obtaining time prediction, $\sigma_{\Delta\tau}(\tau) \sim \tau\sigma_y(\tau)$ is used for estimating range measurement accuracy. Limits to the precision of range-rate and range distance measurements are shown.

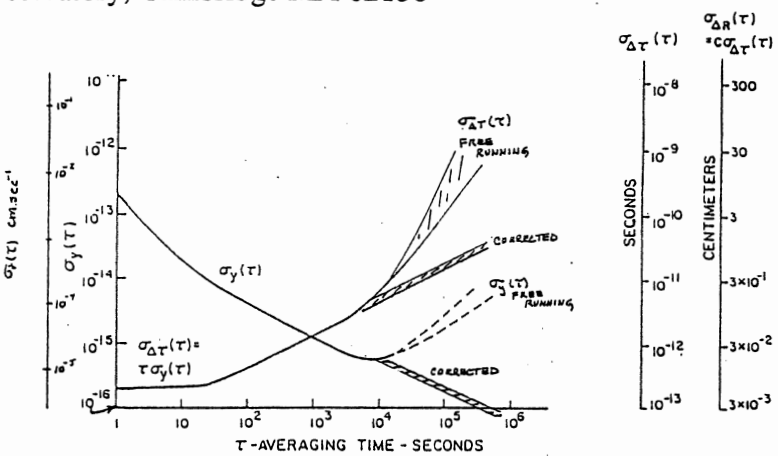


Figure 1 Measurement Capability of a Spaceborne Cesium Fountain / H-Maser System

3. THE 1976 GRAVITATION REDSHIFT EXPERIMENT AND WHAT WE LEARNED FROM IT

Figure 2 schematically describes the phase-coherent S-band Doppler canceling system used to compare the earth and space clocks in the 1976 SAO-NASA test [3] of the gravitational red shift.

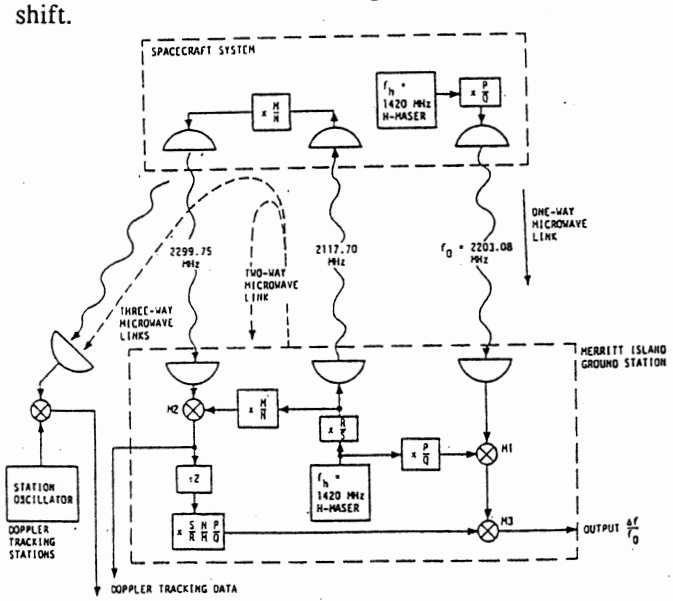


Figure 2 Schematic of the 1976 GP-A Doppler Canceling system

Effects of range-rate (first-order Doppler), ionospheric and tropospheric delay were systematically removed in real time by subtracting one-half of the two-way Doppler phase from the one-way phase received by the earth station. The gravitationally and relativistically induced fractional frequency variations (expanded to order c^{-2}) at the output from mixer M3 in Figure 2 are:

$$\frac{f_s-f_e}{f_0} = \frac{(\phi_s-\phi_e)}{c^2} - \frac{|\vec{v}_e-\vec{v}_s|^2}{2c^2} - \frac{\vec{r}_{se}\cdot\vec{a}_e}{c^2} \qquad (1)$$

Here the total frequency shift is f_s-f_e, and f_0 is the clock downlink frequency, $\vec{r}_{se}$ is the vector distance between the spacecraft and earth station, and $\vec{a}_e$ is the acceleration of the earth station owing to the earth's rotation in an inertial frame. $(\phi_s-\phi_e)$ is the gravitational potential between the two clocks. The term $\frac{|\vec{v}_e-\vec{v}_s|^2}{2c^2}$, is from the second-order Doppler effect of special relativity from the relative velocity of the two clocks. The term $\frac{\vec{r}_{se}\cdot\vec{a}_e}{c^2}$ is the result of the acceleration of the earth station.

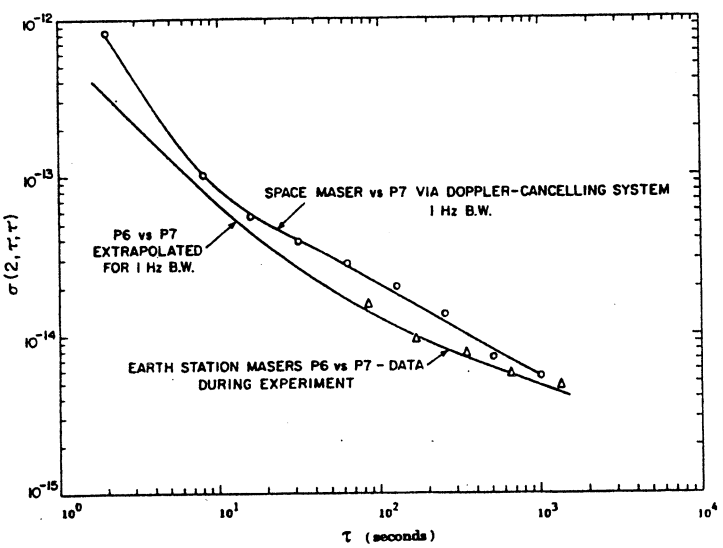


Figure 3 Stability of Doppler Canceled Signals After Removal of Relativistic effects

Predicted frequency variations obtained from tracking information, were fitted to the data. The error in the fit of the data was within $(+2.5 \pm 70) \times 10^{-6}$ of Einstein's prediction. Figure 3 shows the Allan deviation of the frequency residuals after removing the effects in Equation 1.[4]

The stability of the frequency comparison made through the three-link system over signal paths of 10,000 km, in the presence of Doppler shifts of magnitude ±44 kHz, plus the ionospheric and tropospheric noise, is comparable to the frequency comparison made between the two reference masers in the same room.

3.1 A Symmetrical Four-Link System to Provide Time-Correlated Doppler Data at Both ends of the System

By phase coherently transponding the clock downlink signal back to the spacecraft as shown in Figure 4, the system can be made symmetrical. One-way, two-way, and Doppler-canceled data can be received at both stations of the system and can be time correlated. When the light time between stations is long compared to the intervals required for measurements a spatially localized noise process can be canceled *systematically* by time correlation of the data from the two stations. Figure 5 shows the continuum of space-time paths of the four signals in Figure 4.

Here $E_1(t)$ and $E_2(t)$, signify signal outputs representing earth-based one-and two-way data. $S_1(t)$ and $S_2(t)$ represent one- and two-way data recorded in space. By time-correlating the Doppler responses we can systematically cancel a spatially localized noise source such as from the earth's troposphere and ionosphere.[5]

The signal outputs $S_0(t)$ and $E_0(t)$ in Figure 4 contain relativistic and gravitational information that can also be time-correlated.[6] The counterpart to Equation 4 from the space system is shown below as $S_0(t)$.

$$S_0(t) = \frac{f_e-f_s}{f_0} = \frac{\phi_e-\phi_s}{c^2} - \frac{|\vec{v}_e-\vec{v}_s|^2}{2c^2} - \frac{\vec{r}_{es}\cdot\vec{a}_s}{c^2} \qquad (2)$$

By adding these two time-ordered data sets we can cancel the first term and double the magnitude of the second term representing the second-order shift. Conversely, if we subtract the data sets, we double the first-term representing the gravitational shift and cancel the second term.

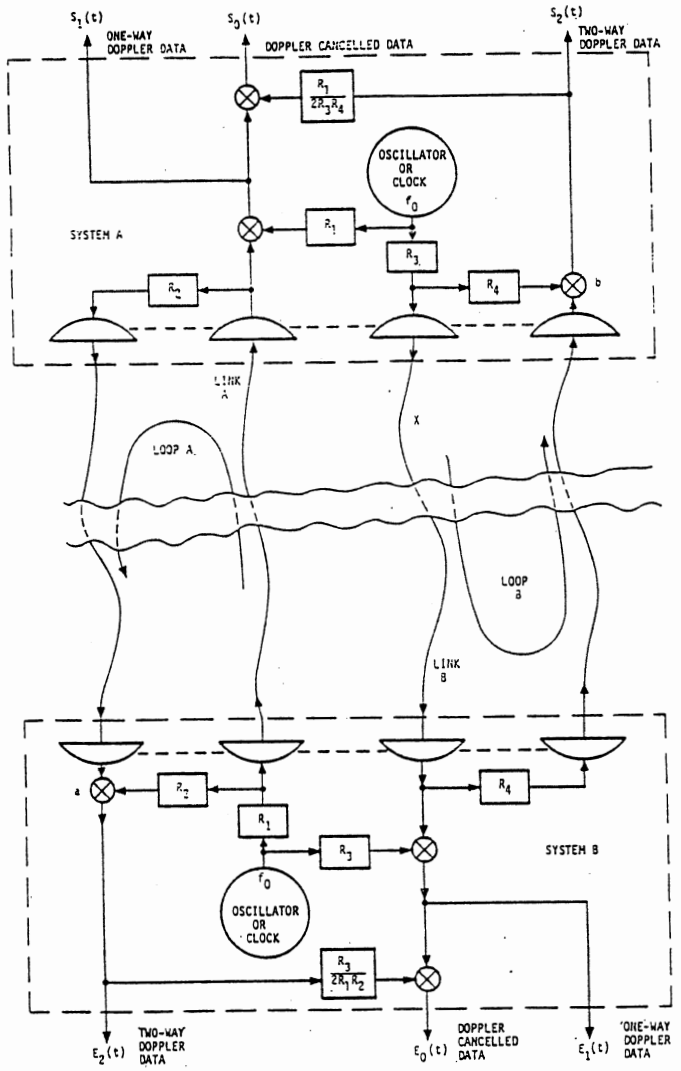


Figure 4 Four-Link Microwave System

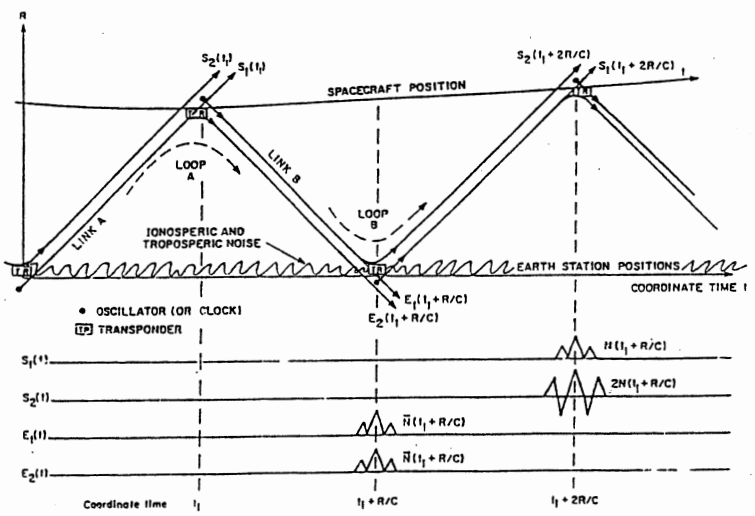


Figure 5 Space-Time Diagram of Four link Signals

4 THE EINSTEIN EQUIVALENCE PRINCIPLE

The Einstein Equivalence Principle includes two additions to the original Newtonian (or weak) Equivalence Principle stating that, regardless of their composition, all falling bodies have the same trajectories. These two additions are: *Local Lorentz Invariance* (LLI).

The outcome of tests of the velocity of light will be the same for unaccelerated systems travelling at any speed and performed anywhere and at any time. And

Local Position Invariance (LPI).

All tests of physics will give the same result no matter where or when they are done. Similar clocks compared at different gravitational potential must show the gravitational redshift, $\Delta\phi/c^2$. The relative frequency between different types of clocks when moving together through a varying gravitational potential must not change.

Clocks and oscillators play very strong roles in verifying Local Lorentz Invariance and Local Position Invariance.

5. TESTS OF LOCAL POSITION INVARIANCE AND LOCAL LORENTZ INVARIANCE IN LOW EARTH ORBIT

5.1 High Accuracy Frequency Discriminators.

Space versions of the terrestrial "Fountain" frequency discriminator are under development for operation on the ISS by the National Institute of Standards and Technology (NIST) for the NASA/PARCS program and by the University of Paris for the ESA/ACES program. Both systems will control the frequency of a spaceborne H-Maser.

The 1 part in 10^{16} accuracy from the combined fountain/maser, when transmitted to earth, will require frequency corrections for the effects shown in Equation 1 (with additional terms for $1/c^3$ effects). The second-order doppler shift is about - 3.3 x 10^{-10}, the gravitational redshift is about +4 x 10^{-11}. Tracking with a precision of ~ 0.5 mm/sec in velocity and of position within 0.5 m the corrections will provide predictions for the relativistic shifts with the required precision. The expected agreement with general relativity will be 100 times better test of General Relativity (LPI) than was done by the 1976 GP-A test.

5.2 High Stability Oscillators for Space operation

Recently, NASA has approved the development of a high stability superconducting microwave cavity stabilized oscillator (SUMO) for operation in the Low Temperature Microgravity Physics Facility to be installed on the International Space Station. Operating in near-zero g the frequency stability this oscillator is substantially better than that of the Hydrogen Maser for averaging intervals below 1000 sec.

By comparing the frequency of two such microwave cavity stabilized oscillators, whose resonator axes are at right angles to each other, greatly improved tests of the LLI can be performed. The Kennedy-Thorndike (K-T) test for a possible velocity dependent preferred reference frame can be improved by a factor of 10,000 and a factor of 870 improvement in a Michaelson-Morely test for a possible position dependent reference frame.

If operated in conjunction with the atomic frequency discriminators such as PARCS or PHARAO substantial further improvement in LPI and K-T could be achieved.

A Space Hydrogen Maser (SHM) is under development by the Observatoire de Neuchatel. The SHM has substantially reduced size and weight by operating with a dielectrically loaded single crystal sapphire cavity resonator. The prototype has successfully met the stringent size, weight and power constraints for operating in the ESA Atomic Clock Ensemble in Space (ACES) program as the oscillator to be controlled by the cesium frequency discriminator, PHARAO.

The H-Maser developed by the Smithsonian Astrophysical Observatory for the 1976 Gravity Probe-A has been redesigned for long term use in space for operation on the European Recoverable Carrier, which was not reflown. In 1995 the system was redesigned for operation on the docking module of the Russian Mir station, and again, progress was stopped due to scheduling problems with Mir. There is renewed interest in HMC use for operating on the ISS as a local oscillator for the NIST/NASA PARCS program's frequency discriminator.

6. TESTS OF RELATIVISTIC GRAVITATION WITH CLOCKS IN A SOLAR PROBE

6.1 Extension of the GP-A Experiment to Test General Relativity with a Solar Probe

Tests of relativistic gravitation with a clock in a space probe approaching within 4 solar radii of the sun's center in a polar orbit have been proposed.[78]

During the 10 hours before and after perihelion, the value of ϕ/c^2 varies from 5.3 x 10^{-7} at perihelion to 2.0 x 10^{-7} at times ±10 hours from perihelion. Over the same time interval, the second-order red shift $[\phi/c^2]^2$, varies from 2.81 x 10^{-13} to 4 x 10^{-14}, as shown in Figure 6.

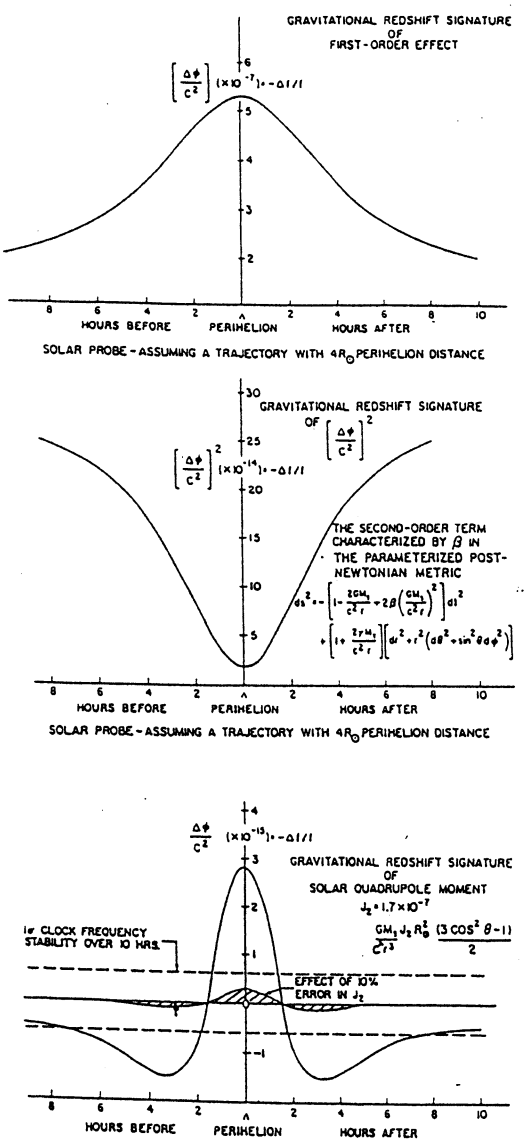


Figure 6 Relativistic Effects of the Sun's Gravity that Could be Observed by a Space Probe in a Polar Orbit Going Within 4 solar Radii

Taking the Allan standard deviation of the space clock over 10 hours averaging time as 6 x 10^{-16}, the accuracy of measurement for the first-order measurement of the gravitational redshift is 1.8 x 10^{-9}, the corresponding accuracy for the second-order measurement of the redshift is 2.5 x 10^{-3}.

The sun's gravitational potential has a number of multipole components, the largest of which is the quadrupole moment, J_2, which must be accounted for in the measurement of the second-order term in the red shift.

In Figure 5 we have shown how the localized noise near earth can be canceled by taking data at the probe. At the earth station, changes in the near earth propagation over the 1000 sec of time delay between the uplink transmission and reception from the transponder will lead to incomplete cancelation of propagation noise.

6.2 Solar probe experiment to measure differential redshift

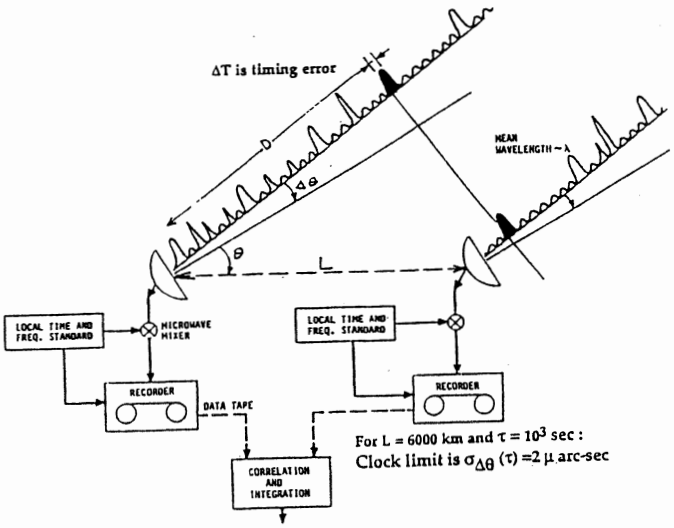
A test of Local Position Invariance can be done by comparing the hyperfine separation frequency of different atoms or ions as they travel through the gravitational potential encountered by a probe going close to the sun. A test has been proposed involving three linear ion traps operating with different types of ions.[9]

Another test would involve comparison with an oscillator whose frequency depends on the dimensions of resonator also can set limits on the gravitational dependence of the fine structure constant that governs these dimensions[10].

These tests do not require the Doppler cancellation system, only the telemetering of the changes in relative frequency differences between these clocks..

7. HIGH RESOLUTION VERY LONG BASELINE INTERFEROMETRY (VLBI) ASTRONOMY IN SPACE

High precision measurement of angle can be made with Very Long Baseline Interferometry. [11,12], Figure 7 shows a schematic of VLBI where two radio telescopes, separated by a distance, L, each detect the arrival of radio noise signals from a distant radio star. The noise signals are recorded as a function of time and the two data sets are time-correlated to obtain the angle between the baseline, L, and the radio star. VLBI measurements have been used in light deflection tests of relativistic gravitation.[13]



Angle measurement $\theta = \cos^{-1}(cT/L)$

Figure 7 Schematic of VLBI with Radio Telescopes

The effect of oscillator instability on successive measurements of angle taken t seconds apart is

$$\sigma_{\Delta\theta}(\tau) \sim \frac{c\tau\sigma_y(\tau)}{L \sin\theta} \qquad (4)$$

where θ is the angle between the propagation vector and the baseline. The spacing between the fringes is $\lambda/L \sin\theta$, where λ is the wavelength of the signals at the antennas.

The error in successive angular measurements owing to the instability of the clocks in a terrestrial system with L = 6000 km, assuming σ_y (10^3 sec) = 1 x 10^{-15}, and $\theta = \pi/2$, is given by $\sigma_{\Delta\theta}$ (10^3 sec) = 5 x 10^{-11} rad or 2 m arcsec.

The present limit is 100 m arcsec level from terrestrial stations operating at 7 mm wavelength[14] with an 8000-km baseline because of tropospheric and ionospheric fluctuations.

Operating VLBI stations in space removes limits in angular resolution and baseline distance.

In 1986[15] NASA's orbital Tracking and Data Relay Satellite System (TDRSS) was successfully operated as a space borne radio telescope with a number of radio telescopes on earth. Today, the Japanese HALCA space borne radio telescope is in successful operation with a number of US tracking stations.[16]

7.1 A Space Borne Four Terminal (VLBI) Array that Establishes its Own Inertial Reference Frame.

Let us consider a space borne system where L = 5 x10^6 km, σ_y (10^4 sec) = 4 x 10^{-16}, and θ = $\pi/2$. In this case $\sigma_{\Delta\theta}$ (10^4sec) = 2 x 10^{-13} rad or 0.05 m arcsec. For λ = 1 mm we have λ/L = 2 x 10^{-13} rad and we see that the limit imposed by clock stability with 10^4 sec integration time is capable of resolving fringes at 1-mm wavelengths in a space borne system with baseline distances of 5 x 10^6 km.

While *relative* angular positions of radio stars, and features of their brightness distribution, can be made with high precision, the *absolute spatial orientations* of the baselines between stations depend on the choice of a reference frame. In earth based VLBI, the earth's orientation is defined from the position of very distant radio sources.

Figure 8. shows a three-dimensional array of four stations, separated by about 5 x 10^6 kilometers in the form of a giant tetrahedron located. I would be located at one of the Lambda point in the earth's orbit.[17] Each station contains a clock and is connected to its three neighbors by the four link system shown in Figure 4. Using one-way and two-way Doppler techniques with time modulation, the six baseline distances can be precisely measured to define the *configuration* of the array as a function of time.

We can monitor changes in the *orientation* of this array using the Sagnac effect[18], which is the basis for laser gyroscopes. Let the area of each of the triangular faces of the tetrahedron be A_i where (i=1,2,3,4). By observing the difference in arrival times, $\Delta\tau$, of light signals sent in opposite senses about the closed path about each A_i, we can determine the rotation vector

$$\Omega_i(\text{rad/sec}) = \frac{\Delta\tau\, c^2}{4A} \text{ perpendiculat to surface, } A_i.$$

From the components of rotation measured from the four triangles that define the tetrahedron we have an over-determination of the rotation of the whole array and obtain an estimate of the accuracy of the measurement.

For the array shown in Figure 8, over about 100 sec of averaging, the accuracy of the detection of changes in rotation rate, $\Delta\Omega$, is 1.2 x 10^{-15} rad/sec. Located at 1 AU from the sun this system would test the Einstein–deSitter precession of 3 x 10^{-15} rad/sec owing to the bending of space-time by the sun's gravity and this precession must be accounted for.

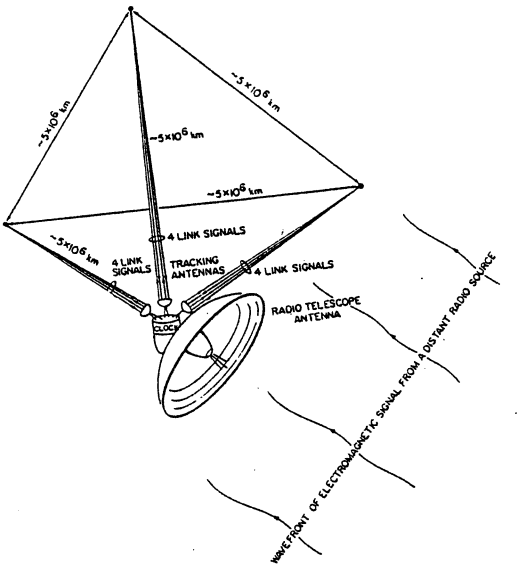


Figure 8 A four station array of radiotelescopes connected with the 4 link system

Distant radio stars define the conventionally used inertial frame. By invoking the Sagnac effect,we can define an inertial frame based on the velocity of light in which to determine *changes* in the orientation of the array with reference to this inertial frame.

This array could enable us to compare an inertial frame defined by the positions of very distant radio sources with an inertial frame locally defined by the velocity of light.

Could this system reveal some aspects of the behavior of the "missing matter" in the universe?

8. CONCLUSION

Since the mid 1960s the frequency stability of atomic clocks (or oscillators) continues to improve by a factor of about ten every decade. The highest levels of clock accuracy and frequency stability are now available by spaceborne operation.

We now make measurements of astronomical and astrophysical objects near the edge of our universe with clocks that operate using quantum processes whose dimensions encompass staggeringly smaller scales of distance. As tests with atomic clocks are extended into space, it is clear to the writer that we can expect to see some surprises about the nature of our universe.

References

[1]Leo Holberg et al., "Primary Atomic Reference clock in Space (PARCS)." Proc.of the 1999 Joint meeting of the EFTF and International Frequency Control Symposium Besancon, France 13-16 April, 1999 pp 141-147.

[2]S Feltham, F Gonzalez and P. Puech, "Atomic Clock Ensemble in Space. "Proc.of the 1999 Joint meeting of the EFTF and International Frequency Control Symposium Besancon, France 13-16 April, 1999 pp 148-151

[3]R.F.C. Vessot, M.W. Levine, et al."Tests of relativistic gravitation with a space-borne hydrogen maser," *Phys. Rev. Lett.*, vol. 45, pp. 2081-2084, Dec. 1980.

[4]R.F.C. Vessot and M.W. Levine, "Performance data of space and ground hydrogen masers and ionospheric studies for high-accuracy frequency comparison between space and ground clocks," in Proc. Twenty-eighth Annual.Symp. Frequency Control, U.S. Army Electronics Command, Ft. Monmouth, NJ, pp. 408-414,1974.

[5]R.F.C.Vessot and M.W. Levine, "A time-correlated four-link Doppler tracking system," in A Closeup of the Sun (M. Neugebauer, and R.W.Davies, Eds.) JPL Publication 78-70, NASA, 1978.

[6]L.L. Smarr. R.F.C. Vessot, C.A. Lundquist, R. Decher, and T. Piran, "Gravitational waves and redshifts: A space experiment for testing relativistic gravity using multiple time-correlated radio signals," *Gen. Rel. and Grav.* vol. 15, no. 2, pp. 129-163, 1983.

[7]R.F.C. Vessot, "Clocks and spaceborne tests of relativistic gravitation," in Advances in Space Research, Eds. (R.D. Reasenberg and R.F.C. Vessot Eds.) Pergammon Press, vol.9, No. 9, 1989.

[8] Kenneth Nordtvedt and RobertVessot "The Clock(s)- to-the- Sun Mission: A New Look in Proceedings of the 8th Marcel Grossman meeting on General relativity, June 22-27,1997: Jerusalem, Israel.

[9]L Maleki et al.,"Relativity at Four Solar Radii" Proceedings of the 1998 IEEE International Frequency Control Symposium May 27-29 1998, Pasadena CA. pp.-329-335

[10]J.D. Turneaure, C.M. Will, B.F. Farrell, E.M. Mattison, and R.F.C. Vessot, "Tests of the principle of equivalence by a null gravitational redshift experiment."Phys. Rev. D. Vol 27, No. 8, 15 April 1983; p. 1009.

[11]A.E.E. Rogers and J.M. Moran, Chapter 5, "Interferometers and Arrays," Chapter 5 in Methods of Experimental Physics, Astrophysics, M.L. Meeks, Ed., vol. 12. New York: Academic Press, 1976.

[12]A.E.E. Rogers, A.T. Moffet, D.C. Backer, and J.M. Moran, "Coherence limits in VLBI observations at 3-millimeter wavelength," Radio Sci., vol. 19, no. 6, pp, 1552-1560, 1984.

[13]E. B. Formalont and R.A. Sramek, "Measurement of the solar gravitational deflection of radiowaves in agreement with general relativity," Phys. Rev. Lett., vol. 36, p. 1475, 1976.

[14] N. Bartel et al, "VLBI imaging with an angular resolution of 100 micro arc seconds,"Nature, vol. 334, no. 6178, pp. 131-135, 1988.

[15] G.S. Levy, et al. "Results and communications considerations of the very long baseline interferometry demonstration using the tracking and data relay satellite system," Acta Astron., vol. 15, no. 6/7, pp. 481-487, 1986.

16 J Ulvestad "VSOP Coherence Measured at the VLBA Correlator," National Radio Astronomy Observatory, technical memorandum,,25 January 1998.

[17] R.F.C. Vessot, "Clocks and spaceborne tests of relativistic gravitation," inAdvances in Space Research, Eds. (R.D. Reasenberg and R.F.C. Vessot Eds.) Pergammon Press, vol.9, No. 9, 1989.

[18] E.J. Post, "Sagnac Effect," Rev. Mod. Phys., vol. 39, p. 475, 1967.

NEW MATERIALS AND NEW RESONATORS

G.D.MANSFELD
Institute of Radioengineering and Electronics, RAS,
Mokhovaya 11, 103907, Moscow, Russia

The first problem discussed in this review is the evaluation of the acoustic parameters of new acoustic materials perspective for the use in acoustic devices mainly in bulk acoustic wave (BAW) resonators. The advantages and disadvantages of four groups of materials are discussed:

-traditional crystals with quartz like symmetry - berlinite, gallium orthophosphate, langasite family crystals;

-group of strong piezoelectric crystals like lithium tantalate ;

-low acoustic waves propagation loss piezoelectric and non-piezoelectric materials like solid solutions based on crystalline matrix of yttrium aluminum garnets, some spinels;

- thin dielectric, metal and piezoelectric films.

The physical mechanisms of acoustic losses and ways of their decrease are discussed.

The second problem discussed in this review is the use of new materials in various types of RF and microwave BAW resonators including high overtone microwave resonators, membrane resonators, stacked structures and solidly mounted acoustically isolated (quarter wavelength multilayer acoustic mirror) thin film resonators.

Some new results of the use of BAW resonators in sensors are analyzed. In particular high sensitivity gas sensors with close to quarter wavelength thickness gas selective layer is discussed.

G.D.MANSFELD
Institute of Radioengineering and Electronics, RAS,
Mokhovaya 11, 103907, Moscow, Russia
TEL 7(095) 203 50 90
FAX 7(095) 203 50 90
e-mail mans@mail.cplire.ru

GALLIUM PHOSPHATE, $GaPO_4$: CRYSTAL GROWTH AND CHARACTERIZATIONS.

P. Yot*, O. Cambon*, D. Balitsky*, A. Goiffon*, E. Philippot* , B. Capelle° & J. Détaint°.
LPMC, CNRS, Université de Montpellier II & ° LMCP, Paris VI

Corresponding author: Philippot, LPMC, UM II, cc 003, Place E. Bataillon, 34095, Montpellier cedex 5
Tel : 33 4 67 14 33 19 Fax : 33 4 67 14 42 90 E-mail : ephilip@lpmc.univ-montp2.fr

ABSTRACT

The investigation of the gallium phosphate is justified for at least two reasons. Firstly, for the same AT compensated cut, the value of its coupling coefficient is higher than that of quartz. This suggests the possibility of obtaining filters of larger pass-bands. Secondly, there is no α↔β quartz transition for this material, so it can be used over a larger range of temperature (cristobalite transition equals 933°C), for filter and sensor devices. We shall present our crystal growth approach with some new results of infrared spectrometry, X-ray topography investigations and resonator measurements.

1. INTRODUCTION

Since the work of Chang and Barsch [1], many studies have been devoted to growing and characterizing quartz-like materials of general formula $M^{III}X^{V}O_4$ (M=Al, Ga; X=P, As), such as $AlPO_4$ and $GaPO_4$, and to the modeling of bulk and surface acoustic waves devices made with these materials. It was successively shown that these materials are similar to quartz and possess analogous cuts with zero frequency temperature coefficients, either for bulk or surface wave applications [2].

The investigation of the gallium phosphate growth is justified for at least two reasons. Firstly, for the same AT compensated cut, the value of its coupling coefficient is higher than that of quartz [2]. This suggests the possibility of obtaining filters of larger pass-bands. Secondly, there is no α-β quartz transition for this material, so it can be used over a larger range of temperature (cristobalite transition equals 933°C), for filter and sensor devices. We shall present our crystal growth approach with some new results of infrared spectrometry, X-ray investigations and resonator measurements.

2. SOLUBILITY INVESTIGATION

The present knowledge of the solubility of $GaPO_4$ in different acids being contradictory, a new investigation, using two complementary methods, has been carried out in phosphoric, sulfuric acid media and their mixtures. From a general point of view, the solubility of $GaPO_4$ in acid media exhibits a retrograde behavior, i.e. the solubility decreases when the temperature is raised, at "low" temperature (T<300°C) and "low" pressure (P<2MPa). In more drastic conditions, the solubility becomes direct. All our investigations were performed in the first range of temperature and pressure. Some results of this new investigation are reported in figure 1.

Conversely to berlinite, $AlPO_4$, the solubility of $GaPO_4$ in the acid mixtures is not the "mean value" between the both acids taken separately. The solubility decreases strongly by adding small quantities of H_3PO_4 (10%) in H_2SO_4 medium and, when the temperature of crystallization increases, the solubility of $GaPO_4$ in the mixtures converges to the solubility in phosphoric acid.

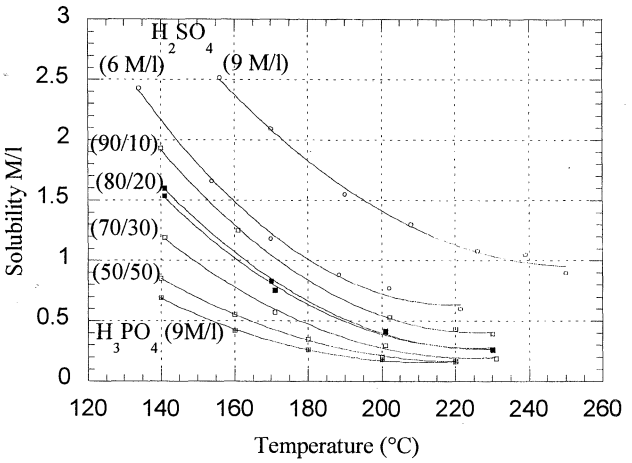


Figure 1 – Solubility of $GaPO_4$ in H_2SO_4 6 and 9 M/l, H_3PO_4 9M/l and in the mixtures H_2SO_4 6 M/l - H_3PO_4 9 M/l (90/10, 80/20, 70/30 and 50/50)

Strong divergences often appear with the classical "single crystal dissolution" method for high acid concentrations and low temperatures. These differences are too important to be due to an experimental inaccuracy of the solubility methods. Indeed, the single crystal dissolution method leads to systematic lower values of the solubility, figure 2. In fact, the region between the two curves is the metastable zone where no crystallization/dissolution process appears. Its accurate knowledge is essential to choose the good crystal growth restart conditions and to avoid seed dissolution (at lower temperature) or nucleations at the seed interface (at higher temperature).

From this accurate knowledge of the different solubility curves, growth restart parameters are chosen to be in the metastable zone and, then, avoid any defect in the first part of the growth which determines the final crystal quality.

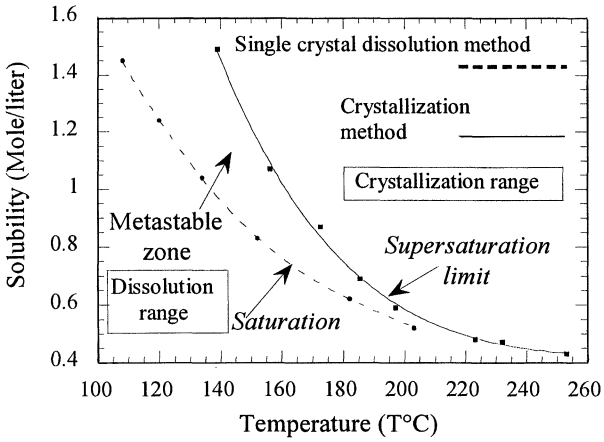


Figure 2 – Compared solubility curves of $GaPO_4$ in H_3PO_4 15 M/l giving the metastable zone.

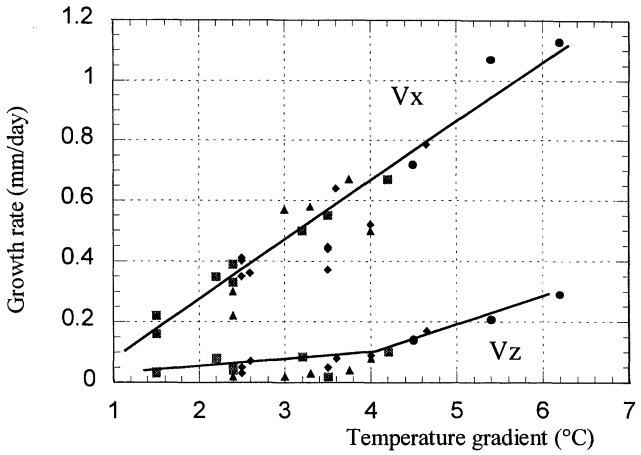


Figure 3 – V_x and V_z growth rates in H_2SO_4 6M/l - H_3PO_4 9M/l medium at 175<Tc<180°C.
(▲ (90/10), ♦ (80/20), ■ (70/30) and • (50/50) mixtures).

3. CRYSTAL GROWTH STUDIES

From a general point of view, different problems, due to contradictory effect of parameters (growth temperature, concentration,...), had to be solved to lead to a reproducible growth of high quality $GaPO_4$ crystals. The improvement of the crystal quality can be evaluated through the OH content by infrared spectroscopy and by X-ray topography.

Two different investigations were performed in order to improve the crystal quality and to lengthen the seed dimension along Y-axis (no natural crystal being available).Until last year, the only possibility of obtaining good $GaPO_4$ crystals through the vertical gradient method in autoclaves [3] was from Z-seeds in phosphoric acid medium with a very low growth rate (V_z < 0.04 mm/day). All our investigations on $GaPO_4$ crystal growth from X-seeds in vertical autoclaves led to disturbed growth giving "*H-crystals*" [4].

3.1 Optimal growth parameters for X-seeds

After many experiments, in different solvents and different growth conditions, the best reproducible quality from X-seeds is obtained in the H_2SO_4 6M/l - H_3PO_4 9M/l mixtures with a sufficient supply of solute at the interface crystal-solution, i.e. a sufficient temperature gradient for a growth temperature and a solvent given.

The different Vx and Vz growth rates are summarized in figure 3. The dispersal of the growth rates, specially for Vx rates, is due to the morphology of the crystals obtained. Almost all of them ended on one side, in the +X direction, exhibit underestimated values. The main feature of this investigation is the growth rates which increase with the temperature gradient, independently of the sulfuric/phosphoric acid ratio. Two examples of crystals obtained in these conditions are showed in figure 4.

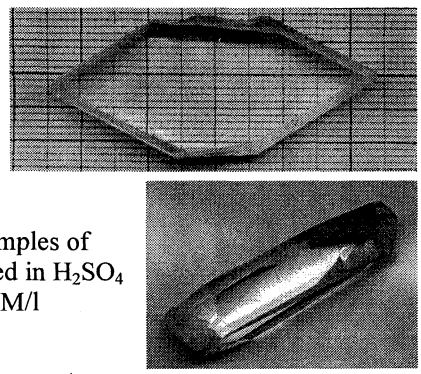


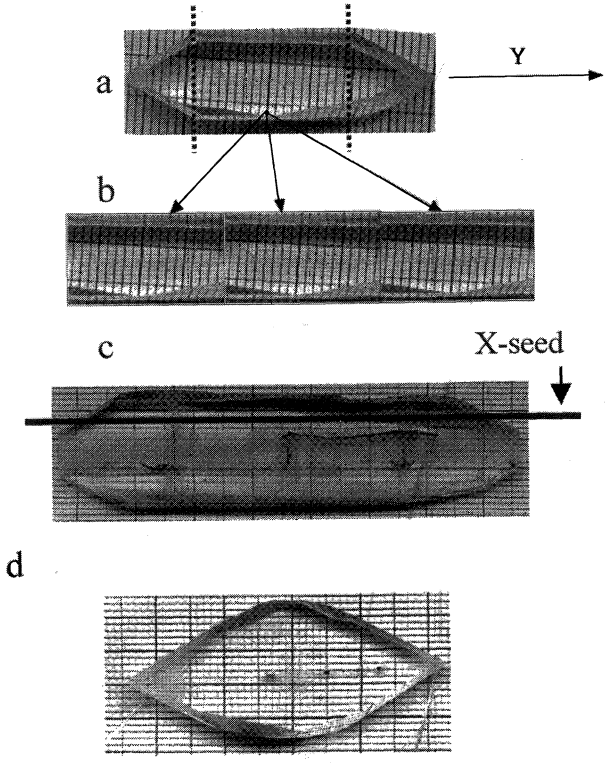
Figure 4 – Examples of crystals obtained in H_2SO_4 6M/l - H_3PO_4 9M/l mixtures.

3.2 Seed lengthening by splicing

Since V_y growth rate is always less than 0.01 mm/d, large crystals cannot be obtained by successive growth of nucleations. This lack of large seeds of $GaPO_4$ requires to develop the lengthening of the crystals by the crystal growth epitaxy on α-quartz [5] and $AlPO_4$ [6] seeds or by the splicing method, which we have preferred to investigate .

In this way, the growth of large crystals was carried out from several small Z (or X) $GaPO_4$ seeds, strictly in line with the Y-direction, spliced on a quartz support. In the beginning, well defined crystals with natural parallel faces were used [6]. This kind of crystals being relatively rare, a new method was perfected from any kind of crystals, figures 5. Firstly, figure 5a, the crystal is accurately sawn perpendicular to the Y-direction, in order to eliminate the ends, afterwards, in plates perpendicular to the Z (or X) direction. Then, these small seeds are spliced on quartz plates, figure 5b. Crystal growth experiments are carried out in glass vessels through a modified horizontal composite temperature gradient method (HCTG). The resulting crystals show a very flat surface without any obvious defects (such as twins) at the interfaces between small seeds, figure 5c. Nevertheless, after this first

crystal growth, the X-ray topographs revealed the importance of very high dislocation density close to the seed junction. These defects can be easily and drastically reduced using the "crossed" crystal growth method. For example, if the initial seed is a Z-oriented one, the resulting crystal must be sawn in X-cuts before a new growth experiment, figure 5d, and so on.



Figures 5: Principle of the splicing method used to lengthen crystals in the Y-direction and crystal issued from the splicing method.

So, the combination of seed lengthening followed by crossed crystal growth allows to obtain good crystals with reasonable size (≈ 8 cm).

4. CHARACTERIZATIONS

Simultaneously, characterizations of crystals have been undertaken to check:

- the crystal morphology;
- the OH content by infrared spectroscopy;
- the crystalline quality by X-ray topography;
- their piezoelectric characteristics.

4.1 Morphology characterization

If the main developed faces in $GaPO_4$ crystals are m, R, r, π & π', two kinds of other orientations are observed at low temperature (Tc<170°C). First, in phosphoric acid medium, the crystals are ended by the faces $\{3,\bar{1},\bar{2},3\}$ & $\{3,\bar{2},\bar{1},\bar{3}\}$ faces in the +X direction. Conversely, in the phosphoric/sulfuric mixtures, the crystals, always along the +X-direction, are ended by the $\{2,\bar{1},\bar{1},1\}$ & $\{2,\bar{1},\bar{1},\bar{1}\}$ faces.

4.2 OH content determination by infrared spectroscopy

The OH content in crystal is favored by the hydrothermal growth process in acid media and influences the piezoelectric characteristics, specially the Q-factors of the resonators. In the absorption range characteristic of OH groups, the infrared spectrum of $GaPO_4$ exhibits 3 bands at 3400, 3290 and 3170 cm^{-1} which are intrinsic to the material [7].

For the OH content, we use the transmittance value at 3400 cm^{-1}, in agreement with the Krempl & al. relation [7]:

$$\alpha_{3400} = \left(\frac{1}{d_{cm}}\right) * \left(\log \frac{T_{3800}}{T_{3400}}\right) - \alpha *$$

where α^* is the absorbance at 3400 cm^{-1} which is not due to OH vibrations (α^*=0.078cm^{-1}).

From our general investigation, three main features can be pointed out. If all the other parameters are constant, the OH content decreases with the acid concentration of the solvent, figure 6, with the growth rate, and when the growth temperature rises, figure 7.

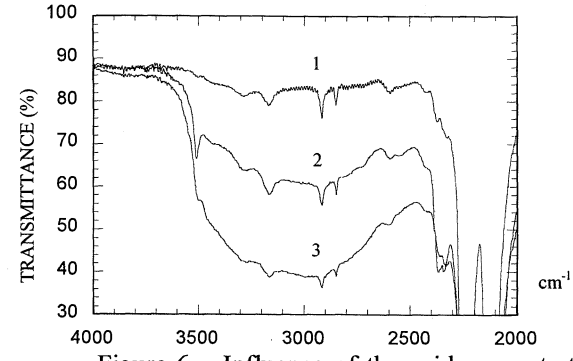


Figure 6 – Influence of the acid concentration on the OH content (the bands at 2920 and 2850 cm^{-1} are due to finger prints). (H_3PO_4 7.5 (1), 9 (2) and 15 (3) M/l at Tc = 160°C).

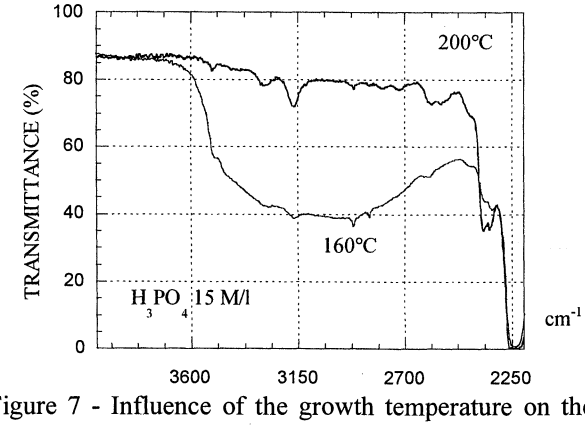


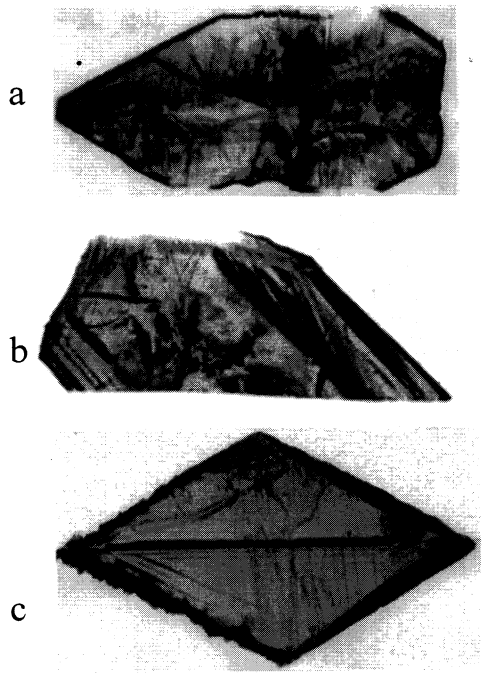
Figure 7 - Influence of the growth temperature on the OH content.

4.3 Characterization of growth defects by X-ray topography.

In the crystals, the most important kinds of defects are dislocations and twins which can be detected by X-ray topography. This study was carried out at the

synchrotron X-ray source of LURE, Orsay, France. Of course, the results are dependent on the seed quality, but also on the growth restart conditions.

The next example, figures 8, gives an idea of the crystal improvement by crossed crystal growth after splicing of Z-seeds. The figure 8a, corresponding to a Z-slice parallel to the initial spliced Z-seeds, exhibits a high density of dislocations due to the splicing. In the figure 8b, X-slice from crystal obtained with the previous Z-seed, the crystal quality becomes better (with growth bands which are often met in this material) and the last one, figure 8c, Z-slice obtained with the previous X-seed shows only some discrete dislocations.



Figures 8 – X-ray topographies obtained at different steps of the splicing process.

4.4 – Piezoelectric properties

AT compensated cut

An initial investigation on slices of crystals obtained at high temperature (240°C) allowed to find a more accurate angle value of the AT compensated cut at room temperature (zero first order temperature coefficient of the resonance frequency of the fundamental mode) Y-15°46' [2]. It was observed that the thermal variations of the resonance frequency for this cut are characterized by a very low value of the second order temperature coefficient of the resonance frequency (-8 to -10.e^{-9}°C^{-2}). Moreover, this cut displays also an angular sensitivity which is smaller than those observed for the AT cut of quartz and berlinite. A shift of the cut angle of -25' produces only an increase of the inversion temperature of about 40°C. Resonators made with plates having orientations situated near Y-13° display inversion temperatures in the range 130-150°C, they can be very useful for high temperature sensor applications.

Lateral field excitation of shear modes in Z plates.

The interest of this type of measurement is to give the unmetallized resonator performances excited by a lateral field in the aim to determine the elastic constant. For any plane resonator excited by a lateral field, the resonance frequencies are equal to the antiresonance frequencies of the same resonator with a thickness excitation. The relation is:

$$Fr(\text{fund.mode})_{\text{long.excit.}} = Fa(\text{fund.mode})_{\text{thick.excit.}}$$

$$= \frac{1}{2.2h}\sqrt{\frac{\hat{C}}{\rho}} \quad \text{with} \quad \hat{C} = \overline{C}\ (\text{unmetall.}) = C_{44} = C_{55}$$

For this experiment, two plano-convexe Z resonators have been made. In this case, the plate curvature must be taken into account to recalculate the antiresonance frequency constant by the Tiersten theory:

$$V_{real} = 2.Kf_{measured(cor).} \approx \sqrt{\frac{C_{44}}{\rho}} \approx \frac{2\frac{KF_n}{n}}{\sqrt{1+\frac{1}{n\pi}\sqrt{\frac{2h}{Rc}}\left\{\sqrt{\frac{Mn}{\hat{C}_{pe}}}+\sqrt{\frac{Pn}{\hat{C}_{pe}}}\right\}}}$$

Where 2h = thickness (in the centre), Rc=curvature radius, Mn et Pn intricate functions of the constants, Cpe=Cpl stiffened (rotated) constant of the mode. V = velocity of plane wave in an infinite (piezoelectric) solid = 2Kfa.

From this correction, it is possible to extract the value of the elastic constant C_{44} for unmetallized plane plates.

The eigen modes of unmetallized plano-convexe plates are excited by a nearly in plane field. For this measurement, two plano-convexe resonators with a radius curvature of R=175mm were investigated. These plates, obtained from lateral berlinite epitaxy in H_2SO_4 6M/l at Tc=230°C, exhibit a low OH content ($0.03<\alpha<0.08$). The results for the fundamental mode are as good as for quartz, next Table.

Table – Main characteristics of both resonators (*Thickness accuracy = ± 1µm, frequency accuracy = a few ±100 Hz).*

	Q (at 5MHz)	$Fr_{exp.}$ (Hz)	$Fr_{calc.}$ K97	$KFa_{cor.}$	C_{44} 1.E10N/m^2
1	88450	2520878	2550528	1647.55	3.8762
2	91600	2748020	2808614	1630.96	3.7985

If we compare the experimental values for the sample N°1 with those calculated from different constant sets, the results show that all the recent values of C_{44} are similar. As predicted, the corrected values obtained are very nearly independent of the constant set chosen. The new value of C_{44} is 3.873 10^{10} N/m^2.

Comparison of experimental and calculated CTFR.

On the other hand, the comparison of experimental and calculated CTFR for different Y-rotated cuts seems to indicate that the best set of known constants is that given by Krempl & al. in 1998 [8], figure 9.

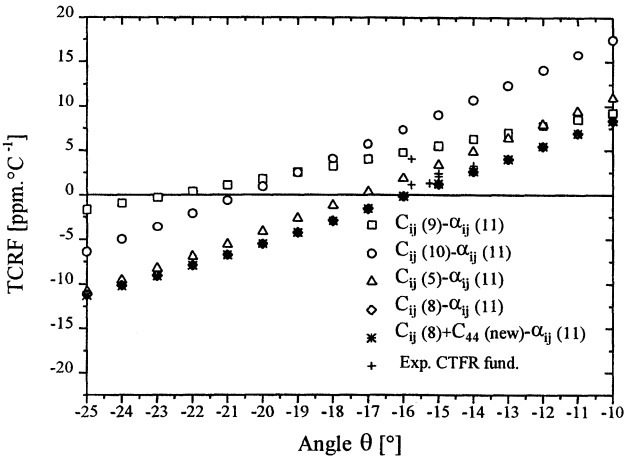


Figure 9 – Comparison of experimental and calculated TCRF.

Resonator characteristics

Although the objective was not to measure the ultimate value of the Q factor achievable at a given frequency, very interesting values were recorded during this study from plano-convexe resonators (curvature radius = 175 mm) of the order of 8 10^5 at 5 MHz for samples with a significant OH content (α=0.8cm^{-1}). On the other hand, conversely to quartz resonators, the Q-factor increases with temperature. So, for a plano-convexe resonator (R=175 mm, Y-15°30'), the Q-factor value at 5MHz for the 5th overtone which is 6.8 10^5 at room temperature rises to 1.1 10^6 at 118.5°C.

Another result of these investigations is the observation of a much lower impedance level as for similar quartz resonators (the inductances are nearly 5 times lower). This is very favorable for oscillators applications (smaller difference between the loaded and the unloaded Q factor for a given oscillator scheme) and also to obtain filters with more useful termination impedances.

5. CONCLUSION

The knowledge of the metastable zone from solubility investigations improves the restart growth conditions and, then, leads to better quality of gallium phosphate crystals. On the other hand, the determination of a necessary minimum supply quantity of solute at the interface crystal/solution allows to predict the crystal growth parameters and to obtain good crystals, even from X-seeds.

For the seed lengthening, the association of the splicing and the crossed crystal growth methods leads to gallium phosphate crystals of which the high quality was controlled by X-ray topographs.

For the piezoelectric properties, the AT cut was accurately specified at Y-15°46'. Even if the objective was not to measure the ultimate value of the Q factor, very interesting values of the order of 8 10^5 at 5 MHz were recorded during this study from plano-convexe resonators with a rather high α value (0.8 cm^{-1}). On the other hand, from lateral field excitation of shear modes in Z plates, it has been possible to specify a very accurate value of the C_{44} value.

Acknowledgements: The authors thank the European Community (Brite EuRam contract, BE96-3105, BRPR-CT97-391) for its financial support

REFERENCES

[1] Y.P. Chang and G.H. Barsch, Elastic constants and thermal expansion of berlinite, IEEE Trans. Sonics Ultrasonics SU-23, 1976, vol. 23(2), pp.127-135.

[2] J. Détaint, B. Capelle, A. Zarka, D. Palmier and E. Philippot, Frequency control and filtering applications of the new piezoelectric materials of class 32, Ann. Chim. Sci. Mat, 1997, vol. 22, pp. 651-660.

[3] E. Philippot, A. Ibanez, A. Goiffon, M. Cochez, A. Zarka, B. Capelle, J. Schwartzel and J. Détaint, J. Cryst. Growth, 1993, vol. 130, pp. 195-208.

[4] E. Philippot, A. Goiffon and A. Ibanez, Comparative crystal habit geometry of quartz and MPO_4 isomorphous compounds (M = Al,Ga), J. Cryst. Growth, 1996, vol. 160, pp. 268-278.

[5] P.W. Krempl, F. Krispel and W. Wallnöfer, Industrial development and prospects of $GaPO_4$, Ann. Chim. Sci. Mat., 1997, vol. 22, pp. 623-626.

[6] D. Palmier, A. Goiffon, B. Capelle, J. Détaint and E. Philippot, Crystal growth and characterizations of quartz-like material: $GaPO_4$, J. Cryst. Growth, 1996, vol. 166, pp. 347-353.

[7] P.W. Krempl, F. Krispel, W. Wallnöfer, and G. Leuprecht, $GaPO_4$: A critical review of material data, Proc. 9th Eur. Freq. Time Forum, 1995, pp.66-70.

[8] C. Reiter, H. Thanner, W. Wallnöfer and P.W. Krempl, Temperature dependence and compensation of $GaPO_4$ thickness resonators, Proc. 12th Eur. Freq. Time Forum, 1998, pp.447-450.

[9] D. Palmier, Thesis, University of Montpellier II, France, 1996, pp. 182-185.

[10] W. Wallnöfer and P. W. Krempl, Determination of the elastic and photoelastic constants of quartz-type $GaPO_4$ by Brillouin scattering, Phys. Rev. B, 1994, vol. 49(15), pp. 10075-10080.

[11] P.W. Krempl, Private communication, 1998.

COMPARAISON BETWEEN FINITE ELEMENTS AND EXPERIMENTAL ANALYSES OF THE FERROBIELASTIC BEHAVIOR OF THE QUARTZ CRYSTAL

J.J. BOY, K. YAMNI*,, A. YACOUBI****

ENSMM / LCEP - 26, chemin de l'Epitaphe - 25000 BESANÇON - FRANCE
* Faculté des Sciences et Techniques, LMCE, BP 146, 20650 MOHAMMEDIA - MAROC
** Faculté des Sciences, LCMA, BP 4010, 53000 MEKNES – MAROC

ABSTRACT :

To study the propagation of the ferrobielastic twinning in quartz crystal, finite elements analyses are performed on various parallelepiped X-cut samples. The pressure has been applied along Z'-axis rotated by ψ degrees towards the crystallographic X-axis. A small defect has been introduced in the geometry of the sample to clearly understand the propagation of the twin walls. The energy threshold, corresponding to the coercive stress, introduced in the analyses has been deduced from new experiments performed at ambient temperature on various samples cut in different quartz crystals.

We report that X-cut samples exhibit the same stripe-shaped patterns with walls parallel to the optical axis as in the experiments. We notice that the coercive stress depends strongly on the orientation of the applied pressure.

1. INTRODUCTION

As it has been already investigated, the application of quartz crystal in pressure and force transducers is drastically limited by electrical twinning due to severe environmental conditions (high stress and/or high temperature). Additionally, the nucleation of electrical twins under mechanical or thermal stresses can be observed during the manufacturing process of quartz resonators, used as well as transducers than "time bases".

The domain states of a ferroic crystal can be structurally reoriented by application of a suitable external driven force **[1]**. α-quartz is a secondary ferroic crystal in which the domains (two Dauphiné twin states) differ particularly by the sign of the s_{14} elastic coefficient **[2, 3]**. Due to this difference in elastic compliance tensor, a properly oriented stress acts to create a strain difference between the two states. Dauphiné twinning corresponds to a 180° rotation about the optical axis resulting in small atomic displacements.

In function of the principle of the lower energy, twinning occurs when the free energy of the alternate twin state is smaller than this of the existing state. The calculated energy difference between free energies of each twin state depends only on the stress direction and the crystallographic orientation of the sample **[2, 3]**.

2. MODELIZATION BY FINITE ELEMENTS METHOD

As it has been shown in **[7]**, the Gibbs' free energy function G is used to describe the ferrobielastic transformation when twin occurs. And the energy difference between the two twin states existing when the sample is submitted to an uniaxial homogeneous compressive stress P, acting along the Z"-direction, can be expressed as :

$$\Delta G = 2 \, . \, s_{14} \, . \, F(\varphi,\theta,\psi) \, . \, P^2$$

with : $s_{14} = 4.5 \, . \, 10^{-12}$ m^2/N and where $F(\varphi,\theta,\psi)$ depends only on the crystallographic orientation (φ,θ) and on the stress direction ψ. We note that this expression exhibits a quadratic variation in terms of P, but it does not contain information about the threshold stress which is the energy necessary to shift the atoms of the lattice from their initial positions (corresponding to the domain I) to the alternate positions (domain II) without breaking the interatomic bonds.

As we know, for X-cut (defined by $\varphi = 90°$ or $-30°$ and $\theta = 0$), twinning occurs when the uniaxial stress is applied in a direction between Y' (or Y'') and Z' (or Z'') as shown by various authors **[4-7]**. Furthermore, twinning occurs more easily for large magnitudes of the energy difference. For example, in case of the X-cut, the maximum of |F| is obtained with :

$\psi = 120°$ (or -60°), for which :

$$\Delta G = -2.8 \, . \, P^2 \text{ (in } N/m^2)$$

2.1 DESCRIPTION OF THE METHOD

The basic assumption of our treatment consists of stating that the stress-strain relationship is locally linear by pieces, the bounds of which are fixed by the threshold conditions on the energy difference of the twin materials. Since the global material behavior of the structure results from summing the behaviors of its elements, it is also linear by pieces. So a linear FEM solver can be useful for a stepped analysis performed through successive runs in which the stress-strain switches from one behavior to the twin one, for all elements newly obeying the threshold condition.

Our calculations have been performed by Finite Elements Method with SYSTUS® program on parallelepiped samples for which the largest parallel faces are perpendicular to the crystallographic X-axis. The pressure is applied on the upper face along Z'-axis which is rotated by ψ degrees towards X-axis. Conforming to the previous observations, ψ will take mainly the successive values : -30, -45 and -60°.

The dimensions of the samples introduced in the calculation are :

3, 9 and 8 mm along X-, Y'- and Z'-axes.

The sample has been meshed to obtain cubic meshes with 1 mm along each side. The pressure is applied on the upper face by the mean of skin meshes, usually used in massive structures.

The boundary conditions are introduced in the lower surface as imposed displacements. So, the displacement along Z'-axis of all nodes belonging to this plane is fixed to zero. Furthermore, to block any rotation, it is necessary to impose other displacements along X- and Y'-axes : we have chosen to impose a displacement equal to zero along Y'-axis for two opposite nodes in the same surface and to fixe the node A which is close to the middle of the plane (see figure 1).

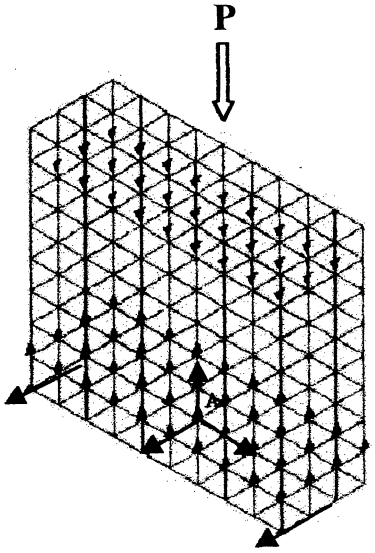


Fig. 1 : Boundary conditions and loading pressure on X-cut sample

Finally, the stiffness are introduced in two tables corresponding to the two domains (I for initial domain and II for the alternate one).

At each step and for each mesh, the energy difference between the twin states is calculated and compared to a chosen threshold which is fixed at :

-450,000 N/m^2 (and called E_T)

according to previous experiments **[8]**. While the energy difference for all meshes is higher than the threshold value, we increase the applied pressure. When a first mesh exhibits a lower value, we change the material properties of its because we consider that it is twinned and we performed another calculation without increasing the applied pressure : we have reached the coercive stress, called S_c.

In fact, as we will explained below, the propagation depends strongly on the form of the sample and of course on the direction of the applied pressure.

2.2 PERFECT CASE

Here, the sample is a perfect parallelepiped. The pressure is homogeneously applied on the upper face and E_T is introduced as threshold value.

In the calculation, we increase the pressure until one mesh exhibits lower value than E_T. We note that the quartz sample twinned completely (i.e. the energy value of each mesh is lower than E_T) as soon as the coercive stress is reached.

The next table (Table 1) gathers values calculated for different ψ-orientations of the X-cut samples. We note that the coercive stress depends strongly on the orientation of the sample.

Table 1 : Coercive stress for different orientations

ψ (in °)	-30	-45	-60
Coercive stress (S_c) (in MPa)	**680**	**448**	**393**

2.3 REAL CASE AND PROPAGATION OF THE TWINNING

Manufactured quartz samples present generally defects of parallelism due to the different machining operations as cutting, grinding and finally polishing. Furthermore, the roughness of the surface in which the pressure is applied being not perfect, the contact between the press and the sample can create stress concentrations. So, it appears as more efficient to introduce, in addition of the sticking condition, a slight defect in the geometry of the sample. We have chosen to induce a 0.1 mm defect in the parallelism between the 2 faces perpendicular to the Z'-axis (see figure 1).

We note that the first mesh exhibiting a free energy smaller than E_T is located in the lower face, close to the node A. The coercive stress calculated for each ψ-orientation are presented in the Table 2, including two new orientations : -15 and -75°.

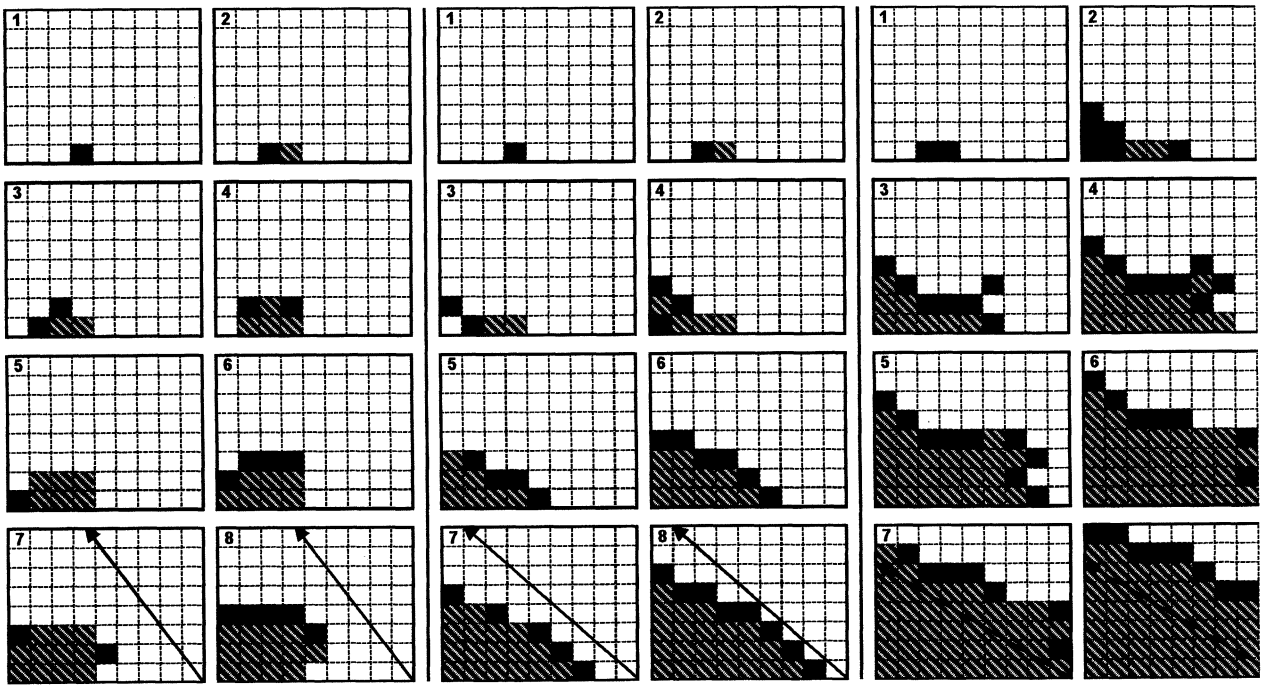
These values are lower than the first one, calculated for the perfect case. So, we deduce that the geometrical defect favors the process of twinning.

Table 2 : Coercive stress for different orientations in a more realistic case

ψ (in °)	-15	-30	-45	-60	-75
S_c in (MPa)	**880**	**575**	**411**	**368**	**381**
Deviation with perfect case		18 %	9 %	7 %	

Unlike the previous case, twinning occurs in just one or two meshes when the pressure reaches the coercive stress. It is so possible here to follow step by step the propagation of the twinning by changing elastic coefficients of the elements newly obeying the threshold condition. The figure 2 presents stepped analysis up to the 8^{th} run for three different orientations of the applied pressure (corresponding to three different values of ψ).

To visualize the propagation, we have just plotted the meshes located in the front X-face for which the normal is coming out. Dashed lines separate the meshes, the newly twinned meshes colored in black and changed in grey in the following steps.



$\Psi = -30°$ $\Psi = -45°$ $\Psi = -60°$

Fig. 2 : Twinning propagation for different orientations of the loading pressure

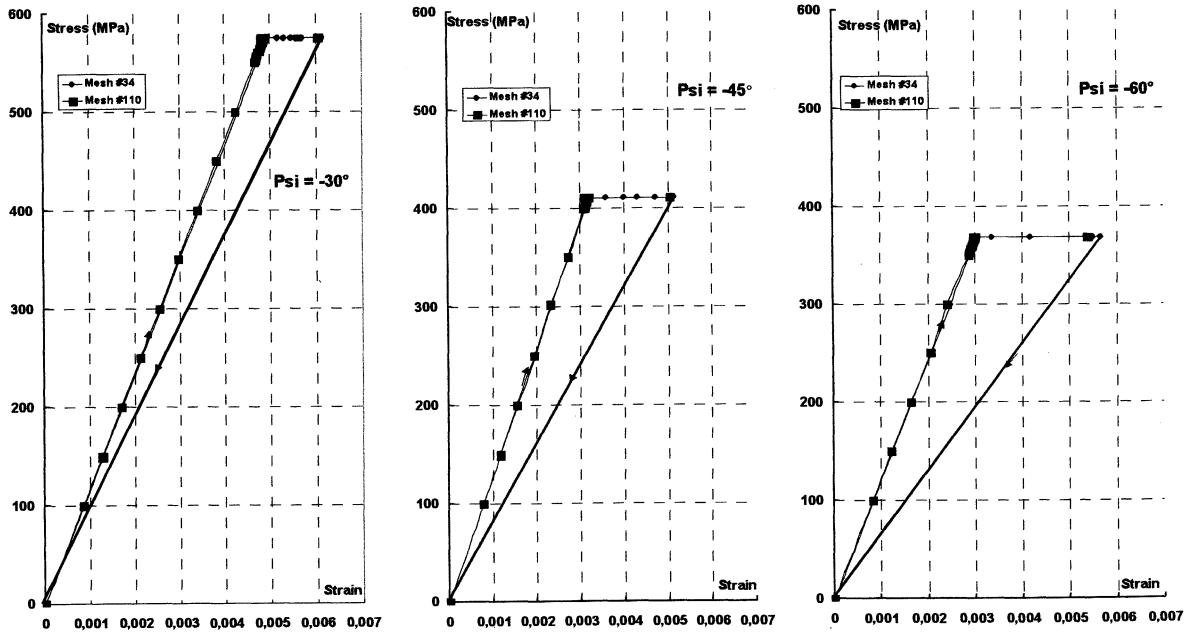


Fig. 3 : Stress-Strain curves for different orientations of the loading pressure

We observe that the propagation of the twin wall is clearly along Y-axis, each mesh newly twinned during one particular step being aligned in the Z-direction of the crystal. Furthermore, at each step, the number of twinned meshes increases when the coercive stress decreases. In other words, it seems easier to completely twin a X-cut quartz sample when the pressure is applied in a direction close to – 60° and so, for a given pressure higher than S_c, twinning occurs more rapidly.

The previous figure (Fig. 3) represents the stress-strain curves for three different orientations and, in each case, for two differently located meshes. For each curve, the pressure is increased up to the coercive stress and decreased when the sample is completely twinned. The first chosen mesh (number #34) is the first for which the energy value reaches the threshold one. As it is explained above, it is close to the point A belonging to the basic surface. We suppose that this is due to the boundary conditions which introduce some very small but existing stresses. The second one (number #110) is in the middle of the sample.

The following table (Table 3) summarizes our main observations concerning particularly the strain variations during twinning propagation (and so at a given pressure inducing coercive stress). We have to note first that the strains measured in each mesh of one given sample are slightly different due to the boundary conditions imposed in the point A, belonging to the mesh #34. The 3 last lines present the strain values calculated during three different runs and for the same pressure corresponding to the coercive stress (recalled in the second line). During the first runs, the strain calculated in the mesh #110 decreases slowly to increase after ferrobielastic transformation. The last line presents the difference between the two extreme states and indicates that this difference increases when the coercive stress decreases.

Table 3 : Strain and coercive stress for different orientations and for 2 meshes

Ψ (in °)	-30		-45		-60	
S_c (in MPa)	575		411		368	
Mesh Nb	# 34	# 110	# 34	# 110	# 34	# 110
Strain ($*10^{-3}$) P= 200 MPa	1.71	1.69	1.56	1.55	1.64	1.63
Strain ($*10^{-3}$) for the 1st run	**4.93**	**4.87**	**3.22**	**3.19**	**3.03**	**3.01**
Strain ($*10^{-3}$) for the 5th run	5.56	4.85	4.72	3.20	5.66	2.98
Strain ($*10^{-3}$) last run	**6.14**	**6.04**	**5.13**	**5.05**	**5.46**	**5.39**
Total strain difference	**1.21**	**1.17**	**1.91**	**1.86**	**2.41**	**2.38**

3. CONCLUSION

These first calculations on the model of the uniaxial stress explain clearly how the twinning can propagate in a very ordinary sample as X-cut parallelepipeds. We show that twin walls are parallel to Z-axis and propagate along Y-axis up to the surface of the sample. It means that twinning can propagate up to the edge of the sample without increasing pressure. We assume that the twinning is initiated by geometrical shape or surface defect or by an internal crystal defect, in fact in a locus where it can exist a concentration of stresses. But, if this model explains how twin walls propagate when the loading pressure is applied along a perpendicular to the crystallographic X-axis, we suppose that a more complicated pattern will be observed for other doubly rotated cut samples.

We have observed too that the crystal orientation plays an important role on the stress level associated to ferrobielastic twinning. Nevertheless, we think that our Finite Elements Method can be used to estimate the stress level for ferrobielastic twinning in more complicated samples with various orientations.

REFERENCES

1- **K. AIZU** : "Possible species of ferromagnetic, ferroelectric and ferroelastic crystals", Phys. Rev. B 2 (1970), pp. 754-772.

2 - **K. AIZU** : "Second order ferroic state shifts", J. Phys. Soc. Japan, 34 (1) (1973), pp. 121-128.

3 - **R.E. NEWNHAM, L.E. CROSS** : "Symmetry of secondary ferroics II", Mat. Res. Bull 9 (1974), pp. 1021-1032.

4 - **T.L. ANDERSON et al.** :"Coercive stress for ferrobielastic twinning in quartz", 30th Ann. Symp. Freq. Cont. (1976), pp. 171-177.

5 - **J.W. LAUGHNER** : "Induced Dauphiné twinning in quartz", Ph. D. Thesis, Penn. State Univ., 1982.

6 - **P.L. GUZZO et al.** : "Influence of the sample geometry on twinning generation in quartz : experimental measurements and finite elements analysis", 8th Europ. Freq. Time Forum (1994), pp. 802-810.

7 - **P.L. GUZZO, J.J. BOY** : "Influence of crystallographic orientation on ferrobielastic twinning in quartz", 9th Europ. Freq. Time Forum (1995), pp. 438-442.

8 - **J.J. BOY, P.L. GUZZO** : "Quartz crystal twinning under mechanical stress : Experimental measurements", 50th Ann. Symp. Freq. Cont. (1996), pp. 155-160.

LANGASITE FOR SURFACE ACOUSTIC WAVE SENSORS AT HIGH TEMPERATURES

R. FACHBERGER [1+3], M. HONAL [1], E. RIHA [1], E. BORN [2], P. PONGRATZ [3]

[1] Siemens AG, Corporate Technology, Germany
[2] Technical University of Munich, München, Germany
[3] Technical University of Vienna, Wien, Austria

Langasite ($La_3Ga_5SiO_{14}$) already has successfully been used as a piezoelecric substrate for surface acoustic wave (SAW) devices operating at high temperatures [1].
Thermal analyses (DSC) on commercially available Langasite (LGS) 3" wafers prove its stability up to temperatures of more than 1000 °C as no phase transition has been observed up to the melting point. Chemical and physical investigations of the stoichiometry (RFA, WDX) and optical properties (UV-VIS spectroscopy) allow the characterisation of different samples.
The crystal quality of different LGS wafers has been examined by X-ray topography (XRT), high resolution X-ray diffraction (HR-XRD) and chemical etching combined with optical microscopy. Most of the substrates show only few defects, however, transverse growth banding (TGB) perpendicular to the pulling axis (c-axis) has been observed in some crystals [2].
High temperature measurements on SAW-delay lines and ID-tags have been carried out up to 950 °C. Scanning electron analyses of the annealed devices indicate that the Ti/Pt interdigital transducers and the SMD-packages are the limiting factors for high temperature applications. The LGS substrate itself withstands these temperatures without a measurable change of the crystal surface and volume.

[1] J. Hornsteiner et al, *Surface acoustic wave sensors for high temperature applications*, Proc. IEEE International Frequency Control Symposium, pp. 615-620, 1998

[2] A. Zarka et al, *New results on high perfection langasite crystals: Studies of crystalline defects and modes shapes*, Proc. IEEE International Frequency Control Symposium, pp. 629-637, 1995

Corresponding author:
Dipl. Ing. René Fachberger
Siemens AG, ZT MS 1
81739 Munich
Germany
Tel.: 0049 89 636 53008
Fax: 0049 89 636 45740
E-mail: rene.fachberger@mchp.siemens.de

CHARACTERIZATION OF DEFECT AND STRAIN CONFIGURATIONS IN LANGANITE AND LANGATATE SINGLE CRYSTALS USING SYNCHROTRON WHITE BEAM X-RAY TOPOGRAPHY, AND ASSESSMENT OF THEIR INFLUENCE ON RESONATOR PERFORMANCE

M. DUDLEY,[1] B. RAGHOTHAMACHAR,[1] H. CHEN,[1] A. KHAN,[2] S. TIDROW, AND C. FAZI[2]
[1]Department of Materials Science and Engineering, State University of New York at Stony Brook, Stony Brook, NY 11794-2275, USA
[2]U.S. Army Research Laboratory, 2800 Powder Mill Road, Adelphi, MD 20783-1145, USA

Synchrotron white beam x-ray topography (SWBXT) has been used to characterize defect and strain configurations in langanite and langatate single crystals that are grown using the Czochralski technique. Both wafers sliced from as-grown crystals and resonator structures fabricated from as-grown crystals have been examined. Defects observed include growth striations, precipitates, and dislocations. The striation contrast is indicative of a slight variation in lattice parameter during growth and is possibly associated with the onset of constitutional supercooling in an off-composition melt. The periodicity of the variation is likely to be related to the crystal rotation rate. Striations observed on topographs recorded from crystal sections cut parallel to the growth axis reveal a cross-section view of the melt-solid interface shape. Typically, this is not observed to be flat but rather concave or convex. On topographs recorded from sections cut perpendicular to the growth axis, the growth striation structures appear as concentric rings and their spacing reveals the periodicity of the striations, and their symmetry reveals the symmetry of the thermal field inside the growth chamber. Clear images of precipitates are also routinely observed, and they provide further evidence of inhomogeneous constituent element or impurity segregation. An investigation of the chemical nature of the inhomogeneities is being done with a combination of techniques, including scanning electron microscopy and infrared spectroscopy. Once identified, strategies for the elimination of the inhomogeneities will be developed. Observations of defect configurations in crystal resonator structures enable the influence of the various defects present on resonator Q factors to be directly assessed. Preliminary results of this assessment will be presented. This work forms part of a larger project directed towards the development of these crystals for various resonator applications.

Introduction

It is well known that langasite, or LGS ($La_3Ga_5SiO_{14}$), and its isomorphs, langanite or LGN ($La_3Ga_{5.5}Nb_{0.5}O_{14}$ and langatate or LGT ($La_3Ga_{5.5}Ta_{0.5}O_{14}$) are of interest for application as bulk wave resonators for precision oscillators, with all three materials exhibiting high piezoelectric coupling, low acoustic loss (high Q), and temperature compensation. It is also generally accepted that crystal defects can influence mode shapes in such single crystals resonators. In langasite, the influence of crystal defects such as dislocations and growth bands (or striations) on resonator mode shapes has been elegantly demonstrated by the work of Capelle et al.[1] In this paper, we provide a preliminary survey of the quality of langanite and langatate crystals carried out using synchrotron white beam x-ray topography (SWBXT). Results obtained from the langasite crystals studied here were very similar to those observed by Capelle et al.[1] and, as such, are not presented here. Results obtained from the langanite and langatate crystals are discussed in light of the published work on langasite crystals.

Results and Discussion

Langanite

Although topographs were recorded from wafers cut from as-grown crystals and resonator structures fabricated on such crystals, the features observed were essentially the same (apart from the strain contrast visible along the edges of the electrode) so that only the former are presented here. Figure 1 shows a SWBXT image recorded in transmission geometry from an (001) wafer cut from a [001] growth-axis langanite crystal boule. Note the concentric growth bands or striations. Precipitates or inclusions can also be observed, as well as numerous surface scratches. The concentric striations indicate that the growth interface is either concave or convex, rather than flat. Chai et al.[2] attributed the origins of growth striations in langasite crystals to the onset of constitutional supercooling in an off-composition melt. It is possible that the growth striations observed in the langanite crystals have similar origins.

Figure 2 shows a pair of stereo SWBXT images, recorded in transmission geometry from a wafer cut parallel to the growth axis of a similar langanite boule. In this example, the curved nature of the growth interface is directly revealed (although whether it is concave or convex could not be discerned since the direction of positive growth was not known). Contrast from the striations is observed to reverse with the sign of the active reciprocal lattice vector, as shown in the enlargement in figure 3. This likely indicates that the striation contrast arises from dynamic diffraction effects close to the crystal surface, as previously observed in the striations in InP crystals.[3] However, in the InP crystals, the striations were clearly identified as being due to inhomogeneous incorporation of dopant during the liquid encapsulated Czochralski growth process. The exact origin of the striations observed here is under investigation. Figure 4(a) is an enlargement from figure 2 that shows the striations in more detail. The figure also reveals small precipitates or inclusions that are distributed throughout the crystal. Figure 4(b) is a similar enlargement that reveals oscillatory contrast along the growth interface that appears to be associated with the onset of interface instability.

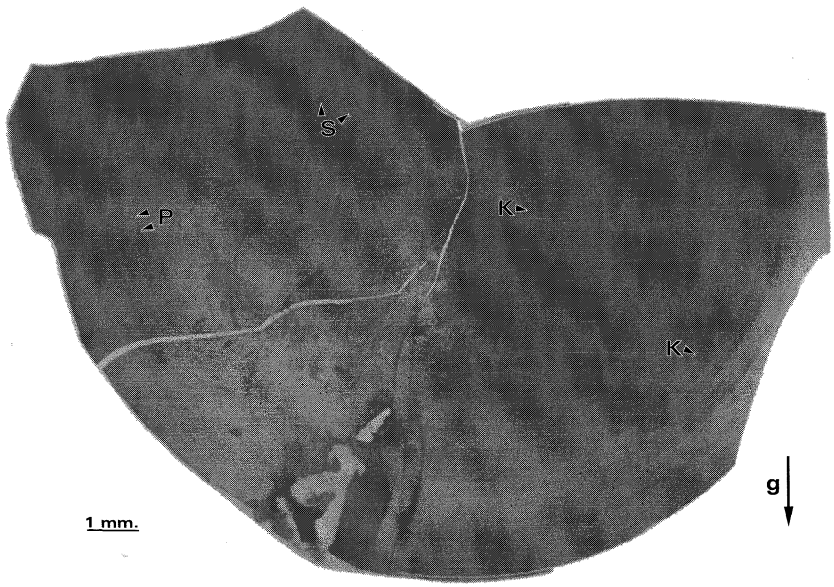


Figure 1. SWBXT image recorded in transmission geometry from a langanite crystal wafer cut perpendicular to the [001] growth axis of the LGN boule.

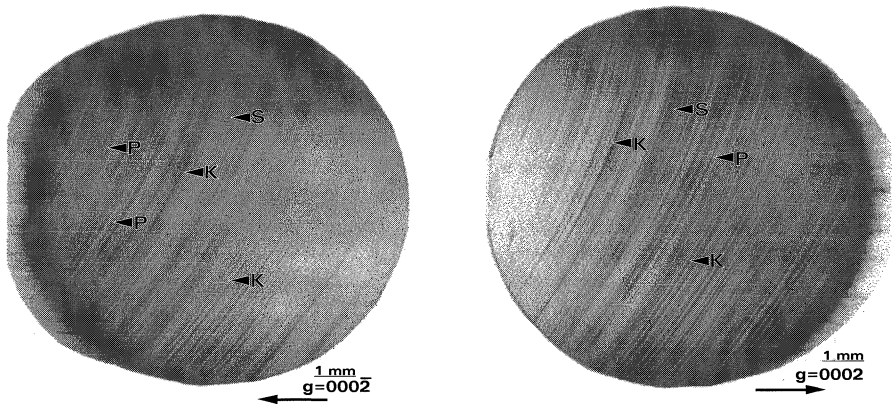


Figure 2. Pair of stereo SWBXT images recorded from wafer cut parallel to the [001] growth axis of a similar LGN boule.

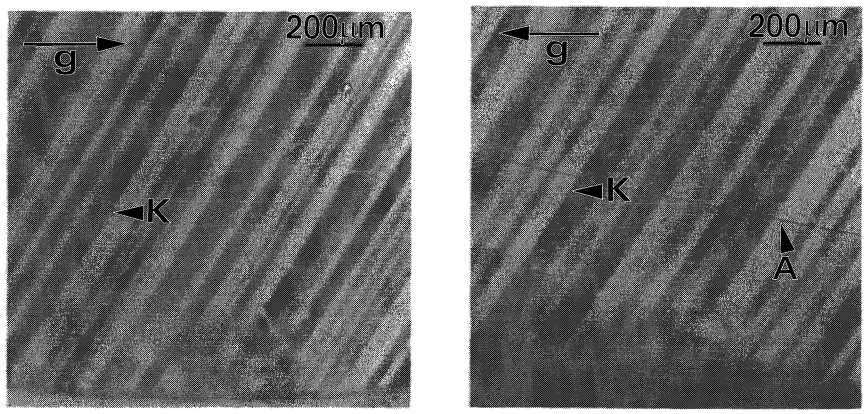


Figure 3. Enlargements from figure 2 confirming the reversal of striation contrast with reversal of sign of reflection vector.

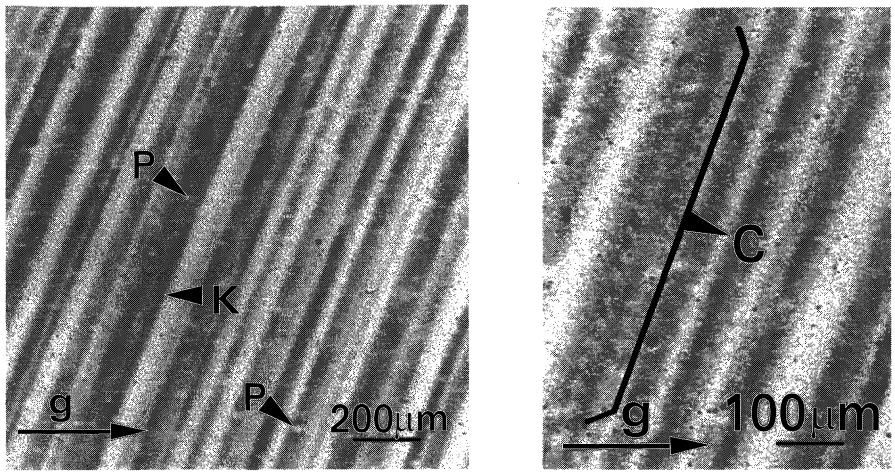


Figure 4. Enlargements taken from different areas of figure 2 showing (a) the presence of a distribution of precipitates or inclusions, P, and (b) oscillatory contrast, C, along the striations, possibly indicating the onset of interface instability.

Langatate

Figure 5 is a reflection topograph recorded from a langatate crystal cut parallel to the growth axis. Deep scratches, S, can be observed, as confirmed by optical microscopy. Also visible are bands of dislocations, D, possibly associated with plastic deformation of the crystal during postgrowth cooling. These are more clearly observed in the enlargement taken from the central region of figure 5, which is presented in figure 6, along with the optical image of the same region.

Figure 7 shows a SWBXT image recorded in transmission geometry from a langatate wafer cut parallel to the growth axis. Note the weak growth striations, F, visible along with precipitates, P, and scratches, S.

As in the case of langanite, the exact nature of the growth striations is under investigation.

Conclusions

Defect configurations somewhat reminiscent of those previously reported for langasite[1] (confirmed in this group) have been observed in both langanite and langatate single crystals. Growth striations, precipitates or inclusions, and dislocations are observed in both systems. Work is underway to determine the influence of these various defect configurations on resonator Q-factors. In the future, we expect that stroboscopic observation of the influence of the defect configurations on mode shape can be carried out as pioneered by Zarka et al (for example, see reference 4).

Acknowledgments

Support is acknowledged from the U.S. Army Research Office under contract number DAAG55-98-1-0392 (contract monitor Dr. John Prater). Topography was carried out at the National Synchrotron Light Source facility, at Brookhaven National Laboratory (BNL), which is supported by the U.S. Department of Energy, under contract number DE-AC02-98CH10886.

References

[1] B. Capelle, A. Zarka, J. Detaint, J. Schwartzel, A. Ibanez, E. Philippot and J. P. Denis, "Study of Gallium Phosphate and Langasite Crystals and Resonators by X-ray Topography," *Proceedings of 1994 IEEE International Frequency Control Symposium* (1994), pp. 48-57.

[2] B. Chai, J. L. Lefaucheur, Y. Y. Ji, and H. Qui, "Growth and Evaluation of 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 1998 IEEE International Frequency Control Symposium* (1998), pp. 748-760.

[3] H. Chung, Ph.D. thesis, SUNY, Stony Brook, NY (1997).

[4] A. Zarka, B. Capelle, J. Detaint, J. Schwartzel, "Stroboscopic Topography of Quartz Resonators," *Proceedings of 42nd IEEE International Frequency Control Symposium* (1988), pp. 85-92.

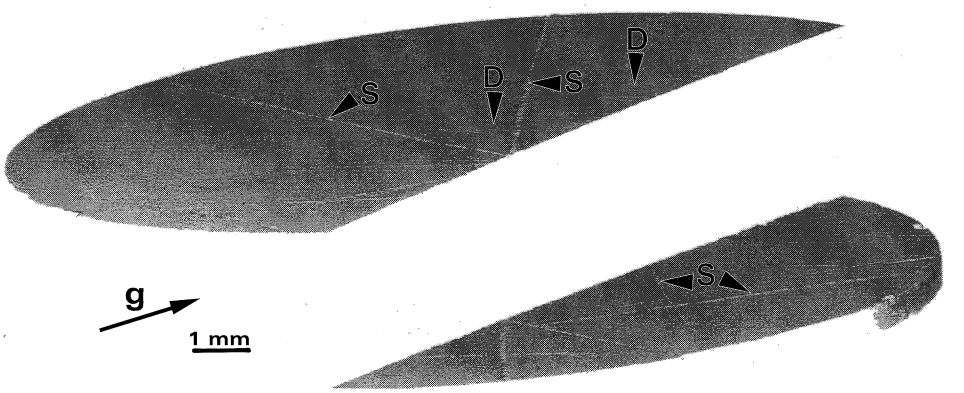


Figure 5. Reflection topograph recorded from a langatate single crystal wafer cut parallel to the boule growth axis.

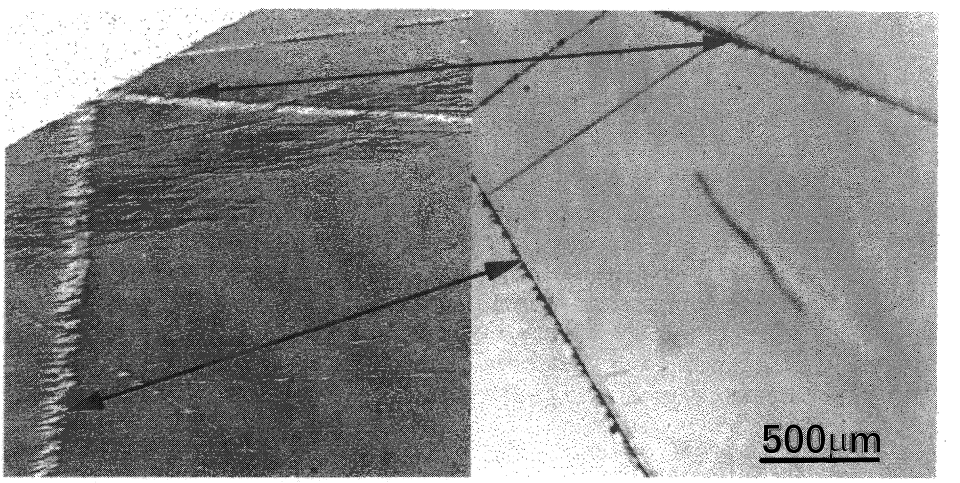


Figure 6. Enlargement taken from figure 7 along with a reflection optical micrograph of the same region showing the deep scratches.

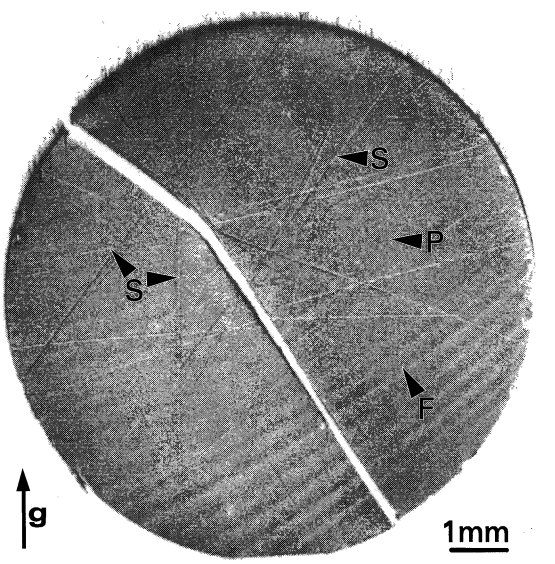


Figure 7. Transmission SWBXT image recorded from a langatate crystal cut parallel to the growth axis.

SELF-CONSISTENT THEORY OF THE CPT MASER

*Aldo Godone, *Filippo Levi, ^S. Micalizio and +J. Vanier
*Istituto Elettrotecnico Nazionale, Torino, Italy
*Politecnico di Torino, Torino, Italy
+Université de Montréal, Montréal, Canada

1 ABSTRACT

The non-linear differential equations which fully describe the behavior of the Coherent Population Trapping (CPT) maser in transient and in continuos operation are solved and analyzed in a pure three level system. The reaction of the field generated by the atomic ensemble through the microwave cavity on the atoms themselves is fully taken into account. The coupling of the atomic ensemble with the microwave introduces an unbalance in the populations of the ground state hyperfine levels which reduces the intrinsic symmetry of a Λ scheme. The analogous equations for the Optically Pumped (OP) maser are also reported and compared with the previous ones, allowing a deeper understanding of the CPT maser operation.

2 INTRODUCTION

The emission of coherent microwave radiation from the ground state of alkali-atoms has been demonstrated in [1], when the two ground state hyperfine levels are coupled to a common excited state by means of two coherent radiations. This emission is due to the oscillating magnetization created in the atomic ensemble due to the strong coherence generated by the Λ scheme excitation, which couples the two ground-state levels. This coherence is also responsible of the dark state which can be observed in the fluorescence spectrum [2].

Beside the physical interest of the above phenomenon called Coherent Population Trapping (CPT), applications have been proposed in the field of atomic cooling, lasing without inversion, magnetometry and atomic frequency standards.

In particular the characteristics of the microwave emission have been examined both theoretically and experimentally in [3], in view of the possible realization of a new atomic frequency standard: the "CPT maser". The absence of a threshold for the microwave emission allowed to simplify the theoretical analysis in the limit of negligible reaction of the microwave field on the atomic ensemble (low-Q cavity or low number of atoms). From the analysis turned out the possible interpretation of the microwave emission as a stimulated coherent Raman radiation allowing to ascribe this device among the group of atomic oscillators such as the hydrogen and rubidium masers [4,5].

In this paper we will develop a fully consistent theory for this atomic oscillator, taking into account the coupling of the atomic ensemble to the microwave cavity. Finally we will deduce, in the frame of the same theoretical approach, the equations describing the intensity pumped maser which allow by comparison a deeper understanding of the CPT maser peculiarities.

3 THEORY

The energy level scheme considered in our analysis is shown in Fig.1; although the following discussion is referred to the D_1 optical transition of ^{87}Rb, it can be extended to the general case of Λ transitions in alkali-metal atoms, observed in a cell with buffer gas, or more in general in a monokinetic ensemble.

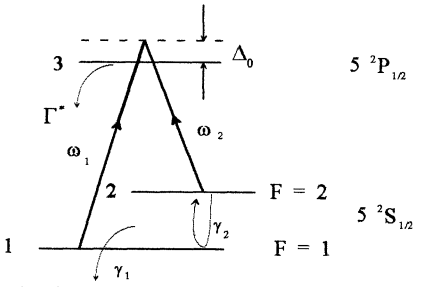


Fig.1. Three-level system considered in the analysis; ω_1 and ω_2 and are the laser angular frequencies and $\omega_{21}/2\pi$ is the hyperfine frequency (6.834 GHz) shifted by various static perturbations described in the text. Γ* is the decay rate from the excited state to the ground state levels, γ_1 and γ_2 are the ground state relaxation rates of the population difference and of the coherence and Δ_0 is the lasers detuning from the optical resonance.

The decay rates Γ^*, γ_1 and γ_2 take into account the effect of the collisions of the Rb atoms with the buffer gas which is used to reduce the linewidth broadening due to the transit time of the atoms across the laser beams (Dicke effect) [6]. The frequency shifts of the unperturbed hyperfine frequency and of the optical transitions due to the buffer gas are included in the static energy level separations as well as the Zeeman shift due to the quantization static magnetic field $\vec{B}_0$, which decouples also the Zeeman sub-levels in the ground state and allows the observation of isolated pairs of these levels. The basic physical arrangement considered in the analysis is shown in Fig.2.

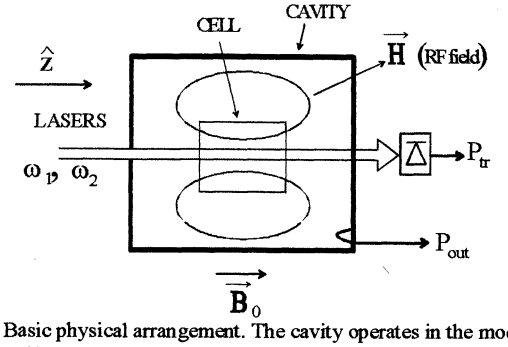


Fig.2. Basic physical arrangement. The cavity operates in the mode TE_{011} and has a loaded quality factor Q_L; Ptr and Pout are the optical transmitted power and the microwave output power.

The master equation describing the atomic ensemble is the generalized Liouville-Bloch equation [6,7]:

$$\frac{d\hat{\rho}(t)}{dt}+\hat{\Gamma}_{rel}(t)=\frac{1}{i\hbar}\left[\hat{H}(t),\hat{\rho}(t)\right]+\hat{\Gamma}_{exc}(t) \qquad (1)$$

where $\hat{\rho}(t)$ is the density operator, $\hat{H}(t)$ is the Hamiltonian operator, $\hat{\Gamma}_{rel}(t)$ is the relaxation operator and $\hat{\Gamma}_{exc}(t)$ is an operator describing all the excitation processes not induced by external fields (mainly spontaneous emission and collisions, in our case).
The total Hamiltonian operator, resulting by the sum of the unperturbed Hamiltonian and the time-dependent perturbation Hamiltonian, in the dipole approximation may be written as:

$$\hat{H}(t)=\begin{pmatrix} \varepsilon_1 & -\mu_{12}B(t) & -d_{13}E_1(t) \\ -\mu_{21}B(t) & \varepsilon_2 & -d_{23}E_2(t) \\ -d_{31}E_1(t) & -d_{32}E_2(t) & \varepsilon_3 \end{pmatrix} \qquad (2)$$

where ε_i (i=1,2,3) are the unperturbed energy eigenvalues, which include all the time-independent perturbations of the atomic levels; $d_{ij}=d_{ji}=\langle i|e\vec{r}\cdot\vec{e}_\lambda|j\rangle$ are the electric dipole matrix elements of the optical transitions and $\mu_{12}=\mu_{21}$ is the projection of the magnetic dipole moment between the ground state hyperfine levels on the microwave magnetic field direction.
The laser field $\vec{E}_1(t)$ and $\vec{E}_2(t)$ are assumed of the form:

$$\begin{cases}\vec{E}_1(t)=\vec{e}_\lambda E_1(t)=E_{01}\vec{e}_\lambda\cos(\omega_1 t+\vec{k}_1\cdot\vec{r}) \\ \vec{E}_2(t)=\vec{e}_\lambda E_2(t)=E_{02}\vec{e}_\lambda\cos(\omega_2 t+\vec{k}_2\cdot\vec{r})\end{cases} \qquad (3)$$

being $\vec{e}_\lambda$ the polarization vector and $\vec{k}_i$ (i=1,2) the wave-vectors ($\vec{k}_i=k_i\hat{z}$). The RF field $\vec{B}(t)$ is assumed of the form:

$$\vec{B}(t)=\hat{z}B_z\cos(\omega_{rf}t+\phi) \qquad (4)$$

where $\omega_{rf}=\omega_1-\omega_2$ and ϕ is a phase angle whose meaning will be cleared up later on.
The relaxation operator may be expressed as (Bloch-Wangness-Redfield theory [6]):

$$\hat{\Gamma}_{rel}(t)=\begin{pmatrix} \frac{1}{2}\gamma_1\rho_{11} & \gamma_2\rho_{12} & \frac{1}{2}\Gamma^*\rho_{13} \\ \gamma_2\rho_{21} & \frac{1}{2}\gamma_1\rho_{22} & \frac{1}{2}\Gamma^*\rho_{23} \\ \frac{1}{2}\Gamma^*\rho_{31} & \frac{1}{2}\Gamma^*\rho_{32} & \Gamma^*\rho_{33} \end{pmatrix}. \qquad (5)$$

The operator $\hat{\Gamma}_{exc}(t)$ has the form:

$$\hat{\Gamma}_{exc}(t)=\begin{pmatrix} \frac{1}{2}\Gamma^*\rho_{33}+\frac{1}{2}\gamma_1\rho_{22} & 0 & 0 \\ 0 & \frac{1}{2}\Gamma^*\rho_{33}+\frac{1}{2}\gamma_1\rho_{11} & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad (6)$$

For the off-diagonal terms of the density matrix (coherences) we assume solutions of the form:

$$\begin{cases}\rho_{12}=\delta_{12}e^{i(\omega_1-\omega_2)t} \\ \rho_{13}=\delta_{13}e^{i\omega_1 t} \\ \rho_{23}=\delta_{23}e^{i\omega_2 t}\end{cases} \qquad (7)$$

where the δ_{ij} are slow time-varying terms.
We make also the long-wavelength and the rotating-wave approximations and we assume that the Doppler terms are null or through the Dicke effect [8] or by using a cold atom ensemble; moreover the buffer gas is assumed to be nitrogen or a mixture containing nitrogen to quench the fluorescence light scattered by the atoms.
Inserting equations (2) to (7) in (1) we obtain the system of coupled differential equations which describes our atomic ensemble:

$$\begin{cases}\rho_{11}+\rho_{22}\approx 1 \\ \dot{\rho}_{33}+\Gamma^*\rho_{33}=-\omega_{R1}\,\mathrm{Im}\,\delta_{13}-\omega_{R2}\,\mathrm{Im}\,\delta_{23} \\ \dot{\Delta}+\gamma_1\Delta=-2b\,\mathrm{Im}(e^{-i\phi}\delta_{12})+\omega_{R2}\,\mathrm{Im}\,\delta_{23}-\omega_{R1}\,\mathrm{Im}\,\delta_{13} \\ \dot{\delta}_{12}+(\gamma_2+i\Omega_\mu)\delta_{12}=i\frac{b}{2}e^{i\phi}\Delta+i\frac{\omega_{R1}}{2}\delta_{32}-i\frac{\omega_{R2}}{2}\delta_{13} \\ \dot{\delta}_{13}+\left(\frac{1}{2}\Gamma^*+i\Delta_0\right)\delta_{13}=-i\frac{\omega_{R1}}{4}(1-\Delta)-i\frac{\omega_{R2}}{2}\delta_{12} \\ \dot{\delta}_{23}+\left(\frac{1}{2}\Gamma^*+i\Delta_0\right)\delta_{23}=-i\frac{\omega_{R2}}{4}(1+\Delta)-i\frac{\omega_{R1}}{2}\delta_{12}^*\end{cases} \qquad (8)$$

In the system (8) ω_{R1}, ω_{R2} and b are the Rabi frequencies defined as:

$$\omega_{R1}=\frac{1}{\hbar}E_{01}\,d_{13},\quad \omega_{R2}=\frac{1}{\hbar}E_{02}\,d_{23},\quad b=\frac{1}{\hbar}B_z\mu_{12}. \qquad (9)$$

The microwave detuning and the population inversion between the ground state levels are defined by $\Omega_\mu=(\omega_1-\omega_2)-\omega_{21}$ and $\Delta=\rho_{22}-\rho_{11}$ respectively.
Equations (8) have been obtained introducing the following approximations widely satisfied in the usual experimental conditions:

$$\Omega_\mu\ll\Gamma^*,\quad b\ll\Gamma^*,\quad \omega_{R1,2}\ll\Gamma^*. \qquad (10)$$

The third approximation of (10) allows to cancel ρ_{33} in the optical coherences equations and uncouples the first two equations of (8) from the other equations.
The system (8) has been analyzed analytically in [3] in the case of $b=0$, leading to the explanation of the dark-line observed in the fluorescence spectrum or to the coherent microwave emission at $\omega_1-\omega_2$ observed when the atoms are coupled to a low-feedback cavity.
The microwave emission is due to the macroscopic magnetization $\vec{M}$ oscillating at $\omega_1-\omega_2$, created by the coherence ρ_{12}:

$$\vec{M}(\vec{r},t)=Tr(\hat{M}\,\hat{\rho}) \tag{11}$$

This magnetization is the "source term" of the classical standing field $\vec{H}(\vec{r},t)$ sustained by the cavity; the complex amplitude $\vec{H}(\vec{r})$ of the field is [3]:

$$\vec{H}(\vec{r})=\frac{-iQ_L}{1+i2Q_L(\Delta\omega_c/\omega_{21})}\vec{H}_c(\vec{r})\int_{V_c}\vec{H}_c(\vec{r})\cdot\vec{M}(\vec{r})dv \tag{12}$$

where $\vec{H}_c(\vec{r})$ is the orthonormal cavity mode, V_c the cavity volume and $\Delta\omega_c=\omega_c-(\omega_2-\omega_1)\cong\omega_c-\omega_{21}$ the cavity detuning.
Being the power dissipated in a cavity near resonance ($\Delta\omega_c \ll \omega_{21}/2Q_L$) given by

$$P_{diss}=\frac{\omega_{21}\mu_0}{2Q_L}\int_{V_c}\left|\vec{H}(\vec{r})\right|^2 dv \tag{13}$$

where μ_0 is the vacuum permeability, from (11), (12) and (13), following the computations reported in [3], we have:

$$P_{diss}=\frac{\frac{1}{2}\hbar\omega_{21}kN|2\delta_{12}|^2}{1+4Q_L^2(\Delta\omega_c/\omega_{21})^2} \tag{14}$$

where k is the number of microwave photons emitted by an atom in one second:

$$k=\frac{\mu_0\mu_{21}^2\eta' Q_L N}{\hbar V_a}. \tag{15}$$

In (15) V_a is the volume containing the interacting atoms and η' is the filling factor defined as:

$$\eta'=\frac{V_a\langle H_{cz}(r)\rangle_a^2}{V_c\langle H_c^2(r)\rangle_c} \tag{16}$$

Here $\langle\ \rangle_a$ and $\langle\ \rangle_c$ mean an average over V_a and V_c of the mode eigenvectors [6]. From equations (13) and (14) we obtain also the fundamental relation between the Rabi frequency b of the microwave generated by the atomic ensemble and the coherence in the ground state (always in the limit of small cavity detuning):

$$b=-2k|\delta_{12}| \tag{17}$$

This last equation, introduced in (8), provides a self-consistent set of equations for the description of the atoms inside the microwave cavity; as far as the phase ϕ of the generated field is concerned, we have from equation (12), in the limit of small cavity detuning:

$$\phi=+\frac{\pi}{2}+\tan^{-1}2Q_L\frac{\Delta\omega_c}{\omega_{21}}+\tan^{-1}\frac{\mathrm{Im}\,\delta_{12}}{\mathrm{Re}\,\delta_{12}}. \tag{18}$$

The last term of (18) is the phase of the complex magnetization $\vec{M}(\vec{r})$, which is proportional to $-\delta_{12}$.

3.1 CPT MASER

We consider at the beginning the case of equal intensity and exactly tuned lasers, that is $\omega_{R1}=\omega_{R2}=\omega_R$ (if $d_{13}=d_{23}$ as far the D_1 line), Δ_0 and $\Omega_\mu=0$; we assume also $\Delta\omega_c=0$ (cavity tuned to the hyperfine frequency). In this case we have from (8), (17) and (18) the following set of coupled differential equations:

$$\begin{cases}\dot{\rho}_{33}+\Gamma^*\rho_{33}=-\omega_R\,\mathrm{Im}\,\Delta^+\\ \dot{\Delta}+\gamma_1\Delta=-4k|\delta_{12}|^2+\omega_R\,\mathrm{Im}\,\Delta^-\\ \dot{\delta}_{12}+\gamma_2\delta_{12}=k\,\Delta\delta_{12}+i\frac{1}{2}\omega_R\,\mathrm{Re}\,\Delta^-+\frac{1}{2}\omega_R\,\mathrm{Im}\,\Delta^+\\ \dot{\Delta}^+-\frac{1}{2}\Gamma^*\Delta^+=-i\frac{1}{2}\omega_R-i\omega_R\,\delta_{12}^r\\ \dot{\Delta}^--\frac{1}{2}\Gamma^*\Delta^-=-i\frac{1}{2}\omega_R\,\Delta-\omega_R\,\delta_{12}^i\end{cases} \tag{19}$$

where $\Delta^+=\delta_{23}+\delta_{13}$, $\Delta^-=\delta_{23}-\delta_{13}$ and $\delta_{12}=\delta_{12}^r+i\delta_{12}^i$.
The optical coherences respond at a rate of the order Γ^* and the hyperfine coherence at a rate of the order of γ_2; in the practical cases $\Gamma^*\approx 10^8\div10^9\ s^{-1}$ and $\gamma_2=10\div1000\ s^{-1}$ so that Δ^+ and Δ^- reach their equilibrium values much faster than δ_{12}. It is than possible to solve (19) taking into account that the optical coherences build up rapidly while during that time the RF coherence stays at a negligible value and after this build-up the optical coherences follow adiabatically the RF coherence. The stationary expressions ($\partial/\partial t\to 0$) of Δ^+ and Δ^-

$$\begin{cases}\Delta^+=-i\frac{\omega_R}{\Gamma^*}(1+2\delta_{12}^r)\\ \Delta^-=-i\frac{\omega_R}{\Gamma^*}(\Delta-2i\delta_{12}^i)\end{cases} \tag{20}$$

may then be substituted in the first two equations of (19). It turns out that $\delta_{12}^i=0$ and that the system (19) reduces to:

$$\begin{cases}\dot{\Delta}+(\gamma_2+2\Gamma_p)\Delta=-4k(\delta_{12}^r)^2\\ \dot{\delta}_{12}^r+(\gamma_2+2\Gamma_p)\delta_{12}^r=k\,\Delta\delta_{12}^r-\Gamma_p\end{cases} \tag{21}$$

where $\Gamma_p=\omega_R^2/2\Gamma^*$ plays the role of "transversal pumping rate".
Equations (21) constitute an autonomous system of coupled non-linear differential equations which fully describes the CPT maser behavior and allows to evaluate the power generated by the atoms through equation (14). The equilibrium solutions $[\Delta_e,(\delta_{12}^r)_e]$ are given by:

$$\begin{cases} \Delta_e = -\dfrac{4k}{\gamma_2 + 2\Gamma_p}(\delta_{12}^r)_e^2 \\ (\delta_{12}^r)_e^3 + \dfrac{(\gamma_2 + 2\Gamma_p)^2}{4k^2}(\delta_{12}^r)_e + \dfrac{\Gamma_p(\gamma_2 + 2\Gamma_p)}{4k^2} = 0 \end{cases} \quad (22)$$

The above system has always only one real solution with $\Delta_e \le 0$ and $(\delta_{12}^r)_e \le 0$ ($\delta_{12}^i \equiv 0$ if $\Omega_\mu = 0$), which represents a stable stationary point of operation of the atomic system coupled to the microwave cavity. The stability of this equilibrium point may be easily analyzed through the eigenvalues of the Jacobian associated to (21): it turns out to be a stable spiral when $k^2(\delta_{12}^r)_e^2 < 2(\gamma_2 + 2\Gamma_p)^2$ or a stable node in the opposite case.

The main physical properties of the stationary solution provided by (22) are:

i) no threshold exists for the coherent emission;

ii) $\rho_{11} > \rho_{22}$: the ground state populations are "anti-inverted" (more population in the lower state).

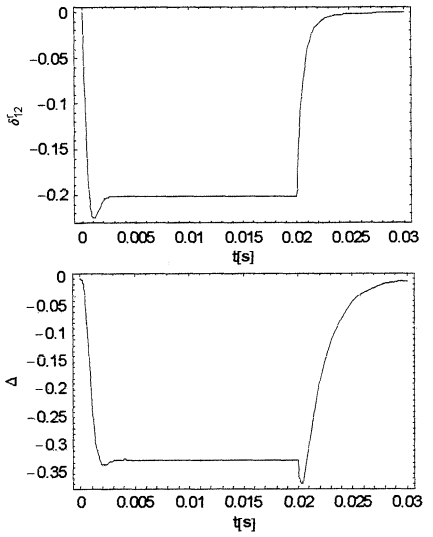


Fig.3. Time response of the ground-state coherence and population difference to a laser pulse with: γ_2 = 500; k = 3000; Γ_p= 500. Laser pulse width 20 ms.

The analytical solution of (21) is not known when $\Gamma_p \neq 0$ and the response of the system to a laser pulse (dynamical behavior) has to be analyzed by numerical integration, as shown in Fig.3 for a particular set of physical parameters.

The effect known as radiation damping is readily observed in the transient response reported in Fig. 3: oscillation in the switching on transient and reduced decay time in the switching off transient.

Before going further on in the analysis of the CPT maser it is useful and instructive to examine the coherent emission due to pure intensity pumping.

3.2 OP MASER

The OP maser may be studied in the frame of the three-level theory so far developed, setting in (8) $\omega_{R1} = \omega_R$ and $\omega_{R2} = 0$. Following the same procedure reported for the CPT case, we obtain:

$$\begin{cases} \dot{\Delta} + (\gamma_2 + \Gamma_p)\Delta = -4k(\delta_{12}^r)^2 + \Gamma_p \\ \dot{\delta}_{12}^r + (\gamma_2 + \Gamma_p)\delta_{12}^r = k\Delta\delta_{12}^r \end{cases} \quad (23)$$

In this case $\Gamma_p = \omega_R^2 / 2\Gamma^*$ plays the role of "longitudinal pumping rate" and represents the main difference with respect to the CPT maser. This difference anyway implies a fundamental change in the behavior of the atomic ensemble, as we are going to describe.

The above system admits one or two equilibrium points:

$$P_1 \begin{cases} \Delta_e = \Gamma_p / (\gamma_2 + \Gamma_p) \\ (\delta_{12}^r)_e = 0 \end{cases} \qquad P_2 \begin{cases} \Delta_e = (\gamma_2 + \Gamma_p)/k \\ (\delta_{12}^r)_e = \pm\dfrac{1}{2}\sqrt{\dfrac{\Gamma_p}{k} - \left(\dfrac{\gamma_2 + \Gamma_p}{k}\right)^2} \end{cases}$$

The point P_1 is always existing and represents an atomic system with inverted population in the ground state but with no emitted radiation; it is a stable point when $k < (\gamma_2 + \Gamma_p)^2 / \Gamma_p$ (otherwise is a saddle). The second point P_2 exists only when $k > 4\gamma_2$ and $k > (\gamma_2 + \Gamma_p)^2 / \Gamma_p$ and corresponds to an active maser emission; it is stable node when $k < 9(\gamma_2 + \Gamma_p)^2 / 8\Gamma_p$ and a stable spiral in the opposite case. The choice between the two solutions (± sign) is defined by the initial conditions.

From the physical point of view we have in this case that: a threshold exists for a stable coherent microwave emission ($k = 4\gamma_2$); $\rho_{22} > \rho_{11}$: the ground state populations are inverted; there is only a well defined pumping rate range where the maser emission can be self sustained.

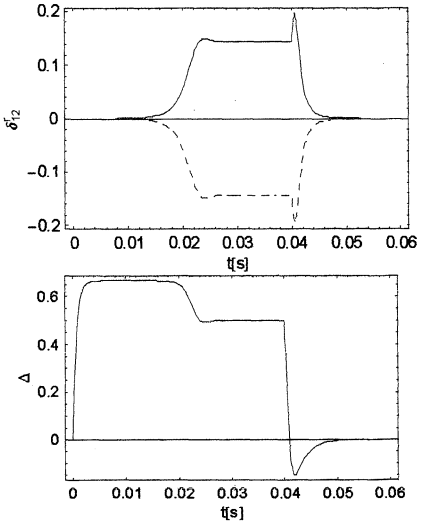


Fig.4. Dynamical behavior of the OP maser over threshold. k = 3000 s^{-1}, γ_2 = 500 s-1, Γ_p = 1000 s^{-1}. Laser pulse width 40 ms.

Even over threshold noise is required in order to move the system from P_1 to P_2 and to define the sign of the

solution; this feature is typical of active oscillators. A numerical solution of (23) showing the response to a laser pulse is reported in Fig.4. We will not go much further on in the analysis of the well-known intensity pumped maser, but we only remark that many results reported in the literature may be easily derived from the three-level theory here reported.

3.3 COMPARISON BETWEEN CPT AND OP MASERS

The fundamental difference between CPT and OP masers is even more evident in Fig.5 which shows the stable operating points for the two oscillators in the plane Δ, δ_{12}^r : in the OP maser no emission takes place as far as the population inversion Δ has reached a significant value, while in the CPT case coherent emission is present since $t = 0^+$ where $\Delta \cong 0$.

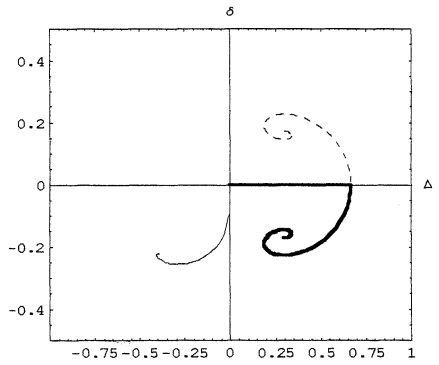


Fig.5. Parametric plot of the coherence versus population inversion for the CPT (—) maser and for OP (---) maser. In the second case a symmetric solution (—) is also possible (depending on the initial value of δ_{12}^r). k = 5000 s^{-1}, γ_2 = 500 s^{-1}, Γp = 1000 s^{-1}. The axes origin corresponds to t = 0.

Figure 6 a) and b) compares the emitted power, proportional to $k(2\delta_{12}^r)_e^2$, as a function of the pumping rate for the two cases considered in this section.
Table I summarizes the main differences between the CPT and the OP masers; even if both of them emits coherent radiation the CPT maser has mostly the features of a passive maser.

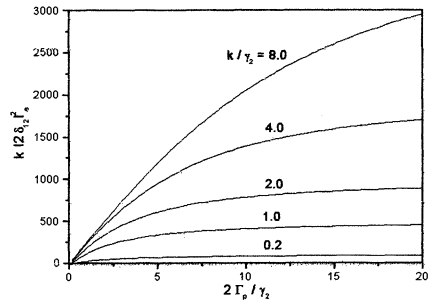


Fig. 6a). Emitted power of CPT maser versus light intensity

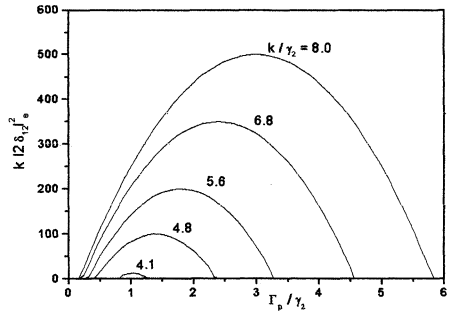


Fig. 6b). Emitted powerOP maser versus the pumping ratethe curves are parabolas of equation $P_a \propto k(2\delta_{12}^r)_e^2 = \Gamma_p - (\gamma_2 + \Gamma_p)^2/k$.

Maser	Δ	Threshold	ω_{out}	P_{out}
CPT	<0	no	$\omega_1 - \omega_2$	Saturation *vs.* pump
OP	>0	$k/\gamma_2 = 4$	ω_{21}	Threshold *vs.* pump

Table I. Main physical differences between the two types of maser considered in the text.

4 CONCLUSIONS

In this paper we have presented a new, self-consistent theory of the CPT maser, that take in account the reaction of the emitted microwave on the atomic system itself. The same theoretical approach, based on a three level system, is used also to analyze the OP maser, allowing a simple and powerful description of its behavior. A direct comparison of the two masers is possible in the same theoretical framework showing similarities and differences between the two phenomena.

5 REFERENCES

[1] A. Godone, F. Levi and J.Vanier, Phys. Rev A 59, R12 (1999)
[2] G. Alzetta, A. Gozzini, L. Moi and G. Orriols, Nuovo Cimento B36, 5 (1976)
[3] J.Vanier, A. Godone and F. Levi, Phys. Rev A 58, 2345 (1998)
[4] D. Kleppner, H. M. Goldenberg and N. F. Ramsey, Phys. Rev. 126, 603 (1962)
[5] P. Davidovits, Appl. Phys. Lett. 5, 15 (1964)
[6] J. Vanier and C. Audoin, The Quantum Physics of Atomic Frequency Standards, (Adam Highler Bristol, England,1989)
[7] M. Sargent, M. O. Scully and W.E. Lamb Jr, Laser Physics (Addison-Wesley, London, 1974)
[8] R. H. Dicke, Phys. Rev. 89, 472 (1953)

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

DETERMINATION OF RESIDUAL CAVITY PHASE SHIFT IN SMALL CESIUM CLOCKS

Luc CHASSAGNE , Frédéric HAMOUDA, Geneviève THEOBALD and Pierre CEREZ

Laboratoire de l'Horloge Atomique, Unité Propre de Recherche du CNRS associée à l'Université Paris-Sud, Bâtiment 221 – Université Paris-Sud, 91405 Orsay Cedex, France

ABSTRACT

It clearly appears at the present time that the frequency accuracy of a small optically pumped cesium beam resonator (OPCsBR) is limited by the end to end phase difference $\Delta\phi$ between the two oscillating fields in the Ramsey cavity arms. In a small and compact OPCsBR, $\Delta\phi$ is unknown because the beam reversal procedure is not implemented.

At the Laboratoire de l'Horloge Atomique, we designed and operated a small and symmetrical optically pumped cesium beam resonator named Cs5 which allows us to implement the beam reversal technique.

The description of Cs5 and its frequency performances are reported in this paper. The clock frequency stability is about $8.10^{-12}\ \tau^{-0.5}$ and less than 4.10^{-14} for $\tau > 1$ day. The residual phase difference of the Ramsey cavity has been measured either by the beam reversal method or by a direct method without beam reversal. The values of $\Delta\phi$ obtained by the two methods and their uncertainties are compared and a complete accuracy budget of the clock is performed.

INTRODUCTION

Small optically pumped cesium beam clocks have been extensively studied and evaluated in our laboratory [1-2]. It appears that they exhibit a very satisfactory level of performance for the stability. Concerning the accuracy, the limiting factor still remains the lack of knowledge of the residual cavity phase difference $\Delta\phi$ between the two arms of the Ramsey cavity. Presently, two methods are available to determine $\Delta\phi$: the well known beam reversal technique on one hand, and the Makdissi-de Clercq method on the other hand. In small commercial cesium beam tubes, the first method is not applied for evident compactness reasons. In order to validate the second method on small optically pumped cesium beam resonator, we have designed a small machine equiped with two ovens which enables the beam reversal. This small and symmetrical resonator called Cs5 will be detailed in a first part. Then its frequency performances will be given. We shall focus our attention on the measurement of $\Delta\phi$ by the two above mentioned methods. Finally an accuracy budget will be established and we shall compare the $\Delta\phi$ uncertainties delivered by the two measurement methods.

1. DESIGN AND FEATURES OF Cs5

Cs5 is a small and symmetrical optically pumped cesium beam resonator. Its design is derived from Cs4 developped and studied earlier in the laboratory. The main characteristics of Cs4 can be found in reference 2. The sketch of Cs5 is shown in Figure 1. Two ovens are implemented on each side of the structure. They are called East oven and West oven. They operate continuously at 100°C. Two graphite blocks can be moved in front of each oven in order to block out the cesium beams. The symmetrical structure of Cs5 requires a set of identical graphite diaphragms for the beam collimation.

The two light-cesium atom interaction zones are identical in the resonator in order to ensure either the pumping function or the optical detection function. Two photodetection systems are outside the resonator. One of them is devoted to the laser frequency stabilization. The other delivers the clock signal. The two systems can be easily interchanged.

The Cs5 Ramsey cavity has been carefully designed in order to avoid microwave leakages [3]. Moreover, its Q_c factor is low ($Q_c \sim 650$), and its resonance frequency is thermally tuned in order to cancel the cavity pulling effect [4]. Finally, this cavity operates with the optimum microwave excitation level [5]. The main default of our Ramsey cavity remains the phase difference $\Delta\phi$ between the two oscillating fields in the two arms. It is well known that the determination of $\Delta\phi$ is difficult.

Cs5 is optically driven by a DBR laser diode tuned to the cesium D_1 line. Most of the experiments have been achieved with light tuned to the F=4→F=3 transition of the D_1 line and σ polarized [6]. The electronic set up is the same as in Cs4 resonator. A square frequency modulation waveform is used to interrogate the atomic beam and a digital PC controlled servo system locks the 10 MHz VCXO [7].

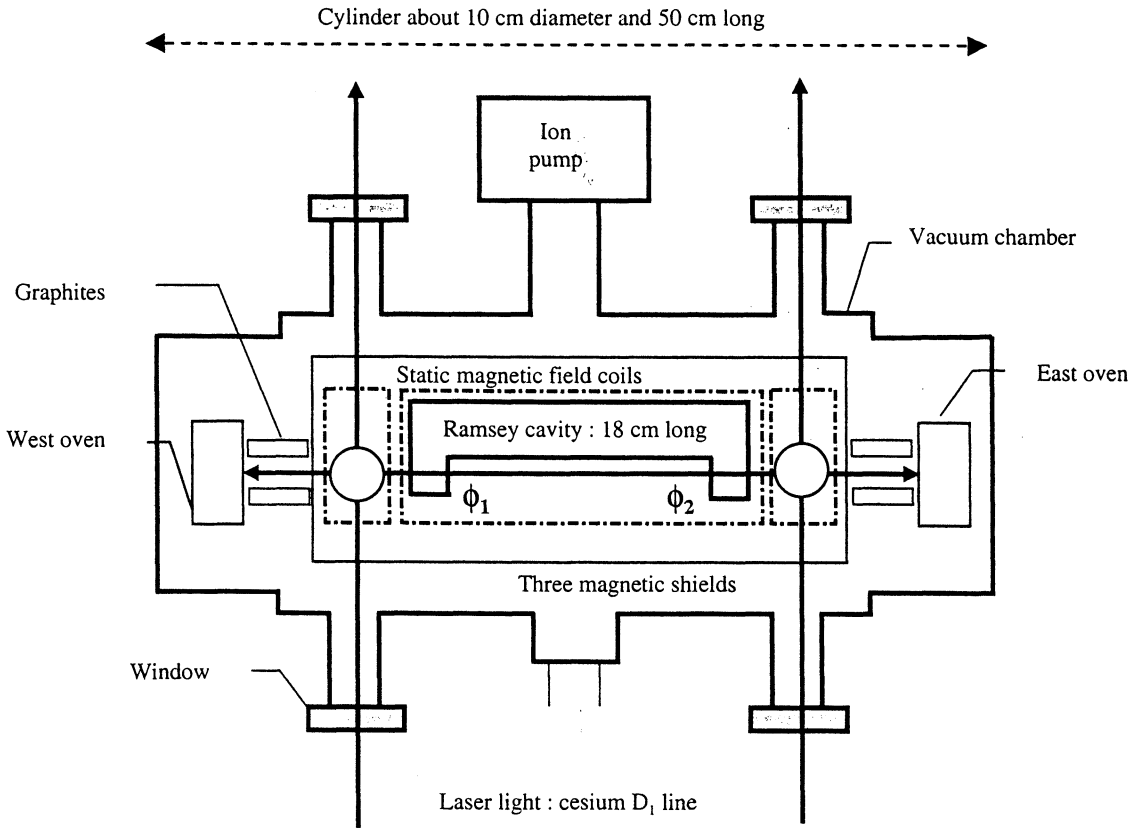


Figure 1 : Sketch of the symmetrical OPCsBR named Cs5

2. FREQUENCY MEASUREMENTS

2.1 Frequency stability

The first evaluation concerns the frequency stability delivered by Cs5. The short term stability has been evaluated first. The measured signal to noise ratio is about 5000 in one Hertz bandwith with the optical configuration described above. This value appears to be reasonable if we take into account the cesium flux decrease due to identical graphite diaphragms. The short term frequency stability of Cs5 has been measured against the H-Maser of the laboratory. Expressed in Allan standard deviation, it is about $8.10^{-12}.\tau^{-0.5}$ (see in Fig 2) for τ between 1 and 4000 seconds. The medium term frequency stability has been measured against a HP5071A high stability frequency standard. We have found about 4.10^{-14} for τ = 1 day see Fig 3.

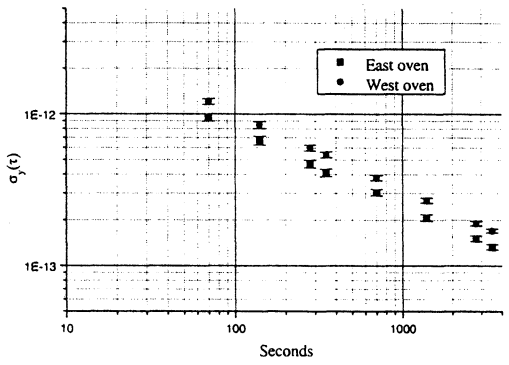


Figure 2 : Short term frequency stability of Cs5

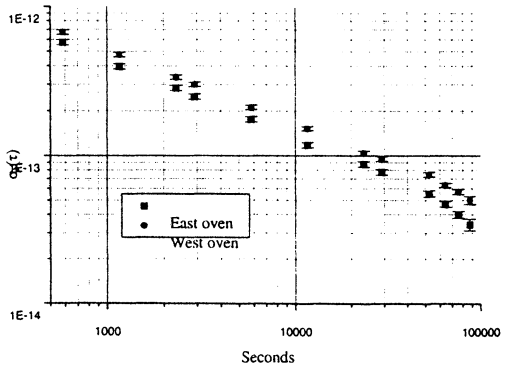


Figure 3 : Medium term frequency stability of Cs5

2.2 Residual phase difference measurements

We define ϕ_1 and ϕ_2 as the phase of the microwave magnetic field in the two cavity arms (see Fig 1). We have $\Delta\phi = \phi_1 - \phi_2$ for example. In the following, $\Delta\phi$ has been measured by two methods. The first one is the classical beam reversal method [8]. A frequency measurement is done when the East oven operates and then when the West oven does. All the operational parameters are the same (modulation depth, Rabi angular frequency, oven temperature,...). All the frequency shifts are theoretically identical for the two beams except the residual phase shift which changes its sign with the beam direction. We have :

$$\Delta\phi = \frac{f_{measured-EAST} - f_{measured-WEST}}{2F(b,\omega_m)} \qquad (3)$$

where $F(b,\omega_m)$ is a given factor which can be calculated along the lines defines in ref 8.

$$\sigma_{\Delta\phi} = \frac{\sigma_{measured-EAST}}{2F(b,\omega_m)} + \frac{\sigma_{measured-WEST}}{2F(b,\omega_m)} + \Delta\phi.\frac{\sigma_F}{F(b,\omega_m)} \qquad (4)$$

The uncertainties $\sigma_{measured\text{-}EAST}$ and $\sigma_{measured\text{-}WEST}$ are related to the frequency stability. All measurements are made with an integration time of about one day. The frequency uncertainty is then estimated to $4.5.10^{-14}$ (0.4 mHz). The uncertainty on $F(b,\omega_m)$, σ_F, is estimated equal to 5 for a nominal value of -253 Hz/rad. It is related to the microwave power measurement uncertainty. Figure 4 shows the results of measurements of $\Delta\phi$ as a function of $F(b,\omega_m)$ by beam reversal technique. This last quantity varies by changing b or ω_m. The dotted line represents the mean value of $\Delta\phi$. We have :

$$\Delta\phi = +\,21.9\ \mu\text{rad} \pm 0.9\ \mu\text{rad}.$$

The associated frequency shift is :

$$\left.\frac{\Delta f_0}{f_0}\right|_{EAST} = \frac{\Delta\phi}{f_0}.F(b,\omega_m) = -\,6.10^{-13} \pm 3.7.10^{-14}.$$

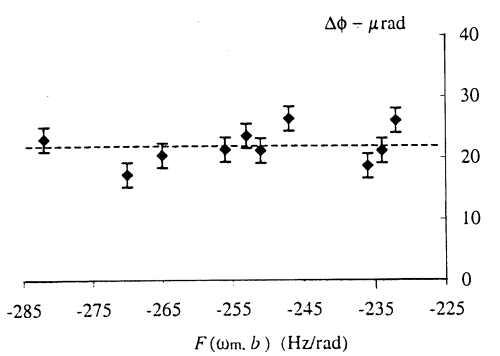


Figure 4 : Measurement of $\Delta\phi$ by beam reversal method.

The second measurement method has been developped first at the LPTF on JPO primary frequency standard [9]. The residual phase difference measurement requires only one Cs beam. This method is based on a set of measurements of the clock frequency at various operating conditions concerning b and ω_m. The measured frequencies are then corrected from the well known frequency shifts (second order Doppler shift essentially ; cavity pulling effect is not taken into account because the cavity is tuned). Let us call f_{EAST} (f_{WEST}) these corrected frequencies for the East beam (West beam respectively). They can be written :

$$f_{EAST} = (f_0 - f_{0\text{-}HP}) + \Delta\phi.F(b,\omega_m) \qquad (5)$$
$$f_{WEST} = (f_0 - f_{0\text{-}HP}) - \Delta\phi.F(b,\omega_m)$$

where f_0 is the theoretical unperturbed cesium frequency and $f_{0\text{-}HP}$ is our reference frequency given by the HP5071A apparatus.

According to eq. 5, f_{EAST} and f_{WEST} vary linearly with $F(b,\omega_m)$. A set of frequency measurements has been done with several values of b and ω_m. The corrected f_{EAST} and f_{WEST} values are then derived and reported as a function of $F(b,\omega_m)$ on Fig 5. A linear regression by the method of least squares identifies the slope and the ordinate at the origin y_0. Following (5), the slope represents $\Delta\phi$ and y_0 the frequency offset of our reference. We obtained :

East beam operation

$\Delta\phi$ = + 25.5 μrad ± 32 μrad and
$y_{0\text{-}EAST}$ = - 4.7 mHz ± 8.1 mHz,

West beam operation

$\Delta\phi$ = + 27.8 μrad ± 29 μrad and
$y_{0\text{-}WEST}$ = - 7.5 mHz ± 7.4 mHz.

The frequency shift corrections are respectively

$$\left.\frac{\Delta f_0}{f_0}\right|_{EAST} = -\,7{,}0.10^{-13} \pm 8{,}9.10^{-13}$$

$$\left.\frac{\Delta f_0}{f_0}\right|_{WEST} = +\,7{,}7.10^{-13} \pm 8{,}2.10^{-13}.$$

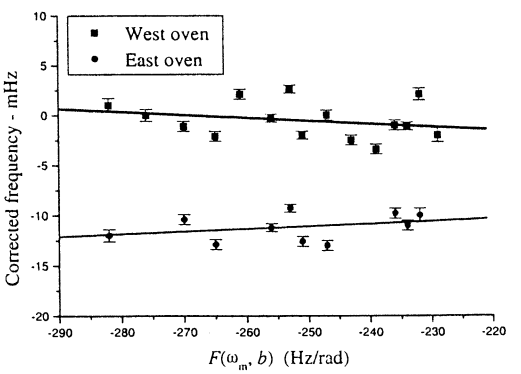


Figure 5 : Measurements of $\Delta\phi$ by a slope method

We note that $\Delta\phi$ and y_0 deduced from this method suffer from a large uncertainty due to large

electromagnetic instabilities in the laboratory which occur during this measurement campaign. It may be also due to the fact that the short term frequency stability is not good enough when unusual values of b and ω_m are set. Concerning the results delivered by this method, it appears that we obtained a good order of magnitude for $\Delta\phi$. The y_0 value is also in agreement with the result of calibration of our reference against the UTC-OP via a GPS link which was $f_0 - f_{0\text{-HP}} = -2.4$ mHz.

3. Cs5 FREQUENCY EVALUATION

The most significant frequency shifts have been evaluated for this clock. The uncertainty on most of them are related to frequency stability limitations which fix large integration times. The main frequency shifts and their uncertainties are gathered in Table 1.

Frequency shift	Rel. value	Uncertainty
Zeeman effect	56386.10^{-14}	7.10^{-14}
Second order Doppler	$-3.9.10^{-13}$	2.10^{-14}
Cavity pulling	0	6.10^{-15}
Rabi/Ramsey pulling	0	2.10^{-14}
Light shift	0	4.10^{-14}
Microwave leakages	0	4.10^{-14}
Phase shift with beam reversal technique	$\pm 6.10^{-13}$	$3.7.10^{-14}$
Phase shift with East beam configuration	$-7.0.10^{-13}$	$8.9.10^{-13}$
Phase shift with West beam configuration	$7.7.10^{-13}$	$8.2.10^{-13}$
Clock accuracy Beam reversal technique Slope measurement technique		 **1.10^{-13}** **9.10^{-13}**

Table 1 : Frequency shifts and their uncertainties

With the beam reversal technique, the frequency accuracy is 10^{-13}. It is obtained by a sum of effects which have the same order of magnitude. In the other case, the $\Delta\phi$ value is less accurate and the frequency accuracy is clearly limited by the phase difference uncertainty. The clock frequency accuracy is slightly better than 10^{-12}.

CONCLUSION

Two methods of measurement of the residual phase difference between the two arms of the Ramsey cavity have been implemented in a small OPCsBR. One of them used the Makdissi and de Clercq method. This method is time consuming. Morevover the results were spoilt by unusual large uncertainties due to electrical perturbations present during the course of the measurements. The second method is the classical beam reversal technique. It leads to a more accurate determination of $\Delta\phi$. It is worth noting that only the last method brings a significant accuracy improvement.

ACKNOWLEDGEMENTS

The authors are grateful to Délégation Générale de l'Armement DSP-STTC for its financial support. One of us L. Chassagne would like to acknowledge a PHD grant from Tekelec Telecom. We also thank M. Baldy, D. Gin, F. Locre, C. Sallot from this company for helpful discussions.

REFERENCES

[1] P. Petit, V. Giordano, N. Dimarcq, P. Cérez, C. Audoin, G. Théobald, « Miniature Optically Pumped Cesium Beam Resonator », Proceedings of the 6th EFTF in Noordwijk, 1992, p. 83.

[2] C. Audoin, V. Giordano, N. Dimarcq, P. Cérez, P. Petit, G. Théobald, « Properties of an Optically Pumped Cesium-Beam Frequency Standard with $\Phi=\pi$ between the Two Oscillary Fields », IEEE Transactions on Instrumentation and Measurement, Vol 143, N°4, August 1994, p. 515.

[3] B. Boussert, G. Théobald, P. Cérez, E. de Clerq, « Frequency Shifts in Cesium Beam Clocks Induced by Microwave Leakages », IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control, Vol 45, N°3, May 1998, p. 728.

[4] L. Chassagne, "Amélioration de l'exactitude d'une horloge atomique compacte à jet de césium pompé optiquement", Thèse de Doctorat Paris XI, 2000.

[5] C. Audoin, F. Hamouda, L. Chassagne, R. Barillet, « Controlling the Microwave Amplitude in Optically Pumped Cesium Beam Frequency Standards », IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control, Vol 46, N°2, March 1999.

[6] N. Dimarcq, V. Giordano, G. Théobald, P. Cérez, « Comparison of Pumping a cesium beam tube with D_1 and D_2 lines », J. Appl. Phys., Vol 6, N°3, February 1991, p. 1158.

[7] F. Hamouda, R. Barillet, P. Cérez, « AVersatile PC Controlled Servo-System for High Performance Optically Pumped Cesium Beam Frequency Standard », Proceedings of the 12th EFTF in Warsaw, 1998, p. 553.

[8] J. Vanier, C. Audoin, « The quantum physics of atomic frequency standards », Bristol and Philadelphia, IOP Publishing, 1989.

[9] A. Makdissi, E. de Clercq, « Estimation of the End-to-End Phase Shift Without Beam Reversal in Cs Beam Frequency Standards », IEEE Transactions on Instrumentation and Measurement, Vol 48, N°2, April 1999, p. 500.

COMPLETE EVALUATION OF A PerkinElmer RAFS IN THE GALILEO CONTEXT

Jérôme DELPORTE, Michel BRUNET, Thierry TOURNIER

Centre National d'Etudes Spatiales
18 avenue Edouard Belin - 31401 TOULOUSE Cedex 4 - France
Phone : 33+5.61.27.44.36 - Fax : 33+5.61.28.26.13
Email : jerome.delporte@cnes.fr

ABSTRACT

In navigation satellite systems, it is necessary to determine the difference between the onboard time and the reference time for each satellite. This offset can be estimated in real time by filtering time measurements collected over a ground station network. The estimate is used to generate a model for predictions when the satellite is out of visibility of this network and for the generation of the clock parameters which are inserted into the navigation message. This analysis has been carried out at CNES in the GNSS-2 context and leads to specifications for adjustment and prediction errors of the onboard time.

The purpose of this paper is to present our evaluation of a PerkinElmer Rubidium Atomic Frequency Standard Evaluation Unit (hereafter referred to as RAFS EU) in comparison to these specifications. This evaluation takes into account the environmental conditions estimated at MEO. Temperature variations, magnetic field and radiation have been successively applied to the RAFS EU in order to determine its sensitivity to each parameter and the impact on its performance.

First, CNES current specifications for Galileo will be briefly presented and explained. Second, the CNES test bed and test plan for the clock evaluation will be reviewed. Third, the PerkinElmer RAFS EU will be presented and, last, CNES evaluation results in stable and expected environment will be provided in comparison to specifications.

1. INTRODUCTION

Since 1996, CNES has conducted studies on global navigation satellite concepts for GNSS-2, one with a LEO/GEO combination, the other based upon a MEO constellation. Within CNES, several departments are involved in these studies : the precise orbit determination group works on the system analysis and the ODS (orbit determination and synchronization) overall performance, the radionavigation department is in charge of the signal and payload aspects and the time-frequency department works on the characterization of ultra-stable oscillators and atomic clocks.

For several reasons, the MEO option was chosen by EC, so our studies continue only in this direction. In 1998, CNES decided to evaluate the performances of the existing atomic clocks. The PerkinElmer Rubidium Atomic Frequency Standard was borrowed for this purpose.

2. ONBOARD CLOCK SPECIFICATIONS

In any navigation satellite systems, the ability to transmit the satellite position (ephemeris) and the time of the emission (synchronization) very accurately to the users is a key point for the performance of the system. To perform this function, the following architecture has been studied :

- dual-frequency pseudo-range and carrier phase data are collected over a ground station network
- these data are upload in quasi real time through a dedicated uplink
- onboard software using a Kalman filter computes these measurements to provide a precise position-velocity-time estimation
- these data are then used to compute the navigation message parameters.

This architecture allows the navigation message to be refreshed as often as needed to meet the performance requirements, reduces the operational constraints on the ground and it avoids expensive inter-satellite links to update the navigation message. Moreover, the same measurements are used for orbit and clock estimation, as they are correlated it reduces some onboard biases for the user.

A covariance analysis was carried out at CNES in 1997 [1] with the objective of determining the mean precision

of onboard synchronization estimation and propagation. Position, velocity and time are estimated by a least-square adjustment of pseudo-range measurements collected over a ground station network. The least-square model is then applied to a time interval corresponding to the maximum non-visibility period for the ground network considered.

The assumptions for this covariance analysis are the following :

- 5 synchronization stations
- orbital clock variations < 2 ns
- station synchronization : 1.9 ns bias, 1.2×10^{-14} drift
- orbit errors : 1m radial, 3.3m along track, 3.1m across track

The table below presents the sensitivity of the performance budget (in nanoseconds) :

	at the end of the adjustment interval	at 3.5 h extrapolation
Troposphere	0.17	0.28
Station synchro	2.11	3.48
Orbit	3.32	5.34
Station position	0.68	1.17
Receiver noise	0.27	0.47
Rb noise	1.48	3.67
Thermal clock variations	1.64	3.89
Station delay variations	0.98	1.69
Total (RSS)	**4.68**	**8.58**

Table 1 : Sensitivity of the performance budget

In this analysis, the orbit is supposed to be determined by a separate filter, which means that it isn't correlated to time. This hypothesis is pessimistic for the URE budget because clock and radial errors are mixed in the pseudo-range and carrier phase measurements.

The clock time difference is adjusted by a least-square parabolic function during a time interval T_{adj} and the resulting model is extrapolated during a time interval T_{ext}. So the onboard clock requirements are expressed in terms of :

- S_{adj} : quadratic mean of adjustment residuals over the whole adjustment interval T_{adj}
- S_{ext} : maximum extrapolation error after T_{ext}.

The following graph illustrates these onboard clock specifications :

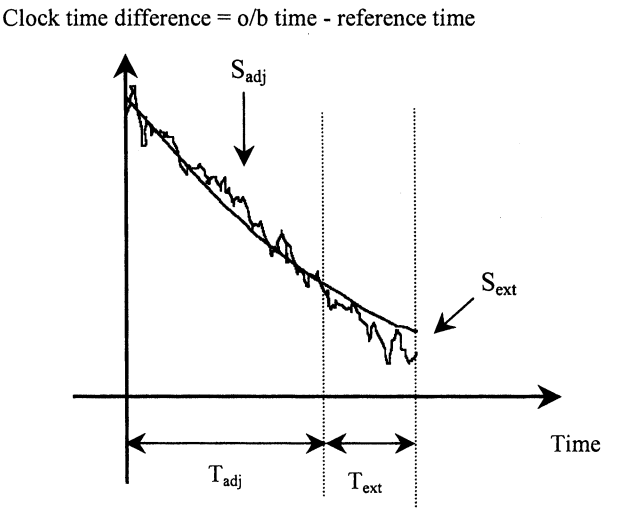


Figure 1 : Onboard clock model

CNES preliminary time-frequency specifications for GNSS-2 payload are the following :

T_{adj}	S_{adj}	T_{ext}	S_{ext}
3 h	< 1.1 ns	-	-
24 h	< 2.1 ns	3.5 h	< 5 ns

Table 2 : Onboard clock specifications

3. CNES TEST BED AND TEST PLAN

In order to verify if an atomic clock meets these requirements, the time-frequency department has developped a specific navigation test bench.

This test bench performs simultaneous time (every second) and frequency (every 10 seconds) measurements. A PC monitors the whole test bench, acquires the time and frequency measurements as well as environmental, telemetry and consumption measurements. A synopsis is given below :

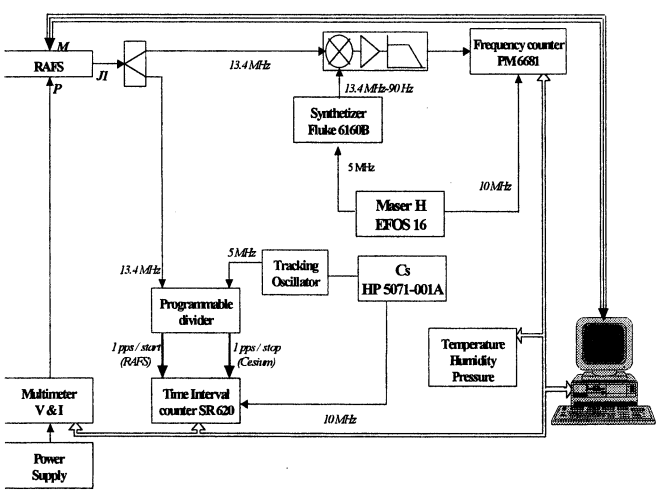


Figure 2 : CNES test bed

Using the same HP 5071-001A as reference and clock under test, we have measured a very low residual noise of the test bench :

T_{adj}	$<S_{adj}>$	T_{ext}	$\sigma(S_{ext})$
3 h	0.13 ns	-	-
24 h	0.2 ns	3.5 h	0.3 ns

Table 3 : Residual noise of the test bench

The test plan for the RAFS evaluation falls into 4 parts :

- stable environment
- thermal vacuum
- magnetic field in vacuum
- radiation test in vacuum

The variations of each environmental parameter follow CNES/ASPI studies of the MEO environment for the GNSS-2 payload [1]. Temperature and magnetic field are expected to follow a sinusoidal variation over the orbital period (12 hours), their magnitude being ± 1°C and ± 1μT. As for radiation, ONERA-Toulouse studies have determined the MEO radiation profile :

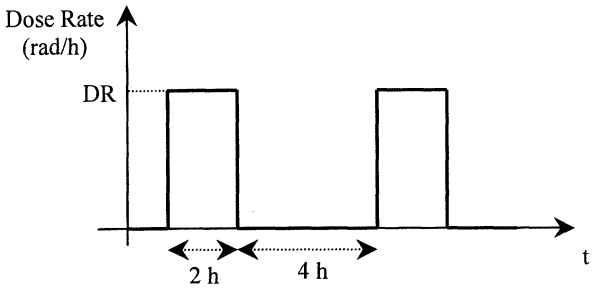


Figure 3 : MEO dose profile

The dose rate DR depends on the Al equivalent thickness of the clock and was adjusted to simulate the expected operating environment.

4. PerkinElmer RAFS

The PerkinElmer RAFS [3] has been developped specifically for satellite applications. The goal was to offer an aging rate in the low $pp10^{14}$/day range and a drift-corrected 1-day stability in the low $pp10^{14}$ range.

The RAFS has a natural frequency output of 13.40134393 MHz that is a submultiple of the Rb atomic resonance. This concept eliminates the RFS secondary loop and ground tuning sections, reducing size, weight, power consumption and complexity while improving performance, radiation hardness and reliability.

The physics package is based upon classical Rb gas cell frequency standard principles. The EU lamp operates in the Xe-Rb mixed mode and is excited by 0.45W of RF power at 105 MHz. A separate oven contains the filter cell, the temperature of this oven being adjusted for ZLS (Zero Light Shift). The absorption cell buffer gases are chosen for narrow bandwidth and low TC (Temperature Coefficient). The absorption cell length is optimized for maximum signal/noise ratio at the chosen light intensity and temperature, the latter being the coolest practical for operation at a baseplate temperature of +45°C. The 3 ovens are supported by a fiberglass structure that makes effective use of the vacuum environment for thermal insulation. The RAFS operates at uniform low fixed magnetic field (provided by a 2-section C-field coil configuration) for increased stability. The RAFS chassis is controlled at +45°C by an integral baseplate temperature controller.

The radiation hardening has been a critical aspect of the design : all critical parts and subassemblies have undergone radiation testing (in particular the physics package). The results of these tests have been analyzed by PerkinElmer and have revealed that the RAFS can meet the stringent clock performance requirements under natural or manmade radiation.

5. RESULTS OF THE EVALUATION

CNES evaluation of the RAFS has followed the previously described test plan. The axes of the RAFS are defined as follows :

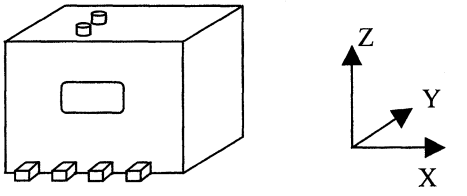


Figure 4 : Axes of PerkinElmer RAFS

5.1 Stable environment

This first test was performed in vacuum at 20°C and lasted more than 10 days. Frequency stability is presented below : $\sigma_y(\tau) = 6.5 \times 10^{-13}.\tau^{-1/2}$ for $0.1 < \tau <$ 5000s :

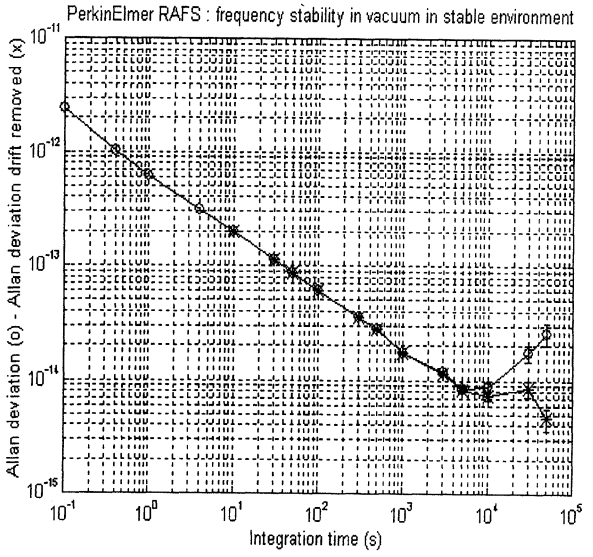


Figure 5 : PerkinElmer RAFS Allan deviation

Datation performance is summarized in the table below:

	Datation performance
$<S_{adj}>$ for T_{adj} = 3 h	< 0.18 ns
$<S_{adj}>$ for T_{adj} = 24 h	< 0.42 ns
$\sigma(S_{ext})$ for T_{ext} = 3.5 h	< 1.17 ns

Table 4 : Datation performance in stable environment

This performance is deemed to be excellent. We have now to estimate the sensitivities to environment and the ensuing degradation in performance.

5.2 Thermal vacuum

The temperature of the thermal chamber was varied so that the RAFS underwent a sinusoidal variation of temperature of 20 ± 1°C with a 12 hour period. This test lasted approximately 10 days.

No frequency sensitivity to temperature variations was observed. This is in agreement with PerkinElmer specifications : temperature sensitivity is specified below the noise level for ± 2.5°C temperature variations with the baseplate temperature controller.

The table below summarizes datation performance during this test :

	Datation performance
$<S_{adj}>$ for T_{adj} = 3 h	< 0.17 ns
$<S_{adj}>$ for T_{adj} = 24 h	< 0.37 ns
$\sigma(S_{ext})$ for T_{ext} = 3.5 h	< 1 ns

Table 5 : Datation performance during the thermal vacuum test

5.3 Magnetic field in vacuum

For this test, the RAFS is still in its vacuum chamber but no longer in the thermal chamber so its temperature is not controlled as well as for the previous tests. The temperature measured by our probe on the baseplate stayed between 29 and 31.5°C during the whole test, around 10°C above the mean temperature of the previous test. So the thermal sensitivity may not be negligible and the effect of thermal and magnetic variations can not be separated on the datation performance.

The magnetic field is provided by Helmholtz coils and was applied on each axis of the RAFS successively. The sensitivity specified by PerkinElmer is $< 1x10^{-12}$/G that is $1x10^{-14}/\mu T$. The Earth magnetic field at MEO is estimated at ± 1μT, this value is so small that the resulting frequency shift couldn't be measured by our test bed. So we decided to apply a larger value : ± 50 μT with a 12 hour period.

The only axis in which sensitivity was detected is the X axis, the most sensitive one because it is in line with the C-field. This sensitivity was estimated to be around $3x10^{-13}$/G.

The table below summarizes the datation performance measured on each axis :

	X axis	Y axis	Z axis
$<S_{adj}>$ for T_{adj} = 3 h	< 0.17 ns	< 0.18 ns	< 0.17 ns
$<S_{adj}>$ for T_{adj} = 24 h	< 0.80 ns	< 0.49 ns	< 0.43 ns
$\sigma(S_{ext})$ for T_{ext} = 3.5 h	< 2.03 ns	< 1.56 ns	< 1.48 ns

Table 6 : Datation performance during the magnetic field test

This table confirms that the RAFS has a very low sensitivity to magnetic field : the datation performance is slightly degraded during applied variations of magnetic field 50 times larger than expected.

5.4 Radiation in vacuum

The radiation test was performed at ONERA-Toulouse. The radiation used was γ-rays provided by a MILGA Cobalt 60 source. The time and frequency reference was a HP Cs 5071A-001 with a Tracking Oscillator. This test lasted around 2 months and the radiation lasted 10 days.

Taking into account the equivalent Al thickness of the test unit, the vacuum chamber wall thickness, and the characteristics of the Co^{60} source, the proper distance between the RAFS and the Co^{60} source was calculated.

The RAFS was positioned at the appropriate distance. The actual dose rate was recorded using a radiation probe that was located beside the RAFS. A slightly higher than expected dose rate was measured during the test.

Unfortunately, the room temperature was very high (between 35 and 37.5°C), during part of the test and exceeded the limit over which thermal regulation within the clock can be maintained. However during a limited period, we did have radiation with temperature below the limit : during this period, no frequency sensitivity was observed.

The table below summarizes the datation performance for 2 periods during which the temperature was below the limit :

	Radiation	No radiation
$<S_{adj}>$ for T_{adj} = 3 h	< 0.18 ns	< 0.19 ns
$<S_{adj}>$ for T_{adj} = 24 h	< 0.36 ns	< 0.44 ns
$\sigma(S_{ext})$ for T_{ext} = 3.5 h	< 1.32 ns	< 1.4 ns

Table 7 : Datation performance during the radiation test

This table clearly shows that the radiation we applied had no measurable influence on the datation performance.

CONCLUSION

A complete evaluation of the PerkinElmer RAFS in the Galileo context as foreseen by CNES has been performed.

We have measured a frequency stability of $\sigma_y(\tau) = 6.5 \times 10^{-13} . \tau^{-1/2}$ for $0.1 < \tau < 5000$s.

Measurements have showed no sensitivity to environment except for the X axis in the magnetic field test (because we have applied 50 times the field specified). We have evaluated the sensitivities to environment by a least mean square method : we have adjusted in the same time a linear frequency drift and a sensitivity coefficient. This method leads to upper bounds of thermal and radiation sensitivities relative to the experiment observability capability.

The table below summarizes these upper bounds :

	Sensitivity
Temperature (*)	$< 2.4 \times 10^{-14}$ / °C (2σ)
Magnetic field (*)	2.7×10^{-15} / µT $\pm 5 \times 10^{-16}$ (2σ)
Radiation (**)	$< 1.6 \times 10^{-14}$ / rad (2σ)

Table 8 : RAFS frequency sensitivities to environment

(*) : sinusoidal variations with a period of 12h.

(**) : following the profile described in section 3.

The datation performance is slightly degraded by the environments. The table below shows the upper bound of the performance of the RAFS in the overall representative environment (that is accumulating temperature, magnetic field and radiation variations) and compares it to the CNES preliminary time-frequency specifications for the future Galileosat payload :

	CNES spec	RAFS performance
$<S_{adj}>$ for T_{adj} = 3 h	*< 1.1 ns*	< 0.2 ns
$<S_{adj}>$ for T_{adj} = 24 h	*< 2.1 ns*	< 0.5 ns
$\sigma(S_{ext})$ for T_{ext} = 3.5 h	*< 5 ns*	< 1.5 ns

Table 9 : Estimation of the datation performance in representative environment

Since the datation performance of the RAFS is excellent, our Cs reference may contribute to the above figures. This contribution is under evaluation, that's why we announce only upper bounds of the RAFS datation performance. Anyway, this table shows that the PerkinElmer RAFS is at least three times better than the CNES preliminary time/frequency specifications for the future Galileosat payload.

ACKNOLEDGMENTS

The authors wish to thank Jean-François Dutrey and Philippe Guillemot (CNES time-frequency department) for their constant and essential support, and Jacques Bourrieau (ONERA/CERT/DESP) for his help in the radiation tests. We also wish to thank PerkinElmer for the loan of the RAFS and for reviewing the paper.

REFERENCES

[1] "A precise onboard real-time ephemeris and clock estimation for GNSS-2 concepts studied at CNES", T. Tournier - GNSS 98 , IX-O-12

[2] "CNES preliminary time/frequency specifications for GNSS-2 payload", T. Tournier, J. Delporte - Note CNES/DTS/AE/TTL/TF 99-082.

[3] "Rubidium Atomic Frequency Standards for GPS Block II R", William J. Riley- ION GPS-92, 5th International Meeting, September 16-18, 1992

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

DEVELOPMENT OF FLIGHT TECHNOLOGY FOR FUTURE LASER-COOLED SPACE CLOCKS

R.J. Thompson, W.M. Klipstein, J. Kohel, D.J. Seidel, and L. Maleki, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, 91109

Mailing address: R.J. Thompson, MS 298/100, Jet Propulsion Laboratory, 4800 Oak Grove Drive, Pasadena, California 91109, U.S.A.

ABSTRACT

In this paper we present an overview of two missions to fly ultra-high precision atomic clock experiments aboard the International Space Station (ISS) and perform tests of general relativity theory. In particular, we address the role played by the Jet Propulsion Laboratory (JPL) in developing the core technologies to support these and future flight projects. We also give an overview of the Japanese Experiment Module External Facility (JEM-EF) platform and the challenges of performing a laser cooled atomic clock experiment in this environment.

1.0 INTRODUCTION

Because of the limitations imposed by gravity on terrestrial laser cooling experiments, NASA's Fundamental Physics program has identified Laser Cooling and Atomic Physics (LCAP) as one of three disciplines (along with low-temperature and gravitational physics) which are poised to take advantage of the micro-gravity environment offered by the ISS and the Space Shuttle. Two clock experiments have been selected by NASA for flight aboard the International Space Station (ISS): the Primary Atomic Reference Clock in Space (PARCS), with principle investigators at the National Institute of Standards and Technology (NIST) and the University of Colorado; and the Rubidium Atomic Clock Experiment (RACE), with principle investigator at Yale University. In addition, 12 ground-based investigations have been funded to date on topics including atomic clocks, Bose-Einstein Condensation (BEC), Electric Dipole Moment (EDM) searches, and atom interferometry.

The Time and Frequency Sciences and Technology Group at the Jet Propulsion Laboratory plays a key role in this program, supporting LCAP missions through the design, construction and integration of instruments capable of meeting the science goals. Ground testbeds in our laboratory are used to refine our designs and aid in the development of new technologies required for the flight missions.

2.0 OVERVIEW OF THE PARCS AND RACE CLOCK EXPERIMENTS

Two atomic clock experiments will form the first generation of LCAP flight projects. Each will rely on micro-gravity for their performance and, in addition, will utilize the difference in the gravitational potential between the earth's surface and the International Space Station to perform a variety of tests of the theory of general relativity. The first of these, PARCS, has as Co-Principle Investigators Dr. Don Sullivan, and Dr. Bill Phillips of NIST, and Professor Neil Ashby of the University of Colorado. PARCS is planned to fly aboard the space station late in 2004. The other flight experiment, RACE, is led by Principle Investigator Kurt Gibble, of Yale University. This experiment differs from the PARCS experiment in that it utilizes a different atomic species (rubidium as opposed to cesium) which can result in a dramatic reduction of the collisional frequency shift, a significant accuracy limitation for cold atomic clocks. RACE is expected to fly 16 months after PARCS. We are currently targeting each of these missions to fly on the Japanese Experiment Module's External Facility (JEM-EF), an external platform aboard the International Space Station (ISS). The missions will have a duration between six months and one year.

Each of these missions is currently in its definition phase, in which their scientific goals, and the instrument requirements to meet these goals, are being developed. PARCS is scheduled to move into the design/build phase in the fall of this year, with RACE expected to begin design and build in spring of 2002.

Detailed descriptions of the two instruments have been presented elsewhere.[1,2] We present here an overview of some of the common features. Each of the two clocks will utilize a lin⊥lin optical molasses source (RACE will feature an additional collection region, to obtain larger numbers of atoms). In order to obtain high stabilities while minimizing the effect of the spin-exchange collision shift, each clock will launch multiple balls into the microwave cavity in a quasi-cw manner. This requires the instruments to utilize a system of mechanical shutters, so that those atoms within the microwave cavity are not effected from radiated light from either the source, state selection, or

detection regions. The local oscillator for the PARCS mission is a space qualified hydrogen maser built by the Harvard Smithsonian Center for Astrophysics (SAO), with a short-term stability of better than $\sigma_y(\tau)=5\times10^{-14}\tau^{-1/2}$. [5] A local oscillator has not yet been selected for the RACE mission. Each mission will utilize GPS carrier-phase measurements both for time transfer and precise determination of the ISS orbit and velocity.

3.0 INSTRUMENT DEVELOPMENT

Instrument development and PI support for both of the current flight projects and for later LCAP flights is provided by the Time and Frequency Sciences and Technology Group of the Jet Propulsion Laboratory. Work in our ground test bed has focused initially on demonstrating the feasibility of performing laser cooling experiments in space. Here the challenges are to dramatically reduce the volume and mass requirements of a typical laser-cooling experiment, while improving the reliability and ruggedness of the apparatus. The hardware must be capable of surviving a typical space shuttle launch, and of operating autonomously for several months.

One of the most challenging technologies is the development of rugged compact laser systems, capable of producing high power single frequency laser light with the stability and frequency tunability needed to meet the demands of a laser cooling experiment. We have demonstrated a compact laser system built within the mass and volume constraints of the PARCS mission. Our current baseline laser system for PARCS will have a single master laser (a New Focus Vortex model extended cavity diode laser) that will then be used to injection lock two higher power slave lasers. Acousto-optical modulators positioned between the master and slave laser allow one to have a very fine control of the laser frequency, while similar modulators following the slave laser allow the laser intensity to be controlled as well. We have vibration tested a variety of optical components that will be included in the PARCS laser system up to the qualification level required for a shuttle launch. These components include a Vortex master laser, and a variety of acousto-optic modulators, optical isolators and mounts. Each of the components passed the tests successfully, giving us high confidence that these devices are rugged enough to survive launch.

Another example of the unique technologies being developed at JPL is the novel non-magnetic shutter system shown in figure 1, which is required in order to utilize the launching of multiple balls of atoms without introducing light shifts. The requirements on this shutter are severe—it must be capable of performing reliably for one year in ultra-high vacuum without perturbing the micro-gravity environment, or producing any significant magnetic field, which might perturb the clock. The shutter must be capable of closing from a 1cm aperture in about 2ms.

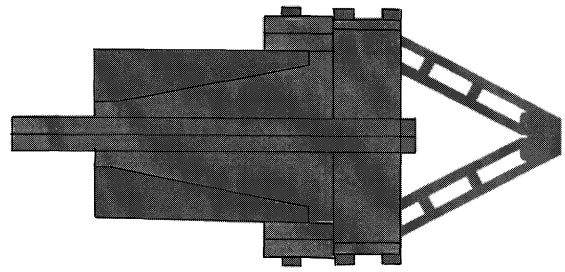


Figure 1. Conceptual design of non-magnetic shutter for the PARCS and RACE missions.

Other components and subsystems are also under development to provide ISS experimenters with the same capability currently available in ground-based laboratories. These include ultra-high vacuum systems, and sources of alkali atoms.

A key to minimizing the overall cost of the LCAP program is to develop experiments with as much modularity and reusability as possible. Thus the laser and optics subassembly for the PARCS experiment will also be suitable for a variety of future LCAP flights using atomic cesium. The RACE laser and optics subassembly, will have a similar design to the PARCS one, but will utilize different components in order to match the rubidium wavelength. Again, this system will be designed to be easily refurbished for future flights involving rubidium. Currently it is believed that most LCAP flights will utilize one of these atomic species.

4.0 THE JEM-EF PLATFORM

The ISS carries a number of different platforms, both internal and external, for science payloads. Of these, we have selected the Japanese Experiment Module's Exposed Facility (JEM-EF), shown in Fig. 2, as our preferred platform for both clock missions. This decision is based on a number of criteria, primarily the access to site that would be acceptable for placement of a GPS receiver for time transfer and precision orbit determination. It is also highly desirable to have both the receiver and the clock co-located. Co-location of equipment is required to determine the position of the clock relative to a ground station to within 10 cm, an amount that is less than the expected flexing of the Space Station structure over a typical orbit. In addition it is difficult to obtain a high quality rf or optical link between locations on the station that are widely separated (and in particular between locations inside and outside the station). The JEM-EF also allows both zenith and nadir pointing instruments, affording us the flexibility of adding a

direct downlink frequency transfer system. Finally, a location on the JEM-EF also gives proximity to the Super-conducting Microwave Oscillator (SUMO) instrument. A link between this highly stable oscillator and an atomic clock would dramatically enhance the science return of each instrument.[3] A link to the Atomic Clock Ensemble in Space (ACES) experiment would also be highly desirable[4], but might prove challenging if there is no direct line of sight for an optical link between the instruments (ACES is expected to be flown on either a nadir pointing Express Pallet located on the starboard side of the Space Station truss, or on the side of the Columbus orbiting facility).

Figure 3 shows a picture of a typical JEM-EF payload bus. Instruments mount into a volume 1.8× 1.0× 0.8m, and must have a mass < 500 kg. A maximum of 3000 W of power is allowed per payload (in practice this might be considerably reduced, because of the limits to the total amount of station power available). Up to 2000 W of heat can be taken away by a closed water loop. These numbers are extremely liberal compared to other ISS platforms.

A potential disadvantage of the JEM-EF platform arises from its proximity to various space station structures (in particular the solar panels), which limit the viewing of the GPS constellation, and may also produce multipathing problems--reflections of the GPS signal at the –70dBm level (~20 dB below the expected signal amplitude) may disturb the carrier phase solution. Finally there is a powerful Ku band communications antenna on the platform that may cause interference problems.

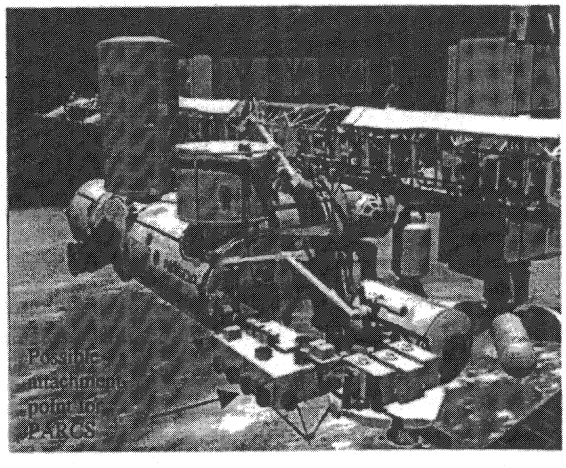


Figure 2. View of the Japanese Experimental Module's External Facility, with no payloads attached.

5.0 ACKNOWLEDGEMENTS

This work was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. We acknowledge many fruitful discussions with Don Sullivan, Leo Hollberg, Steve Jefferts, John Kitching, Hugh Robinson, Fred Walls, Tom Heavner, Tom Parker, and Dawn Meekhof of NIST,with Andrea De Marchi of Politecnico di Torino, with Neil Ashby of the University of Colorado, and with Kurt Gibble of Yale University. The picture of the PARCS/RACE shutter was produced by Fabien Nicaise of JPL.

6.0 REFERENCES

[1] S. R. Jefferts et al. "PARCS: A Primary Atomic Reference Clock in Space" Proceedings of the Joint Meeting EFTF- IEEE IFCS, 1999, pp. 141-144.
[2] Chad Fertig et al. "Laser-Cooled Microgravity clocks" Proceedings of the Joint Meeting EFTF-IEEE IFCS, 1999, pp. 145-147.
[3] J. D. Lipa, et al. "A precision test of the local position invariance principle of general relativity using precision clocks on space station" Proceedings of the STAIF conference, 2000, pp. 695-700.
[4] S. Feltham, "ACES project overview" from the proceedings of this conference.
[5] E. M. Mattison and R.F.C. Vessot, "Design of a Hydrogen Maser for Space," in Proceedings of the 11th EFTF, 1997, pp 525-528.

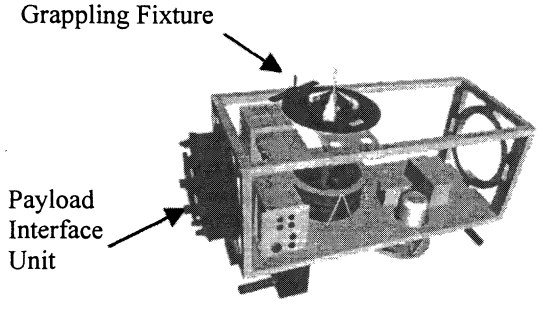


Figure 3. A sketch of the JEM-EF payload bus. (Courtesy NASA-Goddard)

DIFFERENTIAL CALIBRATION OF ASHTECH Z12-T RECEIVERS FOR ACCURATE TIME COMPARISONS

Gérard Petit, Zhiheng Jiang
BIPM, Pavillon de Breteuil, 92312 Sèvres CEDEX, France
Pierre Uhrich, François Taris
BNM-LPTF, Observatoire de Paris, 61 Ave. de l'Observatoire 75014 Paris, France
e-mail: gpetit@bipm.fr

ABSTRACT

Dual frequency carrier-phase and code measurements from geodetic like receivers are a promising tool for frequency and time transfer. We have carried out the differential calibration of two such receivers of Ashtech Z12-T type, by comparing their raw pseudo-range measurements with those of classical NBS-type time receivers operated side by side. We then perform a time comparison between the BNM-LPTF and the BIPM using both classical time receivers and Ashtech Z12-T and compare the results.

1. INTRODUCTION

Time comparisons are usually carried out with GPS time receivers using C/A code measurements. A few number of these receivers have had their electrical delays absolutely calibrated with an uncertainty of a few nanoseconds. By differential measurements of receivers operated side by side, it has been possible to calibrate over the years most of the receivers operating in time laboratories worldwide so that time comparisons used e.g. for TAI are expected to be accurate to within a few nanoseconds. Such an exercise of differential calibration is also being carried out for multi-channel code receivers.

Recently the use of dual frequency carrier-phase and code measurements from geodetic like receivers has emerged as an outstanding tool for frequency comparisons. It has been shown that the relative frequency stability between two independent Ashtech Z12-T systems is below 1×10^{-16} for averaging durations of half a day [1], and that an intercontinental frequency comparison could be achieved with an uncertainty of 1×10^{-15} over a day [2], a necessary requirement to compare recently developed frequency standards. Numerous groups have initiated time and frequency experiments with such geodetic receivers, including comparisons with other time transfer techniques [3,4].

In order to be used also for time comparisons, such receivers should be calibrated to obtain the absolute values of the electrical delays. Although a direct measurement of the propagation delays is in principle the preferred calibration technique, it is in practice easier to carry out differential calibration so only this technique is considered here. The carrier-phase measurements cannot be used for time comparison, because of the ambiguity, but they are used to smooth the pseudo-range measurements which, therefore, should be calibrated. The outline of the paper is the following: First we define in section 2 what we consider the time reference of the Ashtech Z12-T receivers. Then we present in section 3 the differential calibration of two such receivers obtained by comparing their raw pseudo-range measurements with those of classical NBS-type time receivers operated side by side. We then perform a time comparison between the BNM-LPTF and the BIPM using both NBS-type time receivers and Ashtech Z12-T, the results of which are presented in section 4.

2. THE INTERNAL REFERENCE OF THE ASHTECH Z12-T RECEIVER

The Ashtech Z12-T receiver performs pseudo-range and carrier phase measurements which are referred to an "internal reference". This reference is derived from an externally provided 20 MHz signal. An important modification of the Z12-T version with respect to the Z12 is that an externally provided 1 PPS signal allows the receiver to unambiguously choose one particular cycle of the 20 MHz to form the internal reference, therefore providing repeatability of this reference in case of any interruption of the tracking or operation of the receiver. According to the Ashtech documentation, the internal reference is the input 20 MHz inverted and delayed by 15.8 ns. Since the usual time reference in time transfer is a 1-PPS signal, it is necessary to compare the Z12-T internal reference to this signal. To do so, we direct the 1-PPS and the 20 MHz signals on two channels of a digital oscilloscope where the 1-PPS signal triggers the data acquisition. By direct

measurement on the oscilloscope display (Figure 1), it is possible to determine the relative phase of the two signals with an uncertainty lower than one nanosecond. It is then possible to ensure that, when the equipment is put into operation in a new location or after a power off, the relative phase of the internal reference is known with respect to the local reference with that same uncertainty. In addition, the relative phase of the 20 MHz and 1-PPS input signals may not be set to an arbitrary value: for proper operation of the receiver, the relative phase must take a defined value (within a range of about +/- 5 ns). The oscilloscope display also allows to check the proper configuration of the signals (see an example in Figure 1).

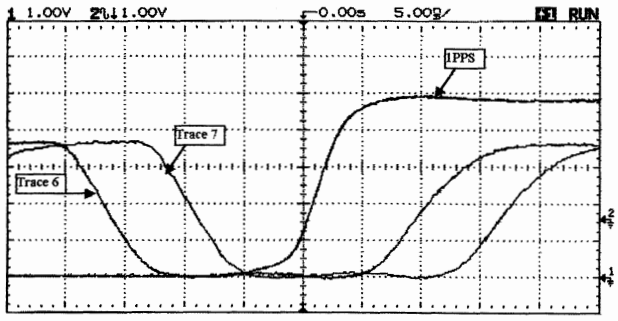


Figure 1: Oscilloscope display of two different valid configurations for the relative phase of the 1-PPS and 20 MHz (Trace 6 and 7) signals. Trace 7 is offset by 7.1 ns from Trace 6.

3. DIFFERENTIAL CALIBRATION OF ASHTECH Z12-T VS. NBS RECEIVER

In order to perform the differential calibration of the Ashtech Z12-T with respect to the NBS receiver at the BIPM and the BNM-LPTF, in each laboratory the two units have been set up with one reference clock as indicated in Figure 2. The Z12-T pseudo-range measurements (P1 and P2) for all satellites in view are taken from the RINEX files, at the nominal sample rate of 30 s while the NBS pseudo-ranges (C/A) are obtained every second as an auxiliary output, for one given satellite at a time according to the BIPM international tracking schedule. Before forming the difference of the pseudo-ranges, two effects have to be taken into account: the geometric effect due to different positions of the phase centres of the antennas and that of the different timing of the measurements (emission time for the NBS, reception time for the Z12-T). In the configuration described above, 1248 differential measurements may be obtained for each day which, owing to a measurement uncertainty of order 10 ns, provides an uncertainty on the daily mean of order 0.3 ns. We could greatly increase the number of measurements by taking 1-second measurements with the Z12-T and by optimising the NBS schedule, but the above mentioned uncertainty is sufficient given the uncertainty in the Z12-T reference (see section 2) and in the NBS calibration (see below).

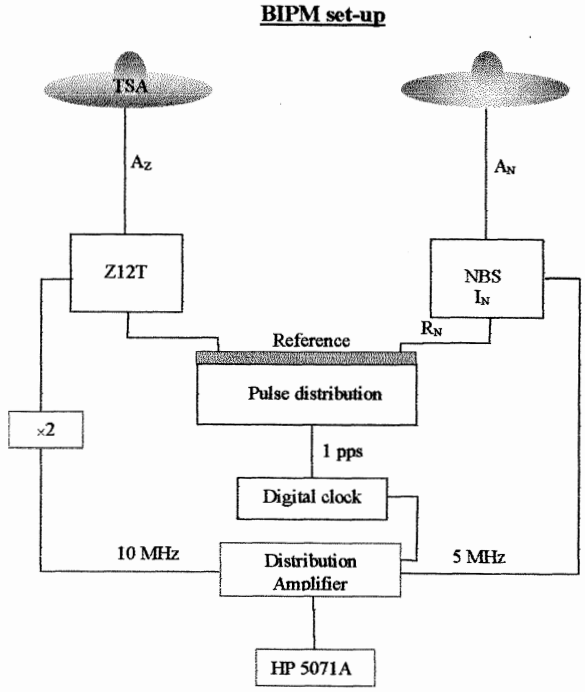


Figure 2: Experimental set-up at the BIPM. The set-up at the BNM-LPTF is similar.

The calibration of the two NBS units has been obtained as follows: The BNM-LPTF NBS51A unit has been compared to the NIST NBS10 through seven calibration trips (differential calibration with a travelling receiver) between 1986 and 1996 and the dispersion of the seven results is 2 ns (1 σ). The BIPM NBS51B unit has been directly differentially calibrated with respect to NBS10 in May 1998 with an uncertainty of 2 ns. Occasional calibration trips have been conducted between the BIPM and the BNM-LPTF and the results are in agreement with the above mentioned uncertainties. It is to be noted that the NIST has sent to the BIPM results of absolute calibrations of its NBS10 carried out in 1986, 1987 and 1998, the uncertainty of the latest being 2.8 ns. In addition, both the BIPM and the BNM-LPTF NBS receivers are continuously compared to other time transfer receivers in each laboratory. The daily mean of these differences show

some long term signatures that may be attributed to a sensitivity to the environment but the dispersion of the values is only of order 1 ns (1 σ) in both laboratories.

To conclude, we believe that the electrical delays of the NBS receivers at the BIPM and the BNM-LPTF are known within an absolute uncertainty of 3 ns but their difference is known somewhat better, with an uncertainty estimated to 2 ns. Therefore a time link between the BIPM and the BNM-LPTF performed with NBS receivers is accurate to that same uncertainty.

The differential calibration of the Z12-T vs. NBS has been performed for the BIPM units over 18 days in October 1999, January and February 2000 and for the BNM-LPTF units over 5 days in February 2000. The results obtained are the following:

P1(Z12-T) – C/A(NBS51) = +148.2 ± 0.3 ns (LPTF)

P1(Z12-T) – C/A(NBS51) = +269.1 ± 0.3 ns (BIPM)

The 18 daily means obtained at the BIPM have a normal distribution and a standard deviation of 0.5 ns. A study is under way in both laboratories in order to assess the long term stability of this differential calibration.

It should be stressed that this calibration concerns only the P1 measurement of the Z12-T. In general, when operating such geodetic receivers, the ionosphere-free linear combination of P1 and P2 (so called P3) is used. The differential calibration of the P2 vs. the P1 measurement, allowing a differential calibration of a P3 link, will be the subject of a later work.

4. TIME TRANSFER COMPARISON

With the equipment described above, it is possible to compare the BNM-LPTF and BIPM clocks (over a distance of 8.5 km) by several different methods (see Figure 3).

The first method is the classical "Common View" technique (CV) using the NBS-type receivers. Owing to the knowledge of differentially calibrated internal delays and measured delays due to the cables, this method directly provides the difference (Clock(BIPM)-UTC(OP))$_{CV}$ for about 48 13-min measurements per day. The other methods are clock solutions based on processing Z12-T phase and code measurements with the Bernese V4.1 software [5]. This may be performed using either single frequency (L1/P1) measurements or a ionosphere-free linear combination of dual frequency measurements (L3/P3). The Bernese clock solutions provide the difference between the internal references of the two receivers at any multiple of the sampling rate of 30 s, we choose here 15-minute intervals. By using the differential calibration described above, the L1/P1 solution yields the difference (Clock(BIPM)-UTC(OP))$_{Z12/P1}$. The three computations (CV, L1/P1, L3/P3) of the BIPM-LPTF time link have been carried out over a period of 5 days from MJD 51578 to 51582 (days 35 to 39 of year 2000). Figure 4 shows the comparison of (Clock(BIPM)- UTC(OP))$_{Z12/P1}$ with (Clock(BIPM)- UTC(OP))$_{CV}$ after application of the differential calibration of section 3 and Figure 5 shows the two Bernese clock solutions L1/P1 and L3/P3, without any calibration applied.

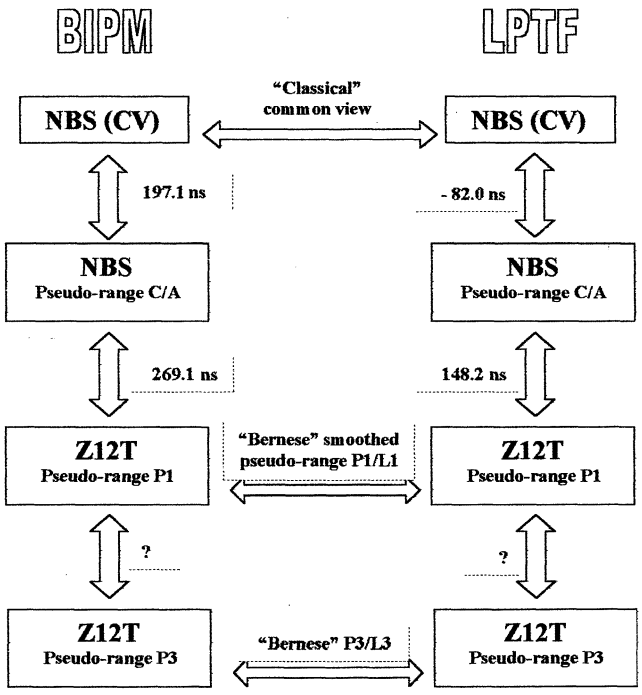


Figure 3: Clock comparison techniques between the BNM-LPTF and the BIPM. The known or measured calibration delays are indicated between each set of measurements.

In order to compare the different methods, we first have to estimate the possible effect of different uncertainty sources. In all methods, the antenna phase centre coordinates are considered known to within a few centimetres and contribute less than 0.1 ns to the time link uncertainty. The differential effects of uncertainties in the tropospheric delays on the time link are expected to be lower than 0.1 ns due to the short distance, but will cancel even better in the comparison of different methods. The differential effects of uncertainties in satellite ephemerides are also estimated to be lower than 0.1 ns. Finally, also due to the short distance, differential ionospheric effects present in the CV and L1/P1 methods should affect the time link by

less than one nanosecond, on average. They should cancel even better when comparing CV to L1/P1 and should affect the comparison between the L1/P1 and L3/P3 solutions only at the sub-nanosecond level, on average.

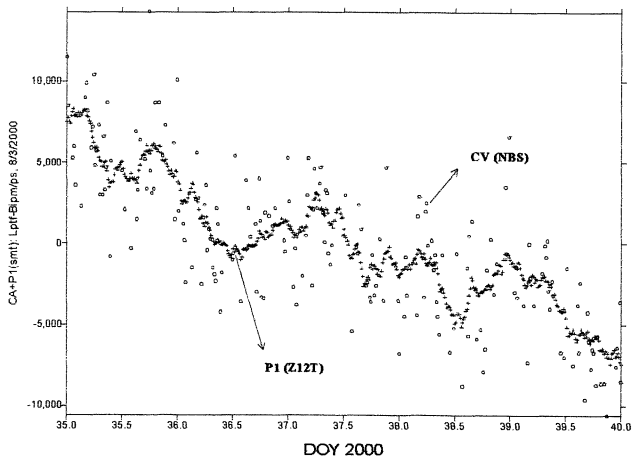


Figure 4: Comparison of the time link (Clock(BIPM)-UTC(OP)) by Z12-T L1/P1 (crosses) and NBS CV (open circles) after application of the differential calibration.

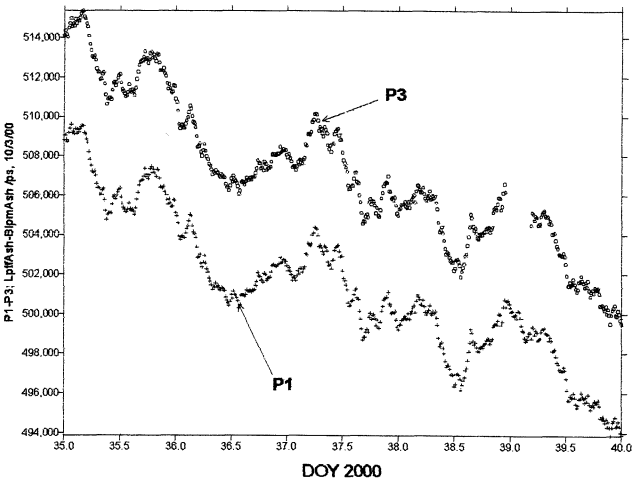


Figure 5: Comparison of the time link between the internal references of the BIPM and BNM-LPTF Z12-T by the L1/P1 and L3/P3 methods.

As shown in Figure 4, we observe that the CV method and the L1/P1 method, after differential calibration, are equivalent at the level of 0.1 ns to compute the time link between Clock(BIPM) and UTC(OP) over the indicated period. It remains to be seen whether this level of agreement may be reached over any time interval because the long term stability of the differential calibration has not been studied yet. In practice, we may expect that the level of 1 ns could be reached because this corresponds i) to the observed level of the long-term stability of calibration delays of NBS receivers (section 3) and ii) to the level to which the Z12-T internal reference may easily be compared to the NBS reference (section 2).

As shown in Figure 5, we observe that a differential calibration of the P2 and P1 measurements is necessary to use the L3/P3 method for a time link, as is required for any distance above a few tens of kilometres. We could use the result obtained here as a calibration of the difference of the two receivers: $[P3(Z12_{BIPM})-P1(Z12_{BIPM})] - [P3(Z12_{LPTF})- P1(Z12_{LPTF})] = 6.1$ ns. However it is not advisable to do so because it would be valid only for the link between these two receivers in the configuration used for this experiment. Differential calibration of the P2-P1 delay for each instrument (main unit + antenna) should be performed. There are different ways to achieve this, one being to use a calibrated signal simulator, another one being to correct the P1 and P2 measurements for the ionospheric delays by using ionosphere models, such as those that may be derived from the IGS ionospheric maps [6]. Work is being carried out in these two directions.

5. CONCLUSION

We have performed the differential calibration of two Ashtech Z12-T systems with respect to NBS time transfer receivers, which allows to use the L1/P1 measurements of the Z12-T for clock comparisons. We have shown that, over a short distance link, the NBS and Z12-T clock comparisons are equivalent at the sub-nanosecond level. We expect that the long term stability of such differential calibrations may reach 1 ns. Work is under way to calibrate the Z12-T P2 measurements so as to be able to use ionosphere free measurements for long distance time links. When this is achieved, it will be possible to use such GPS receivers for clock comparisons, e.g. used in TAI computation.

ACKNOWLEDGEMENTS

We acknowledge the NIST for the loan of the NBS51 receivers located in the BNM-LPTF at Paris Observatory and at the BIPM. We are indebted to Philippe Moussay and André Campos for the technical operations at the BIPM and BNM-LPTF, respectively.

REFERENCES

1. G. Petit, C. Thomas, "GPS frequency transfer using carrier phase measurements", Proc. 50th IEEE FCS, 1151-1158, 1996.
2. G. Petit, C. Thomas, Z. Jiang, P. Uhrich, F. Taris, "Use of GPS Ashtech Z12T receivers for accurate time and frequency comparisons", IEEE Trans. On Ultrason., Ferroelectr., and Freq. Control, Vol. 46, No. 4, pp. 941-949, July 1999.
3. L. Nelson, J. Levine, "Comparison of atomic frequency standards at NIST and PTB using carrier-phase GPS", Proc. 31st PTTI meeting, 1999, in press.
4. T. Schildknecht, R. Dach, T. Springer, G. Dudle, L. Prost, "Recent results with the transatlantic GeTT experiment", Proc. 31st PTTI meeting, 1999, in press.
5. Astronomical Institute University of Bern, "Bernese GPS software version 4.0", Rothacher M. and L. Mervart Ed., 1996
6. P. Wolf, G. Petit, "Use of IGS ionosphere products in TAI", Proc. 31st PTTI meeting, 1999, in press.

CONTINUOUS FILTERING IN GPS CARRIER PHASE TIME TRANSFER ESTIMATION AND PRELIMINARY TIME SCALE RESULTS

KEN SENIOR AND DEMETRIOS MATSAKIS

Time Service Department
U.S. Naval Observatory

Time transfer solutions produced routinely by many IGS Analysis Centers show half-nanosecond level discontinuities between daily solutions. We present evidence that these discontinuities are related to the loss of phase-ambiguity information, and discuss operational methods of handling the problem. Using one method in particular, improved time-transfer results from a variety of baselines are presented, and where possible are compared to Two-Way Satellite Time Transfer data. Using these time-transfer data, preliminary timescale results are also presented.

Ken Senior
U.S. Naval Observatory
3450 Massachusetts Avenue NW
Washington, D.C. 20392
United States
(202) 762-1456
fax: (202) 762-1511
ksenior@usno.navy.mil

COMPARISONS OF THE TWO-WAY, GPS CARRIER-PHASE AND GPS COMMON-VIEW TIME TRANSFER METHODS

Setnam L Shemar, Jon D. Clarke, John A Davis, Andrew J. Lowe and Paul J. Chapman

National Physical Laboratory,
Centre for Electromagnetic and Time Metrology, Queens Road, Teddington, Middlesex, TW11 0LW, UK

ABSTRACT

In the last 15 years international time transfers have traditionally been carried out between national timing laboratories using the GPS Common-View method. More recently, there has been significant activity in the Two-Way Satellite Time and Frequency Transfer (TWSTFT) and GPS Carrier-Phase techniques to achieve an even higher precision of time transfer. In this paper, data obtained using each method are compared over one European and one Transatlantic link. These are the links between NPL and Physikalisch-Technische Bundesanstalt (PTB) and NPL and the US Naval Observatory (USNO). Previous comparisons carried out between Two-Way and GPS Carrier-Phase have sometimes suggested a significant seasonal variation in the differences between their data. We investigate if such an effect is present in our results. We also present statistical comparisons of the data by use of variance and covariance techniques. Finally, we describe an algorithm for combining data from different time transfer links to obtain an optimised composite time transfer.

1. INTRODUCTION

In recent years there has been significant activity in developing higher precision methods of time transfer than the traditional method of GPS Common-View. This has been motivated by caesium fountain atomic frequency standards which are becoming operational in national timing laboratories around the world. The Two-Way time transfer method has now been demonstrated to show a higher precision of time transfer than GPS Common-View. Consequently, this method has already been introduced into TAI for a small number of international links since 1999. Studies of the GPS Carrier-Phase technique are presently also underway. Previous comparisons carried out between Two-Way and GPS Carrier-Phase have sometimes suggested a significant seasonal variation in the differences between their data [1]. We investigate if such an effect is present in our results by studying the NPL-USNO and NPL-PTB links. We also carry out statistical comparisons of the data. Finally, we describe an algorithm for combining data from different time transfer links to obtain an optimised composite time transfer.

2. THE DATA

Time transfer data for each of the three methods have been studied for the NPL-PTB and NPL-USNO links in this paper. Table 1 shows the time spans of data that have been used. The Two-Way (TW) data have been obtained from regular international time transfers which take place three days per week (Monday, Wednesday and Friday) at times between 1422 UTC and 1439 UTC. Each measurement corresponds to data obtained over two minutes from which a representative measurement is obtained at the mid-point of the data-set. Satre modems manufactured by TimeTech are used by NPL and PTB to take the measurements whilst a Mitrex modem is used by USNO.

The GPS Carrier-Phase (CP) data have been obtained using Geodetic Time Transfer (GeTT) terminals owned by the Swiss Federal Office of Metrology (OFMET) and located at each of PTB (named PTBA) and USNO in Washington (named USNB). NPL has used an Ashtech Z-12T (named NPLB). The data have been processed and supplied by the Astronomical Institute of the University of Berne (AIUB). The reference for GeTT PTBA at PTB is the H-maser H2 which is also the reference for TW. Corrections to UTC(PTB) are available at 30 minute intervals and have been obtained via the AIUB. We have used these to correct the PTBA data to UTC(PTB). The TW data have also been corrected to UTC(PTB) within the convention of the TW data format and processing. The reference for USNB at USNO is the H-maser MC2 which is also the reference for TW and GPS Common-View enabling direct comparisons between the three methods. The CP data used in this study correspond to a daily mean taken for all epochs from a period of +/-0.2 days around 1400 UTC of each day. This corresponds to the approximate time of the TW measurements.

The GPS Common View (CV) data have been obtained using an Allen Osborne Associates (AOA) TTR5A receiver at NPL, an AOA TTR6 at USNO and a NBS type receiver made by Rockwell Collins at PTB. The data used in this paper correspond to daily mean values taken from data for each available epoch.

NPL's TW, CP and CV receivers are all referenced to the same H-maser enabling direct comparisons of data. However, the TW and CP receivers are located in a different laboratory to the CV receivers. A distribution amplifier change was carried out for the TW and CP

receivers on MJD 51344 (15th June 1999). This change has necessitated an offset of +1.4 ns to be added to the CP data. The same change has been taken account of in the convention of the TW data format and processing.

Each of the data-sets have been calibrated using Circular-T. The NPL-USNO data have been calibrated using the mean of data from MJD 51434 to 51439 and the corresponding period of Circular-T data. The same has been carried out for the NPL-PTB data except that the data used cover a period from MJD 51433 to 51435.

Table 1: Time spans of data studied for each link.

Method	NPL-PTB (MJD)	NPL-USNO (MJD)
TW	51252-51585	51252-51585
CV	51148-51570	51178-51570
CP	51150-51560	51150-51584

3. DIRECT COMPARISONS

Data for the NPL-PTB and NPL-USNO links for each of the three methods are shown in Figures 1 and 3. To aid visual comparison of the data-sets offsets of +10 and -10 ns have been added to the CP and TW data respectively. Figures 2 and 4 show the differences between measurements for each method. Offsets of 10, -10 and -20 ns have been added to the differences CV-TW, CV-CP and TW-CP respectively to aid visual comparison. The outdoor temperature at NPL at the time of measurements has varied seasonally by 30°C whilst the indoor temperature of the laboratories containing the receivers has been kept constant to within 3°C.

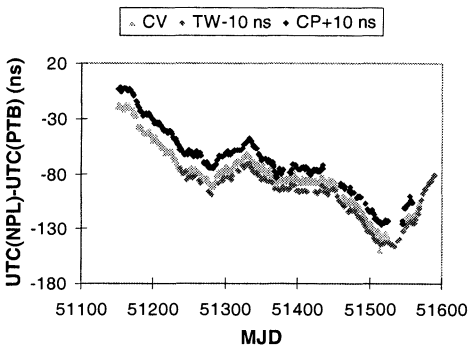


Figure 1: Data for the CV, CP and TW NPL-PTB links.

Figure 2 shows that there is good agreement between the TW and CV links, suggesting that the peak-to-peak noise in each is < 4 ns over one year. However, the deviations of CP from both TW and CV are greater with similar long-term peak-to-peak deviations of ~8 ns. This indicates that the noise in the CP link is greater than the other links and may be as high as 8 ns peak-to-peak.

There is much less agreement between the TW and CV data for the NPL-USNO links. In this case, the relative deviations are 10 ns. This is also the case for the deviations between CV and CP suggesting that there is noise of up to 10 ns peak-to-peak originating in the CV link over one year. The most likely explanation is that this results from a degradation of the CV link caused by ionosphere differences over the Transatlantic baseline. The best agreement is found between TW and CP where the relative peak-to-peak deviations are of magnitude 6 ns.

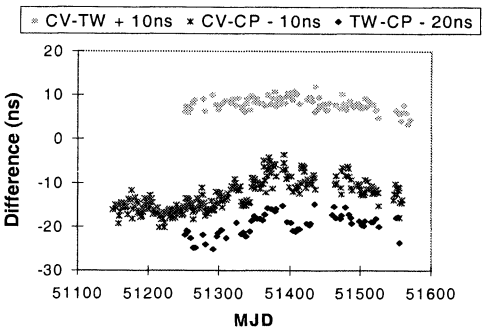


Figure 2: Differences in CV, CP and TW data for the NPL-PTB links.

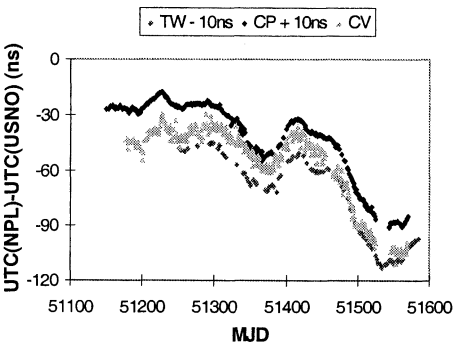


Figure 3: Data for the CV, CP and TW NPL-USNO links.

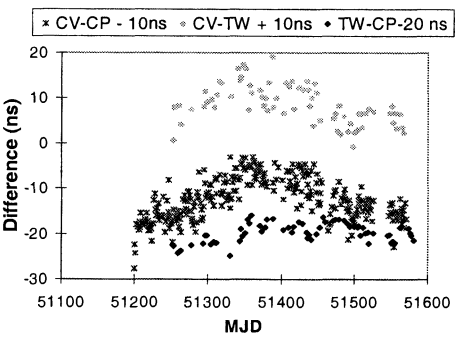


Figure 4: Differences in CV, CP and TW data for the NPL-USNO links.

It has previously been reported that it is possible to fit a sine wave with amplitude 2.7 ns and period of about one year to the differences between TW and CP for the PTB-USNO link [1]. We find there is no strong evidence for such an effect in the data we present.

However, there is some clear structure with peak-to-peak variations of 5 ns occurring over timescales of months. We find that although some of these short-term deviations are found to correlate with NPL outdoor temperature, there is less of a correlation on longer timescales such as a year. The major part of the TW temperature coefficients are found to arise in outdoor 50 Ω PTFE coaxial cables. Previous studies of these cables [2] suggest that the uplink and downlink temperature coefficients are 0.10 and 0.005 ns/°C respectively. These could explain variations of 3 ns over one year. There are also likely to be significant temperature coefficients for the CP hardware [3] which need to be quantified. This presently uses outdoor cables of RG-8/U-type. It is planned to replace the present outdoor cables for both TW and CP with those that are phase stabilised. This may lead to lower variations between the two links in the future.

4. STATISTICAL COMPARISONS

It is not always possible to directly compute Allan variances and covariances of the data presented in Section 3. This is for two main reasons. Firstly, the data may not be evenly sampled as in the case of the TW data where the sampling intervals are either two or three days. Secondly, there may be a gap in the data due to hardware being off-line for certain periods. In order to overcome these problems we apply a previously described method [4] involving linear interpolation between available data in order to estimate measurements that are not available. Previous studies [4] have found that in the case of TW data the Allan variance is underestimated for small values of τ using this method.

The Allan variance and covariance are given by the equation

$$\sigma_y^2(\tau)=\frac{1}{2\tau^2(N-2n)}\sum_{i=1}^{N-2n}(x_{i+2n}-2x_{i+n}+x_i)(y_{i+2n}-2y_{i+n}+y_i)$$

where N is the total number of measurements, x_i is the ith measurement separated by a time interval τ_o and $\tau = n\tau_o$ is the averaging time. In the case of the variance x and y represent the same data-set. This enables us to investigate the distribution of noise amongst a group of data-sets. For the covariance x and y represent two separate data-sets and this enables us to measure the amount of correlated noise between them. For two independent time transfer methods, the main behaviour that we would expect to correlate would be the clock behaviour. Other sources of correlation may be environmental conditions such as temperature.

We have used NPL-PTB and NPL-USNO data presented in Section 3 from MJDs 51471 to 51525 to obtain variances and covariances for each method. In order to carry out consistent comparisons we have sampled each data-set at the same days as 19 available TW measurements. This corresponds to a mean interval of about 3 days between measurements. Linear interpolation has then been applied to obtain estimated measurements for the remaining MJDs. The resulting $\sigma_y(\tau)$ values for the variances and covariances are shown in Figure 5 and 6. Values have not been plotted for averaging times $\tau < 3$ days as these are less than the mean data interval and expected to be unreliable. The CP/TW and CV/TW covariance values for $\tau = 3$ days are found to be negative and have not been plotted.

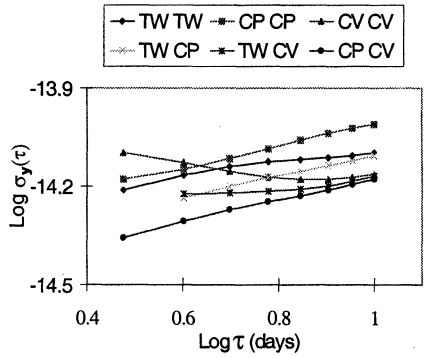


Figure 5: Variances and covariances for each of the three methods using data for NPL-PTB.

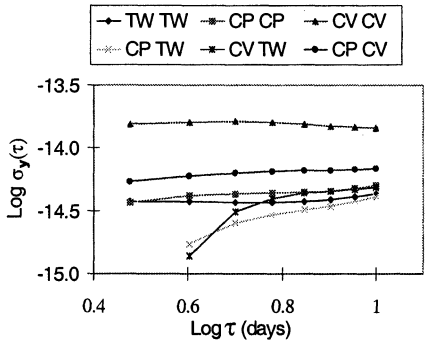


Figure 6: Variances and covariances for each of the three methods using data for NPL-USNO.

The NPL-PTB data in Figure 5 show evidence that for $\tau < 5$ days TW and CP perform better than CV, but that for greater averaging times CV performs best. For $\tau > 5$ days the variance and covariances involving CV converge to a lower level than the variances of TW and CP indicating that CV is revealing the underlying noise of the clocks at such averaging times. The CP/CV covariance gives the lowest values. This indicates the true underlying noise type of the clocks. The slope suggests that over averaging times of a few days the dominant noise type is random walk frequency

modulation. It is significant that for $\tau > 5$ days both CV and TW are significantly better than CP.

There are greater differences in the variance and covariance results in the case of the NPL-USNO links shown in Figure 6. It is clear that the greatest level of noise is shown by the CV data. TW and CP are found to show similar levels of noise that are substantially lower. The TW/CV and TW/CP covariance data show low levels of correlated noise. The level is particularly low for TW/CP with the slope indicating that as for NPL-PTB, random walk frequency modulation is the dominant noise type. The CV/CP covariance data suggest that there is additional correlated noise in the GPS-based techniques. It is interesting that the variances for TW and CP and the covariances for TW/CV and TW/CP converge for $\tau \sim 10$ days, indicating that this is the true level of clock noise at such averaging times.

5. COMBINING TIME TRANSFER LINKS

To make optimum use of all of the available time transfer links between any two locations we have combined them into a composite time / frequency transfer. We wish to minimise the composite link noise $\sigma_x^2(L)$ where

$$\sigma_x^2(L) = \sigma_x^2(C - N)$$

for a given averaging time (τ). C is the clock differences being measured, in the example here (UTC(USNO)-UTC(NPL)) and N is a linear combination of the time transfer measurements (containing noise originating from both the time transfer links and the clocks under comparison) where:

$$N = \sum_{i=1}^{n} w_i T_i$$

T_i are the individual time transfer data-sets and **w** is a vector representing the link weights such that

$$\sum_{i=1}^{n} w_i = 1$$

where n is the total number of time transfer links between NPL and USNO.

Let **R** be the covariance matrix formed from the individual time transfer data sets T_i. The leading diagonal terms of the matrix give the sum of the variance of the individual time transfer links $\sigma_x^2(T_i - C)$ plus the underlying clock noise $\sigma_x^2(C)$. The off-diagonal covariance terms give the noise of the clock differences $\sigma_x^2(C)$ plus any correlated time transfer link noise.

The total noise of the composite time transfer $\sigma_x(N)$ (Links plus clocks) is given by [5]

$$\sigma_x^2(N) = \mathbf{w}^T \mathbf{R}\, \mathbf{w}$$

We assume that there are correlations between the individual time transfer links but not between the clocks and the time transfer links. Minimising the total noise of the composite time transfer $\sigma_x^2(N)$ will then also minimise the composite link noise $\sigma_x^2(L)$. The following equation was chosen to determine the weights **w** [5]

$$\mathbf{w'} = (\mathbf{R})^{-1}\, \mathbf{1}$$

where **w'** is a non-normalised weight. The i th normalised weight w_i is given by $w_i = w_i' / \Sigma w_i'$. The above equation was chosen because the resulting weighting procedure has the following properties:

a) In the absence of off-diagonal covariance terms in the matrix **R** (no correlation between data sets T_i) the weights are proportional to $1/\sigma_x^2$, where σ_x^2 terms form the leading diagonal (variance terms) of the covariance matrix **R**.

b) Correlated links are de-weighted in comparison to uncorrelated links with the same magnitude variance terms.

c) Changing the level of clock noise (adding the same value to each element of the covariance matrix) does not change the resulting weights.

A composite time transfer link has been formed from TW, CP and CV measurements of UTC(USNO)-UTC(NPL). Six time transfer links have been used, including the TW, CP and CV links discussed earlier. In addition three GPS common-view links have been included, using an additional R100-40 GPS receiver at USNO and an AOA TTR5A receiver at NPL. The first two rows / columns of the covariance matrix represent the TW and CP links in that order. The remaining four rows / columns represent common-view links. The GPS common view receivers used are as follows for links 3 to 6 respectively: USNO(R100) and NPL(AA276), USNO(AOA) and NPL(AA276), USNO(R100) and NPL(AA235), and USNO(AOA) and NPL(AA235). The data used cover a period from MJD 51471 to MJD 51525. Unlike in Section 4, all available daily measurements are used for each link. Linear interpolation is used where there are intervals in data of > 1 day. Values of σ_x^2 are calculated with an averaging time of 3 days and shown in Figure 7.

$$\begin{pmatrix} 0.118 & 0.061 & 0.230 & 0.070 & 0.203 & 0.082 \\ 0.061 & 0.123 & 0.255 & 0.275 & 0.231 & 0.280 \\ 0.230 & 0.255 & 3.28 & 1.70 & 2.13 & 1.68 \\ 0.070 & 0.275 & 1.70 & 3.46 & 1.47 & 3.25 \\ 0.203 & 0.231 & 2.13 & 1.47 & 3.14 & 1.74 \\ 0.082 & 0.280 & 1.68 & 3.25 & 1.74 & 3.69 \end{pmatrix}$$

Figure 7: The UTC(USNO)-UTC(NPL) link σ_x^2 (ns^2) covariance matrix **R**.

The following weights have been obtained $[0.537 \quad 0.523 \quad -0.041 \quad 0.006 \quad -0.018 \quad -0.007]$. As expected the majority of the weighting is being given to the TW and CP links.

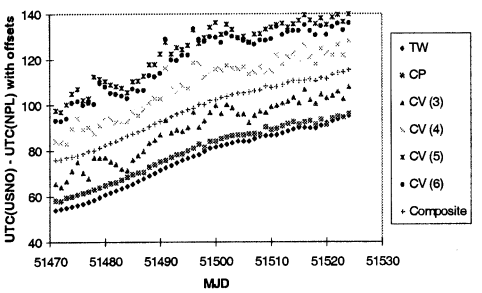


Figure 8: The composite time transfer link and its components.

The composite time transfer link and its components are shown in Figure 8. A value of σ_x^2 ($\tau = 3$ days) of 0.084 ns^2 was obtained. This was significantly lower than the scatter on any of the component links.

6. CONCLUSIONS

The results presented in this paper lead us to conclude the following.

- The NPL-PTB TW and CV data show good agreement suggesting that the noise in each link is < 4 ns peak-to-peak over one year. CV is also found to be the best link for averaging times > 5 days. There is evidence that the CP link shows the greatest noise of up to 8 ns peak-to-peak over one year.
- The NPL-USNO TW and CP data show agreement to within 6 ns over one year. However, it is apparent that the noise in the CV link is greater and may be up to 10 ns peak-to-peak. The most likely explanation is that this is due to degradation of the CV link caused by ionosphere differences over the Transatlantic baseline.
- There is some clear structure with peak-to-peak variations of 5 ns occurring over timescales of months present in the differences between TW and CP in both the NPL-PTB and NPL-USNO links.
- An algorithm has been devised and applied to combine data from different time transfer links and methods to obtain an optimised composite time transfer.

7. ACKNOWLEDGEMENTS

We would like to thank the Astronomical Institute of the University of Berne (AIUB) for processing of the GPS Carrier-Phase data used in this paper, the Swiss Federal Office of Metrology (OFMET) for use of the data collected by their GeTT terminals at PTB and USNO and PTB and USNO for use of their TWSTFT and GPS Common-View data. We also thank Lara Schmidt for supplying various USNO data-sets, Rolf Dach at the AIUB and J. Becker at PTB for supplying corrections between the GeTT PTBA source clock to UTC(PTB) and John Laverty, Rolf Dach, Peter Hetzel and Lara Schmidt for helpful comments on the manuscript.

8. REFERENCES

[1] G. Dudle, F. Overney, th. Schildknecht, T. Springer, L. Prost, 'Transatlantic Time and Frequency Transfer by GPS Carrier Phase', Proceedings of the 1999 Joint Meeting of the European Frequency and Time Forum (EFTF) and International Frequency Control Symposium (FCS), Besancon, France, pp243-246, 13-16 April 1999.

[2] S. L. Shemar and J. A. Davis, "A study of delay instabilities within a two-way satellite time and frequency transfer earth station", Proceedings of the 1999 Joint Meeting of the European Frequency and Time Forum (EFTF) and International Frequency Control Symposium (FCS), Besancon, France, pp243-246, 13-16 April 1999.

[3] J. D. Clarke, J. A. Davis and A. J. Lowe, "Characterisation of NPL's Geodetic GPS Time Transfer Receivers", Proceedings of the 1999 Joint Meeting of the European Frequency and Time Forum (EFTF) and International Frequency Control Symposium (FCS), Besancon, France, pp225-229, 13-16 April 1999.

[4] P. Tavella and M. Leonardi 'Noise Characterisation of Irregularly Spaced Data', Proceedings of the 12th European Frequency and Time Forum (EFTF), 10-12 March 1998, pp 209-214.

[5] J. A. Davis., Rueff B., "A study of correlated noise within a clock ensemble", EFTF Torino, 2000.

SOME TESTS OF GPS MULTI-CHANNEL TIME TRANSFER UNITS TTS-2 BASED ON A MOTOROLA RECEIVER AND USING CCTF STANDARDS

J. NAWROCKI
Astrogeodynamical Observatory, Polish Academy of Sciences, 62-035 Kornik, Poland

M. MIRANIAN, E. POWERS
United States Naval Observatory, Washington D.C., USA

J. AZOUBIB, W. LEWANDOWSKI
Bureau International des Poids et Mesures, Pavillon de Breteuil, F-92312 Sèvres Cedex, France

Abstract

An interesting new method of time transfer utilizes the inexpensive eight-channel Motorola Oncore VP receiver, together with a time interval counter and a microcomputer. Several of such measurement system units called TTS-2 (Time Transfer system –2) has been built and tested over two years at the Time Section of the BIPM and the Astrogeodynamical Observatory in Poland. More recently two TTS-2 were built for the needs of the US Naval Observatory. The novelty of the method lies in its treatment of the raw pseudo-ranges provided by the Motorola module, in accordance with Consultative Committee for Time and Frequency (CCTF) standards. The structure of the new time transfer unit is described in brief, and the results of some multi-channel time transfer tests as one-site comparison, tilr transfer over 1200 km 7000 km are presented. The multi-channel method provides an improved stability of time transfer compared with the traditional single channel method.

Of particular interest is a one-site comparison. Four TTU-2's were tested: two at the BIPM and another two at the USNO. At each site they were connected to the same antenna and relative to a common time source. This one-site test was carried out to characterize the receiver hardware noise. All noise due to travel of the signal through space, possible imperfection of receiver software, and antenna multipath reflections cancels out in this configuration. Only the differential noise of the two receivers is observed. A single frequency standard supplied at each site a 1 pps signal and a 10 MHz reference signal to both TTU-2s. The results of the measurements are similar at at the BIPM and at the USNO. The rms of a single pass is 0.6 ns, with all the data lying within a band about 2 ns wide. These results, obtained in an air-conditioned laboratory, are satisfactory. Under these good conditions the internal noise of the receivers is very low.

The tests described in this paper prove that the TTS-2 units provide results similar, or better in the case of the multi-channel mode, to those obtained using classic single-channel time receivers.

THE TRANSPORTABLE PHARAO FOUNTAIN CLOCK

M. Abgrall, P. Lemonde, Ph. Laurent, G. Santarelli, P. Petit and A. Clairon
BNM-LPTF, 61 Av. de l'Observatoire, 75014 Paris, France

We present a first evaluation of a transportable Cs fountain. Originally, the device was designed as a prototype of a space clock and tested in aircraft parabolic flights [1]. It has been modified to a fountain geometry while preserving its transportability.

We focus on the frequency shifts due to cold collisions and to the first order Doppler effect. We have measured the collisional shift for atoms captured in an optical molasses. With 4 10^5 detected atoms, it amounts to 3.3 10^{-15}. The present relative uncertainty is better than 10 %.

According to numerical simulations, the Doppler effect is expected to be much lower than our present 10^{-15} resolution. This results from the intrinsic symmetry of the interrogation in a fountain and doesn't account for possible defects of the cavity. It is important to confirm the reliability of the model not only for the accuracy evaluation, but also for the design of a space clock where the symmetry is broken. To perform local measurements of the field phase, one possibility is to pulse the microwave. We describe a phase stable microwave switch that uses a Mach-Zehnder interferometer. It should give access to the map of the phase inside the cavity. When accounting for collisions, the quadratic Zeemann effet, the black body shift and the gravitational red shift, the agreement with the FO1 fountain is better than 2 10^{-15}. The measured relative frequency stability is 1.4 $10^{-13}\tau^{-1/2}$.

Finally, we have demonstrated the transportability of the clock. In a collaboration with the Max Planck Institut für Quantenoptik in Garching (Germany), it has been used as a reference for the frequency measurement of the $1S-2S$ transition of hydrogen [2].

References

[1] P. Laurent, P. Lemonde, E. Simon, G. Santarelli, A. Clairon, N. Dimarcq, P.Petit, C. Audoin, and C. Salomon, Euro. Phys. J. D, **3**, 201 (1998).

[2] T. Udem, A. Huber, B. Gross, J. Reichert, M. Prevedelli, M. Weitz, and T. Hänsch, Phys. Rev. Lett. **79**, 2646 (1997).

Corresponding author: M. Abgrall, Fax: +33 01 43 25 55 42, E-mail:michel.abgrall@obspm.fr

INVESTIGATION OF CONTRIBUTIONS TO THE STABILITY AND UNCERTAINTY OF THE PTB ATOMIC CAESIUM FOUNTAIN CSF1

S. Weyers, U. Hübner, B. Fischer, R. Schröder, Chr. Tamm, A. Bauch

Physikalisch-Technische Bundesanstalt, Lab. 4.32, Bundesallee 100, D-38116 Braunschweig

Corresponding author: S. Weyers, Physikalisch-Technische Bundesanstalt, Lab. 4.32, Bundesallee 100, D-38116 Braunschweig, Tel. ++49 (0)531 592 4415, Fax ++49 (0)531 592 4479, email: Stefan.Weyers@PTB.de

ABSTRACT

We report results of a first uncertainty evaluation of an atomic fountain frequency standard at PTB. We also give a short discussion of the influence of different noise sources on the frequency instability.

1. INTRODUCTION

Due to the success of the first atomic caesium fountain frequency standard at the Laboratoire Primaire du Temps et des Fréquences (LPTF) [1], there are worldwide efforts to build similar devices.

In the fountain CSF1 of PTB, laser cooled caesium atoms are collected in a magneto-optical trap (MOT) and launched after a 100 ms molasses phase into a magnetically shielded titanium vacuum chamber. The TE011 microwave cavity used for Ramsey excitation has a loaded quality factor of Q≈2000 and was designed in order to combine a low transversal phase gradient and a low sensitivity of the resonance frequency to temperature-induced detunings [2]. After the microwave interaction the atomic state is analyzed in a region above the MOT. Ramsey fringes with a full width at half maximum (FWHM) of $W = 0.86$ Hz are obtained by launching the atoms to a maximum height of 41 cm above the microwave cavity center (83 cm above the cooling region). The laser system for cooling, repumping and detection consists of two extended-cavity diode lasers and one injection-locked slave laser. A more detailed description of the complete setup is given in [3,4].

In Fig. 1 relative frequency differences between the CSF1 and two hydrogen masers are shown for a period of almost 6 weeks. In the course of these measurements the power of the microwave interrogation field was increased by up to 8 dB relative to the optimum value for π/2-pulse excitation. The relative sublevel population difference

$$A = \frac{N(F=4,m_F=+1) - N(F=4,m_F=-1)}{N(F=4,m_F=0)}, \quad (1)$$

with $N(F, m_F)$ denoting the number of launched atoms in the indicated sublevels, was changed between less than 5% and 30% (see below). Also the total number of launched atoms was varied by up to a factor of 20. The resulting collisional frequency shift was corrected for as described below. The linear least squares fits to the data points in Fig. 1 indicate frequency difference drifts of $(1.6 \pm 0.4)\cdot10^{-16}$/day and $(0.6 \pm 0.4)\cdot10^{-16}$/day, respectively, which we attribute to frequency drifts of

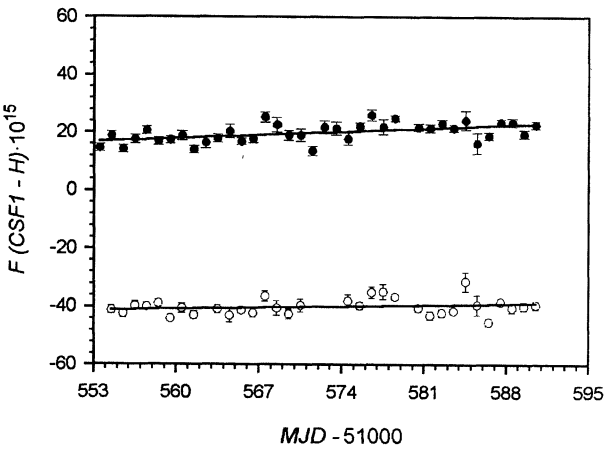


Fig. 1: Relative frequency differences between CSF1 and the PTB hydrogen masers H1: o (VCH-1005) and H2: • (KVARZ CH1-75). The averaging time for most of the data points was about 22 hours. The error bars reflect the statistical uncertainty due to the noise of the fountain signal and of the hydrogen masers. MJD is the Modified Julian Day (MJD 51553 = 10 January 2000). The full lines are linear fits to the data points.

the employed hydrogen masers. The observed drifts are consistent with the drift between the two hydrogen masers which was independently determined as $(1.7 \pm 0.5)\cdot 10^{-16}$/day.

2. MAIN UNCERTAINTY CONTRIBUTIONS

In the following paragraphs we discuss the major uncertainty contributions to the fountain frequency as they have been evaluated as yet.

C-field. Typically, the dominant frequency correction in caesium frequency standards is the frequency shift due to the static magnetic field (C-field) which is applied in order to lift the degeneracy of the magnetic substates. In the PTB fountain an axial mean static magnetic field $\langle B \rangle = 0.1$ μT is produced by a long coil that is supported by an aluminium cylinder enclosing the titanium vacuum chamber. The resulting relative frequency correction is $-46.5\cdot 10^{-15}$ [3]. The temporal stability of the C-field yields a relative frequency uncertainty contribution of less than $4\cdot 10^{-17}$ if the C-field strength is measured once a week.

The inhomogeneity of the magnetic flux density for maximum launch height was measured to be 0.14 nT (rms), only 0.14 % of the mean value of 0.1 μT used in normal operation. As a result the uncertainty contribution due to the inhomogeneity of the C-field is negligible ($\approx 10^{-19}$) [3,4].

Collisional shift. In our fountain the atoms are launched in the state $F = 4$ at a temperature of 2 μK. In this case, the scattering process is dominated by s-wave scattering and the collisional shift depends linearly on the mean density $\bar{n}$ of the atomic cloud between the two microwave interactions [5,6].

One possibility to correct the frequency for the collisional shift is to determine the shift as a function of the absolute value of $\bar{n}$ [6]. In order to avoid the difficulties of an absolute density measurement, we choose an alternative way. We assume that $\bar{n}$ is proportional to the number of detected atoms N_{at}, if other experimental conditions such as detection efficiency, launch height and the atomic temperature are left unchanged. The relative variation and the absolute value of N_{at} can be determined with uncertainties of less than 0.2% and 15%, respectively, from the variation of the relative strength and from the signal-to-noise ratio of the detector signal (see below). The results of a corresponding measurement are shown in Fig. 2. N_{at} was changed by varying the loading time of the MOT. Each data point in Fig. 2 is the result of a measurement lasting on average about 12 hours. The indicated error bars reflect the statistical uncertainty due to the frequency noise of the CSF1 and of the hydrogen maser and do not take into account the long-term hydrogen maser frequency drift. To avoid systematic errors due to this drift, the data was collected using a random sequence of N_{at} values in the range between $6\cdot 10^4$ and $1.3\cdot 10^6$. The full line in Fig. 2 indicates a linear least-squares fit to the data with a resulting slope of $m = -(3.13 \pm 0.23)\cdot 10^{-20}$.

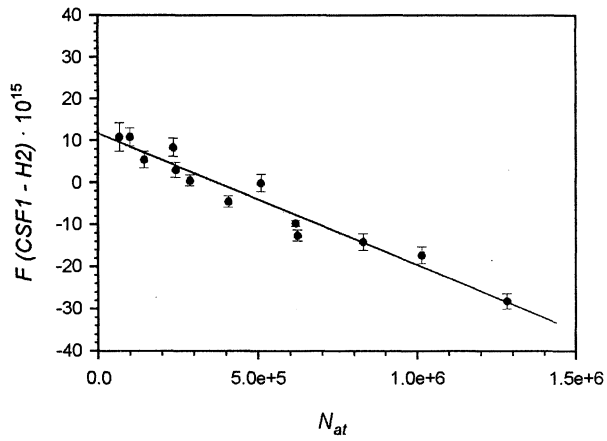


Fig. 2: Measured relative frequency differences between the fountain and hydrogen maser H2 as a function of the total number of detected atoms N_{at}. The full line is a linear fit to the data points.

Fig. 2 indicates a linear dependence of the observed frequency differences on N_{at}. This dependence can be attributed to the collisional shift of the fountain frequency and confirms our expectation that N_{at} is proportional to $\bar{n}$. The constancy of the slope of the least-squares fit of Fig. 1 was checked by another measurement performed two months later. Here the fountain frequency was alternately measured for N_{at} values of about $0.2\cdot 10^6$ and $1.2\cdot 10^6$. The resulting mean slope of $m' = -(2.92 \pm 0.14)\cdot 10^{-20}$ agrees with the previously determined m value within the respective 1σ uncertainties. This permits the conclusion that a possible change of the proportionality factor between N_{at} and $\bar{n}$ can certainly be controlled to better than 10%.

A compromise between low instability and low uncertainty due to the collisional frequency shift was realized by operating the fountain at $N_{at} \approx 2\cdot 10^5$. In this case, taking the average of m and m' we calculate a relative frequency correction due to collisions of $6.1\cdot 10^{-15}$, which corresponds to an average density $\bar{n} \approx 6.1\cdot 10^6/\mathrm{cm}^3$ using the result of [6]. This density value is in reasonable agreement with the value of $\bar{n}$ that is estimated from the signal size without precise knowledge of the atomic cloud size.

Assuming a maximum change of the proportionality factor between N_{at} and $\bar{n}$ of 10% and taking into

account the statistical uncertainty of the measured slopes m and m', we obtain by taking the square root of the quadratic sum of both contributions a relative frequency uncertainty contribution of less than $1 \cdot 10^{-15}$. This contribution can be further reduced by more extended investigations of the collisional frequency shift under various experimental conditions.

Black-body shift. The fountain is placed next to the other PTB primary clocks in a hall whose temperature is stabilized to better than ± 0.2 K. Three Pt-100 thermoresistors are mounted along the C-field coil support cylinder.

The absolute uncertainty of the temperature measurement using the Pt-100 thermoresistors is 0.3 K. The temperature gradient along the cylinder indicated by the thermoresistors is between 0.1 K and 0.2 K. Taking into account potential radial temperature gradients and the unknown emissivity of the titanium chamber, an additional temperature uncertainty of 1 K is assumed for the interaction region. During two months of frequency measurements, each of the indicated temperatures remained stable to within 0.3 K with an unchanged temperature gradient. The measured mean temperature is 297.0 K, resulting in a relative frequency correction of $16.7 \cdot 10^{-15}$. Since the absolute uncertainty of the temperature is 2 K, the resulting relative frequency uncertainty contribution is $0.5 \cdot 10^{-15}$.

Majorana transitions, Rabi-pulling, Ramsey-pulling, microwave leakage. In order to avoid Majorana transitions, an arrangement of axial correction coils prevents zero transitions and abrupt changes of the magnetic field in the region between MOT and microwave cavity. The population difference A (see Eq. 1) can be reduced to less than 5% by careful adjustment of the polarization state of the vertical molasses beams, which minimizes frequency shifts caused by Majorana transitions and by Rabi- and Ramsey-pulling. A calculation of the effect of Rabi-pulling for $A = 20\%$ predicts a shift of less than $1 \cdot 10^{-17}$ at optimum microwave power. In fact, first observations indicate that even for $A = 30\%$, a possible total shift is smaller than $2 \cdot 10^{-15}$ if the microwave power is increased by a factor of 9 with respect to the optimum power. The signal of the $\Delta m_F = \pm 1$ transitions is 1% of the clock transition signal at optimum microwave power. We conclude that the relative frequency shift effect of Rabi- and Ramsey-pulling, Majorana transitions, and microwave leakage is less than $0.5 \cdot 10^{-15}$.

Residual first-order Doppler effect. A general advantage of atomic fountain microwave frequency standards is that the atoms cross the same microwave cavity twice so that a symmetric excitation of spatially separated cavity sections is not required. A nonvanishing transverse phase variation of the cavity field can however give rise to a residual first-order Doppler effect frequency shift.

The microwave field in the cavity can be described as a superposition of various field modes which are excited with different amplitudes and phases. In our cavity design, a single microwave feed excites two symmetrically arranged coupling slits [2]. In this case the predominant first-order Doppler shift contribution is due to the TE111 mode, which is excited if the coupling by the two slits is asymmetric.

For a coupling asymmetry of 10%, which is much larger than the asymmetry expected in our case, one calculates a residual relative first-order Doppler effect frequency shift of less than $0.5 \cdot 10^{-15}$. In this worst-case estimate it is assumed that all atoms ascend through the cavity aperture next to one coupling slit and descend next to the other slit. This is not the case if well-aligned laser beams are centered to the cavity axis and the beam intensities are balanced. The actual residual Doppler shift is also reduced by the additional averaging over cavity phase gradients that results from the expansion of the atomic cloud during the ballistic flight.

Second-order Doppler effect. For launching the atoms to 41 cm above the microwave cavity they have to be launched with an initial velocity of 4.05 m/s. Thus the mean quadratic velocity of the atoms above the microwave cavity is $\langle v^2 \rangle = 2.67\ \mathrm{m^2/s^2}$. Therefore, one calculates a corresponding relative frequency shift contribution of $-1.5 \cdot 10^{-17}$. This shift and the associated uncertainty contribution of much less than 10^{-17} are neglected here.

Electronics and microwave spectral impurities. The microwave interrogation field of the fountain is synthesized from a low-noise 5-MHz quartz oscillator (Oscilloquartz OCXO 8611/02). The oscillator signal is multiplied to 9.2 GHz and mixed with the signal of a DRO (dielectric resonator oscillator). The beat signal is mixed with the output of a frequency synthesizer locked to the employed quartz oscillator. The output of this mixer is used to phase-lock the DRO, whose attenuated output signal excites the fountain microwave cavity.

Calculations show that in our case sidebands of the microwave interrogation field with frequency separations in the range of 50 Hz are most important to consider as a cause for a possible frequency shift. In the present system, such sidebands are suppressed to -60 dBc. For a single 50 Hz sideband at -60 dBc and optimum microwave power, one calculates a maximum shift of $4.4 \cdot 10^{-17}$. A corresponding relative frequency uncertainty contribution of $0.1 \cdot 10^{-15}$ is assumed.

At each launch the output frequency of the synthesizer is switched between $\nu_0 + W/2$ and $\nu_0 - W/2$. Here

$\nu_0 = 9192631770$ Hz is the clock transition frequency. The evaluation of the digitized photodetector time-of-flight signals indicating the population of the levels $F = 4$ and $(F = 3, m_F = 0)$ yields a time-integrated frequency error signal which is to first order independent of fluctuations of the total atom number [1]. The error signal passes through an analogue circuit containing a proportional and an integral path, and is finally fed as the control voltage to the quartz oscillator. Currently no relative frequency uncertainty contribution larger than $0.1 \cdot 10^{-15}$ is expected due to the electronics.

Background gas collisions. The vacuum pressure in the PTB fountain is in the low 10^{-7} Pa range. In this pressure range, the estimated relative frequency shift due to collisions with background gas atoms and molecules is less than $0.5 \cdot 10^{-15}$ [1].

Other effects. There are several other frequency shifting effects (shift due to cavity pulling, DC Stark shift, Bloch-Siegert shift), which can be estimated to be on the order of less than 10^{-17} in our case so that they are neglected here.

Tab. 1: Current uncertainty budget of CSF1.

Effect	Correction [10^{-15}]	Uncertainty [10^{-15}]
C-field	– 46.5	< 0.1
Collisional shift	6.1	1
Black body shift	16.7	0.5
Majorana transitions, Rabi-pulling, Ramsey-pulling, microwave leakage	–	0.5
First order Doppler effect.	–	0.5
Microwave spectral impurities	–	0.1
Background gas collisions	–	< 0.5
Total 1σ uncertainty		1.4

Gravitational red shift. As the gravitational redshift is not significant for the realization of the *proper* second, it needs not to be included in the total uncertainty budget. However, its knowledge is important for comparisons with distant clocks. The mean height of the atoms above the geoid during their ballistic flight above the microwave cavity center is 80.1 m. The corresponding relative frequency correction with respect to zero height is $-8.7 \cdot 10^{-15}$. The relative uncertainty of this correction is of the order of 10^{-16}.

We conclude that the main uncertainty contributions of the PTB fountain should be of the order of $1 \cdot 10^{-15}$ or less and that the expected current total 1σ uncertainty is $1.4 \cdot 10^{-15}$ (Tab. 1).

3. INSTABILITY

In the present fountain setup we do not preselect atoms in the substate $m_F = 0$ [1]. Hence we have to take into account a substate partition noise contribution which reflects the statistical shot-to-shot fluctuation of the number of atoms in the respective substates (F, m_F) as described by the binomial distribution law [4]. With the employed optical detection system, photon shot noise is negligible, so that we are mainly concerned with the partition noise, the quantum projection noise, electronic noise of the photodetector and the local oscillator noise.

The known dependence of the total noise on the absolute number of detected atoms N_{at} enables us to infer this number from noise measurements [4]. This avoids a determination based on the time-of-flight signal which requires an accurate knowledge of the fluorescence collection efficiency.

Fig. 3 shows the resulting Allan standard deviations $\sigma_y(1s)$ that are observed in frequency measurements performed with different numbers N_{at} of detected atoms. N_{at} was changed by varying the MOT loading time and thus the ratio of interrogation and cycle time. The solid curve in Fig. 3 shows the calculated $\sigma_y(1s)$ given by [4]:

$$\sigma_y(\tau) = \frac{W}{\pi \nu_0} \sqrt{\frac{15}{N_{at}} + \frac{1.71 \cdot 10^6}{N_{at}^2}} \sqrt{\frac{T_c}{\tau}}, \qquad (2)$$

where T_c is the cycle time. Eq. (2) takes into account substate partition and quantum projection noise (term proportional to $1/N_{at}$), and detector noise (term proportional to $1/N_{at}^2$). The noise contribution due to local oscillator noise is omitted. Eq. (2) takes further into account that the total number of detected atoms in the sublevels $(F = 4, m_F = 0)$ and $(F = 3, m_F = 0)$ is $N_{at}/8$.

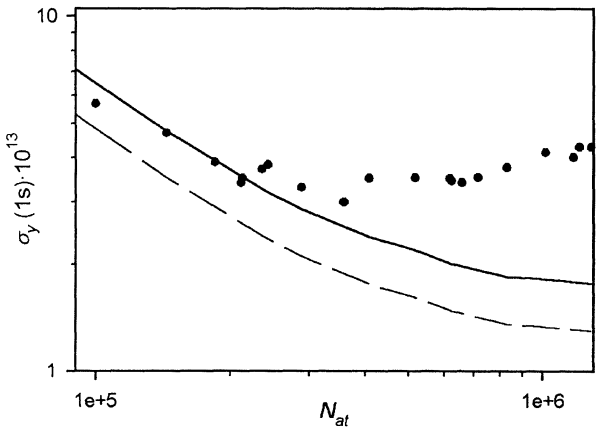


Fig. 3: •: Measured instability of relative frequency difference measurements between the fountain and the hydrogen maser H2 for different total numbers of detected atoms N_{at}.
Full line: Instability calculated according to Eq. (2) (see text).
Broken line: Instability calculated for the same conditions with preselection of the atoms in $m_F = 0$. In this case the scale N_{at} represents the number of detected atoms multiplied by 8.

The comparison of measured and calculated (full line) instabilities in Fig. 3 indicates that the noise sources taken into account in Eq. (2) well describe the observed frequency instability behaviour for $N_{at} < 3\cdot10^5$. For larger N_{at}, the measured instability slightly increases and deviates from the calculated instability presumably due to local oscillator noise [1]. Fig. 3 shows that for $N_{at} \approx 2\cdot10^5$, the measured instability of relative frequency differences is $\sigma_y(\tau) \approx 3.6\cdot10^{-13}\,(\tau/s)^{-1/2}$ which is close to the presently obtainable instability minimum. At this point the noise contribution due to the local oscillator noise is almost negligible. A detected atom number in the range of $N_{at} = 2\cdot10^5$ was also used in the evaluation of the relative uncertainty contribution due to collisions (see above).

For comparison the broken line in Fig. 3 indicates the calculated instability for the case that the substate partition noise contribution in Eq. (2) is zero. This situation is realized if atoms are preselected after launch so that only atoms in the substate ($F = 3$, $m_F = 0$) or ($F = 4$, $m_F = 0$) remain. In this case the scale N_{at} represents the number of detected atoms multiplied by 8.

With preselection, the number of launched atoms can be increased by a factor of 4 relative to the unselected case without increase in collisional shift, if 1/8 of the atoms is in the substates $m_F = 0$ [6]. As demonstrated in Fig. 3, an interrogation oscillator with reduced frequency fluctuations would be required in this case in order to take advantage of the resulting enhanced intrinsic short-term frequency stability.

4. CONCLUSION

We have reported a first evaluation of the uncertainty contributions of an atomic fountain frequency standard developed at PTB. Operating conditions providing an advantageous compromise between low uncertainty and low instability were identified. Even without preselection of atoms in a substate $m_F = 0$, a total relative frequency uncertainty at the low 10^{-15} level seems to be feasible at these operating conditions. This uncertainty can be obtained at an instability of $\sigma_y(\tau) \approx 3.6\cdot10^{-13}\,(\tau/s)^{-1/2}$.

We plan to install a second microwave cavity in the present fountain setup in order to preselect launched Cs atoms in the sublevel ($F = 3$, $m_F = 0$) using a scheme similar to that described in [1]. If an interrogation oscillator with a short-time stability $\sigma_y(1s) < 10^{-13}$ is employed, the preselection allows to further increase the stability of the fountain frequency standard without increasing the collisional shift. In addition, the state preselection also would allow to further reduce some uncertainty contributions (e.g. due to Ramsey pulling).

5. REFERENCES

[1] A. Clairon, S. Ghezali, G. Santarelli, Ph. Laurent, S. N. Lea, M. Bahoura, E. Simon, S. Weyers, K. Szymaniec, "Preliminary accuracy evaluation of a cesium fountain frequency standard", in Proc. 5th Symp. Freq. Standards and Metrology, Woods Hole, MA, Singapore: World Scientific Publishing, 1997, pp. 49–59.

[2] R. Schröder, U. Hübner, "The microwave cavity in the PTB caesium fountain", in these Proceedings.

[3] S. Weyers, A. Bauch, D. Griebsch, U. Hübner, R. Schröder, Ch. Tamm, "First results of PTB's atomic caesium fountain", in Proc. Joint Mtg. Europ. Freq. Time Forum and the IEEE Int. Freq. Contr. Symp., Besançon, France, pp. 16–19, 1999.

[4] S. Weyers, A. Bauch, U. Hübner, R. Schröder, Ch. Tamm, "First performance results of PTB's atomic caesium fountain and a study of contributions to its frequency instability", to be published in IEEE Trans. Ultrason., Ferroelect., Freq. Contr., vol. 47, No. 2, March 2000.

[5] E. Tiesinga, B. J. Verhaar, H. T. C. Stoof, and D. van Bragt, "Spin-exchange frequency shift in a cesium atomic fountain", Phys. Rev. A, vol. 45, No. 5, pp. R2671–R2673, March 1992.

[6] S. Ghezali, Ph. Laurent, S. N. Lea, A. Clairon, "An experimental study of the spin-exchange frequency shift in a laser-cooled cesium fountain frequency standard", Europhys. Lett., vol. 36, No. 1, pp. 25–30, October 1996.

DEVELOPMENT OF THE ON/OFMET CONTINUOUS CS FOUNTAIN STANDARD: A PROGRESS REPORT

Gaetano MILETI, Alain JOYET, Patrick BERTHOUD, and Pierre THOMANN*
Observatoire cantonal, 58 Rue de l'Observatoire, CH-2000 Neuchâtel, Switzerland
*Tel: 41 32 889 6870 Fax: 41 32 889 6281 email: Pierre.Thomann@ne.ch

Gregor DUDLE, Federal office of metrology, Lindenweg 50, CH-3003 Bern-Wabern, Switzerland

ABSTRACT

We report on the design, construction, and preliminary measurements on the resonator of a continuous Cs fountain frequency standard. The construction of the resonator is described, and preliminary measurements of the available atomic flux, of the beam temperature, and of their sensitivity to magnetic field are presented, along with the first Ramsey fringes obtained in this new type of fountain.

1. INTRODUCTION

Most prototypes of Cs fountain frequency standards rely on a pulsed mode of operation [1]: a cloud of cold atoms is first captured in an optical molasses or in a magneto-optic trap, cooled to low temperatures (a few μK), sent in a vertical ballistic flight for microwave interrogation and analyzed on its return in the source region. This process is repeated at a rate of roughly one cycle/second. As a result of the pulsed operation, the difference between the interrogating frequency and the atomic frequency, which is probed during the half-second near the apogee, is sampled with a duty cycle significantly smaller than one.

Cold fountains can lead to large improvements of both the short-term stability and the accuracy of cesium standards. However, their operation in a pulsed mode may introduce difficulties to achieve the ultimate short-term stability and accuracy that the use of cold atoms allow in principle. The best accuracy in cesium fountains is achieved for moderate atomic densities to limit the contribution of the collisional shift, which is proportional to the cold cloud density. The short-term stability, however, improves as the square-root of the number of atoms probed in each cycle, and its optimization would require as high an average atomic flux as possible. (The resulting compromise between stability and accuracy is almost completely relaxed in the case of Rb, which shows a much lower collisional cross-section than cesium [2]). Moreover, the short-term stability of all pulsed passive standards is affected by aliasing effects [2], which impose stringent requirements on the quality of the local oscillator used to produce the interrogating microwave signal.

Several methods have been suggested to overcome these potential limitations, including the use of tandem pulsed resonators [3], or of multiple clouds sent in such a way that atoms are present in the interrogating region at all times [4].

The continuous beam approach [5] provides a simple way of reducing by about two orders of magnitude both the collisional shift and the aliasing effects tied to pulsed operation. Based on the experimental demonstration of a continuous beam of $2 \cdot 10^8$ atoms/s [6], the Allan deviation of a continuous fountain standard using the same initial atomic flux is expected to be $7 \cdot 10^{-14}\ \tau^{-1/2}$ and the potential for accuracy below 10^{-15} [7].

We report here on the design, construction, and preliminary operation of a continuous fountain as the first steps in the development of a primary standard for the Swiss federal office of metrology.

2. FOUNTAIN DESCRIPTION

2.1 Atomic beam and fountain

The basic principle of the continuous fountain is closely related to that of the now well-known pulsed fountain [1]. The main difference, illustrated in the simplified schematics of fig. 1, lies in the parabolic - rather than linear vertical - shape of the average atomic path, which provides the required physical separation between atomic preparation and detection regions. Fig. 2 gives a more detailed picture of the fountain configuration.

The lower part of the vacuum vessel is separated in two compartments: source and detection. They are efficiently isolated from each other for stray light and cesium atoms by appropriate traps and getters while being connected to the same ion pump. Cold atoms are produced in the source part by capture and cooling from a thermal cesium vapour at the intersection of three pairs of mutually orthogonal laser beams. Each beam has an e^{-2} diameter of 25mm, a power of 15mW, and a circular polarisation opposite to that of the counterpropagating beam (σ^+-σ^- molasses). One pair is horizontal (Ox) and is retroreflected by a roof prism inside the vacuum system. The other two pairs are in the Oyz plane and propagate at 45° with respect to the

vertical. The atomic beam is launched vertically upwards at an initial velocity v_0 by the moving molasses technique if the frequency of the upward {downward} cooling beam is shifted by an amount $\Delta f = +\{-\} v_0/(\lambda \cdot \sqrt{2})$.

In the present state of the fountain, the only atomic preparation step following the extraction from the source is optical pumping from the F=4 hyperfine level to the F=3 level by a weak laser beam tuned to the F=4 – F'=4 frequency. Atoms then pass three layers of mu-metal magnetic shielding before entering the microwave cavity (see below), where the Ramsey interrogation process takes place in two separate, symmetrical interaction regions. The height of the parabola above the cavity is about 30 cm, yielding a 0.5s transit time between the microwave interactions, and a resonance line Q of 10^{10}.

In the detection region, the atomic beam crosses a retroreflected probe beam from a narrow linewidth (200 kHz), extended-cavity diode laser tuned to the cycling F=4 - F'=5 transition. The F=4 population, which is proportional to the microwave transition probability, is monitored by the collected fluorescent light. The cycling transition yields up to 10^4 fluorescence photons per atom. With a collection efficiency of >2%, the number of photons detected per atom in the F=4 level is larger than one, ensuring shot-noise limited detection whenever permitted by the thermal noise floor, presently at $2 \cdot 10^{-14}$ W/Hz$^{1/2}$. Shot-noise detection can thus be obtained for the expected atomic beam flux of 10^5 atoms/s.

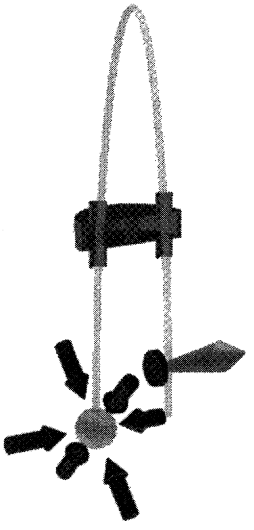


Fig. 1: Basic configuration of the continuous fountain, showing the 6 cooling beams, the microwave cavity and the detection region.

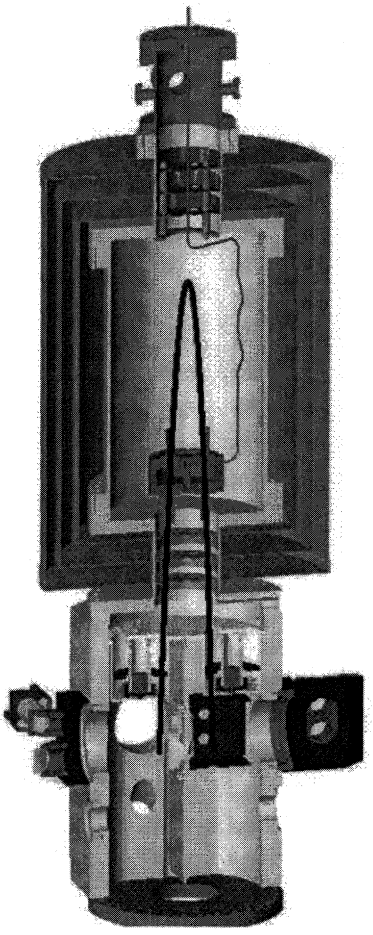


Fig. 2: Schematic view of the fountain resonator

2.2 Microwave cavity

The microwave cavity (fig. 3) is made from oxygen free, high conductivity (OFHC) copper. It is a cylindrical, coaxial cavity oscillating in a TE 021 mode. As shown in fig. 4, the RF magnetic field in each interaction zone is vertical (i.e. axial), the field distribution in the (r,z) plane is independent of the azimuthal angle. The microwave radiation can be coupled into and out of the cavity through two irises placed at equal distances of the atomic paths and connected to waveguide-to-coaxial transitions pieces.

The two end plates are fitted with cut-off waveguides on each of the four atomic beam openings to prevent microwave leakage out of the cavity. The calculated attenuation for the dominant evanescent mode is more than 150dB, neglecting the finite wall conductivity. The measured attenuation is >110 dB, presently limited by the measurement technique. All components of the cavity are assembled with indium seals. The electrical contact between body and end plates is ensured at the end of a quarter-wave section of narrow-gap coaxial cavity (choke).

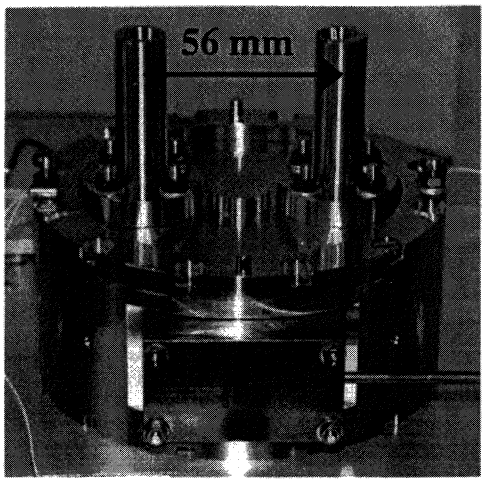


Fig. 3: Photograph of the microwave cavity. One of the coax-to-waveguide feeds is recognizable in the foreground.

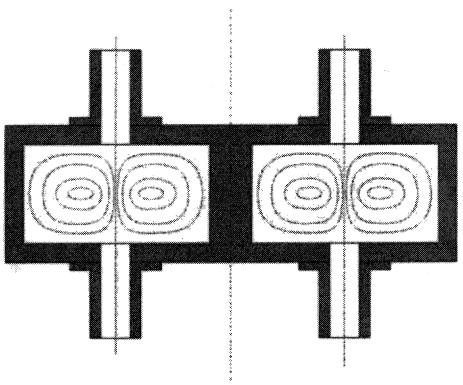


Fig. 4: Cut through the microwave cavity in the (r,z) plane of the parabola. The lines inside the resonator are the H-field lines. Not drawn to scale are the cut-offs which are in reality tilted to keep loss of atoms due to geometrical reasons to a minimum. The central post is on the axis of rotational symmetry for the field mode (Oz).

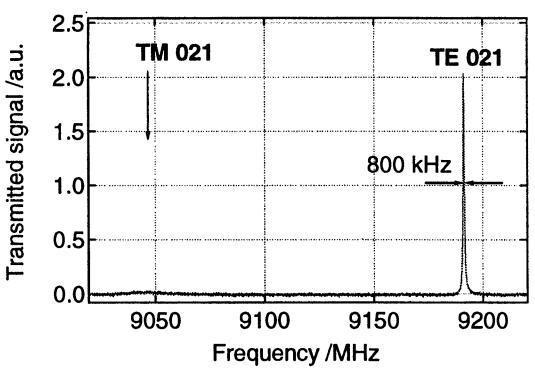


Fig. 5: Transmission spectrum of the cavity. In addition to the TE 021 mode one makes out the TM 021 mode, 140 MHz below the cesium resonance.

This feature efficiently shifts the frequency of the degenerate and unwanted TM mode and inhibits its oscillation.

The measured loaded Q is 11'500, and the measured temperature coefficient in vacuum (-140 kHz/K) is close to a calculated value based on the linear expansion coefficient of copper (-152 kHz/K).

2.3 Magnetic field and shields

Three cylindrical magnetic shields enclose the Ramsey interaction region alone. They are made from 1mm thick Mumetal. A fourth magnetic field will be added around the whole fountain. The static shielding factor, defined as the ratio of an ambient axial field variation of 10μT to the corresponding axial variation at the center of the three shield structure, is 3'500.

The C-field is wound on an aluminum cylinder inside the inner shield. The shields and the microwave cavity construction allow an in-situ measurement of the axial field in the whole interaction region after assembly and demagnetization. Such a field profile is shown in fig. 6. The rms inhomogeneity in the microwave interaction region, dominated by the penetration of the outside field, is 1.5 nT , which is expected to contribute less than 10^{-16} to the inaccuracy.

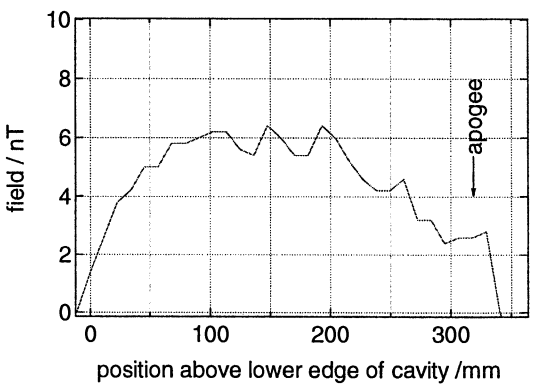


Fig. 6: Axial profile of the static magnetic ("C") field. Inhomogeneity is 1.5nT rms over the atomic path.

3. PRELIMINARY MEASUREMENTS

3.1 Continuous fountain

Preliminary flux measurements have been performed on the fountain without microwave cavity. Two different methods, one based on the optical calibration of the detection system, the other on the measurement of the atomic shot noise, lead to a total flux of $>10^5$ atoms/s (all m_F values) in the detection region , which is still short by close to an order of magnitude of the expected value.

A "pusher" laser beam tuned to the F=4 – F'=5 transition has been implemented to cut thin slices (9.5 ms) of the atomic beam. The time-of-flight (TOF)

distribution is recorded using the fluorescence in the probe beam. A TOF spectrum is shown in fig. 7. The longitudinal temperature deduced from the TOF distribution is 55μK.

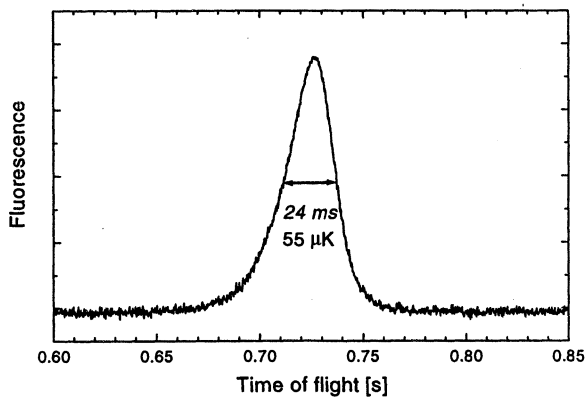


Fig.7: TOF spectrum of the continuous fountain beam.

We have measured the beam temperature as a function of the vertical component B_z of the magnetic field used to compensate the earth's field in the source region (fig. 8). The temperature shows a clear minimum that can be used to adjust the compensating fields.

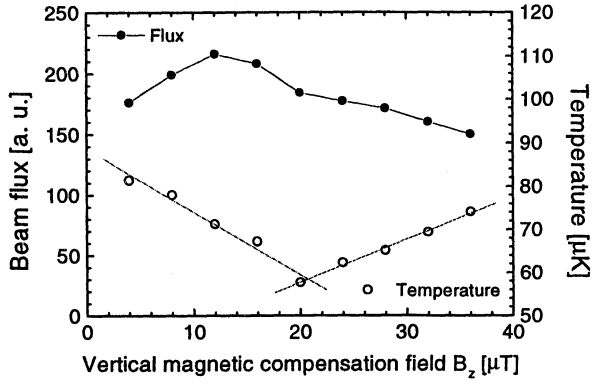


Fig. 8: Longitudinal temperature and beam intensity (signal) vs vertical magnetic field B_z in molasses.

The horizontal component B_x of the magnetic field in the parabola plane has a similar effect on the beam temperature (fig.9). However, the beam intensity shows a sharp dip on top of a broad maximum. This can be explained by two simultaneous effects of the transverse fields: heating and acceleration. The useful flux along the parabola is at an angle of 1.5° to the vertical direction of the moving molasses. As the transverse field is varied, both the divergence of the atomic beam (transverse temperature) and its average tilt vary. A numerical simulation of this combined effect, based on experimental heating and drift coefficients [8], is shown in fig. 9 and reproduces qualitatively the observed feature.

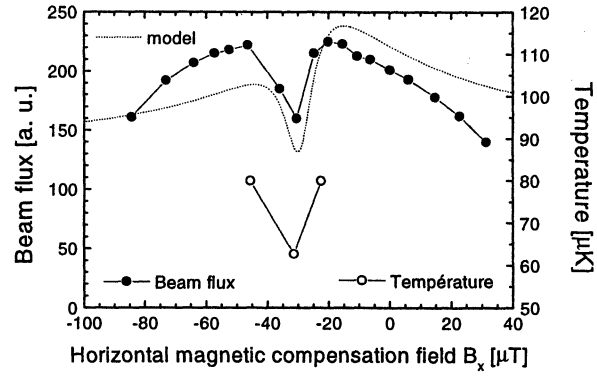


Fig. 9: Beam temperature (◦) and beam intensity (• : measurement; dotted line: model) vs. transverse magnetic field B_x.

Although most efforts are now concentrated towards optimization of the atomic flux, it has been possible under the present operating conditions to obtain preliminary Ramsey fringes near the clock transition frequency (C-field = 1μT) by scanning the frequency of the microwave power in the cavity. Figure 10 shows a scan near the nominal Cs frequency. The 9.2 GHz signal was obtained from a quartz oscillator locked to an H-maser and a frequency multiplier from a commercial Cs standard. The S/N ratio does not yet allow a detailed analysis of the fringes. However, their width (1Hz) is as expected and the decreasing contrast far from the central fringe is explained by the beam temperature, as demonstrated in fig. 11 by the theoretical Ramsey pattern in a 50μK fountain.

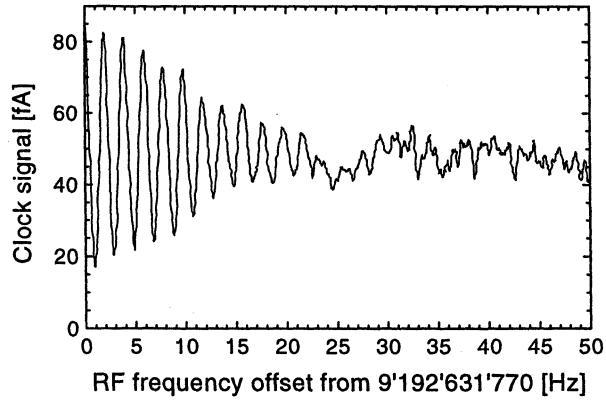


Fig. 10: Ramsey fringes of the clock transition (F=3;m=0 – F=4;m=0) in the continuous fountain. The spacing is 1Hz. The decreasing contrast with increasing detuning may be used to estimate the temperature of the atomic beam (see fig 11).

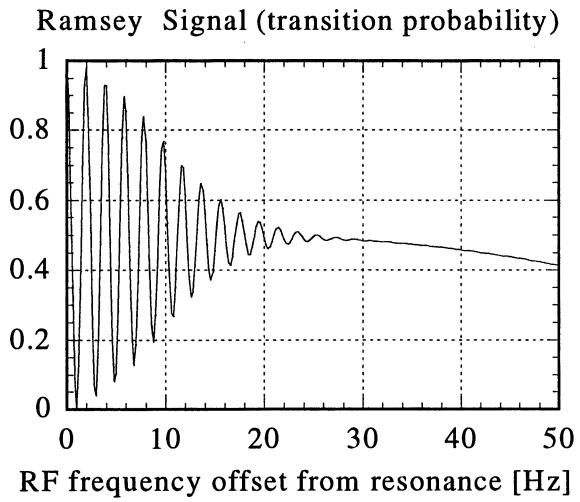


Fig.11: Calculated Ramsey pattern at optimum RF power of a fountain beam with 50μK longitudinal temperature (10ms RF interactions, 500ms free precession time).

4. CONCLUSION AND OUTLOOK

A continuous fountain resonator has been constructed and first Ramsey fringes have been obtained. Optimisation of the beam flux is necessary in a first step in order to increase the signal to noise ratio. 2D transverse cooling will be implemented toward this end.

Implementation of the frequency synthesis under construction at NIST and of the signal processing software under development will allow to start frequency measurements and accuracy evaluation.

Rotating light traps to prevent stray light in the microwave interrogation from the source and detection regions will be added after measurement of the unattenuated light-shift.

5. REFERENCES

[1] A. Clairon et al. "Ramsey resonance in a Zacharias fountain", Europhysics Lett., vol. 16, p165, 1991. See also: Sessions on "laser cooling and atomic clocks", in Proceedings of the joint Meeting of the EFTF and IEEE FCS , Besançon, 13-16 April 1999, pp. 9-156.

[2] K. Gibble and T. Verhaar, "Eliminating cold-collision frequency shifts...", Phys. Rev A, Vol. 52, 1995, p.3370.

[3] J. Dick et al., Proceedings of the 22nd annual PTTI Meeting, 1990, pp. 487-508.

[4] C. Fertig and K. Gibble, "Laser-cooled Rb87 clock", IEEE Trans. Instr. Meas., Vol. 48, April 1999, pp 520-523.

[5] G. Dudle et al., Generation of a continuous beam of cold atoms, Symposium on frequency standards and metrology, pp 121-126, World Scientific, 1995.

[6] P. Berthoud, E. Fretel, P. Thomann, "Bright, slow, continuous beam of laser-cooled atoms", Phys. Rev. A, R4241-4244, Dec. 1999.

[7] P. Berthoud et al., Toward a frequency standard based on a continuous fountain of laser-cooled atoms, IEEE Trans. Instr. Meas., vol 48, pp516-519, April 1999.

[8] P. Berthoud, thesis, Université de Neuchâtel, March 2000.

ATOM NUMBER DEPENDENT FREQUENCY SHIFTS IN A ^{87}Rb FOUNTAIN

S. Bize*, Y. Sortais*, C. Nicolas*, C. Mandache†, G. Santarelli*, C. Salomon‡ and A. Clairon*

ABSTRACT

We describe a measurement of the collisional frequency shift in a ^{87}Rb cold atom fountain with a frequency resolution of 3×10^{-16}. In the clock configuration, when only the $F = 1, m_F = 0$ and $F = 2, m_F = 0$ levels are populated, the collisional frequency shift is found to be 300 times smaller than for ^{133}Cs. Our measurement is in disagreement by 6.7 standard deviations with theory published to date. In addition, we measure, for the first time in cold atom fountains, the cavity pulling frequency shift. This shift well-known in hydrogen masers varies linearly with the number of atoms passing through the microwave cavity, as does the collisional shift. We have therefore subtracted its contribution to deduce the collisional frequency shift.

1 INTRODUCTION

In the past 10 years, the accuracy of microwave frequency standards has been improved by using laser cooling techniques. The relative accuracy of LPTF's cold atom fountains is now equal to 1.1×10^{-15} for ^{133}Cs [8] and 2×10^{-15} for ^{87}Rb [1]. With a current short term stability of $1.2 \times 10^{-13}\tau^{-1/2}$, which is limited by the interrogation oscillator phase noise, 1.4×10^4 s are needed to reach a 10^{-15} relative frequency resolution with these devices. Increasing the resolution by a factor 10 would require a 16 days long integration. The evaluation of all systematic effects with a 10^{-16} relative accuracy seems hard to achieve in a reasonable amount of time.

As shown in [9], the short term stability of cold atom fountains can be improved by using an interrogation oscillator with a lower phase noise, such as a sapphire cryogenic oscillator (SCO), instead of a BVA quartz oscillator. When using a SCO, the short term stability of a Cs fountain can be as low as $4 \times 10^{-14}\tau^{-1/2}$, the best reported value for primary frequency standards. In this case, the short term stability is limited by the quantum projection noise which is inversely proportional to the number N_{at} of detected atoms. Unless samples of spin correlated atoms are used, increasing the atomic flux is the only way to further improve the short term stability in this regime. Prospects for increasing N_{at} obviously raises the question of the collisional frequency shift, which has been measured to be $-2\times10^{-21}/(\text{at.cm}^{-3})$ for Cs [5]. It produces a significant frequency shift in cold atom fountains even at the 10^{-15} level, which means that the average atomic density has to be controlled below the 10% level, a hard task. Based on photoassociation cold collision experiments [2] and theoretical calculations [7], the collisional frequency shift in cold ^{87}Rb clocks is expected to be 15 times smaller than for Cs. Rubidium fountains thus represent an attractive alternative to Cs frequency standards.

2 ATOM NUMBER DEPENDENT EFFECTS IN COLD ATOM FOUNTAINS

In this paper, we present a measurement of the two main atom number dependent effects in a ^{87}Rb fountain, the collisional frequency shift and the cavity pulling effect. The collisional frequency shift is due to the interaction between atoms inside the atomic cloud during the Ramsey interaction. At low atomic density, the effect of the interaction can be described in a two-body mean-field approximation. The total mean-field energy shift of the atomic cloud is equal to $(4\pi\hbar^2/m)\times a\times\bar{n}\times N$ for a non-degenerated gas where $\bar{n}$ is the average atomic density and N the total number of atoms. a is the s-wave scattering length. In the fountain, we drive a transition between two internal states $|1\rangle$ and $|2\rangle$, separated by an energy $\hbar\omega_{21}$, in such a way that all the population is transferred from $|1\rangle$ to $|2\rangle$ when the field is tuned on resonance. The driving field frequency has to match the energy $\hbar\omega_{21}$ corrected by the difference between the mean-energies per atom in the initial state $|1\rangle$ and the final state $|2\rangle$. The collisional frequency shift is thus given by $2\pi\delta\nu = (4\pi\hbar/m) \times (a_{22} - a_{11}) \times \bar{n}$, where a_{ii} is

*Laboratoire Primaire du Temps et des Fréquences, 61, avenue de l'Observatoire, 75014 Paris, France, email : sebastien.bize@obspm.fr

†Institutul National de Fizica Laserilor, Plasmei si Radiatiei, P.O.Box-MG36, Bucuresti, Magurele, Romania

‡Laboratoire Kastler Brossel, 24, rue Lhomond, F-75231 Paris cedex 05, France, email : salomon@physique.ens.fr

the scattering length for atoms colliding in the internal state $|i\rangle$. In a more general case, the collisional frequency shift is given by the following expression [7]:

$$\delta\nu = -(2\hbar/m)\sum_j \bar{n}_j(1+\delta_{10,j})(1+\delta_{20,j})Re(a_{10,j}-a_{20,j}) \quad (1)$$

In this formula, the j index refers to the atomic Zeeman substate F, m_F. $j = 10$ and 20 represent the lower and the higher clock levels ($F = 1, m_F = 0$ and $F = 2, m_F = 0$). $a_{10,j}$ and $a_{20,j}$ are the scattering lengths for atoms colliding in the $10+j$ and $20+j$ internal states. The Kronecker symbol δ_{ij} accounts for quantum statistics.

The second effect which depends on atom number is the cavity frequency pulling, well-known in hydrogen masers [10]. In our fountain, the atomic transition is probed by a 6.8 GHz microwave field in a TE_{011} resonator with a quality factor $Q = 10000$. The pulling effect is due to the interference inside the microwave resonator between the field radiated by the input coupler and the field radiated by the atomic magnetic dipoles, when the atoms pass through the cavity. This interference induces a time dependent phase shift between the field inside the resonator and the signal delivered by the interrogation oscillator. It produces a clock shift that exhibits a dispersive dependence as a function of the cavity detuning with respect to the atomic resonance. Both the width and the amplitude of this dispersion curve are proportional to Q. The amplitude is also proportional to the number of atoms crossing the cavity. When the atoms enter the cavity in the upper state (resp. lower state) and deposit energy in (resp. remove energy from) the cavity, the clock frequency is *pulled* towards (resp. *pushed* away from) the cavity resonance frequency.

3 EXPERIMENTAL SETUP

The cold atom Rb fountain has been described in a previous paper [1]. We only underline the most important features. The fountain operates in a sequential mode. Rubidium atoms are first loaded into a magneto-optical trap (MOT). The magnetic field is then switched off and the atoms are launched upwards at $\sim$ 3.5 m.s^{-1} using the moving molasses technique and cooled at $1.4\mu K$ in the moving frame. After launch, atoms can be selected in a specific Zeeman sublevel by means of microwave and laser pulses. On their way up, atoms then cross a copper resonator fed with a very low phase noise microwave source near 6.8GHz, the ^{87}Rb ground state hyperfine frequency. On their way down, they cross the resonator a second time after a ballistic free flight lasting $\sim$ 500ms. This reproduces the well-known Ramsey interrogation scheme. After the Ramsey interaction, atomic populations $N_{F=1}$ and $N_{F=2}$ in both hyperfine levels are measured by laser induced fluorescence. From the transition probabilities $N_{F=2}/(N_{F=1}+N_{F=2})$ measured on both sides of the central Ramsey fringe, we compute an error signal to lock the microwave interrogation oscillator on the atomic transition using a numerical servo loop.

After launch, atoms are spread among the various m_F substates of the $F = 2$ manifold. The clock has been operated in three configurations. i) normal clock operation : atoms in the $F = 2, m_F = 0$ substate are transfered to $|1,0>$ using a microwave π pulse of duration 432 μs. Atoms remaining in $|2, m_F \neq 0>$ are pushed away by radiation pressure. In this case, only three scattering lengths $a_{10,10}$, $a_{10,20}$ and $a_{20,20}$ in equation 1 are involved. ii) no selection : all Zeeman sublevels in $F = 2$ are equally populated to $2-3\%$. The frequency shift of the $|2,0> \rightarrow |1,0>$ transition is measured in the presence of $|2, m_F \neq 0>$ populations. iii) Atoms are pumped in $F = 1$ by a 9.2 ms laser pulse tuned to the $F = 2 \rightarrow F' = 2$ optical transition. Atoms remaining in $F = 2$ state ($\leq 2\%$) because of incomplete optical pumping are pushed away by radiation pressure. The launched atoms are equally spread among the $F = 1$ substates to within 1%. In this case, the temperature is increased from $\sim$ 1.4 μK to $\sim$ 2.2 μK. The frequency shift of the $|1,0> \rightarrow |2,0>$ transition is measured in the presence of $|1, m_F \neq 0>$ populations.

4 DIFFERENTIAL MEASUREMENTS

The atom number dependent frequency shifts are measured using a differential method. The number of launched atoms is changed every 200 fountain cycles ($\sim$ 350 s). This duration is much longer than the time constant of the servo loop used to lock the microwave source on the atomic transition ($\sim$ 4 s), and yet shorter than the typical time scale of the fluctuations of the clock environment. We vary either (1) the loading duration of the MOT from 100 to 1000 ms or (2) the laser intensities during the loading phase while keeping the loading duration constant and equal to 800 ms. This technique rejects all long term fluctuations which are not associated with atom number or density, such as the maser frequency drifts, and room temperature, black-body radiation and long term magnetic field fluctuations. We have checked by comparing method (1) and (2), that the magnetic field from the MOT coils in method (1) does not produce any systematic frequency shift in the differential measurement. When the number of $m_F = 0$ atoms in the detection region N_{at} exceeds 3×10^6, the clock frequency stabil-

ity is $1.5 \times 10^{-13}\tau^{-1/2}$ where τ is the averaging time in seconds. In the differential method, the measurement resolution is degraded to $2 - 4 \times 10^{-13}\tau^{-1/2}$ depending on the experimental parameters. Each differential measurement point lasts $\sim$ 4 days and the relative frequency resolution on the effects which depend on atom number reaches $8 - 15 \times 10^{-16}$.

5 ATOM NUMBER AND ATOMIC DENSITY CALIBRATION

For a given fountain geometry, three pieces of information are requested in order to make an absolute calibration of the atomic density : the number of launched atoms, the spatial distribution just after launch, and the atomic velocity distribution.

The initial size of the atomic cloud is first measured by imaging the atomic cloud on a CCD camera. In order to discriminate between various spatial distributions, we use an absorption method performed at the end of the cooling process, just after launch time. This method gives information on both number and spatial distribution of the atoms initially. We monitor the transmission through the atomic cloud of a weak probe beam. We check that the absorption signal is not polarisation dependent : all Zeeman sublevels are equally populated, and the coupling coefficient is taken equal to the average value of the Clebsch-Gordan coefficients of the $F = 2 \rightarrow F' = 3$ transitions. The absorption signal is measured for various diaphragm diameters (20 mm down to 1 mm). We fit the results with a model taking into account the non-linearity of the absorption signal with atom number due to the optical thickness of the cloud, as well as the intensity profile of the probe beam. We find that the spatial distribution is well described by a $p(\vec{r}) = A/(1 + (\vec{r}^2/r_c^2))^\alpha$ law, where $r_c = 2.1\,\text{mm}$ and $\alpha = 2.5$, and we deduce a typical number of 3×10^8 launched atoms, spread among the $|F = 2, m_F >$ levels. We estimate that both size and number of atoms are known within 20%, over the 6 month duration of our measurements.

We then use the time of flight signal in the detection region to determine the atomic velocity distribution. These signals provide 1D information and exhibit a non-gaussian shape (fig. 1). In order to further probe the velocity distribution, we perform a 2D test by measuring the number of detected atoms as a function of the launch velocity, the atomic cloud being truncated by various circular diaphragms, such as the cavity cutoffs. We compare these experimental data to a Monte-Carlo simulation, that takes into account the saturation profile of the laser beams in the detection region, the initial spatial and velocity distributions. As shown in fig. 2, the fraction of detected atoms strongly depends on the velocity distribution model. The two models (gaussian and lorentzian-like) fitting the time of flight in fig. 1 lead to a discrepancy by a factor $\sim$ 2 on the fraction of detected atoms. It leads to a distribution which is in good agreement with a $p(\vec{v}) = B/(1 + (\vec{v}^2/v_c^2))^\beta$ law, as predicted in [3].

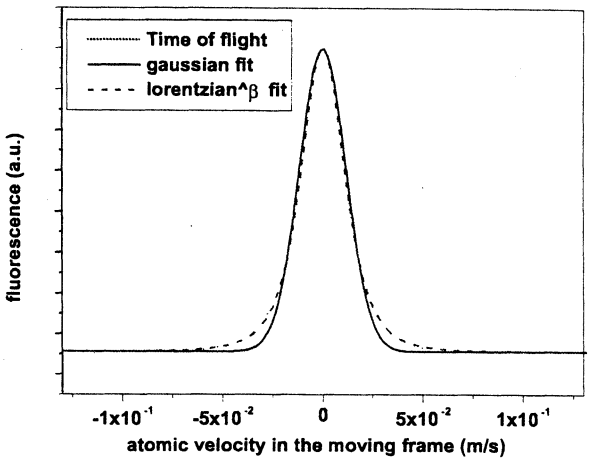


Figure 1: Time of flight signal as a function of atomic velocity in the moving frame. The initial signal is a function of time, and horizontal axis has been rescaled in order to read the atomic velocity in the frame of the falling atoms. The dotted line corresponds to atoms detected in the $F = 2$ hyperfine level crossing the upper detection beam. The time of flight has been fitted by a gaussian law (solid line) with $\sigma_v = 11.7\,\text{mm/s}$, and by a $p(\vec{v}) = B/(1 + (\vec{v}^2/v_c^2))^\beta$ law (dashed line) with $v_c = 18\,\text{mm/s}$ and $\beta = 2.07$. The second law gives much better agreement with experiment.

The number of detected atoms as deduced from the Monte-Carlo simulation and the absorption method is in agreement with an absolute calibration of the detection region within 6%.

From the Monte-Carlo simulation, we also deduce the number of atoms crossing the microwave cavity, N_{cav}. This value is used as an input for a numerical calculation of the cavity pulling effect. By changing the parameters of the spatial and velocity distributions used in the Monte-Carlo simulation, we estimate that N_{cav} is model dependent within 10%. Together with the 20% uncertainty on the initial number of atoms, we obtain a 25% uncertainty on the number of atoms in the cavity.

The Monte-Carlo simulation finally gives a model of the atomic cloud evolution with time from which we deduce an absolute calibration of the atomic density. Due to the atomic cloud expansion, the atomic density varies by a factor $\sim$ 30 during the Ramsey interaction. We thus need to take into account the non uniform sensitivity of the clock frequency to a time dependent perturbation of the atomic frequency. For each atomic level i, the time dependent atomic density is weighted by the sensitivity function

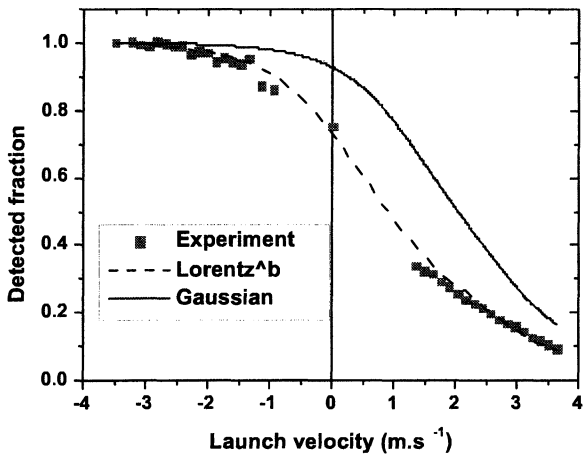


Figure 2: Detected fraction of atoms as a function of the launching velocity, for the velocity distributions fitting the time of flight on fig. 1. Square points : experimental data. Dashed line : results of the Monte-Carlo simulation for a lorentzian like law ($\beta = 2.07$ and $v_c = 18\,\mathrm{mm/s}$). Solid line : Monte-Carlo simulation for a gaussian law ($\sigma_v = 11.7\,\mathrm{mm/s}$). We normalize each point by the number of detected atoms when launching downwards at 3.5 m/s. The absolute fraction of detected atoms is then equal to 95% in the case of the Lorentz like distribution, and $\sim 100\%$ in the case of a gaussian distribution.

$g(t)$ and the transition probability $P_i(t)$. We then average over the duration of the Ramsey interaction and over the atomic trajectories that actually cross the detection region (*i.e.* we only take into atoms that contribute to the clock frequency corrections). This leads, for each level i, to an *effective* atomic density $\bar{n}_i$, that is used in eq. 1 in order to find the predicted collisional frequency shift. Typically, when $|F = 1, m_F = 0 >$ is initially selected (case i)), the total effective density $\bar{n} = \bar{n}_{F=1,m_F=0} + \bar{n}_{F=2,m_F=0}$ reaches $2 \times 10^7\,\mathrm{at.cm^{-3}}$. The effective atomic density has a sharper sensitivity to the models used for the spatial and velocity distributions. The overall uncertainty on $\bar{n}$ is 60%.

6 EXPERIMENTAL RESULTS

Figure 3 shows the measurement of the clock frequency shift as a function of the cavity detuning $\Delta\nu_{cav} = \nu_{cav} - \nu_{at}$ with respect to the atomic resonance, for atoms initially selected in the $F = 1, m_F = 0$ Zeeman sublevel. Data points have been rescaled to a number of atoms crossing the microwave resonator equal to 3×10^7, the highest value achieved in this experiment. The frequency and the quality factor of the microwave cavity have been independently measured as described in [4], with an uncertainty of 20 kHz and 4% respectively. The plain line on this plot represents a least square fit of the

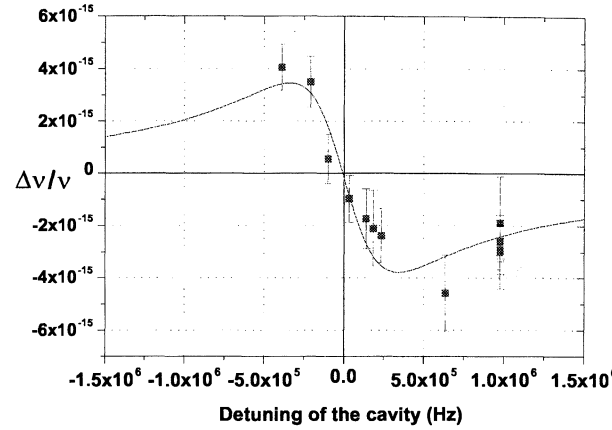


Figure 3: Frequency shift as a function of the microwave cavity detuning with respect to the atomic resonance, for atoms initially selected in the $F = 1, m_F = 0$ state. Data points have been rescaled to a number of atoms crossing the microwave resonator equal to 3×10^7.

data points using the following theoretical model : $\delta\nu/\nu = A(K_{1,0}^{col} \times \bar{n} + f(\Delta\nu_{cav}))$. $f(\Delta\nu_{cav})$ represents the dispersive shape of the cavity pulling effect as expected from theoretical calculations for the measured quality factor when 3×10^7 atoms enter the microwave resonator. $\bar{n}$ is calculated for each data point using our Monte Carlo simulation, given the measured spatial and velocity distributions, and the number of detected atoms. The collisional frequency shift coefficient $K_{1,0}^{col}$ is left as an adjustable parameter. Coefficient A is a global adjustable scale factor on the calibration of the number of launched atoms. We find $A = 1.33 \pm 0.15$ which is in good agreement with our estimation of the number of atoms in the cavity. From this least square fit, we find $K_{1,0}^{col} = (-6.4 \pm 16.4) \times 10^{-24}/(\mathrm{at.cm^{-3}})$ by taking only the frequency measurement uncertainty. This corresponds to a frequency resolution of 3×10^{-16}. Taking into account the 60% uncertainty on the effective atomic density, we find $K_{1,0}^{col} = (-6.4 \pm 19.2) \times 10^{-24}/(\mathrm{at.cm^{-3}})$. For comparison, $K_{1,0}^{col}$ is predicted to be $-11.7 \times 10^{-23}/(\mathrm{at.cm^{-3}})$ in this case [7]. Our result is 6.7 standard deviation away from the theoretical value. Figure 4 shows our measurement of the collisional frequency shift as a function of the effective atomic density $\bar{n}$. Each data point in 4 corresponds to a data point in fig. 3, from which we have subtracted the cavity pulling contribution.

We have also measured the collisional frequency shift of the $|F = 1, m_F = 0 > \rightarrow F = 2, m_F = 0$ transition when selecting the atoms in $|F = 2, all\, m_F >$ (case ii)), and also in $|F = 1, all\, m_F >$ (case iii)). After we subtract the cavity pulling effect (only the two clock states that couple to the microwave field contribute to this effect), we

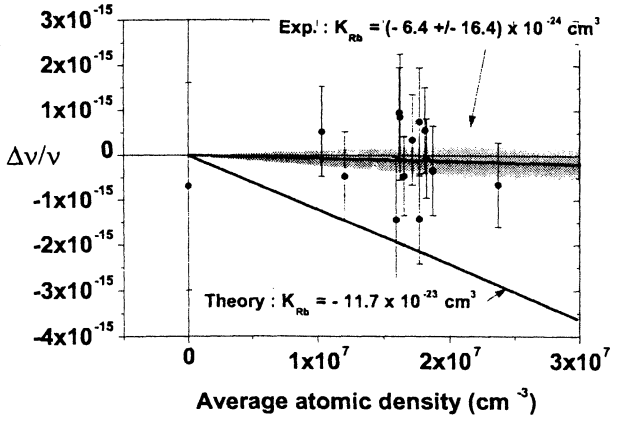


Figure 4: Frequency shift as a function of the atomic density for atoms initially selected in $F = 1, m_F = 0$.

find $K_2^{col} = -(5.3 \pm 1.9) \times 10^{-23}/(\text{at.cm}^{-3})$ and $K_1^{col} = -(5.0 \pm 1.8) \times 10^{-23}/(\text{at.cm}^{-3})$. Both values are in agreement with theoretical predictions within a few our error bars (predicted values are $K_2^{col} = -5.59 \times 10^{-23}/(\text{at.cm}^{-3})$ and $K_1^{col} = -6.76 \times 10^{-23}/(\text{at.cm}^{-3})$).

7 CONCLUSIONS

In summary, we have measured the collisional frequency shift in a ^{87}Rb cold atom fountain. For atoms initially selected in the $|F = 1, m_F = 0 >$ state, we find that this shift is equal to $(-6.4 \pm 19.2) \times 10^{-24}/(\text{at.cm}^{-3})$, a factor 300 smaller than for Cs, in disagreement with the predicted value of [7] by 6.7 standard deviations. When using an optical molasses instead of a MOT as a source for the fountain, the atomic density is reduced by a factor ~ 5 compared to the highest density achieved in this experiment while keeping the same number of detected atoms. It is thus possible to make the collisional frequency shift a negligible source of uncertainty at the 10^{-16} level in ^{87}Rb fountains. However, our measurement of the cavity pulling effect shows that the resonator frequency should be tuned on the atomic resonance within ~ 1 kHz in order to take all the benefits of the smallness of the collisional shift. Obviously, further investigation is necessary in order to explain the disagreement between theory and experiment in the regular clock configuration. At this point, it is worth mentioning that collision induced frequency shift theory also has discrepancy with experiment in room temperature and cold hydrogen masers [6, 11]. These discrepancies could come from a modification of the collisional process by the microwave field. Further improvement of the short term stability of cold atom fountains using very low phase noise cryogenic sapphire oscillator [9] could provide a sufficient frequency resolution to test a possible dependence of the collisional shift with the power inside the microwave resonator.

References

[1] S. Bize, Y. Sortais, M.S. Santos, C. Mandache, A. Clairon, and C. Salomon. "High-accuracy measurement of the ^{87}Rb ground-state hyperfine splitting in an atomic fountain". *Europhys. Lett.*, vol. 45, p. 558, March 1999.

[2] H. M. J. M. Boesten, C. C. Tsai, B. J. Verhaar, and D. J. Heinzen. "Observation of a shape resonance in cold-atom scattering by pulsed photoassociation". *Phys. Rev. Lett.*, vol. 77, p. 5194, December 1996.

[3] Y. Castin, J. Dalibard, and C. Cohen-Tannoudji. "The limits of Sisyphus cooling". in Proceedings of the Workshop, held in Marciana Marina, Elba Island, Italy, May 2–5, 1990, pp. 5-24.

[4] K. Fertig and K. Gibble. "Laser-cooled ^{87}Rb clock". *IEEE. Trans. Instrum. Meas.*, vol. 48, pp. 520–523, 1999.

[5] S. Ghezali, P. Laurent, S.N Lea, and A. Clairon. "An experimental study of the spin-exchange frequency shift in a laser cooled cesium fountain standard". *Europhys. Lett.*, vol. 36, p. 25, 1996.

[6] M.E. Hayden, M.D. Hürlimann, and W.N. Hardy. "Atomic hydrogen spin-exchange collisions in a cryogenic maser". *IEEE Trans. Instrum. Meas.*, vol. 42, p. 314, April 1993.

[7] B. Kokkelmans, B. Verhaar, K. Gibble, and D. Heinzen. "Predictions for laser cooled Rb clocks". *Phys. Rev. A.*, vol. 56, p. R4389, 1997.

[8] P. Laurent *et al.* "Cold atom clocks on earth and in space". in *Proceedings of the 14th Int. Conf. on Laser Spectroscopy*, Innsbruck, Austria, 1999.

[9] G. Santarelli *et al.* "Quantum projection noise in an atomic fountain: A high stability cesium frequency standard". *Phys. Rev. Lett*, vol. 82, p. 4619, June 1999.

[10] J. Vanier and C. Audoin. The Quantum Physics of Atomic Frequency Standards. Adam Hilger, 1989.

[11] B.J. Verhaar, J.M.V.A. Koelman, H.T.C. Stoof, O.J. Luiten, and S.B. Crampton. "Hyperfine contribution to spin-exchange frequency shift in the hydrogen maser". *Phys. Rev. A*, vol. 35, p. 3825, May 1987.

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

DEVELOPMENT OF A NEW REFERENCE CLOCK FOR TELECOMMUNICATIONS

Richard Percival & Mark Rainer
Quartzlock (UK) Ltd, Gothic, Plymouth Road, Totnes, Devon, TQ9 5LH, UK
Tel: +44 1803 862062 Fax: +44 1803 867962 email@ quartzlock@quartzlock.com

ABSTRACT

Quartzlock (UK) Ltd has recently been awarded a European Commission 'CRAFT' award to design a new reference clock for telecommunications; the ALPHA 1. The clock will find application within the fields of metrology, telecommunications and radio transmitter referencing.

In addition to the standard passive hydrogen maser and GPS elements, the ALPHA 1 will thus incorporate GLONASS, Rubidium, BVA crystal oscillator and VLF tracking receiver technology. A unique software clock will drive all elements in a modular high redundancy ring or grid. This allows existing user clocks to be joined at any time. A very high resolution and low noise floor frequency and phase measurement system will also be included.

The ALPHA 1 is intended to be a modular, flexible system featuring high-level redundancy. The objective is to combine and reduce a considerable number of diverse frequency standards to a single output. This will have improved performance (short, medium and long term frequency stability of the source), and high reliability in that the output should not show any sudden phase glitches if the output of one or more of the sources should fail. In space, for example, the ability of the clock to survive - and maintain stable timing in - extreme environmental conditions is at least as important as the actual frequency stability of the device. The vibrations encountered by such a clock during launch far exceed anything experienced in the laboratory.

To monitor the varying environmental conditions the ALPHA 1 may face, an in-built environmental monitor is being developed to continuously display parameters such as humidity, atmospheric pressure, magnetic field, vibration, radiation levels, static and lightning risk. In addition, it is not feasible to repair failed components and so normally satellite systems would carry several clocks in case one failed. The multiple redundancy of the ALPHA 1 renders this unnecessary.

This paper will include a full description of the projects 10 elements, with the improvement expected/already made outlined. The latest test results in tabular and graphical formats will be presented, along with a progress report for each element. A short discussion on anticipated applications for the Alpha 1 will conclude the paper

1. INTRODUCTION

Quartzlock (UK) Ltd has recently been granted €400k of European Commission money to design a new 'primary reference clock', the ALPHA 1. It hoped that such a clock would find applications within the fields of telecommunications, metrology and radio transmitter referencing. The clock will not be a primary frequency standard like NIST F1, the newly developed Caesium Fountain Frequency Standard, but will be a commercially available 'top-level' clock ensembling system for multiple applications. Quartzlock (UK) Ltd has proven to be the leading British company in the American dominated Time and Frequency field. The expertise gained during the previous 20 years in developing and producing highly accurate, stable and affordable Time and Frequency generation, distribution and measurement instrumentation will be used in this project.

This project will bring together the expertise of several European companies. Quartzlock (UK) Ltd is to be project leader and co-ordinator of the project. K&K Messtecknik Gmbh, a small German company will develop innovative software for the GPS elements of the ALPHA 1. A consortium of small companies comprising Simek, Cathodeon, Eltek, Menvier Hybrids and Farran technologies will provide expertise in thick/thin film technology, scientific glass blowing, crystal oscillators and ceramic substrate hybrids. The Physikalische-Technische Bundesantalt (PTB) Germany's national metrology institute will do the primary research into each individual element of the proposed 'Primary Reference Clock'. RF solutions and Mark Rainer will give expert advice in the fields of Hydrogen Maser and Rubidium electronics, GPS and LF Tracking receivers. [1] [2]

The ALPHA 1 is intended to be a modular, flexible system featuring high levels of redundancy and implementing "state of the art" electronics whilst taking care to allow for future developments to be incorporated in an evolutionary way rather than requiring replacement with consequent disruptions. This in effect means that should one element of the system fail, other components will ensure continuous high-level performance of the overall system. To monitor the varying environmental conditions the ALPHA 1 may face, an optional, in-built environmental monitor will continuously display parameters such as temperature, humidity, atmospheric pressure, magnetic field, vibration, radiation levels and incident EMP. In

addition, failed components will be automatically isolated without unnecessary disruption of the primary time-scale. The multiple redundancy of the ALPHA 1 will enable effective "hot-swapping to be accomplished with ease. In addition to the new passive hydrogen maser and GPS elements, the ALPHA 1 could easily incorporate GLONASS, Galileo, Rubidium, BVA crystal oscillator and LF tracking receiver technology. A unique ensembling algorithm will enable all such elements to be combined in a modular high redundancy ring or grid structure. This allows existing user clocks to be joined or removed at any time. The failure or late delivery of any element within the system will not preclude the optimum performance of the overall system.

The priorities in the performance domain are excellent short, medium and long-term stability. The overall performance will be that of the best clock or clocks, which it includes. This may vary depending on the measurement interval such that, for instance, where a Caesium and a Rubidium clock are combined, the short-term stability of the Rubidium may be enhanced with the longer-term accuracy of the Caesium. The ALPHA 1 system may have international traceability to the highest authorities when appropriately referenced, as offered by the BIPM in France and NIST in the USA. A significant advantage of this system is its strategic independence from any one country's military (e.g. GPS in the US, Glonass in Russia and Galileo in Europe).

2. PASSIVE HYDROGEN MASER

One of the aims of the Hydrogen Maser element of this project is to substantially modernise and globalise the electronics within the existing Passive Hydrogen Maser, currently being produced by IEM Kvarz, Russia. To date a significant amount of work has already been undertaken to analyse the areas where work may be necessary. Quartzlock (UK) Ltd have been working with IEM Kvarz for many years now and the expertise demonstrated by Quartzlock (UK) Ltd in developing new electronics for Rubidium oscillators, in distribution amplifiers and GPS disciplined oscillators will be put to good use in this project. One of the current worries for potential users of Passive Hydrogen Masers is the availability of components upon failure of even minor elements of the Maser, say a capacitor. Quartzlock understand this concern, and as Passive Masers become more widely accepted as a viable alternative to high performance Cesium, this issue comes to the fore. It is anticipated that this new maser will be substantially smaller sized, improving its appeal to Telecoms and other non-metrology industries.

The target specification has already been laid down, with a core 5 MHz and 1pps being the main physical output of this new maser. Unlike the current maser, a frequency converter will be incorporated, enabling the user to obtain maser quality outputs at 0.1, 1.0 and 10MHz. To date this has only been possible using an external frequency converter. In addition to these output frequencies, a 1pps GPS/Glonass synchronised timing output is envisaged. An IRIG-B module will appeal to timing users. In unsteered mode (i.e. with no GPS disciplining) an output accuracy of 1E-12 over a period of 5 years is anticipated. When GPS/Glonass disciplining is incorporated, the accuracy should improve to 5E-14, determined by the accuracy of the on-board Navstar clocks. Performance for many users is the short-medium term stability (1s to 1 day). Short-term stability specification is currently $\sim 1.5E\text{-}12/\tau^{1/2}$, and the improved specification is hoped to be $6E\text{-}13/\tau^{1/2}$ up to about 10000s. Due to the nature of the whole system, the GPS-Glonass disciplining will control the long-term stability of the system. A definite figure for this has yet to be finalized. Phase Noise at 10 kHz from the carrier will be reduced to below -155dBc at least with -160dBc the objective. In addition to the metrological improvements anticipated, it is also intended to make the controls and interfaces substantially user-friendlier. To that end RS232 control will be implemented.

Quartzlock will continue to work closely with Kvarz on fundamental advances, to simplify the use of Hydrogen Masers as stable and accurate frequency standards. The wall shift problem will be further investigated. It is hoped to make some further advances in understanding this problematic phenomenon.

3. LF TRACKING RECEIVER

Quartzlock are currently manufacturers of medium stability LF tracking receivers (formerly called off-air frequency standards), which use the 198 kHz signal from Droitwich in the UK or the 162 kHz France Inter signal to discipline either a VCXO or an OCXO. However, it has been difficult in the past to make this receiver work on other LF signals, such as DCF77 on 77.5 kHz in Germany and WWVB on 60 kHz in the US. In order to increase the export appeal of the unit and to provide a lost cost alternative to the use of GPS as a method for disciplining an oscillator, Quartzlock have started work on developing multiple input frequency instruments. This opens up a worldwide market for low cost frequency standards.

One of the problems with using the 198 kHz signal from Droitwich is that it suffers from severe phase shifting due to sky-wave propagation at night, which significantly reduces its short-term accuracy. This makes operation even a few tens of kilometres from the transmitter at night almost impossible due to ionospheric reflections. However, this was preferable to using the other LF and VLF signals because of their time code modulation, which caused discontinuities and made frequency lock difficult with simple PLL

technology. They also have periods of outage due to maintenance, which is inconvenient. Existing receiver technology suffers from indeterminate temperature coefficients, particularly where ferrites are used, which may be the prime cause of medium term frequency error. Also receivers are constrained to the use of only one frequency by the use of costly RF crystals. Analogue filtering techniques as currently employed restrict the use of long time constants causing medium-term stability degradation.

It is proposed to employ a microprocessor in the new receiver, which will enable the use of a synthesizer in the unit. This will have the advantage of removing the restriction on input carrier frequency, a significant advance. This use of a novel form of direct conversion will enable a reduction in the bandwidth of the receiver to <1 Hz, removing the need for narrow band tuned circuits and consequently improving the thermal stability of the unit. The narrow bandwidth of the receiver also improves the range of the unit and the microprocessor permits the use of significantly longer time constants. This should improve the short-term stability by several orders of magnitude, with the limiting factor being local oscillator quality. The long-term accuracy will be determined by the frequency reference employed at the originating transmitter. A Kalman Filter may be implemented at an early stage.

The design of the new unit will initially allow the use of any carrier signal whose frequency is a multiple of 100 Hz. Also the receiver architecture enables the receiver to remain unaffected by all forms of pulse modulation currently employed. This includes the A1 pulse modulation of MSF rugby. This is the factor, which has hitherto made the use of these carrier signals for oscillator frequency disciplining extremely difficult.

It is early days in the design of this unique LF tracking receiver, but potential applications include calibration and referencing of RF test equipment. The current UK market alone is anticipated to be significant (Quartzlock currently sell LF tracking receivers in the UK). A much larger potential market exists in the rest of Europe, the USA and the Far East where LF transmitters exist. The customers will be found in and amongst utilities, standards laboratories, Telecoms and RF engineering. Future papers dedicated to the explanation of the new design will be given within the next year, when preliminary test results become available. **[3]**

4. RUBIDIUM

Quartzlock are a manufacturer of Rubidium Frequency Standards. A major research element of this project is to evaluate through detailed technical and scientific analysis the following: Rubidium Isotope specification, Buffer gas specification, pressure in Rb lamp and cell, volume of Rb lamp and cell, environmental specification of lamp and cell, MTBF calculation, definition of Rb lamp and cell operation, isotope placement, inert gas filling, pressure testing, plasma evaluation and cell tuning. The overall aim of this element is to investigate the reason for Rubidium lamp failure and make progress towards extending the lamp lifetime. The aim is to increase the reliability and environmental stability of the units. This is generally measured in the Mean Time Before Failure (MTBF) and the temperature stability. Temperature stability is increasing in importance due to the placement of Rubidium in areas of extreme environmental conditions. The aim is to increase the range (ambient) over which the unit may perform to -65 deg C to +65 deg C, with a potential base plate temperature of over 85 deg C. The stability over this range should be in the range of 1E-10 to 1E-11 giving a temperature coefficient of ~ 1.... 5E-13/ deg C, a significant advance over units currently in operation. The size of the unit will also be of importance if several are to be integrated into the final PRC. An enclosure of ~220cc has been proposed with potential for 150cc version with OCXO footprint. This competes favourably with many of the miniature rubidium currently being developed and produced in the US. The rubidium oscillator will also be sold separately (as will hopefully every other element of the system).

Unfortunately the full specification for this element is currently subject to a NDA due to the sensitive nature of the work and the high level of competition that currently exists in producing rubidium oscillators.

5. MICROWAVE SATELLITE RECEIVERS

Quartzlock have been looking at integrating a Glonass/Galileo module into their current high performance GPS disciplined Rubidium Time and Frequency standard for some time now. In context of the Craft project the aim is to do several things. The purpose of the GPS/Glonass/Galileo element of the PRC will be to improve the long-term stability, removing the (small) drift element associated with Passive Hydrogen Masers (typically of the order of 1E-15 per day) cf Cesium. As part of this element an extensive improvement will be made in the software packages to algorithmically weight the satellite information. This potential re-design of the algorithms is to achieve the following goals: high level redundancy (unlimited) by modular arrangement, limitless generator connection, cellular division ability in the event of a crisis, hot swap capability with zero transients, nodal connection arrangement, inter-node communication ensuring robust operation and improved system integrity and security. Kalman filtering may be employed.

Ultimately the long-term performance (offset) of the clock must be better than 5E-14 over 5 days. This is currently the frequency offset displayed by the present engine utilized. This will help improve the overall

accuracy of the clock. Temperature controlling of the antenna and downconvertor will naturally improve the timing capabilities of the clock, essential for many users. Features like the Quad Helix antenna will continue to be used in the new design.

6. MEASUREMENT SYSTEM

Quartzlock currently produce a very high resolution Frequency and Phase Comparator measurement system. The system has an integrated distribution amplifier and rubidium oscillator. The system is limited to 1 Hz sampling rate in phase mode, which makes it impossible to look at Allan Variance for t less than 1s. This is inconvenient for many people producing Rubidium Oscillators and high quality OCXO's, where very short term stability is of interest. The very high resolution of the current system is a result of the techniques employed in the phase comparator (frequency different multiplier), which is the heart of the system. Quartzlock are currently developing a new Phase Comparator that will have several advantages. At the moment input frequencies are limited to 5 and 10 MHz, which is inconvenient for Telecoms users using 1544 kHz, 13 MHz, 2048 kHz, 800 kHz. The input range will therefore be broadened. The unit will also be a 1U system dedicated solely to measurement.

7. TIME TRANSFER MODULE

Whilst GPS one way is a very cost-effective and useful system, it cannot compare the world's best clocks with high enough accuracy. The system that is routinely used throughout the world is the method of common view, whereby users track particular satellites simultaneously and note the differences in readings between their clocks and the satellite clocks. If two or more users do the same then the satellite clock may be subtracted out of the equation and the two users have made an accurate comparison between their accurate clocks. As part of the total package, which the PRC will offer, a time transfer module has been proposed. This is to have 1ns resolution with an accuracy of at least 5ns. The module will be similar to those employed for the creation of TAI/UTC, in that the inputs will be a 1pps of the PRC and a reference 1pps from common view GPS receiver. The software employed in the TT module will enable automatic tracking of the BIPM schedule. This enables autonomous operation of the module, with automatic satellite selection and data collection. Free running atomic clocks and not those steered by GPS can only make contribution to TAI. GPS disciplined oscillators are merely replications of remote timescales (in the case of GPS this will be GPS time). Therefore using one timescale replication to contribute towards another is of no use. If the PRC wishes to follow the BIPM schedule and send readings to the BIPM for the creation of TAI/UTC, then some method must be found of taking the raw unsteered Passive Maser readings and measuring these. Due to the unique modular and nodal structure envisaged for the PRC, this should not be a difficult thing to do. This will enable more laboratories to participate. As more and more countries set up standards laboratories a cost effective way for them to set up a system will find increasing use. The cost for this system is not expected to be high as the individual elements are really only a basic GPS receiver engine, a simple time interval counter and the BIPM software.

8. SOFTWARE CLOCK

Ring/Grid Phase Alignment Combiner

In principle this should have two integrated functions

i) Phase aligning interface modules enabling the coupling of each clock to the interconnecting lines.

ii) A measurement system that enables clock drift to effect weighting in the algorithm that realizes averaging of clocks in a robust system

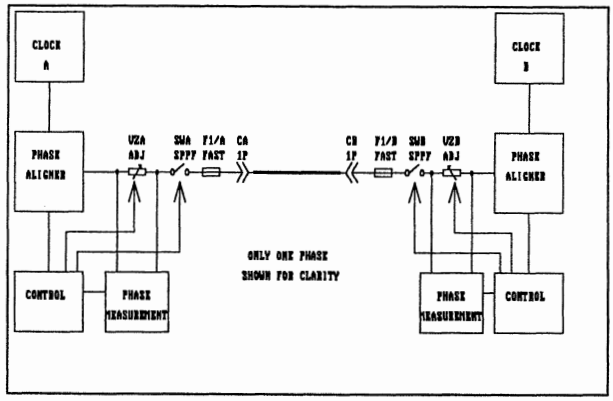


FIGURE 1: SIMPLIFIED PRC DIAGRAM

This will be the innovative element of the clock. Many of the individual elements of the clock will have been improved but the significant advance in resilience will not be possible if this element is not researched and implemented properly.

The objective of this software clock is to combine a (considerable) number of diverse frequency standards so as to result in a single time-scale, which has improved performance (the measure of performance is the short, medium and long term frequency stability of the source), and hitherto unsurpassed reliability in that the time-scale should vary by any more than a negligible amount if one or more of the frequency sources should fail. In addition, it would be desirable that the random behaviour of the output phase of a source should be detectable by comparison with the phase of the other sources.

The sources may be divided into two classes: The first is the free running oscillator type of source or primary clock, which may show a frequency offset from an internationally defined time scale. The hydrogen maser

and the rubidium oscillator are in this class. They may have excellent intrinsic stability, but can still have an unknown, although small, frequency offset.

The second class of source is the type, which provides a reference to a time scale elsewhere. The quality of the link controls the short and medium term stability of the standard. In the long term the stability of this class of source will approach that of the host time scale. **[2]**

The redundancy requirement poses several problems. The outputs of the various standards can only be combined if they are all in phase and operating at the same frequency. If sine waves are added without phase alignment, the output sine wave may have zero amplitude with some combinations of input phase. If we consider a simple system where one source is considered to be the master, then the other sources may be phase aligned to the master, and all the sources may be added together with equal amplitude weighting. If one source should now fail, the amplitude of the output will only drop by the ratio of the number of sources to be combined. This simple system would work quite well if all the sources were of the same class and the same degree of frequency (phase) stability. If we were combining three rubidium oscillators, we would want to give them equal amplitude weighting, as the expected short-term stability's are similar. One will even derive performance benefit as the phase noise (instability) of each will not be correlated, and thus the output noise of the ensemble will be lower than that of the individuals.
If we are combining sources of quite different performance levels, the output should be substantially that of the best standard, for example a hydrogen maser (short to medium term). Any addition of output from the other standards can only degrade the performance.

One could argue that a good method would be to improve the short-term stability of the lesser sources by phase comparison with the hydrogen maser, and then phase adjustment (of these standards). This is certainly true, and one could then use a higher percentage of the improved (phase adjusted) standard in the final output. However, if the hydrogen maser did fail, then the reference source for the phase adjustment would suddenly vanish, and the basic (inferior) noise characteristics of the lesser sources would suddenly reappear. The next best source, for example the rubidium oscillator, could then be designated the master and used for phase alignment. However, it would be difficult to achieve a smooth change over.

A better solution might be to derive a notional time scale based on phase measurement of all the sources and the to phase-align all the sources to the notional time scale. In this way, the need for a master is overcome. Careful consideration of the weighting of each source contribution to the notional time scale would be required, with a different contribution at different averaging times. To clarify this, a source of the second class would have more weight over a long period of time as it is directly referenced to an international time scale.

A major part of this project is in the computer algorithms used to derive the notional time scale, and to control the phase aligners that condition the outputs of each standard before they are added together. It is also important that the phase aligners themselves do not contribute noise and phase drift. The successful implementation of the PRC will combine state of the art electronic design, software analysis and control algorithms **[2]**

9. CONCLUSION

The ALPHA 1 is intended to improve reliability for time standards of all kinds, whether for standard laboratory calibration, telecom network synchronisation at the Stratum 1 level or national standards laboratories. Whilst the advantages of this project are deemed to be high, the risks associated with it are very low. The ALPHA 1 will be a user-friendly system, with the modular structure enabling customer choice and cost flexibility, ease of service and the ability to upgrade when required. The extreme reliability will be derived by virtue of multiple redundancies with all processing being accomplished in a distributed fashion. No single module or component will be vital for the continued operation of the system wherever three or more clocks are used and with greater numbers, it is even possible to imagine that the system could be separated into two or more totally independent systems each of which could continue in operation independently.

10.REFERENCES

[1] M.Rainer, 'A high reliability master reference clock', Quartzlock Internal Application Note, March 2000

[2] C.R.Green et al, Documents supplied to European Commission in support of CRAFT proposal, Feb 1999

[3] M. Rainer, A new LF Tracking Receiver. Quartzlock Internal Application Note, March 2000

A PICOSECOND TIMING SYSTEM FOR THE *LASER MEGAJOULE*

P.LECLERC, JY.SALMON, JP.ARNOUL (CEA)
Commissariat à l'Energie Atomique
CEA-DIF
BP 12
91680 BRUYERES LE CHATEL
FRANCE
Tel. : +33.1.69.26.47.71 - Fax. : +33.1.69.26.70.43
e-mail : leclerc@bruyeres.cea.fr

Abstract

The *Laser MégaJoule* is a 240 beam laser facility that will be built in France to study plasma and to achieve inertial confinement fusion. For proper operation, it needs a timing system capable of distributing 8,000 trigger signals over an area exceeding 30,000 m², with an accuracy ranging from 1 µs down to 5 ps. We developed a system based on very accurate delay generators connected to a master clock through a fiber optic time distribution network to face this requirement. We mainly focused on the study of the distributing network and the delay generators. An important effort was also made to develop time measurement methods in the near-picosecond range. A first prototype system was built by the CEA, then an industrialized system with enhanced performances was realized for the LIL facility (a 8 beam prototype of the LMJ). It can deliver up to 1,000 trigger signals. We are now developing new picosecond measurement methods and keep working to expand the capability of the timing system to reach the LMJ requirements.

1. Introduction

The *Laser MégaJoule* is a 240 beam laser facility that will be built in France to study plasma and to achieve inertial confinement fusion. It will be operating in 2008. In this article, we describe the principle of its Timing System that must achieve picosecond range accuracy. We present the first studies, the prototype developed by the CEA (*French Atomic Agency*), the first system for a 8 beam laser, and the still necessary effort to reach a near picosecond system.

2. The Laser MégaJoule project

The *Laser MégaJoule* will concentrate 1.8 MJ of UV light on a small ball, filled with hydrogen isotopes, that will be compressed and heated up to the ignition point. The optical power will be provided by 240 laser beams that must be precisely focused, power-balanced and synchronized. Each laser pulse will have a particular temporal response to generate specific effects such as mechanical compression, plasma heating, and ignition. To control the proper operation of the facility, several kinds of measurement and control systems have to be developed. The timing system is a critical one.

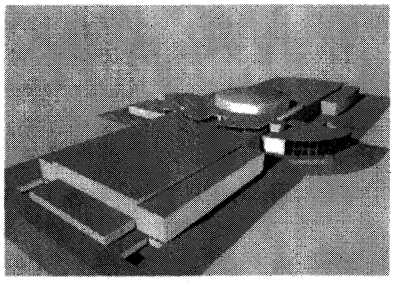


Figure 1 : The LMJ building 30,000 m² ground surface

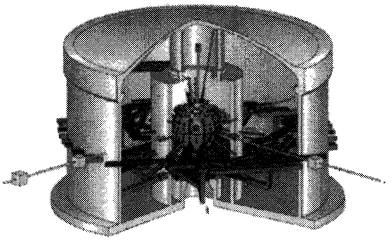


Figure 2 : The target chamber (at center) : 10 m diameter

3. Requirements of the Timing System

To guaranty its proper operation, the facility needs a timing system capable of delivering 8,000 trigger signals in single shot mode over an area exceeding 30,000 m², with a dynamic range of up to 2 s, an accuracy ranging from 1 µs down to 5 ps and a jitter ranging from 100 ns down to 5 ps (rms).

Fortunately an analysis shows that the needs for high accuracy and low jitter triggers are close to the time the light impacts the target ($T0_{target}$), while devices needing soon or late trigger are less sensitive. Therefore, to reduce the cost of delay generators, the 8,000 channels of the timing system were divided in several classes, each of them having a different accuracy and delay range.

Class	Dynamic range	Resolution	Rms jitter	Accuracy
Ultra-precise	+/- 10 µs	< 5 ps	< 5 ps	5 ps
Precise	+/- 50 µs	< 15 ps	< 15 ps	15 ps
Fast	+/- 100 ms	< 1 ns	< 100 ps	1 ns + 5 E-7 x delay
Slow	+/- 1 s	< 100 ns	< 100 ns	100 ns + 5 E-7 x delay

Table 1 : class specifications

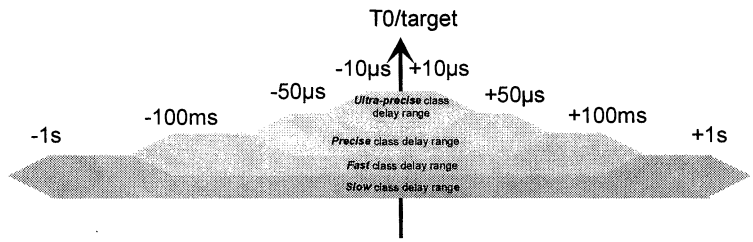


Figure 3 : class delay ranges

4. The Design of a Distributed Timing System

To face our specific need, we had to develop a system with the following capabilities :
- a highly stable reference clock
- a time distribution network
- many high quality low cost delay generators
- a self-calibration tool

We conceived a system based on very accurate delay generators connected to a master clock through a bi-directional fiber optic time distribution network

4.1 The Master Clock

The master clock is unique. It is the only time reference for the facility. It runs a very stable OCXO that is calibrated with an exactitude better than 10^{-8} by comparison to a GPS clock. It generates an optical numerical message whose modulation frequency is connected to its internal quartz, while the data stream is used to send trigger information to delay generators.

4.2 The Optical Distribution Network

A distribution network had to be realized to distribute the reference clock and trigger information through the entire LMJ building. We chose to use a fiber optic network. A careful technical study and economic considerations led us to select a SDH-SONET type link. With a proper choice of components, this permits the use of 155.52 MHz standard components from telecommunications.

4.3 Slave Delay Generators

Slave units are connected to the optical network. They locally recover the reference 155.52 MHz clock and trigger information from the data stream.

Each slave device also includes several independent delay generators that drive electrical or optical trigger outputs. Their design was probably the most difficult point we encountered. They are based on real time counters that count each period of the 155.52 MHz reference clock, linked to analog time interpolators to generate arbitrary delays with a picosecond resolution. A very careful study had to be conducted to avoid any problem such as diaphony, non linearity, noise effects, etc.

4.4 A Bi-Directional Link

Fiber optics are easy to use and well fitted to our need, however their propagation delay time varies with temperature changes (from 40 ps/°C/km to 170 ps/°C/km for the fiber samples we characterized). These variations were not acceptable for our timing system. Therefore, we designed a bi-directional link using two slightly different wavelengths that are separated with WDMs. An emitter was added to some slave units, while an optical converter and a time interval counter were added to the master unit, so that it is possible to monitor propagation delay drifts in the fiber optics. The same functionality, coupled to a specific installation procedure also gives us the exact absolute transit time of each fiber. In order to reduce the cost of the system, only long fibers are monitored with a bi-directional link. Inside a local area of about 10 m, time offsets between slave systems can be considered as constants, and do not require to be monitored. So only one slave system per local area needs to drive a bi-directional link. We named it the reference slave unit.
The design of this bi-directional link was the second most critical point of the study.

4.5 Overview

Figure 4 summarizes the architecture we described.

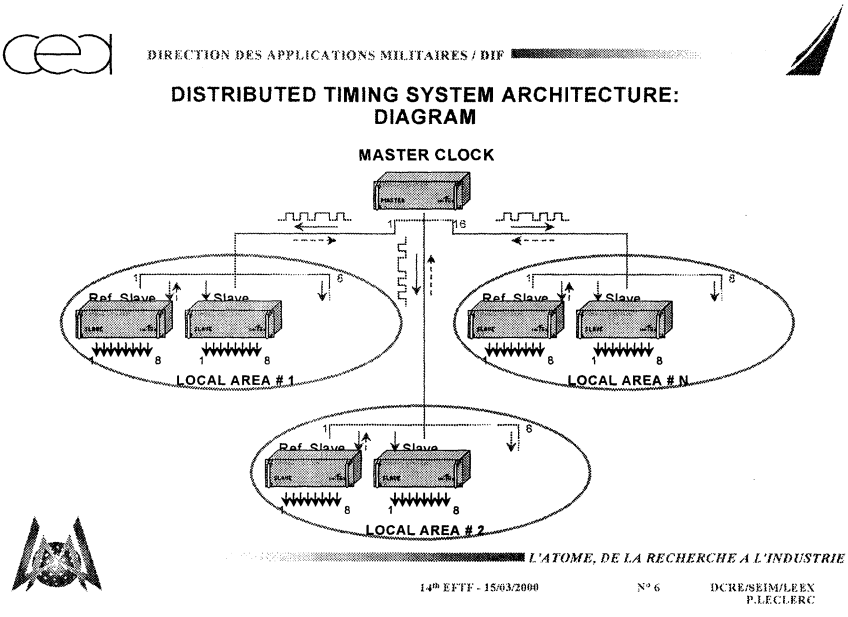


Figure 4 : The timing system distributed architecture

5. First implementation

A first timing system was built by the CEA. It is based on one master unit and 10 slave units, but it does not include the bi-directional link.
Its performances are :

- trigger channels : 80
- jitter < 50 ps rms (35 ps typ.)
- accuracy : 100 ps +5 E-8 x delay

This system was achieved in 1999. It was designed to be used in the early phases of exploitation of a 8 laser beam prototype of the LMJ currently under construction, the L.I.L. (*Ligne d'Intégration Laser*). It will be operated until the end of the year 2000.

6. Implementation for the L.I.L.

The second step was to industrialize the prototype system and to enhance its performances for the future of the LIL facility. New components were used for delay generators and bi-directional self compensating optical links were implemented. The new system reaches the following performances :

- trigger channels : 800
- jitter < 15 ps rms
- accuracy :<15 ps +5 E-7 x delay

This new system is scalable and can deliver more than 1,000 electrical or optical triggers. It is currently being built and its performances have been validated. It will be deployed in the LIL facility during years 2000 to 2002.

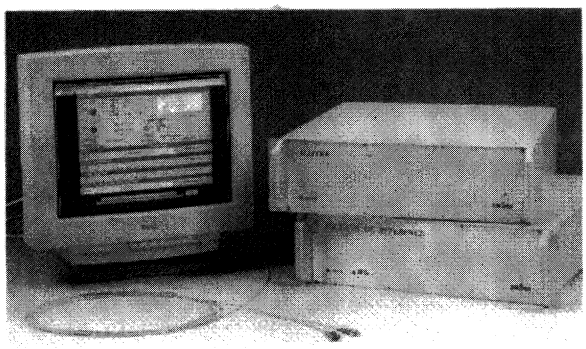


Figure 5 : Picture of some LIL timing system components (the master unit and a slave unit)

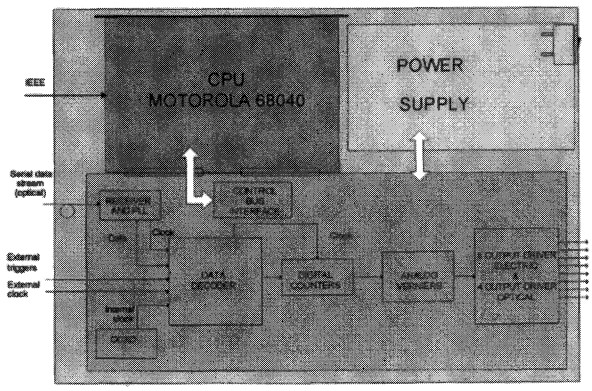


Figure 6 : Synoptic of a slave delay generator unit

7. Toward a Picosecond Timing System

The LMJ timing system requires a few more studies to be achieved.

- the average cost per channel must be decreased
- high end delay generators and optical links with 5 ps accuracy and jitter must be studied. They are at the limit of current technologies, or must rely on new concepts
- new qualification tools for this timing system have to be developed as they must overpass capabilities of current instrumentation. Our laboratory is working on time measurement in order to achieve 1 ps accuracy over a 10 µs range, with new methods, by combining different kinds of existing instruments, looking for new ones, ... or by developing some new products such as picosecond streak cameras for our own needs.

These efforts must allow us to qualify the LMJ timing system within 3 years.

8. Conclusion

The picosecond timing system of the Laser MégaJoule still requires important studies, both from the CEA to enhance its capacities of time measurement in the picosecond range, and from the company in charge of the development of the system to reduce costs of low-end delay generators while enhancing performances of high-end ones.

However the important study effort that has already been made, led to the realization of distributed timing systems and delay generators whose performances overpass those of other products currently available on the market, for lower prices. They can be of interest for other physics facilities.

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

UNCERTAINTY AND PREDICTION OF CLOCK ERRORS IN SPACE AND GROUND APPLICATIONS

Patrizia Tavella
Istituto Elettrotecnico Nazionale G. Ferraris
Strada delle Cacce 91, 10135 Torino, Italy
Tel. +39-11-3919235, Fax +39-11-3487046, tavella@tf.ien.it

Monica Gotta
Politecnico di Torino, Dip. Elettronica
C.so Duca degli Abruzzi 24, 10129 Torino, Italy
Tel. +39-11-3919238, Fax +39-11-3487046, tesimat@tf.ien.it

ABSTRACT

Random fluctuations on the signal of atomic clocks onboard satellites in modern radionavigation systems represent one of the causes of positioning errors. This paper examines the problem of prediction of clock errors due to random fluctuations of time signal. By using optimum prediction theory, theoretical results are compared with simulations and experimental measurements of IEN oscillators. As an example, the errors in clock prediction are evaluated in the case of the currently studied Galileo GNSS.

INTRODUCTION

Clock error onboard satellites impacts navigation error as, once a clock is synchronized at a certain istant t=0, it is left stand alone for a certain period of time and it accumulates an inevitable error $x(t)$. For example a satellite could be out of view of the Earth control stations for some hours: in that period the clock onboard is used to generate the navigation message and the clock error (unknown) influences the error in the localisation of the user. According to the accuracy goal of the navigation system, the clock error should be lower than a certain threshold and is therefore important to evaluate the possible time deviation $x(t)$ according to the type of clock used.
In case of frequency standards it is known that, apart from "deterministic" variation, the leading behaviour in the performance of a clock can be ascribed to five different stochastic noises. The last are distinguished by the slopes of the single sided power spectral densities (PDSs) of frequency fluctuations $y(t)$ as functions of Fourier frequency in a log-log plot. In the standard power-law model, the PDS is assumed to be proportional to f^{α}. The five different noises are white phase, flicker phase, white frequency, flicker frequency and random walk frequency, corresponding to integer values of α ranging from 2 to –2 respectively. The standard power-law model is written as [1]:

$$S_y(f) = \sum_{\alpha=-2}^{2} h_\alpha f^\alpha$$

where h_α is the intensity coefficient of the noise. Furthermore, as phase deviation $x(t)$ is the quantity we are interested now, we can translate the power-law model to a form that works with $x(t)$. It is known that PSD of $x(t)$ is related to the PSD of $y(t)$ by:

$$S_x(f) = \frac{S_y(f)}{(2\pi f)^2}$$

We will use then new parameters β and g_β defined by:

$$S_x(f) = \frac{h_\alpha}{(2\pi)^2} f^{\alpha-2} = g_\beta f^\beta \qquad (1)$$

with $\beta=\alpha-2$ ranging from 0 to-4.

In the followig, two approaches are considered dealing with the problem of prediction: stochastic theory and prediction theory. The last approach was already partially developed by Bernier in [12] and by Barnes et Allan in [11]; semiempirical formulas were derived by Kartaschoff in [13], together with simulation results.

1. CLOCK ERROR MODEL

The problem of determining the probable time departure $x(t)$ of an oscillator at a certain time t after synchronization at initial time t=0, is related to the statistical model of noises. Infact the best prediction for $x(t)$ is the expected value of the process: $E\{x(t)\}$ while the prediction uncertainty is estimated by the standard deviation of the probability density describing the process at time t. For this reason, we examine the statistichal properties of the five noises categories that can be used in clock error prediction.

In the case of *white phase noise* (WPM, α=2, β=0), the process $x(t)$ is characterized by a spectral density of phase fluctuations which is indipendent of frequency, nominally [2]:

$$S_{x_{+2}}(f) = q_w \qquad (2)$$

As real systems are characterized by a frequency cut-off f_c (related for example to measurement system bandwidth), the variance of measured white phase noise is not infinite and, since $E\{x(t)\}=0$, it is equal to the mean square value:

$$\sigma_{x_{+2}}^2(t) = E\{x^2(t)\} = \int_0^{+\infty} S_x(f)df = q_w f_c \quad (3)$$

The variance (3) is related to the Allan variance ad to the intensity coefficient h_{+2}, by the following relashionship [3]:

$$\sigma_y^2(\tau) = \frac{3\sigma_x^2(\tau)}{\tau^2}$$

so that:

$$\sigma_{x_{+2}}^2(t) = \frac{h_{+2}}{4\pi^2} f_c \quad (4)$$

by considering a rectangular filter [4] in (3).

Therefore $q_w = \frac{h_{+2}}{(2\pi)^2}$.

The best prediction of time departure x(·) at time t in presence of white noise is $E\{x(t)\}=0$; the uncertainty of this prediction is equal to the square of classical variance, i.e. $u_x(t) = \sqrt{q_w f_c}$, indipendent of the time elapsed after synchronization.

In the case of *white frequency* (WFM, $\alpha=0$, $\beta=-2$), *x(t)* becomes a random walk and, from the theory of Wiener process, the probability density of *x(t)* is a gaussian $N(0, q_1 t)$ [2]:

$$f_x(x;t) = \frac{1}{\sqrt{2\pi q_1 t}} e^{-x^2/2q_1 t}$$

where q_1 is the diffusion parameter measured for example in $[ns^2/s]$ if frequency fluctuations are measured in nanoseconds over seconds. The variance of *x(t)* is therefore $\sigma_{x_0}^2(t) = q_1 t$ and it can be related to h_0 by means of [5]:

$$\sigma_y^2(\tau) = q_1 \tau^{-1} = \frac{h_0}{2}\tau^{-1}$$

as:

$$\sigma_{x_0}^2(t) = \frac{h_0}{2} t \quad (5)$$

The best prediction of x(t) in presence of WFM is $E\{x(t)\}=0$; x(t) on the average remains at the synchronization error $u_x(t) = \sqrt{q_1 t}$, increasing with the square root of time.

In the case of *random walk frequency* (RWFM, $\alpha=-2$, $\beta=-4$), *x(t)* is interpreted as an integrated random walk, and its probability density is a gaussian $N(0, q_2 \frac{t^3}{3})$ [6]:

$$f_x(x;t) = \frac{1}{\sqrt{\frac{2\pi}{3} q_2 t^3}} e^{-3x^2/2q_2 t^3}$$

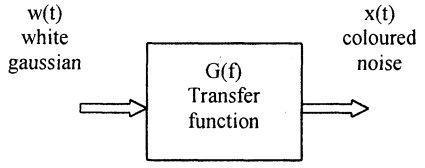
where q_2 is the diffusion parameter of the frequency random walk ($[ns^2/s^3]$ for example). From the relation of Allan variance with q_2 [5], we can derive:

$$\sigma_{x_{-2}}^2(t) = q_2 \frac{t^3}{3} = \frac{2\pi^2}{3} h_{-2} t^3 \quad (6)$$

The best prediction of time departure in presence of RWFM is the synchronization error $E\{x(t)\}=0$; the prediction uncertainty is $u_x(t) = \sqrt{\frac{2\pi^2}{3} h_{-2} t^3}$, increasing with 3/2 power of time.

2. COLOURED NOISE AS FILTERED WHITE NOISE

The two remainig cases, flicker phase and flicker frequency, cannot be theoretically modeled. To study the behaviour of time signal affected by these kind of noises, we applied the theory of filtered white noise [7]. If we model discrete coloured noise as:



The PDS of *w(t)* is equal to $S_w(f) = q_w$. We'd like to obtain *x(t)* such that $S_x(f) = S_w(f)|G(f)|^2 = g_\beta f^\beta$ with the correct slope β of eq. (1), corrisponding to the desidered output noise. The transfer function of the filter should then be of the form:

$$G(f) = f^{\beta/2}$$

whose corrispondent impulse response is [8]:

$$g(t) = \frac{u(t)}{\Gamma(-\beta/2)} t^{-(\beta/2)-1} \qquad |\beta| \geq 1 \quad (7)$$

where *u(t)* is unit step and Γ is Gamma function.

From optimum prediction theory [7], if $e(t) = x(t) - \hat{x}(t)$ is the prediction error, defined as the difference between the real value and predicted one, we can find [9] $E\{e(t)\} = 0$ and, since estimated coloured noise $\hat{x}(t)$ is represented by a linear combination of white gaussian samples with zero average, it can be stated:

$$E\{x(t)\} = 0 \tag{8}$$

The mean square prediction error is found to be [8]:

$$E\{e^2(t)\} = \sigma_x^2(t) = \frac{q_w}{2}\int_0^t |g(\tau)|^2 \, d\tau \tag{9}$$

By substituting (7) in (9) and considering [10]:

$$q_w = \frac{h_\alpha}{(2\pi)^\alpha} \tag{10}$$

we can derive the variance of the prediction for each of the five noises.

In particular, for *flicker phase noise* (FPM, α=+1, β=-1), function to be integrated presents a singularity in t=0 but we can write (9) as:

$$\sigma_x^2(t) = \frac{q_w}{2}\int_\varepsilon^t \frac{1}{\Gamma^2(1/2)}\tau^{-1}d\tau = \frac{h_{+1}}{4\pi^2} ln\left(\frac{t}{\varepsilon}\right)$$

and with the aid of simulation, the value of the constant ε is fixed giving:

$$\sigma_{x_{+1}}^2(t) = \frac{h_{+1}}{4\pi^2} ln(2\pi f_c t) \tag{11}$$

α	**Prediction uncertainty $u_x(t)$**
+2 white phase	$\sqrt{\frac{h_{+2}}{8\pi} f_c}$ *lorentian filter* * $\sqrt{\frac{h_{+2}}{4\pi^2} f_c}$ *rectangular filter*
+1 flicker phase	$\sqrt{\frac{h_{+1}}{4\pi^2} ln(2\pi f_c t)}$
0 white frequency	$\sqrt{\frac{h_0}{2} t}$
-1 flicker frequency	$\sqrt{2h_{-1}t^2}$
-2 random walk freq.	$\sqrt{\frac{2\pi^2}{3} h_{-2}t^3}$

Table 1: prediction uncertainties

* If in (3) a Lorentian filter is assumed, $u_x(t) = \sqrt{h_{+2} f_c / 8\pi}$

which presents the same time dependence found in [1] and[11]. *We can conclude that best prediction of time departure x(t) in presence of FPM is the synchronization error E{x(t)}=0, with uncertainty* $u_x(t) = \sqrt{\frac{h_{+1}}{4\pi^2} ln(2\pi f_c t)}$ *, increasing with the square root of logarithm of t.*

In the case of flicker frequency (FFM , α*=-1,* β*=+1) substituting (7) and (10) in (9), we can obtain that best prediction is the synchronization value with uncertainty* $u_x(t) = \sqrt{2h_{-1}t^2}$ *, increasing linearly with time.*

In the other three cases, from this prediction theory we would obtain the same results found in part 1 by means of stochastic theory. Uncertainty prediction formulas for all cases are resumed in Table 1.

3. SIMULATIONS AND EXPERIMENTAL RESULTS

To test the theoretical results, we investigated the evolution of $x(t)$ by means of simulations. We used Kasdin's algorithm [10] to generate clocks with sampling time $\tau_0 = 100s$ and we examined the statistical properties (data mean and variance) of $x(t)$ at different instants ($t = 10\text{^}4s, 10\text{^}5s, 10\text{^}6s$). At time t, each clock has accumulated a time departure $x_i(t)$. The distribution of $x_i(t)$ should be normal; the mean value should be $E\{x_i(t)\}$=0 and variance should follow the theoretical formulas of Table 1. To verify that the distributions are normal, comparison between experimental and theoretical distributions was carried out by means of χ^2 test; an example is illustrated in Fig.1. The simulated distributions were all found in agreement with the theory. Fig.2 reports five of 1000 simulated clocks, in the cases of white FM, with the theoretical gaussian distribution, at different instants,

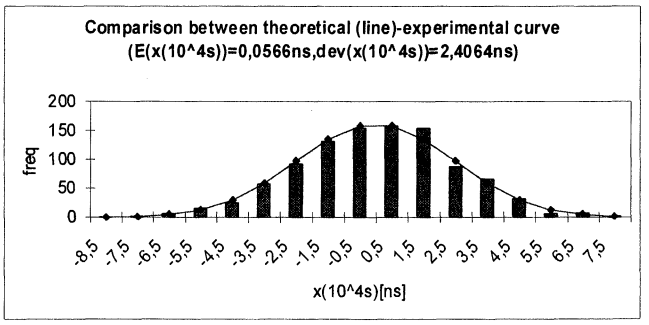


Fig.1: experimental distribution of simulated values of x(10^4s) (1000 data) compared with theoretical gaussian distribution, in the case of flicker FM

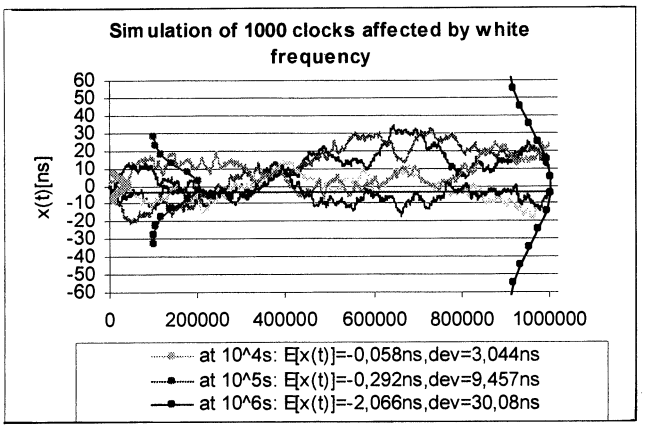


Fig 2: distribution of phase deviation in the presence of white phase frequency ($h_0 = 3{,}2E-20$)

superimposed. From the plot a good agreement is visible and in Table 2 the experimental standard deviations of simulated data are compared with the theoretical formulas of Table 1.It can be seen that the behaviour observed over 1000 simulated clocks verify to a large extent the developed theory.

Real measurement data were are also used to support theoretical prediction. Deviations *x(t)* were measured with sampling interval $\tau_0 = 100s$, from Cs HP 5071A (characterized by WFM, $h_0 = 3{,}8E-22$) and Rb HP 5065A (characterized by FFM, $h_{-1} = 5{,}28E-25$). Then the measures were analized to obtain mean and variance of the distribution of *x(t)* at different instants. Fig.3 and Fig.4 presents distribution of *x(t)* at instants *t=10^5s* and *10^6s* for the two oscillators. In Table 3 we compare theoretical results with experimental (measured) ones.

α	dev(t=10^4s)	dev(t=10^5s)	dev(t=10^6s)
0	Exper. 3044,09ps Theor. 3000ps	Exper. 9457,13ps Theor. 9486ps	Exper. 30089ps Theor. 30000ps
-1	Exper. 2406,48ps Theor. 2402,2ps	Exper. 25308ps Theor. 24022ps	Exper. 243161ps Theor. 240220ps
-2	Exper. 52,12ps Theor. 49,99ps	Exper. 1565ps Theor. 1581ps	Exper. 50533ps Theor. 50000ps

Table 2: comparison between theoretical standard deviation of x(t) and simulated ones

	dev(t=10^5s)	dev(t=10^6s)
Cs	Exper. 4,52 ns Theor. 4,37 ns	Exper. 20,06 ns Theor. 13,84 ns
Rb	Exper. 137 ns Theor. 102 ns	Exper. 1061 ns Theor. 1028 ns

Table 3: comparison between theoretical and measured standard deviations

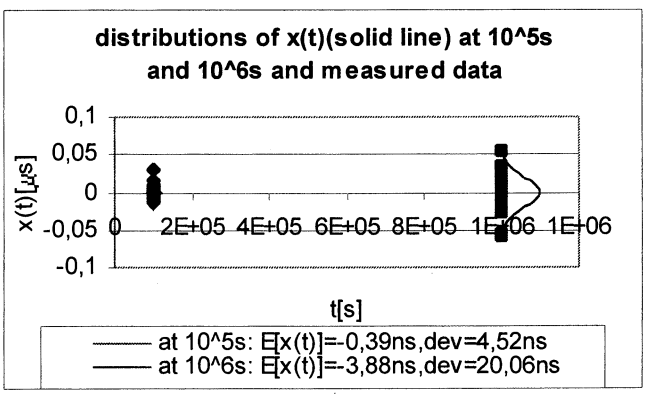


Fig.3: real measurements on Cs HP 5071 A

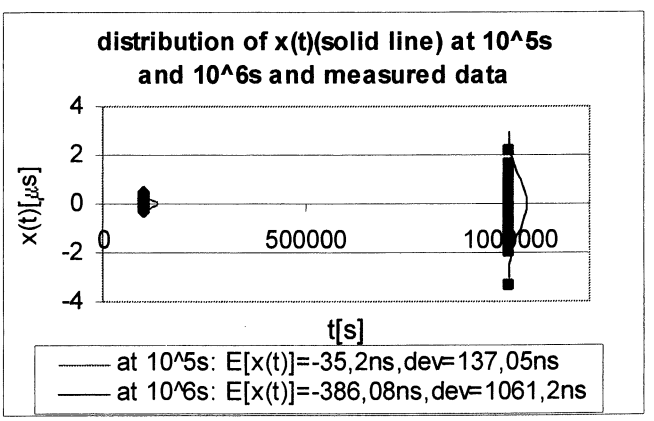


Fig.4: real measurements on Rb HP 5065 A

4. GNSS APPLICATION

We can apply the developed theory to evaluate errors in the case of clocks which have been proposed for Galileo Space Segment. For example,

NOISE	$\sigma_y(\tau)$	h_α
WFM	$\sigma_y(\tau) = 5E-12 \cdot \tau^{-1/2}$	$h_0 = 5E-23$
FFM	$\sigma_y(\tau) = 5E-14 \cdot \tau^{0}$	$h_{-1} = 1{,}8E-27$
RWFM	$\sigma_y(\tau) = 3E-16 \cdot \tau^{+1/2}$	$h_{-2} = 1{,}76E-32$

Table 4: stability of the space qualified Rubidium

technical characteristics in terms of stability are resumed in Table 4 for a space qualified Rubidium by Temex (Neuchatel). Appling formulas in Table 1 corrispondent to each kind of noise, we can obtain the resulting total uncertainty as:

$$u_{tot}(t) = \sqrt{\frac{5E-23}{2} \cdot t + 2 \cdot 1{,}8E-27 \cdot t^2 + \frac{2\pi^2}{3} 1{,}76E-32 \cdot t^3}$$

which is illustrated in Fig. 5.
We can conclude that, with an error treshold of 1,5ns, the synchronization interval should not be longer than 5 hours.

5. CONCLUSIONS

In this work we have theoretically derived prediction formulas from statistical theory of stochastic processes or from the prediction theory of filtered white noise. We obtained the same theoretical results from both approaches. We compared theoretical prediction with experimental behaviour of a large number of simulated clocks, and with real measurements on IEN oscillators. We can conclude that prediction formulas reflect satisfactorily real clocks behaviour and therefore, from the knowledge of the intensity level of noise there is a mean to predict *x(t)* at a certain time after synchronization and to evaluate the uncertainty of this prediction.

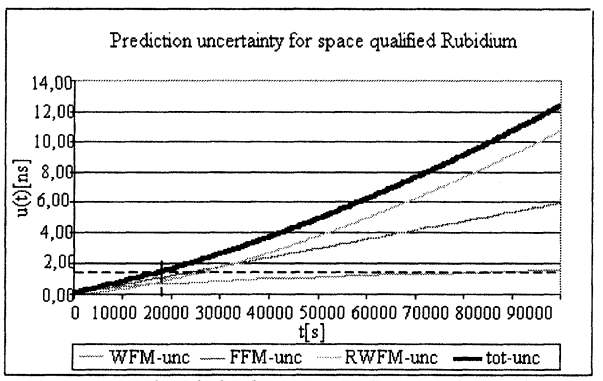


Fig.5: example of clock error evaluation for Galileo

ACKNOWLEDGEMENTS

The authors are thankful to Dr. Kasdin of Princeton University and to Ing. Godone of IEN for their kind support during the preparation of this work.

REFERENCES

[1] D.W.Allan, "Time and Frequency (Time-Domain) Characterization, Estimation and Prediction of Precise Clocks and Oscillators", IEEE Trans. Ultras. Ferroel. Freq. Control, vol UFFC-34, no.6, November 1987, pp.647-654

[2] L. Lo Presti, F.Neri, Introduzione ai Processi Casuali, CLUT Editrice, 1992, pp.124-125, 50-55 (in Italian)

[3] CCIR Recommendations, ITU 1992, p.35, Rec.538-2

[4] C. Audoin, *"Frequency Metrology"*, in Metrology and Fundamental Constants, International School of Phisics Enrico Fermi, North Holland Publishing Company, 1980, pp.169-191

[5] J.W. Chaffee, "Relating the Allan Variance to the Diffusion Coefficients of a Linear Stochastic Differential Equation Model for Precision Oscillators", IEEE Trans. Ultras. Ferroel. Freq. Control, vol. UFF C-34, no.6, pp.655-658, November 1987

[6] S. Karlin, H. Taylor, A First Course in Stochastic Processes, Academic Press, San Diego, CA, 1975, pp.384.

[7] W. B. Davenport, W. L. Root, An introduction to the Theory of Random Signals and Noise, Mc Graw Hill, 1958, pp.219-238

[8] G. A. Campbell, M. Foster, Fourier Integrals, Van Nostrand, New York, 1948, p.50, pair 521

[9] Sergio Bittanti, Teoria della Predizione e del Filtraggio, Pitagora Editrice Bologna, 1990, pp.116-149, (in Italian)

[10] N.J. Kasdin, T. Walter, "Discrete Simulation of Power Law Noise", IEEE Frequency Control Symposium,1992, pp.274-283

[11] J.A. Barnes, D.W. Allan, "An Approach to the Prediction of Coordinated Universal Time", IEEE Frequency Control Symposium, November 1967

[12] L.G. Bernier, "Linear Prediction of the Non-Stationary Clock Error Function", 2nd European Frequency and Time Forum Neuchatel, 1988

[13] P. Kartaschoff, "Computer Simulation of the Conventional Clock Model", IEEE Trans. Instr. Meas., vol. IM-28, no.3, pp.193-197, September 1979

CELLULAR SYNCHRONIZATION NETWORKS FOR TELECOM APPLICATIONS BASED ON THE GPS AND ON SDH/SONET NETWORKS

Dominik Schneuwly
schneuwly@oscilloquartz.com

Oscilloquartz SA, CH-2002 Neuchâtel, Switzerland

ABSTRACT

Synchronization is usually distributed throughout a telecommunications network using the traffic links of an SDH or a SONET transport network. The alternative synchronization distribution method presented here is based on the Global Positioning System (GPS) for coarse distribution to a number of subnetworks called cells, and on the SDH or the SONET transport network for fine distribution to all nodes within the subnetworks. The internal structure of the cells and the way the cells are interconnected provide for protection against node and link failures, and also against local corruption of the GPS radio signal.

1. INTRODUCTION

The most widely used method for distributing synchronization in telecommunications networks uses SDH (Synchronous Digital Hierarchy) or SONET (Synchronous Optical Network) transport networks as the synchronization carrier. Traditionally a synchronization reference signal is generated by a central Primary Reference Clock (PRC) with high frequency accuracy. The synchronization signal is then distributed to all nodes of the network via a tree of slave clocks connected by SDH/SONET traffic links. This paper presents an alternative method based on the Global Positioning System (GPS) for coarse distribution to a number of subnetworks, and on the SDH/SONET network for fine distribution to all nodes within the subnetworks. The method consists in partitioning the network at hand into subnetworks of predefined internal structure, and to interconnect the subnetworks according to well-defined rules. These rules determine the inter-subnetwork or 'external' structure. The obtained subnetworks are called 'cells', the resulting synchronization network is called a 'cellular synchronization network'. The chosen internal and external structures result in some interesting properties, which will be discussed in detail.

2. CLASSICAL SYNCHRONIZATION NETWORKS

As stated above, synchronization is usually distributed throughout a telecommunications network using the traffic links of an SDH or a SONET transport network. Often the frequency reference is generated by a Primary Reference Clock installed in a centrally located node. The PRC usually consists of a redundant set of atomic Cesium clocks and/or GPS-receivers. The SDH/SONET network provides functions for the distribution of synchronization from the PRC to all nodes of the network via a tree of slave clocks interconnected by SDH/SONET traffic links. There is also a simple messaging protocol for communicating a reference source quality information from clock to clock throughout the synchronization distribution tree. This protocol is called Synchronization Status Message (SSM). SSM is also useful for automatic protection switching in case of clock or synchronization link failures.

Let us consider a linear chain of SDH/SONET nodes with their slave clocks. Synchronization reference signals are fed to the chain in at least two nodes, typically at the two ends of the chain. The direction in which synchronization is distributed inside the chain depends on priority table settings in the SDH/SONET equipment. Let us assume that under normal conditions, synchronization is distributed from the west end node, which is also a synchronization feeding point, to the east. In case of a link failure, all nodes east of the failed link are cut off in terms of synchronization. Within the cut off half of the chain, the synchronization flow will automatically change direction. After the automatic reconfiguration, the synchronization will flow from the east end node, which is the second synchronization feeding point, to all nodes east of the failed link. This automatic protection switching makes use of the SSM and of the priority-based reference selection function.

These classical synchronization distribution techniques are illustrated schematically in figure 1. The upper part of the figure shows a portion of the network which is synchronized by a tree of master-slave clocks. The lower part of the figure shows one instance of a

linear chain of clocks which uses the SSM-controlled automatic reconfiguration mechanism. The numbers next to the arrows (= synchronization links) indicate the priorities assigned to the synchronization reference inputs. The dashed arrows inside the cirlces (= clocks) show the synchronization flow under normal operating

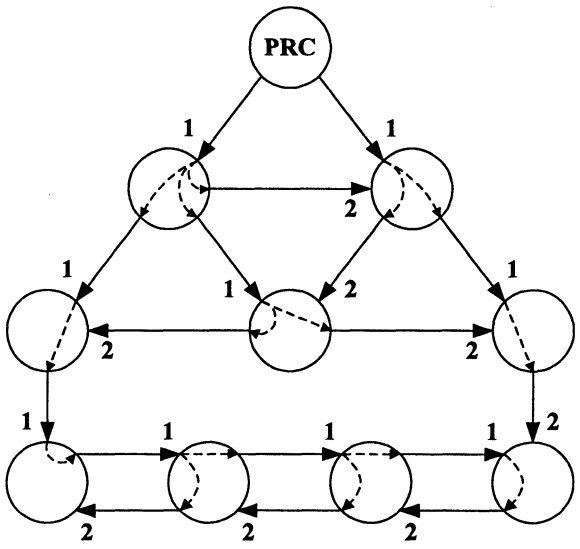


Figure 1: Classical synchronization network

conditions: in absence of failures each clock selects the priority 1 reference input. For detailed information on these synchronization techniques refer to [1] to [5].

3. CELLULAR SYNCHRONIZATION NETWORKS

3.1. General approach

The cellular synchronization network concept was developed with the intention of synchronizing telecommunications network to the GPS. Its architecture was designed with the following criteria in mind:

- The GPS is the sole primary reference source (no atomic Cesium clocks)
- The synchronization distribution to each node is protected (redundancy) against link failures, clock failures, GPS-receiver failures, and against local corruption of the GPS radio signal
- The protection mechanism is automatic (no human intervention required)

These criteria reflect a concern about the possibility of failing GPS reception due to radio interference or intentional jamming. Because of the stringent availability requirements in telecommunications networks, a synchronization distribution network must provide protection against failure modes with non-negligible probablility of occurrence.

The chosen architecture consists of an aggregation of synchronization subnetworks called 'cells'. There exists a limited number of standard cell structures called cell patterns. The cells are interconnected to form a synchronization distribution network that covers the entire SDH/SONET transport network. The way cells are interconnected follows a set of simple rules.

3.2 The cell's internal structure

There are two standard cell structures called pattern A and pattern B. Figure 2 shows the typical internal structure of a pattern A cell.

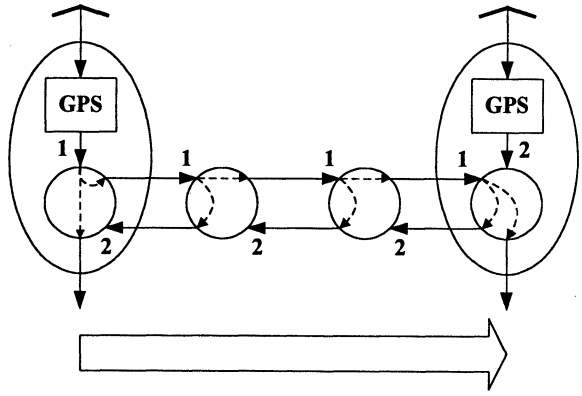


Figure 2: Cell pattern A

It consists of a linear chain of clocks using the SSM protocol for automatic protection switching. At both ends of the chain there is a GPS-receiver which delivers a synchronization signal to the node's co-located clock. This synchronization signal is derived from the GPS system time. The GPS-receiver's output signal must

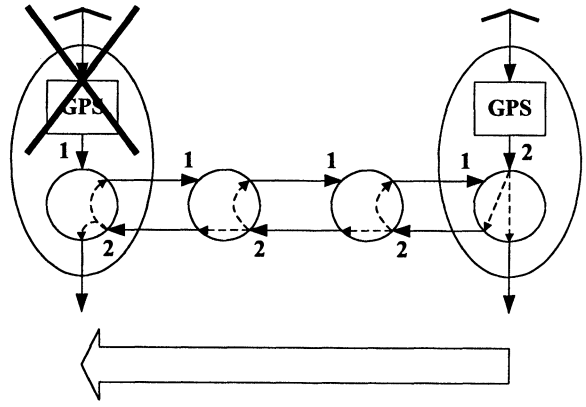


Figure 3: Failure of a GPS-receiver in an A-type cell

comply with the specifications of ITU-T Recommendation G.811 (see [6]). The clocks at the two ends of the chain provide outgoing synchronization

links for interconnection with other cells. Figure 2 also shows the direction of synchronization flow under normal operating conditions (i.e. no failures). Figure 3 shows what happens in case of failure of the GPS-receiver in the west end node: the automatic reconfiguration mechanism causes the GPS-receiver in the east end node to become the active primary reference source. Figure 4 shows the automatic reconfiguration caused by a link failure. After the automatic reconfiguration, all the clocks east of the failed link will get synchronization from the GPS-receiver located in the east end node. The failure of a clock inside the chain has similar consequences.

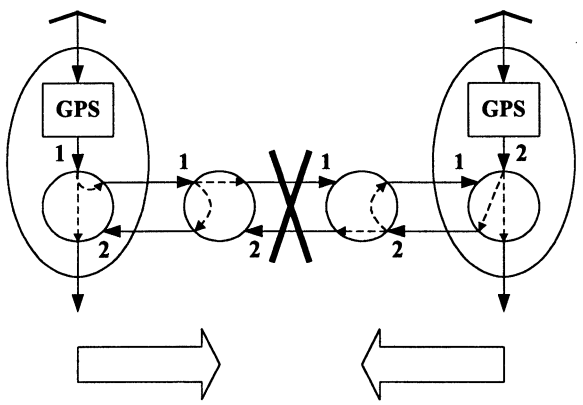


Figure 4: Link failure in an A-type cell

Cell pattern B resembles pattern A in the sense that it also consists of a linear chain of clocks using SSM-based automatic reconfiguration. Unlike pattern A, pattern B does not contain any GPS-receiver. Instead, there are incoming synchronization links connected to the chain's western and eastern end clocks. This is shown in figure 5. A B-type cell receives synchronization from other cells to which it is connected via the just mentionned incoming synchronization links. This will be shown in section 3.3. In case of failures, a pattern B cell essentially behaves the same way as a pattern A cell.

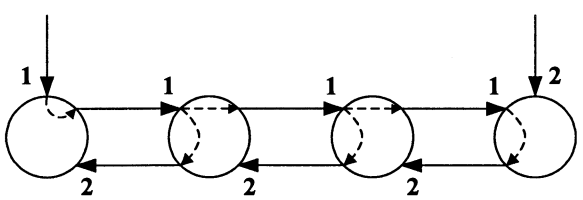


Figure 5: Cell pattern B

The maximum length of a chain, i.e. the maximum number of clocks in a chain, is limited by considerations regarding the maximum amount of jitter and wander that is allowed at any point in the network. The jitter and wander network limits for synchronization interfaces are specifed in ITU-T Recommendation G.823, chapter 6 (see [7]). The maximum chain length for the two standard cell patterns will become clear once the interconnection rules are defined. Therefore we will resume the discussion about the maximum chain length at the end of section 3.3.

3.3 The interconnection rules

A cellular synchronization distribution network consists of a set of A- and B-type cells interconnected according to the following rules:

1) For all B-type cells of the network: each of the two synchronization inputs must be connected to a synchronization output of an A-type cell; morever these inputs must be connected to different synchronization outputs (however, the two synchronization inputs of a B-type cell may be connected to the two synchronization inputs of a single A-type cell).
2) An A-type cell's synchronization output may be connected to one or many synchronization inputs of B-type cells.

Rule 1) prohibits connecting the two synchronization inputs of a B-type cell to one and the same synchronization output of an A-type cell. However, the two synchronization inputs of a B-type cell may be connected to the two distinct synchronization outputs of one and the same A-type cell.

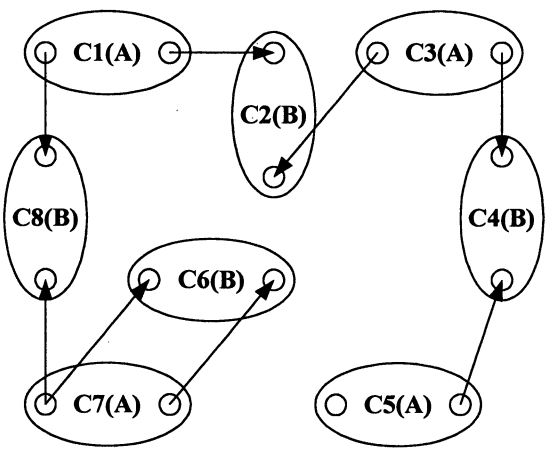


Figure 6: Cellular synchronization network

Figure 6 shows a simple example of a cellular synchronization distribution network (or simply 'cellular synchronization network'). Each ellipse

represents a cell. The cells are numbered: C1, C2, etc. The letter within the parentheses indicates whether the cell is an A-type or a B-type cell.

Here we need to resume the discussion about the maximum length of the clock chains in the cells. ITU-T Recommendation G.803, chapter 8 (see [4]) describes the architecture of a classical SDH-based master-slave synchronization distribution network. Section 8.2.4 states that, in a chain of clocks, there should be no more than 20 SDH equipment clocks between the PRC and the first so-called 'node clock' (a node clock is high performance slave-clock featuring a bandwidth that is three orders of magnitudes narrower than that of an ordinary SDH equipment clock; node clocks are used to attenuate accumulated jitter and wander). Since there are no node clocks in a cellular synchronization network, the maximum of 20 SDH equipment clocks applies to the maximum number of clocks that may be found in the longest possible chain following a GPS-receiver. As a consequence of the interconnection rules described ealier, the longest possible chain of clocks following a GPS-receiver is given by the chaining of one A-type cell and one B-type cell. It follows that the rule from ITU-T Rec. G.823 is fulfilled, if the chain length in both A- and B-type cells is limited to 10 SDH equipment clocks. Of course, this limit can be extended by the deployment of narrow bandwidth node clocks, as described in ITU-T Recommendation G.803; but this approach is outside the scope and the motivation of this paper.

3.4 Extended architecture

It is possible to define a somewhat richer syntax for the cellular synchronzation network architecture.

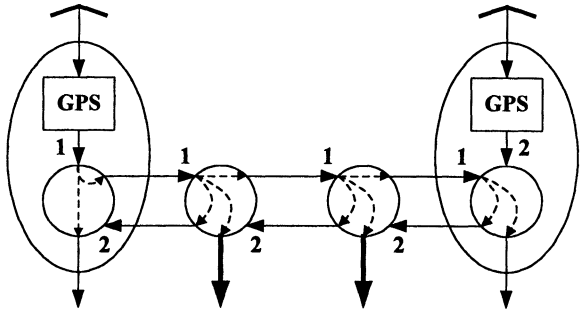


Figure 7: Cell pattern A*

Figures 7 to 9 give some indications on how this can be done. Figure 7 shows a cell pattern A* which is derived from pattern A. Compared to the latter, pattern A* features additional synchronization outputs indicated by the boldface arrows. Similarly, figure 8 shows a cell pattern B* which is derived from pattern B. Compared to the latter, pattern B* has additional synchronization inputs. It can be shown that the SSM-based automatic reconfiguration works also with pattern B*. These new cell patterns allow for more complex topologies to be covered by the cellular architecture. This is illustrated only schematically in figure 9. Compared to the example of figure 6, there is an additional synchronization link (indicated by the boldface arrow) connecting cells C1(A) and C6(B). In this case the additional link provides additional protection against failures affecting the synchronization distribution between C1 and C6, and within C6.

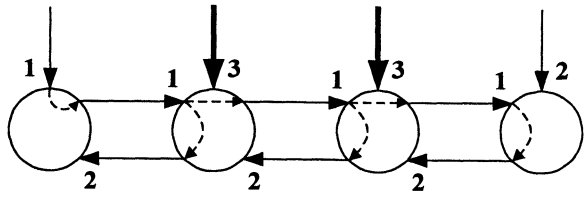


Figure 8: Cell pattern B*

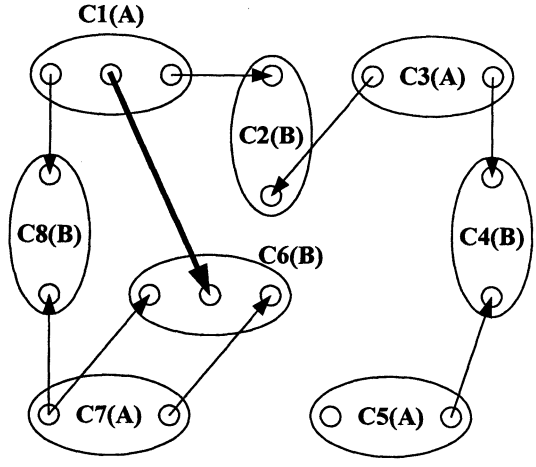


Figure 9: Extended cellular synchronization architecture

4. PROPERTIES

The chosen network structure results in some interesting network properties.

4.1 Protection against failures

An interesting feature is the fact that all nodes of a cellular synchronization network get two reference signals coming from two GPS-receivers located in different sites via geographically separate routes. Thus a cellular synchronization network is at least protected against any single failure of a link, a clock, or a GPS-

receiver. Of course, there is also protection against many multiple failure cases. The protection against GPS-receiver failures includes the case of local corruption of the GPS radio signal caused by radio interference or intentional jamming. This failure mode would not be covered by simply installing two GPS-receivers and two antennas on that site, since both receivers would capture the same corrupted radio signal.

4.2. Timing loop prevention

The topologies resulting from the described cellular architecture have the property of necessarily excluding timing loop formation. Timing loops occur, when a chain of slave-clocks forms a loop, so that the active reference input of any of the clocks is actually locked indirectly to the output of that same clock. The occurrence of a timing loop can have dramatic consequences. The clocks on the timing loop are not locked to a PRC anymore. Because of the dynamic characteristics of the loop, the loop's clock frequency diverges, thus causing catastrophic degradation of traffic performance (e.g. bit error rate) in the telecommunications network. Timing loops occur because of human error during the network design process or during equipment commissioning. Making sure that there is no possibility of timing loop formation in a classical redundant master-slave tree network is a complex task. One must not only check the primary synchronization distribution paths, but also all possible path combinations resulting from protection switching events. Failures in the network will cause some of the clocks to switch to another reference input (automatic reference selection function). These switchover events modify the topology of the synchronization distribution paths. The network designer must make sure that these topology changes never lead to the creation of a timing loop. Verifying that a network is exempt of timing loops is much easier with cellular synchronization networks. The internal structure of the cells and the cell interconnection rules necessarily exclude timing loops. A cell either delivers synchronization to other cells (cell patterns A and A*), or receives synchronization from other cells (cell patterns B and B*). There are no cells which would simultaneously deliver and receive synchronization to or from other cells. It is thus impossible to interconnect cells in such a way, that a timing loop is created. Inside a cell the creation of timing loops is prevented by the SSM protocol. The way SDH equipment clocks process the SSM is standardized in ITU-T Recommendation G.781 (see [5]). The standard clock behaviour prevents creation of timing loops in chains of SDH equipment clocks using SSM. Finally, the simplicity of the design rules minimize the risk of human error during the network design process.

4.3 Simple and scaleable network design

One of the most beneficial advantages of cellular synchronization networks over classical master-slave tree networks lies in the reduced complexity of the network planning process. Network planning includes the maintenance of the synchronization network plan when the traffic network grows or when it undergoes topology, equipment, or configuration changes. Experience shows that adapting the synchronization network to these modifications is a time-consuming task with classical master-slave tree designs. Sometimes it is necessary to redesign the entire tree topology. Also, these changes entail a lot of equipment reconfiguring. Adapting a cellular synchronization network is less complex. All it takes is appending new cells to the existing design or changing the internal design of cells. Adding new cells to a cellular network does not affect the structure of the already existing network part. Changing the internal design of a cell is easy because of the cell's simple chain structure. Moreover, a change of the internal design of a cell does not affect the inter-cell topology. The only point that requires some attention is the length of the cell's clock chain, since there exists an upper limit (see section 3.3). But again, if the number of equipment clocks in a chain grows beyond the upper limit, all that is required, is splitting the cell into several smaller cells. In short, one can say that cellular synchronization networks are easily scaleable.

5. CONCLUSIONS

This paper presents an novel method for synchronizing telecommunications networks. It is based on the Global Positioning System (GPS) for coarse distribution to a number of subnetworks, and on the SDH/SONET transport network for fine distribution to all nodes within the subnetworks. Two types of subnetworks with different internal structures are defined. A set of rules describes how these subnetworks, also called cells, can be combined in order to build a network that distributes synchronization to the entire telecommunications network under consideration. This set of rules determines a class of inter-cell or external structures. The chosen internal and external structures result in some interesting network properties. Synchronization networks of this type are scalable: they can be expanded very easily by adding new cells, hence the name « cellular synchronization networks ». This kind of scalability greatly facilitates the maintenance of a synchronization network plan during times when the telecommunications network is growing. Furthermore the chosen architecture has the property of excluding the formation of timing loops. Another interesting feature is the fact that all nodes of a cellular synchronization network get two reference signals

coming from two GPS receivers located in different sites via geographically separate routes. Thus the network provides protection against link failures, clock failures and corrupted GPS radio signals. These properties show that the cellular synchronization network concept is an interesting alternative to the more traditional designs based on a tree of master-slave clocks locked to a central Primary Reference Clock. Cellular synchronization networks are suitable for a wide range of telecommunications networks, both large and small.

6. REFERENCES

[1] M. Sexton and A. Reid, Broadband Networking: ATM, SDH and SONET, Boston: Artech House, 1997, ch. 11, pp. 415-424.

[2] St. Bregni, "A Historical Perspective on Telecommunications Network Synchronization", IEEE Communications Magazine, vol. 36, pp. 158-166, June 1998.

[3] S. Wong, "How to Synchronize Telecommincations Networks", presented at the 11th European Frequency and Time Forum, Neuchâtel, Switzerland, March 4-6, 1997.

[4] ITU International Telecommunications Union, ITU-T Recommendation G.803: Architecture of transport networks based on the synchronous digital hierarchy (SDH), Geneva, 1997, ch. 8, pp. 14-25.

[5] ITU International Telecommunications Union, ITU-T Recommendation G.781: Synchronization layer functions , Geneva, 2000.

[6] ITU International Telecommunications Union, ITU-T Recommendation G.811: Timing characteristics of primary reference clocks , Geneva, 2000.

[7] ITU International Telecommunications Union, ITU-T Recommendation G.823: The control of jitter and wander within digital networks based on the 2048 kbit/s hierarchy, Geneva, 2000.

[8] E. Butterline, J. Abate, G. Zampetti, "Timing a National Telecommunications Network Using GPS", presented at the 6th European Frequency and Time Forum, Noordwijk, Netherlands, 17-19 March, 1992.

Transmission Line Stabilised Monolithic Oscillators

K. S. Ang, M. J. Underhill and I. D. Robertson
School of Electronic Engineering, Information Technology and Mathematics, University of Surrey, Guildford, Surrey GU2 5XH, UK

ABSTRACT

A novel technique for frequency stabilisation and phase noise reduction of monolithic oscillators is presented. It employs simple transmission line resonators which are many wavelengths long to increase the oscillator quality factor. Monolithic oscillators at 20GHz and 40GHz are realised for the application of this technique. Phase noise reduction of more than 20dB was achieved for both oscillators. The single side-band phase noise obtained was -100dBc/Hz at 100KHz offset for the 20GHz oscillator and -90dBc/Hz at 1MHz offset for the 40GHz oscillator. The approach does not need accurate resonator dimensions or its precise placement and no additional circuitry is required. Thus, it presents significant potential in the development of low cost MMIC oscillators with enhanced noise performance.

1. INTRODUCTION

In recent years, an increasing number of oscillator circuits have been implemented using monolithic technology due to the overwhelming advantages of size, reliability and cost [1]. These oscillators, however, generally have poor phase noise performance as the monolithic circuitry has relatively low quality factor (Q). To meet the low noise requirements of practical systems, some form of oscillation stabilising scheme has to be employed.

The most common frequency stabilisation and phase noise reduction techniques are the use of dielectric resonator oscillators (DROs) [2,3] and phase locked oscillators (PLOs) [4,5]. Due to the extraordinarily high Q, DROs offer excellent low noise performance. However, they require accurate dimensions of the dielectric puck and its precise placement on the monolithic or package substrate as these parameters determine the oscillating frequency and stability. This demands careful post-fabrication attention, resulting in high manufacturing cost and low repeatability. The alternative approach of using PLOs requires many components such as voltage controlled oscillators, frequency dividers and phase/frequency comparators. Furthermore, additional frequency multipliers and amplifiers are required due to the limited operating frequency of frequency dividers. These requirements result in a complex and high cost multi-chip scheme at millimeter-wave frequencies.

In this paper, a stabilising technique using simple transmission line resonators is presented. It does not require accurate dimensions or precise placement of the resonator. In fact, the transmission line resonator is connected directly to the monolithic oscillator without any additional circuitry. This results in a simple, low cost and robust stabilising scheme. A simplified derivation of the improvement in oscillator Q and measurement results of K and Ka-band monolithic oscillators are presented to demonstrate the technique.

2. TRANSMISSION LINE RESONATORS

The most effective way of improving an oscillator's noise performance is to increase its Q factor, as it determines the oscillator's operating parameters such as phase noise, frequency stability and pulling figure. In this section, the feasibility of using long transmission line resonators to increase the loaded Q of oscillator circuits is investigated.

Fig. 1 shows a functional diagram of a microwave oscillator. Maintaining the resonator plane a-a as a reference, the oscillator circuit and the load constitute a microwave generator, as shown in Fig. 2.

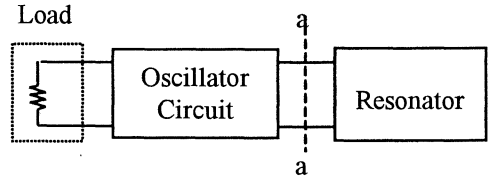


Fig. 1: Functional diagram of a microwave oscillator.

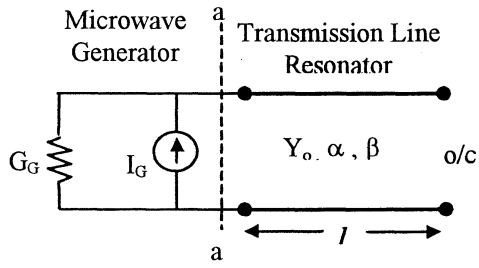


Fig. 2: Equivalent circuit of the oscillator at the resonator plane.

A constant current source I_G in parallel with source admittance G_G is used to represent the generator. The resonator is realised using a transmission line with characteristic admittance *Yo* and electrical length *l*. Open-circuited transmission

line resonators will be considered although equivalent results can be obtained with short-circuited resonators. The input admittance of the transmission line resonator can be expressed as

$$Y_{in} = G_{in} + jB_{in} \tag{1}$$

where

$$G_{in} = Y_0 \frac{\sinh(\alpha\ell)\cosh(\alpha\ell)}{\sinh^2(\alpha\ell)+\cos^2(\beta\ell)} \tag{2a}$$

$$B_{in} = Y_0 \frac{\sin(\beta\ell)\cos(\beta\ell)}{\sinh^2(\alpha\ell)+\cos^2(\beta\ell)} \tag{2b}$$

α and β are the attenuation and phase constants of the transmission line.

At frequencies where l is an integer number of half-wavelengths,

$$\beta\ell = N\pi \tag{3}$$

where N is an integer.

The input admittance is then given by

$$Y_{in} = G_{in} = Y_0 \tanh(\alpha\ell) \tag{4}$$

At frequencies in the vicinity of $\beta l = N\pi$, the input characteristics of the transmission line can be modelled by a parallel resonant circuit. The loaded Q factor of the resonant circuit can be evaluated as [6]:

$$Q_L = \frac{\omega_r}{2\left(G_G + G_{in}\big|_{\omega=\omega_r}\right)} \left.\frac{dB_{in}}{d\omega}\right|_{\omega=\omega_r} \tag{5a}$$

$$= \frac{Y_0 N\pi}{2\left(G_G + Y_0 \tanh\left(\frac{\alpha N\lambda_r}{2}\right)\right)\left(\cosh^2\left(\frac{\alpha N\lambda_r}{2}\right)\right)} \tag{5b}$$

where λ_r is the resonant wavelength.

From (5b), it appears that by increasing the length of the open-circuited transmission line, Q_L can be increased. This is valid if the transmission line is sufficiently low loss such that the terms of $(\alpha N\lambda_r)$ are negligible. As Q_L increases with increasing length, the associated line losses will become increasingly significant. This will result in decreasing improvement with increasing length. Eventually, as the line losses become excessive, Q_L will start to decrease.

Another effect of increasing line length is that the transmission line will have many closely spaced resonant frequencies. Thus, the circuit will have spurious response close to the desired resonant frequency, which is unacceptable in many applications. Therefore, for the approach to be feasible, the oscillator circuit must retain some selectivity, other than that provided by the transmission line. This is to ensure oscillation occurs only at a single mode whereby other alternative modes are sufficiently suppressed once a single oscillation mode is excited [7]. In addition, it must be able to operate under dual loading conditions, being connected both to an external load and a transmission line resonator. This requires two output ports with sufficient isolation. If these oscillator requirements are met and the transmission line resonator is sufficiently low loss, increasing the resonator length is an effective way to increase Q_L, and reduce phase noise. A simple oscillator circuit fulfilling these requirements is the balanced oscillator circuit presented in [8].

3. OSCILLATOR CIRCUIT

The schematic diagram of the oscillator circuit is shown in Fig. 3. It consists of two FETs whose gates are interconnected by a transmission line. Capacitive source feedback is applied to both FETs to generate a negative conductance at the desired oscillation frequency. Looking into the gates of each device is a positive susceptance jB. The length of the transmission line is chosen such that the two devices resonate one another. This is achieved by selecting the length so that each device's susceptance is transformed to a susceptance with the same magnitude but opposite in sign. Thus, the device susceptances and the transmission line form a resonant circuit, which determines the oscillating frequency. The oscillator outputs for the external load and transmission line resonator connections are taken from the drains of the two FETs, providing the required output port isolation.

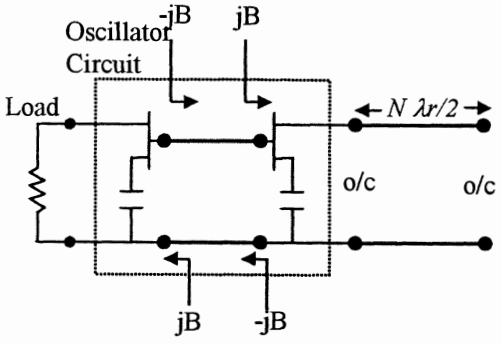


Fig. 3: Schematic diagram of the oscillator circuit.

4. MONOLITHIC OSCILLATOR REALISATION

Based on the oscillator circuit described, monolithic oscillators at K and Ka-Band were designed to investigate the technique experimentally. These chips were fabricated using

GEC Marconi's standard F20 and H40 processes. The F20 process employs 0.5μm gate length MESFETs on 200μm thick GaAs substrate. The H40 process employs 0.25μm gate length AlGaAs/InGaAs pHEMTs on 100μm thick GaAs substrate. The oscillator frequencies for the F20 and H40 processes were chosen to be 20GHz and 40GHz respectively, to fully exploit the capabilities of the devices and to demonstrate the technique for millimeter wave applications.

Photographs of the 20GHz and 40GHz oscillators are shown in Figs. 4 and 5. The chip dimensions are 1.2x1.6mm and 0.9 x 1.3 mm, respectively. Each circuit employs two devices with capacitive source feedback, laid out in rotational-symmetry. Mesa resistors of 3kΩ were used for the gate bias while quarter-wavelength transmission lines were used for the drain bias and source grounding.

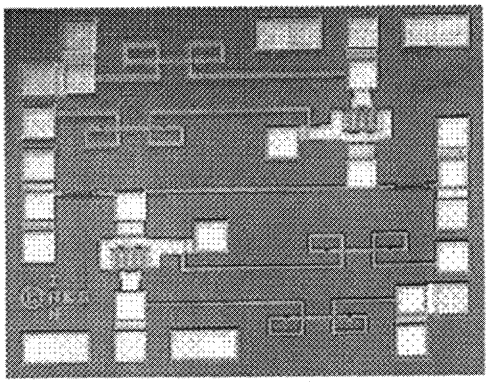


Fig. 4: Photograph of K-band oscillator chip.

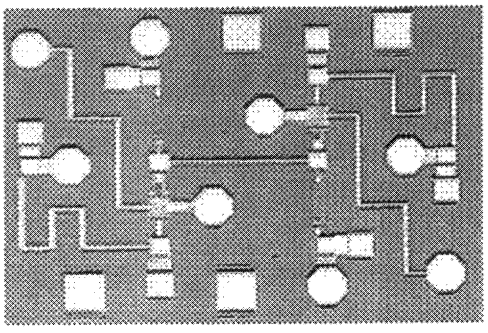


Fig. 5: Photograph of the Ka-band oscillator chip.

5. EXPERIMENTAL RESULTS

The oscillator chips were measured on a Cascade Microtech probestation. All RF connections to the chips were made using 200μm pitch coplanar probes while DC biasing was applied though probe needles.

The K-band oscillator was first measured by connecting one of the output ports to a HP8563E spectrum analyser, leaving the other output port open-circuited. Stable oscillations at around 20GHz with 0dBm output power were observed at Vgs=0V and Vds=3V. Without altering the biasing conditions or the analyser settings except the center frequency, the oscillator was then measured with open-circuited coaxial cables connected to the other output port. Three coaxial cables of different lengths were used and their measured parameters are given in Table 1:

Table 1: Measured parameters of the 3 coaxial cables.

Reso-nator	Cable Length (cm)	No. of half-wavelengths, N		Insertion Loss (dB)	
		20GHz	43GHz	20GHz	43GHz
R1	≅ 10	23	49	0.35	0.7
R2	≅ 50	89	191	1.7	3.2
R3	≅ 100	171	368	3.2	6.5

Fig. 6 shows the measured spectra obtained for the four different resonator loading conditions. The displayed power levels include probe and cable losses, which are about 4dB in total. Similar spectra were observed when the transmission lines were terminated in a short-circuit instead of an open-circuit. Spurious oscillations at unwanted signal frequencies were not observed in any of the cases.

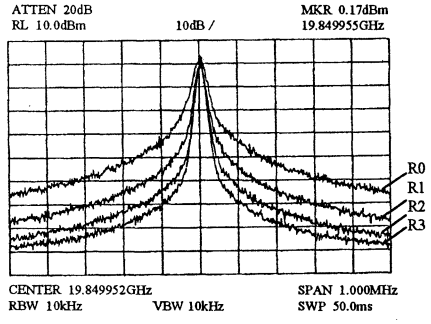


Fig. 6: Frequency spectrums of K-band oscillator with different resonators:
R0: No external resonator
R1: Open-circuit resonator of length ≅10cm
R2: Open-circuit resonator of length ≅50cm
R3: Open-circuit resonator of length ≅100cm

Fig. 6 depicts a significant 10dB reduction in phase noise with the connection of a short transmission line resonator, R1. Increasing the resonator length by five times, using R2, results in a further improvement of about 8dB. Finally, using resonator R3, which is twice the length of R2, results in a further 3dB improvement. The measured phase noise for R3 at 100kHz offset is about –100dBc/Hz.

This trend of decreasing phase noise improvement with increasing length is depicted in Fig. 7. The resonator length is plotted in terms of N, the number of half-wavelengths. Also plotted, is the increase in Q_L as the resonator length increases, given by (5b). Neglecting the effect of non-linear reactances, phase noise near the carrier is inversely

proportional to Q_L^2 [9]. Therefore, $20log(Q_L)$ is plotted in Fig. 7 for direct comparison with improvement in phase noise. The two curves are in general agreement. Therefore, it is evident that increasing resonator length effectively increases the oscillator Q_L and equivalently reduces phase noise.

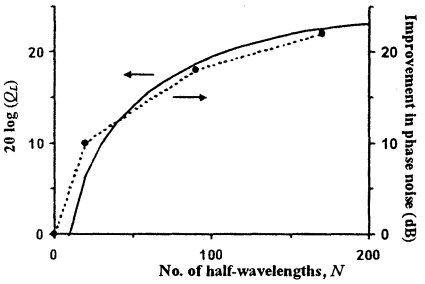


Fig. 7: Improvement in phase noise and 20 log(Q_L) vs N for the K-band oscillator.

The Ka-band oscillator was measured on the spectrum analyser using external mixers. The same coaxial cables given in Table 1 were used. Fig. 8 shows the frequency spectra obtained under the different resonator loading conditions, with Vgs=0V and Vds=2.0V. Note that the displayed power level includes probe and cable losses which are about 7dB, as well as a 40dB attenuator, which was included at the input of the external mixer. Fig. 8 depicts a dramatic 20dB decrease in phase noise with the addition of resonator R1. The lowest phase noise is obtained using resonator R2, with -90dBc/Hz at 1MHz offset. Increasing the resonator length to R3 actually causes 3dB degradation in phase noise.

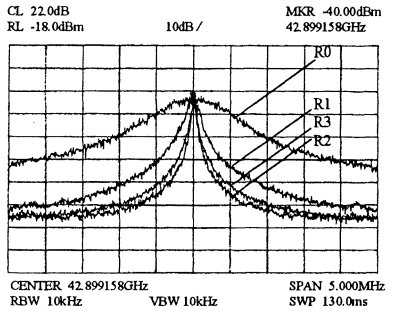


Fig. 8: Frequency spectrums of Ka-band oscillator with different resonators:
R0: No external resonator
R1: Open-circuit resonator of length ≅10cm
R2: Open-circuit resonator of length ≅50cm
R3: Open-circuit resonator of length ≅100cm

Fig. 9 shows the measured phase noise improvement and the predicted increase in Q_L as a function of resonator length. The curve of $20log(Q_L)$ predicts a maximum phase noise improvement of 35dB with resonator length of about 200 half-wavelengths. Beyond which, the associated line losses become increasing significant such that Q_L starts to decrease and phase noise degrades with increasing length. This trend is in agreement with the measured results.

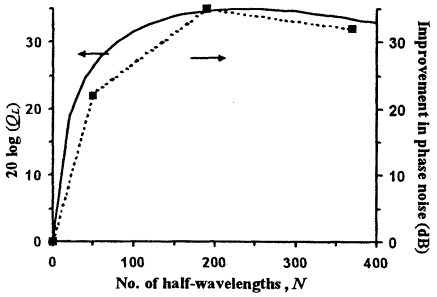


Fig. 9: Improvement in phase noise and 20log(Q_L) vs N for the Ka-band oscillator.

In addition to phase noise reduction, an increase in frequency stability was also observed with increasing resonator length. As an indication of frequency stability, frequency pushing of the two oscillators for the different resonators was measured. The variation of oscillation frequency with drain bias for the K and Ka-band oscillators are shown in Figs. 10 and 11, respectively. With the addition of external resonators, a dramatic decrease in frequency pushing was observed for both oscillators. The minimum frequency pushing measured was 5MHz/V for the K-band oscillator and 30MHz/V for the Ka-band oscillator.

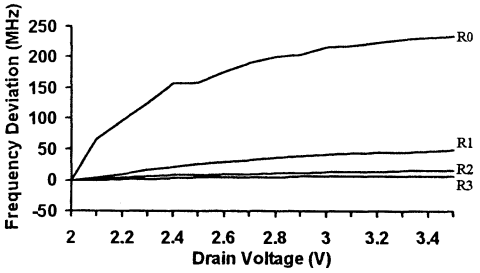


Fig.10: Frequency pushing of the K-band oscillator under different resonator conditions:

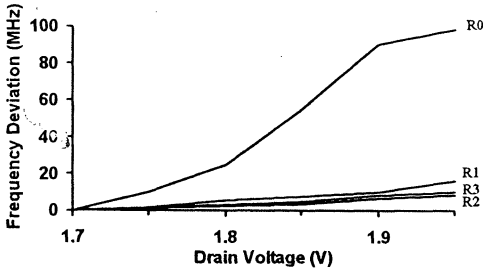


Fig. 11: Frequency pushing of the Ka-band oscillator under different resonator conditions:

6. CONCLUSION

This paper has demonstrated that oscillator frequency stabilisation and phase noise reduction can be achieved using simple transmission line resonators. A simplified derivation of the improvement in phase noise has been verified through measured results of K and Ka-band monolithic oscillators. More than 20dB improvement in phase noise was obtained for both oscillators. For simplicity, widely available coaxial transmission lines were used. However, any other low loss transmission lines can be employed. In fact, further reduction in phase noise can be achieved by using extremely low loss transmission line media, such as those fabricated using micromachining[10]. The simplicity and robustness of the technique offers great potential in the development of low-cost MMIC oscillators with significantly improved noise performance.

REFERENCES

[1] I. D. Robertson, Ed., *MMIC Design* (Inst. Elect. Eng. Circuits and Systems Series 7). London, U.K.: IEE, 1995, pp 337-370.

[2] P. G. Wilson, "Monolithic 38GHz dielectric resonator oscillator," in *1991 IEEE Int. Microwave Symp. Dig.*, pp. 831-834.

[3] M. Funabashi *et al.*, "V-band AlGaAs/InGaAs heterojunction FET MMIC dielectric resonator circuit," in *1994 IEEE GaAs IC Symp. Dig.*, pp.30-33.

[4] M. Madihian and K. Honjo, "GaAs monolithic IC's for an X-band PLL-stabilised local source," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-34, pp. 707-713, June. 1986.

[5] A. Kanda, T. Hirota, H. Okaxaki and M. Nakamae, "An MMIC chip set for V-band phase-locked local oscillator," in *1995 IEEE GaAs IC Symp. Dig.*, pp.259-262.

[6] P. A. Rizzi, *Microwave Engineering*, New Jersey: Prentice-Hall, 1988, pp. 411-428.

[7] M. J. Underhill, "Fundamentals of oscillator performance," *Electronics & Communication Engineering Journal,* vol. 44, pp. 185-193, Aug. 1992.

[8] K. S. Ang, S. Nam and I. D. Robertson, "A balanced monolithic oscillator at K-band," in *1999 IEEE Radio Frequency Integrated Circuits Symp. Dig.*, pp.163-166.

[9] D. B. Leeson, "A simple model of feedback oscillator noise spectrum," *Proc. IEEE*, vol. 54, pp. 329-330, Feb. 1966.

[10] A. R. Brown and G. M. Rebeiz, "A Ka-band micromachined low-phase-noise oscillator," *IEEE Trans. Microwave Theory Tech.*, vol. 47, no. 8, pp. 1504-1508, Aug. 1999.

NOVEL DESIGN OF HIGH-Q WHISPERING GALLERY SAPPHIRE-RUTILE DIELECTRIC RESONATORS FOR A SECONDARY FREQUENCY STANDARD

M. E. Tobar*, J. G. Hartnett*, E. N. Ivanov*, D. Cros#, P. Blondy#, P. Guillon#

* Department of Physics, University of Western Australia, Nedlands, 6907, WA, Australia
# IRCOM, UMR 6615 CNRS, Faculté des Sciences, Av. A. THOMAS, 87060 Limoges, France.
Email: mike@physics.uwa.edu.au Phone: +61-8-9380-3443 Fax: +61-8-9380-1014

ABSTRACT

The highest short-term frequency stable microwave resonator-oscillators utilize liquid helium cooled sapphire dielectric resonators. The temperature coefficient of frequency of such resonators is very small due to residual paramagnetic impurities canceling the Temperature Coefficient of Permittivity (TCP). At higher temperatures, which are accessible in space or with liquid nitrogen, the effect is too weak, and if extra impurities are added the loss introduced is too great. An alternative technique involves using two low-loss dielectric materials with opposite signs of TCP. Following this approach a sapphire-rutile resonator was designed with mode frequency-temperature turning points between 50 to 80 K, with Q-factors of order 10^7. Previous designs use thin disks of rutile fixed to the ends of the sapphire cylinder. Because of the high permittivity of rutile such resonators have a high density of spurious modes. By placing rings at the end faces instead of disks, the majority of the spurious modes are raised above the operation frequency and the requirement for thin disks is removed. Finite element analysis has been applied and compares well with experiment. The application to the design of high stability "fly-wheel" oscillators for atomic frequency standards is discussed.

1. INTRODUCTION

Low-noise high-stability resonator-oscillators based on high-Q sapphire Whispering Gallery (WG) mode resonators have become important devices for telecommunication, radar and metrological applications. The extremely high quality factor of sapphire, $>10^5$ at room temperature, $>10^7$ at liquid nitrogen temperature and $>10^9$ at liquid helium temperature has enabled the lowest phase noise[1, 2] and most frequency stable[3-5] oscillators in the microwave regime. To create an oscillator with exceptional frequency stability, the resonator must have the frequency-temperature dependence annulled, as well as a high quality factor. The Temperature Coefficient of Permittivity (TCP) for sapphire is quite large, and of the order of 10 ppm/K at 77 K. This mechanism allows temperature fluctuations to transform to resonator frequency fluctuations.

The usual electromagnetic technique of annulment is due to the effect of paramagnetic impurities contributing an opposite temperature coefficient (due to magnetic susceptibility) compared to the TCP. This technique has only been realized successfully in liquid helium environments[6-8]. It is important to raise this temperature of compensation to 40-80 K if this technology is to be developed for space applications and for liquid nitrogen cooled devices. To raise the temperature of annulment large concentrations of paramagnetic impurities are required, which in turn significantly degrade the Q-factor[9]. Recently a new technique incorporating dielectric compensation to a WG sapphire resonator was developed[10, 11]. The method consists of placing two dielectric disks at the end of the sapphire cylinder as shown in fig. 1.

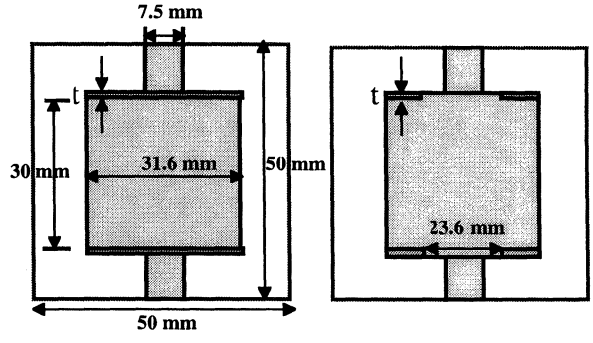


Fig. 1. Sapphire resonator with rutile disks/rings of thickness t, held to the end faces.

A good choice of compensating material is rutile, which has low loss and an opposite sign of TCP. The values of TCP for sapphire and rutile have been measured accurately from 300 K to below 10 K[12, 13], and we used these results to calculate the characteristics of a sapphire-rutile resonator as a function of temperature. Dielectric compensation was achieved in a sapphire-rutile composite WG resonator with a measured Q-factor of 3×10^7 at 56 K.

From this result and state-of-the-art frequency stabilization circuitry, we demonstrate that the current technology is suitable for the construction of an ultra-stable oscillator with fractional frequency instability of order 10^{-14}. This is the 'fly-wheel' oscillator requirement necessary for an atomic fountain or cold atom clock to reach the performance set by the quantum limit[14]. Even though a liquid helium cooled clock has two orders of magnitude better stability[5], the liquid nitrogen clock is more easily transportable and much cheaper to maintain. The oscillators based on composite dielectric resonators can also be considered for use as flywheel oscillators for space applications (such as the Atomic Clock Ensemble in Space "ACES" project on board the International Space Station). State-of-the-art quartz oscillators are an order of magnitude worse ($\sim10^{-13}$) and limit the performance of an atomic clock due to the Dick Effect[15, 16].

2. COMPOSITE RESONATOR PROPERTIES

2.1 Finite Element Analysis

Rigorous analysis of 12 GHz modes in the structure (see fig. 1) was achieved by implementing finite element software developed at IRCOM, specifically designed to solve resonant anisotropic dielectric systems[17]. We analyzed the frequency-temperature behavior of the $WGH_{10,0,0}$ (or $N1_{10}$) and the $WGE_{8,0,0}$ ($S2_8$) modes. The WG notation is the same as introduced at Limoges[18], and the N-S notation means Non (or anti-) -symmetric and Symmetric magnetic field in the axial direction respectively, and the following number denotes the ascending order in frequency[13, 19]. The annulment temperature versus thickness graphs are shown in fig. 2, and the electric field density plots are shown in fig. 3.

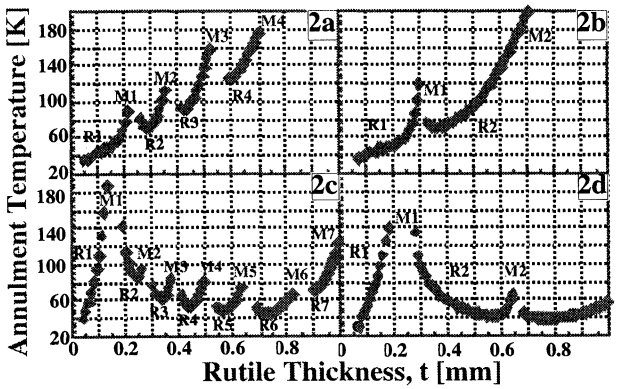


Fig. 2. Annulment temperature versus thickness for various modes. Regions of mode compensation are labeled R and regions with spurious mode interaction are labeled M: 3a: $WGH_{10,0,0}$ ($N1_{10}$) mode with disks of rutile fixed to the end faces. 3b: $WGH_{10,0,0}$ ($N1_{10}$) mode with rings rather than disks fixed to the ends. 3c: $WGE_{8,0,0}$ ($S2_8$) mode with disks of rutile fixed to the end faces. 3d: $WGE_{8,0,0}$ ($S2_8$) mode with rings rather than disks fixed to the ends.

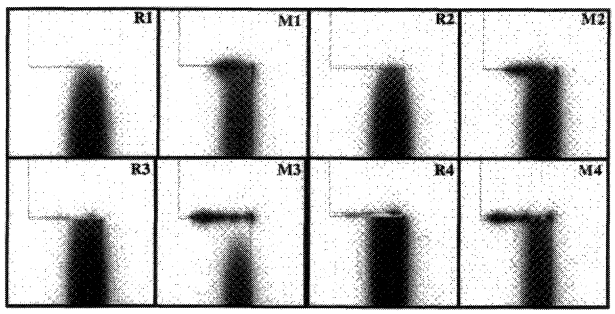


Fig. 3a. Electric field density plot for the $WGH_{10,0,0}$ mode in the labeled regions of fig. 2a.

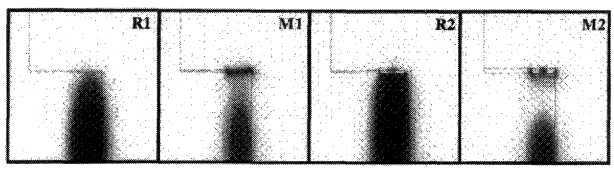


Fig. 3b. Electric field density plot for the $WGH_{10,0,0}$ mode in the labeled regions of fig. 2b.

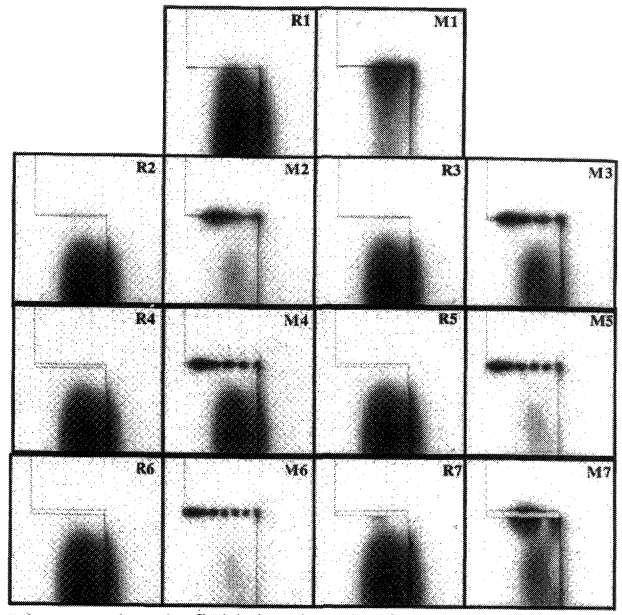


Fig. 3c. Electric field density plot for the $WGE_{8,0,0}$ mode in the labeled regions of fig. 2c.

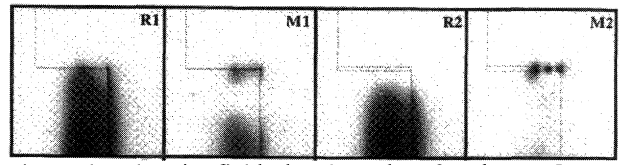


Fig. 3d. Electric field density plot for the $WGE_{8,0,0}$ mode in the labeled regions of fig. 2d.

To calculate the frequency-temperature dependence the anisotropic permittivity and expansion coefficients of sapphire and rutile need to be known as a function of temperature[12, 13]. Using these values an automatic program was written to calculate the frequency of the resonator as a function of temperature at intervals of 2.5 K between 20 to 62.5 K and intervals of 10 K above 70 K. A polynomial fit was then applied to this data to find the annulment temperature. This procedure was carried out for rutile thickness between 0.03 to 1.0 mm, and the results of calculations are plotted in fig. 2. Due to the symmetry of a cylindrical resonator, it was only necessary to calculate the frequency and field patterns in 1/4 of the resonator as shown in figure 3a to 3d.

All the frequencies of the analyzed modes were close to 12 GHz at cryogenic temperatures. The frequency of the same mode in the composite resonator in comparison to the bare sapphire resonator was higher for WGE modes as the axial boundary condition at the top and bottom sapphire-rutile interface causes the field to be squashed further into the sapphire (a Bragg effect). Conversely, the frequency is lower in the WGH modes as the axial boundary conditions cause the field to be stretched by the rutile disks due to the larger permittivity of rutile.

In general for a small thickness of rutile (~0.03 mm) there is not enough rutile to fully compensate the resonator. Instead a point of inflection occurs at about 20-30 K for both WGE and WGH modes. When the

thickness gets large enough the inflection point turns into an annulment point and separates in two. At the point of separation the second derivative is matched as well as the first and a flat annulment point of zero curvature is created. At larger values of thickness excess spurious modes exist due to modes mainly in the rutile. To lessen the spurious mode density we introduced rings (with an inner diameter of 23.6 mm) held to the ends. This had the effect of tuning the spurious rutile modes to higher frequencies with correspondingly larger frequency separations, and reduced the effect. However, the effect of the spurious modes on WGE and WGH modes are significantly different. Thus, we describe them separately in the following two subsections.

2.2 Temperature Characteristics for WGE Modes

Fig. 2c and 2d show the annulment temperature versus thickness for the $WGE_{8,0,0}$ mode. The rutile acts to only perturb the resonant frequency when the thickness is less than about 0.12 mm. This is the linear regime where the annulment temperature is proportional to the thickness. However, at about 0.12 mm in the disk structure and 0.15 mm in the ring structure a spurious mode starts to interact with the sapphire $WGE_{8,0,0}$ mode. This interaction with the new spurious mode is caused from the tangential boundary conditions, which require the transverse electric field between the rutile and sapphire to be continuous at the top and bottom of the sapphire resonator. The electric field density plots of the mode as a function of rutile thickness are shown in figures 3c and 3d. The first rutile modes that effect the sapphire mode are dominately TE due to the smaller permittivity of rutile in the perpendicular plane of the crystal.

After the thickness becomes large enough the spurious mode is tuned lower in frequency than the $WGE_{8,0,0}$ mode and the annulment temperature starts to decrease in temperature (anti-resonant effect). The rutile disks can support many WG modes, and in the range of thickness from 0 to 1 mm five more spurious mode interactions exist. If we substitute the disk for a ring structure, the frequency and separation of these modes are raised high enough that only one spurious mode remains out of the five.

It is interesting to note if we ignore the interacting WG modes in rutile, the temperature versus thickness characteristic with the disk and ring in fig. 2c and 2d are very similar. The first resonance due to the new mode occurs only at a slightly different thickness, unlike the other rutile WG modes, which are shifted greatly. Also, a local minimum in the annulment temperature-thickness characteristic occurs in both cases close to 0.75 mm. This suggests an effect that is mainly due to the boundary conditions at the top and bottom sapphire-rutile interface, between the sapphire WGE mode and the first TE Bragg mode in the rutile. The WGE modes are quasi-TE and hence have the majority of the electric field tangential to this boundary. Thus, WGE modes couple strongly to this new mode. This phenomenon is a manifestation of the Bragg effect[20].

2.3 Temperature Characteristics for WGH Modes

WGH modes are quasi-TM, therefore the majority of the electric field is normal to the sapphire-rutile boundary and the modes do not couple strongly to the new Bragg modes, which are quasi TE. Inspecting fig. 2a and 2b closely, small kinks at 0.12 mm in 2a and 0.15 mm in 2b are present. This is due to the small hybrid TE component coupling to the new TE Bragg mode. The coupling is too small to see the Bragg effect dominate due to the dominant TM structure. Thus, in general as the rutile thickness is increased, so does the annulment temperature as long as the interactions with spurious modes are ignored. When we replace the disk with the ring, three spurious rutile WG modes are reduced to one, in the range of thickness from 0 to 1 mm.

2.4 Comparison with Experiment

Two rutile rings 0.42 mm thick with inner diameter of 23.6 mm were held to the ends of the sapphire by sapphire holders incorporating a spring mechanism. Sapphire is the only material that we considered, so the Q-factor would not to be degraded. Measurements of frequency and Q-factor were achieved using standard techniques, similar to that as described by Luiten et. al.[21]. Experimental results and calculation are compared in table 1.

Table 1. Comparison of experimental results with finite element calculation

Mode	$WGE_{8,0,0}$	$WGH_{10,0,0}$
Measured frequency at annulment temperature [GHz].	12.031	11.916
Calculated frequency at annulment temperature [GHz].	12.071	11.947
Measured annulment temperature [K].	55	72
Calculated annulment temperature [K].	54	76
Measured curvature [ppm/K^2]	0.0365	0.44
Calculated curvature [ppm/K^2]	0.0395	0.37
Measured Q-factor [10^6]	4	6
Calculated Q-factor [10^6]	30	9

Even though we have not modeled the support system for the rutile rings, results of experiment and finite element modeling are in good agreement. This shows that finite element analysis is an excellent technique to design such a resonator for both WGE and WGH modes. This type of resonator is very difficult to model accurately with other techniques. Also it is important to note, if it was not for the ring structure (rather than a disk structure) the turning point temperature for WGE modes could not have been designed for accurately due to the large spurious mode density, without having

extremely thin disks. For example, previous results with 0.2 mm thin disks enabled the design of the turning point temperature for WGH modes, but not WGE modes [22].

Q-factors were also calculated by solving for complex frequencies. Losses due to the dielectric and copper cavity were taken into account. It should be noted here that the measured Q-factor of the $WGE_{8,0,0}$ mode is degraded due to the influence of a nearby weakly interacting spurious mode not calculated in the software, which also slightly reduces the curvature. The $WGE_{9,0,0}$ mode has been measured at 13 GHz (but not modeled) and also had a compensation point close to 55 K with a Q-factor of 30 million. This result is evidence that the Q-factor of the $WGE_{8,0,0}$ mode should be of the order of 30 million as predicted, if the near by mode did not degrade the performance. The mode is believed to be coupled to the support structure and with redesign we anticipate an improved Q-factor for this mode.

3. OSCILLATOR PERFORMANCE CONSIDERATIONS

3.1 Frequency Variations due to Electronic Noise

It is customary to characterize the frequency instability of an oscillator by the square root of Allan variance (SRAV), σ_y[23]. This is the primary measure of oscillator frequency instability in the time domain. Taking the minimum value of σ_y and multiplying it by the resonator Q-factor: $\delta = \sigma_y^{\min} Q$, another characteristic of an oscillator frequency stability, termed the line splitting factor is defined. Assuming the oscillator is frequency stabilized with a frequency noise suppression system, the line splitting factor can be interpreted as fraction of the resonator bandwidth within which the oscillator remains locked, with respect to the center of resonance.

Our goal is to build a microwave oscillator with a short-term frequency instability of order 10^{-14}. This is a necessary requirement to achieve the potential of a typical atomic Cs fountain or cold atom frequency standard for space applications. The UWA sapphire clock has been locked with a line splitting factor of 10^{-7} to obtain an instability of order 10^{-16}[5]. The JPL 87 K mechanically temperature compensated clock also locked with a line splitting factor of 10^{-7} to obtain a stability of 10^{-13}[24]. The voltage noise floor in an optimized Pound frequency discriminator was measured and the SRAV calculated. For a resonance with a Q-factor of 10^7 the measurements translate to a discriminator noise floor of 3 x 10^{-15} from 1 to 10 seconds of averaging time, rising to 3 x 10^{-14} at 100 seconds, which is suitable for an atomic frequency standard.

3.2 Frequency Variations due to Temperature

A copper cavity, in an evacuated can of a design similar to the liquid helium clock, was cooled to 77 K by liquid nitrogen and fractional temperature fluctuations were measured. Based on the results and the curvature at the annulment point of 2 x $10^{-7}/K^2$, the SRAV due to temperature fluctuations, was calculated to be 1.5 x 10^{-15} at 1 second rising to 3 x 10^{-14} at 30 seconds of averaging time. This assumes a temperature control maintaining the resonator within a 100 μK of the turning point, which is achievable with current temperature control technology. This calculation however, was without active temperature control, which when implemented, will even further reduce this source of noise.

4. CONCLUSION

One would prefer the rutile to act solely perturbatively on the sapphire resonator. Clearly this does not always occur. However our analysis clearly shows how to avoid resonant effects, as well as furnish an understanding of the modes in such a structure. Specifically we have shown by holding rutile rings to the end faces of a cylindrical sapphire resonator good designability of the annulment temperature above 30 K can be achieved with a low spurious mode density. These temperatures are easily accessible by closed cycle refrigerators, liquid nitrogen and radiative coolers for space applications. Also, we have shown that a frequency stabilized oscillator based on this resonator has the potential to pump an atomic frequency standard at the quantum limit.

ACKNOWLEDGEMENT

This work was supported by the Australian Research Council large and small grant scheme, and the French Centre National de la Recherche Scientific.

5. REFERENCES

[1] M. E. Tobar and D. G. Blair, "Phase noise analysis of the sapphire loaded superconducting niobium cavity oscillator," *IEEE Trans. on Micrw. Theory Tech.*, vol. 42, pp. 344-347, 1994.

[2] 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 Micrw. Theory Tech.*, vol. 46, pp. 1537-45, 1998.

[3] A. J. Giles, A. G. Mann, S. K. Jones, D. G. Blair, and M. J. Buckingham, "A very high stability sapphire loaded superconducting cavity oscillator," *Physica B*, vol. 165, pp. 145-146, 1990.

4] A. N. Luiten, A. G. Mann, and D. G. Blair, "Power Stabilized Exceptionally High Stability Cryogenic Sapphire Resonator Oscillator," *IEEE Trans. Instr. Meas.*, vol. 44, 1995.

[5] S. Chang, A. G. Mann, and A. N. Luiten, "improved cryogenic sapphire oscillator with exceptionally high frequency stability," *submitted to Elec. Lett.*, 2000.

[6] S. K. Jones, D. G. Blair, and M. J. Buckingham, "The effects of paramagnetic impurities on the frequency of sapphire loading superconducting resonators," *Electron. Lett.*, vol. 24, pp. 346-347, 1988.

[7] A. G. Mann, A. J. Giles, D. G. Blair, and M. J. Buckingham, "Ultra-stable cryogenic sapphire dielectric microwave resonators: mode frequency-temperature compensation by residual paramagnetic impurities," *Journal of Physics D,* vol. 25, pp. 1105-10, 1992.

[8] A. N. Luiten, A. G. Mann, and D. G. Blair, "Paramagnetic susceptibility and permittivity measurements at microwave frequencies in cryogenic sapphire resonators," *J. Phys. D: Appl. Phys.*, vol. 29, pp. 2082-2090, 1996.

[9] J. G. Hartnett, M. E. Tobar, A. G. Mann, E. N. Ivanov, J. Krupka, and R. Geyer, "Frequency-temperature compensation in Ti^{3+} and Ti^{4+} doped sapphire whispering gallery mode resonators," *IEEE Trans. on UFFC.*, vol. 46, 1999.

[10] 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.

[11] M. E. Tobar, J. Krupka, J. G. Hartnett, E. N. Ivanov, and R. A. Woode, "Sapphire-rutile frequency-temperature compensated whispering gallery microwave resonators," *Proc. IEEE Int. FCS.*, pp. 1000-8, 1997.

[12] M. Tobar, J. Krupka, E. Ivanov, and R. Woode, "Anisotropic Complex Permittivity Measurements of Mono-Crystalline Rutile Between 10-300 Kelvin," *J. Appl. Phys*, vol. 83, pp. 1604-9, 1998.

[13] J. Krupka, K. Derzakowski, A. Abramowicz, M. E. Tobar, and R. Geyer, "Whispering gallery modes for complex permittivity measurements of ultra-low loss dielectric materials," *IEEE Trans. on MTT.*, vol. 47, pp. 752-759, 1999..

[14] G. Santarelli, P. Laurent, P. Lemonde, A. Clairon, A. G. Mann, S. Chang, A. N. Luiten, and C. Salomon, "Quantum projection noise in an atomic fountain: A high stability cesium frequency standard," *Phys. Rev. Lett.*, pp. 4619-4622, 1999.

[15] G. J. Dick, "Local oscillator induced instabilities in trapped ion frequency standards," *in Proc. Precise Time and Time Interval*, vol. Redondo Beach, CA, pp. 133-147, 1987.

[16] G. Santarelli, C. Audoin, A. Makdissi, P. Laurent, G. J. Dick, and A. Clarion, "Frequency stability degradation of an oscillator slaved to a periodically interrogated atomic resonator," *IEEE Trans. Utrason. Ferroelect. Freq. Contr.*, vol. 45, pp. 887-894, 1998.

[17] M. Aubourg and P. Guillon, "A mixed finite element formulation for microwave device problems. Application to MIS structure," *Journal of Electromagnetic Waves and Applications*, vol. 45, pp. 371-386, 1991.

[18] P. Guillon and X. Jiao, "Resonant frequencies of whispering gallery dielectric resonators modes," *Proc. IEE, Part H*, vol. 134, 1987.

[19] J. Krupka, D. Cros, A. Luiten, and M. Tobar, "Design of very high-Q sapphire resonators," *IEE Electron. Lett.*, vol. 32, pp. 670-671, 1996.

[20] C. J. Maggiore, A. M. Clogston, G. Spalek, W. C. Sailor, and F. M. Mueller, "Low-loss microwave cavity using layered dielectric materials," *Appl. Phys. Lett.*, vol. 64, pp. 1451-1453, 1994.

[21] A. N. Luiten, A. G. Mann, and D. G. Blair, "High resolution measurement of the temperature-dependence of the Q, coupling and resonant frequency of a microwave resonator," *IOP Meas. Sci. Technol.*, vol. 7, pp. 949-953, 1996.

[22] M. E. Tobar, J. G. Hartnett, D. Cros, P. Blondy, J. Krupka, E. N. Ivanov, and P. Guillon, "Design of high-Q frequency-temperature compensated dielectric resonators," *Electron. Lett.*, vol. 35, pp. 303-5, 1999.

[23] D. W. Allan, "Time and frequency (time domain) characterization, estimation, and prediction of precision clocks and oscillators," *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, vol. 34, pp. 647-654, 1987.

[24] D. G. Santiago, G. J. Dick, and R. T. Wang, "Frequency stability of 10^-13 in a compensated sapphire oscillator operating above 77 K," *Proc. IEEE Int. Freq. Cont. Symp.*, pp. 772-775, 1996.

The anharmonic mode vibrating SC-cut resonators excited by lateral field - recent results

Krzysztof Weiss, Wiesław Szulc
Tele and Radio Research Institute, Ratuszowa 11, 03-450 Warsaw, Poland
E-mail: kweiss@itr.org.pl

Abstract
The SC 301 and 310 mode resonators have been developed over past several years in Besancon (perpendicular field) and in Warsaw (lateral field).The first results of 301 mode resonators excited by lateral field were presented on 12th EFTF in Warsaw. The Q factor of 8,2 MHz resonator was 1,3 x 10^6 and motional resistance was about 400Ω. The 311 mode resonators were presented on the Joint Meeting in 1999 in Besancon. Two different electrodes configurations were presented. The first gives the Q factor of 5 MHz resonators about 2 x 10^6 and motional resistance over 700 Ω, with the 310 mode resistance of only about 900 Ω. The second gives Q factor over 1,7 x 10^6 and motional resistance 340 Ω, with another modes vibrations resistance over 1000 Ω. At Tele and Radio Research Institute investigations on the lateral field excited resonators working on 301 or 311 modes are still being continued. By the electrodes configuration modification it was possible to obtain Q factor of 5 MHz resonators over 2 x 10^6 and resistance about 200Ω. These parameters were obtained as a result of investigations of anharmonic modes excitation areas by lateral and perpendicular field. 311 mode 10 MHz resonators were also designed and Q factor over 1,2 x 10^6 and motional resistance about 800 Ω were obtained. The 311 mode resonators exhibit about four times higher motional capacitance than classic lateral field 300 mode resonators. It means that for these resonators the frequency adjusting tolerance can be four times wider.

1. INTRODUCTION

The first resonators utilizing an anharmonic modes were developed in 1996 [1].These were resonators excited by perpendicular field with four electrodes configuration. The next were resonators excited by lateral field for 301 or 310 anharmonic mode vibrations presented in 1998 [2] and for 311 mode presented in 1999 [3]. These were the initial constructions presenting only the possibility of these modes excitation. In last year were carried out next investigations giving results in resonators parameters possible for practical utilization.

In the case of lateral field excited resonators the main problem is very narrow range of frequency adjustment. The classic lateral field resonators exhibit very low motional capacitance and are in practice non adjustable in oscillator. The anharmonic mode utilizing resonators exhibit higher motional capacitance and significant wider range of frequency adjustment in oscillator. But the first models of resonators exhibited significantly lower values of Q factor and in result higher level of noise. Presented below results of investigations show possibility of Q factor and motional resistance improvement by electrodes configuration and thickness modification.

2. EXPERIMENTAL RESULTS

2.1. 301 mode resonators

The first lateral field excited anharmonic mode resonators presented on 12th EFTF were made of round plano-convex SC cut plates designed for classic resonators with frequency 8,192 MHz on third overtone main mode. The best of these resonators exhibited motional resistance 390Ω, Q factor 1,3 x 10^6 and motional capacitance 0,038 fF. Comparing its parameters to classic lateral field resonator, Q factor and motional resistance were the same and motional capacitance was two times higher.

Next the 301 mode 5MHz resonators with the same electrodes configuration were designed. The resonators were manufactured and their parameters were measured. The electrical parameters of these resonators in comparison with classic LFE resonators are presented in tab. 1.

Tab. 1 Parameters of 300 and 301 mode resonators

Resonator type	No	T_e [C]	R1 [Ω]	Q [1E6]	C1 [E-18F]
ANH_{301}	54	85	562	1.53	37.0
ANH_{301}	58	70	565	1.56	36.0
LFE_{300}	M	70	574	2.35	24.0

The Q factor of these resonators was over 30% lower with about 50% motional capacitance increase.

The main works over 301 mode resonators improvement concerned Q factor and motional capacitance increase and motional resistance lowering. For 301 mode vibrations excitation the electrodes presented in fig. 1 were used. The first stage was the inner electrode widening and electrode thickness optimization. The next stage was a slight electrodes configuration modification. Two versions of the inner electrode with 0,5 and 1,2 mm width were checked and electrode thickness 35 and 50 nm were used (in previous

experiments electrode thickness 70 and 100 nm were checked).

The lateral field excitation regions measurement for 5MHz 3rd overtone plates did show that excitation regions were not symmetric. These measurements were repeated many times for different quartz plates samples and every time the eccentricity of 0,5 - 0,7 mm was observed (see fig. 2). These results were an inspiration for the electrodes modification. The electrodes were shifted from a plate center 0,7 mm. The electrodes shape is presented in fig. 3. The results of these experiments are presented in tab. 2. The resonators with resonant frequency 8 MHz and 10 MHz were also made and measured. The results are presented in tab 3 and 4. In tables, for comparison the same frequency perpendicular field and lateral field 300 mode resonators parameters are presented.

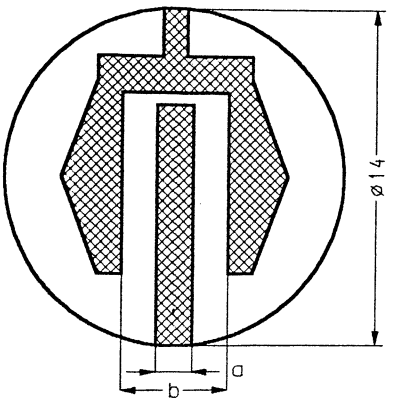


Fig. 1. The electrodes configuration for 301 mode resonators

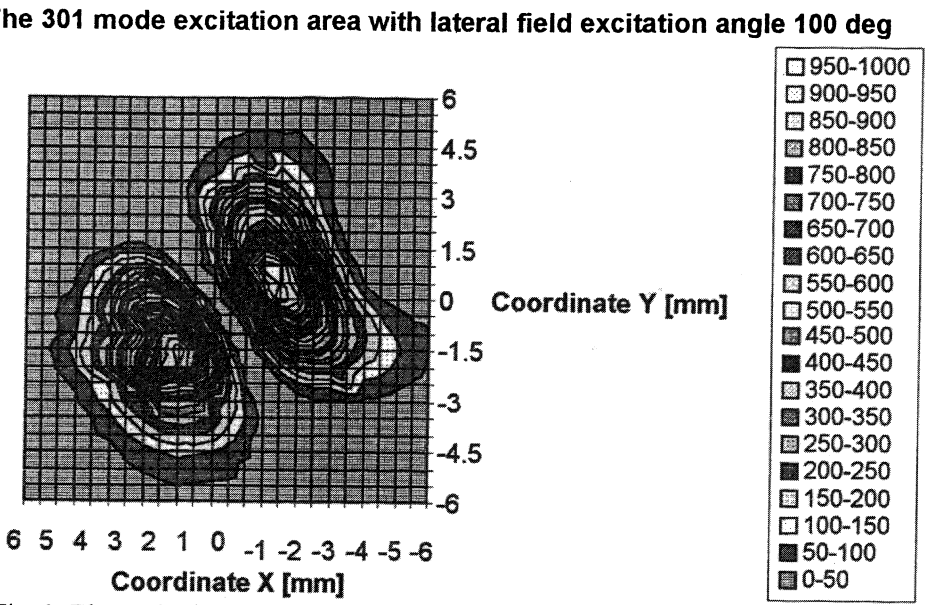


Fig. 2. The excitation areas distribution example in SC cut quartz plate for 301 mode

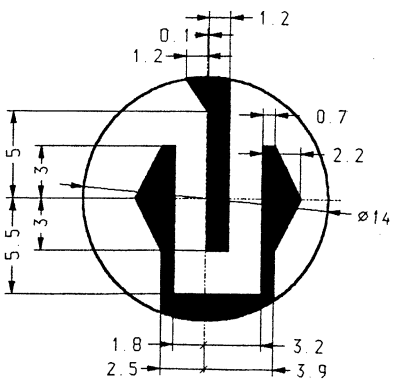


Fig. 3. The asymmetric electrodes configuration for 301 mode excitation.

The inner electrode widening gives higher value of motional capacitance and lower value of motional resistance. The electrode thickness 50 nm gives optimum value of Q factor. The electrode shift gives significant improvement of resonators with thick quartz plate (5 MHz 3rd ov.). In thinner plates the electrode shift effect is not too significant.

These experiments results showed possibility to obtain anharmonic 301 mode resonators Q -factor values comparable with main mode resonators. The motional capacitance is about two times higher than in LFE 300 mode resonators. It means that frequency adjustment tolerance can be two times wider.

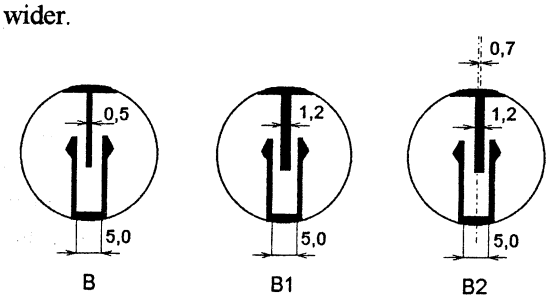


Fig. 4. The electrodes configuration for 301 mode experimental resonators excitation

Tab. 5. Parameters of 5 MHz 110 mode resonator

Mode	Frequency [Hz]	R_1 [Ω]	C_0 [fF]	C_1 [aF]	Q [E6]
110	5 129 409	117	855	240	1,36
100	5 024 050	7900			
101	5 103 607	1029			
112	5 278 022	457			
PFE_{100}	4 999 865	5,9	7452	5683	1,33
LFE_{100}	5 022 718	280	681	112	1,15

2.3. The 311 resonators

The two kinds of electrode configurations of 311 mode anharmonic resonators were presented on Joint Meeting EFTF and FCS in Besancon [2]. These electrodes shape vs. X ray topography of vibration regions comparison showed the electrodes positioning on vibrating areas. As a result one can not expect the aging effect induced by electrodes lowering. For removing electrodes from active regions it was necessary to recognize vibrations direction in every active region. The active regions shape of 311 mode is presented in fig. 5.

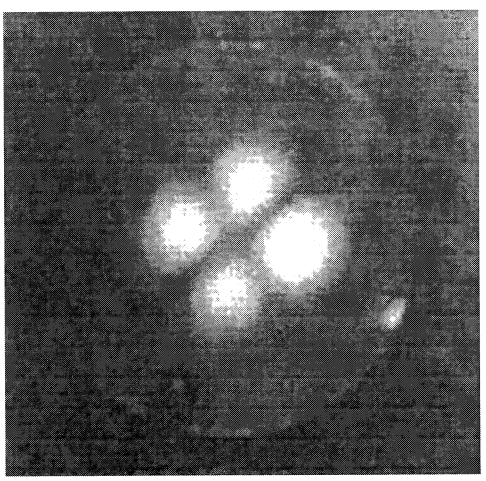


Fig. 5. 311 mode pattern [2]

For the vibration regions excitation polarity recognize the simple electrodes configuration on printed circuit board, presented in fig. 6 was designed. In the gapes between electrodes electric field directions are almost opposite. The quartz plate movement along the electrodes in central position gives excitation decay in position corresponding to excitation area centers over the gapes between electrodes. The quartz plate movement in direction perpendicular to electrodes gives significant signal increase in position where two excitation regions in the plate axis are positioned over one gap and one of other excitation region is positioned over the second gap.

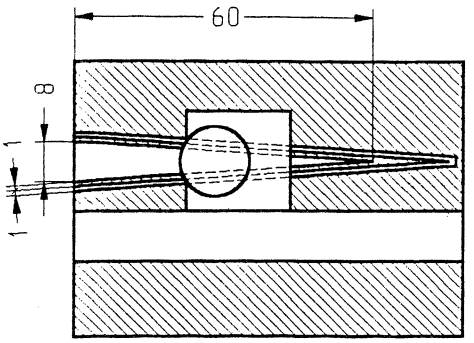


Fig. 6. Electrodes configuration for the 311 mode vibration regions excitation polarity recognize

These results show the same polarity of regions laying in one symmetry axis and opposite polarity of regions laying on perpendicular symmetry axis. On the basis of these results the four strip electrodes configuration was designed. This electrodes configuration is presented in fig. 7. The 5MHz and 10 MHz resonators, with this electrodes configuration were made. These resonators parameters are presented in tab. 6 and 7.

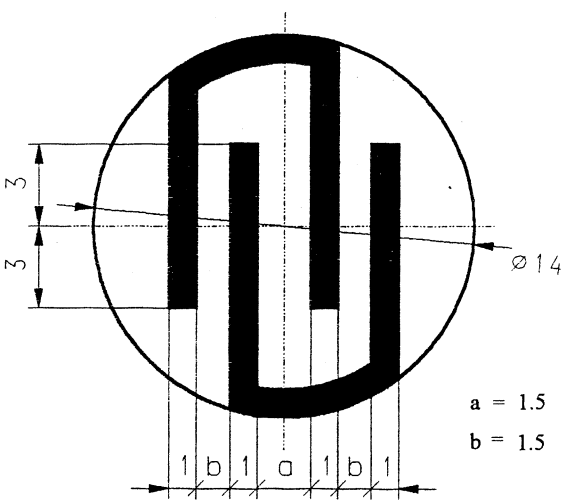


Fig. 7. The four strip electrodes configuration for 311 mode excitation

Comparing presented results with 301 mode resonators one can notice the low repeatability of electric parameters of 5 MHz 311 mode resonators. The parameters spread is wide. In one group there are existing resonators with diametrical different values of Q factor, motional resistance and motional capacitance. The Q factor values are significant lower than in classic resonators. These are the first samples of this kind of resonators. Probably it will be necessary also in this case to move electrodes from the center of plate as in 301 mode resonators. In the case of 10 MHz resonators repeatability of electric parameters is better. The Q factor values are comparable with classic resonators.

One of 5 MHz resonators with B1 (fig. 4) electrode configuration was mounted in oscillator and it's aging was measured. It exhibited long term stability in the first month 8E-11/day. After six month of aging the long term stability was 5E-11/day.

Tab. 2. The 5MHz anharmonic f_{301} resonators parameters

No	R [mm]	Electrode thickness [nm]	f_s [Hz]	R_1 [Ω]	C_1 [aF]	Q 10E6	Electrodes
		35	5 067 659	1015	28,56	1,08	fig. 4, B
18	100	50	5 067 606	933	32,29	1,11	B
		50	5 067 493	344,8	48,91	1,86	B1
		35	5 066 043	746	28,87	1,46	B
19	100	50	5 065 974	500	38,05	1,65	B
		50	5 065 784	408	42,81	1,80	B1
4	100	50	4 999 591	370,4	46,05	1,87	B2 (fig. 3)
5			4 999 773	489,5	32,66	1,99	B2
07	150	50	5 068 286	292	62,17	1,73	B2
08			5 072 546	228	68,32	2,01	B2
P22			5 072 055	270,8	61,74	1,88	B1
2			5 052 391	638,5	34,75	1,42	B
3	200	50	5 052 920	575,5	38,67	1,42	B
05			5 052 829	278,0	65,02	1,74	B2
06			5 052 841	265,0	68,02	1,75	B2
PFE_{300}	200	100	5 000 000	65-80	230	2,0-2,3	ϕ8
LFE_{300}	150	50	5 000 003	520	27	2,3	ϕ5/2,3

Tab. 3. The 8,192 MHz anharmonic f_{301} resonators parameters

No	R [mm]	Electrode thickness [nm]	f_s [Hz]	R_1 [Ω]	C_1 [aF]	Q 10E6	Electrodes
B1	500	100	8 247 977	584	28,7	1,15	fig. 1. a=0.5, b=4
B4			8 244 897	544	33,6	1,06	
B2	500	100	8 247 696	564	30,5	1,12	fig. 1. a=0.5, b=5
B3			8 248 326	511	32,6	1,16	
P3	500	50	8 247 840	420	34,6	1,33	fig. 1. a=1.5, b=4,5
14			8 191 724	334	45,5	1,28	
P5	500	50	8 248 192	255	54,0	1,38	fig. 3
P7			8 248 320	310	47,2	1,28	
P8			8 245 848	300	50,2	1,26	
P10			8 247 762	290	49,0	1,30	
PFE_{300}		100	8 192 000	40-55	260	1,2-1,55	ϕ8

Tab. 4. The 10 MHz anharmonic f_{301} resonators parameters

No	R [mm]	Electrode thickness [nm]	f_s [Hz]	R_1 [Ω]	C_1 [aF]	Q 10E6	Electrodes
P24	500	50	10 000 664	440	33,53	1,07	fig. 4, B
P25	500	50	10 000 842	350	40,56	1,08	B1
P26	500	50	9 999 125	475	28,69	1,13	
P20	1000	50	10 056 362	375	36,51	1,13	B
P21	1000	50	10 054 129	474	32,52	1,0	
PFE_{300}	1000	100	10 000 000	40-55	260	1,1-1,3	ϕ 6,5
LFE_{300}	750	50	9 999 996	650	19	1,28	ϕ 6/1,5

2.2. The 110 mode resonator

It was checked that this mode is excited stronger as 101 mode. Only one resonator on fundamental anharmonic mode 110 with electrodes presented in fig. 1 was made. The 5MHz resonator was made with curvature radius 200 mm. The parameters of this resonator are presented in tab. 5. This resonator exhibits higher Q value in comparison with classic resonator and over two times lower value of resistance and over two times higher value of motional capacitance in comparison with LFE 100 mode resonator.

Tab. 6. The parameters of 311 mode 5MHz resonators with four strip electrodes

No	R [mm]	Electrode thickness [nm]	f_s [Hz]	R_1 [Ω]	C_1 [aF]	Q [10E6]	Electrodes
P 33	100	50	5 069 758	680	37,25	1,26	fig. 7
P 54		50	5 069 913	765	31,96	1,23	
P 56	150	50	5 130 555	212	90,40	1,66	
P 58		50	5 119 171	400	54,57	1,41	
P 31	200	50	5 102 035	435	49,76	1,44	
PFE_{300}	200	100	5 000 000	65-80	230	2,0-2,3	ϕ8
LFE_{300}	150	50	5 000 003	520	27	2,3	ϕ5/2,3

Tab. 7. The parameters of 311 mode 10 MHz resonators with four strip electrodes

No	R [mm]	Electrode thickness [nm]	f_1 [Hz]	R_1 [Ω]	C_1 [aF]	Q [10E6]	Electrodes
P 63			10 002 866	369,1	40,22	1,01	fig. 7
P 64	500	50	10 000 474	346,1	42,32	1,09	
P 65			10 000 459	403,3	38,15	1,01	
P 66			10 081 499	380,2	34,44	1,21	
P 67	750	50	10 077 127	302,8	46,09	1,13	
P 68			10 074 478	328,0	41,74	1,15	
P 69			10 074 540	257,2	51,07	1,20	
P 51			10 097 127	373,3	39,98	1,06	
P 52	1000	50	10 097 315	362,3	42,50	1,02	
P 53			10 096 570	456,0	32,71	1,06	
PFE_{300}	1000	100	10 000 000	40-55	260	1,1-1,3	ϕ 6,5
LFE_{300}	750	50	9 999 996	650	19	1,28	ϕ 6/1,5

3. CONCLUSIONS

Comparing presented results with the previously presented it is possible to notice the progress in anharmonic resonators parameters, but it is not easy. The higher frequency resonators exhibit better parameters in comparison with classic resonators. A specially good parameters exhibits resonator of fundamental mode 110. The 311 mode resonators are not optimized yet. The optimization of their parameters has to be continued. This optimization is very difficult because of low repeatability of resonators parameters. Probably it is caused by very large sensitivity of resonator parameters to little changes in quartz plate and electrode geometry. On the basis of presented above experimental results one can confirm without any doubt that anharmonic mode lateral field excited resonators can reach parameters similar as classic resonators and in comparison with LFE main mode resonators they exhibit lower motional resistance and few times higher motional capacitance. The main problem in their technology is that resonant frequency of quartz plate measured in conventional electrodes configuration is not the same as resonant frequency utilized in resonator. But this problem is possible to overcome by special electrodes configuration use. The second problem is an accurate determination of electrode direction. It needs very precise quart plate positioning in the electrode depositing masks. This problem has not been solved yet and probably it is the main source of resonator parameters low repeatability.

4. REFERENCES

1. K.Weiss, W.Szulc, The SC - cut lateral field resonator working on anharmonic antisymmetric mode, Proc. of 12th EFTF, 1998
2. K. Weiss, W.Szulc, B.Dulmet, R.Bourquin, The 311 anharmonic mode vibration SC-cut resonator excited by lateral field, Proc. Joint Meeting EFTF - FCS, 1999

THE INFLUENCE OF THE SURFACE STRUCTURE ON THE VIBRATION AMPLITUDE AND THE EQUIVALENT ELECTRICAL DATA OF QUARTZ CRYSTALS - A THEORETICAL OVERVIEW

Eberhard Seydel

Alfa Quartz C.A.
Urbanizacion Industrial Rio Tuy, Parcelas 187-189 Carretera Charallave-Cua Edo. Miranda, 1210, Venezuela
Tel. +58-39-980936 / Cell-Phone +58-16-8340833 / Fax. +58-39-986856 E-mail: eseydel@telcel.net.ve

ABSTRACT

This paper will give a theoretical overview of the influence of different surface parameters on the behaviour of the quartz crystal components. It is possible to calculate different types of crystal shapes with various structures of the surface. The results give tendencies and very important values for the tolerances in blank production to keep the specified values of the final product, e.g. dips, C1-tolerances or the attenuation of inharmonic modes.

1. INTRODUCTION

It is very important for the blank manufacturer to know which mechanical tolerances are usable to determine electrical parameters of the final crystal.
A FDM (finite differences method) program was written which calculates the influence of different surface parameters on the behaviour of the quartz crystal components with high resolution. In this way it is possible to calculate different types of crystal shapes, mean plane, bevelled, convex or special structured blanks (e.g. inverted mesa) with various structures of the surface, mean roughness, wedge or concave surface.

2. PROGRAM DESCRIPTION

The program has been developed over the time, see [4] to [8]. The basis for the calculation program is the calculation of vibration-modes of quartz crystal resonators with different dimensions and surface contours. It is based on the publication of Beaver [1].

The starting points are the differential equations

$$h(\Psi_1 + \Psi_{2,1}) - (\rho h^3 \omega^2 / D_{60})\Psi_1 = 0 \qquad (1)$$

$$(h(\Psi_1 + \Psi_{2,1}))_{,1} - (\rho k h \omega^2 / D_{60})\Psi_2 = 0 \qquad (2)$$

and the boundary conditions

$$\Psi_{1,1} = 0 \quad \text{at} \quad x_1 = \pm d/2 \qquad (3)$$

$$\Psi_1 + \Psi_{2,1} = 0 \quad \text{at} \quad x_1 = \pm d/2 \qquad (4)$$

In this way it is possible to calculate the vibrations of crystal resonators of various shapes, of different cut angles, with different electrode dimensions and at different temperatures. The input parameters are the mechanical values of the mounted finish blank. They are :

- shape factors including:
  - thickness
  - electrode diameter and material
  - surface roughness (different models)
  - surface wedge (plano-parallel)
  - surface concave
- cut angle (three axis rotations available)
- temperature

The output values of the program are:

- eigenvalues and frequencies
- amplitude distribution
- relative amplitude ratio per mode
- relative C1 (integrated amplitude) per mode
- graphical display

The program works in steps and adds the mechanical values step by step to the structure (i.e. starting from shape type (plano, convex, mesa, etc.), add roughness, add wedge, add concave part, add electrode on the surface, add holder damping). For the determination of the cut angle, a three dimensional tensor rotation and the temperature dependency of the physical constants are used. Calculation accuracy must be better or equal the real*8 notation of Fortran. In the case of lower accuracy the resolution of the calculation is not sufficient.

The calculated values are comparable with the measured values. Although there is an offset between calculation and measurements, the results are useful for calculation of tolerances, for the direct comparison in one or two points and for the adjustment of ranges.

The structure of the calculated elements is very important for such types of FEM or FDM calculations. For the calculations with shear vibrations, the ratio of

thickness to width of the element should be larger than 2:1 to obtain good results. In the case of smaller ratios, the stiffness of the shear modes is too high.
This is necessary to provide adequate minimum resolution for different blank dimensions, e.g. 17 MHz blank with 8 mm diameter, mesh width 0.05mm, 160 meshes in one direction.
The program is written in FORTRAN and uses a DIGITAL Fortran Compiler. The resolution is adjustable in the start-phase of the program and depends on the used type of computer. It runs on normal PC's.

3. CALCULATIONS AND THEIR RESULTS

Different results for plano blanks and contoured blanks are published in [5] and [6]. For more information on the influence of the surface roughness see [7] and on inverted mesa blanks [8]. Mechanical values of the blank for the best energy trapping, temperature frequency behaviour and inharmonic mode damping are determined. In addition, calculation of the influence of wedge and concave blank forms is also computed.

An example of an 8mm diameter 0,1 mm thick AT-blank with a calculation in one direction is used to show the utility of the program. In this case, face shear modes are not included. The effects of the wedge shape, the roughness, concave part and the effect of electrode thickness on concave blanks are included in the calculations. Figures 1 to 3 show mode chart dependencies for different diameters and also the temperature dependency. These values determine the tolerances of the diameter or width and length, respectively. The blank dimensions diameter, electrode and mean thickness in the following calculations correspond to the blank of Figure 2. Figures 4, 5 and 6 show the influence of a wedge, of roughness, or of a concave part, respectively, on the temperature mode chart of the blank.
The wedge influences the amplitude deviation of the main mode and the inharmonic modes and also the values of R1 and C1 and the damping of the inharmonic modes. The roughness influences also the R1 and C1 values and the amplitude ratio of the inharmonic modes and the main mode, but the frequency mode charts are not changed significantly. The surface roughness influences over scattering effects on the surface also the temperature frequency curves (dips) [7], [8]. Figure 7 shows the mode chart dependent on the concave part for a blank with a fixed film thickness. In this case the coupling parameters for the unwanted mode show a significant change versus small concave variations this means that the blank is very sensitive for dips. Figure 8 shows the mode chart of this blank for different film thickness. The influences on the inharmonic modes and the instability of the vibration in the range of thin film thickness are visible very well. In Figure 9, the normal amplitude deviation of the main mode and the first inharmonic mode for this blank without concave part is shown. In Figure 10, it is seen that the concave part influences these modes and Figure 11 shows the last stage for a normal vibration. For thinner films the main mode is splitted and not usable. In that case the blank will run on an inharmonic mode with high R1 values. Figures 12 to 14 show the dependencies of two modes from the Z- and Y-rotation of an AT-blank. This influences very strongly the coupling parameters (dips) and is caused by the cutting process. Table 1 shows principal dependencies of the electrical values of the finished crystal on the mechanical parameters, possibilities to use the FDM-Calculation, and the determined mechanical tolerances.

4. CONCLUSIONS

Table 1 shows the conclusions for the production tolerances in the blank manufacturing area. Special values can be fixed by production parameters, like frequency and electrode. Others are used to calculate each new specification. This program is a very good tool in the design phase of a new blank with delicate specification in the electrical parameters or in the temperature frequency behaviour. Test calculations have shown that this program is also useful for other cuts such as SC. Also is it possible to calculate the influence of the material, like changing of constants or etch channels as well.

References

[1] W.D. Beaver DESIGN EQUATIONS FOR BI- AND PLANO-CONVEX AT-CUT RESONATORS 27TH ANNUAL FREQUENCY CONTROL SYMPOSIUM 1973 P.11
[2] R.Lerch SIMULATION PIEZOELEKTRISCHER KÖRPER MIT DER FINITE- ELEMENT-METHODE FORTSCHRITTE DER AKUSTIK -DAGA '86 P.797
[3] J.Stoer EINFÜHRUNG IN DIE NUMERISCHE MATHEMATIK SPRINGER-VERLAG 1979
[4] E.Seydel et.al. ELEKTRISCHE EIGENSCHAFTEN VON SCHWINGQUARZEN BEI VARIABLEN BETRIEBSBEDINGUNGEN PIEZO 1988
[5] Seydel E., Priess U.,1994, OPTIMIZATION OF QUARTZ CRYSTAL DESIGN FOR TCXO- AND VCXO-APPLICATIONS; EFTF 1994
[6] Seydel E., Priess U.;1994; QUARTZ CRYSTAL FOR MICROBALANCE; PIEZO'94
[7] Seydel E., Landmann A.,1995, THE INFLUENCE OF SURFACE STRUCTURE ON AMPLITUDE DEVIATION AND ELECTRICAL PARAMETERS OF QUARTZ CRYSTALS; EFTF 1995
[8] Seydel,E., Weitzel,L., 1996, HIGH FREQUENCY FUNDAMENTAL MODE RESONATORS; EFTF 1996

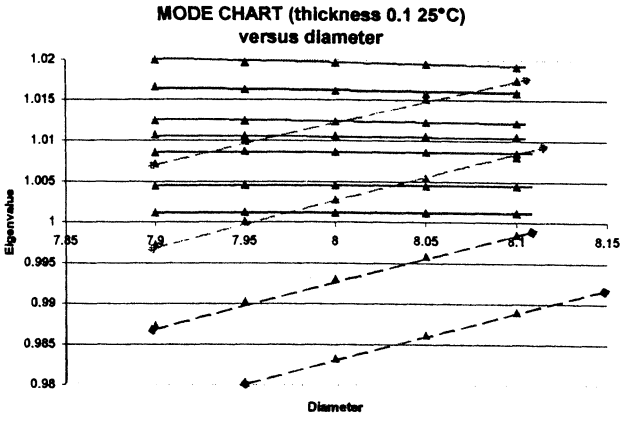


Figure 1

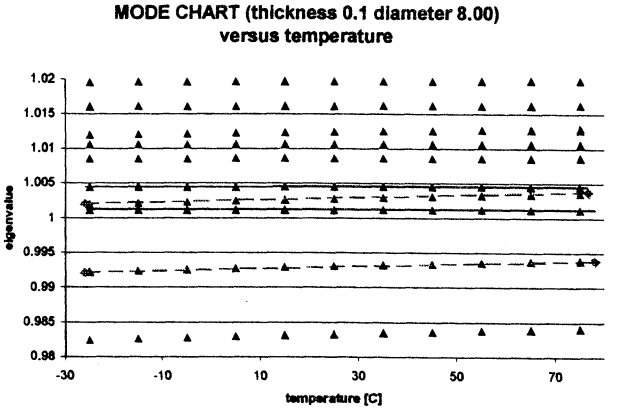


Figure 2

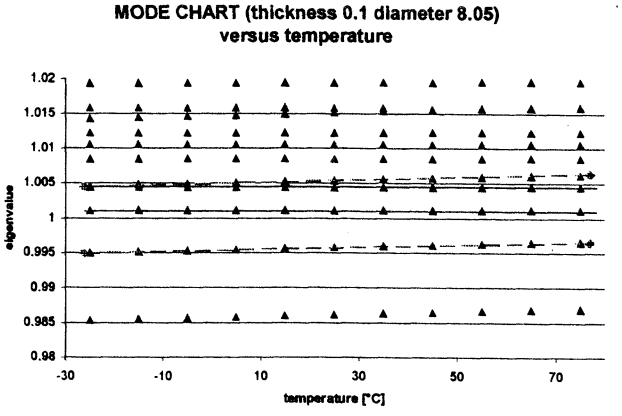


Figure 3

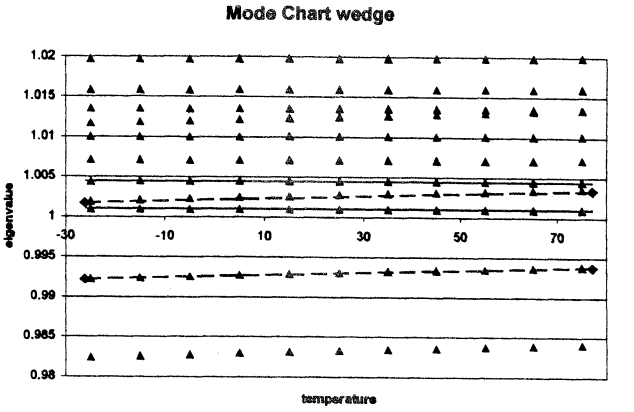


Figure 4

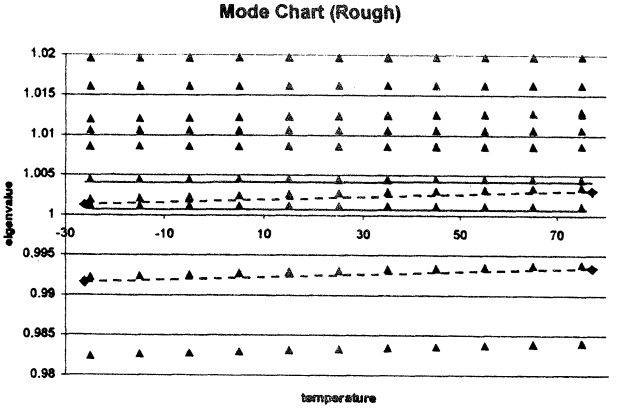


Figure 5

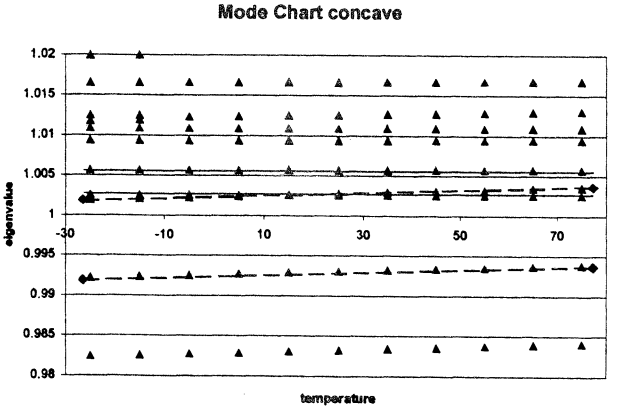


Figure 6

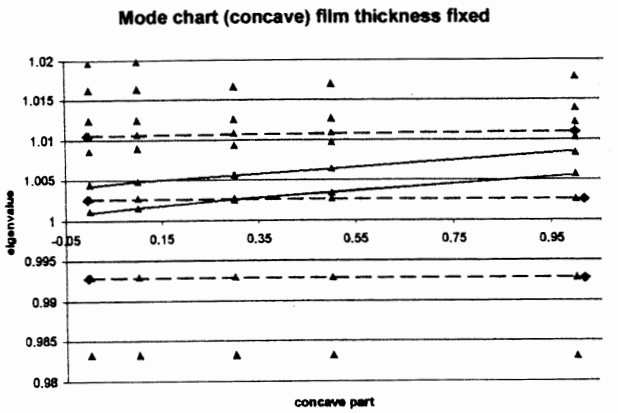


Figure 7

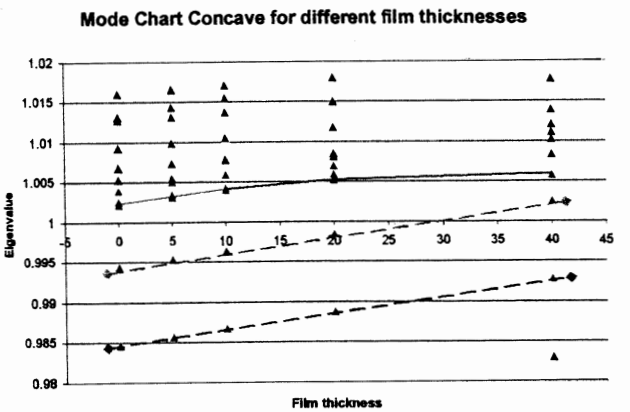


Figure 8

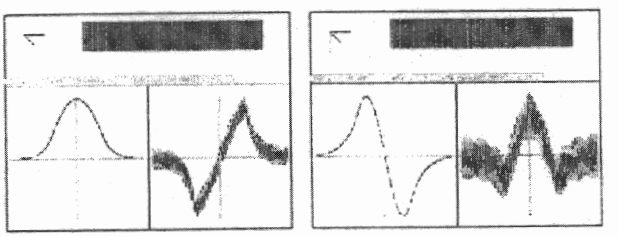


Figure 9

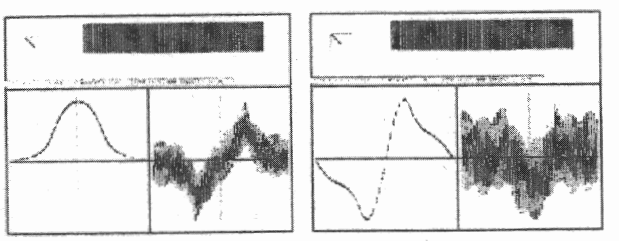


Figure 10

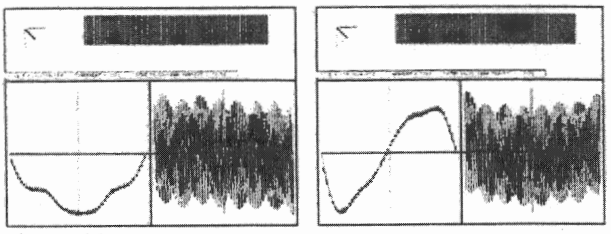


Figure 11

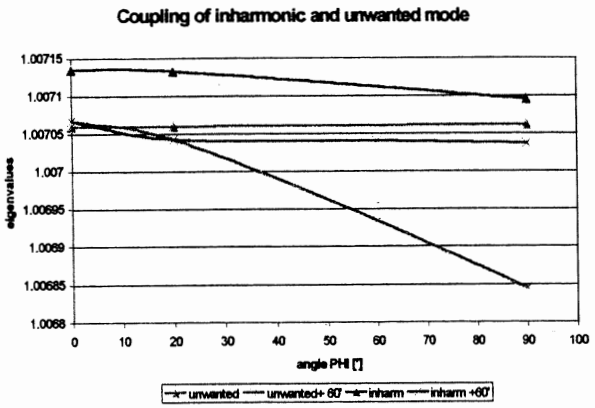


Figure 12

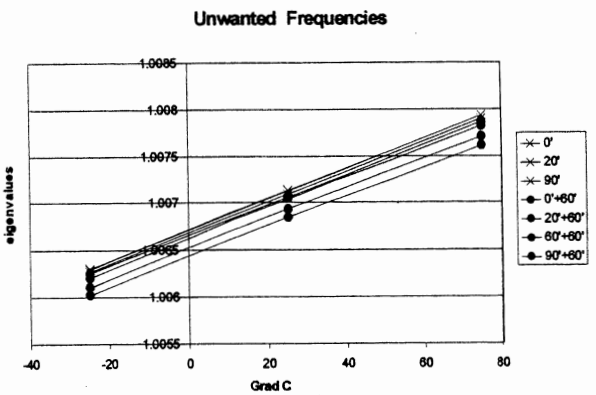


Figure 13

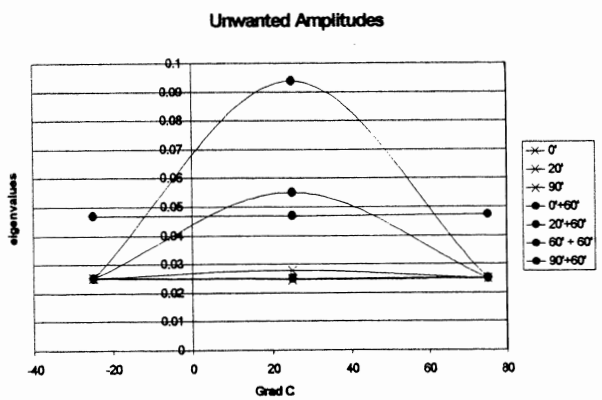


Figure 14

Table 1 Principal dependencies and results

Mechanical parameter	Electrical effect	Calculated values	Production Tolerances
Mechanical dimensions			
Diameter or width /length for strips	Temperature frequency (cut angle) and dips and Q-value	Mode chart and amplitude deviation	20 % of the thickness
Thickness (blank frequency-plate back)	Temperature frequency dips Inharmonic modes	Mode chart	Thickness tolerance < 10 % of F^2 of the electrode
Shape and surface			
Shape (plano, bevel, lenses)	Q-value C1 Inharmonic modes Dips Temperature frequency	Amplitude deviation Mode chart	To be calculated
Flatness (wedge ,concave)	Inharmonic modes C1 Q-value	Amplitude deviation Mode chart	Wedge < 1% blank thickness Concave < 10% equivalent filmthickness
Surface roughness	Q-value C1 Inharmonic modes Coupling parameters for dips	Amplitude deviation	< 1% blank thickness divide by overtone
Electrode			
Electrode diameter	C1 Inharmonic modes Coupling parameters for dips	Amplitude deviation Mode chart	To be calculated
Electrode material (mass, stress)	Temperature-frequency Stress dips	Only mass dependency	To be calculated
Mounting			
Mounting points	Temperature-frequency Dips	Amplitude deviation with damping points	To be calculated
Material			
Cut angle	Temperature frequency Dips	Mode chart	To be calculated AT (Z 20' / Y 20')

MEASUREMENTS ON DOUBLY ROTATED QUARTZ RESONATORS WITH VERY LOW ISOCHRONISM DEFECT

N. Gufflet, R. Bourquin, J.J. Boy

ENSMM / LCEP – 26 Ch. de l'Epitaphe – 25030 BESANÇON Cedex – France
Fax : 33 381 88 57 14 – E-mail : ngufflet@ens2m.fr, rbourquin@ens2m.fr, jjboy@ens2m.fr

ABSTRACT

This paper deals with the study of quartz crystal cuts with low isochronism defect. We have already shown that it exists at least one cut, working in C-mode, for which the isochronism defect is practically null. The C-mode is thermally compensated in order to use the resonator in Ultra Stable Oscillator for Time-Frequency applications.

Here, we entirely characterize this new cut (called *LD-cut*), mounted in a BVA structure. We justify the choices of the design parameters and compare theoretical and measured electrical parameters.

Further, we present measurements of isochronism defect on other temperature compensated doubly rotated cuts working on B-mode.

Key words : Isochronism defect, contoured resonator, doubly rotated cut, BVA resonator.

1. INTRODUCTION

In a previous paper [1], we have checked the $R^{-1/2}$ dependence of the amplitude frequency effect of plano-convex quartz resonators versus the radius of curvature and we have found that existing theoretical investigations give a relatively good description of this effect except in the case of too much trapped vibrations. Moreover, we have pointed out that it exists at least one temperature compensated cut working on C-mode with a very low isochronism defect.

In the first part we present new resonators which have been manufactured in BVA technology with cut angles close to those for which the isochronism defect is null. This new cut (called *LD-cut* standing for Low isochronism Defect) has been investigated in order to know precisely its motional parameters as well as all characteristics necessary to use it in an Ultra Stable Oscillator.

In the second part we study temperature-compensated resonators operating on B-mode because we wonder whether it exists an other cut with low isochronism defect. The θ angle of these cuts ranges in [-47°,-23°] (according to the IRE 49 convention). We report results about their measured isochronism defect and thermal sensitivity.

2. CHARACTERISTICS OF THE NEW *LD-CUT*

2.1 Design Parameters :

We have reported [1] some measurements on the isochronism defect of thermally compensated C-mode resonators for cuts with a φ angle ranged between SC- and X+30-cuts (with θ close to +34°). The curve of the isochronism defect versus φ cross the zero-axis for a value close to 27° (see Figure 1).

Numerous *LD-cut* BVA resonators, designed in order to work on the 3rd overtone, have been manufactured. Their mean parameters are compared with these of the well-known SC-cut.

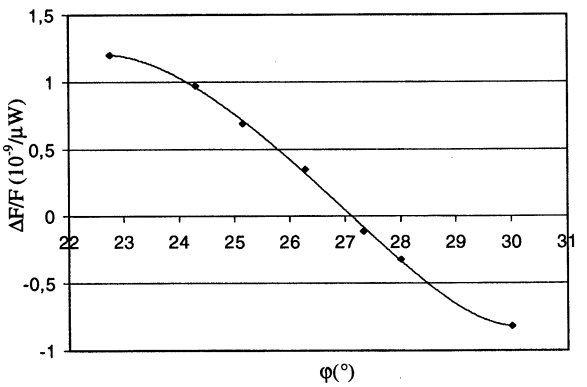


Figure 1 : Isochronism defect versus φ

To obtain a resonant frequency at 10 MHz for the C-mode third overtone, the calculation of the effective elastic constant leads to a thickness of 0.563 mm, slightly higher than this of the SC-cut (0.559 mm).

For convenience reason, we have decided to keep the same values of the vibrating part and electrodes diameters than these of the SC-cut currently manufactured at the laboratory, i.e. an inside diameter of 10.6 mm (which is the diameter of the vibrating part) and an electrode diameter of 6 mm.

One of the key advantages of the BVA resonator is the possibility of choosing the orientation of the four "bridges" in order to reduce the frequency shift due to the in-plane stresses induced by the mounting structure. These small bridges connect the vibrating part of the crystal and the dormant part and it is well known that those of the SC-cut resonators are aligned along the projections of X- and Z-axes in the plane of the cut. For the *LD-cut*, we have calculated the Ratajski coefficient

K_f versus the azimutal angle ψ. The curve, shown on Figure 2, computed from [2], indicates that the best orientations for the two pairs of bridges are 0 and 80°.

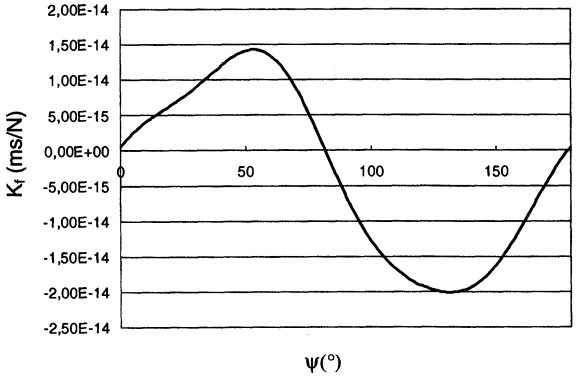


Figure 2 : Radial stresses sensitivity

So, the design of the new *LD-cut* is such as indicated on the Figure 3.

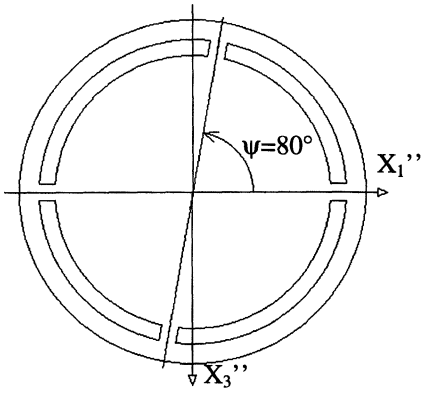


Figure 3 : Position of bridges

The well known theory of energy trapping [3] has been used to calculate the radius of curvature in order to have an optimal trapping of vibration, i.e. to obtain a vibration at least 10^{-6} times lower in the periphery of the vibrating part than in the center, and the calculation of the M_n and P_n coefficients [3] leads respectively to radius of curvature of 480 mm and 320 mm. So, the best compromise (depending too of our mechanical possibilities) is R_c = 300 mm which leads to a theoretical motional resistance of 125 Ω. Our choice as also been leaded by the fact that a too much trapped vibration leads to an higher motional resistance.

2.2 Resonator Properties

Six *LD-cut* BVA resonators have been realized and enclosed under vacuum in a HC-40/U can. The typical value of their motional resistance and the frequency spectrum around the 3rd overtones of B- and C-modes are presented in table 1, in which we have also included the resonance frequencies of the SC-cut for comparison purposes.

Table 1 : Frequency spectrum of *LD-* and SC-cuts

	LD-cut		*SC-cut*	
Mode	*F (kHz)*	*R (Ω)*	*F (kHz)*	*R (Ω)*
C 300	10000	138	10000	85
C 320	10107	266	10128	130
C 302	10122	276	10140	140
C 340	10220	-	10258	-
C 322	10228	-	10267	-
C 304	10245	-	10282	-
B 300	10755	60	10914	55

We should note that B- and C-modes are nearer on *LD-cut* than on SC-cut.

We have performed an X-ray topography of the C-mode 3rd overtone to confirm that the vibration is correctly trapped (drive level +6 dBm). As we can see on the Figure 4 the vibration is well confined in the middle of the resonator, preventing from edge effects.

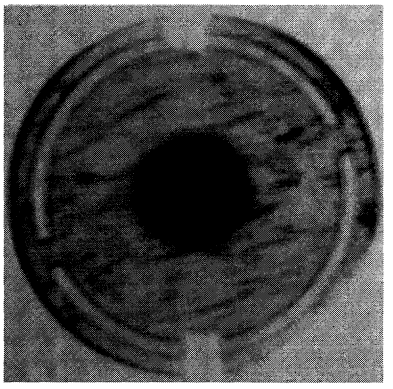


Figure 4 : X-ray topography of the C300

The motional parameters of the C- and B-modes, 3rd overtone, are reported on Table 2. As comparison we have also included the motional parameters of the same modes for the SC-cut.

Table 2 : Motional parameters of the C300

C300	*R (Ω)*	*L (H)*	*C (fF)*	*C_0 (pF)*	*Q (*1.E6)*
SC	85	1.76	0.14	2	1.3
LD	138	2.85	0.09	2.9	1.3

The motional resistance is a little bit greater but the self too and then the quality factor stay very good.

Below (figure 5), we present the frequency-temperature curve, in the range [20, 110 °C], of a *LD-cut* (with a cut angle adjusted to obtain a turnover point at 80°C) which is, here too, compared to the SC-cut one. If the ΔF between 25°C and the turnover temperature is close to 200 Hz for the SC-cut, it is rather close to 300 Hz for the *LD-cut*. So, the curve is

not so "flat" around the turnover point but, since the inflexion point is located at an higher temperature, resonators with higher turnover point can be made.

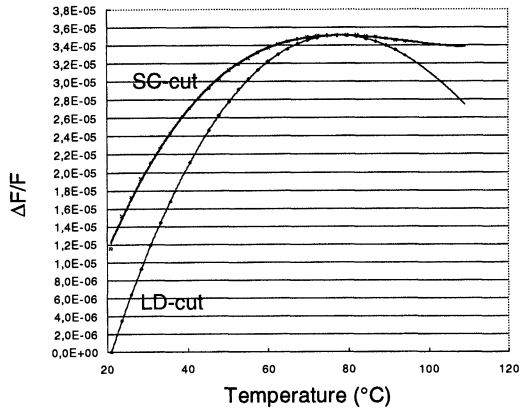


Figure 5 : Frequency-temperature curves for *LD*- and SC-cuts

2.3 Isochronism defect :

Precise measurement of the isochronism defect of the C-mode 3rd overtone is not easy because of its very small value. Indeed thermal effects and precision of the bench make the measurement difficult. But we are able to insure that the absolute relative frequency shift is smaller than :

$$5.10^{-11} / \mu W$$

These measurements have been performed with a power on the resonator ranged between 10 to 500 μW. So, the drive level sensitivity of this cut is 20 to 30 times lower than this of the SC-cut. Moreover the power limitation due to the instability of the resonance at high level is about 7.5 mW instead of 1.3 mW for the SC-cut if the resonators have the same trapping of vibration.

3. EXPERIMENTAL STUDIES OF OTHER CUTS

3.1 Resonators Parameters

In order to study the isochronism defect, we have been interested in cuts working on thermal compensated modes.

It is well known that the temperature frequency dependence can be fitted by the cubic law:

$$\Delta F/F=a\Delta T+b\Delta T^2+c\Delta T^3 \qquad (1)$$

where $\Delta T=T-T_0$, T_0 being the temperature reference (usually 25°C). A mode is said temperature compensated when the a coefficient is close to zero. Only the B- and C-modes can be compensated. The Figure 6 shows the loci a = 0 of these 2 modes for φ ranged in [45, 60°]. It should be noted that for convenience we have taken φ in the range [45,60] and θ in [20,50] instead of [0,15] and [-50,-20] respectively as it should be according to the usual convention, but due to the symmetries of the quartz crystal, both ranges are the same.

We have also indicated, for some particular cuts, the ratio between the frequencies of the C- and B-modes and the electromechanical coupling factor k^2 (in 10^{-3}).

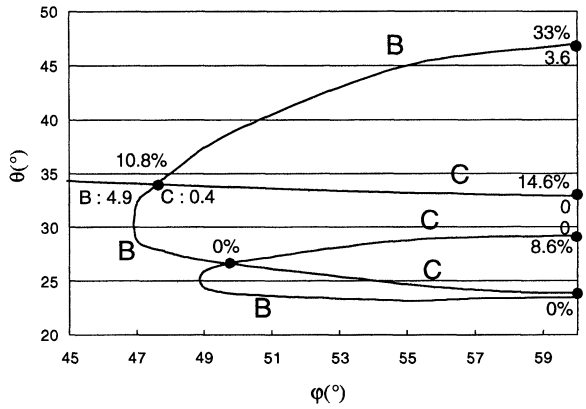


Figure 6 : Loci of temperature compensated C- and B- modes

Since the C-mode coupling factor is very low for those cuts (null for φ = 60°), we have been just interested in the B-mode.

Moreover the small frequency shift between B- and C-modes for compensated cut with θ<30° impose us to limit our measurements on cuts with θ>30°.

So, following the above discussion, we have chosen 4 orientations on the upper part of the B-mode locus with θ ranged in [34°, 47.5°].

For manufacturing convenience, the resonators of each orientation are electroded. The electrodes and active part diameters are respectively of 7 and 13.2 mm for all resonators. The radius of curvature has been calculated to trap the 3rd overtone. Table 3 gathers design parameters values of these resonators.

Table 3 : Resonators Parameters

Res.	*φ (°)*	*θ (°)*	*e (mm)*	*R_c (mm)*	*$\varnothing_e$ (mm)*	*$\varnothing_a$ (mm)*
1	47.6	34	0.67	300	7	13.2
2	50	39.1	0.56	100	7	13.2
3	55	45.2	0.67	300	7	13.2
4*	60	45.3	1.26	150	7	13.2

(* The resonator #4 is a BT-cut resonator and usually defined as a single rotated cut by θ = -45.3°.)

3.2 Resonators Properties

We have studied these resonators by measuring their resonance frequencies, motional resistances, quality factors Q, positions of the turn-over points and anisochronisms. Results are presented in Table 4 where we have also included the first order temperature coefficient a.

Table 4 : Resonators Properties

Res.	*F* (kHz)	*R* (Ω)	*Q* (10^6)	*T.O.P.* (°C)	*a* (10^{-6})	*I.D.* (10^{-9}/μW)
1	10365	22	1.2	25	-0.1	-0.6
2	12910	38	1.8	50	2.0	-1.1
3	11215	23	1.5	50	2.0	-0.7
4	6105	49	2.2	70	3.94	-0.3

An expression of the isochronism defect is given by (see [1]) :

$$\frac{\Delta F}{F} = \frac{\sqrt{2}\rho\alpha P}{\pi^2 h^{1/2} n (M_n P_n)^{1/4} \eta^{(1)} \bar{c}^{(1)3/2} R_c^{1/2}} \quad (2)$$

where

$$\alpha = \frac{9}{64}\gamma - \frac{\bar{c}^{(1)}}{4}\left(\beta_2^2 \nu_2 + \beta_3^2 \nu_3\right) - \frac{53}{512}\frac{\beta_1^2}{\bar{c}^{(1)}} + \frac{h\bar{c}^{(1)}}{n^2\pi^2}\left(\beta_2^2 \zeta_{n2} + \beta_3^2 \zeta_{n3}\right) \quad (3)$$

γ is an intrinsic constant of the cut for a given mode. The other terms are defined in reference [4]. γ include the fourth order elastic coefficients and is the preponderant term in the expression of α, however the loci of γ=0 are not exactly the same than those for which the α (and then the isochronism defect) is null.

For the realized cuts, we have calculated the γ and α coefficients for a resonator working on the 3rd overtone with a thickness of 0.5 mm. Results are summarized in Table 5.

Table 5 : γ and α coefficients

Res.	*1*	*2*	*3*	*4*
γ (10^{12} N/m²)	-0.53	-0.56	-0.69	-0.57
α (10^{11} N/m²)	-0.78	-0.81	-0.99	-0.84

The corresponding plots of γ and α versus φ are shown on figure 7. We can see that according to the α coefficient there is no thermally compensated cut on B-mode that exhibit an anisochronism null for θ>34 °.

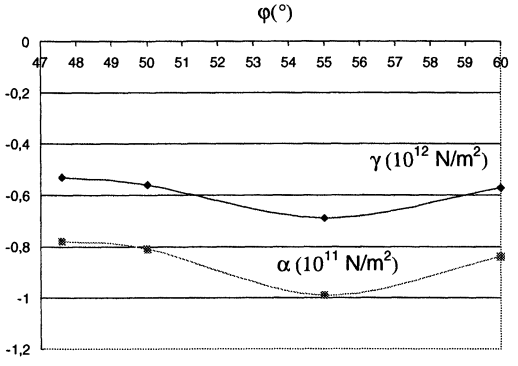


Figure 7 : γ and α coefficients versus φ cut angle

4. CONCLUSION

The properties of the new *LD*-cut resonators manufactured in a BVA structure are very correct and it will certainly be possible to make very good oscillators with it. For the moment test in oscillators has just been made with prototypes and the short term stability has been measured close to 3.10^{-13} at 1s. The phase noise of a *LD*-cut resonator has also been measured (passive method) and leads to a typical value $\sigma_y(\tau)$ of $1.5.10^{-13}$ at 1s (Allan variance). Moreover the high value of the limitation power for the LD-cut (7.5 mW) present at least a great interest for radar-type applications. Indeed it is well known that the long term stability is worth for high driving level but for those kind of applications only the short term stability is of interest and then the increasing of the driving level can improve it.

Although measured isochronism defect are not so high, the study of new cuts has shown that there is no usable cuts with null isochronism defect on B-mode.

ACKNOWLEDGMENTS

Authors wish to thanks Eric Andrey, Séverine Michel and Daniel Thiebaud for resonators manufacturing and Jacques Groslambert from LPMO for phase noise measurements on resonators.

REFERENCES

[1] N. Gufflet, R. Bourquin, J.J. Boy, "Isochronism defect for various doubly rotated cut quartz resonators", Proc. Joint Meeting EFTF-IEEE IFCS, p784-787,Besançon 1999

[2] M. Mizan & A. Ballato, "The Stress Coefficient of Frequency of Quartz Plate Resonators", Proc. 37 th AFCS, p194-199,1983

[3] 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

[4] H.F.Tiersten, D.S.Stevens, "An analysis of non-linear resonance in contoured quartz resonators", J.A.S.A., Vol 80, (4), pp 1122-1132, Oct. 1986

Direct mapping of the vibration modes of piezoelectric devices using white beam stroboscopic topography at ESRF.

B. Capelle, J. Détaint & Y. Epelboin
LMCP, UMR 7590 CNRS, Universités P. et M. Curie and D. Diderot, IPGP
Case 115, 4 place Jussieu 75252 Paris Cedex 05, France

X-ray topography is one of the most powerful techniques to assess the vibration modes of the piezoelectric devices [1] widely used for frequency generation and filtering. We have previously shown [2] that polychromatic beam with an extremely small divergence as available at third generation synchrotron sources such as ESRF, allows the direct determination of the mode shape of devices having a vibrating amplitude in the range actually used in most of their applications ranging from some μW up to several mW. The width of the diffraction image obtained on the film in the case of section topography, is proportional to the local rotation of the lattice plane, thus we obtain a direct recording of the vibration amplitude.

The theoretical analysis of this technique will be detailed in the communication together with the results concerning plano-convexe resonators with different parameters (overtones rank and constitution parameters). Maps of the amplitude of the vibration along the surface of the device will be shown. They will be precisely compared with computed ones.

This method will also be applied to determine the amplitude of vibration in surface wave devices. The case of the two diffraction geometry will be examined. Rayleigh wave propagating on Y cut lithium niobate will be considered for the Bragg geometry and we will present the results of a direct and quantitative measurement of the vibration amplitude. In a similar maner as for the bulk wave, the amplitude is extracted from the value of the deflexion of the diffracted beam due to the vibration around its position at zero displacement. For the Laue geometry we will consider the case of ST quartz devices and present a study of the influence of several parameters on the diffracting image (amplitude, ratio of the acoustical wave length to the thickness of the plate).

[1] A. Zarka, B. Capelle, J. Détaint & J. Schwartzel (1987) Proc. 41st Annual Frequ. Contr. Symp. pp 236-240, NY, USA
[2] Capelle B., Détaint J & Epelboin Y. (1999), to be published.

THEORY OF MICROWAVE PARAMETRIC EFFECTS IN BULK ACOUSTIC WAVE COMPOSITE RESONATORS

G.D.Mansfeld, N.I.Polzikova, I.G.Prokhorova
Institute of Radioengineering and Electronics, Russian Academy of Sciences
Mokhovaya 11, 103907, Moscow, Russia

ABSTRACT

The theory of parametric amplification of acoustic waves in bulk acoustic composite resonator has been developed. Two mechanisms of nonlinearity such as AC controlled boundary conditions and nonlinear piezoeffect were analyzed. Numerical calculations revealed the possibility of parametric amplification of acoustic waves in structures involved.

1. INTRODUCTION

This work is devoted to the investigation of different mechanisms of parametric excitation and interaction of acoustic oscillations and electromagnetic fields in bulk acoustic wave (BAW) resonators containing a piezoelectric layer or a plate as an acoustic load. The interaction between the acoustic oscillations and microwave fields applied to the load - piezoelectric layer - occurs due to the change in the electrical boundary conditions and physical phenomena appearing in the piezoelectric media in an AC electric field.

In a nonlinear piezoelectric media subjected to time periodical electric field or in a media with periodically changing boundary conditions the waves of alternating electric fields both in time and space are exited. Some of these electrical field waves may enhance additional damping or gain of acoustic waves. It happens under synchronism in space and time between electric field waves and acoustic modes (parametric resonance).

It was found that the input electric impedance of the composite resonator depends on the frequency and amplitude of the external electric load variations (in practice it can be realized by varactor load controlled by applied AC voltage). Parametric resonance takes place when the frequency of the applied voltage Ω equals $2\omega_m/n$, where ω_m is the frequency of one of the composite resonator modes, n is an integer. It is the amplitude of the voltage that controls the rearranging of the spectrum of the resonator. At some amplitude of the voltage parametric amplification of acoustic waves may take place and in this case the sign of the real part of the input electric impedance changes.

The parametric phenomena also take place when electric field is directly applied to the piezoelectric media. In this case the propagation parameters of the acoustic waves are periodically changed by the field via non-linear piezoeffect.

The parametric phenomena considered effect the operating parameters of bulk acoustic wave resonators and they can be used for the development of parametric acoustoelectron amplifiers and generators for microwaves.

2. PARAMETRIC EFFECTS DUE TO VARIATION OF BOUNDARY CONDITIONS

The model of BAW composite resonator made of two piezoelectric layers is shown in Fig.1. It consists in essence of two piezoelectric transducers. One of them (1) - the electrical input of the structure - is connected with the source of signal of frequency ω. The second one (2) – output - is connected with an electrical capacitor. Its capacitance depends on applied voltage and changes periodically with frequency Ω as $C(\Omega) =$, $C_0(1+m\cos\Omega t)$, where m is a modulation amplitude. Metallic electrodes are located at x = $-l$, 0, d and considered as infinitesimally thin with ideal conductivity.

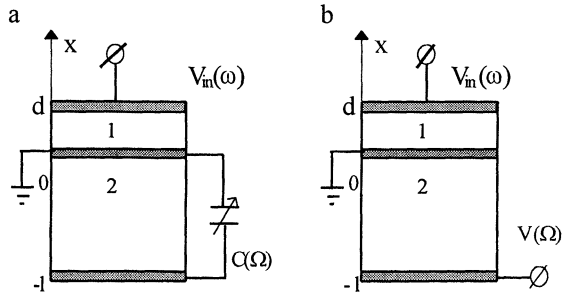


Fig.1. The schematic diagram of the BAW composite resonators: *a* - with electrically loaded by AC controlled capacitance; *b* - subjected into alternative electric field

The calculation of an electrical input impedance of this structure is based on the solution of the system of equations: wave equation for mechanical displacement u and Poisson equation for electrical displacement D [1,2]

$$\rho^{(s)}\frac{\partial^2 u_i^{(s)}}{\partial t^2}=\frac{\partial T_{ij}^{(s)}}{\partial x_j},$$

$$\frac{\partial D_m^{(s)}}{\partial x_j}=0,$$

$$I_i^{(s)}=S\frac{\partial D_i^{(s)}}{\partial t}. \qquad (1)$$

where

$$T_{ij}^{(s)} = C_{ijkl}^{(s)} u_{kl}^{(s)} - e_{ijk}^{(s)} E_k^{(s)} - e_{ijklm}^{(s)} u_{kl}^{(s)} E_m^{(s)},$$

$$D_m^{(s)} = \varepsilon_{mn}^{(s)} E_n^{(s)} + e_{ijm}^{(s)} E_{ij}^{(s)} + \frac{1}{2} e_{ijklm}^{(s)} u_{ij}^{(s)} u_{kl}^{(s)}, \quad (2)$$

Here T_{ij}, C_{ijkl}, e_{ijk} and e_{ijklm} are stress, elastic, piezoelectric and nonlinear piezoelectric tensors respectively, u_{ij} and ε_{mn} are deformation and dielectric permeability tensors; $I_i^{(s)}$ is an electric current, S is a beam crossection (transducer aperture), $i,j,k,l,m,n=1,2,3$. Superscript $s=1$, indicates the layer involved.

These equations should be added by appropriate mechanical boundary conditions corresponding to continuity of stress and deformation on interface between layers 1,2 and the absence of stresses on free boundaries

$$T_{ij}^{(1)}(d) = T_{ij}^{(2)}(-l) = 0,\ T_{ij}^{(1)}(0) = T_{ij}^{(2)}(0),$$
$$u_i^{(1)}(0) = u_i^{(2)}(0). \quad (3)$$

From here now we shall consider 1D problem and neglect nonlinear piezoeffect. Then

$$u_i^{(s)} = u_x^{(s)} = u^{(s)},\ E_i = E_x = E,\ T_{ij}^{(s)} = T_{xx}^{(s)},$$
$$e_{ijm}^{(s)} = e_{xxx}^{(s)} = e^{(s)},\quad e_{ijklm}^{(s)} = 0,\quad C_{ijkl}^{(s)} = c_0^{(s)},$$

where x is a coordinate in the direction of wave propagation. In this case the Eqs.1 will take a form

$$\rho^{(s)} \frac{\partial^2 u^{(s)}}{\partial t^2} - c^{(s)} \frac{\partial^2 u^{(s)}}{\partial x^2} = 0,$$
$$\frac{\partial D^{(s)}}{\partial x} = 0,\ I^{(s)} = S \frac{\partial D^{(s)}}{\partial t}, \quad (1')$$
$$c^{(s)} = c_0^{(s)}(1 + k^{(s)2}),$$

where $k^{(s)2} = (e^{(s)})^2 / (\varepsilon c_0^{(s)})$ is electromechanical coupling constant.

Voltages across the layers are

$$V^{(2)} = \int_0^{-l} E^{(2)} dx,\ V^{(1)} = \int_0^{d} E^{(1)} dx, \quad (4)$$

The current in branch 2 depends simultaneously on $C(\Omega)$ and $V^{(2)}$

$$I^{(2)} = \frac{\partial \left(C(\Omega) V^{(2)} \right)}{\partial t}. \quad (5)$$

Eqs.4,5 are equivalent to electrical boundary conditions. The product of $C(\Omega)$ and $V^{(2)}$ produces nonlinearity connecting acoustic and electric subsystems and generating excitations of the system on combined frequencies $\omega_n = \omega - n\Omega$. Then the solutions of Eqs.1',2 will be found in the form

$$u^{(s)} = \sum_n e^{-i\omega_n t} \left(u_n^{+(s)} e^{iq_n^{(s)} x} + u_n^{-(s)} e^{-iq_n^{(s)} x} \right) + c.c$$

$$E^{(s)} = \sum_n e^{-i\omega_n t} [E_{0n}^{(s)} + i \frac{e^{(s)}}{\varepsilon^{(s)}} q_n^{(s)} \left(u_n^{+(s)} e^{iq_n^{(s)} x} - u_n^{-(s)} e^{-iq_n^{(s)} x} \right)] + c.c \quad (6)$$

Here $E_{0n}^{(s)}$ is a space independent amplitude, $q_n^{(s)} = |\omega_n| / v^{(s)}$ and $v^{(s)} = \sqrt{c^{(s)} / \rho^{(s)}}$ are wave vector and wave velocity respectively.

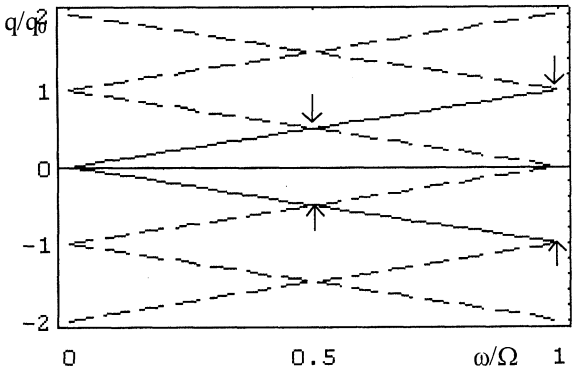


Fig.2. The normalized dispersion laws

In Fig.2 the frequency dependences of $q/q_0 = \pm q_n / q_0$, where $q_0 = \Omega / v^{(s)}$ are represented. Arrows show the points, where intersections of the acoustic mode (solid line) with one of the induced electroacoustic excitations (dashed line) occur. The connection between modes results from boundary conditions due to space independent electric fields $E_{0n}^{(s)}$. These amplitudes for different n satisfy an infinite system of coupled Eqs.4 resulting from Eq.5. Strictly speaking we should take into consideration all harmonics. The effective mutual harmonics influence rapidly vanishes with the increase of harmonic number n, so we shall confine ourselves to the case of n=0, ±1. Then the uniform fields corresponding to these modes in parametric layer may be represented as follows

$$\begin{pmatrix} E_{00}^{(2)} \\ E_{01}^{(2)} \end{pmatrix} = \frac{e}{\varepsilon l} \frac{C_0}{C_2 + C_0} \|A_0\| \cdot \begin{pmatrix} u_0^{+(2)} \\ u_0^{-(2)} \end{pmatrix} + \frac{e}{2\varepsilon l} \frac{C_0^2 m}{(C_2 + C_0)} \|A_1\| \cdot \begin{pmatrix} u_1^{+(2)} \\ u_1^{-(2)} \end{pmatrix}, \quad (7)$$

where

$$\|A_i\| = \begin{pmatrix} e^{iq_i^{(2)}l} - 1 & e^{-iq_i^{(2)}l} - 1 \\ e^{iq_i^{(2)}l} - 1 & e^{-iq_i^{(2)}l} - 1 \end{pmatrix},$$

$C_2 = \dfrac{S}{\varepsilon^{(2)} l}$ - capacitance of layer 2.

For transducer layer one can obtain

$$E_{00}^1 = \frac{iI_0^1}{\omega C_1 d}, \quad E_{01}^1 = 0, \quad C_1 = \frac{S}{\varepsilon^{(1)} d}. \qquad (8)$$

After substitution Eqs.6-8 into mechanical boundary conditions for each harmonic one can express amplitudes $u_{0,\pm1}^{\pm(s)}$ in terms of current

$$I^{(1)} = I_0 e^{-i\omega t} = -i\omega C_1 V_{in}(\omega). \qquad (9)$$

The end result is a connection between the voltage $V^{(1)}(\omega) = V_0(\omega) + V_1(\Omega - \omega)$ and the current (Eq..9):

$$V^{(1)}(\omega) = Z_0(\omega) I_0 e^{-i\omega t} + Z_1(\omega) I_0 e^{-i(\Omega-\omega)t} \qquad (10)$$

The impedances in Eq.10 take a form

$$Z_0(\omega) = \frac{1}{i\omega C_1} + \frac{1}{i\omega C_1} \cdot \frac{(k_t^{(1)})^2}{q_0^{(1)} d} \cdot \frac{iZ \sin q_0^{(1)} d - 2Z'(1 - \cos q_0^{(1)} d)}{\sin q_0^{(1)} d - iZ \cos q_0^{(1)} d}$$

where

$$Z' = \frac{1 - 2\Gamma(\omega) \operatorname{tg} q_0^{(s)} l \dfrac{(1 + \cos q_0^{(1)} d)^2}{2 \cos q_0^2 d}}{1 - \Gamma(\omega) \operatorname{tg} q_0^{(s)} l \cos q_0^{(1)} d},$$

$$Z = i \frac{Z^{(2)}}{Z^{(1)}} \operatorname{tg} q_0^{(2)} l \frac{1 - 2\Gamma(\omega) \operatorname{tg} \dfrac{q_0^{(2)} l}{2} \cos q_0^{(1)} d}{1 - \Gamma(\omega) \operatorname{tg} q_0^{(2)} l \cos q_0^{(1)} d}$$

and

$$Z_1(\omega) = -\frac{(k_t^{(1)})^2}{\omega C_1} \left(1 - \frac{1}{\cos q_0^{(2)} l}\right)(1 - \cos q_0^{(2)} l) \cdot \left(1 - \frac{1}{\cos q_0^{(1)} d}\right)^2 \frac{\cos q_0^{(1)} d}{\sin q_0^{(1)} d - iZ \cos q_0^{(1)} d} \cdot \frac{Z^{(2)} g_1 F(\omega)}{Z^{(1)} q_0^{(1)} d},$$

where

$g_1 = \dfrac{(k^{(2)})^2}{q_0^{(2)} l} \dfrac{C_0^2 m}{C_2 + C_0}$ is a dimensionless connection parameter. The expressions for $\Gamma(\omega, \Omega, g_1^2)$ and $F(\omega, \Omega, g_1^2)$ are too complicated and not presented here. Out of resonance point $\Omega - \omega \neq \omega$ one can define input electric impedance Z_e as

$$Z_e = \frac{V_0(\omega)}{I(\omega)} = Z_0(\omega)$$

At the resonance point $\Omega - \omega = \omega$ one can get

$$Z_e = \frac{V_0(\omega) + V_1(\omega)}{I(\omega)} = Z_0(\omega) + Z_1(\omega). \qquad (11)$$

2.1. Numerical simulations

The dependences of input electric impedance of composite resonator structure consisting of input (ZnO) and parametric ($LiNbO_3$) piezoelectric layers as a function of frequency at different modulation amplitudes were calculated. Dissipation losses are taken into account by replacing $c_0^{(s)} \to c_0^{(s)} + i\eta^{(s)}\omega_n$.

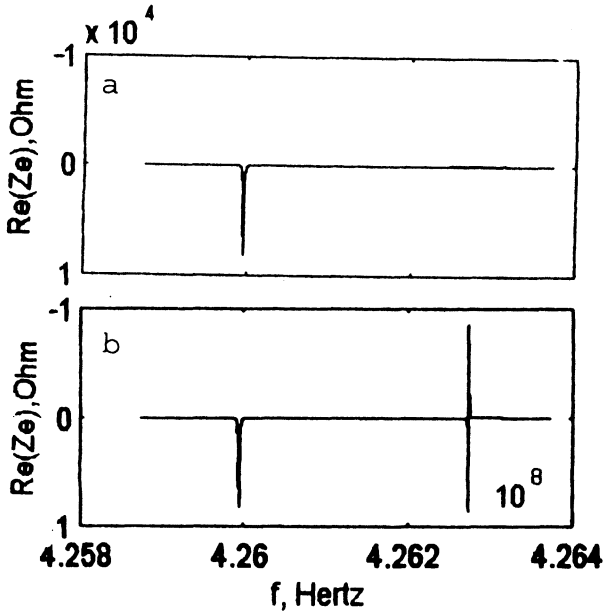


Fig.3. The reactance frequency dependences at different modulation amplitude: *a* - *m* = 0; *b* - *m* = 0.8

In Fig.3a the position of the resonant peak at the absence of capacitance modulation is shown. In Fig.3b the appearance of additional peak in the presence of modulation is demonstrated. It is seen that real part of impedance is becoming negative in some frequency range. The change of the sign means, that there is a signal amplification instead of damping.

In Fig.4 the frequency dependencies of $\mathrm{Re}Z_e$ near the resonant peak at different amplitude modulations are illustrated. As it is seen, with the increase of *m* the amplitude of $\mathrm{Re}Z_e$ also increases.

The threshold value of *m* for sign change onset is very small (*m* = 2.8 10^{-6}) and the frequency interval where $\mathrm{Re}Z_e$ becomes positive increases with *m*, as it is shown in Fig.5.

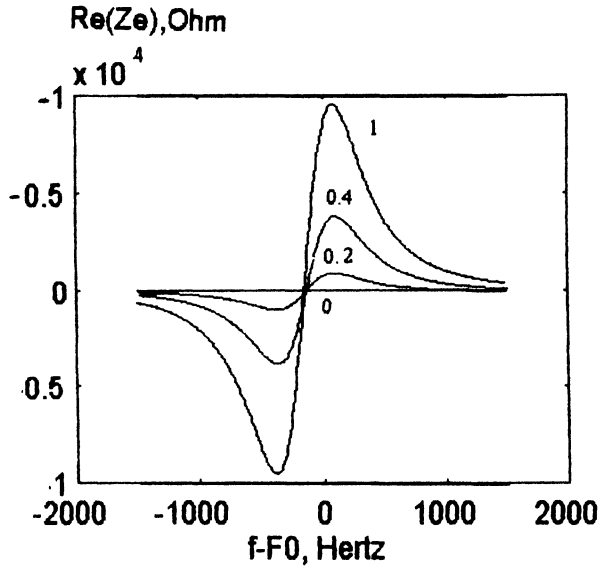


Fig.4. . Frequency dependences of resonant peak at different modulation amplitudes m =1; 0.4, 0.2; 0. F_0 = 426.27 MHz

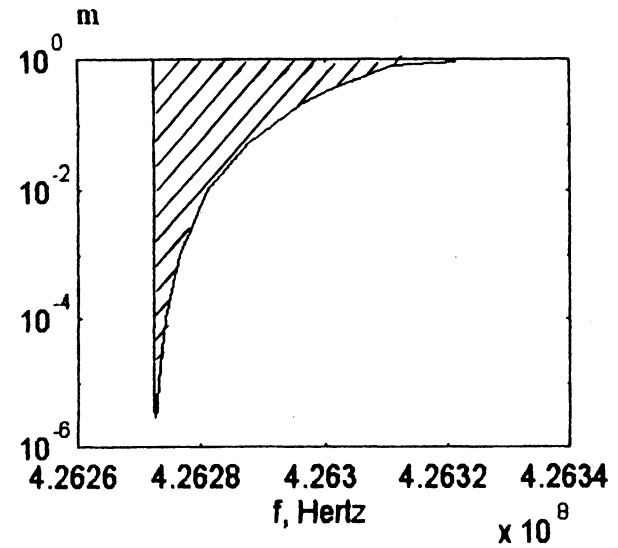


Fig.5.The region of negative reactance at (m, f) plane

In numerical simulations we used conventional parameters of materials for acoustic waves propagation in z-axis of the crystals involved. The diameter of the electrodes was 1mm, thickness of ZnO layer was 4μm and for $LiNbO_3$ layer - 400μm.

3. PARAMETRIC EFFECTS DUE TO NONLINEAR PIEZOEFFECT

Now we consider the situation when AC electric field is applied directly to the piezoelectric media. Nonlinear interaction between the field and acoustic waves propagating in the crystal is provided by nonlinear piezoeffect. The voltage applied to parametric layer and the electric field in it are $V(\Omega)=V_2\cos\Omega t$, $E(\Omega)=E_2(e^{i\Omega t}+e^{-i\Omega t})$, $E_2=V_2/2l$.

We will neglect parametric interaction in the transducer layer and suppose that $e^{(1)}_{ijklm}=0$. Let's assume that in the parametric layer $e^{(2)}_{ijklm}=\bar{e}$. We will take into account in the stress tensor only the nonlinear terms, containing electric field $E(\Omega)$. Then we get

$$T^{(2)}=C^{(2)}u^{(2)}-l^{(2)}E^{(2)}-\bar{e}E(\Omega)\frac{\partial u^{(2)}}{\partial x}, \qquad (12)$$

$$\rho^{(2)}\frac{\partial^2 u^{(2)}}{\partial t^2}-[C^{(2)}-\bar{e}E(\Omega)]\frac{\partial^2 u^{(2)}}{\partial x^2}=0. \qquad (13)$$

Eq.13 is a Mathieu equation with variable coefficients and its solution may be found as

$$u^{(2)}=e^{-i(\omega t-q^{(2)}x)}\sum_n u^{(2)}_n e^{in\Omega t} \qquad (14)$$

Substitute Eq.14 into Eq.13 one can get infinite chain of coupled equations for amplitudes $u^{(s)}_n$

$$[v^{(2)}(q^{(2)})^2-(\omega-n\Omega)^2]u^{(2)}_n=$$

$$=\frac{\bar{e}E(q^2)^2}{\rho^{(2)}}(u^{(2)}_{n+1}+u^{(2)}_{n-1}). \qquad (15)$$

In the absence of nonlinear piezoeffect or event $E=0$ Eq.15 resolves itself into case of noninteracting harmonics $u^{(s)}_n$ considered above. The appropriate dispersion curves are also shown in Fig.2.

Confine ourselves as before to the first resonance only we can cut the chain Eq.15. Then we get characteristic equation for coupled wave u_0 and u_1

$$[(\omega^2-v^{(2)}(q^{(2)})^2][(\omega-n\Omega)^2-v^{(2)}(q^{(2)})^2]=$$

$$=\left(\frac{\bar{e}E^{(2)}(q^{(2)})^2}{\rho^{(2)}}\right)^2 \qquad (16)$$

The solution of Eq.16 is

$$q^{(2)}_{1,2}=\sqrt{\frac{(\omega-\Omega)^2+\omega^2\pm\sqrt{\left((\omega-\Omega)^2+\omega^2\right)^2+4\omega^2(\omega-\Omega)^2\gamma^2}}{2(v^{(2)})^2(1-\gamma^2)}}$$

where

$\gamma=-\dfrac{\bar{e}E_2}{z^{(2)}v^{(2)}}$ is a dimensionless couple parameter.

Two other roots are $q^{(2)}_{3,4}=-q^{(2)}_{1,2}$.

As a result the solution for amplitude may be written as

$$u^{(s)}=\sum_{\substack{j=1,2\\ n=0,1}}\left(u^{+}_{0j}e^{iq^{(2)}_j x}+u^{-}_{0j}e^{-iq^{(2)}_j x}\right)K_{jn}e^{-i\omega_n t},$$

where

$$K_{j0}=1,\ K_{j1}=\frac{q_j^2 v^{(2)}\gamma}{(\omega-\Omega)^2-q_j^2 v^{(2)}}.$$

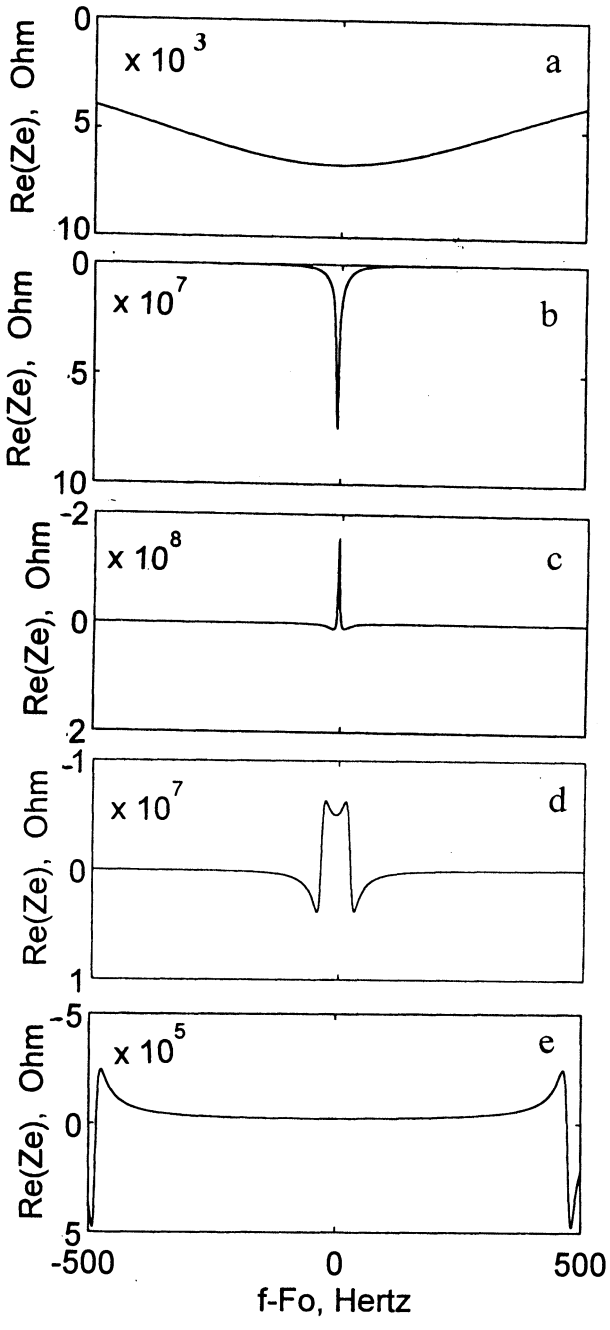


Fig.6. Frequency dependences of the resonant peak near F_0 = 444.135 MHz at different electric fields: *a* - 0; *b* - 77.54; c - 77. 55; *d* - 77.65; *e* - 100 *kV/m*

For impedance Z_e defined in Eqs.10, 11 we obtain $Z_e = Z_0(\omega) + Z_1(\omega)$, where

$$Z_0(\omega) = \frac{1}{i\omega C_1}\left(1+\frac{(k_t^{(1)})^2}{q_0^{(1)}d}\frac{iZ_N\sin q_0^{(1)}d-2(1-\cos q_0^{(1)}d)}{\sin q_0^{(1)}d-iZ_N\cos q_0^{(1)}d}\right),$$

$$Z_N(\omega)=i\frac{Z^{(2)}v^{(2)}}{Z^{(1)}\omega}\cdot \frac{q_1^{(2)}(1-\gamma K_{11})\sin q_1^{(2)}l-q_2^{(2)}N(\omega)(1-\gamma K_{21})\sin q_2^{(2)}l}{\cos q_1^{(1)}l-N(\omega)\cos q_2^{(1)}l}$$

and

$$Z_1(\omega) = -\gamma\cdot\frac{(k_t^{(1)})^2 Z^{(2)}}{\omega C_1 Z^{(1)} q_0^{(1)}d}\cdot \frac{(1-\cos q_0^{(1)}d)(1-\cos q_1^{(1)}d)\Phi(\omega)}{(\sin q_0^{(1)}d - iZ_N\cos q_0^{(1)}d)\sin q_1^{(1)}d}$$

The complicated expressions for $N(\omega,\Omega,g_1^2)$ and $\Phi(\omega,\Omega,g_1^2)$ are and not presented here.

3.1. Numerical simulations

The dependence of input electric impedance of composite resonator structure consisting of input (ZnO) and parametric ($LiNbO_3$) piezoelectric layers as a function of frequency at different electric fields with $\Omega = 2\omega$ applied to the parametric layer is calculated. The results are shown in Fig.6. With the increase of the field up to threshold value 77.55 *kV/m* the resonant peak at first is sharpened (Fig.6a,b) and then change its sign (Fig.6c). It means that parametric effects at initial stage decrease total acoustic attenuation and increase the effective quality factor of the structure. At some electric field all losses are compensated by parametric gain – this field corresponds to the threshold for pararmetric amplification. Subsequent increase of the field leads to the frequency interval broadening where parametric amplification occurs (Fig.6d,e).

The work was partly supported by ISTC Grant №1030.

REFERENCES

[1] J.F.Nye, Physical properties of crystals: their representation by tensors and matrices, Oxford: Claredon Press, 1964.

[2] H.F.Tiersten "Nonlinear electroelastic equations cubic in small field variables", J.Acoust.Soc.Amer. vol.57, p.660, 1975.

Interpreting AM and PM Noise Measurements in Two-channel Interferometric Measurement Systems

E. N. Ivanov* and F. L. Walls**

*Frequency Standards and Metrology Group, Physics Department, The University of Western Australia, Nedlands, 6907, WA

**Time and Frequency Division, National Institute of Standards and Technology, 325 Broadway, Boulder, CO 80303

ABSTRACT

In this work we show experimentally and theoretically that cross-correlation analysis of the output of a two-channel noise measurement system (NMS) yields the differential PM and AM noise of one input relative to the reference port. We also show that the statistical uncertainty, which sets the ultimate spectral resolution in the thermal-noise-limited regime, is approximately the same for both single- and two-channel NMS.

1. INTRODUCTION

In this paper we examine two-channel interferometer noise measurement systems (NMS) that use cross-correlation to measure PM or AM noise in a device under test (DUT). We use a simplification of the entire measurement system to facilitate the measurements at X-band. By changing the noise temperature of the two inputs we show that

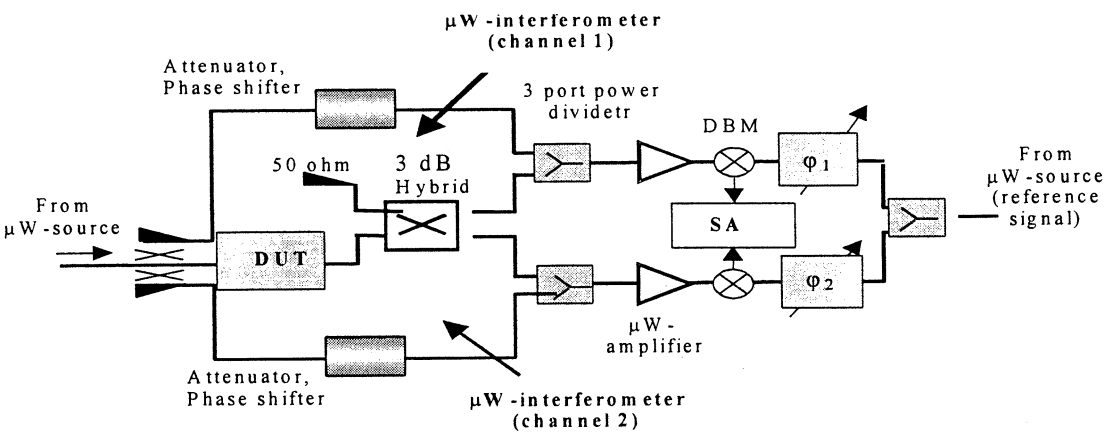


Figure 1a. Two-channel interferometric noise measurement system (NMS). DUT is the device under test, SA is a spectrum analyzer, DBM is a double-balanced mixer, and $\varphi_{1,2}$ are phase shifters.

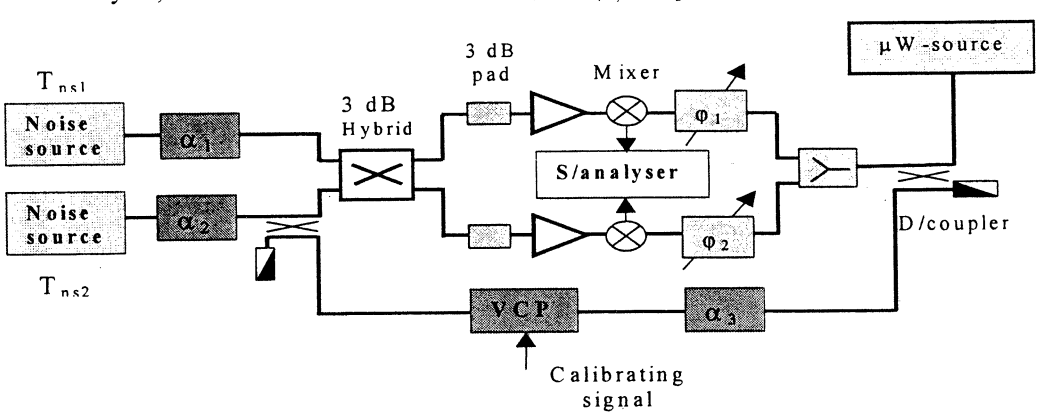


Figure 1b. Experimental setup for studying the basic properties of cross-correlation noise measurement systems. VCP is a voltage controlled phase shifter.

this approach measures only the difference in input noise between the two inputs. We show that the statistical uncertainty is approximately the same for single- and two-channel systems. Practical considerations, however, indicate that the two-channel systems will provide better confidence for the measurement when the noise added by the DUT is near or below the ambient thermal noise in the reference arm of the interferometer.

2. EXPERIMENTAL STUDY OF THERMAL NOISE CANCELLATION EFFECTS IN TWO-CHANNEL NOISE MEASUREMENT SYSTEMS

Figure 1a shows a two-channel NMS. Figure 1b shows the readout portion (low-noise homodyne down-converter) of this system that was used to study the noise processes in a two-channel NMS [1-5]. Two 3 dB pads simulate the loss in the power combiners used in the double interferometer for carrier suppression. The experimental setup also contains two microwave amplifiers and two double balanced mixers. Phase shifters φ_1 and φ_2 are used for optimizing the sensitivities of both channels with respect to either phase or amplitude variations of the incoming signal. The latter is derived from the same source, which drives the mixer's local oscillator (LO) ports and enables the calibration of the readout system. Two noise sources with effective temperatures T_{ns1} and T_{ns2} are attached to the inputs of the measurement system (ports 1 and 2 of 3 dB hybrid, respectively). Variable attenuators α_1 and α_2 allow the effective temperatures of the input noise, T_{inp1} and T_{inp2}, to be varied from the highest values of T_{ns1} and T_{ns2} to the level of ambient temperature T_o. For example,

$$T_{inp1} = T_{ns1}\,\alpha_1 + T_o(1-\alpha_1). \quad (1)$$

Variable attenuator α_3 in the path of the calibration signal was used to set the level of the carrier at the inputs of the microwave amplifiers to approximately -47 dB below 1 mW (-47 dBm), which was a typical operating condition. To simplify the data analysis, the sensitivities of both channels of the NMS were made approximately equal. This was achieved by adjusting the gain of low noise amplifier G_{LNA}, in one of the channels.

The spectral densities of various voltage fluctuations measured with the above setup as a function of the intensity of the input noise are shown in Fig. 2. These data were collected at a Fourier frequency of 35 kHz, which was chosen to minimize the effect of microwave amplifier flicker noise on the voltage noise floor of the NMS (see noise spectra in Fig. 3). The error bars in Fig. 2 show the statistical uncertainty in the three measurements, (see discussion below).

The left-hand vertical scale shows the output in dB below 1 V in a 1 Hz bandwidth (dBV/Hz). The right-hand vertical scale shows the equivalent temperature.

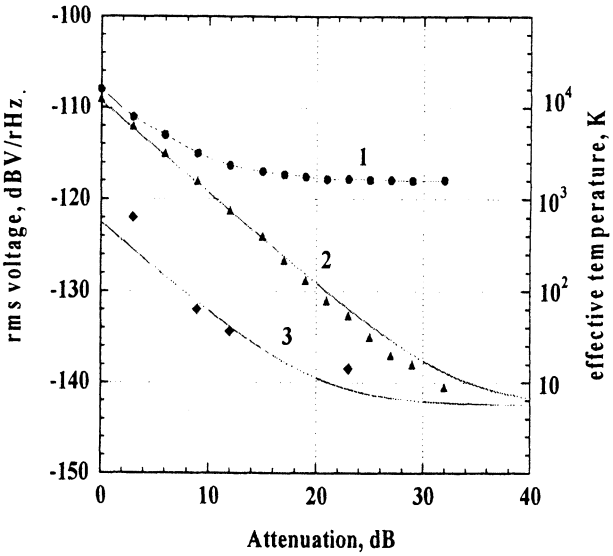


Figure 2. Voltage fluctuations (or equivalent temperature) at the output of a PM or AM NMS as a function of intensity of the input noise T_{inp1} at a carrier frequency of 9 GHz and a Fourier frequency of 35 kHz. Curve 1 is the single-channel voltage noise spectral density for T_{inp2} = T_o. Curve 2 is the cross-channel voltage noise spectral density for T_{inp2} = T_o. Curve 3 is the cross-channel voltage noise spectral density when T_{inp2} is *increased* from T_o to T_{inp1}.

Curve 1 in Fig. 2 shows the dependence of the rms voltage noise at the output of a single-channel NMS on attenuation α_1 with the temperature of port 2 $T_{inp2} = T_o$ (attenuation α_2 is set to a maximum). From these data we see that the noise floor of the single-channel NMS expressed in temperature units was approximately 1300 K.

Curve 2 shows the cross-channel rms voltage as a function of α_1. It was acquired when both channels were identically tuned (phase sensitivities were maximized) and $T_{inp2} = T_o$. These data indicate that, to within the experimental error, the cross-spectral density of output voltage

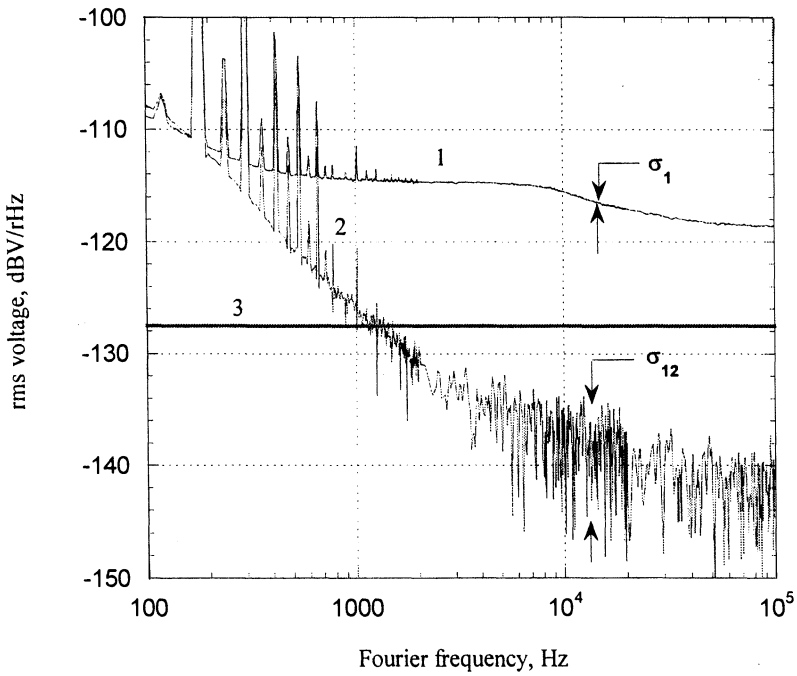


Figure 3. Curve 1 is the measured rms voltage noise floor of a single-channel NMS, curve 2 is the measured rms voltage noise floor of a two-channel NMS, and horizontal line 3 is the calculated rms voltage noise of a single-channel measurement system due to ambient temperature fluctuations.

fluctuations $S_{u1,2}$ is proportional to the difference of effective noise temperatures:

$$S_{u1,2} \propto T_{inp1} - T_{imp2}. \qquad (2)$$

Curve 3 was acquired under the same conditions as above, except that the noise temperature at port 2 was *increased* to match that of the port 1. These data are also consistent with Eq. (2).

The data of Fig.2 provide conclusive experimental evidence that using cross-correlation analysis of the two identically tuned (either phase or amplitude sensitive) outputs of a two-channel NMS yield the difference of the spectral density of the noise between the two inputs.

Fig. 3 shows the dependence of the data from curve 1 and 2 from Fig.2 on Fourier frequency offset f. The heavy horizontal line corresponds to the thermal noise floor of a single-channel NMS at T = 300K. For f above a few kHz the mean value of cross-spectral voltage noise is 12 to 15 dB below the thermal noise floor of a single-channel system.

The fractional statistical uncertainty of the measurement in curve 1 of Fig. 3 is $\pm 1/\sqrt{N_{avg}}$, where N_{avg} is the number of averages [6]. When the DUT noise is much less than the single-channel noise, the average value of the cross-correlated output also falls as $1/\sqrt{N_{avg}}$ [6,7]. The result is that the fractional statistical uncertainty of the cross-correlated two-channel measurement in curve 2 of Fig. 3 is approximately 1 when the effective temperature of the input noise T_{inp} is close to the ambient temperature. The large uncertainty in voltage noise cross-spectral density at Fourier frequencies above 3 kHz indicates that the noise originates from the single-channel noise reduced by the cross-correlation signal processing. This residual noise sets an upper limit to the smallest variations in the measurement system noise floor (or DUT) that can be resolved. In Fig. 3 we compare the spectral resolutions of a single- and a two-channel NMS by considering their voltage noise floors in curves 1 and 2 respectively. Each noise floor looks like a 'fuzzy' trace due to the scatter of experimental data. By measuring the width of such a trace in the vicinity of a given Fourier frequency, one can empirically estimate the spectral resolution of the measurement system. Intervals σ_1 and σ_{12} approximately characterize the spectral resolutions of the single- and two-channel measurement systems, respectively. At Fourier frequencies above 35 kHz: $\sigma_1 \approx 7x10^{-8} V/\sqrt{Hz}$ and $\sigma_{12} \approx 1.2x10^{-7} V/\sqrt{Hz}$ for $N_{avg} = 10^5$. These results are consistent with the fluctuations in both noise measurements

decreasing as $1/\sqrt{N_{avg}}$. See Section III for further discussion.

Similar effects of single-channel noise suppression were observed in the early experiments with conventional two-channel cross-correlation NMS that used double balanced mixers as phase detectors [8]. However, the differential noise floors did not drop below k_BT_o/P_{imp} until much higher Fourier frequencies due to the high level of flicker noise exhibited by these mixers.

3. THEORETICAL ANALYSIS OF CROSS-SPECTRAL DENSITY OF VOLTAGE FLUCTUATIONS IN THE THERMAL-NOISE-SUPPRESSION REGIME

The complex amplitudes of the output signals of an ideal 3 dB hybrid coupler are given by

$$U_3 = \left(U_1 e^{j\varphi_H} + U_2\right)/\sqrt{2}$$
$$U_4 = \left(U_1 + U_2 e^{j\varphi_H}\right)/\sqrt{2}, \qquad (3)$$

where φ_H is a hybrid differential phase shift and U_1 and U_2 are the complex amplitudes of the input signals (see Fig. 1). The differential phase shift in (3) is $\pi/2$ to satisfy the energy conservation conditions.

If there is a loss in the hybrid, equations (3) are no longer valid and a complete set of S-parameters are required to describe the relationship between the complex amplitudes of input and output signals [9]. Apart from that, the differential phase shift in a real hybrid coupler is frequency dependent. For example, for commonly used coplanar stripline couplers, φ_H varies by $\pm 7^o$ with respect to $\pi/2$ in the operating frequency range. In the following analysis we assume the above description of an ideal 3 dB coupler, remembering that $\varphi_H = \pi/2$.

Considering a noise source with a white power spectrum at the input of the measurement system (port 1 in Fig. 1), we can write analytical expressions for the cross-spectral densities of the output voltage fluctuations that are a result of both PM and AM components.

The effect of PM components is given by

$$S_{u1,2}^{(1)PM} = \chi^2 \frac{k_B T_{inp1}}{8}\left\{\cos(\varphi_H - \Delta_{21}) - \cos(\varphi_H - \Sigma_{21})\right\} \qquad (4)$$

where T_{inp1} is the effective temperature of the noise entering port 1, χ^2 is the calibration factor, and Δ_{21} and Σ_{21} are phase angles calculated as $\Delta_{21} = \varphi_2 - \varphi_1$ and $\Sigma_{21} = \varphi_2 + \varphi_1$.[10]

The voltage noise cross-spectral density due to AM components of input noise is

$$S_{u1,2}^{(1)AM} = \chi^2 \frac{k_B T_{inp1}}{8}\left\{\cos(\varphi_H - \Delta_{21}) + \cos(\varphi_H - \Sigma_{21})\right\} \qquad (5)$$

Combining (4) and (5), the cross-spectral density of output voltage fluctuations due to the first noise source is obtained:

$$S_{u1,2}^{(1)} = \chi^2 \frac{k_B T_{inp1}}{4}\cos(\varphi_H - \Delta_{21}). \qquad (6)$$

By analogy, the cross-spectral density of voltage fluctuations caused by the second noise source attached to port 2 in Fig. 1 is

$$S_{u1,2}^{(2)} = \chi^2 \frac{k_B T_{inp2}}{4}\cos(\varphi_H + \Delta_{21}). \qquad (7)$$

Combining (6) and (7) results in the total cross-spectral density of voltage fluctuations at the output of the NMS:

$$S_{u1,2} = \chi^2 \frac{k_B}{4}\left\{\begin{matrix} T_{inp1}\cos(\varphi_H - \Delta_{21}) \\ + T_{inp2}\cos(\varphi_H + \Delta_{21}) \end{matrix}\right\}, \qquad (8)$$

which at $\varphi_H = \pi/2$ becomes

$$S_{u1,2} = \chi^2 \frac{k_B}{4}\left\{T_{inp1} - T_{inp2}\right\}\sin\Delta_{21}, \qquad (9)$$

where the phase angle Δ_{21} is close to $\pi/2$, for identically tuned channels of the NMS.

The above result confirms that the two-channel NMS performs a differential temperature measurement. From previous discussion it also follows that the low voltage noise cross-spectral density must be observed with arbitrary noise sources no matter how high their intensities, provided that (i) noise sources are stationary and (ii) their intensities are matched.

4. DISCUSSION AND CONCLUSION

We have clearly shown that, in contrast to the traditional single-channel NMS, the two-channel NMS with cross-correlation yields the differential PM or AM noise between the two input ports. Since this noise difference is determined on a narrow bandwidth basis, these results apply whether the global character of the noise is white or varies with Fourier frequency.

The resolution of a two-channel NMS with good carrier suppression is typically $2/\sqrt{N_{avg}}$ lower than that of a single-channel system. For example, assuming that both channels are phase sensitive, the cross-correlation analysis yields

$$S_{\varphi}^{1x2}(f) = \left(S_{\varphi}^{dut}(f) - \frac{k_B T_o}{P_{imp} L_{DUT}} \right) + \frac{2k_B T_o}{P_{imp} L_{DUT}} \left(\frac{2}{\sqrt{N_{ave}}} + \delta \right), \quad (10)$$

where S_{φ}^{dut} is the spectral density of PM fluctuations in the DUT, k_B is Boltzmann's constant, T_o is the ambient temperature, P_{inp} and L_{dut} are the power at the input of the DUT and its insertion loss, δ is the parameter ($|\delta| << 1$) characterizing the asymmetry between the two-channels, and N_{avg} is the number of averages. This is to be compared with the results for a single-channel system of

$$S_{\varphi}^{1}(f) = S_{\varphi}^{dut}(f) + \frac{2k_B T_o}{P_{imp} L_{DUT}}. \quad (11)$$

The statistical uncertainty of the measurements is

$$\sigma = \pm \left(S_{\varphi}^{dut}(f) + \frac{2\beta k_B T_o}{P_{imp} L_{DUT}} \right) \frac{1}{\sqrt{N_{ave}}}, \quad (12)$$

where $\beta \approx 1$ for single-channel NMS and $\beta \approx 2$ for two-channel NMS.

These results demonstrate that both single- and two-channel measurement systems are capable of measuring the additive PM (AM) noise in the DUT with an effective temperature smaller than the ambient temperature. Moreover, the two-channel measurement system does not, in principle, offer an advantage over the single-channel one, as far as the spectral resolution of noise measurements is concerned, and it is more complicated. However, in practical terms, the non-stationary nature of the noise, the temporal separation of calibration and measurement, and the difficulty of reproducing the calibrations for two measurements make it extremely difficult to resolve noise which is more than 10 dB below the noise floor in a single-channel NMS.

5. ACKNOWLEDGEMENTS

The authors are grateful to L. Hollberg (NIST), Steve Jefferts (NIST), and A. Luiten (UWA) for useful discussions and valuable advice. This work is supported in part by the Australian Research Council.

6. REFERENCES

1. E.N. Ivanov and F. L. Walls, "Effect of anomalously low voltage noise in two-channel measurement systems", submitted to IEEE UFFC in November 1999.
2. E. N. Ivanov, M. E. Tobar and R. A. Woode, "Applications of interferometric signal processing to phase noise reduction in microwave oscillators', IEEE Transactions on MTT, **46**, 1537-1546, (1998).
3. E. N. Ivanov, M. E. Tobar and R. A. Woode, "Microwave interferometry: application to precision measurements and noise reduction techniques," IEEE Trans. UFFC, **44**, 1526-1537, (1997).
4. E. Rubiola, V. Giordano and J. Groslambert, "Double correlating interferometer scheme for measuring PM and AM noise," Electronic Letters, vol. **34**, 93-94, (1998).
5. E. Rubiola, V. Giordano and J. Groslambert, "Improved interferometric method to measure near-carrier AM and PM noise," IEEE Trans. Instrum. Meas. **48**, 642-646 (1999).
6. F. L. Walls, D. B. Percival and W. R. Irelan, "Biases and variances of several FFT spectral estimators as a function of noise type and number of samples," Proc. of IEEE Freq. Cont. Symp., 1989, 336-341.
7. J. S. Bendat and A. G. Piersol, Measurement and analysis of random data, Wiley & Sons, 1967.
8. F. L. Walls, S. R. Stein, J. E. Gray and D. Glaze, "Design considerations in the state-of-the-art signal processing and phase noise measurement systems," Proc. of the 30th Ann. Symp. on Freq. Cont., 1976, pp. 269-274.
9. P. A. Rizzi, Microwave Engineering Passive Circuits, Prentice Hall, pp. 347-411 (1988).
10. F. L. Walls, A. J. D. Clements, C. M. Felton et al, "Extending the range and accuracy of phase noise measurements", NIST Technical note 1337, pp 129-137(1990)

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

NOISE LIMITS OF CROSS-CORRELATED TWO-CHANNEL NOISE MEASUREMENT SYSTEMS

A. De Marchi, F. Andrisani, G. P. Bava
Dipartimento di Elettronica, Politecnico di Torino
e-mail: demarchi@polito.it
A. De Marchi and G. P. Bava are affiliated with INFM

ABSTRACT

Very interesting sensitivity can be obtained with noise measurement systems built in two channels when their outputs are cross-correlated [1, 2]. This concept was recently applied to microwave interferometers [3], and it was suggested that this approach produces a peculiar noise cancellation phenomenon due to its structure [4].

In this paper, a scattering matrix based analysis of such systems is spelled out, which sheds light on their noise limits and their criticalities. In particular, it is shown that the compensation of contributions from different noise sources to the output cross-correlation is intrinsic of the two-channel structure, subject to certain conditions, and is not peculiar of interferometry.

These conditions are sufficient to obtain the noise compensation, and it's not necessary to double up the interferometer, if that is the technique of choice. A novel scheme of correlated interferometer is hence proposed, in which compensation is obtained without the complications introduced by the double instrument.

1. INTRODUCTION

Two channel noise measurement systems have been widely used for years, in conjunction with cross correlation techniques, in order to discriminate between the noise contributions of the Device Under Test (DUT) and those of the measurement apparatus [1].

In doing so, it was commonly assumed that the noise signal of the DUT, equally divided between the two channels, would show up totally correlated at their output, while the noise generated by the two measurement channels would show up totally uncorrelated and thus be discriminated against by the system. However, no detailed discussion of this assumption is easily found in the literature of papers presenting cross correlation results in two channel noise measurements.

It was recently reported that only the extra DUT noise (the fraction exceeding thermal noise) is detected by cross correlating the two outputs of a double interferometer [4]. Theoretical support of this finding was given, which showed that the contributions of different noise sources to the output cross correlation are indeed expected to compensate, in the particular circuit structure considered, in such a way that the thermal noise is obliterated under thermal equilibrium conditions.

In this paper a more general approach is adopted, in which no predetermined structure is assumed for the system. It is shown that, under certain conditions, no correlation is expected between the noise observed at two ports of any system in thermal equilibrium, irrespective of its architecture. Ideal interferometers happen to meet these conditions, but they are not the only circuits that do.

By the same token, as systems can be closer or further away from these conditions, attention must be paid in assessing the degree to which thermal noise is compensated, in the output cross correlation, when working with any kind of two channel measurement system. A critical review is then proposed here of the above mentioned assumption, which is at the basis of double channel cross correlated measurements.

A simple interferometric system is also proposed, in which the task of compensating for thermal noise is performed outside of the interferometer. In this way the two functions can be optimized separately.

2. THEORY

The discussion of noise cross correlation in double channel measurement systems will be carried out in three successive steps. In this section the first two will be discussed, which assume a perfect system. In the next section the types of relevant non-idealities that exist in a real system will be addressed.

3. Lossless systems in equilibrium

A lossless N-port linear system will be first considered, that is one whose scattering matrix S meets

the unitarity condition $SS^{T*} = 1$, or

$$\sum_k S_{ik} S^*_{jk} = \delta_{ij}. \tag{1}$$

Such a system will not generate noise because it is lossless, and therefore the noise observed at the generic port i is solely a combination of contributions from all N ports. These are here assumed to be closed on matched terminations. By using incoming power waves a_i and outcoming power waves b_i, it is

$$b_i = \sum_k S_{ik} a_k, \tag{2}$$

and this formula can be used to figure out the cross spectrum ${}^b\mathbf{S}_{ij}(f)$ of the noise power waves coming out of ports i and j. The latter, by definition, is the Fourier transform of the cross-correlation ${}^bR_{ij}(\tau)$, which should be calculated as

$${}^bR_{ij}(\tau) = < b_i(t) b^*_j(t+\tau) > . \tag{3}$$

Supposing that all incoming power waves are totally uncorrelated to one another (which means that ${}^a\mathbf{S}_{kr}(f) = {}^a\mathbf{S}_{kk}(f)\delta_{kr}$ identically), it is easily found that

$${}^b\mathbf{S}_{ij}(f) = \sum_k S_{ik} S^*_{jk} {}^a\mathbf{S}_{kk}(f), \tag{4}$$

where ${}^a\mathbf{S}_{kk}(f)$ is the power spectrum of a_k.

An important simplification is obtained in (4) if all the incoming noise power waves are similar, in the sense that they have the same power spectrum (${}^a\mathbf{S}_{kk}(f) = {}^a\mathbf{S}(f)$), because the latter can then be moved as a multiplier outside of the sum, yielding

$${}^b\mathbf{S}_{ij}(f) = (\sum_k S_{ik} S^*_{jk}) {}^a\mathbf{S}(f) = 0. \tag{5}$$

It can be easily seen from the unitarity conditions (1) that b_i and b_j (with $i \neq j$) are then totally uncorrelated. Both their cross correlation and cross spectrum are identically null!

What this means is that, in any non dissipative N-port structure, the cross-correlation between noise power waves outgoing any two ports is always zero, provided all incoming noise power waves are uncorrelated and have the same power spectrum. This condition is filled if only thermal noise is present and the whole system is in thermal equilibrium, although it can be satisfied also in the presence of different noise processes (e.g. flicker or random walk) if the power spectra are the same for all sources.

An easy generalization of this property can be devised in the case where the system is not totally lossless, provided losses are concentrated in such a fashion that they can, in a *gedanken* way, be considered as external to the system and connected to it through hidden matched ports. The conclusions of this paragraph would apply to that case also.

4. Two-channel measurements

While the case dealt with in the previous paragraph is particularly easy to discuss, the model used is not adequate to cover the case of two-channel measurement systems.

One problem with these is that detection front-ends, whether they are amplifiers or mixers or otherwise, usually feed backward (into the system) noise signals with spectral distributions that are different from those of the other noise sources. Then the property demonstrated above would not hold true any longer.

Another problem is that, in a real system, noise voltages or currents are detected, rather than power waves. In order to study a more realistic situation, the quantity $b_i \pm a_i$ should then be considered at each port, and the relevant cross-spectrum between ports i and j should be thought of as obtained from the cross-correlation $R_{ij}(\tau) = < (b_i(t) \pm a_i(t))(b^*_j(t+\tau) \pm a^*_j(t+\tau) >$ rather than from eq.(3). As a result, the observable cross-spectrum $\mathbf{S}_{ij}(f)$ at ports i and j becomes

$$\begin{aligned}\mathbf{S}_{ij}(f) = &\sum_k S_{ik} S^*_{jk} {}^a\mathbf{S}_{kk}(f) + \\ &S_{ij} {}^a\mathbf{S}_{jj}(f) + S^*_{ji} {}^a\mathbf{S}_{ii}(f),\end{aligned} \tag{6}$$

This expression does not vanish, even in the assumption that all incoming noise power waves have the same spectrum, because of the existence of the last two terms.

Moreover, supposing that i and j identify the two output ports of a two-channel noise measurement system, it is to be expected that the noise fed back by the detection front-ends be different from that generated by other sources, so that not even the first part of (6) should be expected to vanish.

However, if all other sources are equal and in equilibrium, (6) can be rewritten as

$$\begin{aligned}\mathbf{S}_{ij}(f) = &(\sum_{k \neq i,j} S_{ik} S^*_{jk}) {}^a\mathbf{S}(f) + \\ &(1 + S^*_{jj}) S_{ij} {}^a\mathbf{S}_{jj}(f) + \\ &(1 + S_{ii}) S^*_{ji} {}^a\mathbf{S}_{ii}(f),\end{aligned} \tag{7}$$

where the contributions of the front-ends have been made explicit. It can then be noticed that all the contributions to the output cross correlation coming from the detection front-ends include the connecting scattering matrix element S_{ij} (or S_{ji}) as a multiplier. As a consequence, such contributions vanish if $S_{ij} = S_{ji} = 0$. In this case, due to (1), the output cross-spectrum would also vanish according to (6) or (7).

The conclusion of this discussion is that, in a lossless system matched at all ports and in equilibrium,

the condition that the output (i and j) ports be pefectly isolated ($S_{ij} = S_{ji} = 0$) is sufficient for the output cross correlation to vanish.

At this point it must be noticed that a two-channel noise measurement system, even if ideal, differs intrinsecally from the depicted model because of the DUT that it hosts within. In fact, the noise generated by the latter, which is exactly the object of the measurement, does not come from the outside of the system. The generalization proposed at the end of section 3 does not help in this case because the DUT usually has two ports, and the noise that it presents at both can hardly be assumed uncorrelated.

However, it was shown in [5] on purely thermodynamical considerations that thermal noise generated within a system matched at all ports does not induce correlation between any two perfectly isolated ports, if the system is in thermal equilibrium. This result can obviously be generalized to cases where all noise sources are similar, even if they are not just thermal. The initial assumption that all noise sources are external to the system can thus be removed.

The conclusion is that only excess DUT noise is theoretically expected to contribute to the output cross-correlation if the two channels of the measurement system are perfectly isolated, irrespective of the system architecture. Excess noise should be understood here as the difference between the DUT's noise power spectrum and the system's equilibrium noise (e.g. the thermal noise spectrum kT). Incidentally, it should be noticed that the sign of this difference gets easily lost in the cross-spectrum measurement, with the effect that it may be hard to distinguish between DUT noise spectra above or below equilibrium noise by the same amount.

This conclusion applies to any two-channel system, and points to possible mistakes in the interpretation of results obtained with cross-correlated two-channel measurement systems. In fact, the already mentioned assumption that is at the basis of their adoption is usually that such a system can detect all the DUT noise, because it is shared by the two channels. It turns out instead that only excess DUT noise is detected in an ideal system, with the additional possible ambiguity on whether the DUT noise is above or below the system's equilibrium noise.

5. IMPERFECTIONS

The effects of the limited compliance of a real system with the conditions assumed for an ideal one in the previous section will be examined in this section. Deviations from the three conditions of perfect equilibrium, perfect output isolation, and perfectly matched loads at all ports will be considered separately, in the assumption that their effects can be superimposed linearly. This assumes that deviations

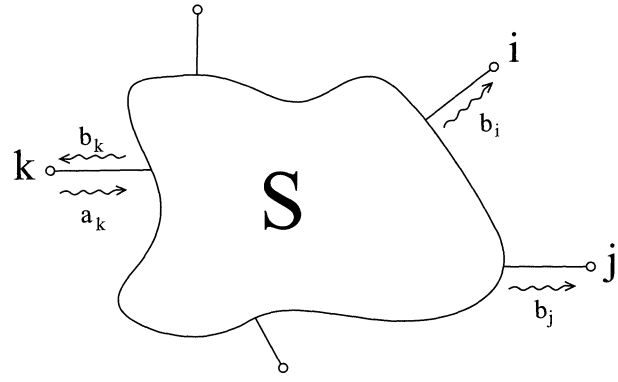


Figure 1: Simple model of the N-port system which is used in the text as a support for discussion.

are small.

Since the focus here is on determining the noise limits for the considered systems, the DUT will be removed in the following.

The discussion will be carried on with reference to Fig. 1, where a simple model of such a system is given.

For this system, assuming that the two outputs are at ports i and j, in perfectly ideal conditions (that is when $S_{ij} = S_{ji} = 0$ and noise spectra available at all other ports are similar and uncorrelated), equation (7) is reduced simply to

$$\mathbf{S}_{ij}(f) = (\sum_{k \neq i,j} S_{ik} S_{jk}^*)^a \mathbf{S}(f) = 0. \qquad (8)$$

While (8) vanishes, modifications have to be introduced in it for the different non-idealities.

6. Imperfect equilibrium

A noise imbalance at port k would prevent the vanishing of the output cross-spectrum. By defining

$$^a\mathbf{S}_{kk}(f) = {}^a\mathbf{S}(f) + {}^a\Delta(f) \qquad (9)$$

the latter can be written as

$$\mathbf{S}_{ij}(f) = S_{ik} S_{jk}^* {}^a\Delta(f). \qquad (10)$$

For example, this may raise from a temperature difference ΔT between the two terminations. In this case it would be ${}^a\Delta(f) = k\Delta T$. If $\Delta T = 0.3$K, the noise limit resulting from this imbalance would be some 30 dB below thermal noise.

7. Imperfect isolation

If the output isolation is not perfect ($S_{ij} \neq 0$), then the detection front-ends contribute to the cross-spectrum, and (8) must be written again in full as in (7).

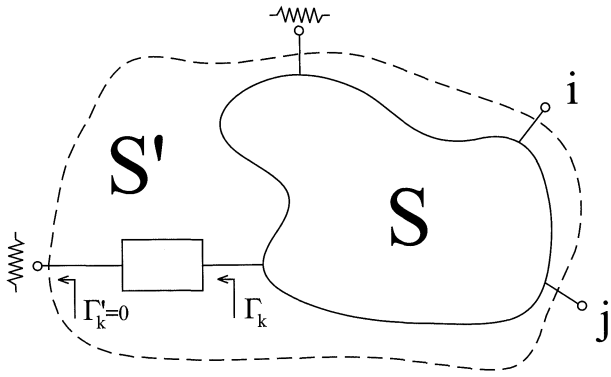


Figure 2: Model of a real system with a mismatched load, showing how the mismatch can be included in the system's modified scattering matrix S'

In the unlikely event that the noise fed back by the detection front-ends be equal to that of the other noise sources (${}^a\mathbf{S}_{ii}(f) = {}^a\mathbf{S}_{jj}(f) = {}^a\mathbf{S}(f)$), equation (7) would reduce to

$$\mathbf{S}_{ij}(f) = (S_{ij} + S^*_{ji}){}^a\mathbf{S}(f). \qquad (11)$$

As a matter of fact, if the noise fed back by the front-ends is not too far from equilibrium, equation (11) may be not such a bad approximation anyway, provided S_{ij} and S_{ji} are not too far from zero, because then it can be assumed that $S_{ii}S^*_{ji}$ and $S_{ij}S^*_{jj}$ be much smaller than $S_{ik}S^*_{jk}$ for any k, with the consequence that (11) is almost the same as (7).

In any case, if for example the output isolation is -30 dB, the resulting noise limit will be found about 30 dB below the noise fed back into the system by the detection front-ends.

8. Imperfectly matched loads

The case of imperfect matching at one or all ports can be rather easily reconducted to the previous case.

In fact, a mismatch at port k, producing a reflection coefficient Γ_k, can be thought of as introduced by a two-port structure inserted between the load and an ideal system's port, as shown in Fig. 2. By characterizing this structure with its two-by-two scattering matrix, an adjusted scattering matrix S' can be calculated for the whole system. What is relevant is the term S'_{ij} obtained in this way, which is given by

$$S'_{ij} = S_{ik}S_{kj}\Gamma_k. \qquad (12)$$

The meaning of (12) is soon recognized by inspection, as it represents the fraction of the noise, fed back into the system by the front-end of channel i, that couples into channel j by reflection on the mismatched load in port k.

If the mismatch is, for example, -30 dB, the resulting noise limit is again of the order of 30 dB below the noise coming from the detection front-ends.

On the bright side, it can be pointed out that this contribution can be manipulated by acting on the load matches, and can even be played against the previous two kinds of contributions, with the aim of minimizing the noise limit of the whole two-channel measurement system. In fact, for example, if S_{ij} is the output coupling of a reciprocal non-ideal system, it appears from (12) that the modified output coupling S'_{ij} can be trimmed down to vanish by inserting a load mismatch at port k, such that

$$\Gamma_k = -\frac{S_{ij}}{S_{ik}S_{kj}}. \qquad (13)$$

This kind of fix can obviously be expected to work only in narrow band operation, but this is exactly the situation in systems designed to measure modulation noise around a given carrier.

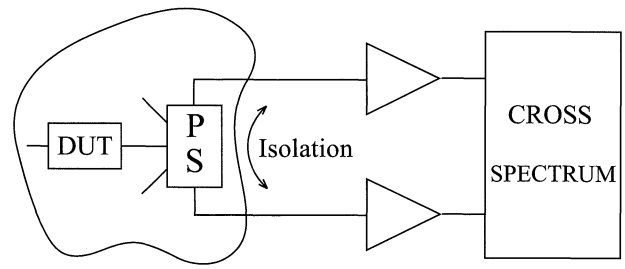


Figure 3: Model of the two-channel noise measurement system, in which it is underlined how the system's performance depends critically on the characteristics of the power splitter

9. POWER SPLITTERS AND NOISE LIMITS

At some point of a two-channel system, downstream from the DUT, some kind of power splitter has to be employed, as illustrated in Fig. 3. If everything that exists after the splitter is considered as part of the detection devices, then the detection front-ends are right next to the splitter's output ports, and the characteristics of the splitter itself, in particular its output isolation, are most important to determine the noise limits of the two-channel system, according to the discussion of section 5. The latter will determine how much the noise fed back by the front-ends contribute to the output cross-correlation.

If the device used is just a common off-the-shelf power splitter, with no particular isolation property, then the analysis offered in section 7 may not even hold, as it was done in the assumption of small devi-

ations from ideality. A realistic analysis would then be much more complicated than the present one.

If instead a coupler is used, it would typically be a 3 dB (e.g. hybrid) coupler, in order to present similar signal levels to the two front-ends. Values of the order of a few percent can then be assumed for the isolation S_{12} if a good quality broadband off-the-shelf device is employed. According to (11), a noise limit of the order of 20 dB below the noise fed back by the detection front-ends would then be experienced (for example 20 dB below thermal noise if such back-fed noise is not too far from equilibrium).

It must be pointed out, however, that even a commercial coupler can be trimmed up externally for optimum isolation, if only narrow band operation is needed, which is usually the case in noise measurement systems. In this case, it might be possible to improve significantly the situation.

Of course, two isolators could alternatively be inserted at the output of the splitter in order to increase the isolation and reduce the front-end contribution. In this case, however, the (excess) noise of the isolators themselves would also give a contribution, albeit small.

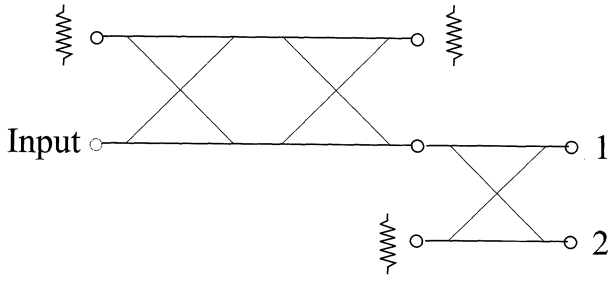


Figure 4: Proposed scheme of a simple interferometer with cross-correlation

10. INTERFEROMETERS

Interferometers have been used in many disciplines for a long time [6,.7], and were recently intensively developed for application in noise measurement systems (for some ot the best results see [8]). Their main advantage in this perspective is the reduction in detection (amplifier) noise that they achieve by suppressing the carrier, and consequently their capability of realizing very low noise limits in real time operation, without the need of resorting to time consuming statistical post-processing, and without the disadvantage of becoming blind to part of the DUT noise. For these properties of theirs, they are suitable for use in noise reduction servo loops.

As already mentioned, a two-channel cross-correlation system based on a double interferometer has been proposed [4] as a means to further decrease their noise limits.

After the discussion given above, however, it appears questionable whether combining the two techniques would bear any advantage, apart from decreasing averaging times for the cross-correlation (cross-spectrum) convergence because the carrier suppression reduces the detector noise, and possibly relaxing the specifications for output isolation, for the same reason.

Because of the complications involved in combining the two techniques, and because the major advantage of real-time operation is lost in the process, the right answer is probably positive only when these two gains are considered very important.

In particular, it should be pointed out that obtaining satisfactory operation from a two-channel interferometric system with cross-correlation requires at least the optimization of two features at the same time: one is the carrier extinction ratio (or fringe contrast), which is the mark of a good interferometer, and the other is the output isolation, which was shown here to be needed for best operation of the two-channel system.

Careful adjustments must therefore be made in the interferometer to fill the first need (commercial 3 dB couplers are typically not good enough to this aim), and also on the power splitting device in order to maximize the isolation of the two outputs.

The latter task is no different than in any two-channel system. However, if the power splitter is also part of the interferometer, it is to be expected that the two optimization tasks might not be easily performed at the same time.

For this reason a simple system, in which the correct operation of the interferometer and the isolation of the two outputs can be optimized separately, is proposed here, in Fig. 4, as possibly the best solution for marrying interferometry with the two-cannel cross-correlation technique.

In this scheme, the interferometer can be optimized in order to minimize the noise coming from the input to its useful output, and the output coupler can be optimized in order to maximize the output isolation.

11. CONCLUSIONS

In this paper it was shown theoretically that no cross-correlation (i.e. a vanishing cross-spectrum) is expected between the two outputs of any "ideal" two-channel noise measurement system. By "ideal" it is intended here that all noise is in perfect equilibrium, the two output channels are perfectly isolated, and all ports are closed on perfectly matched loads. The detailed structure of the system is otherwise irrelevant. Only the excess noise of a DUT can be ob-

served in such systems.

Noise limits appear in the output cross-spectrum if one or more of these ideality conditions are met only to a certain level. These limits were evaluated in the paper. It turns out that they can derive from two possible sources: a noise imbalance from the loads, and noise fed back by the front-ends of the detectors that couple into the other output channel due to lack of isolation. This latter contribution may well be the main one in most cases, particularly if the detector noise is not very low.

Trimming of the system for noise floor optimization can be tried, and is expected to work in a narrow band, although it appears not trivial to obtain levels much lower than say 30 dB below the thermal noise.

Interferometers can obviously be arranged in a two-channel configuration and combined with the cross-correlation technique, although it is argued here that the advantages of this technique appear questionable, while the price of giving up real-time operation and becoming blind to equilibrium noise is certain.

If marrying the two techniques is nevertheless considered desirable for some particular application, the problem arises of optimizing at the same time the output isolation and the carrier suppression of the interferometer. This appears hard if the two tasks are performed by the same device. As a solution for this problem, a very simple scheme of a two-channel system using interferometry was proposed here, in which the two optimization tasks can be performed separately on two different parts of the system.

12. REFERENCES

[1] F. L. Walls, S. R. Stein, J. E. Gray, D. J. Glaze, "Design considerations in state-of-the-art signal processing and phase noise measurement systems", Proc. 30th FCS, pp. 269–274, Atlantic City, 1976.

[2] D. Fest, J. Groslambert, J. J. Gagnepain, "Individual characterization of an oscillator by means of cross-correlation or cross-covariance methods", IEEE Trans. on I&M, vol. 32, n. 3, Sept. 1983.

[3] E. Rubiola, V. Giordano, J. Groslambert, "Double correlating interferometer scheme to measure PM and AM noise", Electronics Letters, vol. 34, n. 1, pp. 93–94, Jan. 8th 1998.

[4] E. Rubiola, V. Giordano, "Correlation-based noise measurements below the thermal noise floor", Proc. 1st EFTF-FCS joint meeting, pp. 1125–1128, Besançon, 1999.

[5] A. G. Orlov, "Calculation of the intrinsic noise of a lossy passive multipole from its scattering matrix", Radio Engineering, Vol. 24, n. 9, 1969.

[6] K. H. Sann, "The measurement of near-carrier noise in microwave amplifiers", IEEE Trans. on MTT, vol. 16, n. 9, pp. 761–766, Sept. 1968.

[7] C. H. Horn, "A carrier suppression technique for measuring S/N and carrier/sideband ratios greater than 120 dB", Proc. 23rd FCS, pp. 6–8, Atlantic City 1969.

[8] 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. on UFFC, vol. 44, n. 1, pp.161–163, 1997.

A NOVEL BI-DIRECTIONAL INTERFEROMETRIC DISCRIMINATOR FOR LOW NOISE MICROWAVE APPLICATIONS

John G. Hartnett, Michael E. Tobar, Eugene N. Ivanov.

Department of Physics, University of Western Australia, Nedlands, WA, Australia.
Tel: +61 8 9380 3443 Fax: +61 8 9380 1014 Email: john@physics.uwa.edu.au

ABSTRACT

We report on a novel frequency discriminator, which employs a bi-directional (BD) configuration (patent pending). Microwaves pass in both directions through each arm of the interferometer. The discriminator sensitivity was found to be 6 dB higher than that for the conventional interferometer. Also, as no circulator is used inside the interferometer its phase noise contribution becomes negligible.

1. INTRODUCTION

A frequency discriminator (FD) comprises a dispersive element such as a resonator or a delay line and a phase sensitive read-out system. The FD is used to detect and cancel oscillator frequency excursions from the frequency of operation. The sensitivity ultimately determines the performance of a frequency-stabilized oscillator.

Galani et al [1, 2], combined a part of the incident signal with the reflected signal from a dielectric resonator. Using a double balanced mixer phase detector, a phase error was fed back to reduce the phase noise in an oscillator. This technique is limited by the noise temperature of the mixer, which is typically of the order 10^4 K. By critically coupling the reflection port of the resonator (suppressing the carrier), the effect of the high mixer noise temperature on the sensitivity can be reduced. This can be further improved by placing an RF amplifier in front of the mixer. This method necessarily relies on near perfect critical coupling to avoid the production of power dependent flicker noise in the amplifier. G.J. Dick et al [3] implemented this method to achieve a phase noise of -125 dBc/Hz, at 1 kHz offset from the carrier, in a 8.6 GHz quartz oscillator by locking it to a 77 K sapphire loaded cavity FD.

Significant improvements in the sensitivity of FDs have recently been achieved with applications of microwave circuit interferometry to the design of noise measurement systems [4-6]. The interferometer is tuned to a "dark" port by balancing phase and attenuation in the arms. This technique overcame the problems associated with the critical coupling technique. Interferometric FDs are capable of operating with almost thermal noise limited sensitivity. Making use of an interferometric FD based on a room temperature sapphire loaded cavity resonator, the phase noise of an X-band oscillator was reduced to –150 dBc/Hz at 1 kHz offset from the carrier[4].

In this paper, we outline an alternative interferometer configuration, where a 6 dB better sensitivity than the conventional interferometer can be achieved. Apart from that, the circulator necessary in [4] may be moved outside the interferometer. Transfer functions of the FD output and voltage-to-frequency conversion ratios are compared with conversion ratios for the conventional interferometric FD. Experimentally measured results are compared with those calculated and found to be in good agreement.

2. EQUIVALENT CIRCUIT MODEL

In order to model the FD circuit presented in this paper, a high-Q resonator was modeled as a parallel LCR circuit (Fig. 1). The probes were modeled as resistive with small series inductance (ρ_1 and ρ_2) due to probe design. Coupling on ports 1 and 2 are defined by $\beta_x = R/R_x$ where x = 1 or 2 and the effective coupling $\beta_1^* = \beta_1/(1+\beta_2)$ and visa versa.

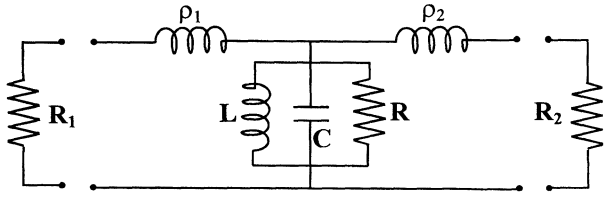


Figure 1: Equivalent parallel LCR circuit model for a high-Q resonator with small series inductance (ρ_1, ρ_2) from the loop probes. R_1 and R_2 represent probe resistance.

3. BI-DIRECTIONAL INTERFEROMETRIC FD

The BD FD is shown schematically in Fig. 2. It comprises an interferometer with phase shifter (φ) and attenuator (α) in one arm to meet the necessary conditions to suppress the carrier at the output of a 3-dB hybrid. A microwave signal is applied at the sum (Σ) port of the hybrid (input in Fig. 2), and travels through both arms of the interferometer. The four amplitude components are combined in the hybrid as a sum (Σ) or difference (Δ) of the returned signals. Generally, when a "dark" port is created at one output port of the hybrid a "bright" port is created at the other. Two cases are

illustrated; where the carrier is suppressed at the Δ-port (output 1 in Fig. 2) and where it is suppressed at the Σ-port (output 2 in Fig. 2).

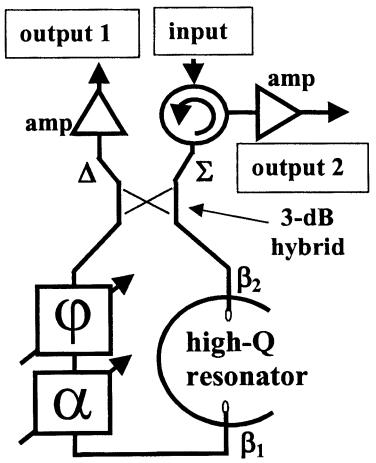


Fig. 2:Schematic of BD interferometric FD.

4. CARRIER SUPPRESSION AT HYBRID Δ-PORT

When carrier suppression is achieved at the hybrid Δ-port, the on-resonance output amplitude transfer function for the "dark" port (T_Δ) is

$$T_\Delta = \frac{e^{-2\alpha-2j\varphi}(\beta_1-\beta_2-1)+(\beta_1-\beta_2+1)}{2(\beta_1+\beta_2+1)} \qquad (1)$$

where a 180° 3-dB hybrid was assumed. By equating T_Δ to zero, phase and amplitude balance conditions of carrier suppression are determined. The required attenuation is $\alpha = \frac{1}{2}\ln\left|\frac{1-\beta_1+\beta_2}{1+\beta_1-\beta_2}\right|$ with the required phase balance conditions; $\varphi = 0$ when $\beta_1^* \le 1$ and $\beta_2^* \le 1$, and $|\varphi_1| = \pi/2$ when either $\beta_1^* \ge 1$ or $\beta_2^* \ge 1$. The parameter α must be positive representing an additional attenuation required for carrier suppression.

5. CARRIER SUPPRESSION AT HYBRID Σ-PORT

In the BD configuration where the carrier is suppressed at the hybrid Σ-port, the on-resonance output amplitude transfer function for the "dark" port is given by

$$T_\Sigma = \frac{e^{-2\alpha-2j\varphi}(\beta_1-\beta_2-1)+(\beta_2-\beta_1-1)}{2(\beta_1+\beta_2+1)} + \frac{2e^{-\alpha-j\varphi}\sqrt{\beta_1\beta_2}}{\beta_1+\beta_2+1} \qquad (2)$$

but different phase and amplitude balance conditions to (1) are required. Analysis of this configuration was quite complex, however when $\beta_2 \approx 0.5$ it is simplified. Under this constraint, the conditions required for carrier suppression are; firstly, when $\beta_1+\beta_2 \le 1$, $\alpha = \frac{1}{2}\ln\left|\frac{1-\beta_1+\beta_2}{1+\beta_1-\beta_2}\right|$ and $\varphi = ArcTan\left[\frac{\sqrt{1-(\beta_1+\beta_2)^2}}{2\sqrt{\beta_1\beta_2}}\right]$. The latter function evaluates as a continuous function varying from $\pi/2$ to 0, not discrete values as in the Δ-port configuration. Secondly, when $\beta_1+\beta_2 \ge 1$ $\alpha = \ln\left|\frac{2\sqrt{\beta_1\beta_2}+\sqrt{(\beta_1+\beta_2)^2-1}}{1-\beta_1+\beta_2}\right|$ and $\varphi = 0$ when under-coupled or $|\varphi| = \pi$ when overcoupled.

6. FD SENSITIVITY

One of the important figures of merit of a FD is its frequency sensitivity or conversion efficiency. This is expressed as

$$S_{FD} = \kappa\sqrt{P_{inc}K_{amp}}\left|\frac{\partial[\mathrm{Im}\,T]}{\partial f}\right| \qquad (3)$$

where $|\partial[\mathrm{Im}T]/\partial f|$ is the absolute value of the frequency derivative of the imaginary part of the interferometer transfer function (either T_Δ or T_Σ). It is maximized when $f_{res} = f_{osc}$. f_{res} is the resonance frequency of the resonator and f_{osc} the operational frequency of the oscillator, κ is the mixer power-to-voltage conversion ratio, P_{inc}, the power incident on the interferometer and K_{amp}, the gain of the low noise microwave amplifier.

For the carrier suppressed Δ-port

$$\left|\frac{\partial[\mathrm{Im}T]}{\partial f}\right|_{f_{osc}=f_{res}} = \frac{|\beta_1-\beta_2|}{|\beta_1-\beta_2|+1}\frac{1}{\Delta f_{0.5}}$$

where $\Delta f_{0.5} = f_{res}/2Q_L$ and Q_L is the loaded Q-factor. $|\partial[\mathrm{Im}T]/\partial f|$ becomes maximized when $\beta_2 = 1$ and $\beta_1 = 0$ or visa versa. It simplifies in this case to $0.5/\Delta f_{0.5}$, which is identical in the conventional interferometer.

For the carrier suppressed Σ-port

$$\left|\frac{\partial[\mathrm{Im}T]}{\partial f}\right|_{f_{osc}=f_{res}} = \frac{(\beta_2-\beta_1)+(\beta_1+\beta_2)^2}{(1-\beta_1+\beta_2)^2}\frac{1}{\Delta f_{0.5}}$$

which becomes maximized when $\beta_1=\beta_2= 0.5$. In this case, it simplifies to $1/\Delta f_{0.5}$.

Normalized S_{FD} is plotted in Fig. 3 for each of these FDs with optimum coupling. S_{FD} is expressed in dB units relative to the maximum value of the conventional interferometric FD (critically coupled to resonator). The maximum value of frequency sensitivity in the BD_Σ configuration is 6 dB better than that for both the conventional interferometer and the BD_Δ configuration. (Subscripts, Δ and Σ, denotes carrier-suppression port).

When either port becomes uncoupled from the resonator (that is, zero coupling) in the BD_Δ configuration, the resulting topology is identical with the conventional reflection interferometric FD using a hybrid and no circulator [7]. This is the condition for the maximum S_{FD} in the BD_Δ FD (shown by curve 1 in Fig. 3 when $\beta_2 = 0$ and 1). When the coupling on the port 1 (curve 1a in Fig. 4) is unity the coupling on port 2 is zero. The maximum sensitivity of the BD_Σ FD is when $\beta_1 = \beta_2 = 0.5$ (see curve 2 in Figs 7 & 8).

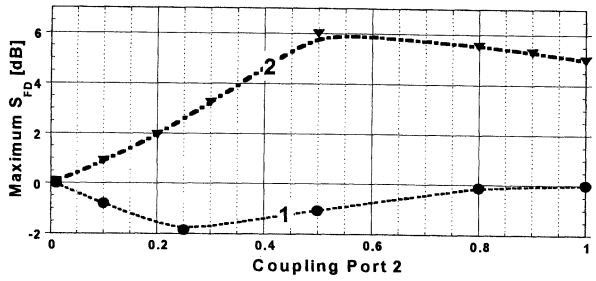


Figure 3: Comparison of normalized S_{FD}, maximized by adjusting β_1 for a range of coupling values β_2. Curve 1 is for the BD_Δ FD and curve 2 for the BD_Σ FD.

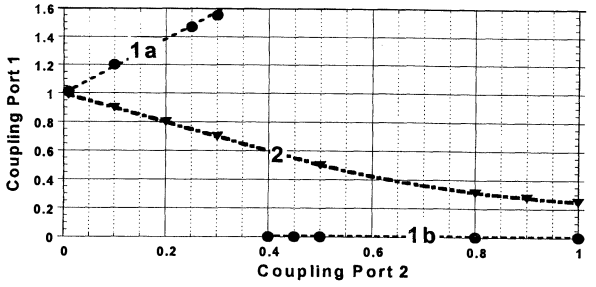


Figure 4: The relationship between the coupling on the resonator ports required to yield the maximum value of S_{FD}. Each curve corresponds to the same numbered curve in Fig. 3, that is, curve 1 for the BD_ΔFD and curve 2 for the BD_Σ FD.

7. EXPERIMENTAL RESULTS & DISCUSSION

Both BD configurations were examined experimentally. A copper cavity housing a single crystal sapphire resonator was coupled to an interferometer via loop probes and coaxial lines. A phase shifter was inserted in one arm and an attenuator in the other (see Fig. 5). The output scattering amplitudes were measured at the sum and difference ports of a 3-dB 180° hybrid by a vector network analyzer. Various coupling conditions were chosen, the carrier suppressed by about 60 dB (Figs 6 & 7). Normalized S_{FD} values were calculated and compared with those modeled. Since the circuit was broken each time a measurement was made some variation from the modeled value is expected (Table 1).

An examination of the output amplitudes of the carrier-suppressed and "bright" ports in both BD configurations revealed that they were shifted in frequency relative to each other. This was seen as the effect of small loop probe inductance. Once this was added to the model the agreement was excellent for the BD_Δ configuration (Fig. 6) but failed to match the "bright" port in the BD_Σ configuration (Fig.7). However the frequency shift was in the predicted direction. The discrepancy in the result may be due to some additional phase shift introduced through the addition of a circulator.

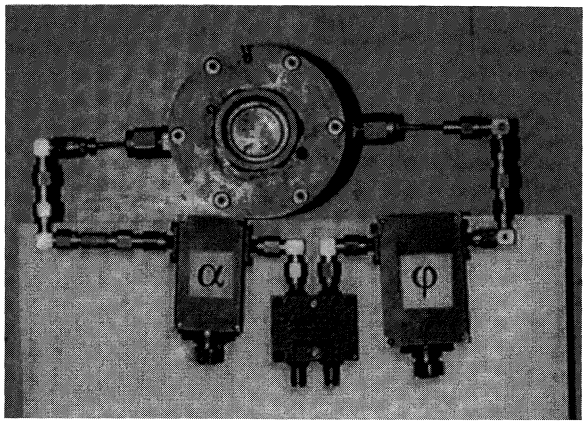


Figure 5: Experimental set up for bi-directional interferometric configuration. This comprised a phase shifter (φ), an attenuator (α), a 180° 3-dB hybrid and a copper cavity housing a single crystal sapphire resonator coupled to the interferometer via loop probes.

Table 1: Modeled and measured values of S_{FD} (normalized to the bold value) for the BD_Δ and BD_Σ configurations. The last column is the percentage difference.

	β1	*β2*	*Model*	*Measured*	*%*
Δ	0.08	0.99	1.000	0.898	11
Δ	0.31	0.88	0.774	0.774	0
Δ	0.15	0.50	0.628	0.533	18
Σ	0.31	0.88	1.752	1.467	19
Σ	0.08	0.64	0.584	0.628	7

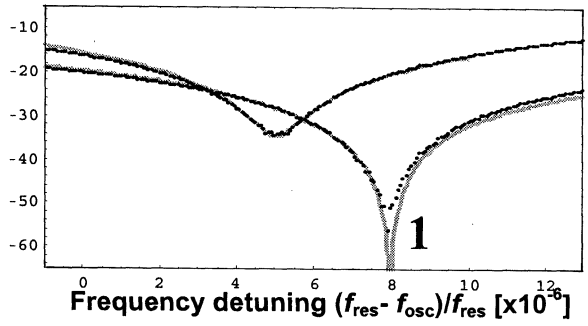


Figure 6: The output power in dBm at the carrier-suppressed port (curve 1) and the "bright" port (curve 2) in the BD_Δ configuration plotted against fractional frequency detuning. Both minima are shifted up in frequency. The modeled curves are the grey solid lines and the black points are the measured data.

8. DISCUSSION

Only the BD_Σ FD, because of its high sensitivity, will be considered in the following analysis. Assuming that a mechanical phase shifter and attenuator is used, zero reference phase shifter error ($\delta\varphi_{ref} = 0$) and the operational frequency of the oscillator is equal to the

resonator resonance frequency, then the phase noise floor of the BD_Σ FD may be written as

$$S_\varphi^{n/f} \approx \left\{ \frac{\kappa \sqrt{k_B (T_0 + T_{amp})}}{S_{FD}} \frac{1}{f} \right\}^2 + S_\varphi^{circ} \left\{ 1 + \frac{1}{CS_P} \left(1 + (\frac{\Delta f_{0.5}}{f})^2 \right) \right\} \quad (4)$$

where $\beta = \beta_1 = \beta_2 \approx 0.5$; CS_P is the level of carrier suppression; k_B is Boltzmann's constant; T_0, the ambient temperature; T_{amp}, the microwave amplifier noise temperature at the carrier-suppressed port; f, the Fourier frequency and S_φ^{circ} is the circulator phase noise. The first term of (4) results from thermal noise generated in lossy components of the FD as well as microwave amplifier intrinsic noise. The second term is the result of noise generation in the circulator. The mixer noise is not taken into account in the first term of (4) because high amplifier gain makes its contribution insignificant. Comparing this with the conventional interferometric FD (employing a 3-dB hybrid at both the input and output of the interferometer), it is seen to be 6 dB less at all Fourier frequencies.

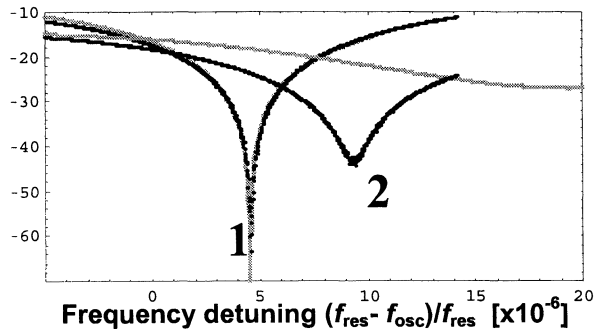


Figure 7: The output power in dBm at the carrier-suppressed port (curve 1) and the "bright" port (curve 2) in the BD_Σ configuration plotted against fractional frequency. Both minima are shifted up in frequency. The modeled curves are the grey solid lines and the black points are the measured data.

9. CONCLUSION

Fig. 8 illustrates how it may be used to reduce the phase noise in an oscillator. A signal is taken from the oscillator and injected into the FD via the input port of the 3-dB hybrid and the carrier is suppressed at the output of the circulator at this port. In practice, either port could be used to inject the signal. In the ideal situation (no loss in the arms or hybrid and $\beta_1 = \beta_2 = 0.5$), when the carrier is perfectly suppressed all power is dissipated in the resonator and no power is seen at the other hybrid port, on or off resonance. Under these conditions zero added phase or attenuation is required. In a real system, however, small attenuation and phase shift is needed to balance transmission line and loop probe losses and phase added by line discontinuities.

As a result, the BD FD's phase noise floor is 6 dB lower than the conventional interferometric FD. The bi-directional FD, where a circulator is placed at the input port and the carrier is suppressed at the output of that circulator, is superior in performance and not limited by the circulator phase noise contribution. In fact, the carrier power, driving the noise generation in the circulator, is suppressed by the compensating arm of the interferometer, without the need of optimum coupling.

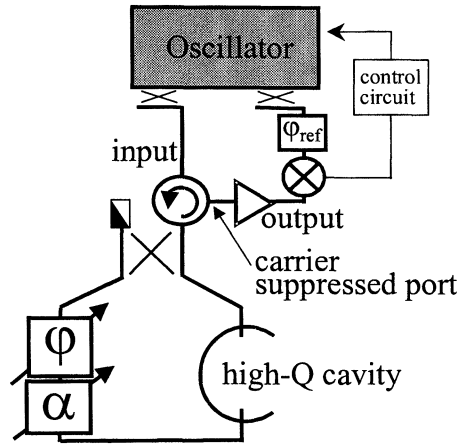


Figure 8: Schematic of BD frequency discriminator with carrier suppression at the injected signal port.

References

[1] Z. Galani, M. J. Bianchini, J. Raymond C. Waterman, R. Dibiase, R. W. Laton, and J. B. Cole, "Analysis and design of a single-resonator GaAs FET oscillator with noise degredation," *IEEE Transactions on Microwave Theory and Techniques*, vol. 32, pp. 1556-1565, 1984.

[2] Z. Galani, M. J. Bianchini, J. Raymond C. Waterman, R. Dibiase, R. W. Laton, and J. B. Cole, "Microwave Oscillator," . US: Raytheon, 1985.

[3] G. J. Dick and D. G. Santiago, "Microwave frequency discriminator with a cryogenic sapphire resonator for ultra-low phase noise," *Proceedings IEEE Freq Control Symp.*, pp. 176-182, 1992.

[4] E. N. Ivanov, M. E. Tobar, and R. A. Woode, "Advanced phase noise suppression technique for the next generation of ultra-low noise microwave oscillators," *Proceedings IEEE Freq. Contr. Symp.*, pp. 314 - 320, 1995.

[5] 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.

[6] 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 Tech.*, 1998.

[7] 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.

Improved Correlation Interferometer as a Means to Measure Phase Noise

E. Rubiola$^{\diamond\star}$ and V. Giordano*

$\diamond$ Politecnico di Torino, c.so Duca degli Abruzzi 24, I-10129 Torino, Italy
fx +39.011.5644099, e-mail rubiola@polito.it

$\star$ LPMO CNRS, 32 av. de l'Observatoire, F-25044 Besançon, France
fx +33.381.666998, e-mail giordano@lpmo.edu

Abstract

A radiofrequency interferometer is combined with the correlation-and-averaging enhancement of sensitivity in a new scheme for the measurement of the phase noise of a component. The method requires the assumption that the non-thermal phase noise of the component being tested (DUT) is dominant compared to the amplitude noise, which is consistent with the general experience in radiofrequency domain. The new scheme is based on the amplification of the DUT noise sidebands, and on the simultaneous measurement of the amplified noise by means of two mixers driven in quadrature, $\pm 45°$ off the carrier. Due to the particular detection phase relationships, the instrument rejects the thermal noise and the noise of the measurement amplifier, despite the fact that this amplifier is shared by the two channels of the correlator. The paper provides the theoretical background of the method and experimental results. The sensitivity of the first 100 MHz prototype, given in terms of the $S_\varphi(f)$ floor, is some 12 dB below k_BT/P_0, i.e. the thermal energy referred to the carrier power.

1 Introduction

The double interferometer, first proposed in [1], was aimed to increase the sensitivity of the interferometric noise measurement system by correlating and averaging the output of two equal interferometers that simultaneously measure the same device under test (DUT). It become soon clear that the sensitivity of the double interferometer is not limited by the thermal energy k_BT, where $k = 1.38\times10^{-23}$ J/K is the Boltzmann constant and T is the absolute temperature of the machine [2, 3]. This feature does not relies upon a specific technology because the reported results were obtained in the VHF and microwave bands with quite a different type of components.

The microwave interferometer as a phase noise measurement tool was proposed in the late 60ties [4], and subsequently improved and extended to the active correction of noise [5]. Further studies [6] provided increased theoretical insight and new design rules that improve the overall performances and extend the operating frequency to the HF and VHF bands, and that are also valid for the double interferometer. In addition, a modified scheme [7] allows the measurement of the frequency stability of quartz resonators with a sensitivity of 10^{-14}.

Finally, it was proved theoretically and experimentally [8, 9] that the high sensitivity of the double interferometer relies upon the difference of spectrum densities, and therefore the thermal noise is rejected if the temperature of the instrument is homogeneous. Later on, the theory was confirmed by other research groups [10, 11]. Furthermore, the same approach makes possible the realization of high sensitivity spectrum analyzers. Prototypes operating at audio and microwave frequencies [12] show a noise floor significantly lower than k_BT.

The following sections recall the basic mechanism of noise rejection, introduce a new scheme for noise measurement and present experimental results.

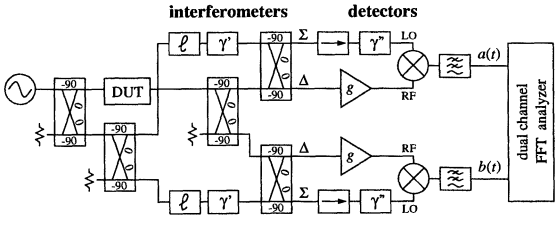


Figure 1: Double interferometer.

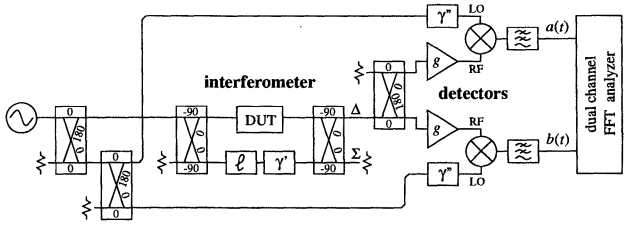


Figure 2: Correlation interferometer.

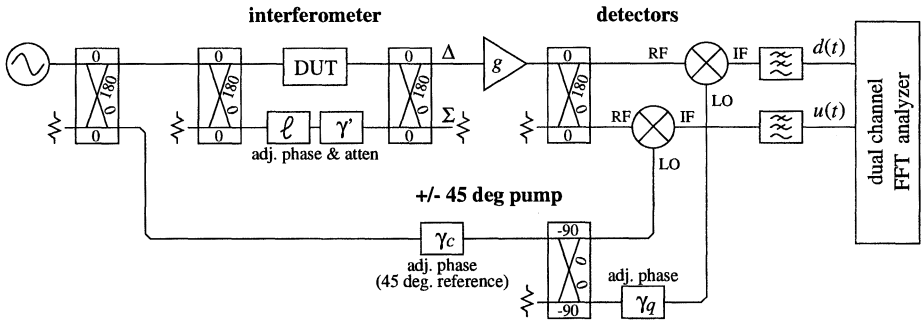


Figure 3: Single-amplifier correlation interferometer.

2 Basics of the Noise Rejection by Correlation

When the sum and the difference

$$a(t) = \frac{1}{\sqrt{2}}\left[n_1(t) + n_2(t)\right] \quad (1)$$

$$b(t) = \frac{1}{\sqrt{2}}\left[n_1(t) - n_2(t)\right] \quad (2)$$

of two independent random signals $n_1(t)$ and $n_2(t)$ are inserted in the cross correlation function

$$\mathcal{R}_{ab}(\tau) = \lim_{\theta\to\infty}\frac{1}{\theta}\int_{\theta} a(t)\, b^*(t-\tau)\, \mathrm{d}t \quad (3)$$

they result in the difference of the two indivindual correlation functions

$$\mathcal{R}_{ab}(\tau) = \frac{1}{2}\left[\mathcal{R}_{n_1}(\tau) - \mathcal{R}_{n_2}(\tau)\right] \quad (4)$$

because $\mathcal{R}_{n_1 n_2}(\tau) = \mathcal{R}^*_{n_2 n_1}(\tau) = 0$.

The cross power spectrum density (PSD) of $a(t)$ and $b(t)$ is, by definition,

$$S_{ab}(f) = \int_{\infty} \mathcal{R}_{ab}(\tau)\exp(-2\pi f\tau)\,\mathrm{d}\tau \ , \quad (5)$$

where the integral performs the Fourier transform. Indicating with $N_1(f)$ and $N_2(f)$ the PSD of $n_1(t)$ and $n_2(t)$, the cross PSD is

$$S_{ab}(f) = \frac{1}{2}\left[N_1(f) - N_2(f)\right] \quad (6)$$

Dividing $N(f)$ in thermal noise $N^{\mathrm{th}}(f) = k_B T$ and extra noise $N^{\mathrm{ex}}(f)$, we get

$$S_{ab}(f) = \frac{k_B\,(T_1 - T_2)\,R_0}{2} + \frac{N_1^{\mathrm{ex}}(f) - N_2^{\mathrm{ex}}(f)}{2} \ . \quad (7)$$

This means that if port 1 is regarded as the input of an instrument and port 2 is terminated to a resistor (hence $n_2(t)$ is a pure thermal fluctuation), and assuming that the temperature is homogeneous ($T_1 = T_2$), the instrument compensates for thermal noise and measures the extra noise $N_1^{\mathrm{ex}}(f)$ only. Alternatively, both ports 1 and 2 are terminated to a resistor. In this condition, the instrument noise floor gives the temperature inhomogeneity $T_1 - T_2$. It is worthly to be remarked that this noise compensation mechanism is based only on the mathematical properties of noise. Thus $n_1(t) + n_2(t)$ and $n_1(t) - n_2(t)$ can be obtained either by means of a lossless 3 dB coupler or just as the result of an algebraic operation.

The above noise rejection mechanism is exploited by the double interferometer, that rejects the thermal noise generated by the three resistive terminations (Fig. 1). In addition, this instrument rejects the interferometer noise because the two interferometers are independent, and it rejects the amplifier noise for the same reason. All the analytical calculations, involving the random signals of the three terminations, are reported in [8] and [9]. A simplified scheme (Fig. 2) uses one interferometer instead of two, which makes the rejection of the interferometer noise impossible. Nonetheless, the rejection of both thermal and amplifier noise still takes place, and works in the same way.

3 Single-Amplifier Correlation Measurements

The scheme of the proposed instrument is shown in Fig. 3. In the presence of noise, the signal at the DUT output is of the form

$$s(t) = V_0\cos\omega_0 t + n_s(t)\sin\omega_0 t + n_c(t)\cos\omega_0 t \quad (8)$$

where V_0 is the peak voltage of the carrier at frequency $\nu_0 = \omega_0/2\pi$, and $n_s(t)$ and $n_c(t)$ are independent random voltages. Taking $\cos\omega_0 t$ as the phase reference, $n_s(t)$ represents the phase noise, and $n_c(t)$ the amplitude noise. Introducing the superscripts 'th' and 'ex' for thermal and extra (i.e. non-thermal) noise, we can divide the noise as $n_s(t) = n_s^{\mathrm{th}}(t) + n_s^{\mathrm{ex}}(t)$ and $n_c(t) = n_c^{\mathrm{th}}(t) + n_c^{\mathrm{ex}}(t)$. While $n_s^{\mathrm{th}}(t)$ and $n_c^{\mathrm{th}}(t)$ are always independent, we assume that also $n_s^{\mathrm{ex}}(t)$ and $n_c^{\mathrm{ex}}(t)$ are independent. This assumption, to be checked for the specific DUT, is consistent with the observation that parametric AM and PM noise types are generally originated by different, and therefore independent, physical phenomena.

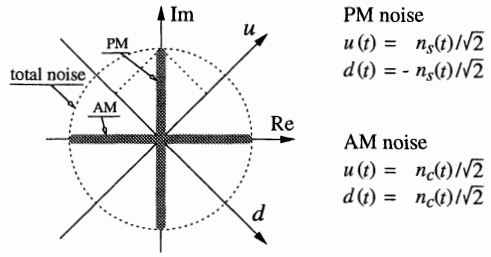


Figure 4: Vector representation of phase and amplitude noise on the complex plane, and the detection of noise on the u and d axes.

Setting the variable attenuation ℓ and the variable phase shift γ' equal to the DUT phase and attenuation, the carrier $V_0 \cos \omega_0 t$ is suppressed at the Δ output of the hybrid, and consequently only the DUT noise is present at the amplifier input. This noise is amplified and down converted to baseband. Because the phase of mixer LO signals are $\pm 45°$, the detected signals (Fig. 4) are

$$u(t) = k\,[n_c(t) - n_s(t)] \quad (9)$$

$$d(t) = k\,[n_c(t) + n_s(t)] \ , \quad (10)$$

where k is the gain that results from all the gain and losses along the signal path. A vector representation of phase and amplitude noise, and their $\pm 45°$ detection, is shown in Fig. 4; the suppressed carrier is on the real axis. The $\pm 45°$ detection is somewhat similar to the FM stereo broadcastings, in which $L \pm R$ (i.e., left $\pm$ right) are transmitted, or to the PAL television systems, in which the signals are $Y \pm C$, i.e. chrominance $\pm$ luminance. L and R, as well as Y and C, are a pair of independent, thus orthogonal, random signals.

From (9) and (9) and Fig. 3, the cross PSD at the instrument output is

$$S_{ud}(f) = \frac{gR_0P_0}{4\ell_h\ell_s\ell_m}\left[S_\alpha(f) - S_\varphi(f)\right] \ , \quad (11)$$

where g is the amplifier gain, ℓ_h is the loss of the hybrid located at the DUT output, ℓ_s is the loss of the hybrid located at the amplifier output, and ℓ_m is the SSB loss of the mixer. Equations (9) and (10) are of the same form of (1) and (2), whereat the mechanism of thermal noise rejection described in Section 2 takes place. As thermal noise is unbiased, i.e. $S_\varphi^{\rm th}(f) = S_\alpha^{\rm th}(f)$, it vanishes in the correlation-and-averaging process. Hence the extra noise only remains.

In most cases of interest the PM extra noise is dominant compared to the AM noise. Thus, taking $S_\alpha^{\rm ex}(f) \ll S_\varphi^{\rm ex}(f)$ as an hypothesis, the output cross PSD is proportional to extra phase noise $S_\varphi^{\rm ex}(f)$ of the DUT according to

$$S_{ud}(f) = -\frac{gR_0P_0}{4\ell_h\ell_s\ell_m} S_\varphi^{\rm ex}(f) \ . \quad (12)$$

It should be stressed that the proposed instrument does not give information on the validity of the assumption $S_\alpha^{\rm ex}(f) \ll S_\varphi^{\rm ex}(f)$. The latter must result from a different analysis, either theoretical or experimental. Finally, the AM noise might be dominant. If this occurrs, the instrument can be used to measure $S_\alpha^{\rm ex}(f)$ in the same way.

In the proposed scheme, the noise of the amplifier is rejected, and this occurrs despite of the fact that a single amplifier is shared by the two channels. In fact, as the carrier is suppressed, the amplifier works in its fully linear regime, thus its noise contribution is of the additive type. Thus, the amplifier noise can be represented as

$$n^{\rm a}(t) = n_s^{\rm a}(t)\sin(\omega_0 t + \gamma) + n_c^{\rm a}(t)\cos(\omega_0 t + \gamma) \ , \quad (13)$$

where γ is an arbitrary phase. As a property of random noise, $n_s^{\rm a}(t)$ and $n_c^{\rm a}(t)$ are independent, and this holds for any γ. Therefore, the amplifier noise vanishes because the axes u and d are in quadrature.

The white noise at one output, not rejected by correlation, is $S_{u0}(f) = S_{d0}(f) = gFk_BT_0R_0/\ell_s\ell_m$, which corresponds to a single-arm phase noise floor $S_{\varphi 0}(f) = 4\ell_h F k_B T_0/P_0$. The cross PSD floor is reduced increasing the number m of averages until the hardware limitation, which is due to the noise of the interferometer components – mainly the variable attenuator and phase shifter responsible of the carrier suppression – and to the accuracy of the quadrature condition between the two detectors. Finally, it should be remarked that the sensitivity of other correlation schemes may be limited by the isolation between the two amplifiers, as discussed in [11], and that this problem is avoided with the single-amplifier scheme.

4 Implementation and Results

The prototype we experimented on works at the carrier frequency $\nu_0 = 100$ MHz, with a carrier power $P_0 = 4$ dBm at the DUT output. The amplifier gain is $g = 41.5$ dB and a noise figure $F \simeq 1.5$ dB. The hybrid losses are $\ell_h = 1$ dB, at the DUT output and $\ell_s = 0.4$ dB at the amplifier output, and the SSB loss of the mixers is $\ell_m = 5.5$ dB. The residual carrier never exceeds -20 dBm at the amplifier output, which is some 35 dB below the 1 dB saturation level. All the design follows the rules given in [6].

The adjustment and calibration procedure is stright-forward. The interferometer is firstly replaced with a synthesizer connected to the input of the amplifier. Thus a signal of power P_s and frequency ν_s close to ν_c is injected, which results in a beat note at the frequency $f_s = \nu_s - \nu_0$ of some kilohertz at the two outputs. The corresponding rms voltage is $V = \sqrt{gP_sR_0/2\ell_s\ell_m}$ independent of the phase setting, which gives the SSB

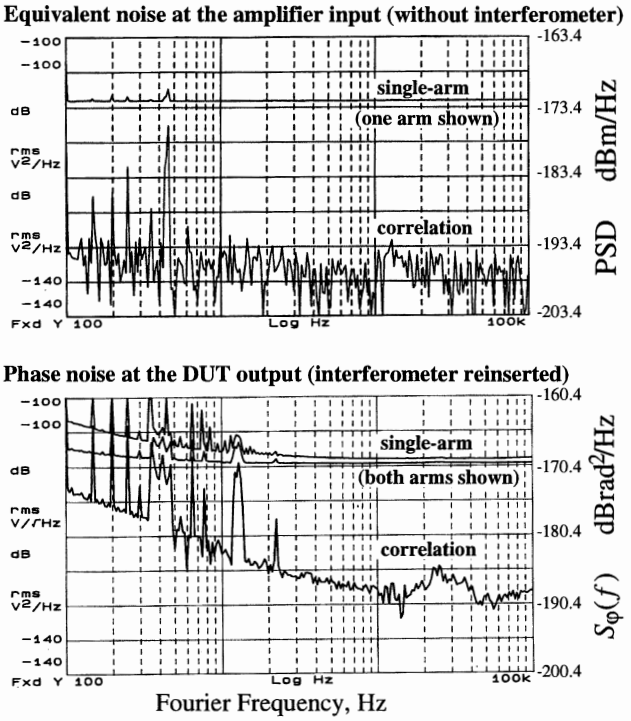


Figure 5: Instrument noise, in the absence of the DUT.

gain $V^2/P_s = gR_0/2\ell_s\ell_m$ of the instrument. Then, the phase γ_q is adjusted for the two output to be in quadrature, which is accomplished observing the output cross spectrum when the amplifier is driven by the synthesizer. As $S_{ud}(f_s)$ is imaginary when the two detectors are in quadrature, the instrument is adjusted for $\Re\{S_{ud}(f_s)\} = 0$. At this stage, the interferometer is reinserted, removing the syntesizer; a phase modulator, calibrated by means of a network analyzer, is inserted as the DUT. Subsequently, the common phase γ_c must be set for the two detectors to be at $\pm 45°$ with respect to the carrier. This is accomplished by provisionally setting the machine for arm u to detect AM and arm d PM, which is observed as a null of $d(t)$ in the presence of a phase modulation, and then replacing the cable in series to γ_c with a 45° longer one. Finally, the instrument gain $S_{ud}(f)/S_\varphi(f)$ is measured injecting a phase modulation $S_\varphi(f)$. The measured gain is of 60.4 dBV/rad, with an asymmetry of some ±0.2 dB between the two arms, including two 40 dB low noise amplifiers inserted between the mixers and the FFT analyzer.

Figure 5 top shows the instrument noise resulting from a first experiment, in which the interferometer is removed and the amplifier is terminated to a resistor. Results are given in dBm/Hz, which is the phisical unit of the PSD of bare random noise, not referred to a carrier. The single-arm noise is of some −172.5 dBm/Hz, which is consistent with the amplifier noise figure $F \simeq 1.5$ dB. The absence of flicker, evident in the left part of the figure, is due to the absence of a

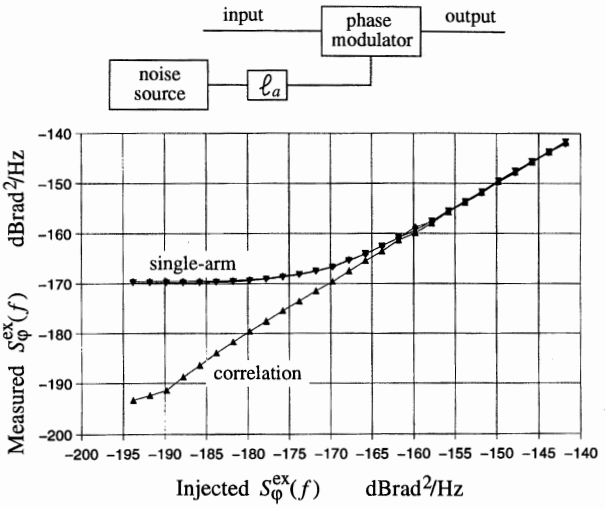


Figure 6: Injection of phase noise in the DUT path.

carrier signal. The cross PSD reveals a noise rejection of 22 dB, which is limited by the averaging capability ($m = 32767$) of the FFT analyzer. The noise floor is of some 20 dB below the thermal energy k_BT, and it is expected to further decrease if the number of averages is increased.

The instrument noise measured on the complete configuration is shown in Fig. 5 bottom. The single-arm noise reveals the presence of flicker around $f = 100$ Hz, and some asymmetry between the two arms. The single-arm white noise is the same as observed in the absence of the interferometer. The cross PSD reveals the presence of flicker $S_\varphi \simeq -174$ dBrad²/Hz at $f = 100$ Hz, which is due to the interferometer. A spurious signal is present in the 15 to 50 kHz region. For $f > 7$ kHz, and depending on frequency, the instrument noise is of some −188 to −190 dBrad²/Hz, which is 10 to 12 dB lower than k_BT/P_0.

A more conclusive experiment consists of the injection of a phase modulation in the DUT path, according to the scheme of Fig. 6 top. A special phase modulator is needed, whose low noise feature is obtained combining a low insertion loss with a small modulation index capability. The modulator is driven by the internal noise source of the FFT analyzer, attenuated by a factor ℓ_a. The measurement results are shown in Fig. 6 bottom. As ℓ_a increases (leftwards in the figure), the injected noise S_φ^{ex} decreases proportionally to $1/\ell_a$. In the left part of the figure, for $S_{\varphi\,inj}^{ex} < -160$ dBrad²/Hz, the single-arm measures are limited by the amplifier noise. On the other hand, the correlation measures are well aligned on the strigh line $S_{\varphi\,meas}^{ex} = S_{\varphi\,inj}^{ex}$ even on the left part of the plot, where the injected signal is lower than the thermal floor $k_BT/P_0 \simeq -178$ dBrad²/Hz.

Acknowledgments
We wish to acknowledge M. Olivier and J. Groslambert from the LPMO, and G. Brida from the IEN for valuable scientific discussions, H. Stoll from the Max Planck Institute and M. Elia from the Politecnico di Torino for their great contribution to the theoretical basis of this work.

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] 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.

[3] E. Rubiola, V. Giordano, J. Groslambert, "Improved interferometric method to measure near-carrier AM and PM noise", *IEEE Transactions on Instrumentation and Measurement* vol. 48 no. 2 pp. 642–646, April 1999.

[4] K. H. Sann, "The measurement of near-carrier noise in microwave amplifiers", *IEEE Trans. on Microwave Theory and Techniques* vol. 16 no. 9 pp. 761–766, September 1968.

[5] E. N. Ivanov, M. E. Tobar, R.A. Woode, "Microwave Interferometry: Application to Precision Measurements and Noise Reduction Techniques", *IEEE Trans. on Ultrasonics, Ferroelectrics and Frequency Control* vol. 45 no. 6 pp. 1526–1535, November 1998.

[6] 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.

[7] E. Rubiola, J. Groslambert, M. Brunet, V. Giordano, "Flicker noise measurement of HF quartz resonators", *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control.* vol. 47 no. 2 pp. 361–368, March 2000.

[8] E. Rubiola, V. Giordano, "Correlation-based noise measurements below the thermal noise floor", Proc. *13th European Frequency and Time Forum / 1999 Frequency Control Symposium* vol. 2 pp. 1125–1128, Besançon (France), 12–16 April 1999.

[9] E. Rubiola, V. Giordano, "Phase noise metrology", invited paper, in M. Planat (Editor) *Noise, oscillators and algebraic randomness*, Springer Verlag, 2000 (expected to be available in Jun).

[10] E. N. Ivanov, F. L. Walls, "Interpreting noise floors in cross-correlated two-channel PM/AM noise measurement systems", Proc. *14th European Frequency and Time Forum* (these proceedings), Torino (Italy) 14–16 March 2000.

[11] A. De Marchi, F. Andrisani, G. P. Bava, "Sensitivity and noise limits of microwave interferometers and correlated double interferometers", Proc. *14th European Frequency and Time Forum* (these proceedings), Torino (Italy) 14–16 March 2000.

[12] E. Rubiola, V. Giordano, "A correlation-based noise measurement scheme showing a sensitivity below the thermal floor", Proc. *15th International Conference on Noise in Physical Systems and* $1/f$ *Fluctuations* pp. 483–486, Hong Kong (Federal Republic of China), 23–26 August 1999.

PHASE NOISE LIMITS OF THE ANTI-JITTER CIRCUIT AND ON-CHIP RC OSCILLATORS

Michael J. Underhill

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

ABSTRACT

The Anti-Jitter Circuit (AJC) can be cascaded to give 60dB or more reduction in phase noise This paper compares the ultimate noise performance limits of the AJC and on-chip oscillators. Except perhaps at high microwave frequencies it is difficult to implement oscillator inductors on an integrated circuit chip and such inductors will in any case have a poor Q factor. For this reason the noise comparison is also made with RC on-chip oscillators at lower frequencies.

The aim of this paper is to show that where the current source is mainly from "majority carriers" the Shot noise can be significantly below the maximum. Theoretical improvements of 25 to 40dB or more would appear to be possible over what previously has been stated. This indicates that the AJC used as or with an on-chip RC oscillator will certainly give better performance than existing on-chip LC oscillators. It now appears that an AJC may also be capable of reaching the performance of even the best LC voltage controlled Oscillators.

INTRODUCTION

The Anti-Jitter Circuit (AJC) was first introduced at the EFTF in 1996 [1] and it has undergone progressive improvement since then[2,3,4]. These improvements have been in highest frequency of operation, frequency range, lowest suppression frequency, settling time, power consumption and ease of integration. It remains as a unique circuit that can be cascaded to produce more than 60dB of phase noise in principle on any frequency source. The limit of suppression is set by the internal noise sources of the circuit and this paper extends previous investigations of these [2].

In the last year the AJC has been redesigned so that it now can be implemented in all bipolar as well as all FET or mixed FET and bipolar technologies. The major problem was that the integrator current source transistors could not be made with low enough gain or transconductance and still deliver the milliamp or so currents that are typically required in the AJC. The original AJC circuits assumed that a FET with sufficiently low transconductance could always be made available by choice of the appropriate gate width-to-length ratio in manufacture. In the bipolar transistor the transconductance is directly related to the transistor current and so this degree of freedom is not then available.

The use of bipolar devices also made it difficult to obtain high enough input impedances to be able to obtain long enough time constants for a low enough frequency limit of sideband suppression; that is without the use of large value capacitors that could not be integrated on a chip. The analysis in this paper shows that a single relatively large (1-100nF) off-chip capacitor can have a highly beneficial effect on the ultimate level of noise suppression.

The acceptance of the necessity of a single off-chip capacitor has made it possible to define an optimum control loop design for the AJC. This optimum defines the best time response and frequency response achievable for the AJC when considered as a linear control system.

A new assessment of the dominant noise sources in the AJC is presented. This shows that the 'shot noise' contribution was previously overestimated [2], and often will not be the limiting source of noise.

The optimal control loop and noise design of the AJC is presented as a few simple design rules,.

REPRESENTATIVE AJC CIRCUIT

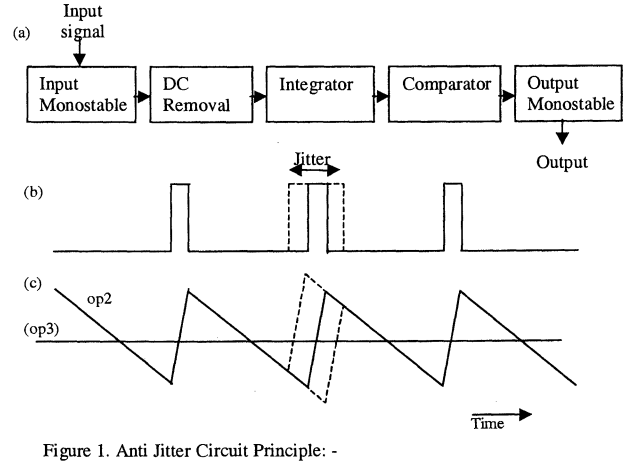


Figure 1. Anti Jitter Circuit Principle: -
(a) Basic Block Diagram
(b) Input monostable with jitter on central pulse
(c) Integrator output (op2) and Comparator switching level (op3)

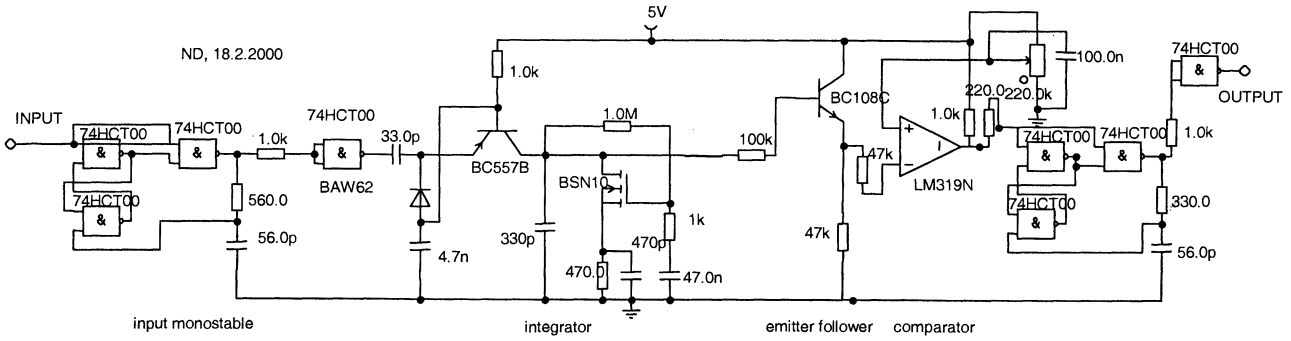


Figure 2 Practical 8 Mhz Adiabatic Anti-Jitter Circuit

The block diagram of Fig 1 shows the basic principle of the AJC. Note that the integrator waveform (long) downslope intersects the comparator switching level at the same place in time irrespective of the time jitter in the (integrated) input pulses. This is the mechanisms by which the jitter is removed.

Fig 2 shows a practical circuit of an 'adiabatic' AJC [4] using off-the-shelf (discrete) components. ('Adiabatic' means that the core part of the AJC is powered from the input pulses.) The single package of 4 nand gates at the input make up a pulse monostable together with the timing circuit of C_a and R_a that sets the pulse length. Strictly speaking, this circuit is not necessary because the 'fixed' packet of charge (current x time) formed by switching C_3 between two voltage levels substitutes for the input pulse and creates the up-slope part of the integrator waveform. The integrator capacitor C_{int} is charged by this charge packet and discharged steadily by the FET acting as a current sink.

After a high impedance emitter follower buffer there is an op-amp comparator and a second triggered monostable to form the output pulse signal. The second monostable uses the same circuit as the first with the values shown this AJC works up to about 10MHz.

AJC CONTROL LOOP DESIGN RULES

The central 'adiabatic' part of the circuit of Figure 2 is shown in Figure 3. Figure 4 shows an 'all-bipolar' version of the central adiabatic core of the AJC. It is the components in this core parts of the AJC that have to be chosen according to a set of 'design rules' so that the AJC has 'best' performance. The approach is an extension of what was in previous papers [2,3,4].

Best control loop performance is when good phase and time jitter suppression is achieved over as wide a range of modulation sideband frequencies as possible. This is achieved if the combination of C_{int} and the FET (or bipolar) current sink behaves as an integrator down to the lowest possible (suppression) frequency.

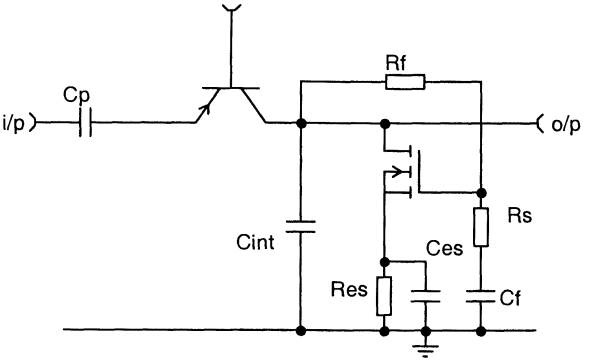


Figure 3 FET-Bipolar AJC Adiabatic Core Circuit

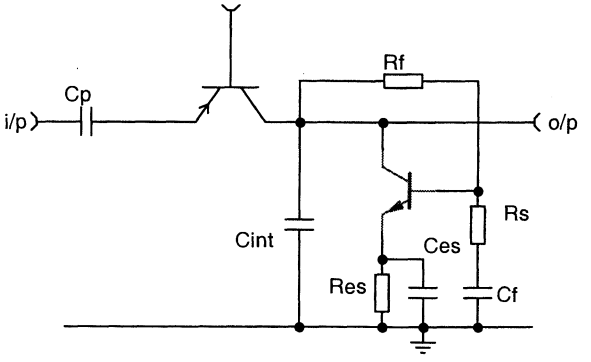


Figure 4 All-Bipolar AJC Adiabatic Core Circuit

The AJC can be treated as a control loop where the current sink is considered as an input. The objective is to design this control loop so that transfer function from the current i input in to the voltage V_{int} on the integrator is $1/s$ because this represents perfect integration with respect to time. The actual transfer function, taking the feedback through R_f into account, is found to be

$$G_{CL}(s) = V_{int}/i = g_m(1 + sC_f(R_f + R_s))/$$
$$(s(C_f + C_{int} + sC_{int}C_f(R_f + R_s)) + g_m(1 + sC_fR_s)) \quad (1)$$

One open loop pole in the denominator of this expression is at $s = 0$ The other is at the important break frequency $|s| = a$ of the components C_{int}, C_f, R_f, and R_s, all connected in series. We therefore define this as

$$a = (C_f + C_{int})/C_{int}C_f(R_f + R_s) \quad (2)$$

Critical damping for the above two closed-loop poles is found to give an optimum compromise between settling speed and lowest suppression frequency. Critical damping occurs for the loop gain condition:

$$(C_f + C_{int} + g_m C_f R_s)^2 = 4 g_m C_{int} C_f (R_f + R_s) \quad (3)$$

As a limiting case we can choose to make the resistor R_s zero and this gives

$$G_{CL}(s) = V_{int}/i$$
$$= g_m(1 + sC_fR_f)/(s(C_f + C_{int} + sC_{int}C_fR_f) + g_m) \quad (4)$$

The critical damping condition then becomes

$$(C_f + C_{int})^2 = 4 g_m R_f C_{int} C_f \quad (5)$$

The usual AJC design has $C_f \gg C_{int}$ and then the optimum loop gain condition is:

$$C_f = 4 g_m R_f C_{int} \quad (6)$$

As another limiting case we can choose the resistor value R_s so that

$$C_fR_s = C_{int}R_f. \quad (7)$$

With $C_f \gg C_{int}$ and hence $R_f \gg R_s$ the transfer function then becomes

$$G_{CL}(s) = V_{int}/i = g_m(1 + sC_fR_f)/$$
$$(sC_f + g_m)(1 + sC_{int}R_f) \quad (8)$$

The position of only one of the denominator poles is a function of the transistor transconductance g_m. As a compromise between settling speed and lowest suppression frequency we choose to place the movable pole at about a quarter of the frequency of the fixed pole and this conveniently gives the same condition as shown in equation (6)

$$C_f = 4 g_m R_f C_{int} \quad ((6)) \quad (9)$$

This condition gives the minimum value for C_f. Smaller values than this will result in a suppression frequency that is unnecessarily higher than the minimum set by the value of a defined above

An intermediate case is with the capacitor C_f slightly too small and R_s increased from zero until critical damping is again achieved. This occurs approximately when

$$C_f = 4 g_m R_f C_{int}/(1 + g_m R_s)^2 \quad (3)$$

and the value for R_s can be found by solving this equation for the given value of C_f.

In practice the main consequence of this rule is that C_f becomes too large to be considered for integration on a chip unless the transconductance g_m is made suitably low. In the case of a FET this can be achieved by choice of transistor geometry on the chip. In the case of a bipolar transistor the transconductance is directly related to the emitter current or the emitter resistor value (if not decoupled) It is then not possible to obtain a low enough g_m with the integrator capacitor discharge current at the values required for optimum AJC performance.

As discussed above the resistor R_s can conveniently be adjusted to prevent underdamped oscillations if the capacitor C_f is not seriously too small. Too small a value of C_f will in general result in a pro rata rise in the lowest frequency of suppression.

The above are the important control loop design rules for ensuring optimum performance of the AJC and they are incorporated in the overall list of design rules given below

AJC NOISE ANALYSIS

Minimum noise design of the AJC is now addressed. Previous papers have identified Shot noise and thermal noise in resistors as the main sources of noise to be minimised [2].

The Shot noise is assumed to occur on the discharge current i_{dis} In the previous paper [2] it was assumed that this noise would for the entire current i_{dis} This would only be true if the current were entirely made up of minority carriers for which the Shot noise partitioning effect would hold. In a bipolar transistor it is argued that only the minority carriers from the base current contribution to the collector current should be taken into account. This means that the Shot noise

should be less by the current gain β of the transistor. For the typical transistor gain of 100 the noise would be improved by 20dB. The Shot noise current i_n is therefore given by

$$(i_n / i_{dis})^2 = 2 q_e / i_{dis} \beta \qquad (10)$$

where the electron charge $q_e = 1.6\times10^{-19}$ coulomb.

When converted by the discharge slope to time and hence to phase deviation, this noise is found to result in an RMS phase deviation of

$$\phi_n = (i_n / i_{dis}) (f_o / f_m) \qquad (11)$$

where f_o is the carrier output frequency and f_m is the sideband frequency.

But we then have for the phase noise power spectral density

$$S_\phi(f) = \tfrac{1}{4} | \phi_n |^2 = (q_e / i_{dis} \beta)(f_o / f_m)^2 \qquad (12)$$

For an AJC operating at 10MHz with only one milliamp of discharge current from a bipolar transistor with a current gain β of 100 we have at 10 kHz from carrier

$$S_\phi(f) = 8 \times 10^{-13} \equiv -121\text{dBc/Hz at 10kHz} \qquad (13)$$

This is found to be better than many synthesised signal generators at this frequency.

The dominant resistor noise for a bipolar transistor is assumed to come from the resistance connected between the base and emitter. In a typical AJC this would be comparable with or less than the typical device noise resistance of about 1k. However if the series resistor R_s is larger than this it becomes the dominant source of resistor noise. Thus R_s should be kept as low as possible by appropriate choice of C_f.

It is useful to compare the Shot noise and resistor noise in this example. For a 1k noise resistance the RMS current noise in 1Hz bandwidth is

$$i_R = (4\ kT / R)^{0.5} = 4 \times 4\times10^{-21} / 10^3 = 4\times10^{-12} \text{ amps} \qquad (14)$$

For a bipolar transistor with a current gain of 100 and collector current of 1milliamp the equivalent RMS base current Shot noise from equation (10) is

$$i_n = (2 q_e i_{dis} / \beta)^{0.5} = (2\times1.6\times10^{-19} / 100)^{0.5} = 5.66\times10^{-11} \text{ amps} \qquad (15)$$

This shows that the estimate of Shot noise may still be too high because it implies an equivalent Shot noise resistance of about 200k This is far higher than observed in practice. .

If then resistor noise is found to be dominant from further measurements we can expect a well designed AJC to be 200 times or 23dB better than given in the above example in equation (13). This would equate to the performance of the best available Voltage Controlled Oscillators (VCOs).

The above noise analysis has been for a bipolar transistor current source. On the basis that FET and bipolar low noise amplifiers in general have comparable performance the same can be expected when these are used in the AJC.

AJC DESIGN RULES

From the above analysis it is now possible to articulate some design rules for the (adiabatic) AJC that will ensure best overall performance.

1. **Choose integrator capacitor C_{int} discharge current, i_{dis} to be as high as allowed by component and system specifications.** This ensures lowest current noise contribution

2. **Increase C_{int} to give smallest ramp peak-to-peak voltage swing V_{p-p} on C_{int} that still ensures reliable comparator switching.** This ensures the lowest suppression frequency limit.

3. **Make input pulses as short as possible remembering that the average pulse current has to be equal to the discharge current i_{dis}.** This maximises the peak phase error that can be cancelled without loss of pulses.

4. **Make the combination of the feedback resistor R_f and the comparator (buffer) input loading impedance (resistance) as high as possible.** This also ensures the lowest possible frequency suppression frequency $1/(R_f C_{int})$.

5. **If AC coupling to the comparator is used ensure the coupling cut-off frequency is at least ten times lower than the lowest suppression frequency.**

6. **Choose the filter capacitor C_f according to the rule**

$$C_f = 4 g_m R_f C_{int} \qquad ((6)) \quad (9)$$

where g_m is the effective transconductance of the discharge device. If an undecoupled source or emitter resistor R_{es} is used g_m cannot be higher than R_{es}^{-1}. Note that lower g_m gives a lower value for C_f. R_f has to be kept high because of rule 4.

7. **Series resistor R_s can be adjusted to compensate for inaccurate C_f but for lowest noise it should not exceed about 1k**

COMPARISON WITH ON-CHIP OSCILLATORS

The phase noise figure of merit for an oscillator is PQ^2/Na, where P is the oscillator power presented at the input to the oscillator amplifier, Q is the tuned circuit quality factor, and Na is the effective thermal noise appearing at the input of the amplifier. Assuming an active device in the oscillator with a noise factor of no worse than about 2 we find that PQ^2 for a good VCO will be no less than about 0.1 watts or volt/amps.

In RC oscillators (with the frequency set by resistors and capacitors only) the effective Q is generally very low. For example a Wien Bridge oscillator has a Q of 1/3 and a two pole "state-space" oscillator will have an effective Q of 1 to 2. To achieve a PQ^2 value of 0.1 watts will then require a power of around 100mW or more. This is considered to be too high for a practical on-chip oscillator.

Analysis of the AJC has shown that this level of performance can be achieved with currents of about 1ma, implying a power consumption of about 10mW. The explanation of this apparent discrepancy is that a current source presents a high value of 'dynamic' resistance that is essentially noiseless. Noiseless resistances imply a highly beneficial lack of thermal equilibrium. Thus the N_a in the PQ^2/ N_a figure of merit is much lower than expected from the assumption of thermal equilibrium.

CONCLUSIONS

An AJC designed according to the design rules given here can be cascaded to suppress the phase noise of source with poor noise and high time jitter to a level of the best LC Voltage Controlled Oscillators.

Shot noise on current sources in the AJC have been overestimated and it may prove to be less important than resistor or thermal noise. Ultimate phase noise performance is now predicted to be 25 to 40dB better than previously stated.

The ultimate on-chip AJC performance is calculated to be better than for existing on-chip RC or LC oscillators.

ACKNOWLEDGMENTS

Thanks go to Neil Downie for the practical discrete component adiabatic AJC circuit shown in Figure 1 and to Mike Blewett for many of the results on which the analysis given here was based.

REFERENCES

[1] M. J. Underhill, M. J. Blewett, "Spectral Improvement of Direct Digital Frequency Synthesisers and Other Frequency Sources", 10th EFTF, Brighton, 5-7 March 1996.

[2] M.J. Underhill and M.J. Blewett, "Performance Assessment of a Delay Compensation Phase Noise and Time Jitter Reduction Method", 11th European Frequency and Time Forum, Neuchatel, 4-6 March 1997.

[3] 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", 12th EFTF, Warsaw, 10-12 March 1998.

[4] M.J. Underhill, "The Adiabatic Anti-Jitter Circuit",Joint IFCS and 12th European Frequency and Time Forum, Besancon 4-6 April 1999.

PROBABILITY PROPERTIES OF ANHARMONIC SENSORS FOR PRECISION CRYSTAL OSCILLATORS

Yuriy S. Shmaliy*, Raymond J. Besson**

*Electronics Dept., Guanajuato Univ., S/N 36730, Salamanca, Gto, Mexico, shmaliy@salamanca.ugto.mx
**LCEP, ENSMM, 26, chemin de l'Epitaphe, 25030 Besancon Cedex, rbesson@ens2m.fr

ABSTRACT

The report addresses the statistical properties of anharmonic modes of BAW crystal resonators in oscillators excited by noise or intended modulation and considered as sensors of environment. The modulated noisy model of a closed loop oscillator is studied for its amplitude-frequency and phase-frequency modulation characteristics caused by an anharmonic influence. We consider in detail the amplitude, phase, and time derivatives of the sensor signal and present their probability distributions for the different SNR.

1. INTRODUCTION

Despite of natural redundancy of a crystal resonator spectrum, design of precision crystal oscillators is focused on excitation of a general BAW mode ("C" mode for shear-thickness vibrations) jointly with suppression of another so called spurious (anharmonic) ones. It brought excellent results [1], [2] and has further potential. At the same time, many efforts are underway to reduce the oscillator operation instabilities caused by its sensitivity to aging, temperature, acceleration, magnetic field, etc., and it is naturally that resonator anharmonics seem to be very attractive precision sensors for dramatically improving an oscillator operational stability.

Once we are operating anharmonic sensors, first, we must define the rules for their excitation in precision oscillators to avoid frequency instability and significant control errors. Fig. 1 shows the signal spectrum in an oscillator of generalized anharmonic sensor responses caused by both a noise and intended modulation. Each frequency $f_{ai}(\vec{\theta}) = \nu_{ai}(\vec{\theta}) - \nu_{os}(\vec{\theta})$, where ν_{ai} is a carrier of i-th anharmonic sensor and ν_{os} is that of an oscillator, depends on an environment destabilizing vector $\vec{\theta}$ [3]. Under the noise floor excitation only, an anharmonic exhibits little power, and significant time is necessary to estimate its frequency with the desired accuracy. External modulation strongly increases the effect so that sensor signal processing may be carried out in real time. This basic concept of the recently proposed modulational method [3], provides for the optimal excitation level of the sensor to avoid its frequency interaction with the fundamental mode through the thermal field, unlike the case of self-excited dual-mode oscillators.

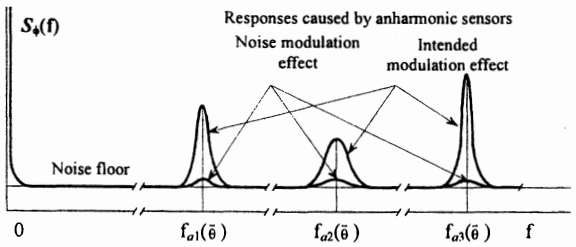


Figure 1. Generalized splashes of phase spectral density of crystal oscillator caused by anharmonic sensors.

In the report, we discuss probability properties of anharmonic sensor amplitude and phase responses.

2. FLUCTUATING MODEL OF A CRYSTAL OSCILLATOR WITH ANHARMONIC SENSORS

Let us form a closed loop oscillator with limiting amplifier of linearized gain G and linear resonant feedbacks of the resonator principle and anharmonics modes with responses $H_p(j\omega)$ and $H_i(j\omega)$, respectively (Fig. 2), where V_{out}, V_p, and V_i are correspondent voltages, V_n is white Gaussian noises with known covariance, and

$$V_{in}(t) = V_s(t) + V_n(t), \qquad (1)$$

where $V_s(t)$ is the deterministic part of the input signal and $V_n(t)$ is its random component. Generally $V_n(t)$ determines the noise floor at large offset frequencies.

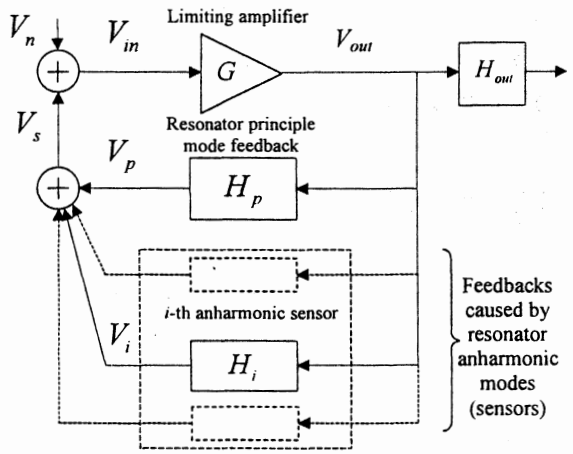


Figure 2. The closed loop model of a crystal oscillator with anharmonic sensors.

2.1. Deterministic Signals

Generally, in oscillators, a noise or intended modulation causes both AM and FM with the following signal representation

$$V_s(t) = [V_0 + \varepsilon_s(t)]\cos[2\pi v_0 t + \phi(t)], \quad (2)$$

where V_0 is a peak voltage, $\varepsilon_s(t)$ is an amplitude time variation, v_0 is a nominal carrier frequency, and $\phi(t)$ is a time variation of phase. For the simple harmonic modulation signal $S(t)$ we get $\varepsilon_s(t) = \varepsilon_{sm}\cos(\Omega t - \phi_A)$, $\phi(t) = m_\phi \sin(\Omega t - \phi_\omega) + \phi_0$, which, thus, brings (2) into the form

$$V_s(t) = V_0[1 + m_A\cos(\Omega t - \phi_A)] \times \cos[\omega_0 t + m_\phi \sin(\Omega t - \phi_\omega) + \phi_0], \quad (3)$$

where $\Omega = 2\pi f$, $\omega_0 = 2\pi v_0$, $m_A = \frac{\varepsilon_{sm}}{V_0} = k_A H_A$, ε_{sm} and $m_\phi = \frac{\omega_0}{\Omega} m_\omega$ are amplitude and phase peak variations, ϕ_0 is initial phase, f is modulation frequency, $m_\omega = k_\omega H_\omega$, k_A and k_ω are coefficients of static AM and FM, $H_A(f)$, $\phi_A(f)$ and $H_\omega(f)$, $\phi_\omega(f)$ are modulation amplitude-frequency and phase-frequency characteristics (AFC) and (PFC), respectively, for AM and FM. Non-uniformities in question for both AFC and PFC are caused by an anharmonic sensor and, for this reason, we may call them its amplitude-frequency and phase-frequency responses in oscillator signals.

Under real conditions of $m_A << 1$, $m_\omega << 1$, and $m_\varphi << 1$, we get AFC and PFC m_{Ai}, $m_{\omega i}$, ϕ_{Ai}, and $\phi_{\omega i}$ with respect to i-th anharmonic sensor:

$$m_{Ai} = \sqrt{m_{A0}^2 + 2m_{A0}k_i\cos(\phi_{A0} - \phi_i) + k_i^2}, \quad (4)$$

$$\phi_{Ai} = \arctan\frac{m_{A0}\sin\phi_{A0} + k_i\sin\phi_i}{m_{A0}\cos\phi_{A0} + k_i\cos\phi_i}, \quad (5)$$

$$m_{\omega i} = \sqrt{m_{\omega 0}^2 - 2m_{\omega 0}k_i g\cos(\phi_{\omega 0} - \phi_i) + k_i^2 g^2}, \quad (6)$$

$$\phi_{\omega i} = \arctan\frac{m_{\varphi 0}\sin\phi_{\omega 0} - k_i\sin\phi_i}{m_{\varphi 0}\cos\phi_{\omega 0} - k_i\cos\phi_i}, \quad (7)$$

where $g = \Omega/\omega_0$, $k_i(f) = V_i(f)/V_0$ is a coefficient of additional AM caused by sensor feedback, m_{A0}, $m_{\omega 0}$, ϕ_{A0}, and $\phi_{\omega 0}$ are values of AFC and PFC beyond anharmonic sensor bandwidth.

2.2. Random Signals

Because of the narrow-band nature of crystal oscillator signals, we present the random signal in (1) as a narrow-band stationary Gaussian process [4] caused by the oscillator noise floor

$$V_n(t) = A(t)\cos[\omega_b t - \Psi(t)], \quad (8)$$

where $\omega_b = \omega_0 + \Omega$, $A(t)$ is the random amplitude distributed by the Rayleigh law on the interval [0,+∞], and $\Psi(t)$ is the random phase uniformly distributed on the interval [-π, π]. Summing the most power upper spectral line of a signal fallen within bandwidth of a sensor of the frequency ω_b with a noise (8), we come from (4)—(7) to

$$m_{Ai} = \sqrt{(b_1 + A\sin\Psi)^2 + (b_2 + A\cos\Psi)^2}, \quad (9)$$

$$\phi_{Ai} = \arctan\frac{b_1 + A\sin\Psi}{b_2 + A\cos\Psi}, \quad (10)$$

$$m_{\omega i} = \sqrt{(d_1 + A\sin\Psi)^2 + (d_2 + A\cos\Psi)^2}, \quad (11)$$

$$\phi_{\omega i} = \arctan\frac{d_1 + A\sin\Psi}{d_2 + A\cos\Psi}, \quad (12)$$

where $b_1 = V_{is} + V_{+f}m_{A0}\sin\phi_{A0}$, $b_2 = V_{ic} + V_{+f}m_{A0}\cos\phi_{A0}$, $d_1 = V_{is} - V_{+f}m_{\phi 0}\sin\phi_{\omega 0}$, and $d_2 = V_{ic} - V_{+f}m_{\phi 0}\cos\phi_{\omega 0}$; V_{+f} and ϕ_{+f} are amplitude and phase of upper spectral line of an oscillator signal; V_{ic} and V_{is} are amplitudes of the orthogonal components of the sensor response; $A\sin\Psi = X_1$ and $A\cos\Psi = X_2$ are independent Gaussian random processes with zero means and equal variances $\sigma_{x1}^2 = \sigma_{x2}^2 = \sigma^2$. Following in (9)—(12) through the Gilbert transform to the analytical signal, we get the orthogonal components

$$U(t)\sin\varphi(t) = y_1(t) + X_1(t), \quad (13)$$

$$U(t)\cos\varphi(t) = y_2(t) + X_2(t), \quad (14)$$

where $y_1 = b_1, d_1$, $y_2 = b_2, d_2$, $U(t) \geq 0$, and $-\pi \leq \varphi(t) \leq \pi$. Finally, it follows from (9)—(10) that $(m_\omega, m_A) = U(t)$ and

$$(\phi_\omega, \phi_A) = \begin{cases} \varphi(t), y_2 + X_2 \geq 0 \\ \varphi(t) \pm \pi, y_2 + X_2 < 0, \begin{cases} y_1 + X_1 \geq 0 \cdot \\ y_1 + X_1 < 0 \end{cases} \end{cases}$$

Thus, noise properties of the sensor may be learned through studies of (9)—(10). Here the main task is to get probability distributions of m_A, m_ω, ϕ_A, and ϕ_ω, and their time derivatives $m'_{A,\omega}(t) = dm_{A,\omega}(t)/dt$ and $\phi'_{A,\omega}(t) = d\phi_{A,\omega}(t)/dt$ as sensor signal is slowly changed due to environment. By this way, we can estimate the optimal sensor parameters needed for optimal design of frequency control systems.

3. PROBABILITY PROPERTIES OF SENSOR RESPONSES

Pursuing the aim, we use here generalized representation of the signals due to anharmonic series tuned circuit in forms $y_1 = y_0/(1+\varsigma^2)$, $y_2 = -y_0\varsigma/(1+\varsigma^2)$, and

$$y = \frac{y_0}{\sqrt{1+\varsigma^2}}, \quad (15)$$

where y_0 is the value of amplitude y at the sensor resonance frequency ν_i, $\varsigma = Q_i(\nu/\nu_i - \nu_i/\nu)$ is generalized detuning, Q_i is a sensor Q-factor.

Let us learn the probability properties of characteristics $m_{A,\omega}(t)$, $\varphi_{A,\omega}(t)$, $dm_{A,\omega}(t)/dt$, and $d\varphi_{A,\omega}(t)/dt$. Noting that functions X_1, X_2, $X_1' = \frac{dX_1}{dt}$, and $X_2' = \frac{dX_2}{dt}$ are jointly Gaussian non-correlated random processes, we get their joint probability distribution in a form of

$$p(X_1, X_2, X_1', X_2') = \frac{1}{4\pi^2\sigma_\varepsilon^2\sigma_1^2}\exp\left(-\frac{X_1^2 + X_2^2}{2\sigma_\varepsilon^2} - \frac{X_1'^2 + X_2'^2}{2\sigma_1^2}\right), \quad (16)$$

where $X_1 = m_{A,\omega}\sin\varphi_{A,\omega} - y_1$, $X_2 = m_{A,\omega}\cos\varphi_{A,\omega} - y_2$, $X_1' = m'_{A,\omega}\sin\varphi_{A,\omega} + m_{A,\omega}\varphi'_{A,\omega}\cos\varphi_{A,\omega}$, and $X_2' = m'_{A,\omega}\cos\varphi_{A,\omega} - m_{A,\omega}\varphi'_{A,\omega}\sin\varphi_{A,\omega}$, σ_ε^2 and σ_1^2 are variances of the input random process and its time derivative, y_1 and y_2 are defined by (16).

To transfer from variables X_1, X_2, X_1', and X_2' to $m_{A,\omega}(t)$, $\varphi_{A,\omega}(t)$, $m'_{A,\omega}(t)$, and $\varphi'_{A,\omega}(t)$, respectively, we find out the Jacobean $\partial(X_1, X_2, X_1', X_2')/\partial(m_{A,\omega}, \varphi_{A,\omega}, m'_{A,\omega}, \varphi'_{A,\omega}) = m_{A,\omega}^2$ and come from (17) to the four-dimensional distribution

$$p(m_{A,\omega}, m'_{A,\omega}, \varphi_{A,\omega}, \varphi'_{A,\omega}) = \frac{m_{A,\omega}^2}{4\pi^2\sigma_\varepsilon^2\sigma_1^2} \times \exp\left\{-\frac{1}{2\sigma_\varepsilon^2}\left[m_{A,\omega}^2 - 2m_{A,\omega}(y_1\sin\varphi_{A,\omega} + y_2\cos\varphi_{A,\omega}) + y^2\right] - \frac{1}{2\sigma_1^2}\left(m'^2_{A,\omega} + m_{A,\omega}^2\varphi'^2_{A,\omega}\right)\right\}, \quad (17)$$

where envelope is in the range $0 \le V = m_{A,\omega} \le \infty$, phase in the range $-\pi \le \varphi = \varphi'_{A,\omega} \le \pi$, and correspondent derivatives are in the range $-\infty \le m'_{A,\omega}, \varphi'_{A,\omega} \le \infty$.

3.1. Modulation AFC and Their Derivatives in Time

To obtain joint distribution $p(m_{A,\omega}, m'_{A,\omega})$ we carry out series integration of (17) over $\varphi'_{A,\omega}$ from $-\infty$ to $+\infty$ and over $\varphi_{A,\omega}$ from $-\pi$ to $+\pi$, and get

$$p(m_{A,\omega}, m'_{A,\omega}) = \frac{m_{A,\omega}}{\sigma_\varepsilon^2}\exp\left(-\frac{m_{A,\omega}^2 + y^2}{2\sigma_\varepsilon^2}\right), \quad (18)$$
$$\times I_0\left(\frac{m_{A,\omega}y}{\sigma_\varepsilon^2}\right)\frac{1}{\sigma_1\sqrt{2\pi}}\exp\left(-\frac{m'^2_{A,\omega}}{2\sigma_1^2}\right)$$

where $I_0(x)$ is a modified Bessel function. Variables in (18) are separated to allow

$$p(m_{A,\varphi}) = \frac{m_{A,\varphi}}{\sigma_\varepsilon^2}\exp\left(-\frac{m_{A,\varphi}^2 + y_1^2 + y_2^2}{2\sigma_\varepsilon^2}\right)I_0\left(\frac{m_{A,\varphi}\sqrt{y_1^2 + y_2^2}}{\sigma_\varepsilon^2}\right), \quad (19)$$

$$p(m'_{A,\varphi}) = \frac{1}{\sigma_1\sqrt{2\pi}}\exp\left(-\frac{m'^2_{A,\varphi}}{2\sigma_1^2}\right). \quad (20)$$

As may be seen, distribution $p(m_{A,\varphi})$ is the Rayleigh-Rice's law, and $p(m'_{A,\varphi})$ is normally distributed with zero mean $\overline{m}'_{A,\varphi} = 0$ and variance $\sigma^2_{m'A,\varphi} = \sigma_1^2$. Fig. 3 and Fig.4 show distributions (19) for $y/\sigma_\varepsilon = 1$ and $y/\sigma_\varepsilon = 9$, respectively. Under internal noise excitation only, there is $y/\sigma_\varepsilon \le 1$, and distribution obeys the Rayleigh's law (Fig. 3). With intended modulation we get $y/\sigma_\varepsilon > 1$, so that with $y/\sigma_\varepsilon >> 1$ the distribution tends to the normal law (Fig. 4) and the variance tends to that of the noise $\sigma^2_{A,\varphi} \to \sigma_\varepsilon^2$.

3.2. Modulation PFC

To define the phase distribution $p(\varphi_{A,\omega})$ we integrate (17) $\varphi'_{A,\omega}$ from $-\infty$ to $+\infty$, over $m_{A,\omega}$ from 0 to ∞ and over $m'_{A,\omega}$ from $-\infty$ to ∞. We get as a result

$$p(\varphi_{A,\omega}) = \frac{1}{2\pi}\exp\left(-\frac{y^2}{2\sigma_\varepsilon^2}\right)\left\{1 + \sqrt{2\pi}\frac{y}{\sigma_\varepsilon}\exp\left[\frac{y^2}{2\sigma_\varepsilon^2}\cos^2(\varphi_{A,\omega} - \beta)\right]\right.$$
$$\left.\times\Phi\left[\frac{y}{\sigma_\varepsilon}\cos(\varphi_{A,\omega} - \beta)\right]\cos(\varphi_{A,\omega} - \beta)\right\} \quad (21)$$

where $\beta = \arctan(y_1/y_2)$. The mean and variance of the phase are defined as

$$\overline{\varphi}_{A,\omega} = \arctan\frac{y_1}{y_2}, \quad (22)$$

$$\sigma_\varphi^2 = \frac{\pi^2}{3} + 4\pi\sum_{n=1}^{\infty}\frac{(-1)^n}{\pi n! n^2 2^{n/2}}\left(\frac{y}{\sigma_\varepsilon}\right)^n\Gamma\left(1 + \frac{n}{2}\right), \quad (23)$$
$$\times {}_1F_1\left(\frac{n}{2}, n+1; -\frac{y^2}{2\sigma_\varepsilon^2}\right)$$

As we may see, expectation of (21) does not depend on SNR and is defined by the sensor feedback only. Fig. 3 and Fig. 4 show correspondent distributions (21) for the SNR = 1 and SNR = 6, respectively. Here, with $y/\sigma_\varphi \to \infty$, variance of the sensor phase response tends zero $\sigma_\varphi^2 \to 0$, and distribution (21) tends to delta function. It means that, with high SNR, the oscillator noise does not influence the sensor phase. This must be pointed as a splendid property. Contrary, with $y/\sigma_\varepsilon = 1$, the phase exhibits extremely small response (Fig. 3), and significant time is necessary for the phase estimation statistically.

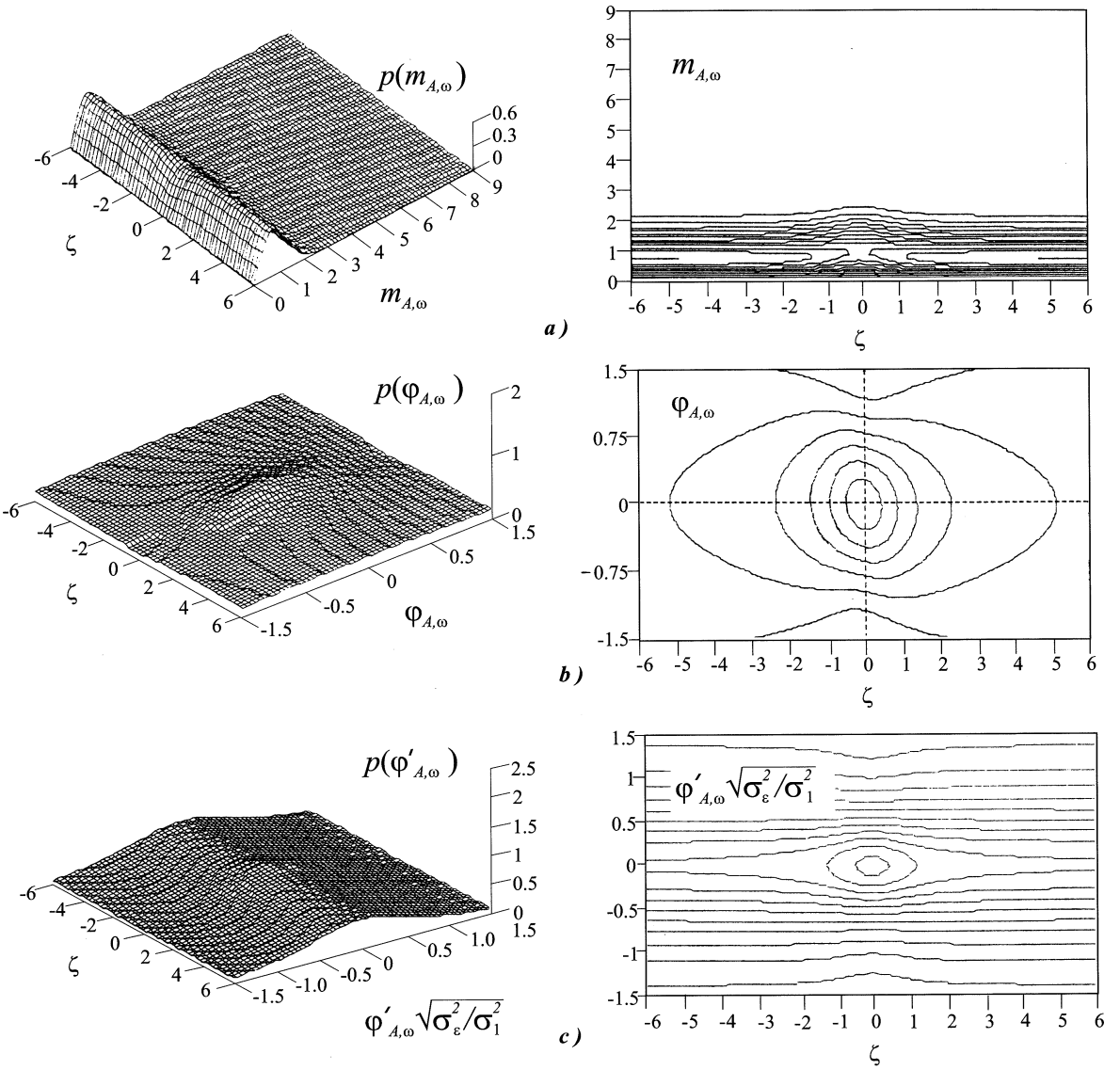


Figure 3. Probability distributions of DMC with noise modulation (SNR ≤ 1): *a*) AFC; *b*) PFC; *c*) Time derivative of PFC

3.3. Derivatives in Time of Modulation PFC

The final task is to define distribution of $\varphi'_{A,\omega}$, accounting that a sensor phase is changing in the environment. First, we integrate (17) over φ from $-\pi$ to π, then over $m_{A,\omega}$ from 0 to ∞, and, finally, over $m'_{A,\omega}$ from $-\infty$ to ∞. We get after routine transformations

$$p(\varphi'_{A,\omega}) = \frac{\sigma_\varepsilon}{2\sigma_1}\exp\left(-\frac{y_1^2+y_2^2}{2\sigma_\varepsilon^2}\right)\left(1+\varphi'^2_{A,\omega}\frac{\sigma_\varepsilon^2}{\sigma_1^2}\right)^{-3/2} \times {}_1F_1\left[\frac{3}{2},1;\frac{y_1^2+y_2^2}{2\sigma_\varepsilon^2}\left(1+\varphi'^2_{A,\omega}\frac{\sigma_\varepsilon^2}{\sigma_1^2}\right)^{-1}\right], \quad (24)$$

The distribution (24) has a symmetric shape, hence, $\overline{\varphi}'_{A,\omega} = 0$. Fig. 3 and Fig. 4 shows distribution (24) for the SNR = 1 and SNR = 6, respectively. We see that steady tendency to variance decrease of function $\varphi'_{A,\omega}(t,\Omega)$ is watched with rise of SNR at the sensor frequency. With $y/\sigma_\varepsilon = 1$, distribution (24) has the faint non-uniformity, however, if the modulation depth is increased and $y/\sigma_\varepsilon > 1$, then one may watch for the brightly pronounced distribution localization at the central frequency (Fig. 4). In analogous to $p(\varphi_{A,\omega})$, the distribution (24) with $y/\sigma_\varepsilon \to \infty$ tends to a delta-function. We pointed it above as the splendid property of a sensor phase if external modulation is introduced into an oscillator loop.

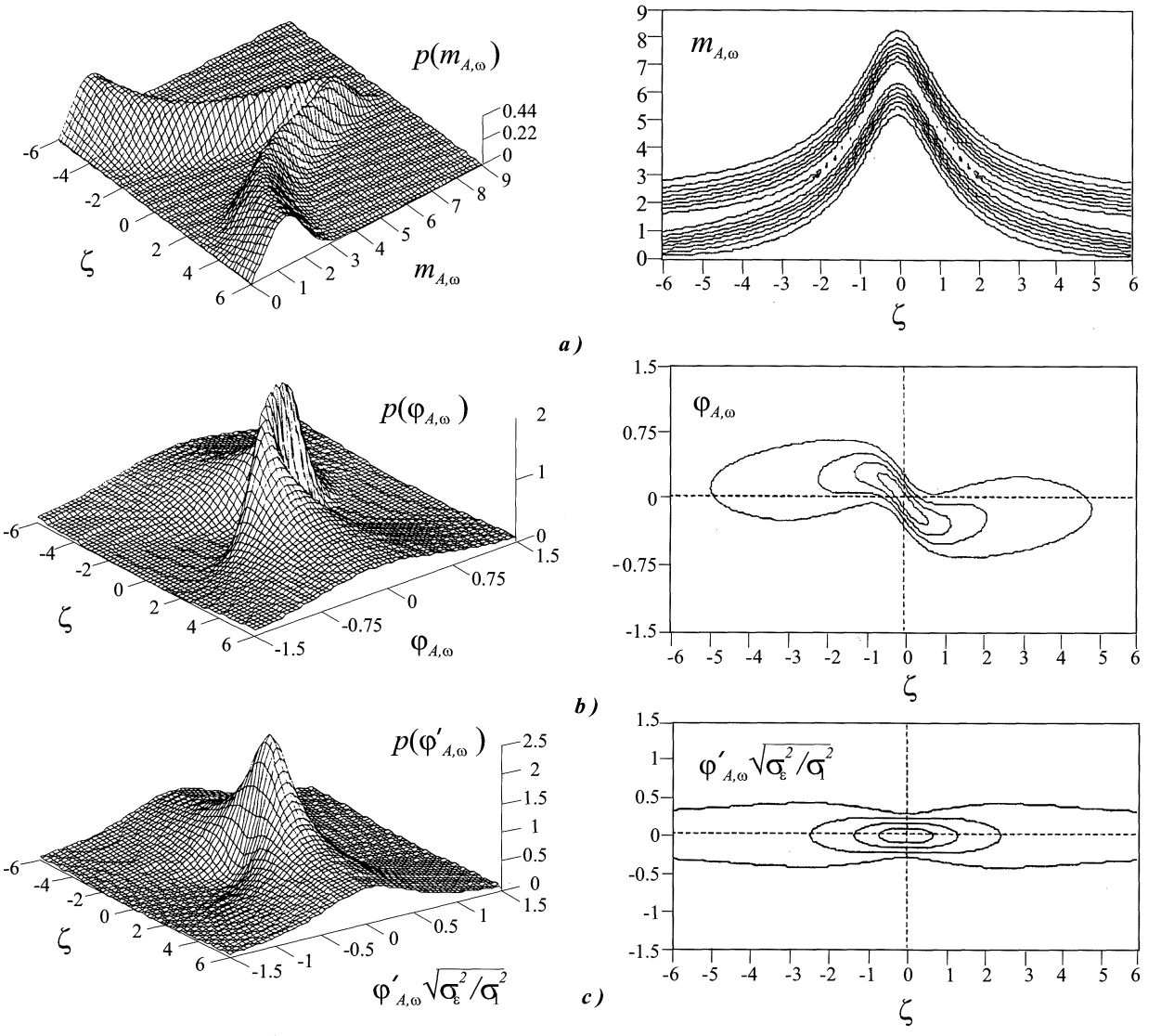


Figure 4. Probability densities with intended modulation (SNR = 6): *a*) AFC; *b*) PFC; *c*) Time derivative of PFC

5. CONCLUSION

The oscillator loop mode with additional resonant feedback caused by anharmonic sensors is characterized by modulation AFC and PFC owing to the sensor response. It strongly depends on the gain GH_i at its frequency and let us assume that it will bring about wide use of anharmonic sensors in precision crystal oscillator frequency control systems and play the critical role for the computerized OCXO and VCXO based on the modulational method.

6. REFERENCES

[1] F. L. Walls and J. R. Vig, "Fundamental Limits on the Frequency Stability of Crystal Oscillators," *IEEE Trans. Ultrason., Ferroelect., Frequency Control,* Vol.42, pp. 576-589, July 1995.

[2] R. J. Besson, "A New Piezoelectric Resonator Design," in *Proc. 30th Freq. Contr. Symp.*, 1976, pp. 78-83.

[3] Yu. S. Shmaliy, "Toward Resonator Anharmonic Sensors for Precision Crystal Oscillators: A Gaussian Model," *IEEE Trans. on Ultrason., Ferroel., and Freq. Contr.*, Vol.47, pp.379-389, March 2000.

[4] Yu. S. Shmaliy, "Probability Distributions of the Envelope and Phase and Their Derivatives in Time of the Sum of a Non-Stationary Sine Signal and Narrow-Band Gaussian Noise," *Journal of the Franklin Institute*, vol. 336, no. 6, pp.1013-1022, 1999.

ON THE ELECTROSTATIC EXITATION OF NONPIEZOELECTRIC PLATES USING NANOTUBE CARBON FILMS

G.D.Mansfeld, S.G.Alekseev, Yu.V.Gulyaev, Z.Ya.Kosakovskaya

The Institute of Radioengineering and Electronics, Russian Academy of Sciences
Mokhovaya 11, 103907, Moscow, Russia

ABSTRACT

The analysis of the possibility of the use of electrostrictive forces for the excitation of acoustic waves in solids and in resonator plates is given. High excitation efficiency is achieved due to the increase in the electric field near the electrodes made of nanotube thin films. Acoustic properties of such films are investigated. Electrical impedance of electrostrictively excited plates with nanotube carbon electrodes is evaluated.

1. INTRODUCTION

From the nonlinear theory of dielectric [1-3] it is known that the effect of electrostriction can be used for the excitation of acoustic waves in nonpiezoelectric solids if the bias DC electric field is applied to crystalline media. The electrostriction constants are usually relatively small and the effective excitation of the acoustic waves is possible if a very high intensity bias electric field is applied [4]. Carbon nanotube films are formed by carbon atoms and the both tops of the tube are of semi-spherical form. The diameters of the nanotubes and the distances between them are ranged from a few angstroms to a few nanometers. Due to very small diameters of the carbon spheres at the tops of the tubes very high values of electric fields can be easily generated near the tops of the nanotubes. The increase in the efficiency (under constant applied voltage) or the possibility to decrease bias voltage applied to the electroacoustic transducers based on electrostriction can be achieved if to use nanotube film as electrodes. But there are some problems when using these carbon nanotube films and with the evaluation of possible parameters of such transducers. The first one is the lack of the information about real acoustic parameters of the nanotubes (previously they were only predicted theoretically [6]). In this work they are found experimentally using HBAR spectroscopy [7]. The second problem is the evaluation of the intensity of the electric field generated by the film consisting of the parallel nanotubes normal to the surface of the dielectric plate. The third one is to reveal the connection between the electric field generated in dielectric by the tubes and resulting deformations. All these problems are discussed in the report. The electrical impedance of nonpiezoelectric crystalline plate electrostrictively excited by both applied constant bias and alternate electric fields.

2. HBAR SPECTROSCOPY OF NANOTUBE CARBON FILMS

The high overtone BAW composite resonator used for the evaluation of the acoustic parameters of the film is schematically shown in Fig.1. It consists of a rather thick plate made of some acoustically transparent material with flat parallel faces, an electrode layer, a piezoelectric film, and another electrode layer. The piezoelectric film deposited on one side of the plate with the electrodes serves as an electromechanical transducer providing electrical connection between the resonator and an external electrical circuit. The thin film of the nanotubes under investigation is deposited on the other face of the plate. The high overtone BAW composite resonator is a multi-frequency resonator . The resonance occurs at the frequency $\omega = \omega_n$ if an integer number n of half-wavelengths are equal to the total thickness of the structure.

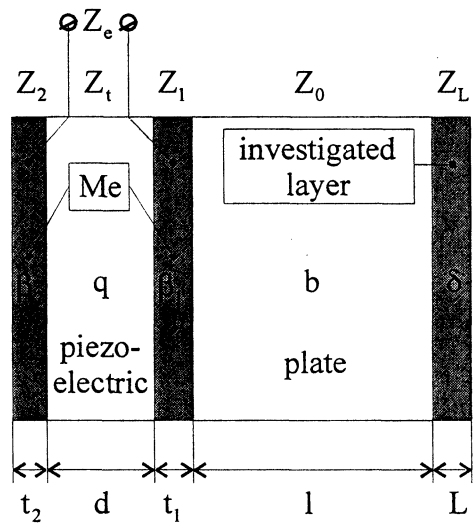


Fig. 1. The schematic diagram of the BAW composite resonator with the layer under investigation.

From the analysis of the resonant properties of the structure acoustic properties of the layer under investigation can be found using

$$\frac{Z_L}{Z_0}\tan\delta_n L = F_n(\omega_n) \qquad (1)$$

where

$$F_n(\omega_n) = -\frac{A}{D},$$

$$A = \frac{Z_t}{Z_0}\tan\left(\phi_2^{(n)} + q^{(n)}d\right) + \frac{Z_1}{Z_0}\tan w_1^{(n)} +$$

$$+ \tan b^{(n)} l\left[1 - \frac{Z_t}{Z_1}\tan\left(\phi_2^{(n)} + q^{(n)}d\right)\tan w_1^{(n)}\right]$$

$$B = 1 - \frac{Z_t}{Z_1}\tan\left(\phi_2^{(n)} + q^{(n)}d\right)\tan w_1^{(n)} -$$

$$- \tan b^{(n)} l\left[\frac{Z_t}{Z_0}\tan\left(\phi_2^{(n)} + q^{(n)}d\right) + \frac{Z_1}{Z_0}\tan w_1^{(n)}\right]$$

$$\phi_2^{(n)} = Arc\tan\frac{Z_2}{Z_t}\tan w_2^{(n)}$$

q, b, β_1, β_2, δ are complex vectors (taking into account the losses in the layers e.g. $q = q' - iq''$) d, l, t_1, t_2, L are the corresponding thickness of the layers.

When the layer under investigation is absent $F_n(\omega_n) = 0$. When the layer under investigation is present the shift in ω_n obeys (2). This expression is convenient for the calculation of the material density of the film. In case $\delta_n L << 1$

$$\rho_l = \frac{\rho_0 F_n(\omega_n)}{b_n L} \qquad (2)$$

One can find the sound velocity in the layer v_L with the numerical solution of the equation

$$\frac{\tan\omega_n L/v_L}{\tan\omega_m L/v_L} = \frac{F_n(\omega_n)}{F_m(\omega_m)} \approx \frac{\Delta\omega_n}{\Delta\omega_m} \qquad (3)$$

All the necessary quantities involved in (2,3) can be taken from the experimental data on the measurements of the frequency shifts on various harmonics ω_n and ω_m.

3. RESULTS AND DISCUSSION

In our experiments we used high overtone bulk acoustic wave resonators consisting of a [100] oriented YAG plate (180 μm thickness) with flat parallel faces with a piezoelectric ZnO thin-film (1,5μm) and Pt-electrodes (100 nm) deposited on one of the plate surfaces. The nanotube carbone layer under investigation was deposited onto the other free surface of the YAG plate. The nanotube films were synthesized by electron-beam evaporation of pure graphite in vacuum [8]. Tunnel electron microscopy image of the film is shown in Fig.2

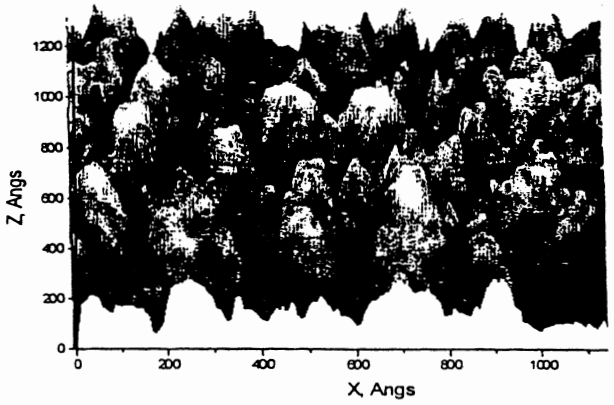


Fig.2. Tunnel electron microscope image of investigated film

In the experiments the frequencies of some series of the resonant peaks of the resonator in the frequency range between 1500 and 2000 MHz were measured before and after the the nanotube film was deposited on it.

To find material densities of the nanotube carbon films «thin» films were used (the thickness was 140 nm). In the frequency band (1500MHz-2000MHz) the thickness of the films was much smaller than a quarter of a wavelength. The obtained (using (3)) values of the material densities for the different nanotube films are: 2.0± 0.1 g/cm^3 and 2.1 ± 0.1 g/cm^3. It is interesting to compare these data with those of graphite: for usual non-ideal samples typical values are near 2.2 g/m^3.

For the velocity measurements we used "thick" films (thickness 246 nm). The best fit of the experimental points to theoretical estimation is achieved for sound velocities near 20 x 10^5 cm/s.

The estimated value of c_{33} is of the order of 1 TPa. This value is in accordance with [5-6].

In our experiments we could not find any increase in the attenuation in the composite BAW resonator connected with the presence of the nanotube film.

4. ON THE USE OF THE NANOTUBES FOR AW EXCITATION

The effect of electrostriction (tensor components are χ_{ijkl}) may be considered as piezoeffect proportional to the DC bias electric field E_l

$$e_{ijk} = \chi_{ijkl} E_l$$

Due to very small diameters of carbon spheres on the tops of the tubes very high values of electric fields may be easily generated near the tops of the nanotubes. It makes it possible to increase the efficiency of capacitive electroacoustic transducers based on electrostriction and to use such transducers for resonant excitation of

nonpiezoelectric plate as it is shown in fig.3. To obtain effective electromechanical coupling such resonator is placed in a bias DC electric field.

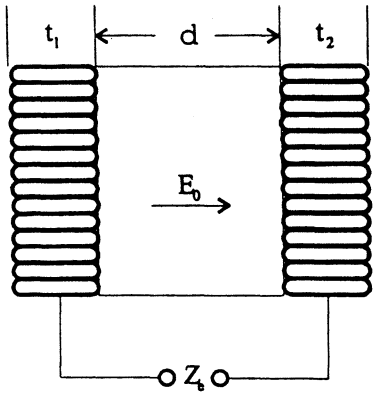
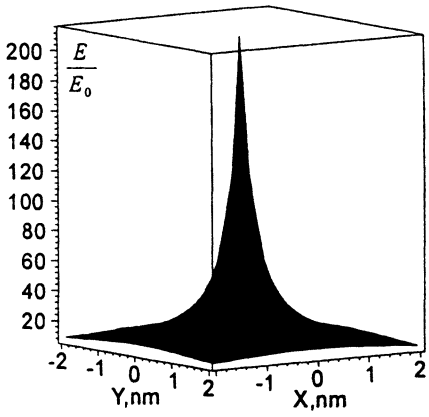


Fig. 3. The schematic diagram of the BAW composite resonator made of nonpiezoelectric plate (thickness d) with carbon nanotube film electrodes (thickness t_1 and t_2).



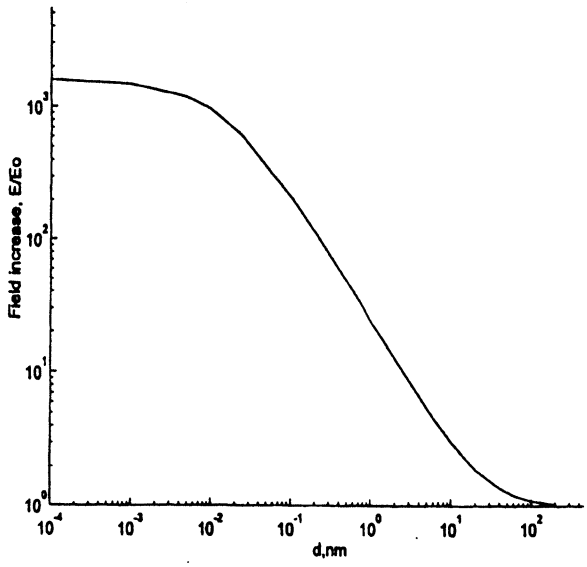


Fig. 4. a) Calculated value of electric field near the single nanotube. b) Dependence of the electric field on the distance from the top of nanotube.

Electric field intensity near the tops of single nanotubes was found from the direct solution of the Laplace equation for electric potential. The single nanotube was modeled as half of a conductive ellipsoid placed perpendicularly to one of the plane capacitor plate. It was found that electric field intensity decays in a very small distance from the tube. So, it was possible to neglect charge redistribution due to mutual influence of the neighbor tubes and to use the superposition principle for the evaluation of total field intensity. The example of such dependence of z component of the electric field on the transverse X and Y coordinates is shown in Fig.4a (the distance from the top z = 0.1 nm) and for the distance from the top is shown in Fig.4b. Here E_0 corresponds to the field in a parallel plate capacitor. The result of the calculation of the resulting electric field for the group of 5 x 5 nanotubes (model of the bundle) is presented in Fig.5. The diameters of the nanotubes are 0.5 nm, the field is calculated for the points located at the distance 0.2 nm from the tops.

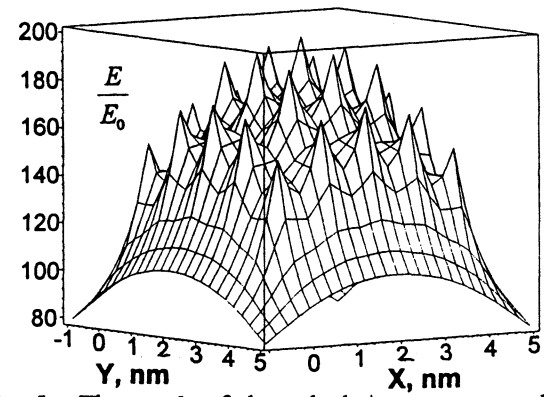


Fig. 5. The result of the calculation of the resulting electric field for the group of 5 x 5 nanotubes .

The fields generated with bundles are more smooth than those generated with one nanotube. So, to obtain high values of the fields one has to place the dielectric layer in which the acoustic waves is generated close to the tops of the tubes. The theory and the experimental study of electroacoustic excitation by electrostatic and electrostriction forces was developed in [4]. Several assumptions were made to modify the theory for electrostrictive excitation by electric fields generated by carbon nanotube films . The spot of the intense of the electric field inside dielectric (the source of deformations) was considered as a place for a piston (or spherical) transducer of radius a much smaller than a wavelength. The resulting acoustic field in Z direction generated by the lattice composed of such transducers was obtained as a result of summurizing. The number of elementary transducers per unit length along X and Y were n_x and n_y. For a wide aperture transducer (D) the parameter $D/\lambda >> 1$, the resulting wave was a plane

wave with the additional phase shift $\pi/2$. The effective electromechanical coupling modulus was found to be $\tilde{\beta} = 2\pi a^2 n_x n_y \gamma^2 E_0$. Here γ is a factor of the increase in the electric field on the boundary of solid due to a small diameter of the tube. E_0 is the value of bias electric field in a plane capasitor. One can use $\tilde{\beta}$ in calculations of transducers and resonators instead of piezomodulus β_{ikl} using expressions derived for piezoelectric devices. As an example in Fig.6 we present the results of the calculation of modulus and the phase of the 3rd overtone electrical impedance Z_e of the YAG plate (thickness d=1290 μm along [100]) excited by the electrostrictive transducer made of the nanotube carbon film providing that $n_x = n_y = 10^6$, γ =33, a= 5 nm, D = 7 mm. Bias voltage V= 30 V was applied. Sharp resonance and antiresonance peaks are seen. $Q = 1.8\ 10^7$ for f_s = 9987500 Hz.

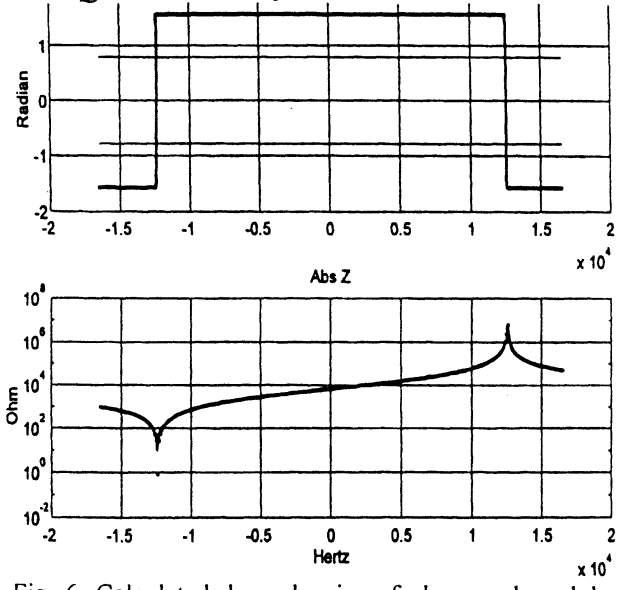


Fig. 6. Calculated dependencies of phase and modulus of electric input impedance of YAG plate electrostrictively excited by electrodes made of carbon nanotube film. Central frequency on the graphs is 10 MHz

Because of the dependence of electromechanical coupling coefficient K^2 on the electric field the resonance frequency is tuned as it is shown in Fig.7.

5. CONCLUSIONS

Using the BAW spectroscopy method material density and the elastic constants for carbon nanotube films were directly evaluated. It was also found that the intrinsic acoustic losses in nanotubes are negligibly small. The effect of the increase in the electric field near the electrode with nanotube thin films can be used for effective acoustic wave excitation in non-piezoelectric plates.

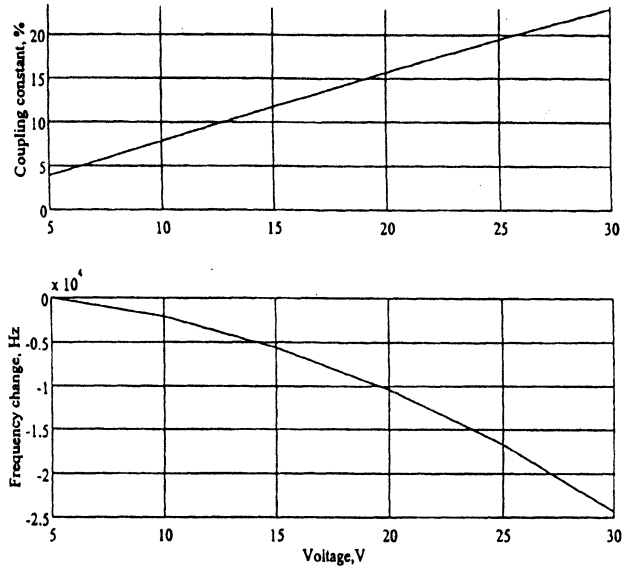


Fig. 7. The dependencies of electromechanical coupling coefficient K and the change of resonance frequency on the electric field.

This work was supported by RFBR GRANT 99-02-17602 and in part by Ministry of Science and Technologies of RF (Fullerenes and atomic clusters N99034)

REFERENCES

1. V.E.Lyamov, Polarization effects and anisotropy in acoustic wave interactions in crystals (Russian), Moscow University Publishers, 1983.
2. Yu.V.Gulyaev, On electron-phonon interaction proportional to the external electric field, Fizika Tv. Tela (sov.), 1967, v.9, N.6, pp.1816-1818.
3. Yu.V.Gulyaev, S.S.Karinskii, V.D.Mondikov, The study of external electric field influence on SAW velocity in lithium-niobate monocrystalls, Pisma Zhurn Tehn. Fiz., v.1, N17, pp.791-794, 1975.
4. K.O.Boltar', I.M.Kotelyanskii, G.D.Mansfeld, Investigation of Dielectric Electroacoustic Transducer, Akusticheski Zhurnal (Sov), 1977,v. 23, N4, pp. 544-549.
5. A.V.Eletskii, Carbon Nanotubes, Physics-Uspekhi 40 (9) 899-924 (1997).
6. D.Sanchez-Portal, E.Artacho, J.M.Soler, A.Rubio, P.Ordejon, Ab-initio Structural, Elastic and Vibrational Properties of Carbon Nanotubes, Phys.Rev., B, v.59, p.12678, 1999.
7. G.D.Mansfeld, BAW composite resonator spectroscopy, 1994 Ultrasonics Symp., proc. v.2, p.655-658.
8. Z.Ya.Kosakovskaya, L.A.Chernozatonskii, E.A. Egorov, Nanofelament Carbon Structure, JEPT Lett (English translation from Pias'ma v Zh.Eksp.Teor.Fiz.), 1992, v.56, N1, pp.26-30.

THEORETICAL ANALYSIS OF MULTILAYER LOVE WAVES DEVICES : EVALUATION OF MASS LOADING EFFECTS FOR (BIO)CHEMICAL SENSING

C. ZIMMERMANN, D. REBIERE, C. DEJOUS AND J. PISTRE
Laboratoire IXL-UMR 5818 CNRS, ENSERB - Université Bordeaux 1
351 cours de la Libération F-33405 TALENCE cedex

Corresponding author : Céline ZIMMERMANN
Laboratoire IXL-UMR 5818 CNRS, ENSERB - Université Bordeaux 1
351 cours de la Libération F-33405 TALENCE cedex
Fax : +33 556 371 545, Tel : +33 556 846 540, e-mail : zimmerma@ixl.u-bordeaux.fr

ABSTRACT

Love-wave devices for detection of chemical species were studied, and a simulation model based on analytical resolution was developed in order to determine the mass sensitivity for several structures. We focused on the simulation of devices equipped with polymeric sensitive layers, as such specific coatings showed their ability for gas/liquid detection. Organophosphorus detection was performed in order to improve theoretical results.

1. INTRODUCTION

Love wave sensors have been receiving a great deal of attention for a few years [1-4], as they can be more sensitive than SAW (Surface Acoustic Wave) devices, and as they are able to work in both gas and liquid media. But, up to now, very few papers concerned the use of Love-wave devices for the detection of chemical species [3,4].

In this paper, we present a model of Love-wave devices which allows to predict mass sensitivity, by calculation of the wave propagation. This model is based on an analytical resolution of motion system and allows to take into account the substrate anisotropy. We focused on the simulation of devices equipped with polymeric sensitive layers as they have demonstrated their ability for such applications [5,6]. We study the influence of sensitive layer parameters on the sensor mass sensitivity, because this knowledge is of primary importance to optimise a sensor and to analyse the test results.

An organophosphorus detection is presented to demonstrate the ability of Love wave devices for gas detection.

2. THEORY

Love mode-based chemical sensors, consist of an acoustic delay line coated with a chemical sorbing layer. The physical properties of this sensitive material (mass, stiffness, conductivity ...) change due to sorption of the chemical compound to be measured. Because of the interaction between the overlay and the acoustic waves, these changes induce wave velocity variations. In many applications, the most significant and best controlled effect is the wave velocity perturbation due to mass loading. Our aim is to develop a model of the Love wave structure to predict its mass sensitivity.

Love wave devices consist in layered structures composed with an anisotropic piezoelectric substrate (e.g. quartz) and an isotropic guiding layer. For gas/liquid detection a third layer called sensitive coating is added (fig 1). The substrate piezoelectricity allows to generate elastic waves. These ones are then coupled into the guiding layer. Because of the waveguiding effect, these waves can be very sensitive to surface perturbations, and a high sensitivity to surface mass loading can be obtained.

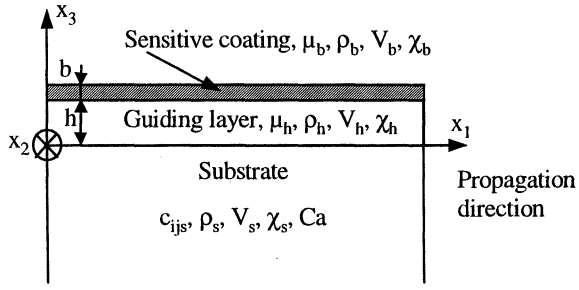


Figure 1 : Love multilayer structure

On the figure 1, μ, c_{ij}, ρ and V are respectively the shear modulus (one of the two Lamé constants), stiffness constants, density and shear bulk wave velocity of each material, with a suffix "s", "h" or "b" referring respectively to the substrate, the guiding layer, and the sensitive layer. $\chi_{s,h,b}$ are variables defined as follow to simplify equations :

- substrate : $\chi_s = Ca\sqrt{1-V^2/V_s^2}$ with anisotropic factor : $Ca = \sqrt{c_{66s}/c_{44s} - c_{46s}^2/c_{44s}^2}$
- guiding layer : $\chi_h = \sqrt{V^2/V_h^2 - 1}$,
- mass loading layer : $\chi_b = \sqrt{V^2/V_b^2 - 1}$

To describe Love-wave propagation, it is necessary to solve motion system in the case of shear horizontal waves. This type of resolution has been developed for

an isotropic Love-wave structure (isotropic layer, isotropic substrate) in ref 7. We propose to resume this resolution method for the Love mode structure in the case of an anisotropic substrate and for the three layered structure (substrate, guiding layer, sensitive layer). It should be noticed that piezoelectricity is neglected : this can be done, as the quartz has a very low electromechanical coupling coefficient. Taking into account mechanical boundary conditions at each interface of the structure (displacement and stress continuity), the resolution of motion equations in each layer leads to the following dispersion relations :

- simple Love structure (substrate and guiding layer) :

$$h = \frac{1}{k\chi_h}\left[\arctan\left(\frac{c_{44s}\chi_s}{\mu_h\chi_h}\right)\right] + \frac{n\pi}{k\chi_h} \qquad (1)$$

- three layered structure (with mass loading layer of thickness b):

$$h = \frac{1}{k\chi_h}\left[\arctan\left(\frac{c_{44s}\chi_s}{\mu_h\chi_h}\right) - \arctan\left(\frac{\mu_b\chi_b}{\mu_h\chi_h}\tan(k\chi_b b)\right)\right] + \frac{n\pi}{k\chi_h} \qquad (2)$$

where n is an integer which represents mode order, and k is the wavelength number.

It can be noted that physical and propagation parameters clearly appear in obtained analytical relations, so their influence could be determined. As the velocity appears in definition of variables χ, equations (1) and (2) allow to calculate the Love wave velocity for each mode (the mode is selected by n value) versus h (guiding layer thickness) and b (mass loading layer thickness). This propagation model was implemented with MapleV software, and permits to predict mass sensitivity for different structures.

3. INFLUENCE OF PHYSICAL PARAMETERS

Our aim is to simulate Love wave structures in order to determine an optimised design and to predict the sensor response in the case of gas detection. In the following the considered structures are composed with AT-cut quartz substrate (Euler angles (0°,122°,90°), anisotropic, piezoelectricity neglected), SiO_2 guiding layer (isotropic) and polymeric sensitive layer (isotropic). The mass loading layer will be thin in order to neglect viscoelastic losses. As we say in paragraph 2, the most important effect in sensor response is the mass loading effect. So we characterise the Love-wave sensor response with the velocity mass loading sensitivity defined as follow :

$$Smv = \frac{1}{\Delta\rho_b b}\frac{\Delta V}{V_0} \qquad (3)$$

where :

$\Delta\rho_b$ is the polymer density increase modelling the mass loading effect during the gas/polymer interaction.

b is the sensitive layer thickness.

$\Delta V = V_1 - V_0$ is the velocity shift due to mass effect and is calculated as follow :

V_0 (without mass loading effect) is obtained with $V_b = \sqrt{\mu_b/\rho_b}$ in χ_b, and V_1 (with mass loading effect) is obtained with $V_b = \sqrt{\mu_b/(\rho_b + \Delta\rho_b)}$ in χ_b.

3.1 Influence of the guiding layer (SiO_2) parameters

As it is now well known for Love-wave devices [1,2], the simulation shows an optimum guiding layer thickness h_{opt} at which the mass sensitivity is a maximum. We demonstrate that for a prescribed polymer (ρ_b and μ_b known), the values of b (up to 100 nm) and of $\Delta\rho_b$ (up to 50% of ρ_b) due to mass loading have negligible influence on h_{opt}. So, for a prescribed polymer, we are able to determine h_{opt} in order to design an optimised device.

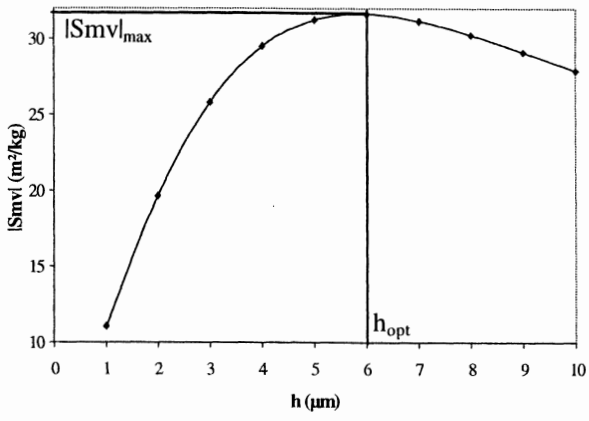


Figure 2 : Velocity mass sensitivity versus guiding layer thickness. Love-wave structure : AT-cut quartz/SiO_2/PMMA (thickness b = 100 nm), wavelenght λ = 40 µm ,distance center to center Lcc = 125 λ, aperture W = 50 λ.

3.2 Influence of the sensitive layer parameters

To predict the sensor response in conditions close to experience, we focus on chemical sensitive coating parameters and their influence on sensor sensitivity. Note that sensitive layer physical parameters clearly appear in equation (2), so their influence is easy to determine.

We present the influence of the sensitive coating density, shear modulus and thickness on the velocity mass sensitivity. It could be noticed that physical parameters of materials are difficult to find in the literature, particularly for polymers which are often developed specifically for one application.

In the following, the wavelength λ is fixed at 40 µm and the guiding layer thickness h is fixed at 6 µm ($\approx h_{opt}$). Unless other specifications, the sensitive layer parameters are those of PMMA [8] : ρ_b = 1100 kg/m^3,

μ_b = 1,1e9 N/m². The sensitive layer thickness is fixed at 20 nm.

3.2.1 Influence of ρ_b

Calculations of the relative phase velocity shift $\Delta V/V_0$ as a function of the polymer density shows a linear behaviour for ρ_b from 1000 to 2800 kg/m³ (typical values for polymeric materials [8]), with $\Delta V/V_0 \approx 100$ ppm at 3000 kg/m³. The |Smv| value is almost constant in this whole range (fig 3), for μ_b higher than 1e7 N/m². Moreover, Smv is in the range of 30 to 32 m²/kg that is higher than the values typically known for SAW sensors (around 20 m²/kg) [9].

3.2.2 Influence of μ_b

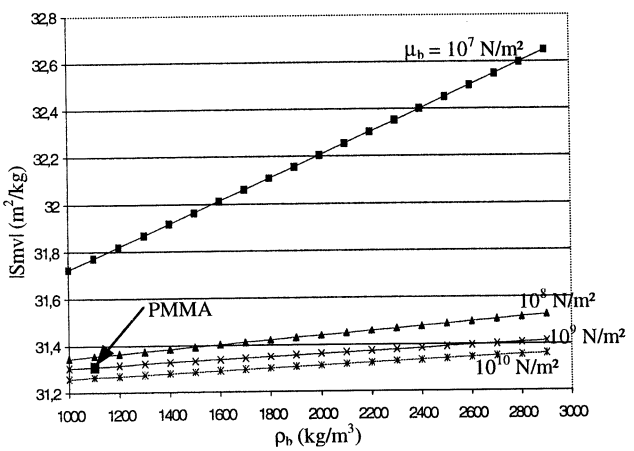


Figure 3 : Velocity mass sensitivity (Smv) versus polymer density (ρ_b) for several shear modulus values (in the range 1e7 to 1e10 N/m²). Love wave device AT-cut quartz [11] / SiO_2 [12]/isotropic sensitive layer.

Figure 3 shows that shear modulus has influence on the relative velocity shift due to mass loading effect. Moreover, polymers shear modulus is strongly dependant on the temperature [10]. However it appears that shear modulus (rarely published for volumic materials, even more difficult to evaluate in case of thin films and at high frequencies) has a low influence if μ_b is higher than a few 1e7N/m² (glassy polymers). If μ_b is smaller than 1e7N/m² (rubbery polymers), the sensitivity increases highly and is also more dependent on the other layer parameters (like ρ_b). The polymer shear modulus increase with frequency [8] is an effect which could limit the influence of μ_b on the mass sensitivity.

3.2.3 Influence of b

As for the density, the relative phase velocity shift $\Delta V/V_0$ versus polymer thickness shows a linear behaviour for b from 5 to 100 nm, with $\Delta V/V_0 \approx 200$ ppm for b at 100 nm. So, the |Smv| value is almost constant in this whole range.

4. ORGANOPHOSHPORUS DETECTION

In order to validate theoretical results presented above, organophosphorus detection with Love-wave devices have been undertaken.

Figure 5 shows an example of DMMP (dimethylmethylphosphonate) detection with a 92 MHz Love-wave devices coated with a polysiloxane polymer (estimated thickness 30 μm). The Love device responses to three different DMMP concentrations (1,2 to 2,4 mg/m³)are represented. We can note that the sensor response is closely related to the presence and the concentration of the organophosphorus polluant.

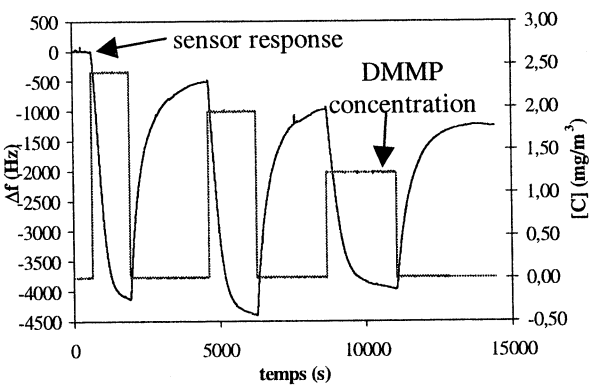


Figure 4 : Detection of DMMP with a Love-wave sensor consisting in AT-cut quartz/SiO_2/polysiloxane polymer. Under N_2 (0,1 l/min) at ambiant temperature.

This first detection dring to light a high level of sensitivity, and more systematic tests are undertaken in order to validate theoretical results.

5. CONCLUSION

A theoretical model of Love-wave propagation was developed and used to predict mass loading effects in case of gas detection. Simulations focus on Love structures with polymeric sensitive coatings. The model allows to find the optimum guiding layer thickness well known for Love structures.

We also studied the influence of polymer physical parameters on phase velocity. It appears that Smv do not shows a strong dependance on polymer density, shear modulus and thickness, all the more if the polymer is not much rubbery. Although, the mass sensitivity increases seriously if the shear modulus is low (< 1e7 N/m²), but energy losses due to polymer viscoelasticity could limit this benefit.

Finally, Love wave devices are more sensitive than SAW devices and demonstrate their potentialities to be applied to chemical detection. To validate these theoretical approach, tests under organophosphorus

compounds have been undertaken. Detection of DMMP shows the ability of Love-wave sensors to detect chemical compounds at sub ppm level.

6. ACKNOWLEDGEMENTS

The authors gratefully acknowledge the helpful discussions and suggestions of J.F. Lipskier and E. Chastaing (Laboratoire Central de Recherches of Thomson C.S.F., domaine de Corbeville France).

REFERENCES

[1] B. Jakoby, M. Vellekoop, "Properties of Love waves: applications in sensors", Smart Mater. Struct., vol. 6, pp.668-679, 1997.

[2] J. Du et al," A study of Love-wave acoustic sensors", Sensors. & Actuators. A, vol. 56, pp.211-219, 1996.

[3] E Gizeli, C.R. Lowe, M. Liley, H. Vogel, "Detection of supported lipid layers with the acoustic Love waveguide device: application to biosensors", Sensors and Actuators B, vol. 34, pp.295-300, 1996.

[4] E Gizeli, "Design considerations for the acoustic waveguide biosensor", Smart Mater. Struct., vol. 6, pp.700-706, 1997.

[5] R.A. McGill, M.H. Abraham, J.W. Grate, "Choosing polymer coatings for chemical sensors", Chemtech, vol. 24 [9], pp.27-37, 1994.

[6] D. Rebière, C. Déjous, J. Pistré, J.F. Lipskier, R. Planade, "Synthesis and evaluation of fluoropolyol isomers as SAW microsensor coatings : role of humidity and temperature", Sensors and Actuators B, vol. 49/1-2, pp.139-145, 1998.

[7] Royer D., and Dieulesaint E., Ondes élastiques dans les solides Tome1, Masson, Paris, 1996.

[8] D.W. Van Krevelen, Properties of Polymers, Third edition, Elsevier, 1990

[9] J.W. Grate, S.J. Martin,R.M. White, "Acoustic wave microsensors", Analytical Chemistry, vol. 65, pp 941-948, 1993

[10] R.W. Murray et al, Chemical Sensors and Microinstrumentation, ACS Symp. Series 403, DC1989, ch. 15, pp 222-236.

[11] Slobodnik A.J., Microwave Acoustics Handbook, Mason, Paris, 1973.

[12] G. Kovacs, G,W Lubking, M.J. Vellekoop, A. Venema, IEEE Ultr. Symp., 1992, pp.280-285.

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

DEVELOPMENT OF A QUARTZ PRESSURE SENSOR FOR EXTREME HOSTILE ENVIRONMENT APPLICATIONS

N. MATSUMOTO, Y. OOHASHI, M. MIYASHITA, G. FUJISAWA, B.K. 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

A new dual-mode quartz pressure sensor HCQG (Hostile environment Crystal Quartz Gauge) is based on the principle of CQG† (Crystal Quartz Gauge) that has been used in the oil and gas industry for the past 10 years. The CQG design is limited to a maximum pressure of 15,000 psi (103 MPa). The maximum pressure limitation is caused by twinning of crystalline quartz that is a precursor to mechanical fracture. The new design for high pressure and temperature applications was obtained by carrying out three-dimensional, finite-element stress analyses of many probe structures. We analyzed probe structure elements with large stress concentrations that are prone to twinning when subjected to high pressures and temperatures. Design modifications to the probe structure enabled us to reduce the stress magnification in those critical areas. Prototypes made with the new design have been successfully fabricated and tested up to 25,500 psi (175 MPa) and 180 °C. This paper will present the development details, resonator characteristics, and performance of HCQG over the entire pressure and temperature ranges.

† Mark of Schlumberger

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 in 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.

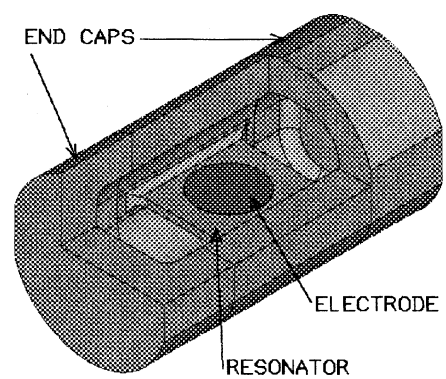


Figure 1: Schematic diagram of the CQG probe

II. DESIGN CHANGE AND STRESS ANALYSIS

Previous three-dimensional stress analyses indicated large stress concentrations at the small fillet of the sensing element. These large stresses are potential causes of failure in a high-pressure environment primarily because of material twinning. Therefore, it is conceivable that reducing these large stresses may enable its use in a high-pressure environment. Figure 2 shows the sensing element of the existing CQG and HCQG with different end shapes. With a larger radius of fillet in the new HCQG design, we have found significant reductions in stress concentrations that enabled a successful design of HCQG for high-pressure applications.

A three dimensional finite element method (FEM) is used to model the stress distributions in the entire HCQG probe structure under a hydrostatic pressure of 103 MPa (15 Kpsi). Ten-node tetrahedral anisotropic elements are employed to simulate stresses in the quartz material, and the number of elements in this modeling is about 20,000. Figure 3 shows the calculated stress distributions in the reonator body. The third principal stress that represents the maximum compressive stress in the quartz material is evaluated. As can be seen from the result shown in Figure 3, the minimum stress (maximum compressive stress) on the sensing element is approximately -280 MPa for an external hydrostatic pressure of 103 MPa (15 Kpsi), and it would be -470 MPa for an external hydrostatic pressure of 172 MPa (25 Kpsi).

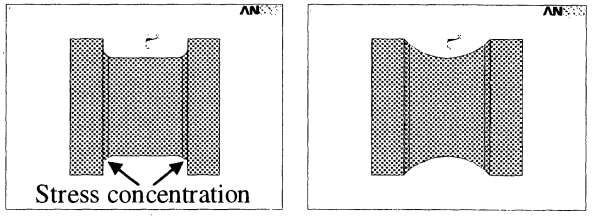


Figure 2: Sensing element of CQG with small radius of fillet (left) and the new design with a larger one for HCQG.

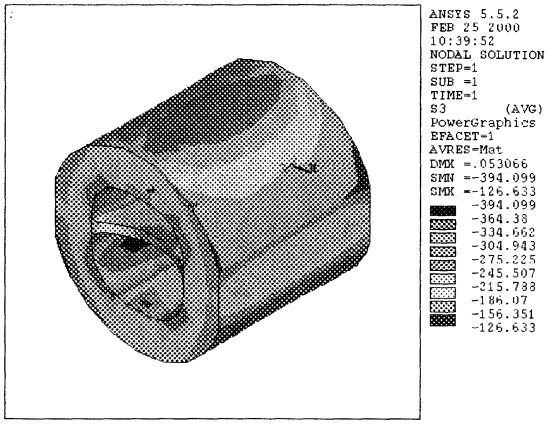


Figure 3: Stress distribution in the resonator body of HCQG under a hydrostatic pressure of 103 MPa (15Kpsi).

III. STATIC CALIBRATION RESULTS

The HCQGs with a modified probe structure were calibrated over a pressure range of atmospheric pressure to 172 MPa (25 Kpsi) and a temperature range of 35°C to 175°C. 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.

$$P_S = \sum_{i,j=0}^{n} A_{ij} f^{i}{}_P f^{j}{}_T \tag{1}$$

$$T_S = \sum_{i,j=0}^{n} B_{ij} f^{i}{}_P f^{j}{}_T \tag{2}$$

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 36 nonzero elements. The largest exponent n of f_P and f_T is chosen to be 5 for both P_S and T_S.

In Figure 4 we show the pressure sensitivities of HCQG at two temperatures. The pressure sensitivity of the C-mode of the new HCQG is only 10% smaller than that of the CQG design over the entire pressure and temperature ranges, and the sensitivity varies almost linearly up to the maximum pressure of 172 MPa (25 Kpsi).

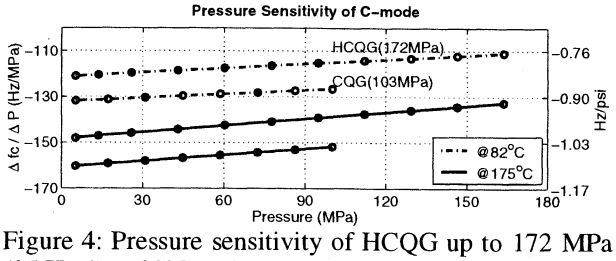


Figure 4: Pressure sensitivity of HCQG up to 172 MPa (25 Kpsi) at 82°C and 175 °C.

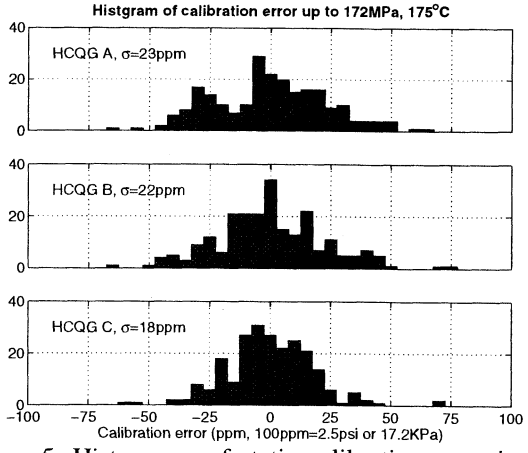


Figure 5: Histograms of static calibration error in 231 pressure and temperature points.

Figure 5 shows histograms of the calibration errors from static calibration of three HCQGs. It can be seen that only a few errors exceeded the calibration error of ±50 ppm (±8.6 KPa or 1.25 psi) and their standard deviations are small to within the order of 20 ppm (3.4 KPa or 0.5 psi) among the 231 points of pressure and temperature combinations. Our results suggest that the HCQG probe can be calibrated with extremely small errors that are equivalent to those of the existing CQGs. The hysteresis values on the pressure readings for the three gauges were small to within 90 ppm (15.5 KPa or 2.25 psi).

IV. PERFORMANCE VERIFICATION

Several static and dynamic tests were performed to verify the performances of the HCQG. In this section we present the results including the resonance data at higher temperatures.

Pressure transient analyses are routinely used for estimating reservoir permeability. 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 6 shows the response of the HCQG and CQG to a pressure step of 34 MPa (5 Kpsi) to atmospheric pressure. Both HCQG and CQG pressure readings after dynamic compensation stabilize at around 10 seconds, while a conventional single mode quartz gauge takes a much longer time. The dynamic response of HCQG to the pressure step is confirmed to be equivalent to those of the existing CQGs.

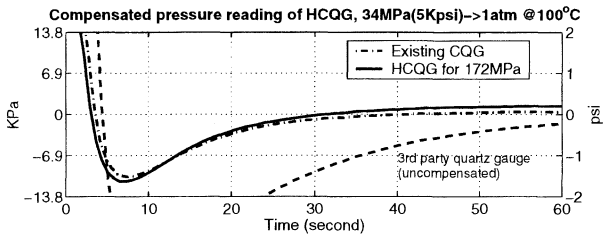


Figure 6: Fast response of HCQG and CQG to a pressure step

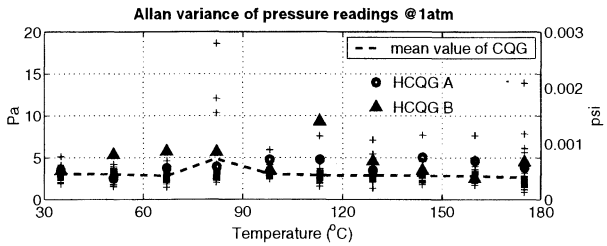


Figure 7: Allan variance of HCQG at various temperatures with the counter gate time of 1 second.

Figure 7 shows the Allan variance [9] of pressure readings of two HCQGs and more than 60 CQGs (+markers). This test was performed at atmospheric pressure and at various temperatures. The counter gate time was set to 1 second. A rather small variance indicates the mean pressure stability of HCQG to be better than 0.003 psi.

At the 13th EFTF, we presented the long-term stability of the CQG probe at 103 MPa, 150°C or 175°C for qualifying a pressure sensor for permanent monitoring systems [10]. The second long-term stability test of the entire CQG system including the pressure gauge, clock for digitization at down hole and their oscillator circuits was carried out in the same pressure and temperature environment. This test has been done with the existing CQG. We expect similar results for the HCQG as well, because there are few differences that might affect the long-term stability characteristics of these two gauges. The pressure reading changes in Figure 8 includes drift of the pressure gauge, clock and the effect of degradation in the electronics. The two CQG A and B in 150°C show excellent stability to well within ±3.4 KPa (0.5 psi) for over 6000 hours (250 days), and another two CQG C and D at 175°C is still within ±6.9 KPa (1 psi). These gauges have been continuously vibrated with oscillators driven at 1 micro-watt.

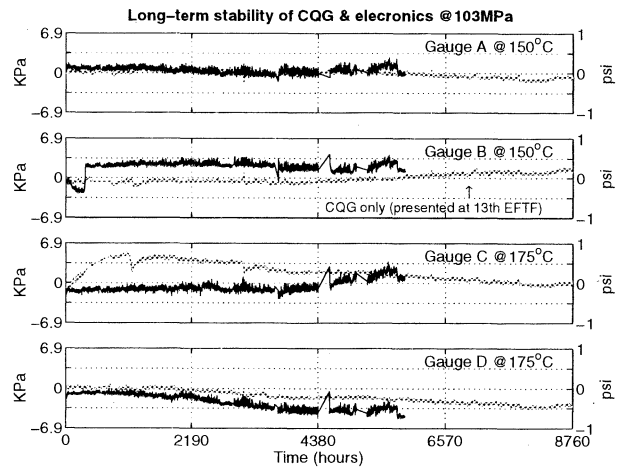


Figure 8: Long-term stability of existing CQG with its electronics at two high temperatures. Gray lines indicate the long-term stability of the CQG probe only (presented at the 13th EFTF in Besançon)

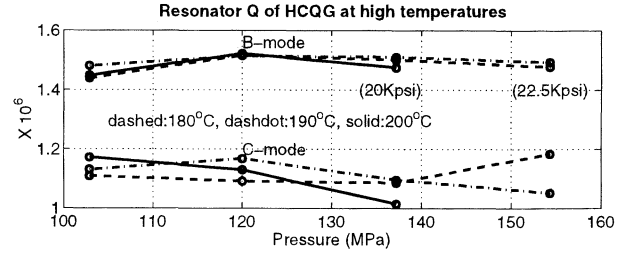


Figure 9: Resonator Q value of HCQG in a high pressure, high temperature environment.

Hostile environment wells characterized by high pressures and high temperatures exist at many places. To investigate the feasibility of using the HCQG in such hostile environments, we measured the resonator characteristics of the HCQG probe with a network analyzer and a π network. Figure 9 shows the resonator Q values of the HCQG up to 155 MPa (22.5 Kpsi), 190°C or 138 MPa (20 Kpsi), 200°C. Among all the test points, the Q values are essentially constant and the C-mode values are at least more than one million (10^6). These results indicate that the HCQG can be successfully used in such hostile environments.

CONCLUSIONS

This paper has described an HCQG development history and the results of its evaluation tests. First of all, very small calibration errors of less than ±50 ppm in full scale suggest that we can claim total accuracy of HCQG at less than ±200 ppm in full scale (34 KPa or 5 psi) including errors from the clock and pressure system uncertainty. This is the best performance among the pressure sensors with the operation range of up to 172 MPa (25 Kpsi), 175°C.

Other test results also exhibit the superiority of HCQG to other quartz pressure sensor technologies. Pressure resolution and dynamic response are found to be equivalent to the existing CQG, and the long-term stability of CQG at 103 MPa (15 Kpsi) with its electronics at high temperatures are excellent to within ±3.4 KPa (0.5 psi) at 150°C and ±6.9 KPa (1 psi) at 175°C for over 6000 hours. In addition, a good resonator Q values up to 155 MPa (22.5 Kpsi), 190°C or 138 MPa (20 Kpsi), 200°C suggests that it will operate in hostile environments.

These results confirm that the HCQG characteristics exceed the demanding specifications for both the well-tests and permanent monitoring systems.

ACKNOWLEDGMENTS

We wish to thank John Ullo and Tarek M. Habashy of Schlumberger-Doll Research for helpful discussions and support of this project.

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.

[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.

[10] N. Matsumoto, Y. Sudo, B.K. Sinha and M. Niwa, "Long-term stability and performance characteristics of crystal quartz gauge at high pressures and temperatures" *Proc. Joint meeting of 13th EFTF and IEEE International Freq. Contr. Symp.*, 1019-1022 (1999).

Schlumberger K.K.
Phone: 81-42-759-2111, Fax: 81-42-759-0905
nmatumot@fuchinobe.skk.slb.com
oohashi@fuchinobe.skk.slb.com
miyasita@fuchinobe.skk.slb.com
fujisawa@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

REALIZATION AND CHARACTERIZATION OF FIBER-PUMPED ROOM TEMPERATURE Tm:YAG AND Tm-Ho:YAG LASERS AT 2 μm WAVELENGTH

C. Svelto[1, 2, 3], G. Galzerano[1, 2, 3], A. Maffiolini[1], and E. Bava[1, 2, 3]

Dipartimento di Elettronica e Informazione – Politecnico di Milano[1], INFM[2], and CNR - CSTS[3]
Piazza Leonardo da Vinci 32, 20133 Milano, Italy
E-mail: Cesare.Svelto@PoliMI.IT Phone: +39 02 2399 3610 Fax: +39 02 2399 3413

ABSTRACT

Different diode-pumped Tm:YAG and Tm-Ho:YAG lasers operating near the room temperature with continuos-wave emission at 2 μm have been developed. A throughout characterization in both multi- and single-longitudinal mode regimes has been performed by experimentally determining the corresponding threshold and slope efficiency values. Continuous-wave output powers of more than 70 mW and 35 mW have been obtained in single longitudinal mode operation near the room temperature for Tm:YAG and Tm-Ho:YAG lasers, respectively.

1. INTRODUCTION

High-resolution spectroscopy and frequency metrology in the 2 μm region of the electromagnetic spectrum are assuming an important scientific interest in order to realize high-accuracy optical frequency standards and frequency connections between standards at different wavelength [1]. Lasers with emission at a wavelength of 2 μm are also useful sources for applications that require eye-safe propagation through the atmosphere. Such applications include coherent Doppler-LIDAR, DIAL measurements of CO_2 and H_2O, and Doppler velocimetry (wind shear detection, global wind speed measurement from satellite) [2, 3]. Moreover these sources may be useful in medical applications because the water molecule presents strong absorption in this wavelength region [4]. All these applications require significant power, single frequency operation, and wide frequency tuning range.

In this work, single-frequency diode-pumped Tm:YAG and Tm-Ho:YAG lasers with continuos wave emission around 2 μm have been developed at room temperature. The achieved results on the characterizations in multi- and single- longitudinal mode regime are hereby reported.

2. FIBER-PUMPED Tm:YAG AND Tm-Ho:YAG LASER

The Tm:YAG and Tm-Ho:YAG crystals used in our experiments, supplied by Scientific Material, have a nominal dopant concentration of 9.3% with Tm^{3+} ions and 5.7% with Tm^{3+} and 0.36% with Ho^{3+}, respectively. A schematic diagram of the relevant energy levels of both the Tm^{3+} and Ho^{3+} ions in a YAG host is reported in Fig. 1. The upper and lower laser levels are $^3F_4$ and $^3H_6$ for the Tm:YAG system whereas in the Tm-Ho:YAG laser the levels are $^5I_7$ a $^5I_8$ in the Ho^{3+} ions. For each absorbed pumping photon from the $^3H_6 \rightarrow {}^3H_4$ levels, two Tm^{3+} ions are raised in the upper laser level by means of a cross-relaxation process between the Tm^{3+} levels $^3H_4 \rightarrow {}^3F_4$ and $^3F_6 \rightarrow {}^3H_4$. The population inversion in the Ho system is obtained by means of an energy transfer process between $^3F_4 \rightarrow {}^3H_6$ and $^5I_8 \rightarrow {}^5I_7$.

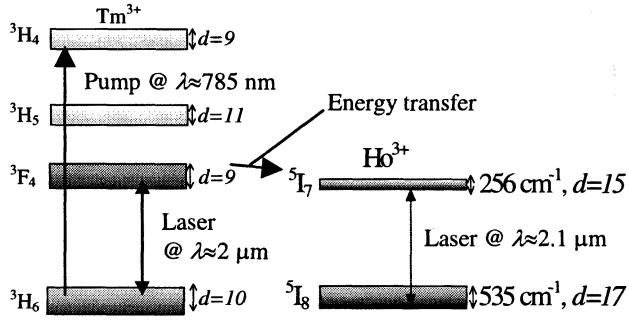


Figure 1. Energy levels of Tm^{3+} and Ho^{3+} ions in YAG.

To build the laser cavity, we used an end-pumped hemispherical resonator as shown in Fig. 2. The end facets of the active material are flat polished and covered with dicroic multidielectric coatings. The first facet of the laser crystal is coated for high reflectivity (HR>99.99%) at 2 μm laser wavelength and for high transmission (HT>96%) at the pump wavelength 781 nm. The second facet is antireflection coated at the laser wavelength (AR<0.2%) and for a high reflectivity at pump wavelength (R>95%). Approximately 69% (42%) of the incident pump power at 781 nm is absorbed during a single pass through the 3 mm long Tm:YAG (Tm-Ho:YAG) laser crystal. Using the double-pass pumping configuration, more than 90% (65%) of the available pump power can be usefully absorbed within the Tm:YAG (Tm-Ho:YAG) active medium. In the longitudinal pumping scheme the gain element cannot be arbitrarily long, thus achieving maximal pump absorption, because in these quasi-three level lasers a significant population inversion always needs to be maintained in order to avoid reabsorption losses.

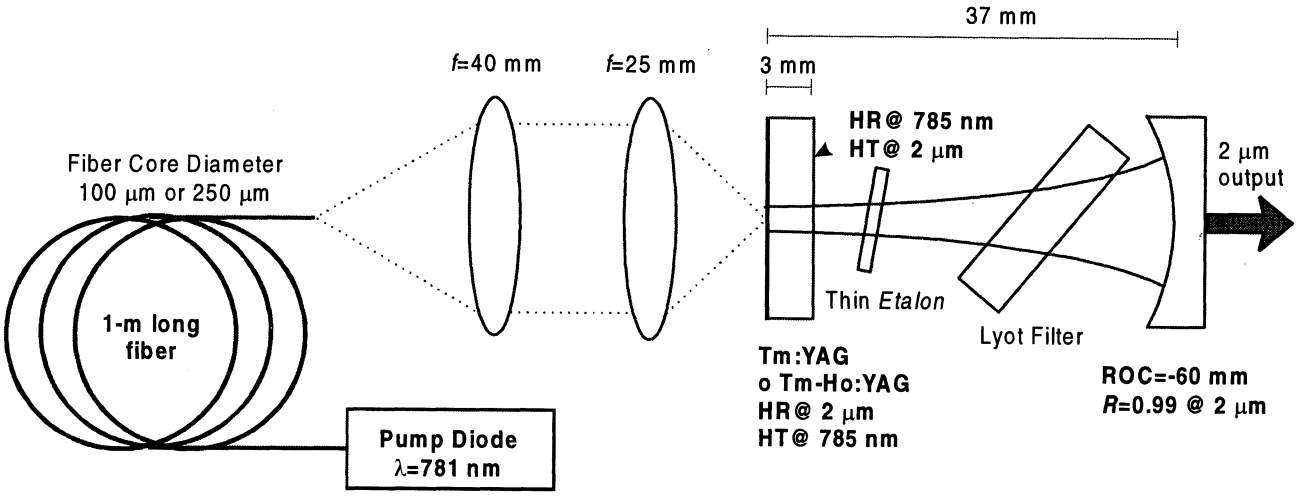


Figure 2. Schematic diagram of the Tm:YAG and Tm-Ho:YAG laser cavity.

Therefore, when choosing the optimal length of the gain medium, there is a trade-off between the pumping efficiency and the optimal profile for the population inversion. The HR coating at 781 nm on the second facet of the active medium, in addition to providing for higher pump absorption, also guarantees a much more uniform inversion, and thus gain, profile in the active medium. A spherical output coupling mirror (T=1%), with radius of curvature of –60 mm, is placed at ~37 mm geometrical distance from the first (pump entrance) facet of the crystal.

The laser was originally end pumped by a fiber-coupled AlGaAs laser diode (OPC-D010-785-HB/250) emitting up to 5 W around 781 nm. The 1 m long optical fiber carrying the pump radiation has a diameter of 250 μm with a numerical aperture of 0.22. The pump beam was focused into the laser crystal, with a spot-size of ~190 μm, by means of two spherical lenses with focal lengths of 40 mm and 25 mm, respectively. Recently, a novel high-brightness fiber-coupled pump diode has been used instead of the previous device in the same end-pumping configuration but using an output coupler with radius of curvature of -40 mm. The new pump source has a 100 μm fiber core (diameter) with a numerical aperture of 0.22 and a maximum output power of 3 W at 781 nm. In this case, an optical imaging system with two spherical lenses of focal lengths f_1=20 mm and f_2=30 mm, respectively, has been adopted. With this more intense pump beam the population inversion within the active medium reaches a much higher value than in the previous case. This turns out in a lower pump threshold and higher laser efficiency. The pump wavelength is tuned to the Tm^{3+} absorption peak, near 781 nm [5], using a thermoelectric cooler (TEC) to control the temperature of the laser diode at ~20.5 °C. The YAG crystal is held in a copper structure and then mounted on a Peltier cooler used to remove pump-generated heat and to actively control the crystal temperature.

3. MEASUREMENTS AND CHARACTERIZATION

The optimized laser configuration (higher brightness pump diode) allowed to achieve rather low threshold values for both the active media: $P_{th,Tm}$=352 mW for the Tm:YAG laser and $P_{th,Tm\text{-}Ho}$≅285 mW for the Tm-Ho:YAG laser. The corresponding slope efficiencies turned out to be η_{Tm}=23% for the Tm:YAG laser and $\eta_{Tm\text{-}Ho}$=19.2% for the Tm–Ho:YAG laser. Figure 3 shows, as an example, the typical Tm-Ho:YAG laser output in multimode regime versus the incident pump power (for the two pump system configurations). The high threshold value achieved with the initial pumping scheme (diamond data point in the diagram) is a direct consequence of the quite large pump spot size realized in the active medium (125 μm) compared to the laser spot dimension (140 μm) [6]. The improvement in terms of performance can therefore observed when a better mode matching is realized using the higher brightness pump diode.

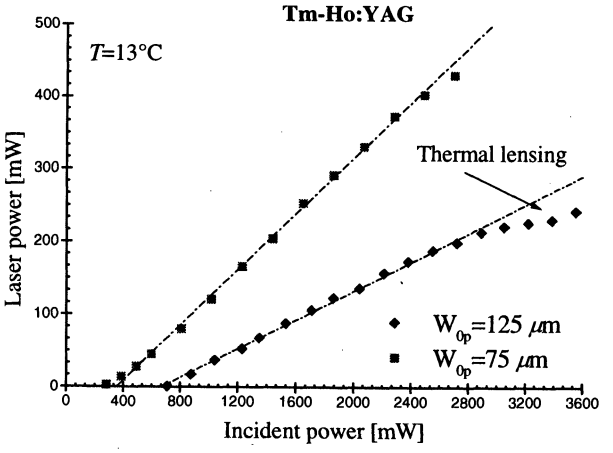


Figure 3. Output power of Tm-Ho:YAG laser as a function of the incident pump power, in single transverse TEM_{00} mode.

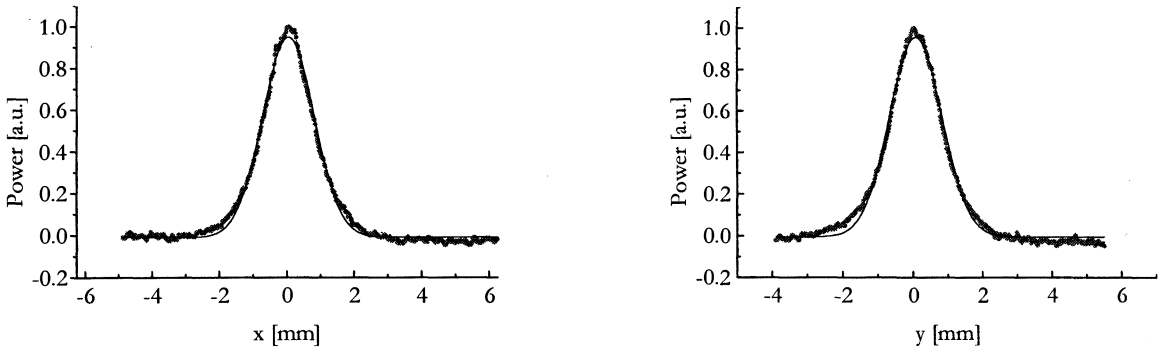


Figure 4. Transverse profile of the Tm:YAG and Tm-Ho:YAG laser beam.

We measured the transverse dimension of both the pump beam and laser mode by matching a 10-times magnified image of the output face of the crystal on a Vidicon camera (λ_{max}=2.2 μm) connected to a beam profiler system. The laser shows a TEM_{00} beam spot, corresponding to a single spatial mode (see Figs. 4). By taking both near-field and far-field measurements, the M^2 value for the laser beam was calculated, turning out into similar values in the two propagation planes, $M^2_x \cong M^2_y$, and always being less than 1.2 (diffraction limited laser beam [7]) for all the operating conditions.

The laser performance have been characterized in terms of the crystal temperature and cavity length (Figs. 5 and 6). For both the laser systems, slope efficiency and threshold pump power have been measured as a function of the operating crystal temperature. Due to the intrinsic behavior of a quasi-three level laser system, the threshold pump power is reduced and the slope efficiency is increased, as the temperature is decreased. Typical measurement results are shown in Fig. 5 for the Tm-Ho:YAG laser (the Tm:YAG system behaves in a similar way but with lower sensitivity to temperature changes).

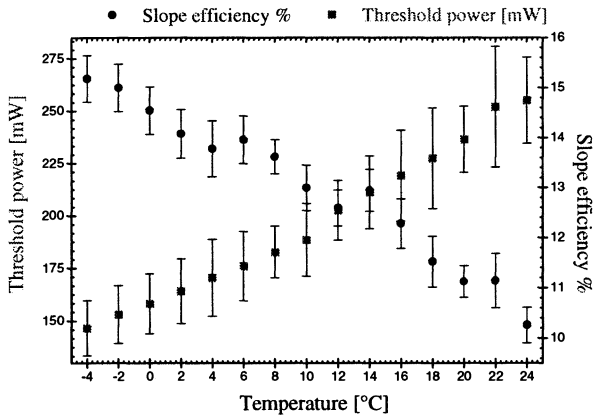


Figure 5. Slope efficiency and threshold pump power as a function of the Tm-Ho:YAG crystal temperature.

The output power versus cavity length characterization, reported in Fig. 6 in the case of the Tm:YAG laser at an operating temperature of 14 °C, shows an optimum cavity length value near the optical instability point (hemispherical cavity configuration) where the pump to laser mode matching is maximized.

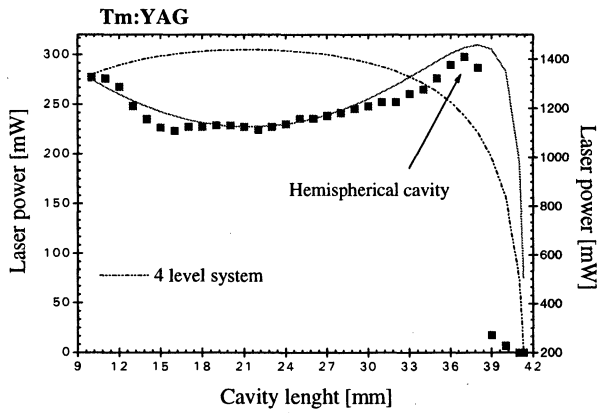


Figure 6. Output power of Tm:YAG laser as a function of the cavity length.

Single frequency operation of the Tm:YAG laser is achieved using an uncoated high index (SF10) etalon with a thickness of 200 μm. Note that a round-trip loss discrimination of less than 1% is enough, in this low-gain active material, to ensure single mode oscillation. A maximum single-frequency output power of 70 mW is achieved with an incident pump power of 2.7 W for the Tm:YAG laser.

Due to the wider gain bandwidth of the Tm-Ho:YAG system, single mode selection is more critical. Figure 7 shows the Tm-Ho:YAG laser spectrum in multimode regime as measured with a scanning Fabry-Perot interferometer with a finesse of 50 and free spectral range of 380 GHz. Seven of the eight observed peaks are different longitudinal mode frequency separated by 28 GHz around 2.097 μm whereas the last peak is near a wavelenght of 2.092 μm. In order to select the single longitudinal mode it was necessary to place within the cavity a 135 μm etalon with 80% reflectivity and a Lyot filter (based on a 5 mm thick quartz plate at Brewster angle) in the same time. In this way, a maximum single-frequency output power of 35 mW is achieved with an incident pump power of 2.7 W. Figure 8 shows the single frequency spectrum of

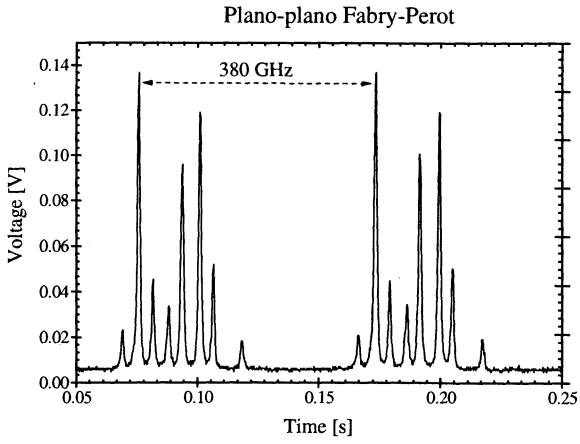


Figure 7. Spectrum of the multimode Tm-Ho:YAG laser.

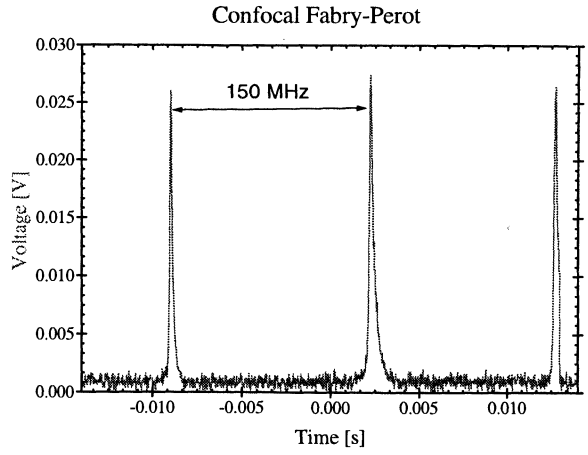


Figure 8. Spectrum of the single mode Tm-Ho:YAG laser.

the Tm-Ho:YAG laser analyzed by means of a confocal Fabry-Perot cavity (Finesse 100) and a photodetector connected to a digital scope.

No special cautions have been taken for long term stability with the current hardware. The laser mirrors are mounted on an aluminum baseplate; in the near future we will replace these mounts with super Invar baseplates.

4. CONCLUSIONS AND DEVELOPMENTS

Characterization of multi- and single-longitudinal mode Tm:YAG and Tm-Ho:YAG lasers with continuos-wave emission around 2 μm has been presented. These types of optical oscillators are interesting sources in high-resolution spectroscopy, metrology, and high-sensitivity optical sensors.

For these applications it is very important to achieve high performance in terms of amplitude and frequency stability. In this sense, and in particular for the Tm-Ho:YAG laser, the next experimental work will be addressed to the active amplitude stabilization (Relative Intensity Noise suppression by means of an optoelectronic control loop) and to the frequency stabilization (against resonant cavities or molecular absorption lines e.g. HBr) of the developed laser system.

REFERENCES

[1] J. E. Bernard, B. G. Whitford, and A. A. Madej, «A Tm:YAG laser for optical measurements: mixing 148 THz with CO_2 laser radiation», *Optics Comm.*, vol. **140**, pp. 45–48 (1997).

[2] S. M. Henderson, C. P. H. Hale, J. R. Magee, M. J. Kavaya and A. V. Huffaker, «Eye-safe coherent radar system at 2.1 μm using Tm,Ho:YAG lasers», *Opt. Lett.* vol. **16** (10), pp. 773–775 (1991).

[3] M. E. Storm, «Coherent 2 μm sources burst into wind shear detection», *Laser Focus World*, pp. 117-122 (April 1991).

[4] J. A. Curcio and C. C. Petty, *J. Opt. Soc. Am.*, vol. **41**, p. 302 (1951).

[5] M. G. Jani, R. J. Reeves, R. C. Powell, G. J. Quarles, L. Esterowitz, «Alexandrite-laser excitation of Tm-Ho:$Y_3Al_5O_{12}$ laser», *J. Opt. Soc. Am. B*, vol. **8**, pp. 741-746 (1991).

[6] D. G. Hall, «Optimum size criterion for low gain lasers», *Appl. Opt.*, vol. **20** (9), pp. 1579-1583 (1981).

[7] T. F. Johnston Jr., «M^2 concept characterizes beam quality», *Laser Focus World*, May 1990, pp. 173-183 (1990).

METROLOGICAL INVESTIGATION OF SF_6 BY A BROADLY TUNABLE DIFFERENCE FREQUENCY SPECTROMETER 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)

Performance of a compact, 9 μm to 11 μm (27 to 33 THz) tunable, high resolution diode-laser-based difference-frequency spectrometer using a type-II cut $AgGaS_2$ nonlinear crystal is reported. This spectrometer is built for high resolution spectroscopy of OsO_4 and SF_6 molecules so as to extend the infrared reference grid of the BNM-LPTF which is based on the stabilization of the CO_2 laser on an OsO_4 saturated absorption line. From 35 mW and 100 mW of the pump powers, respectively at 778 nm and 842 nm, more than 50 nW of the infrared radiation is generated. Although 1 μW power level is necessary for saturated absorption spectroscopy of the OsO_4 molecule, 50 nW are enough if we tighter focus the beam in the Fabry-Perot absorption cell.

1. INTRODUCTION

In frequency metrology, secondary optical frequency standards are based on the stabilization of a laser line to a narrow molecular saturated absorption line. Heavy spherical molecules such as OsO_4, SF_6 or SiF_4 are potential accurate frequency references for an infrared-to-visible frequency measurement chain. Up to now, in the 9-μm to 11-μm (27-33 THz) range, CO_2 secondary frequency standards such as the one in our laboratory (the BNM-LPTF) [1] cannot be continuously tuned. Because CO_2 lasers lack tunability, only some coincidences between laser and molecular lines can be used as frequency references. Up to now, to solve the CO_2 tunability limitations, electro-optic modulators are used. Although the latter allow continuous tunability over more than 10 GHz, the device is heavy, expensive and has low efficiency. Another solution to perform spectroscopy in the 10-μm range is the use of infrared (IR) lead salt diode lasers. But these laser diodes, liquid nitrogen cooled, are very noisy and their tunability is difficult to control. Moreover, the typical linewidth of these IR laser diodes is about 10 MHz. High resolution spectroscopy with such diode lasers is not possible. To extend the infrared frequency reference grid of the BNM-LPTF, it is then necessary to built an alternative, broadly tunable source to CO_2 laser. This source is obtained by difference frequency generation (DFG) of two diode lasers in a nonlinear crystal. DFG spectrometers with nanowatt or microwatt level output have already been reported for high resolution spectroscopy and trace gas detection. Usually, high power lasers, such as diode laser-pumped Nd:YAG, Ti:sapphire, or dye lasers are combined with diode lasers to generate cw IR radiation from 3 μm to 6 μm [2-5]. Recently, for compactness and portability of these spectrometers, many groups are interested in an all-diode-laser based, broadly tunable device [6-8]. Such a portable device can contribute efficiently to local atmospheric gas monitoring. This paper report on our compact and portable, 9-μm to 11-μm tunable, all-diode-laser based difference frequency source and its performance. To our knowledge, this is the first time that such a long wavelength all-diode-laser based difference-frequency spectrometer is being developed. Our aim is to generate a microwatt level power to perform high resolution spectroscopy of OsO_4 so as to replace the CO_2 laser in its spectral gaps in the BNM-LPTF frequency chain.

2. DUAL-CAVITY, HIGH RESOLUTION DFG EXPERIMENTAL SETUP

To perform high resolution spectroscopy, the near infrared diode lasers are two extended cavity diode lasers (ECDL) operating around λ_3=778 nm (pump) and λ_2=842 nm (signal). The pump wavelength (λ_3) is chosen with respect to the BNM-LPTF rubidium two-photon frequency standard. The stabilization of the two ECDL to the same high finesse cavity gives a sub-kilohertz linewidth. The advantage of the use of the same cavity for this stabilization is the correlation between the frequency residual noise of the two ECDL, and so, by difference-frequency mixing, the generated IR radiation would have a better spectral purity than each of the ECDL. The setup of this stabilization is described elsewhere [9, 10]. The power available from these ECDL are

not enough for the difference-frequency mixing setup. The two high purity radiations are then used to injection-lock two other diode lasers with nominal power P_3=50 mW (at λ_3=778 nm) and P_2=200 mW (at λ_2=842 nm). In this injection-locking stage, and to allow fine-tuning of the spectrometer, the 842-nm master radiation is double-passed through a 200-MHz acousto-optic modulator controlled by a radiofrequency synthesizer. This allows fine-tuning over 150 MHz with a step as small as 1 Hz. After beam reshaping and isolation, the powers available for the difference-frequency mixing are 35 mW at 778 nm and 120 mW at 842 nm.

The dual-cavity difference-frequency experimental setup is sketched in Fig. 1. In this final setup and due to its two times higher conversion efficiency, a type-II phase-matching (e-o→e) cut $AgGaS_2$ crystal is chosen with respect to the type-I cut (o-o→e) [9]. This crystal, phase-matchable in its whole transparency window (0.47-12 μm), is the only nonlinear material suitable for our dual-cavity design [11]. The latter is chosen to simultaneously enhance the pump and signal low power. Such a cavity has already been used to control the simultaneous resonance of the pump and idler radiation in a type-II OPO [12, 13], but to our knowledge, it is the first time that it is used in a DFG experiment. The advantage of such a cavity is the possibility to independently stabilize each cavity arm to the corresponding laser frequency. Moreover, compared to a classical linear cavity, the two-arm cavity allows a broader tunability. The two arms of the dual cavity are coupled by a low-loss polarizing beam splitter (PBS). This coupling is possible only because the two laser diode radiation must be cross-polarized to satisfy the collinear phase-matching condition. The type-II phase matching cut $AgGaS_2$ crystal is placed in the common arm of the dual-cavity for the difference-frequency mixing. The optical path length of the 778-nm cavity arm (respectively, the 842-nm cavity arm) is 12.9 cm (resp., 9.8 cm) corresponding to a 75 μm (resp., 96 μm) optimal beam waist. Due to a poor spatial beam profile of the two slave diode lasers, the use of an anamorphic prism pair (APP) is necessary to mode-match the laser mode to the TEM_{00} mode of the cavity: without the APP, the mode-matching rate is about 35% while it is increased to about 75% with the APP. The generated infrared radiation is collimated by a ZnSe lens before being focused in a 300-μm diameter, liquid nitrogen cooled HgCdTe detector.

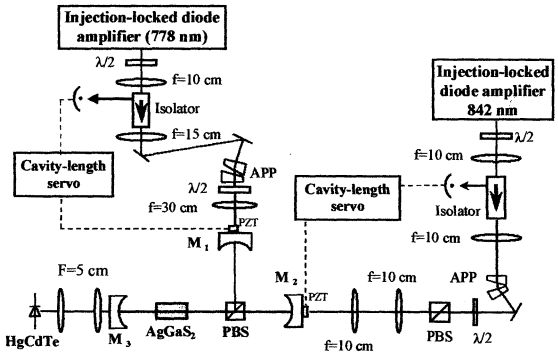


Fig. 1: Dual-cavity DFG experimental set-up. The use of a dual-cavity allows to have an independent cavity length servo to the corresponding laser frequency. The output coupler is made of ZnSe and is coated for high reflection in the near-IR (R>99%) and high transmission in the IR (T=90%). APP=Anamorphic prism pair, PBS=Polarizing beam splitter.

3. DUAL-CAVITY CHARACTERISTICS

The input coupler of the two arms are chosen to impedance-match the intracavity loss. The latter is estimated to about 8% by separately measuring the loss of each intracavity optical component. To optimize the impedance-matching, we have tested three input couplers for each cavity arm. Their characteristics and the performance obtained are presented in Table 1. The best input couplers have 10.8% and 9.2% transmission, respectively at 778 nm and 842 nm. From the measured finesse F of the cavity, one can retrieve the real intracavity loss (p): $p=2\pi/F-T$, where T is the transmission factor of the input mirror. In our case, p is about 15%, that is two times higher than the estimated one. The high transmission factors of the input couplers are then necessary to impedance match this high intracavity loss.

	T (%) @778nm	Finesse	Coupling efficiency (%)	S	T (%) @842nm	Finesse	Coupling efficiency (%)	S
Mirror 1	4.3	33	44	2.8	3.8	33	33	1.7
Mirror 2	7.1	29	55	3.8	9.2	26	47	2.9
Mirror 3	10.8	22	79	4.6	4.1	31	37	2.4

Table 1: Measured characteristics of the 778-nm arm (left hand side) and the 842-nm arm (right hand side) of the dual-cavity. Mirror 2 and mirror 3, with respectively 9.2% and 10.8% transmission at 842 nm and 778 nm, allow to impedance match the intracavity loss. The resulting enhancement factors, S, of the 778-nm arm and the 842-nm arm are respectively 4.6 and 2.9. These small enhancement factors are due to the high intracavity loss (15%).

The output coupler is a ZnSe mirror with a 90% transmission in the infrared and a reflectivity better than 99% in the near-infrared. Due to the high intracavity loss, the estimated enhancement factor is only 2.9 for the 842-nm arm and 4.6 for the 778-nm arm.

4. EXPERIMENTAL RESULTS

In Fig. 2, the infrared powers are plotted against the 842 nm powers for a fixed 778-nm laser power (35 mW) in front of the cavity. The lower curve is obtained when no cavity is used (single-pass DFG experiment). The corresponding conversion efficiency is 1.8 $\mu W/W^2$, which is 35% lower than our previous measured value [14] and three times lower than the theoretical one (5.2 $\mu W/W^2$). This discrepancy can be explained by the elliptical cross section of the infrared radiation (due to walkoff angle) and by the fact that HgCdTe detector is placed far from the nonlinear crystal (>1 meter): part of the focused spot falls outside the 300-µm diameter detector. The measured power is estimated to about two times lower than the real generated power. To remedy to this problem, a ZnSe anamorphic prism will be placed in the beam path so as to circularize the IR beam profile.

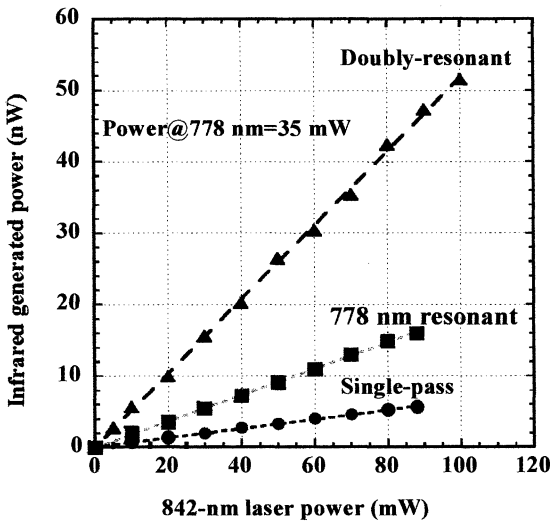


Fig. 2: Compared conversion efficiencies for single pass, singly resonant and doubly-resonant configuration. The measured enhancement factor is about 4.6 for the 778-nm arm and 3 for the 842-nm arm. This factor can be improved by reducing the intracavity loss and by a better impedance-matching.

The intermediate curve is obtained when only the 778-nm radiation is resonating (conversion efficiency=5.8 $\mu W/W^2$). And the upper curve is obtained when the two radiations are resonating (conversion efficiency=15.76 $\mu W/W^2$). By comparing the resonant DFG experimental conversion efficiency to the single-pass DFG conversion efficiency, one can retrieve the measured enhancement factor: 3.2 for the 778-nm arm and 2.8 for the 842-nm arm. This corresponds to the estimated enhancement factors deduced from the cavity finesses (table 1). From P_3=35 mW and P_2=100 mW in front of the dual-cavity, more than 50 nW of IR radiation is generated. To optimize the generated power to the microwatt level, reduction of the intracavity loss is necessary.

5. PRELIMINARY STUDY OF SF_6 WITH THE DOUBLY-RESONANT DEVICE

A first demonstration of our DFG spectrometer capability has been performed with a singly resonant (778-nm resonant) configuration [14]. In the latter reference, we have shown linear absorption spectra of SF_6 in cell while the 778-nm master frequency is continuously tuned over 2 GHz. Presently, our spectrometer is working in a doubly-resonant configuration with more than 50 nW power. The tuning procedure is the following. For a targeted IR wavelength λ_{IR}=10.56 µm, corresponding the P(18) absorption lines of SF_6 for instance, broad tuning can be obtained in three ways: crystal temperature tuning, crystal angle tuning or pump wavelengths tuning. Angle tuning is not convenient, since it requires a re-alignment of the cavity arms. Actually, it is possible to broadly tune the IR wavelength while keeping the same phase-matching crystal angle θ, by changing simultaneously the near-IR wavelengths. Indeed, the phase-matching condition

$$\frac{n_{IR}^{e}(\theta)}{\lambda_{IR}} = \frac{n_{P}^{e}(\theta)}{\lambda_{P}} - \frac{n_{S}^{o}}{\lambda_{S}} \qquad (1)$$

with $1/\lambda_{IR} = 1/\lambda_P - 1/\lambda_S$, can be solved for one of the near-IR wavelengths while fixing θ and λ_{IR}. Hence for each targeted λ_{IR} one can always find a pair (λ_P, λ_S) satisfying (1), without any re-alignment of the cavity. For instance, for the above targeted IR wavelength, the two pumps have to be tuned to λ_P=780.2740 nm and λ_S=842.5434 nm. The λ_P wavelength acceptance bandwidth is quite large (2 nm, e.g.>1 THz). Fine continuous-tuning, without mode-hops, is obtained by sweeping the extended cavity length of the master pump (λ_P) laser (the corresponding dual-cavity arm follows up). Part of the sweeping voltage ramp has to be fed to the slave diode current to obtain about 2 GHz hysteresis-free tuning range. Up to 12 GHz tuning range can be obtained but with an hysteresis on the

descending ramp, causing the dual-arm length servo to unlock.

Fig. 3 shows a spectrum of the P(18)-P(19) manifolds of SF_6, obtained with three consecutive sweeps. The spectra is still contaminated by artifact lines coming from the intensity modulation of the IR baseline. These modulations originate from some etalon effects due to the spurious sub-cavity formed by the prism which reflects the IR power and the output mirror. We are implementing a baseline compensation technique to improve the quality of the Doppler spectra. Presently, the entire 27-33 THz spectral range can then be explored by steps of 2 GHz without hysteresis.

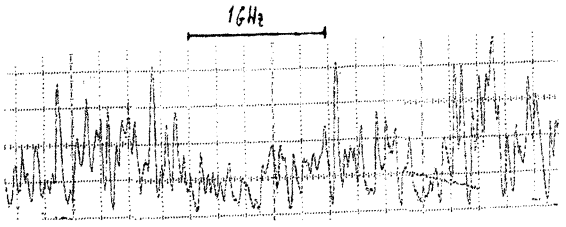


Fig.3: Preliminary Doppler spectrum of SF_6 (100 mTorr, cell length 20 cm). Some of the smaller intensity lines are artifacts.

6. CONCLUSION AND PERSPECTIVES

We have demonstrated the capabilities and performance of a portable and 9-μm to 11-μm tunable difference-frequency spectrometer using diode laser as pump sources. The spectrometer, which delivers more than 50 nW of IR radiation power, can work as a 30-GHZ accurate difference-frequency standard in the near-IR range. When used with a femtosecond (fs) laser comb, this spectrometer allows to directly link the microwave frequency standard to the optical frequency standard (Fig. 4), hence, allowing to avoid the complexity of the classical harmonic frequency chain. Indeed, the frequency-domain comb of such a fs laser is controllable and strictly uniform [15, 16] : its repetition rate is equal to the mode spacing Δ. Hence, by servo-locking the repetition rate to the cesium primary standard, this constant comb spacing Δ can be exactly known. Then by counting number of modes, n, which separate the two near-IR diode lasers, and by measuring the beatnote between each diode laser frequency and the nearest mode of the mode-locked laser, we can measure the absolute frequency of the IR radiation nΔ, so the absolute frequency of an OsO_4 molecule transition on which the IR radiation is stabilized.

Such a new frequency chain also allows transposing the IR frequency standard to the near-IR/visible domain. It will then be possible to disseminate the BNM-LPTF infrared frequency standard to the other laboratories by using only classical optical fibers.

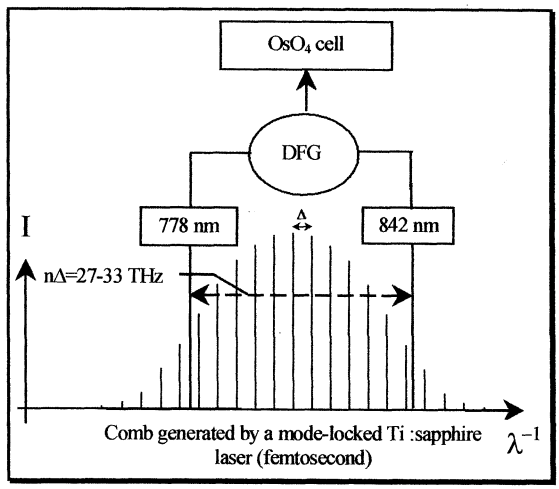


Fig 4 : Schematic of an absolute measurement of the OsO_4 saturated absorption lines by using a broadband mode locked Ti:Sapphire laser. This is a new frequency chain which allows to directly link the microwave and the optical frequencies.

7. REFERENCES

[1] O. Acef, "Metrological properties of CO_2/OsO_4 optical frequency standard", Opt. Commun., vol. 134, pp. 479-486, Jan. 1997.

[2] U. Simon, S. Waltman, I. Loa, F. K. Tittel, L. Hollberg, "External-cavity difference-frequency source near 3.2 μm, based on combining a tunable diode laser with a diode-pumped Nd:YAG laser in $AgGaS_2$", J. Opt. Soc. Am. B, vol. 12, pp. 323-327, Feb. 1995.

[3] D. G. Lancaster, D. Richter, R. F. Curl, and F. K. Tittel, "Real-time measurements of trace gases using a compact difference-frequency-based sensor operating at 3.5 μm", Appl. Phys. B, vol. 67, pp. 339-345, Sept. 1998.

[4] K. P. Petrov, R. F. Curl, and F. K. Tittel, "Compact laser difference-frequency spectrometer for multicomponent trace gas detection", Appl. Phys. B, vol. 66, pp. 531-538, May 1998.

[5] M. Seiter, D. Keller, and M. W. Sigrist, "Broadly tunable difference-frequency spectrometer for trace gas detection with noncollinear critical phase-matching in $LiNbO_3$", Appl. Phys. B, vol. 67, pp. 351-356, Sept. 1998.

[6] B. Sumpf, T. Kelz, M. Nägele, and H.-D. Kronfeldt, "A cw $AgGaS_2$ difference frequency spectrometer with diode lasers as pump sources", Appl. Phys. B, vol. 64, pp. 521-524, May 1997.

[7] B. Sumpf, D. Rehle, T. Kelz, and H.-D. Kronfeldt, "A tunable diode-laser spectrometer for the MIR region near 7.2 μm applying difference-frequency generation in $AgGaSe_2$", Appl. Phys. B, vol 67, pp. 369-373, Sept. 1998.

[8] D. Richter, D. G. Lancaster, R. F. Curl, W. Neu, and F. K. Tittel; "Compact mid-infrared trace gas sensor based on difference-frequency generation of two diode lasers in periodically poled $LiNbO_3$", Appl. Phys. B, vol. 67, pp. 347-350, Sept. 1998

[9] T. Kaing, J.-J. Zondy, A. Yelisseyev, S. Lobanov and L. Isaenko, "Improving the power and spectral performance of a 27-33 THz $AgGaS_2$ difference-frequency spectrometer", IEEE Trans. Instrum. Meas., vol. 48, No. 2, pp. 592-595, April 1999.

[10] L. Hilico, D. Touahri, F. Nez, and A. Clairon, "Narrow line, low-amplitude noise semiconductor laser oscillator in the 780 nm range", Rev. Sci. Instrum., vol. 65, pp. 3628-3633, Dec. 1994.

[11] 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, vol. 67, pp. 363-367, Sept. 1998.

*Corresponding author.
E-mail: To.kaing@up.univ-mrs.fr
Phone.: +33 4 91 28 80 26
Fax.: +33 4 91 28 87 45

[12] F. G. Colville, M. J. Padgett, and M. H. Dunn, "Continuous-wave, dual-cavity, doubly resonant, optical parametric oscillator", Appl. Phys. Lett., vol. 64, pp. 1490-1492, March 1994.

[13] D. Lee, N. C. Wong, "Tuning characteristics of a cw dual-cavity KTP optical parametric oscillator", Appl. Phys. B, vol. 66, No. 2, pp. 133-143, 1998.

[14] T. Kaing, J.-J. Zondy, A. Yelisseyev, S. Lobanov, and L. Isaenko, "Compact, broadly tunable, mid-IR source for the spectroscopic investigation of molecular reference lines in the 27- to 33-THz range", IEEE Trans. Ultrason. Ferroelect. Freq. Contrl., vol. 47, No. 2, March 2000

[15] Th. Udem, J. Reichert, R. Holzwarth, and T. W. Hänsch, "Accurate measurement of large optical frequency differences with a mode-locked laser", Opt. Lett., vol. 24, No. 13, pp. 881-883, July 1999

[16] S. A. Diddams, D. J. Jones, L.- S. Ma, S. T. Cunddiff, and J. L. Hall, "Optical frequency measurement across a 104-THz gap with a femtosecond laser frequency comb", Opt. Lett., vol. 25, No. 3, pp. 186-188, Feb. 2000

CO_2 STABILISED ER^{3+}-DOPED FIBRE LASER AT 1578 NM

H. Simonsen, J. Henningsen and S. Søgaard
Danish Institute of Fundamental Metrology (DFM)
Building 307, Anker Engelunds Vej 1, DK-2800 Lyngby, Denmark
Phone: +45 4593 1144, Fax: +45 4593 1137, Email: jh@dfm.dtu.dk
and
J. Engholm Pedersen
IONAS A/S, CAT Building 347, DK-2800 Lyngby, Denmark
Phone: +45 4525 6314, Fax: +45 4525 6405, Email: jep@ionas.dk

ABSTRACT

A single-mode distributed feedback (DFB) fibre laser has been frequency locked to a CO_2 absorption line at 1578.665 nm. A mechanical lead screw provided coarse wavelength tuning, while wavelength modulation and fast frequency corrections were applied by straining the fibre laser with piezoelectric transducers. With a suitable choice of absorption line such a system may serve as absolute frequency reference for DWDM grids in telecommunication.

1. INTRODUCTION

Single frequency lasers with narrow line width are attractive for applications such as laser spectroscopy, for sensor systems, or as wavelength standards in optical communication. Distributed feedback (DFB) erbium doped fibre lasers meet the requirements, and with the excellent wavelength selectability throughout the telecommunication C and L bands ranging from 1525 to 1610 nm, they represent an interesting alternative to semiconductor lasers. The drawback of fibre lasers in connection with some of these applications is that tuning and modulation is less straightforward to implement than is the case for semiconductor DFB lasers. We here apply the strain approach which has recently been described by Wetjen et al. [1]. The laser is locked by the first derivative technique to an absorption line in CO_2, and the stability of the system is validated by heterodyning against an extended cavity laser (ECL), offset locked to the same line. One possible application of such a system is to serve as an absolute wavelength reference in DWDM grids used for telecommunication [2].

2. EXPERIMENTAL SETUP

The DFB fiber laser is fabricated by introducing a Bragg grating into the core of an Er^{3+} doped fiber [3]. The unpacked fiber laser is epoxy-bonded at two spots just outside the grating, and one end is adjustable by a lead screw and two piezoelectric transducers (PZT). The entire base of the aluminium mount is temperature controlled to ensure passive stability. Output powers up to 4 mW were achieved. The tuneable fibre system was characterised using a 2 GHz scanning Fabry-Perot interferometer, a Burleigh WA-1500 wave meter, and through beat frequency measurement, using a New Focus model 6262 extended cavity laser (ECL) tunable from 1510 - 1590 nm.

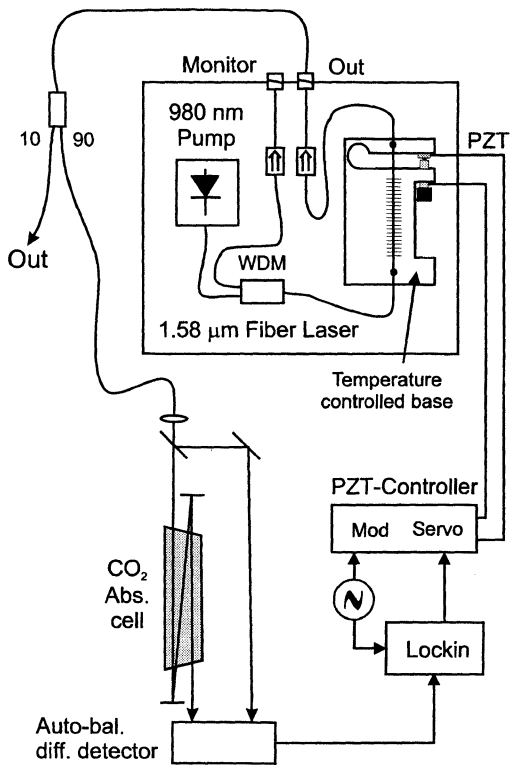


Fig. 1. Experimental setup for a fibre laser locked to a CO_2 absorption line.

3. TUNING AND MODULATION

Coarse tuning as illustrated in Fig.2 is accomplished with the lead screw at a rate of 6.8 nm/mm, referred to the displacement of the screw, corresponding to about 11 nm/mm referred to the stretching of the fibre. To avoid damage to the fibre, tuning was limited to 2.7 nm (325 GHz), corresponding to a relative length change of 0.0023. Based on geometrical considerations alone, this strain would imply a tuning of 3.6 nm, the difference reflecting the strain induced change of the refractive index.

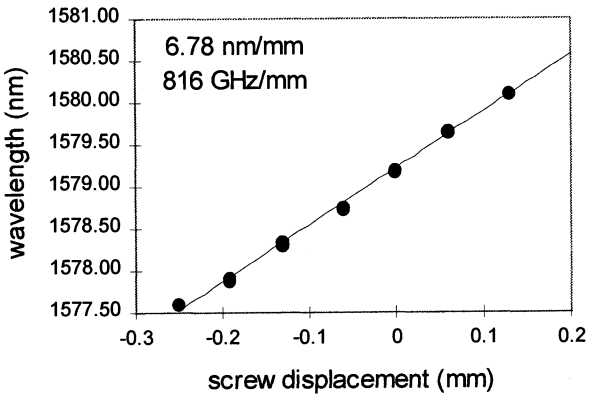


Fig.2 Coarse tuning by mechanical lead screw

Intermediate range tuning is possible by controlling the temperature of the base plate to which the fibre laser is thermally anchored. Over the range 20-30 °C the wavelength tunes linearly with a coefficient of 0.0326 nm/°C, as shown in Fig.3.

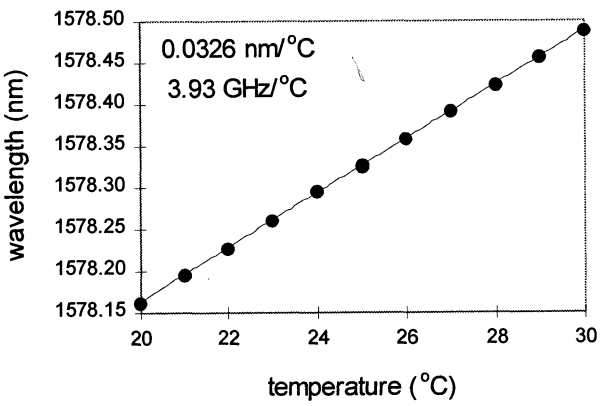


Fig.3 Temperature tuning of fibre laser

Modulation is induced through either of two piezoelectric transducers, a coarse PZT with a nominal displacement of 15 μm at 150 V, and a fine PZT with a displacementof approximately 0.9 μm at 150 V. The frequency response of both piezos as given in Fig.4 shows a uniform fall-off with frequency up to about 2 kHz, while several mechanical resonances were observed at higher frequencies. Maximum tuning was limited to 0.12 nm (14.4 GHz) for the coarse piezo and 0.0055 nm (660 MHz) for the fine piezo by the maximum voltage at the piezos.

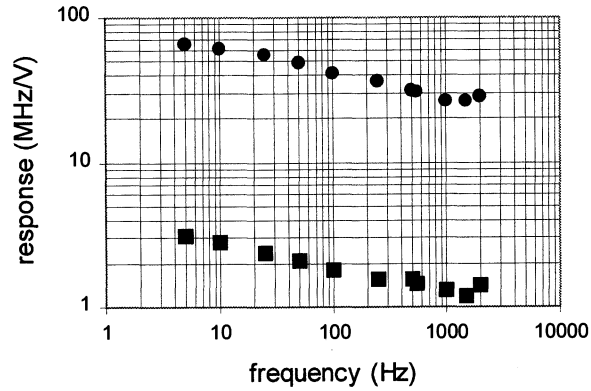


Fig. 4 Frequency response of coarse PZT (circles) and fine PZT (squares).

An undesirable feature of the laser used in these experiments is the presence of an unexpectedly strong dependence of output power on the fibre strain, and an associated strong amplitude modulation induced by strain modulation. Whether this is an inherent property of the fibre, or whether it is induced by the mounting of the laser, is not known at present.

4. FREQUENCY LOCKING

In order to enable locking to an absolute frequency reference, the fibre laser was designed to operate at 1578 nm, in the vicinity of absorption lines of the 30012 ← 00001 combination band of CO_2. For the present work we chose the P16 line at 1578.665 nm [4] with a line strength of $1.56 \cdot 10^{-23}$ cm/mol and a self broadening parameter of 2.96 MHz/mbar [5].

The fibre laser was locked to the line centre of the absorption dip produced by three passes through a 20 cm cell sealed with Brewster windows, and filled with CO_2 to a non-optimised pressure of about 200 mbar. Wavelength

modulation was induced through the fine piezo at 555 Hz with a peak-to-peak modulation width of 130 MHz. The use of an auto-balanced photoreceiver (New Focus Nirvana) enabled elimination of the offset of the 1-f signal, created by the amplitude modulation (Fig.5). The origin of the residual etalon effects has not been identified so far.

In order to produce a beat frequency suitable for counting, the ECL was locked to the side of the same absorption line, using a separate single pass cell of 127 cm length, also filled to a pressure of about 200 mbar. A reference signal was branched off in front of the cell and subtracted from the transmitted signal in a difference detector. By adjusting the magnitude of the reference signal, the zero crossing of the combined signal could be shifted relative to the line centre in order to produce an offset of about 650 MHz.

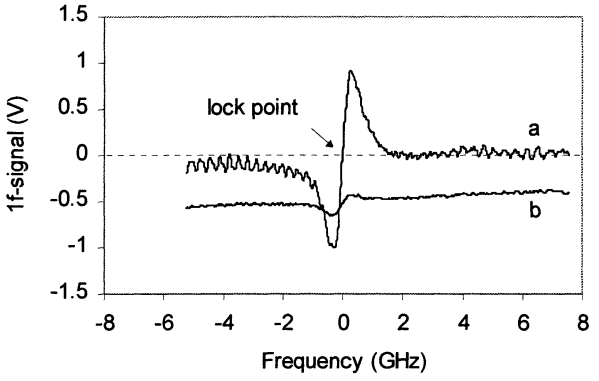


Fig. 5. First harmonic profile of the P16 line of CO_2 at 1578.665 nm obtained with a tuneable fibre laser and auto-balanced detection (a) or traditional balanced detection (b).

5. ALLAN VARIANCE

Signals from the two lasers were combined on a fast detector, and the beat frequency was displayed on a Tektronix 2756P spectrum analyser and counted. An Allan variance analysis of the countings indicates that the lock has no effect for integration times shorter than about 20 s, whereas for longer integration times the stability is improved by about one order of magnitude (Fig.6). The long-term beat frequency stability of around $5 \cdot 10^{-8}$, corresponding to 10 MHz, was limited by the sensitivity of the ECL lock to changes in the level of the signal transmitted through the 127 cm cell. It is therefore reasonable to take 10^{-8} or 2 MHz, as measured for integration times less than 20 s, as an upper limit of the fibre laser stability. From the beat note spectrum we could estimate the line width of the fibre laser to a few MHz. This is somewhat higher than reported for both the ECL and a packed fibre laser [5], and may be caused by acoustical noise in the laboratory as well as by the construction of the fibre laser mount. Another limitation at present may be caused by the significant power variation with frequency tuning of the fibre laser, and the associated amplitude modulation. These effects are under further investigation.

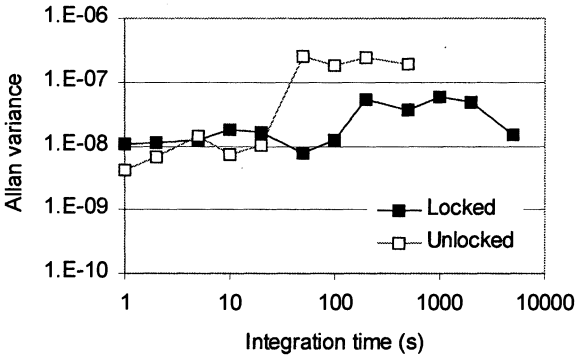


Fig.6 Allan variance analysis of the beat frequency between the fibre laser and the ECL.

6. SUMMARY

Due to the low noise and narrow line width, tuneable DFB fibre lasers are attractive sources for metrologic and spectroscopic applications. An absolute wavelength standard with 2 MHz stability, locked to a Doppler-broadened absorption line in CO_2 and suitable for DWDM applications in telecommunication, has been demonstrated. Other candidate molecules for operation in the 1520 to 1610 nm wavelength range include C_2H_2, HCN, H_2S, and CO. Improvement of the stability of the present wavelength standard may be achieved through optimisation of the absorption cells and reduction of the laser amplitude modulation, and a tighter limit on the stability may be obtained by a more sophisticated locking of the reference laser.

The use of Doppler free techniques would in principle lead to a significant improvement of the stability. However, in the context of present day DWDM grids with minimum channel widths of 50 GHz [2], locking to Doppler broadened lines is sufficient to meet the requirements of telecommunication.

ACKNOWLEDGEMENT

This work has been supported by base funds from the Danish Agency for Development of Trade and Industry

REFERENCES

[1] E.T. Wetjen, D.M. Sonnenfroh, M.G. Allen, and T.F. Morse, "Demonstration of a rapidly tuned Er^{3+}-doped fiber laser for sensitive gas detection", *Appl. Optics*, Vol. 38, pp. 3370-3375, 1998.

[2] "Optical interfaces for multichannel systems with optical amplifiers", ITU-T Recommendation G.692 (10/98)

[3] IONAS Production Information Sheet - "Single Channel DFB Fiber Laser, Model no. IFL01W1E-HP", March 1998.

[4] L.S. Rothman, C.P. Rinsland, A.Goldman, S.T. Massie, D.P. Edwards, J.-M. Flaud, A. Perrin, C. Camy-Perot, V. Dana, J.-Y. Mandin, J. Schroeder, A. McCann, R.R. Gamache, R.B. Wattson, K. Yoshino, K.V. Chance, K.W. Jucks, L.R. Brown, V. Nemtchinov, P. Varanasi, "The HITRAN molecular spectroscopic database and HAWKS: 1996 edition", *J. Quant. Spectrosc. Rad. Transf.*, Vol. 60, 665-710 (1998).

[5] J. Henningsen and H. Simonsen, "The 30012 $\leftarrow$ 00001 Band of CO_2. Line Strengths, Broadening Parameters and Pressure Shifts", *J. Mol. Spectrosc.* (in press).

TOWARDS A NOVEL FREQUENCY STANDARD IN THE INFRARED REGION USING A SUPERSONIC BEAM OF SF_6

A. Amy-Klein, L.F. Constantin, A. Shelkovnikov[1], R.J. Butcher[2], and Ch. Chardonnet

Laboratoire de Physique des Lasers (UMR 7538 CNRS), Université Paris 13,
99 av. J.B. Clément, 93430 Villetaneuse, France
tel : 00 33 1 49 40 33 79 - fax : 00 33 1 49 40 32 00 - amy@lpl.univ-paris13.fr

ABSTRACT

The present 10 μm secondary frequency standards based on the CO_2 laser locked onto a saturated absorption resonance of OsO_4, reach an accuracy of a few $\times$ 10^{-13}. This paper deals with our attempt to increase this performance by using a two-photon resonance and a molecular beam. We report observation of Ramsey fringes on a 10.6 μm two-photon transition of SF_6. We used two interaction zones separated by 8 cm or 20 cm and fringes of periodicity down to 1.5 kHz were recorded. The experimental apparatus might be adapted to increase the resolution by an order of magnitude. This system is thus a very serious candidate for a frequency standard with a potential accuracy in the 10^{-14}-10^{-15} range.

1. INTRODUCTION

CO_2 lasers stabilized on narrow sub-Doppler molecular resonances are widely used as secondary frequency standards in the 9-12 μm spectral region. Using this technique, the CO_2 laser locked onto a saturated absorption resonance of OsO_4 provides a secondary frequency standard in the 10 μm region, with very good stability characteristics [1, 2]. Its accuracy reaches 10 Hz for the optimal lines.

To improve this performance the well-known strategy is to record narrower and narrower lines ; since most of the systematic errors are proportional to the linewidth of the frequency reference any increase in the experimental resolution will result in an increase in the ultimate accuracy of the frequency standard. To push the transit broadening limit down, we tested the method of slow molecule detection [3-5]. The resolution was increased up to a few hundreds of Hz, but with a degraded signal-to-noise ratio (SNR). Alternatively, the possibility of slowing and trapping molecules is very limited for the moment [6].

Alternatively to cell spectroscopy, a well-known technique to reduce the transit broadening is to use a molecular beam and to apply the method of separated fields of Ramsey [7]. Nevertheless the possibility of high resolution spectroscopy with a molecular beam has been little explored at 10 μm. Ramsey fringes have been observed with SF_6 using a saturated absorption resonance [8]. Recently some two-photon lines were observed and calculated in SF_6, opening the opportunity to develop two-photon Ramsey fringes [9-11].

This paper presents some preliminary results of narrow Ramsey fringes obtained using a two-photon Doppler-free absorption.

2. TWO-PHOTON RAMSEY FRINGES

The main point of Ramsey spectroscopy [7] is to provide a signal of which the linewidth is no longer limited by the transit time through the laser beam. For this purpose, the single interaction zone is replaced by two spatially separated zones, with a fixed relative optical phase. When a molecule has passed through the first zone of interaction it is in a coherent superposition of lower and upper levels. The coherence precesses freely between two zones. In the second zone, the molecule is either excited or de-excited, depending on the relative phase between the excitation field and the coherence. Thus, fringes develop in the excitation probability versus the laser frequency, and their spacing depends on the transit time between the two zones.

In the optical domain, this method must be associated with a sub-Doppler technique in order to avoid scrambling of the fringe pattern. The use of saturated absorption and a three- or four-zone configuration [8, 12, 13] imposes severe conditions on parallelism and equidistances which, in practice, limit the distance between zones and, finally, the ultimate resolution. By contrast, in the case of two-photon spectroscopy only two zones are required, each comprising a standing wave, and the Doppler shift is compensated in each one [14]. The experimental signal should exhibit fringes of spacing half the inverse of the transit time between two zones for the mean velocity. The contrast is unity [13, 15] if the observed transition is well isolated. The only constraint is that the relative phase of the standing wave between two zones vanishes. This is easily fulfilled by generating both

[1] Permanent address : Lebedev Institute of Physics, Russian Academy of Sciences, Leninskii pr. 53, Moscow, 117924 Russia.
[2] Permanent address : The Cavendish Laboratory, Madingley Road, Cambridge, CB3 OHE, UK.

standing waves inside the same Fabry-Perot cavity; in addition this guaranties that the central fringe coincides with the molecular resonance. The side fringes will be attenuated due to the longitudinal velocity dispersion and the entire pattern is superimposed on the broader two-photon signal arising from the absorption in one single zone.

3. EXPERIMENTAL APPARATUS

Preliminary 10 kHz two-photon Ramsey fringes were already observed on SF_6 elsewhere [16], with a resolution limited by the hyperfine structure and the laser linewidth. We plan to reach much higher resolution, and have constructed an apparatus with an adjustable interzone distance, to give fringes from 2 kHz to 50 Hz.

The experiment was performed on the P(4) E transition of the $2\nu_3$ band of SF_6. The two-photon resonance is composed of four main hyperfine components spanning approximately 40 kHz [17]. The detuning from the intermediate level is 76 MHz [11].

The experimental set-up is presented in Figure 1. It is composed of three parts: the frequency-controlled laser system, the Fabry-Perot cavity, and the molecular beam apparatus.

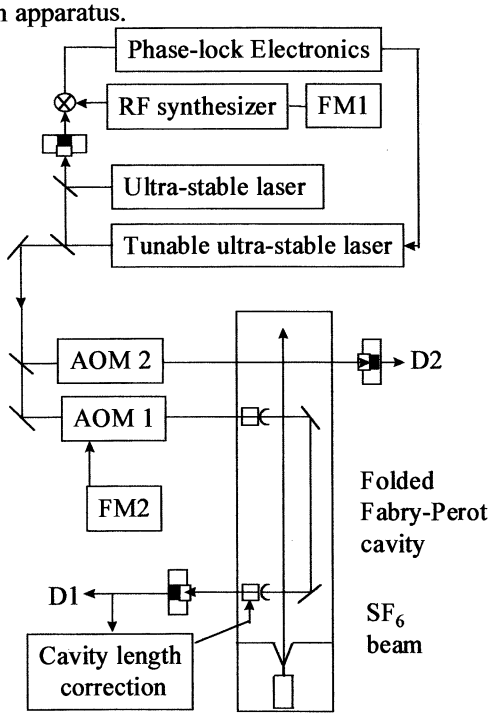


Figure 1 : Experimental scheme (AOM : acousto-optic modulator, FM : frequency modulation)

3.1 The laser sources

The performance of our frequency-controlled laser system has already been described in detail [2]. It employs two CO_2 lasers emitting in the 10 μm region. The first laser is frequency-locked to a strong saturation line of OsO_4, and the beat note between this laser and the second laser is then phase-locked to a tunable synthesizer. Thus, the stability acquired by the OsO_4 lock is transferred to the second laser, which is also tunable around each emission line of CO_2. The main stability characteristics are: a laser linewidth of the order of 6 Hz (FWHM) which is 2×10^{-13} in relative value, and an Allan variance of 0.1 Hz ($\Delta\nu/\nu = 3.5\times10^{-15}$) for a time constant of 100 s. The reproducibility was estimated to be 10 Hz. This stability performance ensures that our experimental resolution will not be limited by the laser itself.

3.2 The folded Fabry-Perot cavity

A key point of the experiment is the generation of the two phase-coherent standing waves forming the two interaction zones for the Ramsey spectroscopy. They are realized within a four-mirror folded Fabry-Perot cavity. The cavity is symmetric; the coupling mirrors are plano-concave with a 50 m curvature, while the folding mirrors are plane. The beam radius is almost constant inside the cavity and is approximately 3 mm for a 10 or 20 cm interzone spacing. The reflectivity of each mirror is 99.8%, leading to a finesse of 500.

The mirrors are mounted on an independent structure and can be translated along this structure to adjust the interzone distance. Each concave mirror is mounted on a piezoelectric transducer (PZT) in order to modulate or adjust the cavity length, while the transmitted signal is used for locking the cavity resonance onto the laser frequency.

3.3 The molecular beam

The molecular beam apparatus was designed for a supersonic expansion of SF_6. This allows a rotational cooling of the beam, which produces a strong enhancement of the population in the J=4 level. We use either a pure SF_6 or a He-seeded beam, with an input pressure of a few $\times10^5$ Pa. In the latter case, the expansion has better performance, but the velocity is faster, so the transit time is shorter, and the resolution is smaller.

The apparatus is 3 meters long with two vacuum chambers separated by a skimmer. For a pure SF_6 beam, the mean velocity is typically 370 m/s with a longitudinal dispersion of 55 m/s. The beam divergence is of the order of 30 mrad. The total flux in the J=4 level is a few 10^{12} molecules/s, and the measured rotational temperature is less than 30 K.

4. EXPERIMENTAL RESULTS

Fringes are recorded on the transmission of the cavity, detection channel D1 in Fig. 1, to take

advantage of the enhancement of the signal contrast due to this cavity. For low-noise detection purpose the laser frequency is modulated through the synthesizer driving the phase-lock loop. Alternatively, the signal can be read on the molecular beam itself by probing its absorption on the upper 1-photon transition using rapid adiabatic passage with an auxiliary beam, detection channel D2 in Fig. 1. If the modulation depth is small enough, the laser frequency modulation can be applied to the radiofrequency driving the AOM 1, and the auxiliary beam is not then modulated.

4.1 8-cm interzone

Fig. 2 displays an ultra-high resolution two-photon spectrum for a 8 cm spacing between zones [18]. The central part exhibits Ramsey fringes with more than 10 oscillations, which are superimposed on the broader signal arising from the two-photon absorption in a single zone.

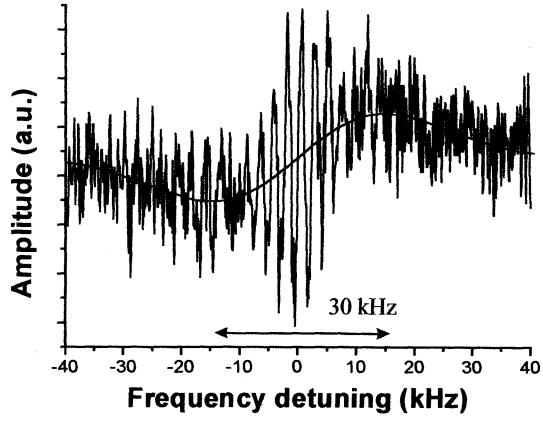


Figure 2 : Two-photon absorption signal detected on the cavity transmission beam. Experimental conditions are : laser FM at 1.012 kHz, depth 500 Hz, 16 mW in the cavity, $5x10^5$ Pa pure SF_6 beam, accumulation time 2.4 s/point. Data are fitted with the derivative of a Gaussian.

The fringes were recorded on the transmission of the cavity for an input pressure of 5 $\times 10^5$ Pa. The optimum power in the folded standing wave was about 16 mW, which corresponds roughly to a $\pi/2$ pulse for the two-photon transition. The background signal fits very well to the derivative of a Gaussian, peak-to-peak width 30 kHz. The width is a convolution of the transit lineshape (peak-to-peak width 28 kHz) with the hyperfine structure. The central part of the spectrum, corresponding to the main hyperfine component, has been fitted with a sine superimposed on a slope. This yields a fringe periodicity of 2.3 kHz and a relative determination of the central frequency with better than 10 Hz precision. This is limited by the SNR, which is approximately 3 for 1 s. The periodicity is consistent with the theoretical value u/2D with u=370 m/s and D=8 cm. The contributions of other hyperfine components are less visible.

Since we are using FM at low frequency, the amplitude of the fringes should be greatly enhanced compared to the two-photon absorption background, and this is obviously not the case. Taking into account the molecular beam divergence and the hyperfine structure is not sufficient to explain this loss of contrast. From comparison with a line shape simulation we deduced that the fringes arise from roughly one order of magnitude less molecules than expected. This loss of signal could originate in an imperfect alignement of the two zones relative to the beam axis, or to a partial molecular decoherence between two zones, due to residual small angles collisions. Different tests are now under progress to discriminate between these processes.

Fringes recorded by detection channel D2, Fig. 1, using rapid adiabatic passage exhibited a little smaller SNR (2 for 1 s) while the signal itself was reduced roughly an order of magnitude. The SNR is limited by the technical intensity noise of the auxiliary beam. Thus this detection channel, which has theoretically no background, should benefit from an active reduction of the laser intensity noise.

4.2 20-cm interzone

As a second step, to increase the resolution, we recorded some fringes with a 20-cm interzone. But with a pure SF_6 beam, the SNR was degraded to about 1 for 10 s averaging time.

To overcome this problem, we decided to seed the beam with He, since its supersonic expansion has better performance. In addition the dynamics of the heavier SF_6 molecules inside the He beam enhance the partial flux in the beam center. Figure 3 displays a typical spectrum performed with 25 % SF_6 in He and a

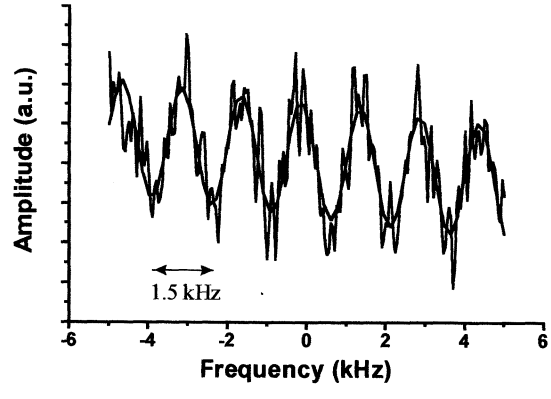


Figure 3 : Fringes detected on the cavity transmission beam. Experimental conditions are: FM applied on the laser PZT at 462 Hz, depth 400 Hz, 26 mW in the cavity, 25 % SF_6 beam at $5x10^5$ Pa, 5 s/point. Data are fitted with a sine superimposed on a slope.

20 cm interzone spacing. Data are fitted with a sine superimposed on a slope, giving a 1.5 kHz periodicity This is consistent with the faster velocity of 590 m/s of the molecules. The fringes were recorded on the transmission of the cavity for a total input pressure of 5 $\times 10^5$ Pa. The SNR is roughly 2 for 1 s averaging time.

When decreasing the SF_6 proportion, the resolution is decreased but the signal is increased. Figure 4 displays a spectrum with 5 % SF_6 recorded on the transmission of the cavity for an input pressure of 4 $\times 10^5$ Pa. The background signal fits with the derivative of a Gaussian, peak-to-peak width 80 kHz. This is broader than in Fig. 2, since the velocity is now 1070 m/s and then the transit time is longer. The central part of the spectrum has been fitted with a sine superimposed on a slope, resulting in a fringe periodicity of 2.8 kHz. The SNR is 2 for 1 s The contributions of other hyperfine components are now evident.

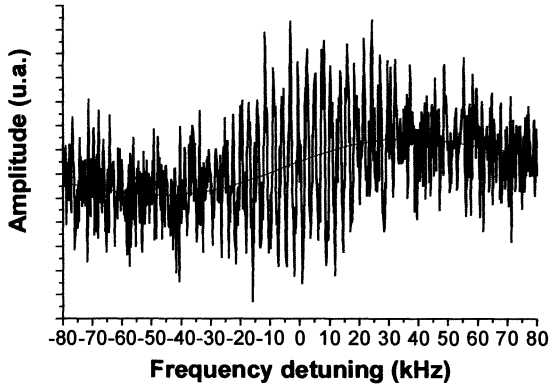


Figure 4 : Fringes detected on the cavity transmission beam. Experimental conditions : FM applied on the laser PZT at 962 Hz, depth 500 Hz, 47 mW in the cavity, 5 % SF_6 beam at 4x10^5 Pa, accumulation time 2 s/point. Data are fitted with the derivative of a Gaussian.

5. METROLOGICAL FEATURES

In this last part we consider the different effects which might affect the accuracy or reproducibility of a frequency standard based on two-photon Ramsey fringes [19].

First we should point out that there is a negligible recoil effect in Doppler-free two-photon spectroscopy [15]. Also there can be no pressure shift for the Ramsey fringes, since there are no collisions in a beam.

The magnetic shift is negligible provided we set a μ-metal shield to cancel the earth's field. We can easily obtain a 0.1 % attenuation, which should result in a shift less than 0.1 Hz. Note that the Zeeman effect was previously measured for a P(39) and a P(46) $\Delta m=\pm 1$ line of OsO_4 [20], and shifts of the order of 100 Hz/10^{-4} T were obtained.

To estimate the shift induced by the second-order Doppler effect (SODE), we have only to take into account the contribution of the velocity component along the molecular beam. The transverse velocity gives a negligible contribution since the beam divergence is small. The shift is $(\nu_{eg}/2)(u^2/2c^2)\left(1+1/2(\Delta u/u)^2\right)$, with $(\nu_{eg}/2)$ the two-photon frequency, u the mean velocity in the beam, and Δu the longitudinal dispersion. We calculate a 23 Hz shift where the second term, corresponding to the velocity dispersion, contributes only 1%. The uncertainty depends on the accuracy of the mean velocity and is of the order of 1 Hz. It could be decreased by repeating the same experiment with a SF_6 beam seeded with different proportions of He, and measuring the different shifts. The experimental value u and Δu might also be measured more precisely from an analysis of the fringes.

The light shift is known to be the weak point of two-photon spectroscopy compared to saturated absorption. However this effect is not so important for two reasons. First, the resulting shift is proportional to the difference of the transition probabilities for the two one-photon transitions, which is small for the rovibrational transitions used here. Second, the effect is much reduced because the interaction with the strong field lasts a negligible time compared to the transit time between two zones. If we consider the dominant contribution of the intermediate level, the shift is [15, 21]: $\delta\nu = (\Omega_R/4)\left(|\mu_{rg}|/|\mu_{er}| - |\mu_{er}|/|\mu_{rg}|\right)(w/D)$; Ω_R is the effective angular Rabi frequency for the two-photon excitation, w is the beam waist, D is the interzone distance, $\mu_{\beta\alpha}$ is the transition dipole for the β<-α one-photon transition. g,r,e represent, respectively: the ground, intermediate and upper levels. For the P(4) E transition, we have $\left(|\mu_{rg}|/|\mu_{er}| - |\mu_{er}|/|\mu_{rg}|\right) = 0.25$ [22]. With a two-photon excitation equivalent to a $\pi/2$ pulse, waist 3 mm, the shift is 19 Hz for an interzone distance of 20 cm. This shift should be easily determined by extrapolating the measurements at different laser powers to zero power. The accuracy will depend on the SNR and will be proportional to the resolution.

Lastly, the black-body radiation might induce a non-negligible shift. For a ro-vibrational transition in a molecule the effect will generally be much smaller than for an electronic transition in an atom. Here, however, the effect will be magnified since the black body emission peak is precisely in the 10 μm region. In fact the Ramsey experiment could be used to measure this effect, which has never been studied in the 30 THz domain.

6. CONCLUSION

In this paper we present our results of two-photon Ramsey fringes in the 10 μm region. Our spectra exhibit a periodicity between 2.8 and 1.5 kHz, which improves by a factor 6 the previous results obtained with a saturated absorption resonance of SF_6 [8]. The fringes arise from a 8 cm or a 20 cm interzone separation, which demonstrates the experimental advantage of a two-photon configuration.

Using the same transition with detection of slow molecules in a cell [5], we achieved a linewidth of 280 Hz (HWHM) to be compared with the present 375 Hz linewidth (HWHM) of the fringes. But the present setup allows a better SNR and has the potential to increase the resolution by an order of magnitude. Considerable room are also available for both signal improvement and noise reduction. Concerning the metrological features, the systematic effects are either very small or can be easily and precisely measured or calculated. This system is thus a very serious candidate for a frequency standard in the infrared with a potential accuracy in the 10^{-15} range.

7. ACKNOWLEDGMENTS

This work has been supported by the Bureau National de Métrologie, the ETCA-DRET and the European Union through two HCM programs. We acknowledge Gonville & Caius College for financial support of R.J. B. and C.N.R.S. for financial support of A.S..

8. REFERENCES

[1] O. Acef, "CO_2/OsO_4 Lasers as Frequency Standards in the 29 THz Range", IEEE Trans. Instrum. Meas., vol. 46, pp. 162-165, 1997.

[2] V. Bernard, C. Daussy, G. Nogues, L. Constantin, P.E. Durand, A. Amy-Klein, A. van Lerberghe, and Ch. Chardonnet, "CO_2 laser stabilization to 0.1-Hz level using external electrooptic modulation", IEEE Journal of Quantum Electronics, vol. QE-33, pp. 1282-1287, 1997.

[3] S.N. Bagayev, A.E. Baklanov, V.P. Chebotayev, and A.S. Dychkov, "Superhigh resolution spectroscopy in methane with cold molecules", Rev. Roum. Phys., vol. 33, pp. 361-367, 1988.

[4] Ch. Chardonnet, F. Guernet, G. Charton, and Ch.J. Bordé, "Ultrahigh-resolution saturation spectroscopy using slow molecules in an external cell", Applied Physics B, vol. B59, pp. 333-343, 1994.

[5] P.E. Durand, G. Nogues, V. Bernard, A. Amy-Klein, and Ch. Chardonnet, "Slow-molecule detection in Doppler-free two-photon spectroscopy", Europhysics Letters, vol. 37, pp. 103-108, 1997.

[6] A. Fioretti, D. Comparat, A. Crubellier, O. Dulieu, F. Masnou-Seeuws, and P. Pillet , Phys. Rev. Lett., vol. 80, pp. 4402, 1999.

[7] N.F. Ramsey, "A molecular beam resonance method with separated oscillating fields", Phys. Rev., vol. 78, pp. 695-699, 1950.

[8] Ch.J. Bordé, C. Salomon, S. Avrillier, A.V. Lerberghe, Ch. Bréant, D. Bassi, and G. Scoles, "Optical Ramsey fringes with traveling waves", Phys. Rev. A, vol. 30, pp. 1836-1848, 1984.

[9] F. Herlemont, M. Lyszyk, and J. Lemaire, "Doppler-Free Two-Photon Spectroscopy of the 2v3 band of SF_6", Appl. Phys., vol. 24, pp. 369-374, 1981.

[10] A. Linskens, S. te Lintel Hekkert, and J. Reuss, "One and two photon spectra of SF_6 molecular beam measurements", Infrared Phys., vol. 32, pp. 259-282, 1991.

[11] F. Herlemont, M. Khelkhal, J. Legrand, and G. Pierre, "Doppler-free two-photon spectrum of SF6 for metrological purposes", Opt. Lett., vol. 23, pp. 957-959, 1998.

[12] Y.V. Baklanov, B.Y. Dubetsky, and V.P. Chebotayev, "Non-Linear Ramsey Resonance in the Optical Region", App. Phys., vol. 9, pp. 171-173, 1976.

[13] Ch.J. Bordé, "Sur les franges de Ramsey en spectroscopie sans élargissement Doppler", C. R. Acad. Sc. Paris, vol. 284B, pp. 101-104, 1977.

[14] Y.V. Baklanov, V.P. Chebotayev, and B.Y. Dubetsky, "The Resonance of Two-Photon Absorption in Separated Optical Fields", App. Phys., vol. 11, pp. 201-202, 1976.

[15] Ch.J. Bordé, in Advances in laser spectroscopy, Plenum, New York, 1983, pp. 1-70.

[16] A.F. Linskens, W.L. Meerts, J. Reuss, and C.J. Bordé, in Thesis, ch. 6, 1994, pp. 101-121.

[17] P.E. Durand, Spectroscopie à deux photons à ultra-haute résolution : sélection des molécules lentes ou franges de Ramsey, thèse de l'Université Paris 13, 1997.

[18] L.F. Constantin, R.J. Butcher, P.E. Durand, A. Amy-Klein, and Ch. Chardonnet, "2.3 kHz two-photon Ramsey fringes at 30 THz", Physical Review A, vol. 60, pp. 753-756, 1999.

[19] A. Amy-Klein, L.F. Constantin, R.J. Butcher, C. Daussy, P.E. Durand, G. Nogues, and Ch. Chardonnet, "Slow molecule detection or Ramsey fringes in two-photon spectroscopy : which is better for high resolution spectroscopy and metrology?", Optics Express, vol. 4, pp. 67-76, 1999.

[20] M.P. Sassi and K. Stoll, in Laser Spectroscopy XIII, World Scientific Publishing, 1997, pp. 111-113.

[21] E. Giacobino and B. Cagnac, in Progress in Optics XVII, North-Holland, 1980, ch. 2, pp. 85-161.

[22] Ch.J. Bordé, in Revue du Cethedec- Ondes et signal-NS 83-1, 1982, pp. 1-118.

Double Oven Quartz Crystal Oscillator Ultimate performances and applications

Ch. LONGET, Th. BLIN, F. LEFEBVRE

OSCILLOQUARTZ S.A
Brévards, 16
CH-2002 NEUCHATEL - SWITZERLAND
Tel : 00 41 32 722 55 07 Fax : 00 41 32 722 55 56
e-mail : osa@oscilloquartz.com

Abstract

A Double Oven Controlled Quartz Crystal Oscillator using an SC-Cut resonator is presented, focusing on the main parameters useful for system applications (either in locked or hold-over mode) like stability versus temperature (static and dynamic), load and power supply changes and aging, as well as the linearity of the frequency control.

We present results of such an oscillator built in an industrial low profile package (51x41x25 mm). This double oven oscillator, which is based on a new thermo-mechanical structure, choosen after extensive simulation and data correlation based on our own long-time production of highly stable quartz crystal oscillators, meets the following performances,:

- Thermal stability better than 2E-10 from 0 to 60°C.
- Aging in the range of few parts of 1E-11 per day.
- Low phase noise.
- Linearity < 10%

Also, absolute performances of a GPS synchronization module equipped with this oscillator is provided (time deviation during temperature variation and stability in hold-over mode).

1/ Oscillator design

The block diagram of this oscillator is shown on figure 1.

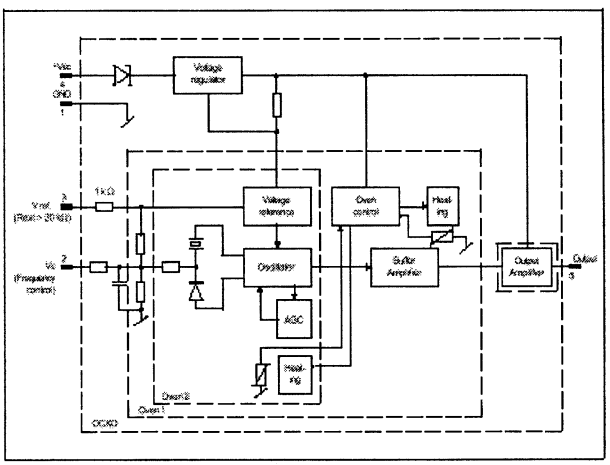


Figure 1

To produce oscillator for telecommunications and synchronization applications, the design must be a compromise between performances, size and cost. For this reason, this part is based on a classical oscillating loop, including an Automatic Control Gain. This allows to use a reasonable quantity of components to obtain high performances. A big effort was done on minimization of the sensitivities to load and power supply changes and stability of the internal voltage references, as well as the one used internally than the one provided for user's application.

In the same state of mind, the thermal control is made through a single loop managing the two ovens.

On figure 2, the thermo-mechanical shows the two ovens, the location of the resonator, glued at the bottom of the internal one to optimize thermal conductivity, and the upper PC board, working like the cap of the internal oven. As far as possible the most critical and sensible components are placed in this area, to prevent thermal effects.

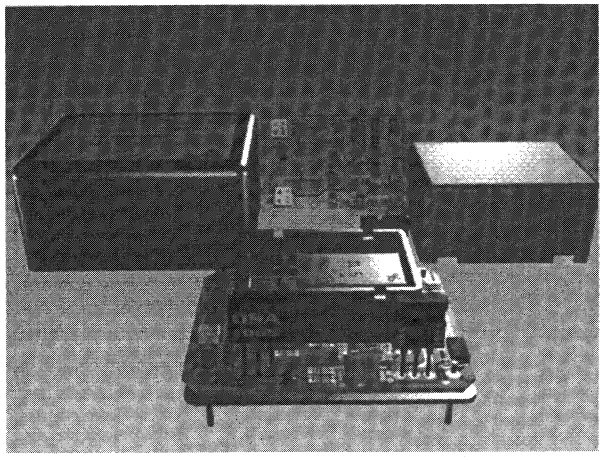


Figure 2

The external oven, powered by two heaters, recovers the heart and the whole is hermetically sealed by the final package.

This structure was defined thanks to the experience of Oscilloquartz in high stability OCXO, calculations and experimental results.
Calculations made with SPICE simulator using the Electrical-Thermal equivalence :

$$U = RxI$$

$$P = Rth \times T$$

allow to define a rough structure, on which impact of modified thermal resistivity are quickly obtained.

For better accuracy, both simulations with PRO-MECHANICA THERMAL or experimental results obtained by placing sensors in the oscillators are necessary.

Hereafter an example of the thermal gradient, measured using sensors and the B-mode of the resonator is shown.

The gradient measured inside the internal oven is better than 0.3°C between top and bottom and the variation of the bottom, where the resonator is placed, is only 0.1°C when the external temperature varies from 0 to 60°C.

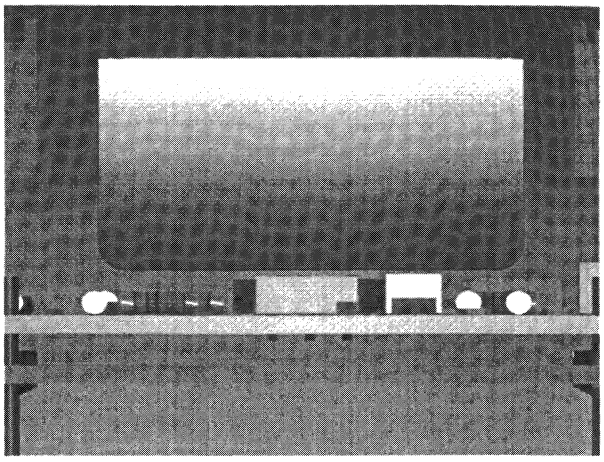


Figure 3

That means that if we can set the oven exactly at the value of the Turn-Over Point of the resonator, its influence is drastically reduced (in the range of E-11). If an offset of 0.2°C is introduced (which can be generated very easily by an error of only 2% on the determination of the TOP), then the resonator influence goes up in the range of E-10 (Figure 4)

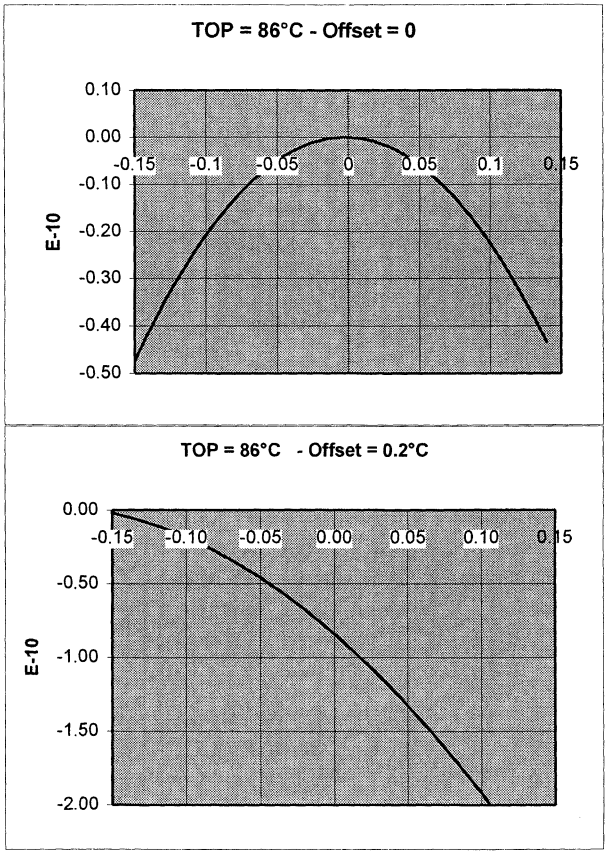


Figure 4

3/ Performances results

Thermal stability :

A typical figure of thermal stability obtained with this oscillator under normalized operating conditions is provided in Figure 5. Maximum deviation is 1.5E-10.

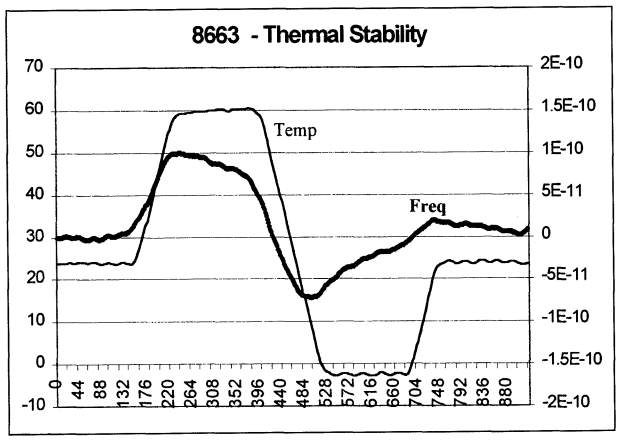


Figure 5

The right way to design is confirmed by results obtained with a single oven oscillator (8743 model). In that case the maximum deviation is 2.5E-9 over 0-70°C

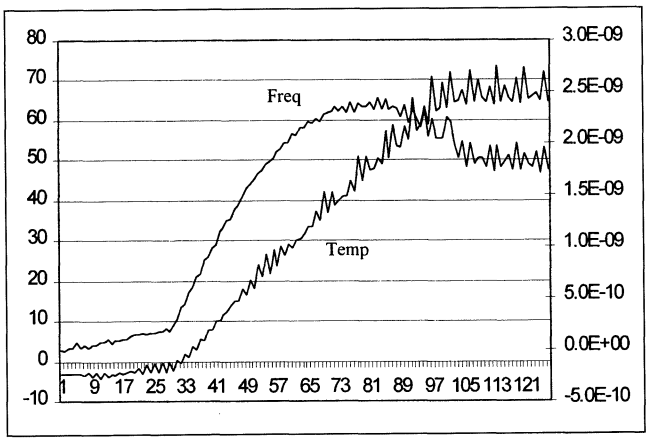


Figure 6

Linearity :

Uses in a system means that the oscillator will work in a locked loop in which a D/A converter will give it voltage step to control and adjust the final frequency. If the frequency voltage slope is huge and highly non linear, this will cause degradations of the long term tracking and noise generation.

This oscillator is designed in order to meet this requirement by providing a linearity of the Frequency Voltage Control better than 2 %. (Figure 7).

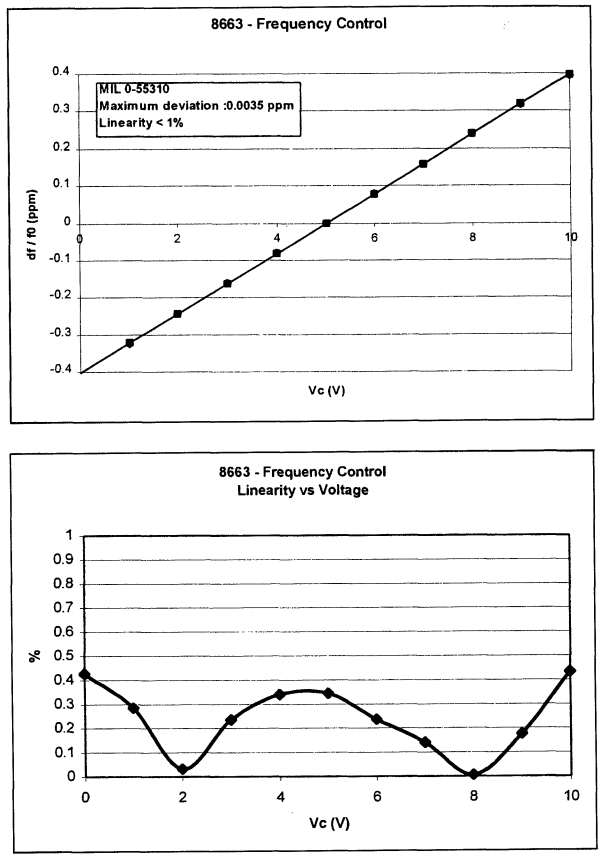
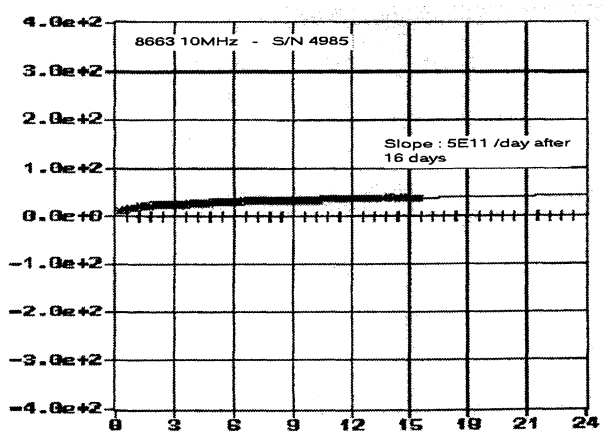


Figure 7

Aging :

Low linearity can be achieved if the frequency excursion is reduced, therefore the aging must be low also. Long term frequency variations in the range of few parts of E-11 are achievable in a very short operating time (typically less than 20 days).



Phase Noise :

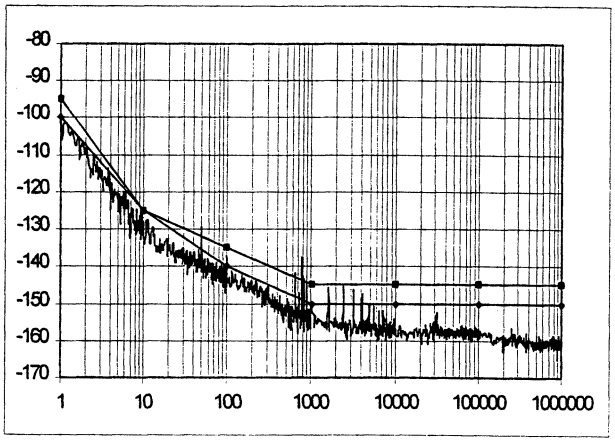


Figure 9

Standard Phase noise figure is –95dBc @1Hz and –145 dBc @100kHz.
A low phase noise option (-100dBc @1Hz and –150dBc @100kHz) is also available with good production yield.

5 / Conclusion

The best conclusion to proove the influence of such an oscillator on the final performances of a system is to make a comparison between two models.

A typical application : The 4510 GPS Module

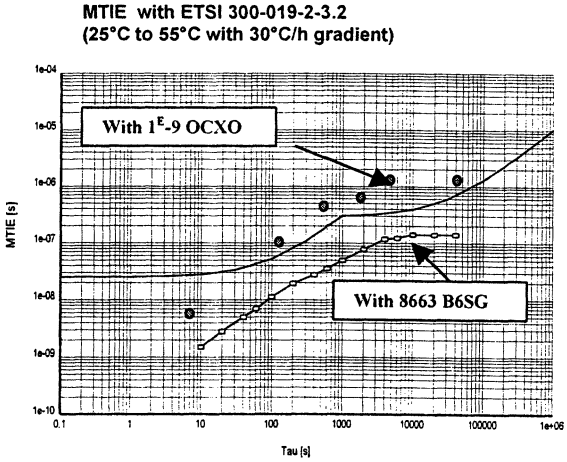


Figure 10

On figure 10, we can see the MTIE (Maximum Time Interval Error) of a GPS Module 4510 using a 8663 B6SG oscillator (Thermal stability = 6E-10), compared to the template of a Level 1 System.

Performance is largely better than the specification.
Moreover we have plotted a result obtained on a similar system using an OCXO with a thermal stability around 1E-9.
This system does not meet the specification.

Acknowledgments

The authors thank Mr Weber and Mr Cosandier, who are the OSA 4510 project headers.

HIGHLY MINIATURIZED SAW FILTERS FOR CDMA MOBILE-PHONES

Ulrich Bauernschmitt, Stefan Freisleben, Andreas Bergmann, Joachim Gerster, Thomas Johannes* and Clemens Ruppel*

EPCOS AG, Munich
* Siemens Corporate Research, Munich

Today's market for mobile communication products with the demand for small-size and high-performance products is putting a high miniaturization pressure on all components invoked in the functionality of a mobile phone. As the analogue front-end of multi-mode CDMA systems is realised using a heterodyne concept, channel filtering and blocking of interfering signals is performed using SAW filters in the IF stage. This papers describes prerequisites for the design of three types of highly miniaturised filter solutions invoked in a multi-band/multi-mode CDMA handheld.

The IS-95 CDMA system in the PCS-band calls for a channel bandwidth of 1.25 MHz as well as for supression of adjacent channels at f_c+-1.25 MHz. We will present an extension of the classical Z-path concept in order to realize an IF-filter for CDMA-1900 at 210.38 MHz with a 5 dB bandwidth of 1.5 MHz on Quartz in a 5x5 mm^2 ceramic package. In the last couple of years Z-path SAW filters have been widely used as IF filters in mobile phones, especially for GSM systems at IF frequencies below 100 MHz. This design technique offers a great potential for size reduction compared to inline structures because the acoustical path is folded. The last SAW IF filter generation for CDMA-1900 realized as inline structures was typically mounted into a 13.3x6.5 mm^2 package. In order to achieve the tremendous size reduction of 70 % the Z-path concept has been combined with the RSPUDT design techniques resulting in a recursive Z-path filter. Both transducers have weighted excitation and reflection functions. The latter exhibit phase reversal thus forming longitudinally coupled resonant cavities leading to steep skirts.As an example a filter having a typical insertion attenuation of 8.2 dB and a selectivity at f_c+-1.25 MHz of typically 39 dB will be presented. The rms deviation from linear phase amounts to 2.3 degree typically in a frequency range of f_c+-0.62 MHz. It will be shown that the theoretical and experimental results are in good agreement.
For a CDMA system in the AMPS-band requirements are more stringent, as very high supression already at f_c+-0.9 MHz is requested. It will be shown how with the Z-path technique the transfer function can be shaped by using the additional selectivity provided by the inclined reflector pair for high ultimate as well as for increased close-in rejection rsp. steepness of the filter's skirts.As an example we will present a filter at a center frequency of 110.0 MHz housed in a 13.3x6.5 mm^2 ceramic package which shows an insertion attenuation of less than 8 dB and a selectivity of more than 38 dB at f_c+-0.9 MHz.
In addition to the digital transmission also analog AMPS with a 30 kHz wide FM channel is supported in CDMA-800 phones. This requires a very narrowband IF filter which provides sufficient selectivity at f_c+-60 kHz for the supression of the alternate channels. We will show a design approach using the technique of transversly coupled resonances (TCF) which enables us to realize the requirements with a filter at 85.38 MHz in a 9x5 mm^2 ceramic package. The underlying method of computation as well as a comparison with experiments will be shown.

DOUBLE OVEN CRYSTAL OSCILLATOR (DCXO) WITH TEMPERATURE COEFFICIENT <1 X 10^{-10} -55° C TO 70° C

MARTIN BLOCH, JOHN FELLNER, JOHN HO, CHARLES STONE

Frequency Electronics, Inc.

A new oscillator designed to meet the requirements of today's CDMA base station will be presented. The design consists of three basic parts, the oscillator, frequency synthesizer and thermal controller. Computer simulations revealed new insights into the temperature characteristics of crystal oscillators and patent pending designs for the frequency synthesizer and thermal control architecture will be presented.

The DCXO architecture is based on an oscillator designed primarily for the lowest and most predictable temperature coefficient possible. Classically, initial frequency calibration and adjustment for aging is accomplished by changing the value of a reactive element placed in series with the crystal. This requires the crystal frequency to be within its tuning range and adds temperature sensitive components such as inductors and varactor diodes to the oscillator circuit. A frequency synthesizer was designed to perform the frequency calibration and aging adjustments. The oscillator and crystal subsequently were designed for the optimal temperature stability, phase noise and aging performance. A 5MHz 5th overtone crystal was designed for low aging, high stability and suppression of the 3rd overtone C and B modes thus further simplifying the oscillator design.

Direct synthesizers (DDS) have some major limitations such as limited frequency resolution and a highly spurious output spectrum. With a unique architecture a PLL was combined with a 32 bit DDS resulting in a frequency synthesizer with resolution less than 3 x 10^{-12} and an output spectrum free of spurious signals.

The thermal control consists of a double oven which has thermal gain greater than 100,000. This is accomplished with a zero temperature gradient outer oven. The zero gradient oven uses a system of satellite heating semiconductors controlled by a main controller. The heat dissipation of the satellite heaters and the controller is optimized over the outer oven surface in such a way as to minimize the temperature change of the crystal to less than 0.001° C over the ambient temperature range of –55° C to 70° C.

Martin B. Bloch
Frequency Electronics, Inc.
55 Charles Lindbergh Blvd.
Uniondale, NY 11553

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

NONLINEAR DIPOLAR MODELLING OF QUARTZ CRYSTAL OSCILLATORS

R. Brendel, D. Gillet, N. Ratier, F. Lardet-Vieudrin, J. Delporte*

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.edu
*Centre National d'Etudes Spatiales - 18, avenue Edouard Belin - 31055 Toulouse Cedex - France

ABSTRACT

A quartz crystal oscillator can be thought of as a resonator connected across an amplifier considered as a nonlinear dipole the impedance of which depends on the amplitude of the current that gets through it.

The nonlinear amplifier resistance and reactance are obtained by replacing the resonator with a sinusoidal current source of same frequency and performing a set of transient analyses giving the current source a larger and larger magnitude.

A Fourier analysis performed on the steady state voltage across the dipolar amplifier allows us to calculate both real and imaginary parts of the dipolar impedance as a function of the current amplitude.

From these curves, it is then possible to accurately calculate the oscillation amplitude and frequency without having to perform unacceptably long transient analysis needed by a direct oscillator closed loop simulation.

1. INTRODUCTION

For many years a large number of analytical and numerical methods have been developed to predict the behaviour of quartz crystal oscillators with higher and higher accuracy. Because of the very high resonator's quality factor, a number of numerical methods fail to converge in a reasonable computer time or lack accuracy.

In the method presented here, the oscillator is considered as a resonator connected across an amplifier that behaves like a dipole whose impedance is evaluated at the resonator's frequency. This method is not new and has been widely used in the past [1, 2]. Nevertheless, when used together with an electrical simulator like SPICE it should become an accurate and powerful tool for the quartz crystal oscillator design.

To describe the nonlinear behaviour of the dipolar amplifier, the resonator is replaced by a sinusoidal current source of a given amplitude at the oscillator resonant frequency and a transient analysis is performed by using an electrical simulator like SPICE. Real and imaginary parts of the dipolar impedance for the actual current amplitude are then obtained by performing a Fourier analysis on the steady state voltage across the dipole. Nonlinear amplifier resistance and reactance are obtained by giving the sinusoidal current source a larger and larger amplitude.

Afterwards, oscillation condition reduces to state that when steady state is reached, the resonator impedance is exactly equal and of opposite sign to the amplifier impedance. The real part of this complex identity gives the oscillation amplitudes while the imaginary part gives the frequency shift with respect to the resonator frequency.

This method is quite simple to implement in a simulation program and can be used to quickly and accurately obtain steady state oscillation conditions, amplitude and frequency transients, oscillator sensitivity to various parameters (component value, temperature, supply voltage, ...) amplitude and phase noise spectra, etc.

It should be emphasized that the oscillator behaviour is fully described by only analyzing the amplifier impedance thus avoiding unacceptably long transient analysis met with the oscillator closed loop simulation.

Eventually, by using slowly varying functions method, amplitude and frequency transients can be easily obtained.

2. DIPOLAR REPRESENTATION

The method used can be summarized by looking at the oscillator represented in Fig. 1a. The resonator is regarded as an impedance strongly varying with the frequency and slightly varying with amplitude (because of the isochronism defect) while the amplifier is considered to be a nonlinear dipole the impedance of which strongly depends on the current amplitude and weakly varying with frequency.

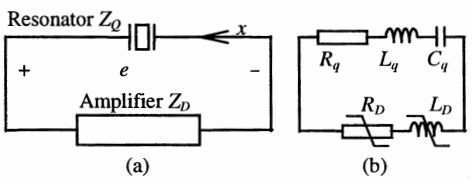


Fig. 1. *Dipolar representation of an oscillator (a) and resonator and amplifier impedance (b)*

It results from these properties that the oscillator frequency is mainly fixed by the resonant frequency of the resonator while the amplitude is determined by the amplifier nonlinearities.

The representation used assumes that the resonator is reduced to its series resonant branch (L_q, C_q, R_q) (Fig. 1b) while other elements like parallel capacitance or pulling capacitance are included in the amplifier dipole. The amplifier dipole (Fig. 1b) can be represented by a nonlinear resistance in series with a nonlinear

inductance that vary with the amplitude of the current and may have negative or positive value.

As it will be shown later, when using high-Q series resonant circuit such as quartz crystal resonator, the loop current x, in the oscillator is almost perfectly sinusoidal while the voltage across the resonator may have a large harmonic distortion. This feature leads to analyze the circuit assuming that the main variable is the loop current x.

3. OSCILLATION CONDITION

When oscillation occurs in the circuit represented in Fig. 1a, the voltage e across the resonator is identical to the one across the amplifier. Calling x the loop current, the oscillation condition is straightforward

$$e = Z_D x = -Z_Q x \;\Rightarrow\; (Z_D + Z_Q)\,x = 0 \qquad (1)$$

The impedance Z_Q of the resonator represented in Fig. 1b can be written as

$$Z_Q = \frac{L_q}{s}\left(s^2 + \frac{R_q}{L_q}s + \omega_q^2\right) \qquad (2)$$

While the impedance Z_D of the amplifier is

$$Z_D = R_D + sL_D \qquad (3)$$

where s is the Laplace's operator. Replacing (2) and (3) in (1) leads to the oscillation equation

$$s^2 + \frac{1}{L_q}(R_q + R_D)s + \omega_q^2\left(1 - \frac{L_D}{L_q}\right) = 0 \qquad (4)$$

where

$$\omega_q^2 = \frac{1}{L_q C_q} \qquad (5)$$

is the series resonant frequency of the resonator. It should be recalled that R_D and L_D depend on the amplitude y of the loop current x.

4. STARTING CONDITION AND STEADY STATE

Equation (4) looks like a second order differential operator that may produce increasing amplitude solution only if the first order term have a negative value. If R_{DS} is the value of the nonlinear dipolar resistance at very low current amplitude ($y \approx 0$), starting condition reduces to

$$R_q + R_{DS} < 0 \qquad (6)$$

As the oscillation amplitude increases, the dipolar resistance increases so that the value of the first order coefficient increases, the steady state is reached when this coefficient becomes null, thus the steady state dipolar resistance $R_D(y_0)$ is given by

$$R_q + R_D(y_0) = 0 \qquad (7)$$

where y_0 is the steady state amplitude of the oscillation. Eq. (4) also gives the steady state frequency of the oscillation that takes the form

$$\omega_0^2 = \omega_q^2\left(1 - \frac{L_D(y_0)}{Lq}\right) \qquad (8)$$

where $L_D(y_0)$ is the value of the steady state dipolar inductance.

As we will see in a next section, Eqs. (6), (7) and (8) give a simple way to determine the starting condition and steady state amplitude and frequency as soon as the dipolar resistance and inductance are expressed as nonlinear function of the current amplitude.

5. AMPLIFIER DIPOLAR IMPEDANCE

Except in a limited number of cases, it is not possible to derive an analytical expression of the nonlinear functions $R_D(y)$ and $L_D(y)$ from the analysis of the amplifier. On the other hand, electrical simulator like SPICE can be used to accurately obtain these functions from a set of transient analyses.

To this end, the resonator in Fig. 1a, is merely replaced by a sinusoidal current source of frequency close to the resonator's frequency ω_q (Fig. 2).

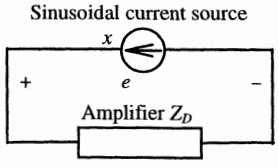


Fig. 2. *Amplifier dipolar impedance characterization*

By giving the amplitude y several values a set of transient analyses with sufficient duration to reach the steady state are performed. Then, by performing a Fourier analysis on the voltage e across the amplifier dipole, it is possible to calculate both real and imaginary parts of the dipolar impedance and derive both dipolar resistance and inductance of the amplifier as a function of the current amplitude y.

6. DEMONSTRATING EXAMPLES

6.1 Van der Pol's oscillator

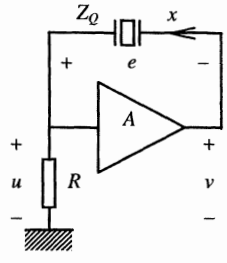


Fig. 3. *Van der Pol Oscillator*

Fig. 3 shows a simple behavioural Van der Pol oscillator that can be analyzed either theoretically or by using electrical simulator.

In this example, the amplifier has a nonlinear gain that is taken under the form :

$$v = Au(1 - \varepsilon u^2) \qquad (10)$$

Replacing the resonator by a sinusoidal current source x of amplitude y and calculating the dipolar impedance at the fundamental frequency we obtain

$$Z_D = (1-A)R + \frac{3A\varepsilon R^3}{4} y^2 \qquad (11)$$

This impedance has no imaginary part so that it is reduced to a nonlinear resistance that is a quadratic function of the current amplitude y (Fig. 4).

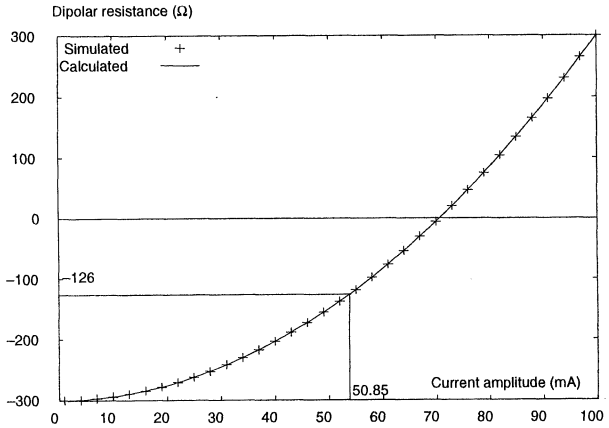


Fig. 4. *Dipolar resistance of the Van der Pol's amplifier obtained by analytical method (solid line) and by simulation (+) using:* $A = 4$, $\varepsilon = 0.02$ V^{-2}, $R = 100\ \Omega$, $R_q = 126\ \Omega$

Starting condition (6) becomes here

$$R_{DS} = (1-A)R < -R_q \qquad (12)$$

Eq. (12) shows that the amplifier must have an initial negative resistance of magnitude larger than the resonator's series resistance for the oscillation to occur.

Steady state amplitude is given by Eq. (7) that has the form

$$(1-A)R + \frac{3A\varepsilon R^3}{4} y_0^2 = -R_q \qquad (13)$$

so that

$$y_0^2 = \frac{4R_M}{3A\varepsilon R^3} \qquad (14)$$

where

$$R_M = -R_{DS} - R_q = (A-1)R - R_q \qquad (15)$$

should be called the "resistance margin" of the oscillator (Fig. 4). Note that since $L_D = 0$ the oscillation frequency is only fixed by the resonator (Eq. 8).

It can be shown in Fig. 4 that the value of R_D obtained by using the simulation procedure described in section 5 gives the same result as obtained by analytical method.

6.2 Transconductance oscillator

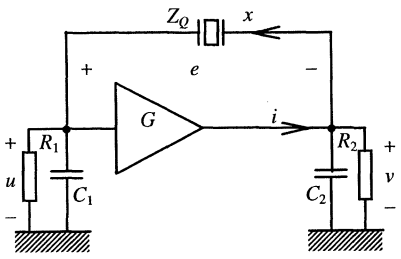


Fig. 5. *Transconductance oscillator*

Fig. 5 represents another simple example of behavioural transconductance oscillator.

Taking the nonlinear transconductance of the amplifier under the form

$$i = -Gu(1-\varepsilon u^2) \qquad (16)$$

the dipolar impedance $Z_D = R_D + jL_D\omega_q$ takes the form

$$R_D = \frac{R}{1+\alpha^2}\left[2 + \frac{1-\alpha^2}{1+\alpha^2} RG\left(1 - \frac{3\varepsilon R^2 y^2}{4(1+\alpha^2)}\right)\right] \qquad (17)$$

$$L_D\omega_q = \frac{-2R\alpha}{1+\alpha^2}\left[1 + \frac{RG}{1+\alpha^2}\left(1 - \frac{3\varepsilon R^2 y^2}{4(1+\alpha^2)}\right)\right] \qquad (18)$$

with $\alpha = \omega_q RC$. This expression is obtained by giving the two capacitors C_1 and C_2 the same capacitance C and the two resistances R_1 and R_2 the same resistance R.

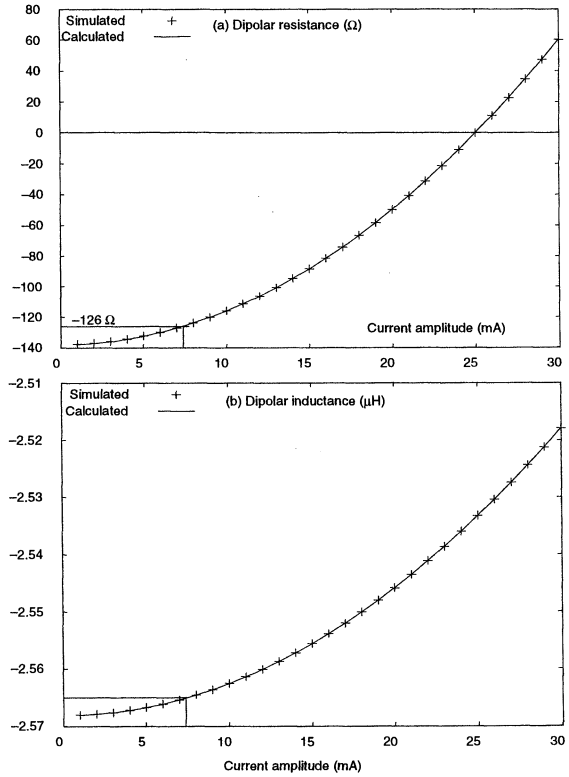


Fig. 6. *Dipolar resistance (a) and dipolar inductance (b) of the transductance amplifier using:* $G = 22$ mA/V, $\varepsilon = 1/3$, $C = 200$ pF, $R = 10$ kΩ, $R_q = 126\ \Omega$

Here the dipolar resistance and the dipolar inductance are both quadratic function of the current amplitude y (Fig. 6).

Eq. (7) gives the steady state amplitude, here we have

$$y_o^2 = \frac{-4(1+\alpha^2)^3 R_M}{3\varepsilon G R^4 (1-\alpha^2)} \qquad (19)$$

where the resistance margin R_M is defined by

$$R_M = -R_{DS} - R_q = -GR^2 \frac{(1-\alpha^2)}{(1+\alpha^2)^2} - \frac{2R}{(1+\alpha^2)} - R_q \qquad (20)$$

As y_0 is determined by the resistance condition (7), the actual value of the dipolar inductance can be

obtained by giving the current amplitude y the value y_0 in the imaginary part of the dipolar impedance (18)

$$\omega_q L_d(y_0) = \frac{-2\alpha R}{(1+\alpha^2)}\left[1+\frac{RG}{(1+\alpha^2)}+\frac{R_M(1+\alpha^2)}{R(1-\alpha^2)}\right] \quad (21)$$

In most cases $\left|\frac{L_d(y_0)}{L_q}\right| << 1$ so that relative frequency shift with respect to the resonator frequency can be obtained from Eq. (8) under the form

$$\frac{\Delta f}{f_q} = -\frac{L_d(y_0)}{2L_q} \quad (22)$$

In the example shown in Figs. 5 and 6, Eqs. (19) and (22) give $y_0 = 7.38\,\text{mA}$ and $\frac{\Delta f}{f_q} = 1.2826\,10^{-3}$, these results are quite close from those obtained by using simulated curves (Fig. 6).

7. AMPLITUDE AND FREQUENCY TRANSIENTS

Substituting the differential operator $\frac{d}{dt}$ for the Laplace's variable s in (4) and applying the operator to the loop current x leads to a nonlinear second order differential equation that takes the form

$$\ddot{x}+\frac{1}{L_q}(R_q+R_d)\dot{x}+\omega_q^2\left(1-\frac{L_d}{L_q}\right)x=0 \quad (23)$$

Because the loop current is a quasi sinusoidal function, Eq. (23) can be considered as a linear second order differential equation with a small nonlinear perturbation term in the right hand side

$$\ddot{x}+\omega_q^2 x=-\frac{1}{L_q}(R_q+R_d)\dot{x}+\frac{\omega_q^2 L_d}{L_q}x \quad (24)$$

The oscillation amplitude and frequency offset can be obtained by using the variable change

$$x = y\cos(\omega_q t+\varphi) \quad (25)$$

where the amplitude y and the phase φ are slowly varying function of time. By using averaging method described in [5], Eq. (24) can be replaced by the approximate associated system [3, 4]

$$\begin{cases} \dot{y} = \dfrac{-y}{2L_q}\left(R_q+R_d(y)\right) \\ v = \dot{\varphi} = -\dfrac{L_d(y)\,\omega_q}{2L_q} \end{cases} \quad (26)$$

Where $R_d(y)$ and $L_d(y)$ are the previously obtained nonlinear functions of the current amplitude y. Solving system (26) by numerical method allows us to quickly obtain the amplitude and frequency transients. Fig. 7 shows the result obtained with the transconductance oscillator described in Fig. 5.

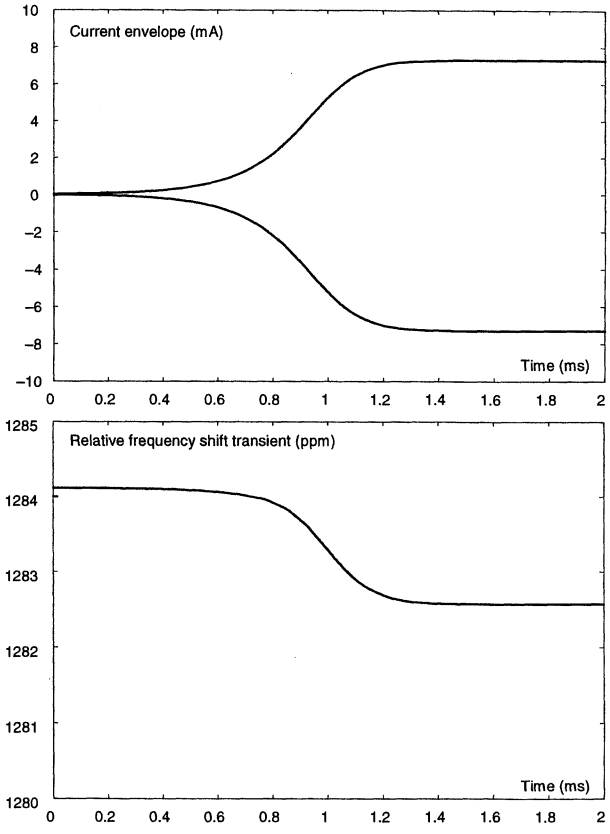


Fig. 7. *Amplitude and frequency transients of the transconductance oscillator*

It should be noted that the steady state amplitude and frequency shift are in perfect agreement with those obtained in the previous section.

To compare results obtained by the dipolar impedance method and by the direct closed loop method, the resonator has been given a rather low quality factor ($Q \approx 500$), such a weak Q-factor explains the quite large frequency shift obtained.

8. COLPITTS OSCILLATOR

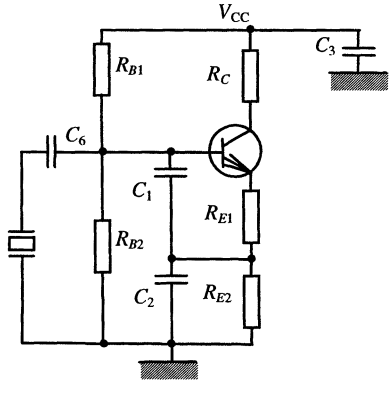


Fig. 8. *Colpitts oscillator*

Fig. 8 shows a more realistic example of Colpitts oscillator. By using the same procedure as described in section 5, the nonlinear resistance and reactance are obtained (Fig. 9) so that the steady state amplitude and

relative frequency shift can be calculated using Eqs. (7) and (22) respectively. In the present case we obtain $y_0 = 552\ \mu A$ and $\frac{\Delta f}{f} = 3475\,\text{ppm}$.

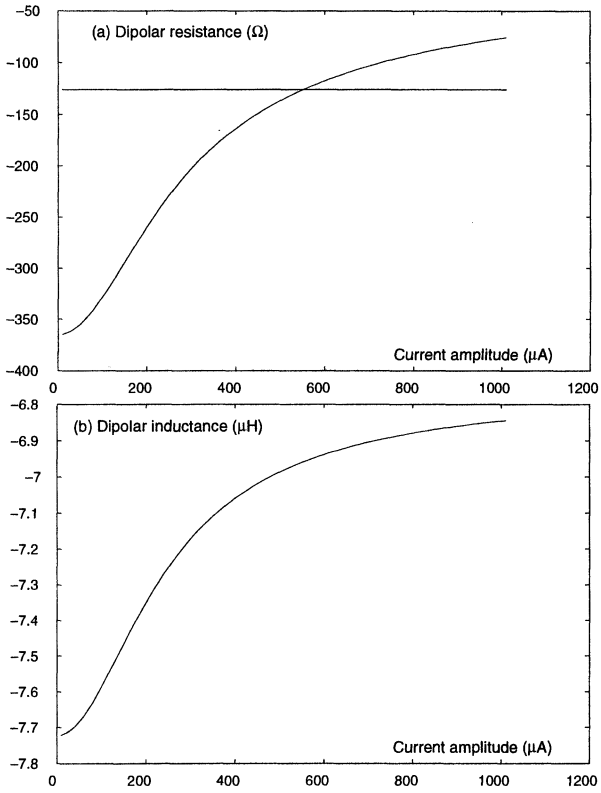


Fig. 9. *Dipolar resistance (a) and dipolar inductance (b) of the Colpitts oscillator*

By solving the associated system (26), the amplitude and frequency transients are easily obtained as shown in Fig. 10. Note that the steady state amplitude and frequency shift are identical to those obtained by using the nonlinear impedance method. As in the previous example, the large frequency shift is due to the low Q-factor of the resonator used.

9. CONCLUSION

By using electrical simulation program like SPICE, it is possible to describe the nonlinear behaviour of the oscillator amplifier by looking at it as a nonlinear dipolar impedance whose real and imaginary parts can be obtained by performing a set of transient analyses.

Applying simple oscillation condition, both steady state amplitude and frequency can be easily and accurately obtained without having to simulate a time consuming high-Q circuit.

In addition oscillation amplitude and frequency transients can be quickly obtained by using slowly varying function method

A perturbation method analog to the one described in [5] can be easily implemented to obtain amplitude and phase noise spectra.

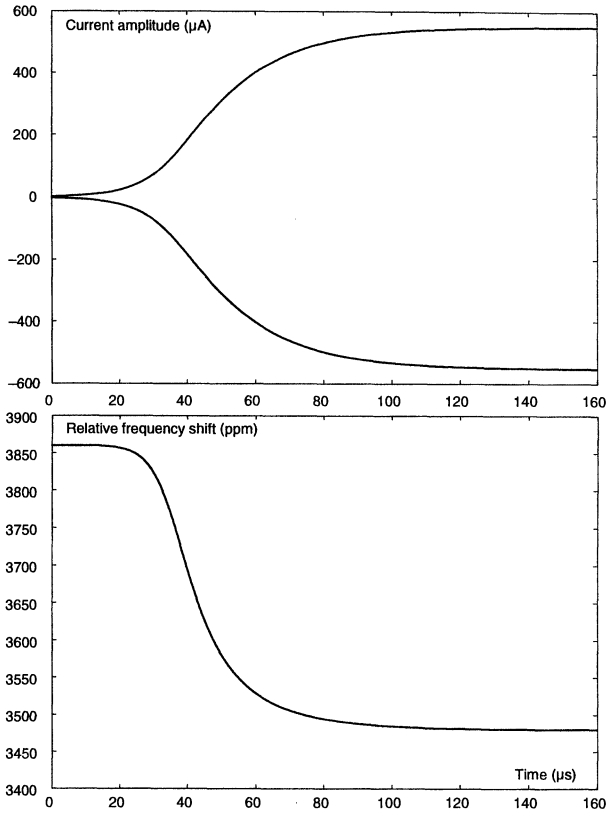


Fig. 10. *Amplitude and frequency transients of the Colpitts oscillator*

It should be emphasized that all the steps of the method presented in this paper are currently being automated so as to develop an efficient and powerful oscillator CAD tool.

ACKNOWLEDGEMENTS

This work has been supported by CNES under contract #714/CNES/99/7671/00.

REFERENCES

[1] E.A. Vittoz, "Quartz Oscillators for Watches", in Proceedings of the 10th International Congress of Chronometry, Geneva, 1979.

[2] B. Parzen, A. Ballato, Design of Crystal and other Harmonic Oscillators, Wiley Intersciences, 1983.

[3] N. Kryloff and N.N. Bogoliuboff, Introduction to Nonlinear Mechanics, Princeton University Press, 1943.

[4] A.A. Andronov, A.A. Vitt and S.E. Khaikin, Theory of Oscillators, New York, Pergamon Press, 1966.

[5] R. Brendel, N. Ratier, L. Couteleau, G. Marianneau and P. Guillemot, "Slowly Varying Functions Method applied to Quartz Crystal Oscillator Transient Calculation", IEEE trans. on UFFC, Vol. 45, No 2, March 1998.

Optical Frequency Synthesis: An update on the UWA/PTB Frequency chain.

A. N. LUITEN, R.P. KOVACICH AND J.J. MCFERRAN
Department of Physics
University of Western Australia
Nedlands, W A 6907, Australia

ORAL PRESENTATION PREFERRED

The University of Western Australia (UWA) and the Physikalisch-Technische Bundesanstalt (PTB) are in the process of developing an absolute optical frequency synthesis chain for the purpose of interrogating the electronic structure of laser-cooled atoms[1]. The frequency chain is based upon the modern approach of frequency-interval bisection[2], combined with an optical frequency comb generated using a mode-locked laser[3], and has a goal resolution of 1 part in 10^{15} over 1 second of integration time. This resolution is only possible because of the extremely high spectral purity, and frequency stability of the UWA-developed sapphire microwave oscillator[4], combined with high bandwidth optical phase-locked loops at each stage of the chain. Although the frequency-interval division approach to synthesis has many advantages over traditional designs, it has potential pitfalls arising from the necessity to generate non-linear signals from low power continuous-wave outputs of laser diodes, the necessity to generate radiation in places that are not necessarily accessible to diode laser technology, and the need for high quality optical phase-locking between weak non-linear signals. This presentation will address each of these difficulties in turn, and present methods for overcoming them. In addition, we will particularly address the operational difficulties that arise because of the need for a high bandwidth linkage between the optical and microwave domains.

The UWA/PTB design commences with the frequency interval that exists between a signal at 1064nm and its second harmonic: this initial wavelength was chosen as it coincides with an existing optical frequency standard at UWA. The interval between 1064 and 532nm is then bisected by three interval bisection stages so as to generate two signals, which are phase-coherent with the 1064nm signal, but lie at 946nm and 851nm. This residual interval will then be bridged using a compact diode-pumped Cr^{3+}:$LiSrAlF_6$ mode-locked laser. One of the modes of the optical comb will be phase-locked to the 946nm signal, while the repetition rate of the laser is simultaneously phase-locked to the microwave reference source output. The phase of the beat-note formed between the 851nm signal and the closest mode of the comb will be stabilised with respect to a radio-frequency reference oscillator, by adjustment of the phase of the 1064nm optical source. This ensures that the 1064nm source is phase-coherent with the microwave reference if each of the bisection stages, and the mode-locked laser, have sufficiently low noise. In this presentation we will introduce our design, show that it is capable of a synthesis resolution in the 10^{-15} range, and describe our progress towards completion of the chain. In particular, we will present experimental measurements of the propagation of frequency fluctuations through a cascade of divider stages, show generation of narrow-linewidth diode-laser radiation for each wavelength needed in the chain design, and demonstrate that sufficient power is generated in each non-linear process to enable high quality optical phase-locking. Tests of a divider unit show that phase-locking of the non-linear signals is stable for periods that are only limited by mode-hopping in the master lasers (~ 12 hours), with residual phase-fluctuations of order $10^{-8}rad^2/Hz$ inside the bandwidth of the loop (0.5MHz) when just ~50nW falls on the photo-detector.

1 A Luiten, R Kovacich, and J McFerran, IEEE Trans. Instrum. Meas. **48**, 558-562 (1999).
2 HR Telle, D Meschede, and TW Hänsch, Opt. Lett. **15**, 532-534 (1990).
3 T Udem, J Reichert, R Holzwarth, and TW Hänsch, Opt. Lett. **24**, 881-882 (1999).
4 AN Luiten, AG Mann, et al, in *Proc. 49th Freq. Control Symp. (FCS)*, San Fransisco, CA, 1995), p. 433-437.

DETECTION AND MIXING OF VISIBLE AND INFRARED RADIATION BY InP AND InSb SCHOTTKY BARRIER DIODES

A. Bertolini[1], N. Beverini[1], G. Carelli[1], S. Chepurov[2], E. Maccioni[1], A. Moretti[1] and F. Strumia[1]
[1] - Dipartimento di Fisica dell'Università di Pisa, GNSM and INFM
[2] - Institute of Laser Physics, Siberian Division of Russian Academy of Sciences

Giorgio Carelli, Dipartimento di Fisica dell'Università di Pisa, via Buonarroti 2, I-56127 Pisa, Italy
Tel. +39—050–844 285, Fax. +39—050–844 333,e-mail carelli@difi.unipi.it

ABSTRACT

We present the results of the experimental investigation of the detection and mixing properties of Schottky barrier diodes in the visible and near infrared regions. We used diodes with different semiconductor substrates, GaAs, InP and InSb, and different contact geometries. We used the diodes in the regions 458 – 633 nm and 1.3–1.5 μm. The video signals from the different devices were measured together with the beat notes from different modes of the visible lasers and high order harmonics of a microwave radiation. In the visible the GaAs-diode proved to be a very effective detector but has a very low frequency conversion efficiency and it generated no beat signal below 510 nm. On the contrary, InP and InSb-SBD operate very effectively as mixers, even with relatively low detection efficiencies, up to the blue. In the region 1.3 –1.5 μm the GaAs diode does not work while InP and InSb continue to be effective.

1. INTRODUCTION

In the past few years Schottky barrier diodes (SBD) have raised great interest due to their advantages over other devices in heterodyne detection in a very large wavelengths range. These diodes operate at room temperature and present a very high operation speed and a much higher stability than conventional Metal-Insulator-Metal (MIM) diodes. For these reasons they have been used extensively for nonlinear frequency mixing in the MM, SMM (FIR), IR ranges, [1-6]. The extension of these operations to the visible region has a large metrological interest, but only a few works dealing with MIM and Schottky diodes in the visible have been published up to now [7-9].
A mixing signal up to 1 THz has been observed with MIM- diodes as a difference of two green laser lines [7] or two dye lasers [9]. The signal-to noise ratio however was quite poor with a large and unpredictable spread of performance between different diodes. Point-contact Schottky diodes have demonstrated more efficiency and more reliability [8]. Recently an attempt was made to utilize the silver plasmon resonance to enhance the mixing signals of a point contact device for visible radiation at 633 nm from an He-Ne laser. [10], but till now there are no significant improvements over earlier devices.

In the present paper we report the results of an experimental investigation on the mixing and detection properties of point-contact diodes at visible frequencies by using as semiconductor substrates new materials (InP and InSb) and comparing their performances to those of conventional GaAs diodes. Our results demonstrate a better behaviour of these devices as detectors and mixers for visible light with respect to conventional MIM diodes and prove a dramatic enhancement in the signal to noise ratio with respect to GaAs Schottky diodes when used as mixers; these results make them very attractive for use in spectroscopy, laser frequency measurements and metrology.

2. Experiment and results

The investigated diodes presented two types of structure: honeycomb and point contact. GaAs and InP diodes were produced in the Semiconductor Institute (Tomsk, Russia), in collaboration with the Institute of Laser Physics (Novosibirsk, Russia) with an honeycomb structure with window diameters of 1.65 μm and 1.5 μm, respectively [1, 5, 6]. InSb and InP diodes were assembled in the Physics Department of the Pisa University with a classic point contact structure with a large semiconductor platelet as the post [11]. The contacts to the semiconductor posts are made by a tungsten whisker, electrochemically etched to the desired tip radius by means of a 2N solution of NaOH.

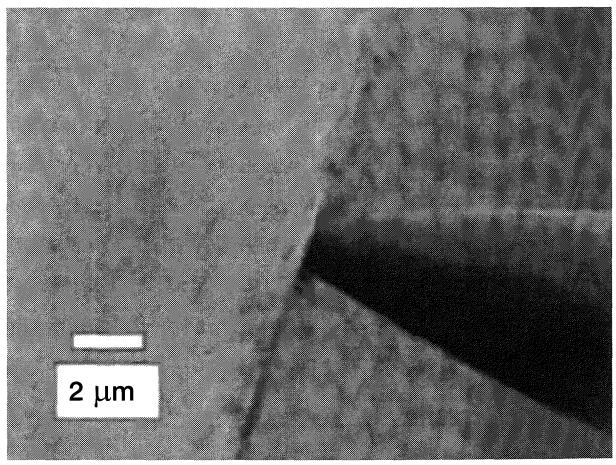


Fig. 1 Photo of the InSb –W junction

We used wires of 30 μm in diameter for the honeycomb diodes while a 127 μm wire was used for the point contact diodes. The curvature radii of the tips were around 50-60 nm, controlled together with the geometry of the contacts by a Scanning Electron Microscope. The InSb crystals have a carrier density of about 10^{14}, InP of 10^{12}, at 77 K. InSb proved to be a soft material and the tip just penetrate the semiconductor surface, in the case of InP, a material harder than InSb, a gold layer was deposed on the surface in order to ensure stable and ripetitive contacts.. A gold coated crystal shows a better video signal in comparison to an uncoated one while the mixing signal does not show signative differences.

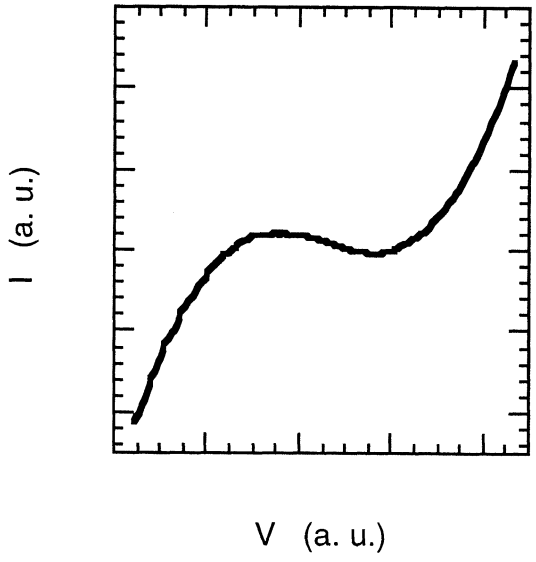


Fig. 2 Non-linear part of the I-V characteristic of an InSb Scottky diode

The I-V characterstic of the different device are completely different and InSb is the only one showing a symmetric one.

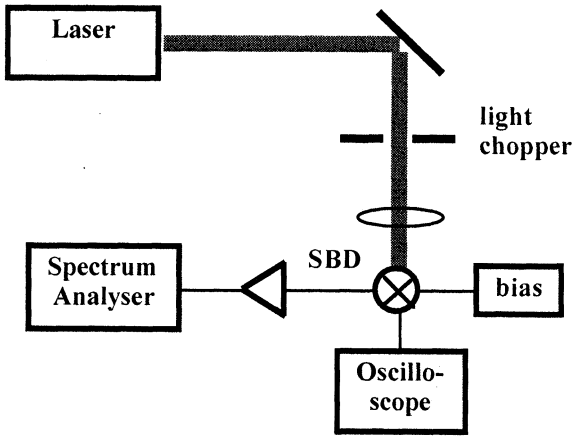


Fig. 3. The experimental setup.

The characteristic of the InP SBD made in Tomsk are classical ones, that of the devise made in Pisa shows an asymmetry between forward and reverse bias, the reverse is closer to a line; the InSb characteristic is close to an ohmic contact, but the non linear part shows a very interesting shape, fig. 2.

The experimental setup for the investigation of the diodes response is schematically shown in Fig. 3. It consisted of a laser radiation source, a microwave source, an SBD and an oscilloscope or a spectrum analyzer as the registering apparatus. The laser radiation was focused onto the SBD by a lens of 10-mm of focal length and the RF radiation was coupled to the diode from an HP 8350B sweep oscillator directly by a microwave horn.

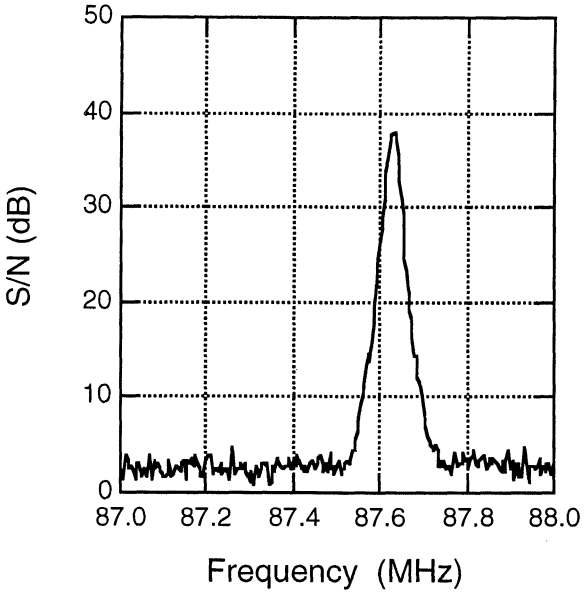


Fig. 4 Beat note betwwen two longitudilal modes of an He-Ne laser, generated by an InSb SBD.

At 633 nm we used as the laser source an He-Ne-laser with an output power of about 7 mW, typically working on two adjacent longitudinal TEM_{00} modes. In the green and the blue we used an Ar-ion laser, with an output power of about 50 mW.. In the infrared diode lasers were used. We measured the rectified videosignals by an oscilloscope. Fig.4 shows the tipical beat detected by the spectum analyzer

In the whole explored range of visible wavelength the GaAs diodes give always the largest video signals, fig.5. We have also tested the frequency mixing properties of the devices by sending the detected signal after amplification in a spectrum analyzer with a bandwidth of 100 kHz and observing the beat notes between different radiation modes of the He-Ne laser and up to the 25th harmonic of the microwave radiation.

In Fig. 6 we show the experimental results for mixing. As mixers, both InP- and InSb diodes proved to be largely more effective than GaAs. The intermode beat signal of the He-Ne-laser at f_b = 875 MHz was observed with a signal-to-noise (S/N) ratio of 70 dB and 68 dB

by using respectively InP- and InSb-diode, against only 40 dB with GaAs-SBD. Similarly, the intermode beat signal at f_b = 130 MHz e Ar-laser was observed with a S/N ratio of about 55 dB, 58 dB, and 20 dB, for InP-, InSb-, and GaAs-diodes, respectively. No beat notes were observed by GaAs diodes below 514 nm, while beat notes up to f_b = 5 GHz (that is in the limit of the gain linewidth of the Ar laser) have been observed with InP and InSb diodes

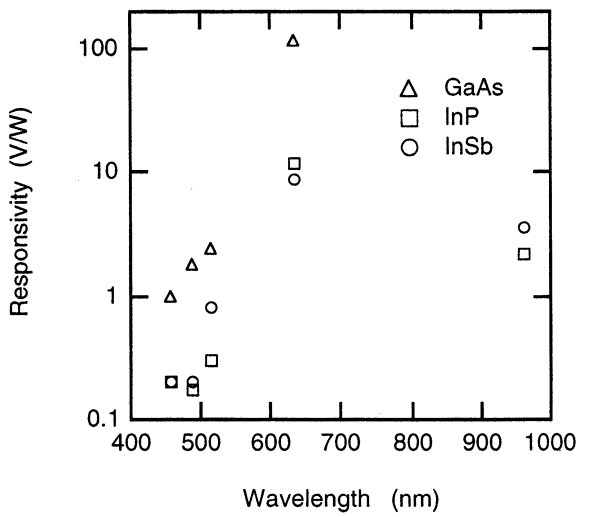


Fig. 5 Video signals obtained by the different SBD

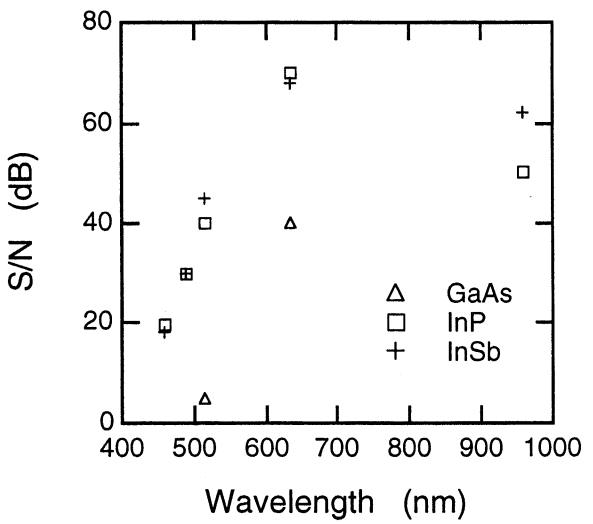


Fig. 5 Mixing signals obtained by the different SBD

In high order mixing experiments, when we applied the microwave power, there was a moderate decrease of the observed beat note signal, but the S/N ratio was still good up to the 27th order of mixing. For example for the InP diode with the He-Ne laser radiation, the application of the MW power led to a decrease of some eight dB of the second order signal; the third order signal presented a S/N ratio of 58 dB and we still had a 18 dB S/N for the twenty-seventh-order signal. It is obviously to be noted that our experiments were carried out in presence of much less power with respect to the previous experiments with GaAs or MIM diodes.

3. CONCLUSION

The results of our experiments showed that the various SBD have different mixing and detection properties. The GaAs-SBD proved to be a very effective detector but has a very low frequency conversion efficiency in the visible. On the contrary, InP- and InSb-SBD operate very effectively as mixers even with relatively low detection efficiencies. This behaviour is quite similar to the results obtained in previous works using infrared radiation [6, 7].

These results can be used in the practical work of construction of different systems of measuring and synthesizing laser frequencies in the visible range and other fields of practical application where there is a need to use nonlinear devices having high sensitivity, high speed of operation, high mechanical stability and ability to work at room temperature

10. REFERENCES

[1] V.M. Klementyev, B.A. Timchenko, V.F. Zakhar'yash, Pribory i Tekhnika Eksperimenta, vol. 1, p. 128, 1992.

[2] O. Acef, L. Hilico, M. Bahoura, F. Nez, P. De Natale, Opt. Commun., vol. 109, p. 428, 1994

[3] H.-W. Hubers, G.W. Schwaab, H.P. Roser, J. Appl. Phys. vol. 75, 4243 (1994).

[4] V.G. Bozhkov, V.M. Klementyev, O.Yu. Malakhovsky, B.A. Timchenko, V.F. Zakhar'yash, Radiotekhnika i Elektronika, vol. 42, p. 489, 1997.

[5] V.G. Bozhkov, S.V. Chepurov, V.M. Klementyev, O.Yu. Malakhovsky, B.A. Timchenko, V.F. Zakhar'yash, Radiotekhnika i Elektronika, vol. 42, p. 631, 1997.

[6] S. N. Bagayev, V.G.Bozhkov, S.V. Chepurov, A.G. Khamoyan, V.M. Klementyev, O.Yu. Malakhovsky, V.S. Pivtsov, V.F. Zakhar'yash, Quantum Electronics, vol. 25, p. 558, 1998.

[7] H.-U. Daniel, M. Steiner, and H. Walter, Appl. Phys. B **25**, 7 (1981); Appl. Phys. B **26**, 19 (1981).

[8] H.-U. Daniel, B. Maurer, M. Steiner, Appl. Phys. B vol. 30, p. 189, 1983.

[9] R.E. Drullinger, K.M. Evenson,D.A. Jennings, F.R. Petersen, J.C. Berquist, L. Burkins, H.-U. Daniel, Appl. Phys. Lett., vol. 42, p. 137, 1982.

[10] T. Kokubo, A. Gallagher, J.L. Hall, J. Appl. Phys., vol. 85, p.1311, 1999.

[11] A. Bertolini, G.Carelli, C.A. Massa, A. Moretti and F. Strumia, Infrared Phys. and Technol. (in press).

SUB-HERTZ BEAT NOTE BETWEEN TWO ELECTRO-OPTICAL PARAMETRIC OSCILLATORS

A. WOLF, B. BODERMANN AND H. R. TELLE

Physikalisch-Technische Bundesanstalt D-38023 Braunschweig, Germany

We demonstrate a 700 mHz wide beat note generated by two laser diode pumped electro-optical parametric oscillators (EOPO) oscillating at λ=850 nm. Such an oscillator [1] is based on an electro-optic phase modulator (EOM), an empty resonator, a photo diode and a rf amplifier. The oscillation loop feeds the beat note between the pump and the resonant field to the EOM. The resulting PM side-band sustains the resonator excitation. We used a quiet, high finesse resonator (~50 kHz FWHM) whose phase slope was much steeper than that of the oscillation loop leading to an intrinsically small frequency noise crosstalk from the pump to the EOPO emission. As a further improvement, we removed even these small frequency pulling effects by (frequency-offset) phase locking the pump field to the EOPO emission.
In this initial experiments, we stabilized both to adjacent longitudinal modes (FSR = 2 GHz) of the high finesse resonator in order to avoid the need an effective acoustical shielding of the cavity.
Effects ultimately limiting the emission line widths of these novel oscillators will be discussed together with possible applications in metrology and spectroscopy.

[1] A. Wolf, H. R. Telle: Opt. Lett. 23, 1775 (1998)

Corresponding author:

Andreas Wolf
Physikalisch-Technische Bundesanstalt
Proj 4.301
P. O. 3345
D-38023 Braunschweig, Germany

Phone: ++49-531-592-4425
Fax : -4423
Email:andreas.wolf@ptb.de

A 3 THz WIDE OPTICAL FREQUENCY COMB FOR THE NPL METHANE-TO-STRONTIUM FREQUENCY CHAIN

H.S. Margolis, S.N. Lea, G. Huang, and P. Gill

National Physical Laboratory, Teddington, Middlesex TW11 0LW, UK

Corresponding author: S. N. Lea, National Physical Laboratory, Queens Road, Teddington, Middlesex, TW11 0LW, UK; Tel. +44 (0)20 8943 6102; Fax. +44 (0)20 8943 2945; E-mail: Stephen.Lea@npl.co.uk

ABSTRACT

In this paper we describe the construction and performance of an optical frequency comb generator for the NPL methane-to-strontium frequency chain. A simple single-cavity design using a commercial electro-optic phase modulator is shown to be sufficient to generate a 3 THz-wide comb in the near infrared.

1. INTRODUCTION

A programme of optical frequency chain development is being undertaken at NPL, to provide a facility for improved absolute frequency measurements of a range of new and existing optical frequency standards [1, 2]. In particular, this work is being pursued in support of the development of optical frequency standards based on forbidden transitions in cold, trapped, single ions of strontium and ytterbium. The initial aim is to relate the frequency of the $^{88}Sr^{+}$ $^{2}S_{1/2} - {}^{2}D_{5/2}$ electric quadrupole transition at 444.8 THz (674 nm) [3] to that of a methane-stabilised helium-neon laser system (HeNe/CH_4) at 88.4 THz (3.39 μm). The strontium transition, which has a natural linewidth of 0.4 Hz, is the first optical transition in a cold, trapped ion to be adopted by the Comité International des Poids et Mesures (CIPM) as a recommended radiation for the realisation of the metre, with a relative standard uncertainty of 1.3×10^{-10} [4], following absolute frequency measurements at NRC, Canada, and at NPL. The determination of the frequency using the NRC phase-coherent frequency chain has recently been improved to give a measurement with a relative standard uncertainty of 4.5×10^{-13} [5]. The NPL measurement used an interferometric technique referenced to an iodine-stabilised HeNe laser standard, and has a relative standard uncertainty of 1.3×10^{-10} [6]. The initial target of our methane-to-strontium frequency chain is to reduce the uncertainty of measurement of the NPL strontium ion traps to the few kHz level.

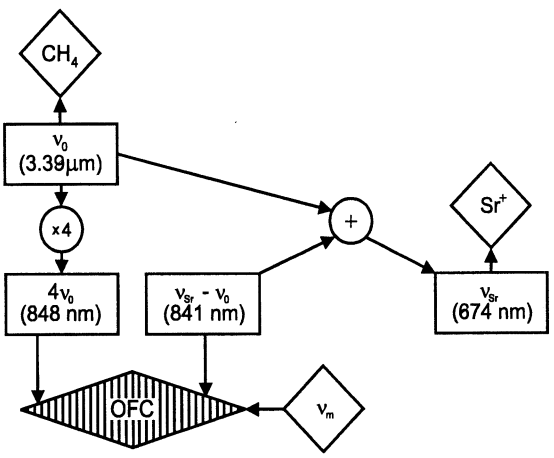


Figure 1. Schematic diagram of the NPL methane-to-strontium frequency chain. ν_0: frequency of the HeNe/CH_4 standard; ν_{Sr}: frequency of the $^{88}Sr^{+}$ ion trap standard; OFC: optical frequency comb; ν_m: comb modulation frequency.

2. METHANE-TO-STRONTIUM CHAIN

The frequency of the strontium transition is only 2.9 THz greater than five times the HeNe/CH_4 frequency, making the latter a particularly convenient reference standard for this measurement and resulting in a relatively straightforward frequency linkage, illustrated in figure 1.

2.1 Methane-stabilised HeNe standard

The HeNe/CH_4 standard which is the initial reference for the chain is a TOFS-60 device constructed by the group of M. Gubin at the P.N. Lebedev Physical Laboratory [7]. The absolute frequency of this laser has been measured using the rf-to-optical frequency chains at PTB and LPTF with relative standard uncertainty 4×10^{-13} and reproducibility over an interval of one year between successive measurements at PTB of 1.2×10^{-13} [8].

2.2 Optical frequency synthesis

The output power of the TOFS-60 system is too low to directly drive any non-linear process. We generate 400 mW of single-frequency 88.4 THz radiation using a singly-resonant optical parametric oscillator (OPO) pumped by a single-frequency Nd:YLF laser [2].

The $^{88}Sr^{+}$ frequency has been synthesised by sum frequency mixing (SFM) of the methane frequency and an extended cavity diode laser (ECDL) at 356.4 THz (841 nm) in potassium niobate. We intend to use periodically-poled lithium niobate (PPLN) to obtain greater non-linear conversion efficiency for this process when sum-frequency mixing with the OPO output. The OPO output is also frequency doubled in $AgGaSe_2$ to provide a frequency reference for a DFB laser at 176.8 THz (1696 nm). The output of the DFB laser is frequency doubled in PPLN to provide a frequency reference for an ECDL at 353.5 THz (848 nm). These non-linear interactions are all performed in single-pass geometry, eliminating the need for build-up cavity locking systems which may introduce additional noise into the locked chain. The frequency difference of 2.9 THz between the two ECDLs is bridged by an optical frequency comb which is described in detail in the remainder of this paper.

3. OPTICAL FREQUENCY COMB

3.1 Theory

An optical frequency comb [9,10] uses an electro-optic phase modulator (EOM) to superimpose a microwave frequency ν_m onto an optical carrier frequency ν_c, producing a comb of equally spaced modulation sidebands. Typically, the single-pass modulation index β is such that energy is shifted mainly into the first-order sidebands at $\nu_c \pm \nu_m$, and the power in higher-order sidebands falls off rapidly with increasing order number. To enhance the power in these modes, and hence the number of modes which can be detected, the EOM is placed inside an optical cavity which is resonant with both the carrier and all the sidebands, i.e. the free spectral range of this cavity is equal to, or is a subharmonic of, the modulation frequency. Thus the sidebands generated from the carrier pass back through the EOM and generate secondary sidebands, which are also resonant with the cavity and generate further sidebands. In this way, a wide comb of modes spaced by the modulation frequency ν_m can be produced. The accuracy of the frequency intervals between these modes is determined by the accuracy of the microwave oscillator used to drive the EOM.

As a result of the finite cavity finesse, the output power of the higher-order sidebands falls off exponentially. The power P_k in the kth sideband of the comb is given approximately by [9]

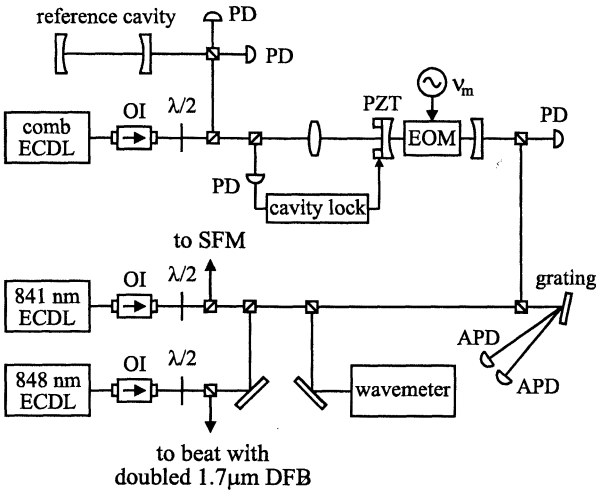


Figure 2. Layout of the optical frequency comb system. (OI: optical isolator; λ/2: half-wave plate; PZT: piezo element; PD: photodiode; APD: avalanche photodiode.)

$$P_k = \eta_{FP} P_{in} \left(\frac{\pi}{2\beta F} \right)^2 \exp\left(-\frac{|k|\pi}{\beta F} \right), \qquad (1)$$

where F and η_{FP} are the finesse and peak transmission, respectively, of the optical cavity with the microwave power to the modulator switched off, and P_{in} is the laser power incident on the cavity. Clearly, a large single-pass modulation index and a high cavity finesse are key to obtaining a wide frequency span.

3.2 Experimental set-up

The optical frequency comb uses a commercial EOM consisting of an anti-reflection coated MgO-doped $LiNbO_3$ crystal embedded in a resonant microwave cavity. The microwave resonance has a bandwidth of around 50 MHz and the resonant frequency can be tuned manually over the range 9.2-9.7 GHz. In our device, the optimum modulation index is obtained near 9.7 GHz. A low noise microwave synthesiser is used to drive the EOM, to prevent collapse of the high-order sidebands due to the multiplied phase-noise amplitude.

This EOM is placed inside a two-mirror optical cavity as shown in figure 2. The length of this cavity is chosen so that its free spectral range is equal to one twelfth of the modulation frequency ν_m. With the microwave power to the modulator switched off, this cavity has a finesse F of around 200 and a peak transmission η_{FP} of around 0.3. The laser used to drive the comb is an ECDL operating at around 355 THz, which is locked to a temperature-stabilised reference cavity using the side-of-fringe technique. The power incident on the comb cavity is about 10 mW.

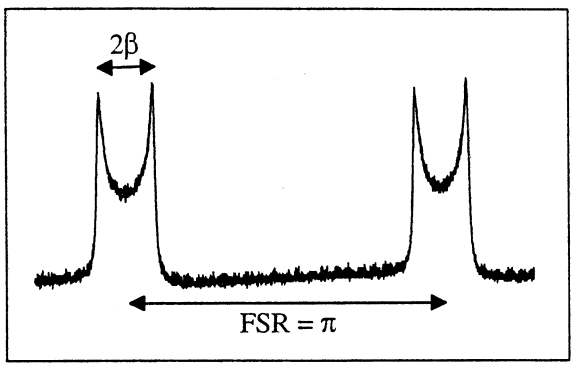


Figure 3. FM lineshape obtained in transmission with the EOM on. The cavity FSR corresponds to a phase change of π radians and the full width of the FM lineshape corresponds to twice the modulation index, β.

When the microwave power to the modulator is turned on, the transmission of the comb cavity as a function of the cavity length changes from the usual Airy function of the unmodulated resonator to a broad frequency modulation (FM) lineshape, figure 3. This lineshape is used to optimise the alignment and the length of the comb cavity, and also enables the modulation index to be determined, because the width of the lineshape, expressed in terms of phase, is 2β.

The comb cavity is servo-locked to the frequency of the ECDL used to drive the comb, with the lock point at the trough of the FM lineshape. To achieve this, the cavity length is dithered at a frequency of 3 kHz by means of piezo-mounted mirror. The light reflected from the cavity is phase-sensitively detected at the dither frequency to generate a discriminator signal, and then integrated to provide feedback to the cavity piezo.

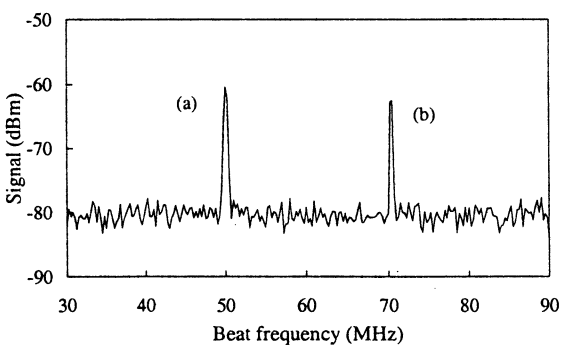


Figure 4. Spectrum analyser trace showing simultaneous beats with two comb modes separated by 2.9 THz: (a) beat between 841 nm laser and 155th comb mode; (b) beat between 848 nm laser and -145th comb mode. The EOM is driven with 0.7 W at 9.683 GHz. Resolution bandwidth 100 kHz, sweep rate 0.5 ms/MHz.

3.3 Beats spectra

The performance of the comb is investigated by looking at beat signals between the comb output and the ECDLs at 353.5 THz and 356.4 THz using avalanche photodiodes. A beat signal is only seen between an ECDL and the comb mode k nearest in frequency, because the bandwidth of the avalanche photodiodes is about 1 GHz. Measurement of the ECDL frequencies on a wavemeter with 10 MHz resolution is sufficient to determine which comb mode is contributing to a beat signal.

The ECDLs can be tuned in frequency to beat with selected comb modes. With a microwave power of about 0.7 W applied to the EOM at 9.683 GHz (which gives a modulation index of about 0.3 radians), combs with a width of 3 THz have been achieved. The power in the higher-order comb modes is observed to fall off at a rate of about -23 dB/THz, in good agreement with the value predicted using equation (1).

Figure 4 shows beat signals between the 841 nm and 848 nm ECDLs and two comb modes 2.9 THz apart. These are simultaneous beat signals observed on a single avalanche photodiode. Beats with an improved signal-to-noise can be observed by using a grating to separate the two beats spatially and to reduce the number of unwanted comb modes which reach the detectors, as indicated in figure 2. This decreases the noise level determined by the dc power because the higher powers in the second beat laser, the carrier and the lower-order comb modes are not detected.

3.4 Limits to the comb width

The limit of the frequency difference at which beats can be detected between comb modes and a second laser depends on the power of the sidebands. From equation (1), it can be seen that the power in a particular high order sideband increases as either F or β is increased. However, this equation holds only in the ideal case that material dispersion does not exist. In practice, the span of an optical frequency comb is limited by the dispersion of the cavity elements, in particular dispersion in the electro-optic crystal. With dispersion, the outer sidebands of the comb are no longer resonant with the cavity and so their intra-cavity fields are too low to generate further sidebands. Thus a sharp drop in sideband power is observed above some maximum sideband number.

It has been shown that the dispersion-limited comb width is given by [11]

$$\Delta\nu_{\text{d-limit}} = \sqrt{\frac{8c^2\beta}{\pi\lambda_0^3 L_c\left(\partial^2 n_e / \partial\lambda^2\right)_{\lambda=\lambda_0}}}, \qquad (2)$$

where λ_0 is the wavelength of the laser used to drive the comb, L_c and n_e are the length and extraordinary refractive index of the EOM crystal respectively, and c

is the speed of light. Using Sellmeier equations for MgO-doped $LiNbO_3$ given in reference [12], we estimate that for a modulation index of $\beta = 0.3$, $\Delta\nu_{\text{d-limit}}$ should be about 3 THz at 845 nm. Hence it appears that the width of our comb is currently limited by dispersion in the EOM crystal.

3.5 Possible improvements

The simplest way to increase the power in the higher-order sidebands of the comb, and also the span of the comb, is to increase the modulation index β. Since β is proportional to the square root of the microwave power into the modulator, the dispersion-limited comb width scales as the fourth root of the power. Using a higher power microwave amplifier, we estimate that we should be able to increase the power by around a factor of four before heat dissipation in the modulator becomes a serious problem. This should increase the comb span by a factor of $\sqrt{2}$, and give an order-of-magnitude increase in the power of the comb modes with $k = \pm 150$.

Because of the practical limits to the modulation index imposed by the microwave power which can safely be used to drive the modulator, several techniques have been devised to overcome the dispersion limit to the comb width. Dispersion-compensating prisms have been inserted into the comb cavity [13], and recently a novel type of comb consisting of an optical parametric oscillator with an intracavity EOM has been demonstrated [14]. These arrangements, however, are significantly more complex than the simple two-mirror cavity described here. Another approach is to broaden the comb spectrum by amplification and four-wave mixing in a dispersion-flattened optical fibre [15, 16].

Although the fraction of the optical power transferred to a high-order sideband of the comb can be increased by increasing the modulation index and the finesse of the optical cavity, this decreases the total optical power in the comb. The modulation transfers power away from the carrier frequency into the sidebands, and this apparent loss at the carrier frequency reduces the coupling efficiency of the laser to the cavity. The coupling efficiency may be improved by using an input coupling cavity which is frequency locked to the carrier frequency [17]. However, the amount of optical power that should be coupled into the comb cavity is limited by photorefractive damage in the EOM crystal, and so if this technique is used great care must be taken not to exceed this damage threshold.

In a similar way, a filter cavity can be used to achieve efficient output coupling of a particular sideband from the comb [18]. If a short, high finesse, filter cavity is used which has a free spectral range greater than the comb width, the full power of a single sideband can be extracted from the comb cavity, whilst keeping the carrier and all the other sidebands trapped inside. This gives an improvement in the signal-to-noise ratio in the beat signals observed. However, it is not clear that it is possible to extract two comb modes on either side of the carrier frequency in this way.

4. CONCLUSION

We have demonstrated that a simple design of single-cavity optical frequency comb generator is adequate to bridge a 2.9 THz-wide frequency mismatch in the NPL methane-to-strontium frequency chain. The passive comb generator design we have developed has a number of applications within our absolute optical frequency metrology programme. After the initial measurement of the strontium frequency it is intended to reconfigure the NPL frequency chain to synthesise the 5th harmonic of the CH_4 standard and to link to the Sr^+ standard by means of an optical frequency comb bridging the 441.9 THz – 444.8 THz interval. This will simplify the future development of the frequency chain to embrace measurements of other optical frequency standards. However, the development of a 3 THz-wide comb at 674 nm is more challenging than at 848 nm, because MgO-doped $LiNbO_3$ is more dispersive at shorter wavelengths.

As part of our programme to develop a compact all-solid-state caesium-to-optical frequency chain we also intend to develop a broad active comb generator based on a repetition-rate stabilised, Kerr-lens mode-locked femtosecond Ti:sapphire laser [19]. Futher broadening of such a comb by self-phase modulation in an optical fibre has recently been demonstrated [20,21], offering the possibility of directly relating the HeNe/CH_4 frequency to the caesium primary standard.

Despite these recent advances, there remains a role for passive comb generators such as that described in this paper, particularly in spectral regions inaccessible to Ti:sapphire lasers, for example around 474 THz for diode-laser spectroscopy of I_2, and in the 1.5 μm optical communications band. A frequency chain incorporating both active and passive comb generators will offer direct linkage to the caesium primary frequency standard for optical frequencies throughout the visible and near-infrared.

ACKNOWLEDGEMENTS

We thank M. Gubin, G. Kramer, O. Acef, and their co-workers for their collaboration in the inter-comparisons of our HeNe/CH_4 standard with the PTB and LPTF frequency chains. This work is supported by the UK National Measurement System Length Programme under contract LE99/A01.

REFERENCES

[1] S. N. Lea, G. M. Macfarlane, G. Huang and P. Gill, "Progress toward an optical frequency chain at NPL", IEEE Trans. Instrum. Meas., vol. 48, pp. 578-582, 1999.

[2] S. N. Lea *et al.*, "Optical frequency synthesis in the near IR with application to trapped ion standards", in Proceedings of the 1999 Joint Meeting EFTF – IEEE IFCS, Besançon, France, 1999, pp. 647-650.

[3] G. P. Barwood, P. Gill, G. Huang and H. A. Klein, "Development of an optical frequency standard based upon the $^2S_{1/2}$-$^2D_{5/2}$ transition in Sr^+", in Proceedings of the 1999 Joint Meeting EFTF – IEEE IFCS, Besançon, France, 1999, pp. 686-691.

[4] T. J. Quinn, "Practical realization of the definition of the metre (1997)", Metrologia, vol. 36, pp. 211-244, 1999.

[5] J. E. Bernard *et al.*, "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] G. P. Barwood, P. Gill, H. A. Klein and W. R. C. 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, pp. 133-136, 1997.

[7] M.A. Gubin *et al.*, "Present performance of the transportable He-Ne/CH_4 optical frequency standards", in Proceedings of the 1999 Joint Meeting EFTF - IEEE IFCS, Besançon, France, 1999, pp. 710-713.

[8] O. Acef *et al.*, "Absolute frequency measurements with a set of transportable methane optical frequency standards", in Proceedings of the 1999 Joint Meeting EFTF – IEEE IFCS, Besançon, France, 1999, pp. 742-745.

[9] M. Kourogi, K. Nakagawa and M. Ohtsu, "Wide-span optical frequency comb generator for accurate optical frequency difference measurement", IEEE J. Quantum Electron., vol. 29, pp. 2693-2701, 1993.

[10] L. R. Brothers, D. Lee and N. C. Wong, "Terahertz optical frequency comb generation and phase locking of an optical parametric oscillator at 665 GHz", Opt. Lett., vol. 19, pp. 245-247, 1994.

[11] M. Kourogi, B. Widiyatomoko, Y. Takeuchi and M. Ohtsu, "Limit of optical-frequency comb generation due to material dispersion", IEEE J. Quantum Electron., vol. 31, pp. 2120-2126, 1995.

[12] R. C. Eckardt, C. D. Nabors, W. J. Kozlovsky and R. L. Byer, "Optical parametric oscillator frequency tuning and control", J. Opt. Soc. Am., vol. B8, pp. 646-667, 1991.

[13] L. R. Brothers and N. C. Wong, "Dispersion compensation for terahertz optical frequency comb generation", Opt. Lett., vol. 22, pp. 1015-1017, 1997.

[14] S. A. Diddams, L. S. Ma, J. Ye and J. L. Hall, "Broadband optical frequency comb generation with a phase-modulated parametric oscillator", Opt. Lett., vol. 24, pp. 1747-1749, 1999.

[15] K. Imai, M. Kourogi and M. Ohtsu, "30-THz span optical frequency comb generation by self-phase modulation in an optical fibre", IEEE J. Quantum Electron., vol. 34, pp. 54-60, 1998.

[16] K. Imai, B. Widiyatmoko, M. Kourogi and M. Ohtsu, "12-THz frequency difference measurements and noise analysis of an optical frequency comb in optical fibres", IEEE J. Quantum Electron., vol. 35, pp. 559-564, 1999.

[17] A. S. Bell, G. M. Mcfarlane, E. Riis and A. I. Ferguson, "Efficient optical frequency-comb generator", Opt. Lett., vol. 20, pp.1435-1437, 1995.

[18] J. Ye, L. Ma, T. Daly and J. L. Hall, "Highly selective terahertz optical frequency comb generator", Opt. Lett., vol. 22, pp. 301-303, 1997.

[19] Th. Udem, J. Reichert, R. Holzwarth and T. W. Hänsch, "Absolute optical frequency measurement of the caesium D_1 line with a mode-locked laser", Phys. Rev. Lett., vol. 82, pp. 3568-3571, 1999.

[20] J. Reichert, R. Holzwarth, Th. Udem and T. W. Hänsch, "Measuring the frequency of light with mode-locked lasers", Opt. Commun., vol. 172, pp. 59-68, 1999.

[21] S. A. Diddams *et al.*, "Optical frequency measurement across a 104-THz gap with a femtosecond laser frequency comb", Opt. Lett., vol. 25, pp. 186-188, 2000.

HYDROGENIC SYSTEMS: CALCULABLE FREQUENCY STANDARDS

J. L. Flowers, H. A. Klein and H. S. Margolis
National Physical Laboratory, Teddington, Middlesex TW11 0LW, UK

K. Gaarde-Widdowson, J. D. Silver and M. R. Tarbutt
Department of Physics, University of Oxford, Oxford, OX1 3PU, UK

S. Ohtani
Cold Trapped Ions Project, JST, University of Electrocommunications, Chofu, Tokyo 182-0024, Japan

D. J. E. Knight
DK Research, 110 Strawberry Vale, Twickenham, Middlesex, TW1 4SH, UK

Corresponding author: H. S. Margolis, National Physical Laboratory, Queens Road, Teddington, Middlesex, TW11 0LW, Tel. +44 (0) 20 8943 6157, Fax. +44 (0) 20 8943 2945, E-mail: Helen.Margolis@npl.co.uk

ABSTRACT

Hydrogenic systems are simple enough that it is possible to calculate the energies of their transitions with an accuracy approaching or exceeding that with which they can be measured. In this paper we discuss the status of measurements and theory for such systems, and consider the current limitations to the accuracy achieved. One source of theoretical uncertainty arises from two-loop quantum electrodynamic corrections to the energy levels. We describe an experiment designed to test these corrections: a measurement of the $2S_{1/2}$-$2P_{3/2}$ transition in Si^{13+} by laser spectroscopy.

1. INTRODUCTION

The study of the hydrogen atom and hydrogenic (one-electron) ions has emerged as a fertile area for research into future frequency standards [1]. Indeed, the two-photon 243 nm $1S_{1/2}$-$2S_{1/2}$ transition in the hydrogen atom has recently been included in a new list of approved radiations for the practical realisation of the metre [2]. The particular interest of this interval lies in the fact that, uniquely among present optical frequency standards, it may be calculated with an accuracy approaching that with which it has been measured.

Transitions with $\Delta n > 0$ in hydrogen span a large frequency range from radiofrequencies into the vacuum ultraviolet. For hydrogenic ions of higher nuclear charge Z, this range extends into the X-ray region of the spectrum due to the Z^2 scaling of the gross energy level structure. These transitions can be regarded as forming a "natural" frequency scale governed by the Rydberg constant $R_\infty = m_e e^4/(8\varepsilon_0^2 h^3 c)$. Thus hydrogenic systems offer the possibility of a set of frequency standards extending from radiofrequencies into the X-ray region, which would be closely linked to fundamental constants via the Rydberg constant.

2. CURRENT STATUS OF HYDROGEN AS A CALCULABLE FREQUENCY STANDARD

2.1 Experiment

Since the advent of Doppler-free two-photon spectroscopy, which enabled the intrinsically narrow $1S_{1/2}$-$2S_{1/2}$ transition to be studied, impressive improvements in the measurement of its frequency have been achieved (see figure 1).

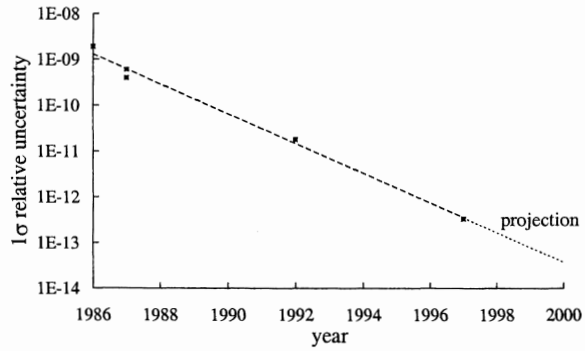


Fig. 1. Accuracy of measurements of the $1S_{1/2}$-$2S_{1/2}$ two-photon transition in hydrogen.

The early absolute frequency measurements of this transition [3-5] used tellurium standards for calibration. A major improvement in accuracy was achieved by Hänsch and co-workers who developed a frequency chain to relate the $1S_{1/2}$-$2S_{1/2}$ interval to a methane-stabilized helium-neon laser standard (HeNe/CH_4) at 3.39 μm [6]. This made use of the fortunate coincidence that the frequency of the laser at 486 nm, which is doubled to excite the two-photon transition, lies close to the 7th harmonic of the HeNe/CH_4 standard. Originally the frequency mismatch of 2.1 THz was bridged interferometrically, but in the most accurate measurement of the $1S_{1/2}$-$2S_{1/2}$ transition frequency to date, it was determined using a phase-coherent chain of five cascaded optical frequency dividers [7]. This enabled the absolute frequency of the

transition to be determined to 3.4 parts in 10^{13}, limited by the reproducibility of the HeNe/CH_4 standard. Recently, techniques have been developed for bridging large frequency intervals in the optical region using mode-locked femtosecond lasers [8], enabling the optical frequency of the $1S_{1/2}$-$2S_{1/2}$ transition to be compared directly with the microwave frequency of a caesium atomic clock. Coupled with higher spectral resolution and an improved theoretical model of the lineshape which allows the unperturbed transition frequency to be determined to an accuracy of 1.5 parts in 10^{14} [9], a significantly improved absolute frequency measurement is anticipated in the near future.

2.2 Theory

To calculate transition frequencies in hydrogen to an accuracy comparable with the experimental precision which has been achieved, it is necessary to take into account a large number of corrections to the values obtained using the Dirac equation. These include quantum electrodynamic (QED) corrections, pure and radiative recoil corrections arising from the finite nuclear mass, and a correction due to the non-zero volume of the nucleus. The evaluation of these corrections is an extremely challenging task (see for example the review of Pachucki *et al.* [10]).

Uncertainties in the calculations arise from a number of sources. Firstly, the calculation of these corrections depends on the values of certain fundamental constants, namely the electron to proton mass ratio m_e/m_p and the fine structure constant α. The values of m_e/m_p and α in the most recent adjustment of the fundamental constants [11] have relative standard uncertainties of 2.1×10^{-9} and 3.7×10^{-9} respectively. The largest correction depending on m_e/m_p is of order $m_e/m_p \approx 5 \times 10^{-4}$, and so is known with an accuracy of around 10^{-12}. The largest term depending on α is included in the Dirac energy, and is of order $\alpha^2 \approx 5 \times 10^{-5}$; the uncertainty of this term is therefore around 2×10^{-13}. Further refinement of the theoretical calculations beyond these levels will therefore require improved measurements of both these fundamental constants.

Secondly, the nuclear size correction to the energy of the S states currently has a large uncertainty because the two most precise measurements of the root mean square charge radius of the proton disagree [12,13]. This uncertainty comes in at the level of 5×10^{-11} for the 1S state of hydrogen, and is reduced by a factor of n for higher-lying levels. Although the Rydberg constant has been determined with an accuracy of 8.3 parts in 10^{12} by considering other transitions in addition to the $1S_{1/2}$-$2S_{1/2}$ transition [7,14], a new measurement of the proton radius is greatly to be desired.

The final type of uncertainty arises from higher-order terms which are neglected in the calculations. For example, whilst the one-loop electron self-energy has recently been calculated with high numerical accuracy (2.4 parts in 10^{16} for the 1S state of hydrogen) [15], the situation for the two-loop binding corrections is less satisfactory. The energy shift due to the two-loop corrections may be written in the form $\Delta E_n = m_e c^2(\alpha/\pi)^2[(Z\alpha)^4/n^3]G_n(Z\alpha)$ where $G_n(Z\alpha)$ is usually written as a double expansion in powers of $(Z\alpha)$ and $\ln(Z\alpha)$. The leading-order term in this expansion has been known for many years [16], but recent calculations have yielded a larger than expected value for the term of order $\alpha^2(Z\alpha)^5$ [17,18], and suggested that higher-order terms may also be larger than anticipated [19]. All the indications are that the convergence of this series expansion is very poor, and that it would be desirable to carry out a high-accuracy numerical calculation which avoids the expansion, as has been done for the one-loop self energy.

3. SINGLY IONIZED HELIUM AND HIGHER-Z ONE-ELECTRON SYSTEMS

3.1 He^+

The theoretical limitations of the proton size uncertainty may be avoided in He^+, where the nuclear size (that of the alpha particle) is well known. This has motivated a project at the University of Sussex, which aims to determine the 2S Lamb shift in He^+ via measurements of two-photon 2S-*n*S,*n*D transitions [20]. By comparing the results with the measurements in hydrogen, it should be possible to separate out the nuclear size and QED contributions to the Lamb shift. There is also a nine standard deviation discrepancy between the latest theoretical calculations and the best previous measurement of the 2S Lamb shift for He^+, carried out using a quenching anisotropy method [21]. However, an unexpected polarization-related systematic effect in that experiment is currently under investigation, and may resolve the situation [22].

3.2 Higher-Z hydrogenic ions

The energy of gross structure transitions in hydrogenic systems scales as Z^2. Thus as Z increases the Lyman-α transitions move from the UV to the X-ray region of the spectrum. Highly charged one-electron ions can be made in a variety of ways. High velocity beams from heavy ion accelerators with velocities 5-10% of the speed of light can be stripped by a thin foil or gas; alternatively recoil ions can be excited by passing such a beam through a dilute gas, e.g. neon or argon. One-electron ions can also be produced in plasma sources or electron beam ion traps.

The best measurements of Lyman-α wavelengths in one-electron ions in the range Z=12-28 have reported accuracies of up to 5 ppm [23-25]. Measurements of lower accuracy have been made right up to U^{91+} [26], for which the energy of the Lyman-α X-rays is about 100 keV. These transitions could one day serve as

calculable frequency standards in the X-ray region, if the QED contributions can be calculated with sufficient accuracy.

4. LASER SPECTROSCOPY OF THE 2S LAMB SHIFT IN HYDROGENIC SILICON

4.1 Motivation

Owing to the rapid scaling of QED effects with atomic number Z, transitions in highly charged ions are much more sensitive than hydrogen to the two-loop binding corrections to the energy levels. Measurements in highly charged hydrogenic ions could therefore test the calculations of these terms, and enable the viability of hydrogen and other hydrogenic systems as frequency standards to be assessed. For example, the $2S_{1/2}$-$2P_{3/2}$ transitions in a number of medium-Z hydrogenic ions lie in a wavelength range which is accessible to high-power tuneable lasers. The most sensitive measurement to date in this range of Z is that for P^{14+} [27], which has a fractional precision of 0.14% of the 2S Lamb shift. This is about the same size as the self-energy correction of order $\alpha^2(Z\alpha)^5$, and hence improved precision is really required to obtain a critical test of theory.

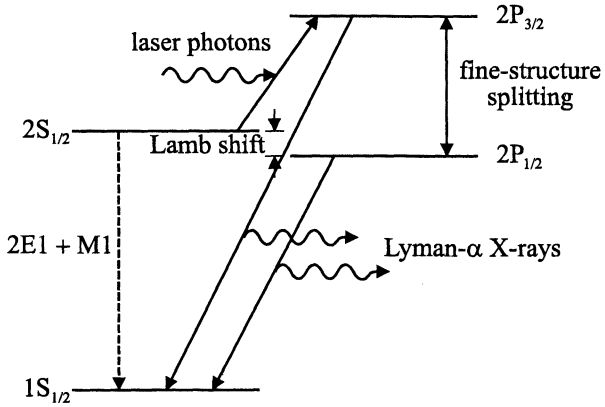


Fig. 2. Partial term diagram of a hydrogenic ion, showing the transition that would be induced in an n=2 laser resonance experiment. (Not to scale.)

4.2 Principle of the experiment

The technique for studying the $2S_{1/2}$-$2P_{3/2}$ transitions in medium-Z ions by laser spectroscopy is illustrated in figure 2. Laser radiation is used to excite ions from the metastable $2S_{1/2}$ state to the $2P_{3/2}$ state, from which they rapidly decay to the ground state via an allowed electric dipole transition. The resonance is monitored by observing the rate of emission of Lyman-α photons as a function of the laser frequency. The 2S Lamb shift may be deduced from such a measurement of the $2S_{1/2}$–$2P_{3/2}$ interval because the n=2 fine structure splitting is more accurately known theoretically.

4.3 Source of highly charged ions

All previous 2S Lamb shift measurements for medium-Z hydrogen-like ions have been carried out using fast ion beams, and uncertainties associated with Doppler shifts form a significant source of error in all these experiments. Various methods have been employed or suggested for reducing the sensitivity of fast beam experiments to Doppler corrections [28-30], A measurement of the $2S_{1/2}$–$2P_{3/2}$ transition frequency in N^{6+} using a fast ion beam is currently under way at Florida State University [31]. Our approach, however, is to reduce such systematic errors by using slow ions trapped in the Oxford electron beam ion trap (EBIT) [32], which have no net centre-of-mass motion.

The operation of an EBIT is described in detail elsewhere [33]. Briefly, neutral atoms or low charge state ions are injected into an electron beam which is compressed to high current density by an axial magnetic field, where they are stripped to high charge states by sequential electron impact ionization. The resulting ions are confined radially by a combination of the space charge of the electron beam and the axial magnetic field, whilst axial trapping is achieved using potentials applied to a series of cylindrical electrodes. In this way, the ions are confined to a cylindrical volume about 70 μm in diameter and 2 cm in length, where they are further excited by electron impact for spectroscopic measurements. The ionization balance obtained within the trap depends on parameters such as the electron beam energy and background gas pressure.

4.4 Choice of ion

We have decided to begin our studies by measuring the $2S_{1/2}$-$2P_{3/2}$ interval in hydrogen-like Si^{13+}. This transition has previously been observed using a fast ion beam [34] but there has been no precision measurement of its frequency to date. Our choice of Si^{13+} is based on the availability of high-power commercial lasers at the appropriate wavelength, the ease of ion production in an EBIT, and the efficiency with which the Lyman-α X-rays can be detected. The $2S_{1/2}$–$2P_{3/2}$ transition in Si^{13+} is at a wavelength of approximately 734 nm, and has a natural width of about 7 nm, so can be studied using a Ti:sapphire laser. Silicon ions have been injected into the Oxford EBIT using a metallic vapour vacuum arc (MEVVA) ion source [35], and an electron beam energy of 3–4 keV has been shown to be sufficient to produce the hydrogen-like charge state [36]. The Lyman-α X-rays have an energy of about 2 keV, and may conveniently be observed using a lithium-drifted silicon (Si(Li)) detector.

4.5 Experimental arrangement

Although the lifetime of the $2S_{1/2}$ state is much longer than that of the $2P_{3/2}$ state, it is still only 16 ns. For this reason, very high laser power (several kW or

more) is required to obtain a reasonable transition rate to the $2P_{3/2}$ state [37]. Such intensities may be obtained by using an extremely high finesse enhancement cavity to build up the output power from a frequency-stabilized laser [38].

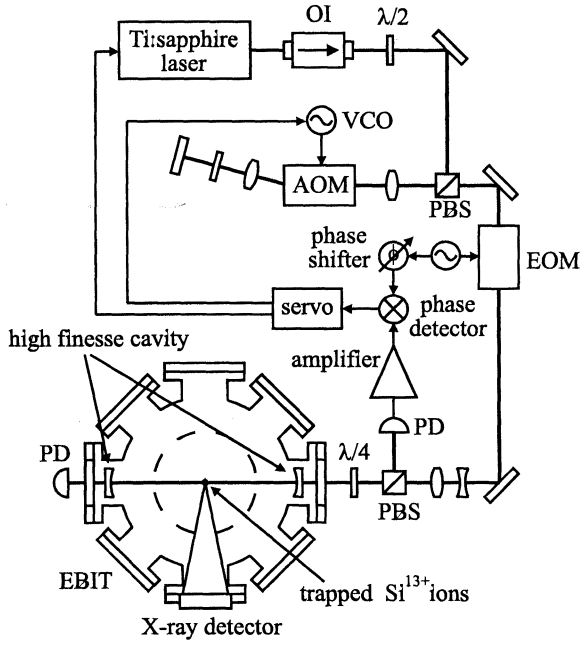


Fig. 3. Schematic diagram of the laser system to be used for a measurement of the $2S_{1/2}$-$2P_{3/2}$ transition frequency in Si^{13+}. (OI: optical isolator, λ/2: half wave plate, λ/4: quarter wave plate, PBS: polarizing beamsplitter, VCO: voltage-controlled oscillator, AOM: acousto-optic modulator, EOM: electro-optic modulator, PD: photodiode.)

The source of 734 nm radiation is a continuous-wave, single-frequency Ti:sapphire laser which will be coupled into a high finesse cavity constructed around the Oxford EBIT as shown in figure 3. The trapped Si^{13+} ions will lie at the laser beam waist within the enhancement cavity, and to keep the finesse of the cavity as high as possible it is necessary for the high reflectivity mirrors to lie within the EBIT vacuum chamber. The Ti:sapphire laser will be locked to the high finesse cavity using the rf sideband locking technique [39]. Fast frequency fluctuations will be corrected using an acousto-optic modulator in a double-pass configuration, whilst the slower branch of the servo loop will use a piezo-mounted mirror in the laser cavity.

Construction of a prototype enhancement cavity and the associated locking electronics is currently underway. This cavity, which will initially operate in air, consists of two identical mirrors with 20 ppm absorption plus scattering loss, and 200 ppm transmission. The cavity dimensions are somewhat constrained by the ion trap geometry; we intend to use mirrors with 0.5 m radius of curvature, positioned in a near-confocal arrangement. This will give a $1/e^2$ beam waist radius of around 250 μm, sufficiently large to avoid serious difficulties in aligning the laser beam to the trapped ions. Such a cavity has a finesse of around 1.4×10^4, which for an incident laser power of 100 mW would give an average intensity of about 4 kW mm^{-2} at the beam waist, sufficient to obtain a reasonable $2S_{1/2}$-$2P_{3/2}$ transition rate.

ACKNOWLEDGEMENTS

This work is supported by the UK National Measurement System under the Foundation Programme project "Hydrogenic systems: calculable frequency standards", and by the Japan Science and Technology Corporation as part of the "Cold Trapped Ions" International Collaborative Research Project.

REFERENCES

[1] B. Cagnac, M. D. Plimmer, L. Julien and F. Biraben, "The hydrogen atom, a tool for metrology", Rep. Prog. Phys., vol. 57, pp. 853-893, 1994.

[2] T. J. Quinn, "Practical realization of the definition of the metre (1997)", Metrologia, vol. 36, pp. 211-244, 1999.

[3] E. A. Hildum *et al.*, "Measurement of the 1S-2S frequency in atomic hydrogen", Phys. Rev. Lett., vol. 56, pp. 576-579, 1986.

[4] R. G. Beausoleil *et al.*, "Continuous-wave measurement of the 1S Lamb shift in atomic hydrogen", Phys. Rev. A, vol. 35, pp. 4878-4881, 1987.

[5] M. G. Boshier *et al.*, "Laser spectroscopy of the 1S-2S transition in hydrogen and deuterium: determination of the 1S Lamb shift and the Rydberg constant", Phys. Rev. A, vol. 40, pp. 6169-6184, 1989.

[6] T. Andreae *et al.*, "Absolute frequency measurement of the hydrogen 1S-2S transition and a new value of the Rydberg constant", Phys. Rev. Lett., vol. 69, pp. 1923-1926, 1992.

[7] Th. Udem *et al.*, "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, 1997.

[8] J. Reichert, R. Holzwarth, Th. Udem and T. W. Hänsch, "Measuring the frequency of light with mode-locked lasers", Opt. Commun., vol. 172, pp. 59-68, 1999.

[9] A. Huber, B. Gross, M. Weitz and T. W. Hänsch, "High resolution spectroscopy of the 1S-2S

transition in atomic hydrogen", Phys. Rev. A, vol. 59, pp. 1844-1851, 1999.

[10] K. Pachucki *et al.*, "Theory of the energy levels and precise two-photon spectroscopy of atomic hydrogen and deuterium", J. Phys. B, vol. 29, pp. 177-195, 1996.

[11] P. J. Mohr and B. N. Taylor, "CODATA recommended values of the fundamental physical constants: 1998", to be published in J. Phys. Chem. Ref. Data; http://physics.nist.gov/constants.

[12] L. N. Hand, D. G. Miller and R. Wilson, "Electric and magnetic form factors of the nucleus", Rev. Mod. Phys., vol. 35, pp. 335-349, 1963.

[13] G. G. Simon, Ch. Schmitt, F. Borkowski and V. H. Walther, "Absolute electron-proton cross-sections at low momentum transfer measured with a high pressure gas target system", Nucl. Phys. A, vol. 333, pp.381-391, 1980.

[14] B. de Beauvoir *et al.*, "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, 1997.

[15] U. D. Jentschura, P. J. Mohr and G. Soff, "Calculation of the electron self-energy for low nuclear charge", Phys. Rev. Lett., vol. 82, pp. 53-56, 1999.

[16] T. Appelquist and S. J. Brodsky, "Fourth-order electrodynamic corrections to the Lamb shift", Phys. Rev. A, vol. 2, pp. 2293-2303, 1970.

[17] K. Pachucki, "Complete two-loop binding correction to the Lamb shift", Phys. Rev. Lett., vol. 72, pp. 3154-3157, 1994.

[18] M. I. Eides and V. A. Shelyuto, "Corrections of order $\alpha^2(Z\alpha)^5$ to the hyperfine splitting and the Lamb shift", Phys. Rev. A, vol. 52, pp. 954-961, 1995.

[19] S. Mallampalli and J. Sapirstein, "Perturbed orbital contribution to the two-loop Lamb shift in hydrogen", Phys. Rev. Lett., vol. 80, pp. 5297-5300, 1998.

[20] M. G. Boshier, "Hydrogen-like systems and quantum electrodynamics", in "Atomic Physics 15", Proceedings of the 15th International Conference on Atomic Physics, Amsterdam, The Netherlands, 1996, pp. 328-342.

[21] A. van Wijngaarden, J. Kwela and G. W. F. Drake, "Measurement of the n=2 Lamb shift in He^+ by the anisotropy method", Phys. Rev. A, vol. 43, pp. 3325-3342, 1991.

[22] G. W. F. Drake, private communication, 1999.

[23] G. Hölzer *et al.*, "Absolute wavelength measurement of the Lyman-α transitions in hydrogenic Mg^{11+}", Phys. Rev. A, vol. 57, pp. 945-948, 1998.

[24] H. F. Beyer, R. D. Deslattes, F. Folkmann and R. E. LaVilla, "Determination of the 1s Lamb shift in one-electron argon recoil ions", J. Phys. B, vol. 18, pp. 207-215, 1985.

[25] H. F. Beyer *et al.*, "Measurement of the 1s Lamb shift in hydrogen-like nickel", Phys. Rev. A, vol. 43, pp. 223-227, 1991.

[26] H. F. Beyer *et al.*, "Measurement of the ground-state Lamb shift of hydrogen-like uranium at the electron cooler of the ESR", Z. Phys. D, vol. 35, pp. 169-175, 1995.

[27] H.–J. Pross *et al.*, "Lamb shift measurement in hydrogenic phosphorus", Phys. Rev. A, vol. 48, pp.1875-1892, 1993.

[28] R. A. Holt, S. D. Rosner, T. D. Gaily and A. G. Adam, "Lamb shift and fine structure measurements in $^7Li^{+}$", Phys. Rev. A, vol. 22, pp. 1563-1571, 1980.

[29] J. D. Silver, "The spectroscopy of hydrogen-like highly ionized atoms", in "The hydrogen atom", ed. G. F. Bassani, M. Inguscio and T. W. Hänsch, Springer-Verlag, 1989, pp. 221-237.

[30] D. Müller *et al.*, "Autocollimation spectroscopy for fast hydrogenic ions", Z. Phys. D, vol. 18, pp. 249-255, 1991.

[31] E. G. Myers, R. Hankins, J. D. Silver and M. R. Tarbutt, "Precision Lamb shift measurement in hydrogenic nitrogen by fast-beam laser spectroscopy", accepted for publication in Hyperfine Interactions.

[32] J. D. Silver *et al.*, "The Oxford electron beam ion trap: a device for spectroscopy of highly charged ions", Rev. Sci. Instrum., vol. 65, pp. 1072-1074, 1994.

[33] M. A. Levine *et al.*, "The electron beam ion trap: a new instrument for atomic physics measurements", Physica Scripta, vol. T22, pp. 157-163, 1988.

[34] S. N. Lea *et al.*, "Intra-cavity laser resonance spectroscopy of hydrogen-like silicon ions", Phys. Lett. A, vol. 185, pp. 327-332, 1994.

[35] I. G. Brown, J. E. Galvin, R. A. MacGill and R. T. Wright, "Miniature high current metal ion source", Appl. Phys. Lett., vol. 49, pp.1019-1021, 1986.

[36] T. V. Back *et al.*, "Laser spectroscopy of the 2s Lamb shift in hydrogen-like silicon using an electron beam ion trap", Physica Scripta, vol. T73, pp. 62-63, 1997.

[37] H. S. Margolis *et al.*, "Laser spectroscopy using the Oxford electron beam ion trap", Hyperfine Interactions, vol. 99, pp. 169-174, 1996.

[38] M. S. Fee *et al.*, "Measurement of the positronium 1^3S_1-2^3S_1 interval by continuous-wave two-photon excitation", Phys. Rev. A, vol. 48, pp. 192-219, 1993.

[39] R. W. P. Drever *et al.*, "Laser phase and frequency stabilization using an optical resonator", Appl. Phys. B, vol. 31, pp. 97-105, 1983.

EXPERIMENTS WITH A CAESIUM-LOCKED UHF TIME&FREQUENCY REFERENCE TRANSMITTER

Pekka Eskelinen
Lappeenranta University of Technology
PL 20, 53851 Lappeenranta, Finland
phone:358-19-638077, email:ari.eskelinen@pp.inet.fi

1. ABSTRACT

The carrier stability and synchronization of cellular or military mobile frequency agile transcievers, producing a radio signal which is hard to track from a stationary site, can be analyzed onboard with the assistance of a wideband time and frequency transmitter. The principle is to bring through the UHF link a reference comparable to that obtained in laboratory conditions to a moving platform which gives the possibility to test the equipment in real operating conditions and limits the effects of internal couplings within a network. Test results indicate an obtainable timing error below 1 μs or a frequency uncertainty comparable to that of the reference in use with just 20 kHz of RF spectrum down to - 105 dBm RX level when using a low-cost communications receiver and basic modulation schemes. With a 1 W EIRP the maximum range is over 10 km on ground and exceeds 100 km at 3000 meters or above and slow fading or drop-outs are tolerated up to 130 ms.

2. INTRODUCTION

Many modern mobile communication systems are partly based on extremely accurate oscillators forming either their transmitting frequencies or synchronizing their bit patterns or both. Typical uncertainties already implemented include the 10^{-9} level of GSM networks and the super performance frequency agile transceivers in military aircraft approaching 10^{-11} for about two to three hours. The obvious problem is how to verify the performance of individual transmitters or receivers without being limited to the artificial laboratory environment or relying on the highly filtered data available through the respective network management computer. Particularly systems where the base station tries to adjust the mobile to follow its frequency or synchronization pattern, may provide misleading interpretations due to the "rubber band" effect. As long as the geographically large control loop is closed and does not reach its adjustment limit the system will work but we don't know the individual offsets or delays. Some observed related problems are documented in e.g. [1.] and [2.].

Although a number of task-sepcific T&F transceiver systems exist they are either at least partly based on the utilization of satellites generally being out of the direct control of the user segment or on very dedicated expensive hardware. Our attempt has been to evaluate the possibilities of using less costly COTS devices and to find both a system architecture and a simple modulation or coding scheme providing a straight forward way to test mobile oscillators. Typical applications would be analyzing the frequency control characteristics in a GSM1800- system or measurements of frequency hopping radios installed in tactical vehicles, both radio designs producing a signal which is hard to track from e.g. a stationary reference site.

3. THE TEST ARRANGEMENT

The suggested idea has been tested with a cesium-locked transmitter - receiver system where the standard locks the carriers and simultaneously modulates them with e.g. a combination of a sine wave and a 1 PPS pulse train. Thus we can in principle measure frequency deviations, estimate the doppler shift and test the syncronization as well. The enhancement in accuracy is mainly based on the quasi-continuous characteristic of the test signal when compared to the complicated hopping or spread spectrum principles employed in the systems under test and being the only thing to track if a measurement were tried from the stationary site. Besides, a very high momentary resolution can be obtained from the UHF carrier itself which has a roughly 1 ns cycle time. The mobile instrumentation includes a cheap commercial communications receiver, a time interval counter and a phase comparator. If accessible in the transceiver under test, its own reference can be compared e.g. with the regenerated 1 MHz signal but a fractional measurement is naturally also possible.

The transmitter system is presented schematically in Fig. 1. A caesium clock [3.] was used as the main reference in an attempt to be compliant with the high stability requirements of e.g. military radios. However, quite soon after the very first field trials, the stationary HP 5071 opt 001 cesium, see Fig. 2., failed due to tube collapse and it took (the late) Hewlett Packard roughly four months to repair the unit. Besides this, the clock

was "lost" by the local HP representative for several weeks.

Some, perhaps less obvious tricks have been tried to get the most out of the simple receiver which prevents the use of sophisticated modulations or complicated coding algorithms. The standard frequency output of 10 MHz from the cesium is used to lock the transmitter carrier and simultaneously to modulate the transmitter with an optional combination of a 1 MHz sine wave and a 1 PPS pulse train. This arrangement gives us the possibility not only to measure frequency deviations but also allows an accurate estimation of the doppler shift due to the relative motion and further gives a way to test the syncronization by a comparison between the actual bit pattern and the 1 PPS waveform.

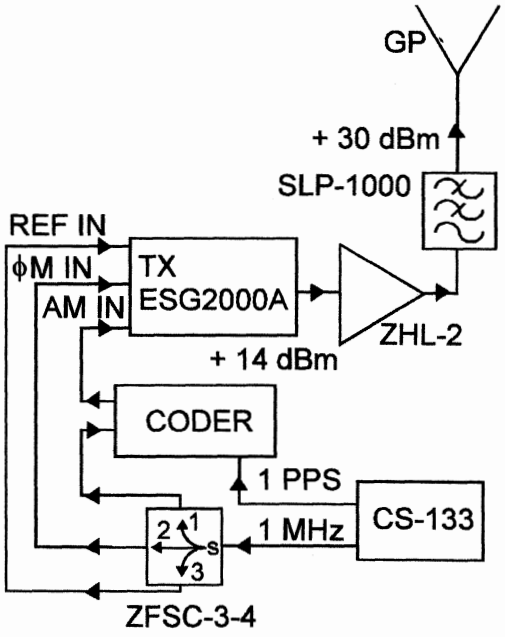


Fig. 1. The test transmitter arrangement includes a caesium clock, a UHF wideband generator and a coder which can be omitted if the 1 PPS signal can be used on the mobile platform as such. The antenna pattern could be directional as well, if needed.

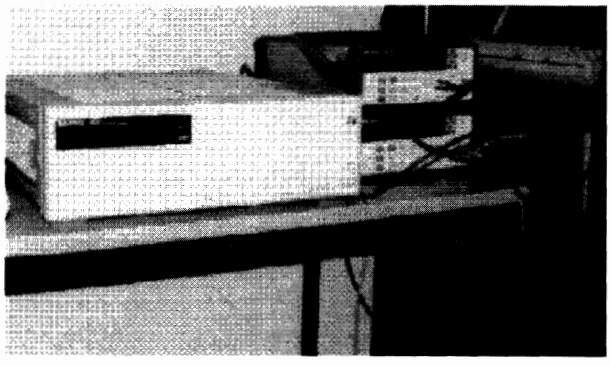


Fig. 2. Here it is, the most unreliable part of the test installation - HP 5071 A opt 001 cesium clock the prologned repair of which caused a four month delay to our field trials.

The measuring equipment on the mobile platform uses an ICOM R7100 receiver, an ovenized crystal oscillator which was documented in [4], a counter and a phase comparator. Fig. 3 demonstrates the instrumentation lay-out in the test van whereas the most vital elements are better visible in Fig. 4. Most of the tests performed up to now have relied on a comparison of either 500 kHz, 1 kHz or the second tic. Key requirements for the "production version" of the RX instrumentation are small size and weight plus a very reliable construction (often a synonym for simple) because of the usually very harsh environment. These were partly not met by the prototype, mainly due to the continuous need of reassembling the equipment into the van for each separate test case.

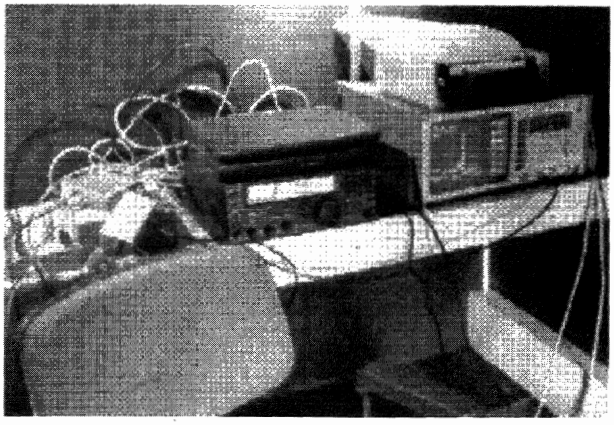


Fig. 3. A view to the test vehicle interior with the TI receiver and the analyzer in the upper corner. Note the somewhat complicated power supply arrangement.

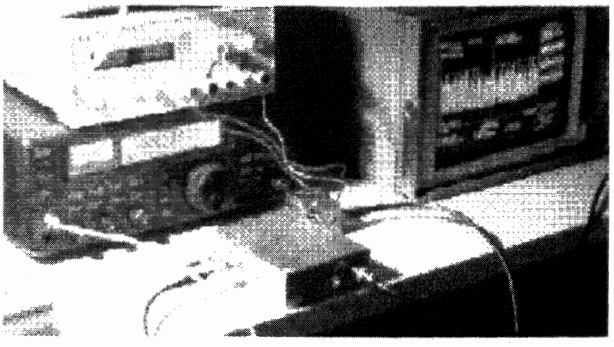


Fig. 4. A close-up of the receiving installation in the lab environment. The phase lock unit is in front of the communications receiver.

4. OPERATING RANGE

Numerous test with both airborne and terrain vehichle mounted receivers indicate a very satisfactory service volume with only + 30 dBm of EIRP. This is partly due to the limited 20 kHz IF bandwidth found suitable for applications where epoch information is not needed but also caused by the low noise figure of an amateur receiver. Fig. 5. is a radar plot from northern Finland showing a test flight with a Hawk jet trainer (Finnish

Air Force) where the top and bottom turning points indicate the sites, about 150 km from the transmitter, where the signal went unusable. The altitude was 3000 meters and the aircraft antenna gain was - 10 dBi. For more aggressive flying, needed e.g. when evaluating equipment characteristics important in ground support tasks (high G values, turbulence), tests have been conducted around 150 meters above average terrain level. Typically the measuring equipment could be used up to 10 km but over hilly areas only to roughly 3-5 km.

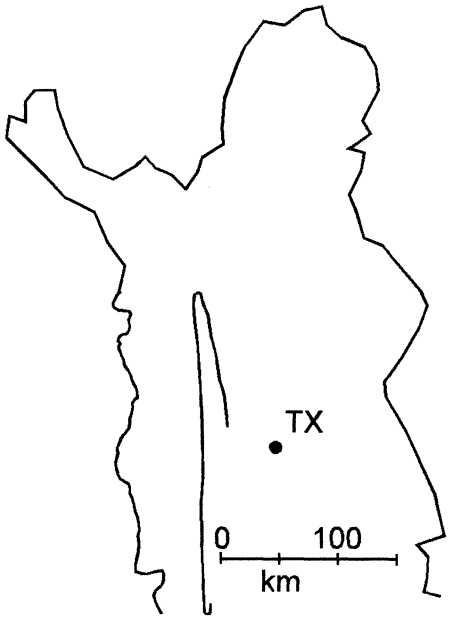


Fig. 5. This radar plot shows the operating range obtained at 3000 meters with just a simple ground plane antenna at the TX site.

5. EFFECTS OF CARRIER LOSS

The first simple configuration tried was to use just pure sinewave modulation and from that extract at the RX site both a frequency reference and time interval information. For this type of system the most obvious problem would be a momentary loss of carrier due to e.g. heavy shading or interference. Because we have to regenerate all the timing information after the receiver, a phase lock seems to be an attractive alternative. Its time constants are crucial for the operation under dynamically changing conditions but are - as known - more or less a compromise with the steady state noise characteristics and partly drift-like phenomena as well.

To start with, Fig. 6 shows the rough test where the carrier was kept silent for one minute which should simulate a reasonably severe propagation problem or even equipment turn-on. As is seen, the design which was optimized for steady state performance, requires about 17 seconds to relock, and what is still worse, the frequency error increases immediately after the swithching of carrier power. Apparently this is avoidable with the addition of a fast hold-mode switch.

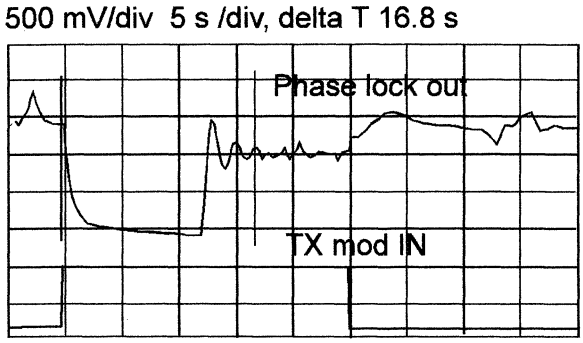


Fig. 6. The high stability phase lock unit unfortunately requires almost 17 seconds to re-stabilize itself after the carrier has been momentarily lost.

In a more realistic scenario we have to accept a continuous stream of interruptions occuring at random intervals. This has been tested with a waveform like the one shown in Fig. 7., where the carrier appears and dissappears consecutively. It is important to note that the state of the lock just before the carrier turn-on is almost as critical as the accumulated frequency error (with reference to the ground oscillator). The typical pattern of 1 second drop outs occuring every ten seconds (approximately, of course) indicates that the selected circuit would be more than 30 percent of the time in an unlocked state.

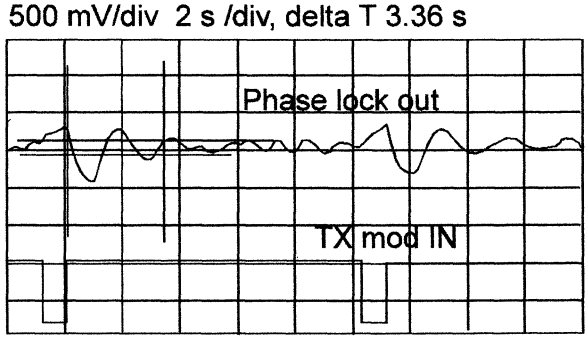


Fig. 7. If the carrier is frequently interrupted, both the ON and OFF time constants become important. With just ten seconds between channel disturbances each lasting 1.0 s it seems adequate to allow 3.4 seconds for stabilization.

If the electronic time constant is adjusted or the RF path can be improved e.g. by using space diversity reception, we are able to do much better, as can be seen in Fig. 8. The system no more notices the absence of carrier if the duration is less than 130 ms. Measured at an average RX input level of -100 dBm the momentary frequency error was here less than 1 kHz (no averaging).

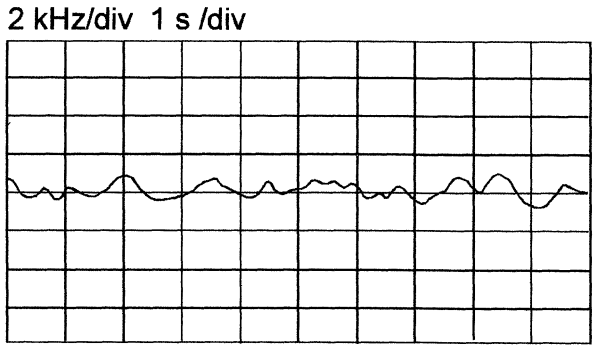


Fig. 8. A further reduction of the carrier OFF time obviously improves the performance. The tested circuit no more noticed the drop-out if it was shorter than 134 ms and the repetition rate was below 100 mHz.

When the UHF path satisfies the predescribed conditions, quite nice performace figures can be obtained by using just 20 kHz of bandwidth. The momentary timing error as obtained from the phase locked crystal oscillator stays at around 500 ns (no averaging) which is clarified is Fig.9. The test receiver input level was kept at - 90 dBm but the signal came via a realistic propagation path having conventional industrial interference and temporary multipath problems due to adjacent moving objects.

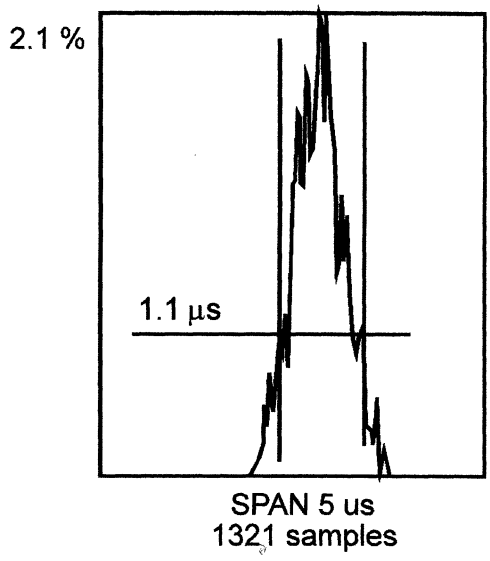


Fig. 9. Phase-locked 1 PPS output at -80 dBm RX level. The receiving bandwidth was only 5.5 kHz, FM modulation.

Long term tests are not very probable with moving oscillators staying within reasonable geographical limits. However, as the necessary instrumentation was there, a several hundred hour recording was taken, this time with the receiver stationary and using the system of [5.] and [6.] as a reference. Obviously we see in the plot of Fig. 10. the error both in the transmitter phase lock and additionally the drifting of the time interval counter's trigger level as a function of elapsed time.

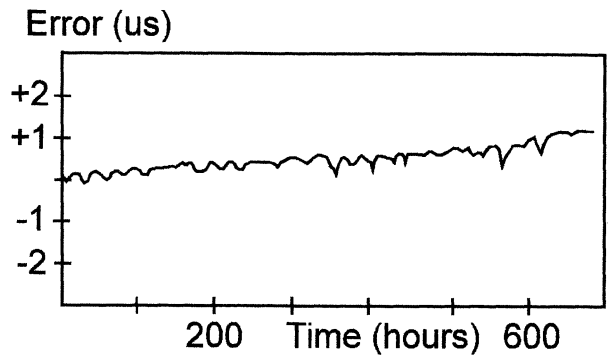


Fig. 10. In the long run, with averaged samples over 100 seconds and a plotted sample spacing of 10 000 s we observe the TX system drift (HP-ESG2000) in this 1 PPS comparison recording at - 90 dBm RX level, FM demodulation and 100 kHz IF bandwidth.

6. COMBINED MODULATIONS

After a satisfactory level was achieved with the continuous FM scheme, a trial was made to include absolute timing information to the transmitted signal as well in an easy form for the receiver. The main problem was to keep the signal within the bandwidth allocated by the telecommunication authorities and at the same time maintain the highest possible timing resolution. It turned out to be possible just to pulse modulate the test generator so that the carrier was suppressed by - 80 dB for 50 microseconds and to detect this from the same audio output where the previous phase locking information came out. To have the sharpest turn-over point for triggering, a careful evaluation of the receiver's AGC performance was carried out. The optimum result is shown in Fig. 11. as the first derivative of the input waveform. Naturally, due to the switching nature of the pulse modulated signal, we had to use the widest available bandwidth of the receiver which caused a 10 dB increase in the pre-detection noise floor.

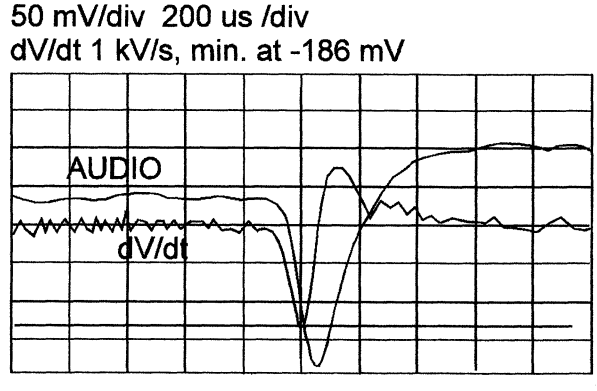


Fig. 11. If using simple carrier keying, an optimum TI reference point was found at the point of highest rate of change of the RX audio output. RX level was -80 dBm, FM detection and 100 kHz IF bandwidth.

The communications receiver was able to simultaneously detect both the quasi-continuous FM signal and the 50 μs absence of the carrier to form a separate 1 MHz frequency output and a true 1 PPS information. However, partly due to the required increase in the IF bandwidth and partly due to occasional spikes in the audio waveform, the uncertainty became a compromise showing a momentary value near 10 μs, as is clarified in Fig. 12.

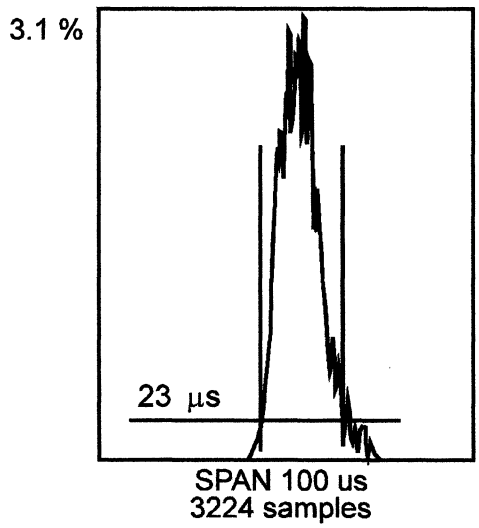


Fig. 12. The cost of epoch information is the worse-than-expected time interval uncertainty. Under otherwise similar conditions than in Fig. 9 the error is twentyfold.

7. CONCLUDING REMARKS

It seems quite possible to assemble a working T&F transmitter - receiver setup for ad hoc tests from common RF hardware without special and generally expensive synchronization equipment. Current results with a low-cost communications receiver and some simple external electronics suggest a worst-case (99.9 %) sync error less than 1 μs, a frequency uncertainty approaching that of the reference with a relative drift of $5.5^{*}\ 10^{-15}$ and a timing resolution better than 10 ns, obtainable down to - 105 dBm received signal level but all this only if there is not a need to transfer the epoch infromation to the mobile installation. The present system produces, depending on selected modulation, a complex spectrum width of 20 kHz, extending to 10 MHz if pulse modulation is added, between the - 60 dBc points. With a 1 W EIRP the maximum range is over 10 km on ground and exceeds 100 km at 3000 meters or above and slow fading or drop-outs are tolerated up to 130 ms. UHF carrier frequencies look most promising due to their propagation characteristics and reasonable antenna alternatives but if available, a licence for at least 1 Mhz of spectral width should be acquired.

9. REFERENCES

[1] de Jong G. and Lewandowski W.,"GPS Time Transfer and the Problem of the Determination of Receiver Delays", *Proc. 12th EFTF'98*, pp. 79-86

[2] Davis J., "Progress Towars Establishing GPS Disciplined Oscillators as Standards Traceable to UK National Time Scale", *Proc. 12th EFTF'98*, pp. 112-116

[3] *HP 5071A Primary Frequency Standard*, HP Product Note, Hewlett Packard, 1996

[4] Eskelinen P.,"A low-cost ovenized crystal oscillator", *Proc. 9th EFTF'95*, pp. 357-360

[5] Eskelinen P.,"Comparison of Cs-, H-maser and GPS-based Timescales",*Proc. 11th EFTF'97*, pp. 477-480

[6] Eskelinen P.,"Tests with externally applied joint electronic and numerical control of commercial atomic clocks", *Proc. 12th EFTF'98*, pp. 54-58

NOVEL HIGH PERFORMANCE SATELLITE TIME & FREQUENCY DISTRIBUTION AND SYNCHRONISATION SYSTEM

G. Garofalo[(1)], G. Busca [(2)]

(1) ESA, Noordwijk, Netherlands, +31 71 656 3054, ggarofal@estec.esa.nl

(2) Observatoire Cantonal Neuchatel, +41 32 889 6870, giovanni.busca@ne.ch

Abstract

The present article, based on ESA Patent 407, describes a novel technique aiming at transferring to a low performance remote clock in space, the time and frequency stability of a state of the art ground atomic clock.
The technique is based upon a standard communication system, a pulsed on board PLL, and a direct measurement, via the communication system itself, of the space-ground clock delay. Preliminary measurements have shown that the on board tracking phase Jitter can be <1° at the clock frequency i.e. the system performances can reach the sub-nanosecond level.
Some potential applications of the proposed new transfer technique are:

- Generation of a world wide real time clock reference in space based upon state of the art ground atomic clock without having to physically place complex, heavy and difficult to maintain atomic clocks on board.
- Capability to perform accurate Time and Frequency comparison with less accurate ground-based clocks for industrial calibrations or real time clock applications.
- Precise synchronisation of switches and networks, wherever located in the world, for digital communications purposes.
- Navigation and positioning in the vicinity of the earth and in deep space.
- Comparison of widely separated ground clocks with state of the art performances (for example: next generation of primary Caesium clocks).

1. System Overview

The (minimum) reference system configuration which allows implementing the method we propose is depicted in Fig. 1 and consists of:

1) one or more Master Stations equipped with a clock of the highest performance conceivable today with the following characteristics:

Ground Reference Clock Characteristics

$\sigma_y(\tau)$	τ
2×10^{-10}	10^{-4}
2×10^{-11}	10^{-3}
2×10^{-12}	10^{-2}
2×10^{-13}	10^{-1}
1×10^{-13}	10^{0}
1×10^{-14}	10
3×10^{-15}	100
1×10^{-15}	1000
3×10^{-16}	10000
1×10^{-16}	100000 (~ 1 day)
1×10^{-16}	$\tau > 10^{5}$

The characterization of clock frequency stability is given as usual in term of the square root of the two-sample Allan Variance:

$$\sigma_y(\tau) = \sqrt{\Sigma_{i=1}^{N} \frac{(\bar{y}_{(i)} \bar{y}_{i+1})^2}{2(N)}}$$

$\bar{y}$ (τ) being the frequency averaged over the time internal τ, N is the number of (the frequency averaged) samples.

The ground reference clock is a composite clock that realizes the best conceivable combination for short, medium and long-term performances corresponding to:

- state of the art quartz crystal oscillator for $10^{-4} < \tau < 10^{1}$ s
- Hydrogen Maser for $1 \leq \tau \leq 1000$s
- cold Cs clock for $\tau \geq 1000$s

2) one synchronization module placed on a geostationary satellite made of:

a) TDMA demodulator
b) clock module, which includes an oscillator with much lower frequency stability than the ground clock, and a data processor for lopback information to the ground
c) TDM modulator

The up-link transmission rate is determining the overall system time stability which can be achieved (see table 1). However the method as such is frequency independent, i.e. it is limited only by the state of the art technology for phase tracking of both on board module and ground station module. This means that the results shown on Table 1 may evolve when better technology allows smaller peak-to-peak phase jitter for on board tracking (e.g. faster comparators, more accurate equalisation of delay paths within the clock module, higher clock frequency).
Table 1 gives an indication of the relative frequency error over one day that can be achieved with this approach in relation to the up-link access rate.
Values of phase jitter less than 1ns have been measured in a laboratory implementation for simulating the key elements of the system where a Caesium Reference Clock for the ground station and a simple quartz oscillator ($\Delta f/f \cong 10^{-6}$) for the on board module were used.
The on board module breadboard was based on a digital PLL and the clock rate utilised was 32.768 MHz.
On the basis of this experiment a further reduction of peak to peak phase difference between the up link clock and the on board clock is deemed possible up to the level 0.1° which is to-day the theoretical technology limit.

This could be achieved by implementing the following improvements to the module:

- use of an Ultra Stable Oscillator as the space clock. This clock should have much improved performances compared to the simple oscillator present in the simulation experiment.
- reduction of loop noise bandwidth
- use of longer clock window duty cycle
- Technological improvement of the phase comparator set up/hold time, D/A resolution, and loop filter gain.

2. Method Outline

The method here proposed is based upon the availability of a very high performance clock at one or more system main stations and of an oscillator placed on board the satellite in a special synchronization module that will be described hereafter.
The reference system for this explanation is supposed to be based on TDMA system running via a geo-stationary satellite.
The synchronization module (see fig. 2) is based on a digital PLL that incorporates: a TDMA Demodulator, a TDM Modulator and a Clock Module including: phase comparators (one per input clock), low pass filter, VCXO (based on an Ultra Stable Oscillator USO) and on board data processor for feedback information to the ground master station.
The main stations ground clocks can be initially considered as reference clocks, even if they are not synchronized with one another. The mutual synchronization will occurs automatically by means of the method described hereafter, realizing one of the objectives of this method.
As shown in fig. 2 the output of the on board TDMA Demodulator provides the data transmitted by the ground main stations and also a clock which is recovered by the Demodulator and is used in the Clock Module as source for synchronization.
The goal of this method consists in obtaining a very accurate phase synchronization between the ground main station clock and the on board clock. The achieved degree of accuracy between these two clocks is such to allow users to rely on the on board clock as system reference clock instead of ground clock. In fact the system will maintain a phase synchronization better than 1° between the ground clock and the on board clock.
A peculiar characteristic of the proposed method is that the clock signal recovered by the on board TDMA Demodulator and produced by the ground main station is used in gated mode, i.e. not continuously.
This implies that only a fixed number N of clock cycles, for each TDMA frame, are used for phase estimation. These N clock pulses are contained in a clock window shown in Fig. 3. The clock window follows the UW window and is independent from the on board burst arrival time variation which only affects the UW position in the frame but not the content of the clock window.
The average phase associated to the N clock pulses contained in the clock window is then compared with the on board oscillator phase and any difference between the two is used in the PLL to reach and maintain the phase difference below 1°.

The method is articulated into four consecutive Phases.

Phase 1: Acquisition Phase

In this initial phase, the on board clock that is in free running mode, has most likely a slight frequency difference with the ground main station clock but, most important, there is an uncertainty on the knowledge of the actual satellite position that needs to be recovered in order establish a common time base between space and ground, which is the basis for this method. The loop back mechanism related to the burst time of arrival makes sure that this frequency difference and the delay are properly recovered. At the end of this phase the burst transmitted from the main station is received on board at nominal time and also the on board and on ground frequencies coincide.
In order to do this an appropriate bandwidth is used for the On board module Low Pass Filter. The loop back information, regarding the on board Time of Arrival of the main station burst is contained in a specific field of the continuous TDM down link frame. This information is recovered by the ground station and used to implement the appropriate corrections to its burst transmission instant until the Acquisition phase is over.

Phase 2: Synchronization Phase

In this phase the on board PLL works in tracking mode, and a much narrower bandwidth is used for the PLL Loop Filter. This allows reaching and maintaining a very accurate phase synchronization between on board and on ground clock.
The phase comparison is done in the Clock Module every TDMA frame (whose duration is in the order of ms). While the space clock is locked in phase to the time shifted ground clock burst, the ground station measures the 2-way delay suffered by the ground clock burst and computes by simple models the one-way delay. This information is used for shifting the ground clock burst timing exactly of the one-way delay. This operation can be performed every TDMA frame. As a result the space clock will run in synchronization with the ground clock within an error of $\leq 1°$ of phase (at the clock frequency). The on board clock can now be used as system reference clock, with the same characteristics and performances of the main clock on ground as long as the tracking of the two clock holds. The on board reference clock is also driving the multitude of on board clock modules used for the TDMA communications.

Phase 3: Synchronization of two or more distant ground clocks via the on board clock.

At this point, a second stations having the same clock characteristics as the first main station, can access the satellite and, after following similar steps as depicted in Phase 1 and 2, can synchronize the associated on-board clock module. The time difference between this clock module and the on board reference clock represent the time difference between the two stations. This information is transmitted to the station which can than be synchronized to the first one. The process can continue infinitely and the synchronization of a multitude of ground stations becomes possible.
This fulfills the main goal of the patent: achieving Time and Frequency dissemination between distant points.

Phase 4: Clock dissemination from space.

At this point a very reliable, stable and accurate clock source is available on board, without having to physically place a high performance atomic clock on the spacecraft.

3. Conclusion

The reference on board clock can be disseminated to a variety of users, which range from frequency calibration users, to telecom network operators or to mobile users. The detail definition of phase 4 depends on the type of system configuration, which in this explanation was assumed to be a geo-stationary satellite. More applications can be covered by constellations of satellites whether in a fixed geo position or orbiting round the earth. For assessing the final performances of the method outlined when used in a specific application, a further investigation of the symmetry of the up-down delay need to performed including the contribution from the ionspheric effects as well as geometrical effects.

References:

Ref. 1: "Procede de synchronisation entre une horologe de reference d'une station au sol et une horologe d'au moins un dispositif distant", ESA Patent 407, by G. Garofalo and G. Busca.
Ref. 2: "The Science of Timekeeping", Hewlett-Packard Application Note 1289, by D.W. Allan, N. Ashby, and C. Hodge.
Ref. 3: "In The Space Age", IEEE SPECTRUM, March 1998, by D.W. Allan, N. Ashby, and C. Hodge.
Ref. 4: "Digital Communications", by K. Feher, Prentice-Hall 1981.
Ref. 5: "Digital Communications by Satellite", by J.J. Spilker, Prentice-Hall 1977.
Ref. 6: "Brown, R.C.: A Textbook of Physics". Longmans, London, 1961, p. 1013.
Ref. 7: "Closed-loop Doppler Correction Scheme". NASA Report X-733-67-282/N67-31384, June 1967.
Ref. 8: "NAVSAT Time Synchronization Simulation Study", ESA Contract 5805/84 NTSSSTU.

U/L CLK [KHz]	On Board Oscillator CLK [KHz]	On Board Output CLK [KHz]	O/B Tracking Phase Jitter peak-peak [°]	Reference CLK versus On Board CLK Relative frequency error (ε) over 1 day averaging time
384	384	384	< 1° < 0.1°	$\varepsilon = 8.37 \times 10^{-14}$ $\varepsilon = 8.37 \times 10^{-15}$
1,024	1,024	1,024	< 1° < 0.1°	$\varepsilon = 3.14 \times 10^{-14}$ $\varepsilon = 3.14 \times 10^{-15}$
2,048	2,048	2,048	< 1° < 0.1°	$\varepsilon = 1.56 \times 10^{-14}$ $\varepsilon = 1.56 \times 10^{-15}$
4,096	4,096	4,096	< 1° < 0.1°	$\varepsilon = 7.85 \times 10^{-15}$ $\varepsilon = 7.85 \times 10^{-16}$
16,384	16,384	16,384	< 1° < 0.1°	$\varepsilon = 1.96 \times 10^{-15}$ $\varepsilon = 1.96 \times 10^{-16}$
32,768	32,768	32,768	< 1° < 0.1°	$\varepsilon = 9.8 \times 10^{-16}$ $\varepsilon = 9.8 \times 10^{-17}$
131,072	131,072	131,072	< 1° < 0.1°	$\varepsilon = 2.4 \times 10^{-16}$ $\varepsilon = 2.4 \times 10^{-17}$

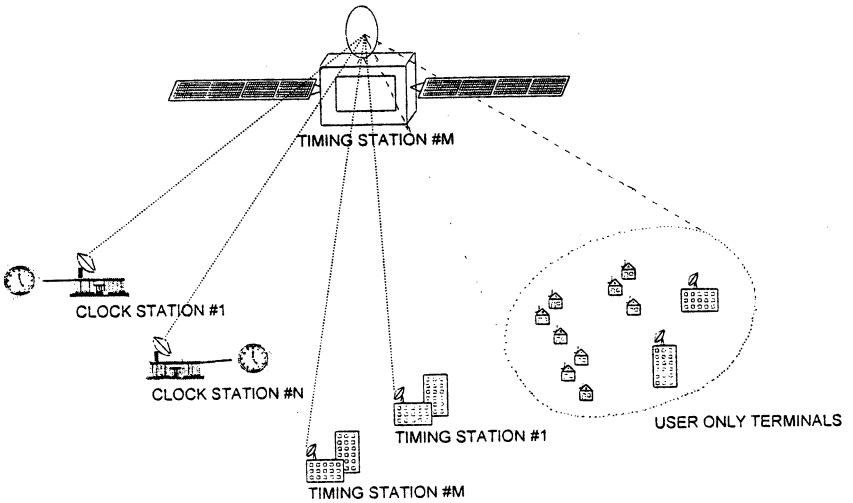


Figure 1

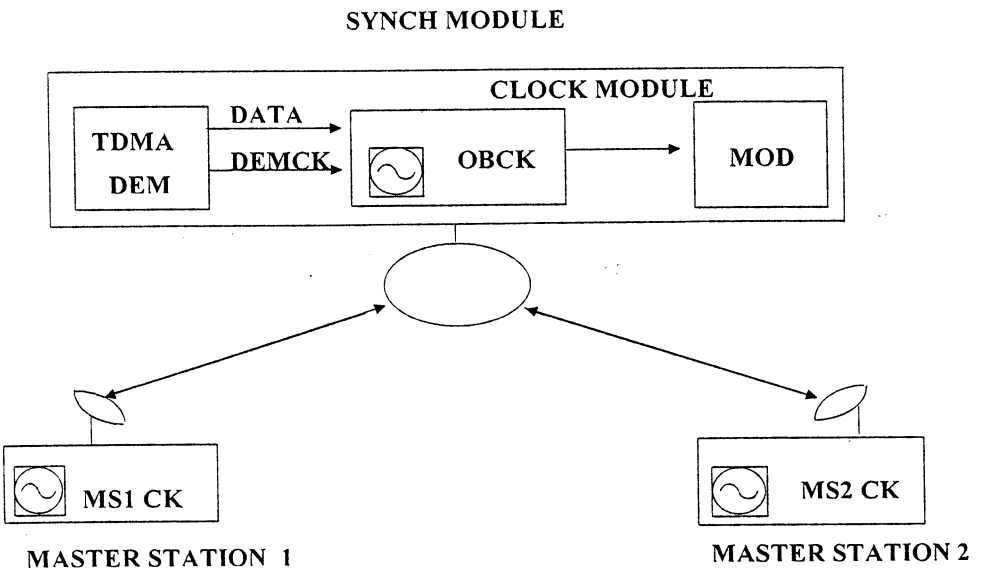


Figure 2

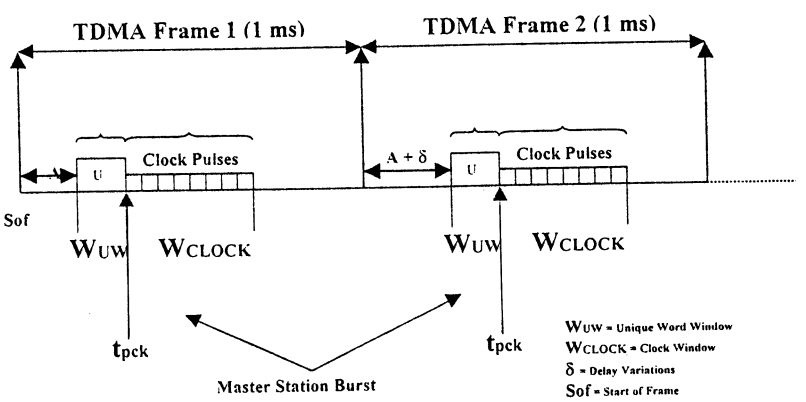


Fig. 3

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

Latest Developments from the CCTF Working Group on Two Way Satellite Time and Frequency Transfer (TWSTFT)

W. J. Klepczynski
Innovative Solutions International
Vienna, Virginia

This report will discuss the latest accomplishments of the CCTF WG on TWSTFT. It will include a brief discussion on the inclusion of TWSTFT data in the formation of TAI, latest developments in the Pacific Rim, the status of stations participating in the WG and future plans.

RESULTS OF THE RECENT INTRODUCTION OF IGS IONOSPHERIC CORRECTIONS FOR SEVERAL TAI TIME LINKS

P. WOLF, G. PETIT
BIPM, Pavillon de Breteuil, 92312 Sèvres CEDEX, France

Several IGS (International GPS Service) analysis centres provide now ionosphere products in the form of global ionosphere TEC (Total Electron Content) maps. Typically, these maps contain the zenithal TEC as a function of longitude and latitude in a 5° by 2.5° grid, renewed every two hours. Starting in August 1999 the ionosphere products of the IGS CODE (Centre for Orbit Determination in Europe) analysis centre are used for the ionospheric correction of several long and medium distance time links in the TAI network. We show the improvements in stability and accuracy resulting from the use of the IGS products rather than the standard STANAG (STANdardisation AGreement) model for the ionosphere. When compared to ionospheric corrections from on site dual-frequency measurements with non calibrated receivers (the method previously used for two long distance links) we see roughly equivalent stability, but improved accuracy and reliability, due to the adjustment of satellite and receiver biases in the IGS solution and to the large number of receivers used by the IGS. We expect that in the near future several more time links will be corrected for ionosphere using the IGS products which should prove advantageous especially in light of the up-coming strong ionospheric activity due to the solar maximum expected around 2001.

Corresponding author:
P. WOLF
BIPM, Pavillon de Breteuil, 92312 Sèvres CEDEX, France
Phone: 33-1-45077075
FAX: 33-1-45077059
e-mail: pwolf@bipm.fr

PREPARATION, CHARACTERISATION AND MICROWAVE DIELECTRIC PROPERTIES OF $RETiNbO_6$ [RE=Ce, Pr, Nd, Sm, Eu, Gd, Tb, Dy, Y, Yb] CERAMICS

R. RATHEESH[1], M.T.SEBASTIAN[2], S. SOLOMON[2] AND P. MOHANAN[3]

[1]Institute of Chemistry, University of Osnabrueck, Barbarastaße 7, D-49069 Osnabrueck, Germany

[2]Structural Ceramics, Regional Research Laboratory, CSIR, Thiruvananthapuram, Kerala India

[3]Department of Electronics, Cochin University of Science and Technology, Cochin, Kerala, India

Microwave ceramic dielectric resonators (DRs) based on $RETiNbO_6$ [RE=Ce, Pr, Nd, Sm, Eu, Gd, Tb, Dy, Y and Yb] have been prepared by the conventional solid State ceramic route. The DR samples are characterised by XRD, SEM and spectroscopic methods. The microwave dielectric properties are measured by resonant methods using a Network Analyser. The microwave measurement results show that the aeschnite-priorite type ceramics with RE= Ce, Pr, Nd and Sm show high ε_r in the range 44 to 55 and positive τ_f. The euxenite-polycrase type ceramics with RE= Gd, Tb, Dy, Y and Yb have low dielectric constant in the range 19 to 22 and negative τ_f. The unloaded quality factor (QXf) of the samples are in the range 5000 to 20000. Among the materials under study, $EuTiNbO_6$ has high quality factor and low temperature variation of resonant frequency and is a possible material for microwave resonator applications. The dielectric constant and temperature variation of resonant frequency of $EuTiNbO_6$ ceramic is in between that of aeshynite and euxenite group of materials. The structure and dielectric properties of these materials depend on the ionic radius of the rare earth ion. The XRD, Raman and FTIR studies reveal a symmetry different from that of aschynite and euxenite group for $EuTiNbO_6$

Corresponding Author:
Dr.R.Ratheesh
Inorganic Chemistry Division
Institute of Chemistry
University of Osnabrück
Barbarastraße 7
D-49069
Osnabrück
GERMANY

Telephone: +49-541-969-2782
Telefax: +49-541-969-2370
E-mail:ratheesh@uos.de

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

A PRACTICAL METHOD TO DETERMINE THE OH CONCENTRATION IN QUARTZ CRYSTALS

M.A. PASQUALI, J.J. BOY
ENSMM / LCEP - 26, Chemin de L'Epitaphe - 25000 BESANÇON – France
Fax : 333 81 88 57 14 - E-mail : jjboy@ens2m.fr

ABSTRACT

We develop here a new method to determine the hydroxyl concentration in quartz crystal. The procedure is deduced from the Paterson method and it consists in obtaining an integral absorption factor **I**, expressed in $cm^{-1}/[N/10^6$ Si], which relates the OH concentration (noticed [OH]) with the integral surface under the IR absorption spectra and the α_{3500} coefficient. This method has been applied to different kinds of samples issued from synthetic and natural quartz and submitted to various treatments as sweeping, irradiation or annealing. A calibration line is obtained for each kind of sample and the [OH] is deduced directly from the curve.

Key-words : IR-FT, OH concentration, quartz crystal.

1 - INTRODUCTION

In the last fifty years, several publications have been written about the hydrogen in quartz. The hydrolytic weaking as well the frequency shifts of electronic quartz components are examples of technological purposes supported by these investigations **[1, 2]**. The hydrogen content is ranged from 10 to about 200 μg/g for cultured quartz used in electronic applications. In natural species as amethyst and some cultured quartz, this rank can achieve an important magnitude (~ 1000 μg/g) **[3]**.

The hydrogen in quartz can be found in many defects centers as :

- $[H_4O_4]^0$ called hydrogenic groups,
- OH^- related defects, as Al-OH, Li-OH and Na-OH,
- $[Si_2O/H^+]$ (atomic hydrogen) where H^+ is trapped by an oxygen site in the *c* axis channels,
- paramagnetic and diamagnetic species as $[AlO_4/H^+]^+$, $[Si_2O/H^+]^+$ and $[AlO_4/H^+]^0$ **[4 - 7]**.

As it can be noticed, the hydrogen in quartz most frequently occurs bonded to oxygen. The resulting OH bond is highly polar and as consequence this dipole absorbs the infrared radiation. Thus, IR spectroscopic methods represent an important tool for the investigation of the OH^- related defects. Usually, IR spectrum of quartz shows as feature a broad absorption band in the region of 3 μm. This band is superimposed upon several sharp bands in the same region, as shown in Figure 1. The sharp absorptions are related to the specific hydrogen point defects as Al-OH and M-OH (M= Li, Na, K...). Then, two kinds of IR absorptions can be identified in the region of 3μm : The broad band and the sharp absorptions. The sharp bands are extensively commented in the specialized literature **[8 - 10]**. Comparatively the investigations about the broad OH band are more restrictive and they have been concentrated in the fields of the deformation and plasticity studies on quartz and its correspondent hydrolytic weakening **[1, 11]**.

The nature and the structure of the defect correlated with the broad absorption band is not yet well understood. For some authors, it has been associated with the group Si-OH **[11]** and for others, with microbubbles of liquid water (liquid inclusions) or clusters of water molecules **[7]**. In some recent studies, it has been proposed the $(4H)_{Si}$ (or $[H_4O_4]^0$) and the H_2O interstitial defects as, maybe, precursors of the 3580 cm^{-1} OH absorption band **[4, 12]**. In this case the substitutional $(4H)_{Si}$ and the H_2O interstitial defects could be included as a kind of defects related with the broad absorption band.

Paterson in **[13]** has proposed a formula to evaluate the hydroxyl (OH^-) content related to this broad absorption band appearing in the 3 μm region of the IR spectrum. So, it is possible to estimate the OH content associated with other defects in spite of the specific impurities as Al, Li, Na, K, Ag, Cu, ...

Our work is based on the application of Parterson's equation in order to evaluate the water-related species associated to the broad OH band. Quartz samples of different origins (natural and cultured crystals) have been selected. After the application of Paterson's formula by numerical integration in the region of 3700 to 3100 cm^{-1}, a calibration line is obtained for each quartz material. The acquisition of this line is also based on the evaluation of the absorption coefficient (α_{3500}) and the area of the spectrum.

2 - METHODOLOGY

The infrared results have been obtained by using a single beam NICOLET Magna 750 FT-IR spectrometer. The samples were put inside a cryogenic cell with CaF_2 windows and cooled to the nitrogen liquid temperature. The spectra were recorded at 0.125 cm^{-1}

resolution at least and 250 scans were taken for each spectrum. The results were evaluated by the OMNIC software : It gives the area of the spectrum and the absorption of each sharp band.

The methods to estimate the bulk hydrogen in quartz or the water content are restricted. In fact they are based on the determination of an absorption coefficient for a specific wavenumber and a molar absorptivity constant from a standard substance **[7]**. Also, some procedures have been evaluated by the determination of an integral absorption in the spectral region of the broad band **[3]**. Others have applied, directly, the value of α_{3500} coefficient to determine the hydrogen content **[14]**. As Paterson, to evaluate the OH content he proposes in 1982 a procedure that takes into account the area of the spectrum. This method is based in the determination of a calibration line to various glasses, quartz, silica, muscovite, pure water and others. The following relationship is obtained :

$$\frac{I}{\gamma} = 150 * (3780 - w) \qquad [2.1]$$

where :

I= integral molar absorption coefficient (cm^{-2}/(mol H/l)

γ = orientation factor (taken equal to 1/3, for isotropic OH orientation distribution).

By means the formula :

$$C = \frac{K}{\varepsilon} = \frac{1}{I}\int K(w)dw = \frac{\Delta}{I} \qquad [2.2]$$

where : C is the molar concentration

K : absorption coeff. at a specific wavenumber w (cm^{-1})

ε : molar absorption coefficient

and Δ, the integral absorption coefficient.

For each w (cm^{-1}) corresponds a K(w) taken from the IR spectrum in the interval Δw. The absorption coefficient can be determined by the Beer-Lambert's equation :

$$K(w) = \frac{1}{t}\log\left(\frac{T_B}{T_A}\right) \qquad [2.3]$$

After the substitution of equation [2.1] into [2.2] and by numerical integration of the entire broad band, we have :

$$C = \frac{1}{50}\sum \frac{[\log(T_B / T_A)] / t}{(3780 - w)}\Delta w \qquad [2.4]$$

C = [OH]

Δw = integration interval (cm^{-1})

t = thickness of the sample (cm)

T_B = transmission (%) on the base base line (Figure 1)

T_A = transmission (%) on the absorption line (Figure 1)

The Paterson's expression is applicable in the interval of 3780 to 2600 cm^{-1}. For this work we have used it in the range 3700 to 3100 cm^{-1}, but in removing the OH strength absorption of each sharp band (Fig. 1).

The numerical integration has been realized with the Basic software. Earlier, other authors have employed this method to evaluate [OH] **[15, 16]**.

The values T_A and T_B are taken from the IR spectrum after defining the two absorption and base lines.

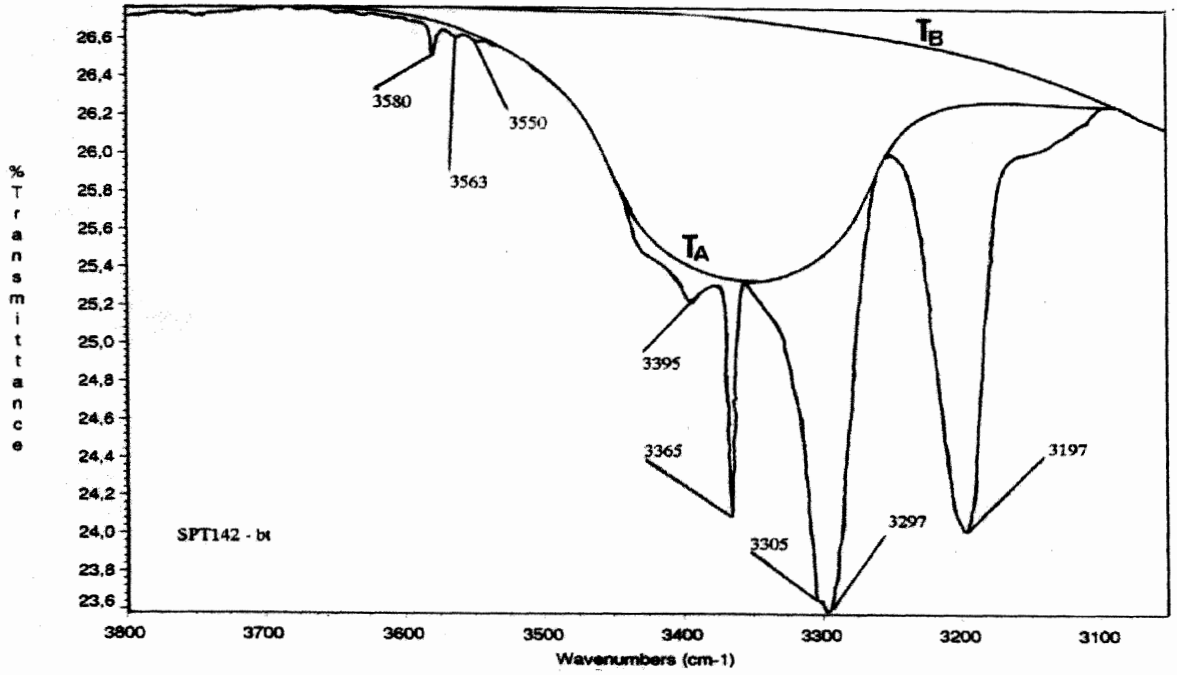


Figure 1 - Low temperature IR spectrum of synthetic sample - Broad OH Absorption Band.

Table 1 gives an example of [OH] determination by numerical integration. The interval Δw has been chosen, arbitrarily, as 50 cm^{-1}. The results can be expressed in mol H/l, μg/g (ppmw) or in relation to the Si atoms, [N/10^6 Si] (ppma).

In order to evaluate the method proposed in this work, other important results should be obtained. They consist in the determination of the areas of the broad-band and of each sharp band, the evaluation of the α_{3500} coefficient and this of **I** which is the integral absorption factor (cm^{-1} /[N/10^6Si]) that relates the [OH] with the effective area and the α_{3500} coefficient. This last coefficient is calculated as :

$$\alpha_{3500} = \frac{1}{t} \log \frac{T_{3800}}{T_{3500}} \qquad [2.5]$$

where,
T_X = transmission (%) at X cm^{-1}
t = sample thickness (in cm)

Table 1 – IR transmission of the sample of the Fig. 1

IR Transmission - [0H]		
W (cm^{-1})	**TB (%)**	**TA (%)**
3700	26.74	26.74
3650	26.74	26.72
3600	26.74	26.70
3550	26.74	26.60
3500	26.74	26.40
3450	26.74	25.90
3400	26.74	25.40
3350	26.70	25.32
3300	26.64	25.48
3250	26.54	26.04
3200	26.45	26.19
3150	26.36	26.22
3100	26.26	26.26
[OH] (mol H/l)	9.66×10^{-4}	
[OH] (μg/g)	0.37	
[OH] [N/10^6Si]	21.88	

Then, we determine the [OH] as follow (Item 1 to 6). The numerical application is deduced from our first example :

(1) **Total Area /t =25.263 cm^{-2}** which was determined by the OMNIC software. This total area corresponds to the surface under the spectrum in the absorbance condition. In order to obtain the cm^{-2} unit, measured areas should be normalized in dividing by the path-length (t). From the equation [2.2], we should note that the integral k(w).dw corresponds to the area evaluated by the OMNIC software.

(2) The broad band area is designated as "effective area" and it represents the total spectrum surface for which the surfaces of each band are deduced. Here we find : **13.296 cm^{-2}.**

Table 2 - IR Measurements - Absorbance for the Spt_142 sample (with t = 0.266 cm)

Band (cm^{-1})	**Area/t** (cm^{-2})
3580	0.034
3563	0.004
3550	
3395	0.538
3365	0.462
3305	0.132
3297	5.267
3197	5.530
Total	11.967

(-) too weak strength

(3) α_{3500} coefficient is determined by the equation [2.5]. Here we find : α_{3500} **= 0.017 cm^{-1}**

(4) We call **Ref**, the ratio between the broad band area of spectrum obtained by the OMNIC software and the α_{3500} absorption coefficient. We have noted a smooth linear relation between these parameters. For our sample, we obtain : **Ref = 782 cm^{-1}**

(5) The OH content is evaluated by the Paterson's method (see Table 1).

(6) $I = \frac{Ref}{[OH]} = 35.74$ cm^{-1} / [N/10^6Si]

Here we propose the determination of the integral absorption coefficient, **I**, from the relation of OH content and the **Ref**. In fact, there is an intensive correlation with the [OH] evaluated by numerical integration and the α_{3500} coefficient. In order to improve the relation above, the "effective area" of the spectrum should be determined.

Originally, the units of **I** are cm^{-2}/ [OH], where the [OH] can be expressed in mol H/l, ppmw or [N/10^6Si]. As indicated by the evaluation proposed above, **I** will be given in cm^{-1} per [OH].

The OH content is obtained indirectly by the calibration line, where **I** and **Ref** values are plotted.

$$\text{Then, } [OH] = \frac{Ref}{I}$$

3 – SAMPLES DESCRIPTION

In almost samples, we have applied different kinds of treatments as irradiation, annealing and electrodiffusion.

Electrodiffusion, called too sweeping, was carried out by applying an electric field of 1500 V/cm along the c-axis of the crystal, when the temperature of 500°C was reached. Two conditions were employed : air and inert (N_2) atmospheres.

For natural quartz, the irradiation doses used are the following : .6, .8 and 1 Mrad.

At least, the heat temperature applied for the annealing was 350 °C under inert conditions.

So, the calibration lines were obtained from the following samples submitted to the above conditions :

- natural quartz, TQ : AT-cuts with 0.5 and 0.15 cm of thickness ;
- cultured quartz, SP, identical to the previous ones ;
- cultured quartz from SAWYER (SW), Y-cut, 0.5 cm of thickness cut in Z and +X growth regions. The samples were exposed at 100 Krad radiation (from Cesium source) and an additional dose of ionizing radiation (from X-rays source) was applied ;
- various brazilian natural quartz Z-cut samples with a thickness of 0.5 cm. These quartz samples were submitted to the same treatment as described for TQ quartz.

4 – RESULTS

The calibration curves are presented in the figures 2 to 5. We can note a good linearity between the numerous experimental measurements in each quartz quality and so a good accuracy. The average fitting factor is about 0.95. Furthermore, this precision can increase if the differences between the samples were reduced as its thickness and its origin.

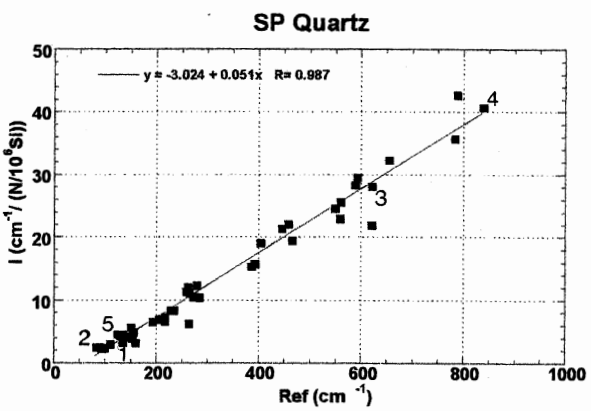


Figure 2 - [OH] Calibration Line – SP synthetic quartz.

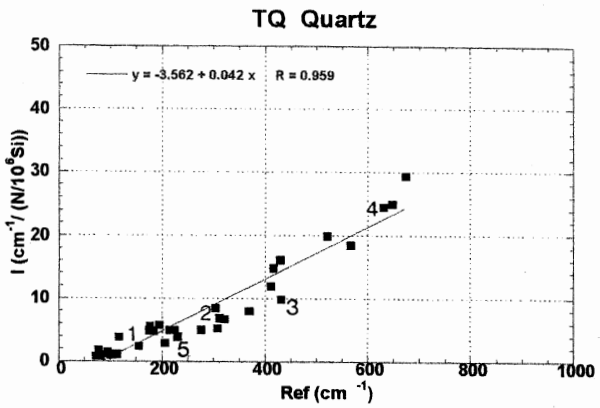


Figure 3 - [OH] Calibration Line - Nat TQ quartz.

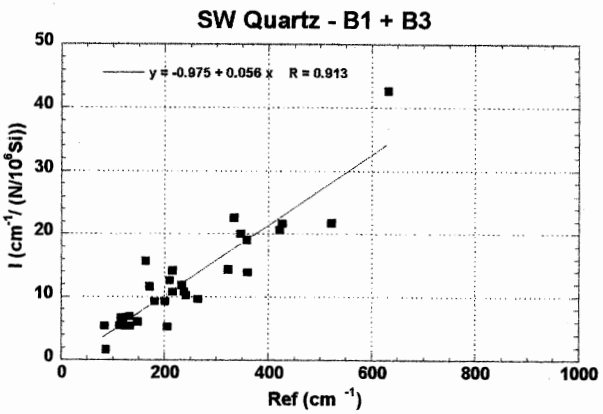


Figure 4 - [OH] Calibration Line - Sawyer synt quartz.

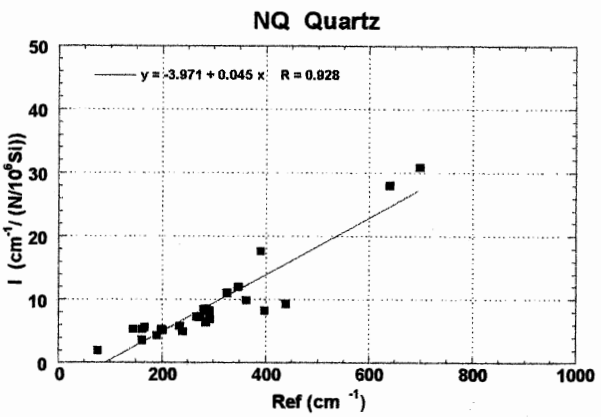


Figure 5 - [OH] Calibration Line – Nat. brazilian quartz.

The present results show that the Paterson's procedure associated with the methodology developped in this work can supply a practical method to estimate the OH content of quartz samples. Based on the assumption that many times the investigations are carried out by only one kind of sample, this method can mean a reduction on the time's analysis.

In order to verify the variation in applying the numerical integration procedure and the method

proposed in this work, the OH concentration was obtained by both methods.

Some results are presented in Table 3. They are numbered too in the previous SP and TQ calibration lines. The mean error is 13 % and the smallest variations are observed for the cultured quartz samples, SP and SW.

Other important feature of this method is that it can be evaluated from the spectra at low temperature or at room temperature. Measurements realized on the broad absorption band, in both conditions,show that the value of I (or Δ) can be taken as being practically the same. However, in order to verify with accuracy this observation, an identical procedure of analysis should be carried out for the IR results taken at room temperature. Also, others publications have mentioned some cases where the broad infrared absorption band remains unchanged on reducing the temperature (Paterson, 1986 ; Aines et al., 1984).

Table 3 - OH Concentration

	[OH] - [N/10^6 Si]			
SAMPLE	**SP**		**TQ**	
METHOD	**(I)**	**(II)**	**(I)**	**(II)**
TREATMENTS	**C.C.**	**Paterson**	**C.C.**	**Paterson**
(1) - As-received	**36.28**	**28.76**	**47.59**	**36.16**
(2) - Sweeping in air	**59.79**	**35.02**	**33.23**	**45.75**
(3) - Sweeping in N_2	**21.49**	**22.06**	**27.86**	**25.70**
(4) - Sweeping in N_2	**20.92**	**20.53**	**30.07**	**43.50**
(5) - Annealing	**35.09**	**29.19**	**38.54**	**59.12**
(I) Calibration Curve -C.C. **(II) Numerical Integration Calibration - Paterson Method**				

ACKNOWLEDGEMENTS

The authors are grateful to the Prof P. Zecchini from the LCCM, Besançon (France), where the IR-FT analyses have been performed. They also acknowledge Telequartz (Brazil) for suppling the natural quartz crystals samples and Dr. P. Guzzo (UFUB – Brazil) by its suggestions. In particular, M.A.P., thanks CAPES (Brazil) for financial support.

REFERENCES

[1] Doukhan J.C., Trepied L. (1985). "Plastic deformation of quartz single crystals". Bull. Minéral., 108 : 97-123.

[2] Halliburton L.E., Chen C.Y., Tapp S.D. (1985). "Radiation effects in quartz : Low doses and defect production mechanisms". Proc. 40th Annual Freq. Control Symposium, 259-265.

[3] Aines R.D., Kirby S.H., Rossman G.R. (1984). "Hydrogen speciation in synthetic quartz". Phys. Chem. Minerals, 11 : 204-212.

[4] Cordier P., Weil J.A., Howarth D.F., Doukhan J.C. (1994). ''Influence of the (4H) Si defect on dislocation motion in crystalline quartz". Eur. J. Mineral.,6,17- 22.

[5] Weil J.A. (1993). "A review of the EPR spectroscopy of the point defects in α-quartz : The decade 1982-1992". Phys. Chemis. SiO_2 and SiO_2 interf. 2., 131-144.

[6] Maschmeyer D., Lehmann G. (1983). "New hole centers in natural quartz". Phy. Chem. Min., 10 : 84-88.

[7] Aines R.D., Rossman G.R. (1984). "Water in minerals ? A peak in the infrared". J. Geophysical Res., 89, B6, 4059-4071.

[8] Kats A. (1962). "Hydrogen in alpha-quartz". Philips Res. Repts., 17, 133-195, 201-279.

[9] Lipson H.G., Euler F., Armington A.F. (1978). "Low temperature infrared absorption of impurities in high grade quartz". 32th Annual Freq. Control Symp., 11-23.

[10] Bahadur H. (1989). ''Infrared characterization of natural and cultured quartz : The effect of electrodiffusion and irradiation". J. Appl. Phys. , 66 (10), 4973-4982.

[11] Griggs D.T., Blacic J.D. (1965). "Quartz : anomalous weakeness of synthetic crystals". Science, 147 : 292-295.

[12] Paterson M.S.(1986). 'The thermodynamics of water in quartz' .Phys. Chem. Minerals, 13: 245-255.

[13] Paterson M.S. (1982). "The determination of hydroxyl by infrared absorption in quartz, silicate glasses and similar materials". Bull. Miné., 105, 20-29.

[14] Brice J.C. (1985). "Crystals for quartz resonators". Rev. Mod. Phys. , 57, 1, 105-146.

[15] Iwasaki F., Shinohara A. H., Iwasaki H., Suzuki C.K. (1990). "Effect of impurity segregation on crystal morphology of Y-bar synthetic quartz". Jap. J. Appl. Physics, 29 (6), 1139-1142.

[16] Guzzo P.L., Iwasaki F., Iwasaki H., Suzuki C. K., Arruda A.C.F. (1993). "Impurity-related point defects and darkening grade in brazilian natural quartz." Rev. Fis. Aplic. Instrumentação, 8(3), 69-83.

TEMPERATURE EFFECT OF THE FERROBIELASTIC BEHAVIOR OF THE QUARTZ CRYSTAL

K. YAMNI*,, J.J. BOY, A. YACOUBI****

* Faculté des Sciences et Techniques, LMCE, BP 146, 20650 MOHAMMEDIA - MAROC
ENSMM / LCEP - 26, chemin de l'Epitaphe - 25000 BESANÇON - FRANCE
** Faculté des Sciences, LCMA, BP 4010, 53000 MEKNES – MAROC

ABSTRACT :

Numerous and previous papers, devoted to the ferrobielastic effect of the quartz crystal, explain how the twinning propagates in a sample when uniaxial pressure is applied in a properly direction. Furthermore, the role of the impurities on the coercive stress has been discussed and compared to the hardness of the material. But any author has been published results about effect of the temperature on the pressure level responsible of the twinning.

Here, we present some experiments realized on X-cut samples built in natural or synthetic quartz crystal having never been submitted to any stress. We show that the role of the temperature cannot be understood only by the temperature variation of compliance coefficients which take place in the "Gibbs free-energy function". Indeed, the coercive stress measured at 100°C is 75% of the coercive stress obtained at ambient temperature. Furthermore, we show that it is not too influenced by structural defects but it is just linked to atomic and molecular thermal vibration energy which increase with temperature.

1. INTRODUCTION

The knowledge of the ferrobielastic twinning and its determination are of prime importance in the framework of the "raw" material qualification. Indeed, electrical twins can be present in the as-grown quartz block or generated by successive mechanical operations necessary to the realization of a resonator : it is what is called "twins induced by ferrobielastic effect", consequence of a property of ferroic materials [1]. It is advisable however to distinguish synthetic quartz and natural quartz because the parameters of the synthetic crystal growth are now sufficiently controlled so that no twins are generated during the growth. On the other hand, the conditions of natural quartz growth having been more unstable, the coexistence of several fields of different orientations is very current and more detrimental for its industrial use. So, in this case it is necessary to make twins detection before any use.

In the same way, to realize force (or stress) sensors, it is essential to control this phenomena of ferrobielasticity which induces appearance of various states generated by the application of pressure. It is significant to know the levels of threshold for which the twin appears.

2. SAMPLES DEFINITION AND EXPERIMENTAL MEASUREMENTS

The samples which are used for our handling were cut out in natural and synthetic quartz blocks. The natural ones come from Madagascar and Brazil. No additional geological information on these crystals has been obtained. The samples have neither visible macroscopic defects, such as fluid inclusions, nor optical twins. The samples were prepared according to the operations of cutting, grinding and polishing usually employed in the LCEP laboratory. The process of cutting was carried out by annular diamond saw. The defect of orientation was always lower than 30''. After cutting, surfaces of the samples were rectified, grinded and polished in order to remove stresses and cutting traces.

We carried out our experiments in applying a quasi uniaxial force on X-cut samples rotated by an angle of –45° towards X-axis. They have a parallelepipedic form with 8 mm along the compressive stress and the surface where the pressure is applied is 3x9 mm^2 (3 mm along X-axis). Once the sample setting up, we apply an increasing pressure on its. When the stress reaches a value which is called threshold (corresponding to the coercive stress), we observe a jump in a stress-strain curve. The pressure is still increased to insure a complete twinning of the sample. According to previous works [2], each experiment has been performed on VIRGIN sample because the pressure decreases when the sample is repeatedly twinned.

The apparatus used to apply almost uniaxial compressive stress was described in a previous work [3]. We recall here that the applied force is measured by a strain gauge load cell, whereas two pairs of displacement gauges indicate the vertical displacement on each side of the sample setting between two hardened steel platens. The axial strain

is evaluated from the average of these two differential displacements.

Furthermore, the temperature is induced by a PID regulation system. Nevertheless, up to now, the maximum of temperature reached is about 120°C.

We will upgrade the system for future experiments to reach higher temperatures, more than 250°C which is the weakening temperature.

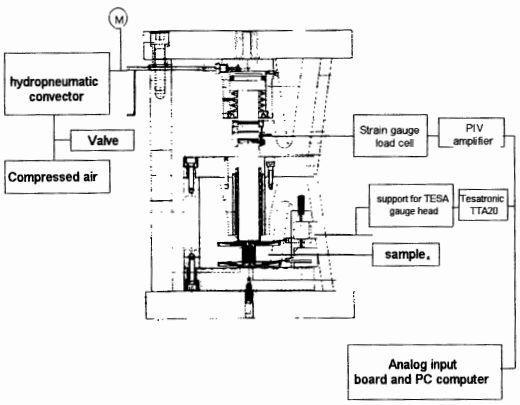


Fig. 1 : experimental arrangement

3. CHARACTERIZATIONS OF SAMPLES

The samples used in our experiments were obtained from the Z-growth region of a synthetic quartz grown by Sawyer Res. Products and natural quartz from Brazil and Madagascar. They are free of macroscopic imperfections as inclusions and twins. Furthermore, the natural samples are free of dislocations. Their characterization consist in using, in one hand, a non destructive method, as IR, to qualify "in situ" the hydroxyl content of each sample, and, in other hand, to quantify the metallic impurities content of each already twinned sample. Furthermore, the IR spectra have been recording before and after experiments in order to observe an eventual change in the repartition of defects. Table 1 gives the values of a well known intrinsic coeff. α.

Table 1 : values of α_{3500} for the different quartz

	Brazil	Madagascar	synthetic
α_{3500}	*0.051*	*0.042*	*0.025*

The following figures (Fig. 2a, 2b, 2c) present the IR spectra for the three kinds of samples obtained at ambient temperature.

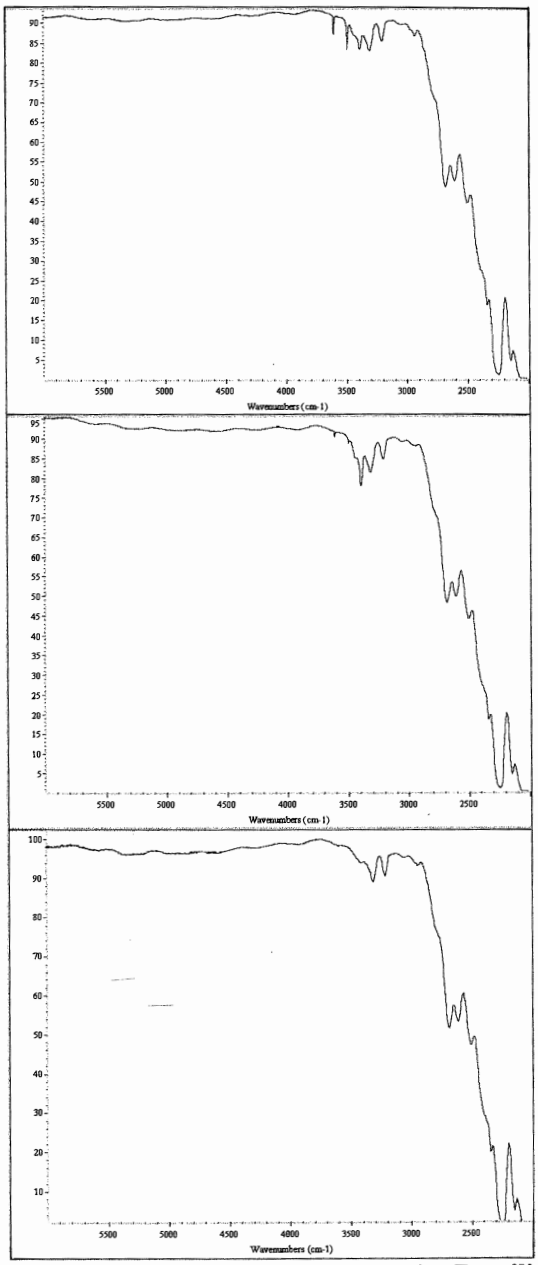


Figure 2a, 2b and 2c : IR-spectra for Brazilian, Madagascan and Synthetic quartz samples at ambient temperature

Table 2 gathers the contents of metallic impurities measured by Inductively Coupled Plasma (ICP) spectrometry following a precise method developed in LCEP laboratory.

Table 2 : impurities contents in ppm (µg / g)

	Fe	B	Mn	Mg	Ti	Al	Ca	Na	Li	K
Synth	**0.55**	0.10	0.03	0.02	0.11	**2.96**	0.19	**0.09**	**0.70**	**<0.01**
Madag	**1.08**	0.20	0.10	0.10	0.30	**31.50**	0.20	**0.19**	**10.57**	**0.50**
Brazil	**0.48**	<0.10	0.01	0.15	0.14	**18.50**	0.40	**0.09**	**5.20**	**<0.05**

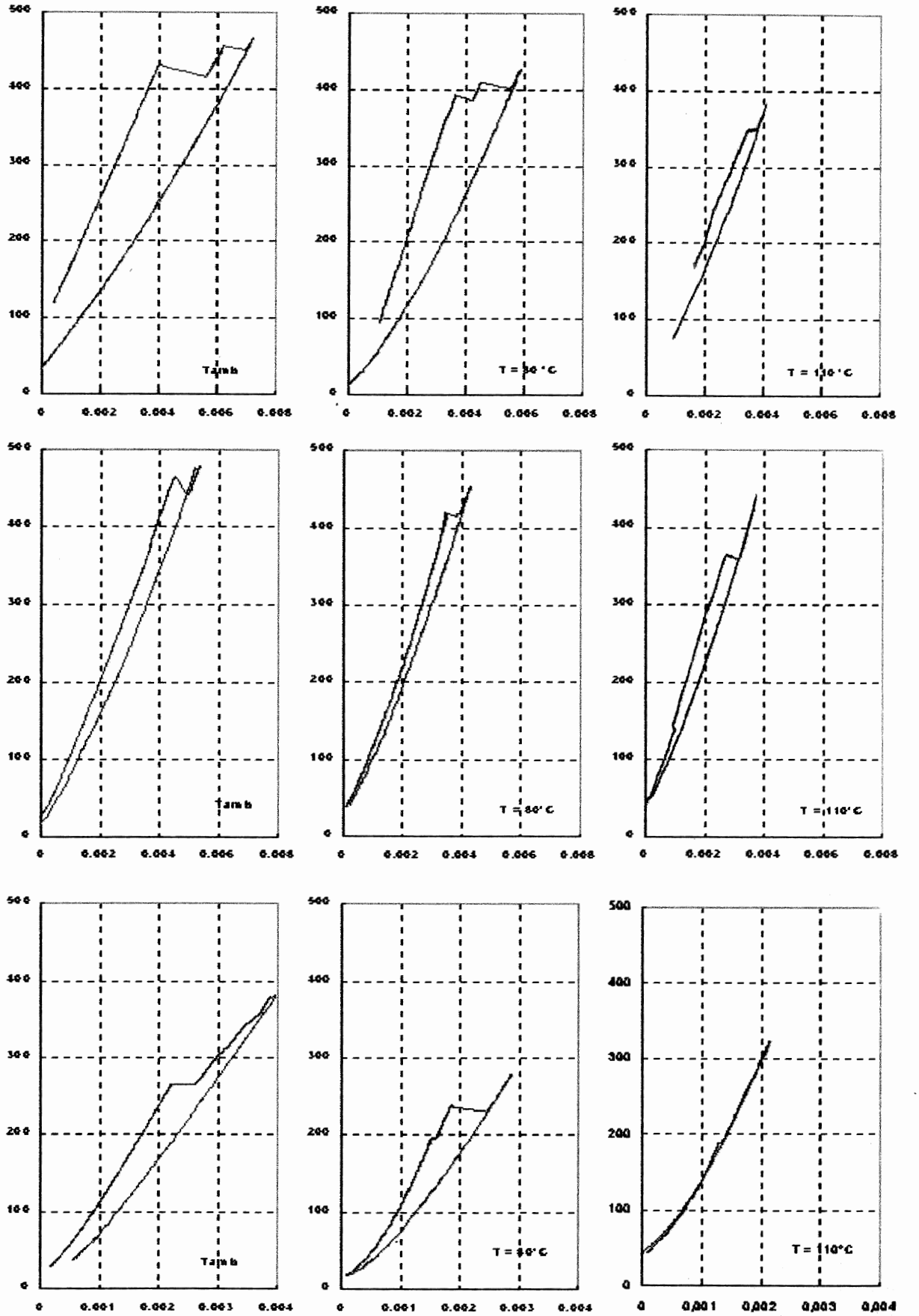


Fig. 3a, 3b and 3c : Experimental Stress-Strain curves versus temperature for : MADAGASCAN, BRAZILIAN and SYNTHETIC Quartz respectively

4. RESULTS :

These experiments make it possible to obtain the curves representing the stress according to the strain. These graphs (Fig. 3a, 3b, 3c) show well the evolution of the crystal structure of the sample during handling. The curves are formed by linear segments characteristic of the linear relation linking the stress to the strain. The setting under pressure of the sample and the unloading are translated on the curves by portions of linear curves having no the same slope. These slopes depend on the elastic coefficients of each state [4]. Therefore the samples exhibit a change in their elastic coefficients. The change from one slope to the other one can be instantaneous or stepped. If the rising curve presents several discontinuities, each step corresponds to the twinning of a part of the sample. In contrary, the entire sample twins instantaneously.

On the previous curves, we note a temperature effect on the threshold of twinning (according to [5, 6]). The coercive stress decreases strongly. Indeed, when the sample is heated, the atomic vibration energy is modified and so the contribution of additional mechanical stresses facilitates the switching in the alternate state. The table hereafter recapitulates this decrease of the coercive stress with temperature and shows that this variation is of the same order for a same family of quartz crystal. In other words, we observe similar behavior for the two kinds of natural quartz and a more important decreasing for the synthetic one which is so more "weak".

Table 3 : Dependency of the coercive stress with the temperature

Coercive Stress (MPa)	T = 25°C	T = 80°C	T = 110°C
Synthetic	**380**	**240 (63%)**	**180 (48%)**
Madagascar	**450**	**405 (90%)**	**350 (78%)**
Brazil	**470**	**410 (87%)**	**370 (79%)**

5. CONCLUSION :

Our study was devoted to the behavior of material ferrobielastic such as quartz. The manufacture of the pressure sensors can carried out only by a good knowledge of material and its behavior under severe conditions of temperature and pressure. Thus the results obtained during this work are :

- The coercive stress decreases when the sample is repeatedly twinned
- The coercive stress depends strongly on the material quality (as proved in previous experiments)
- when the temperature increases, the coercive stress strongly decreases
- the temperature behaviors of the two Natural Quartz are similar.

ACKNOWLEDGMENTS

The authors are grateful to the Prof. A. KIFANI who has permitted the collaboration between LCEP and LCMA laboratories.

REFERENCES

1- **K. AIZU** : "Possible species of ferromagnetic, ferroelectric and ferroelastic crystals", Phys. Rev. B 2 (1970), pp. 754-772.

2- **P.L. GUZZO and Al.** : "Influence of the sample geometry on twinning genaration in quartz : experimental measurements and finite elements analysis", 8th E.F.T.F. (1994), pp. 802, 810.

3 - **J.J. BOY, P.L. GUZZO** : "Quartz crystal twinning under mechanical stress : Experimental measurements", 50th Ann. Symp. Freq. Cont. (1996), pp. 155-160.

4 - **J.J. BOY, K. YAMNI, A. YACOUBI** : "Comparison between finite elements and experimental analyses of the ferrobielastic behavior of quartz crystal", This symposium.

5 - **J.W. LAUGHNER** : "Induced Dauphiné twinning in quartz", Ph. D. Thesis, Penn. State Univ., 1982.

6 - **T.L. ANDERSON et al.** :"Coercive stress for ferrobielastic twinning in quartz", 30th Ann. Symp. Freq. Cont. (1976), pp. 171-177.

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

Influence of Vacuum Rapid Thermal Annealing on the Sheet Resistance of Thin Al Films Deposited on Quartz Crystal

V. Georgieva, D.B. Dimitrov, G. Beshkov, L. Spassov
Institute of Solid State Physics,
Bulg. Acad. of Sciences, 72 Tzarigradsko Chaussee blvd.,
1784 Sofia, Bulgaria, e-mail: Lasarova@issp.bas.bg

ABSTRACT

The influence of Rapid Thermal Annealing (RTA) on the structure and on the sheet resistance of thin aluminium layers – basic material for electrodes in acustoelectronics – has been investigated. The aluminum layers with thickness of 1200 Å have been deposited on quarts substrates AT cut by three different methods: DC magnetron sputtering, RF sputtering and thermal vacuum evaporation. The samples obtained have been subjected to RTA in vacuum of the order of 6.66×10^{-3}Pa, temperature of the heater of 700° C up to 800° C for time durations of 5 s up to 180 s.

The structural investigations of the samples have been carried out by Reflection High Energy Electron Diffraction (RHEEED) with accelerating tension of 50 kV.

The changes in the sheet resistance of the layers have been registered using the four point technique.

It has been established that the sheet resistance of the three kinds of layers is a function of the RTA parameters (temperature and duration) and it has a maximum. This behaviour could be explained with processes taking place at the interface Al-SiO_2 as well as at the Al layer itself.

1. INTRODUCTION

The aging of quartz crystal resonators is the main problem for their long term utilization. According to J. Vig [1], the aging mechanisms include: contamination effects, stress effects, electrode effects, changes in the quartz material. One of the main effects is the effect of changes in the chemical composition and in the structure of electrodes. Thin Al, Ag and Au layers are most frequently used as electrodes in quartz resonators. These layers should meet the requirements for long term stability of their properties, which depends on the fabrication conditions as well as on the sequential temperature treatments. The stress in the layers is also a function of the preparation conditions (substrate temperature, layers thickness, fabrication method). The stress in the electrodes could be either compressive or tensile. In many cases of thin Au electrodes the stress was found to be compressive and a transition to tensile stress was experienced only at annealing [2]. Changes in the quartz substrates themselves are also possible with time and under temperature treatment. The imperfections that are subject to changes include surface defects, point defects, dislocations, impurities, inclusions and twince [3]. Obviously, the stability of the films depends mainly on the deposition temperature and on the temperature subsequent to deposition.

The aim of this work is to investigate the effect of rapid thermal annealing (RTA) in vacuum on the structure, morphology and on the sheet resistance of thin Al films prepared by three different methods of deposition.

2. EXPERIMENTAL RESULTS AND DISCUSSION

Thin aluminum layers have been deposited on polished quartz substrates AT cut. Three deposition methods have been used: thermal vacuum deposition, DC magnetron sputtering and RF sputtering.

The process of Al thermal vacuum evaporation is realized at base vacuum pressure of 6.66×10^{-3} Pa. The thickness of the evaporated Al layer is 1200Å and the deposition rate is 240 nm/min. The distance between the substrate and the evaporator is 10 cm. During the deposition the substrate is not intentionally heated.

The Al layers deposited by DC magnetron sputtering have a thickness of 2000Å and the deposition parameters are: deposition rate 71 nm/min, vacuum 0.3Pa and power – 7 kW.

The RF sputtering process uses Al target with 1% Si content and is done in an Ar ambient. The base pressure of the system was 1×10^{-3} Pa and the working pressure was 3 Pa at power density of

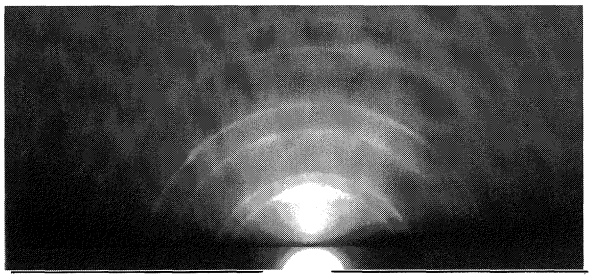


Fig. 1. Electronograph picture of evaporated Al layers after RTA at 800° C, 180 s.

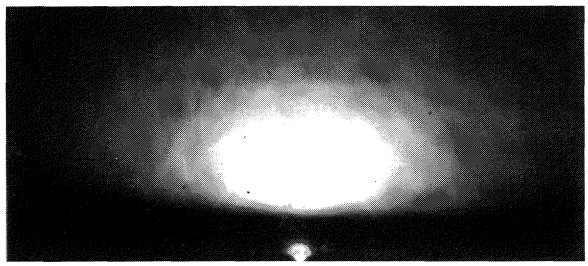


Fig. 2. Electronograph picture of as-deposited DC magnetron sputtered Al layer

5.25 W/cm^2. Under these conditions Al layer with a thickness of 1200 $\mathring{A}$ was deposited at deposition rate of 32 nm/min.

The structures Al-quartz obtained by the three method have been subjected to rapid thermal annealing in vacuum – 6.66×10^{-3}Pa, at temperature of the heater 700, and 800° C for durations ranging between 5 to 180 s.

The structure and the morphology of the Al layers have been investigated before and after the RTA treatment through Reflection High Energy Electron Diffraction (RHEED) and Secondary Electron Microscopy (SEM) using JSM – 5300.

The results from the investigations show that the changes appearing in the layers are a function of the RTA parameters and the initial parameters of the as-deposited layers.

The layers obtained by thermal evaporation are polycrystal. Their RTA treatment up to 180 s at 800° C leads to the appearance of small differences in the Debay lines, which could be explained with the initial phase of texturizing (Fig. 1).

The layers obtained by DC magnetron sputtering are amorphous. They show initial phase of recrystallization after being subjected to RTA even at 700° C. Typical electronogams of the layers before and after RTA are shown in Figs 2 and 3.

The RF sputtered films exhibit polycrystal structure immediately after their deposition. After the RTA treatment in the temperature range investigated no substantial changers have been registered.

The morphology of the Al-layers before and after annealing have been studied using SEM. A typical morphological surface of as-deposited layers is

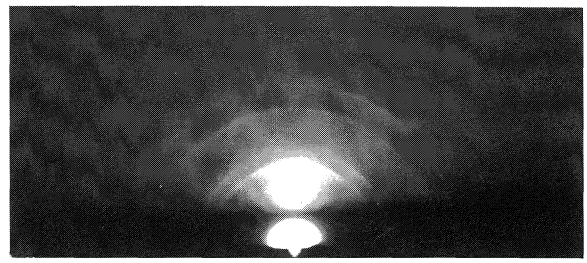


Fig. 3. Electronograph picture of DC magnetron sputtered Al layer after RTA at 700° C, 180 s.

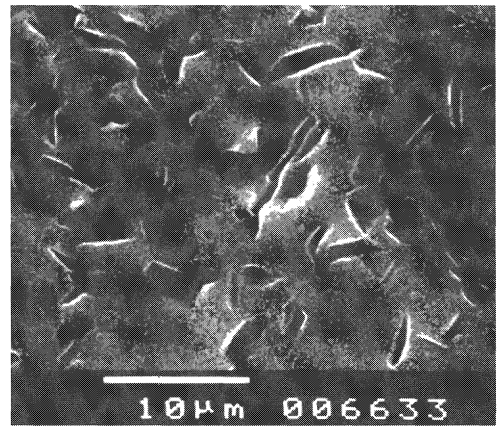


Fig. 4. Electron microscope picture of as-deposited Al layer.

shown in Fig. 4. The surface is characterized by the presence of a great number of pores. As a result of RTA the morphology indicates homogenization of the surface (see Fig. 5). It can be supposed that the morphological changes in Al-films are due to the densification of the layers which have undergone RTA. More details of the morphological changes of thin Al-films deposited by different methods are studied in [4,5].

The changes in the sheet resistance, ρ_s of the Al layers after RTA are registered using four point probe technique.

Depending on the fabrication method, the sheet resistivity of as-deposited thin Al layers may differ by magnitude. While for DC sputtered layers the typical value of ρ_s is 0.158Ω/□, for RF deposited layers it is 2.2Ω/□ and for thermally evaporated layers it may reach 20Ω/□. The difference in ρ_s for as-deposited

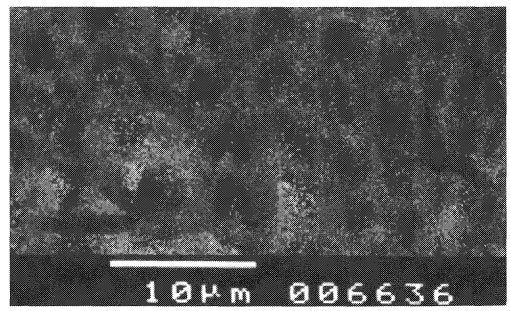


Fig. 5. Electron microscope picture of Al layer after RTA at 800° C, 180 s.

layers is probably due to differences in the mechanisms of deposition for the processes of evaporation and sputtering respectively.

In the case of vacuum evaporation the material to be deposited is evaporated and condensed at the substrate surface at a vacuum sufficient to prevent collisions between the film forming atoms and residual gas atoms as they travel the distance between source and substrate. Thus the vacuum deposition process involves atoms having thermal energies (0.1-0.5 eV) [6].

Sputter deposited films are formed by the condensation of atoms formed by sputtering. These ejected atoms may have energies (1-100 eV) much greater than thermal ones [7].

This is the reason the sputter deposited films to have better adhesion than vacuum deposited ones. Maybe the difference in ρ_s of the layers obtained by the two deposition techniques is due to the differences in the energy of the atoms forming the layers.

Fig. 6 (a,b,c,) illustrates the change of ρ_s for Al layers after RTA at 700 and 800° C for durations up to 180 s. The dependences show one and the same tendency in the change of ρ_s independently of the deposition method.

Fig. 6 a) shows the change of ρ_s for thermally evaporated Al-layers. Two regions can be distinguished for both the temperatures of annealing: the first one when ρ_s increases and the second one when ρ_s gradually decreases its values. It should be noticed that during the first 15 s of the annealing the changes in ρ_s are negligible for both the annealing temperatures. The maximum in the dependence $\rho_s(t)$ for annealing temperatures of 700 and 800° C has been displaced along time axis. At the higher annealing temperature the maximum is reached after 30 s, while at the lower annealing temperature it is reached for 60 s. The decreasing of ρ_s for the annealing temperature interval investigated is gradual and the values for 800° C are lower than these for 700° C. The maximum change of ρ_s at 700° is 30%, and for 800° C RTA it is 40% respectively.

Similar are the results about the RF-sputtered Al films (Fig. 6 b). The maximal change of ρ_s after RTA at 700° C is 19%, and for 800° - 45%. The times for reaching the maximal values are the same as for the evaporated layers. For the DC magnetron sputters Al layers the change in ρ_s is established to be 22% after RTA at a temperature of 700°.

The observed (Fig. 6 c)) dependence could be explained with the changes taking place in the system Al-quartz: in the aluminum layer itself as well as of the interface Al-SiO_2.

The top interface Al_2O_3-Al is simply a result of the exposure of the aluminum metallization to air, which results in the formation of an oxide film approximately 30Å thick [8]. This interface does not

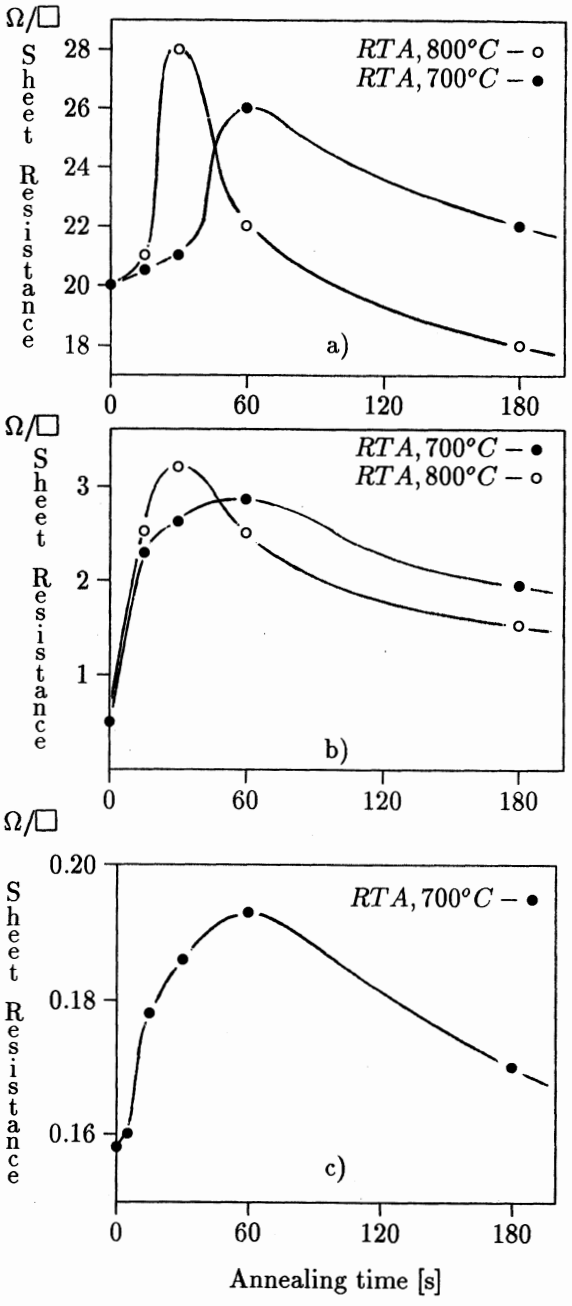


Fig. 6. Sheet resistance of Al films deposited by different techniques on the annealing parameters: a) evaporation; b) RF sputtering; c) DC magnetron sputtering.

affect device characteristics whereas the other interface Al-SiO_2 is complex and has some effect on device performance. It has been experimentally proven, that Al reduces SiO_2 at direct contact. As a result of the reaction Al_2O_3 and free Si are obtained

$$3SiO_2 + 4Al \rightarrow 2Al_2O_3 + 3Si + Qkcal$$

This reaction proceeds as a result of the existing difference between the Al_2O_3 heat formation - 399 kcal/mol and the one for α-quartz formation -201,34 kcal/mol [1]. Thus at the interface of the system Al-quartz a Al, Al_2O_3, SiO_2 and Si are presented. The

effective thickness of the Al layer decreases and as consequence the values of ρ_s increase. This assumption is confirmed as well by the investigations of the Al-SiO_2 system microhardness after RTA treatment [9], which show that a layer of constant hardness has been detected under the initial interface. It is supposed that this results from incorporation of Al atoms into the surface layer of the substrate, reduction of the SiO_2 and diffusion of the atoms of a certain element during the process of preparation and rapid thermal annealing of the structures.

The second region of the ρ_s dependence on the duration of RTA is characterized by gradual decreasing of ρ_s. This decay for the longer annealing times is probably due to recrystallization process related to the temperature treatment.

Similar dependence of the resistivity change as a function of the substrate temperature during the deposition has been registered for thin Au films. It has been established that the ρ_s decrease is weaker when the layer is deposited at temperatures close to the recrystallization ones of Au [10].

3. Conclusion

We have shown in this paper that the rapid thermal annealing of resonator structures in the temperature range 700° C and 800° C of the heater for durations up to 180 s influence the structure, the morphology and the sheet resistance of thin Al layers deposited on quartz. It has been supposed that the change in the sheet resistance is due to processes taking place at the interface Al-SiO_2 as well as of the Al layer itself.

Acknowledgements

This work was partly supported by the *Bulgarian Ministry of Science and Education in the frame of Project* Φ902

References

[1] John R. Vig, T.K. Meeker, "The Aging of Bulk Acoustic Wave Resonators, Filters and Oscillators", Proc. of the 45-th Annual Symposium on Frequency Control, 29-31 May 1991, Los Angelis, pp.77-101

[2] G. Phornell, F. Ericson, G. Portnoff, "Internal Stress in Sputtered Gold Electrodes and Its Significance for Quartz Resonators", Proc. of 1999 Joint Meeting of the European International Frequency and Time Forum and the IEEE International Frequency Central Symposium, 13-16 April, 1999, Beason, France, Vol.2, pp.800-804

[3] J. R. Vig, "Chemically Polished Quartz", Proc. of the 31-st Annual Symposium on Frequency Control, 1977, pp.131-143

[4] L. Spassov, V. Lazarova, M. Marinov, "Behaviour of Al-Quartz System After Rapid Thermal Annealing" Proc. of the 4-th International Symposium on Surface Waves in Solid and Layered Structures (ISSWAS-4), 7-12 June, 1998, St. Petersburg, pp.163-173.

[5] L. Spassov, V. georgieva, M. Marinov, "Structure and Morphology of Thin Al Layers in Resonator's System After Rapid Thermal Annealing" Proc. of 1999 Joint Meeting of the European International Frequency and Time Forum and the IEEE International Frequency Central Symposium, 13-16 April, 1999, Beason, France, Vol.1, pp.421-484

[6] L.I. Maissel and M.H. Francombe, "An Introduction to Thin Films", 1973, pp.200-210, Gordon and Breach Science Publishers

[7] E. Murr and N.C. Inman, "Effects of Vacuum Environment on the Sub-structure of Evaporated F.C.C. Metal Films", Phyl. Mag., Vol. 14, pp.133-153, 1966

[8] Y.E. Strausser, "Observation of Al_2O_3 and Free Silicon at the Interface Between Aluminum Films and SiO_2", Thin Solid Films, 52, pp.203-214, 1978

[9] E.P. E.P. Trifonova, V. Lazarova, L. Spassov, N. Efremova, "Composite Microhardness of Quartz-Al Structures", Cryst. Res. Technol., 34, Vol. 3, pp.391-395, 1999

[10] J.R. Lloyd and S. Nakahara, "Low Temperature Void Grown and Resistivity Decay in Thin Evaporated Gold Films", Thin Solid Films, Vol. 45, pp.411-420, 1977

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

ELASTIC, PIEZOELECTRIC AND DIELECTRIC PROPERTIES OF $La_3Ga_5GeO_{14}$ SINGLE CRYSTALS

I.M.SILVESTROVA*, Yu.V.PISAREVSKY*, B.V.MILL**, A.V.BELOKOPITOV*, N.A.MOISEEVA*

*Institute of Crystallography of Russian Academy of Science,117333 Moscow, Russia, E-mail aopt@ns.crys.ras.ru
** Moscow State University, 119899 Moscow, Russia, E-mail mill@plms.phys.msu.su

ABSTRACT

Full set of independent elastic, dielectric and piezoelectric constants of most close analogue of langasite ($La_3Ga_5SiO_{14}$) - $La_3Ga_5GeO_{14}$ crystals were measured at room temperature by combination of ultrasonic resonance and pulse methods.

Peculiarities of bulk ultrasonic wave propagation were investigated.

Temperature dependencies of several ultrasonic mode velocities were measured and elastic anomalies which ensure temperature compensated cut were established.

1. INTRODUCTION

Crystals $La_3Ga_5SiO_{14}$ (langasite, LGS) belong to group of very promising materials for frequency control devices[1-3]. Up to day more than 40 compounds with similar structure were synthesized and for some of them better than for LGS piezoelectric properties were found [4-6]. But most close analogue of langasite crystals, which can be obtained by substitution Ge instead of Si is still unstudied.

At present paper crystals of $La_3Ga_5GeO_{14}$ were grown and piezoelectric, elastic and dielectric constants were measured at room temperature. On the base derived from above measurements full matrices of elastic, piezoelectric and dielectric constants characteristics of elastic wave propagation were calculated. Temperature dependencies of several mode sound velocities were measured.

2. CRYSTAL, EXPERIMENTAL TECHNIQUE AND MEASUREMENT METHODS

$La_3Ga_5GeO_{14}$ (LGG) melts congruently at the same temperature as $La_3Ga_5SiO_{14}$(LGS) (~1470° C) and its single crystals may be grown from the melt by the Czochralski technique. The growth is complicated by the GeO_2 evaporation. An 200g $La_3Ga_5GeO_{14}$ boule, 23mm in diameter and 90 mm in length used for specimens fabrication, was grown on [001] LGS seed from stoichiometric melt in atmosphere 3%O2+97%N2 at pulling rate 4mm/h and rotation rate 35 rpm. Crystal has an rounded side surface etched by GeO_2 vapour from melt and a flat bottom. Crystal was colourless and transparent up to half length , then some inclusions appeared and crystal quality progressivly decreased up to bottom.

As a LGS crystals LGG which were grown at present work were right handed.

For resonance and dielectric permeability study bars (14 to 21x2.5x1.5 mm3), platea (15 to 20x 16 to 21x 1 mm3) with various orientation were made. Accuracy of orientation was better than 10'. On opposite surfaces of those samples silver electrodes were deposited.

For pulse ultrasonic measurements two parallelepipeds (15x18x19 mm3) one with orientation along main crystallographic directions , other with one face under 45o angle to Y and Z axis(14x15x17 mm3) with orientation accuracy better than 10' were made.

Dielectric constants were obtained by measurement of capacity plates of Y- and Z cuts at the frequency 1 kHz with electric field in specimens of 100 to 200 V/sm.

Electromechanical properties of the LGG crystals were measured by resonance-antiresonance method[4] .

Elastic properties were measured by pulse ultrasonic methods - the synchronous ring and phase pulse methods at a frequences 15 –30 MHz.

3. PROPERTIES OF LGG AT ROOM TEMPERATURE

Electromechanical properties of LGG, which were derived from resonance measurement presented at Table1 with the similiar data of LGS. Big deviation in k_t and k_{26} values had to be point out. It cuold be explained by inhomogenities influence, which usially more strong for thickness modes.

From those data and dielectric constants piesoelectric strain constants were obtained. On the base experimental data for several sound mode velocities elastic stifness modules were calculated.

All independent dielectric, piezoelectric and elastic constants of LGG crystals together with early published data of LGS are presented at Table 2.

Table 1
Elecromechanical data of $La_3Ga_5GeO_{14}$ (LGG) and $La_3Ga_5SiO_{14}$ (LGS) crystals

Mode of vibration	LGG		LGS[1]	
	Fxl	k_{ik}	fxl	K_{ik}
Length-extentional mode of bar (Xytwl)-0°/0° /0°	2180	k_{12} 16.6	2228	K_{12} 16.0
Thickness-extentional mode of plate X-cut	2817	k_t 14	2892	K_t 8
Shear thickness mode of plate Y cut	1362	k_{26} 14	1382	K_{12} 14
Shear contour mode plate of Y cut	176.8	k_{14} 8.9	180.2	K_{14} 8

Table3
Piezoelectric strain $d_{ik}(10^{-12}$ C/N), elastic stiffness c_{ik}^E $(10^{10}$ N/m$^2)$ and dielectric permittivity constants $\varepsilon_{ik}/\varepsilon_0$ and dencity ρ (g/sm^3) of $La_3Ga_5GeO_{14}$ and $La_3Ga_5SiO_{14}$ single crystals

	$La_3Ga_5GeO_{14}$	$La_3Ga_5SiO_{14}$ [4]
d_{11}	-6.59	-6.16
d_{14}	5.51	5.36
c_{11}	18.529	19.092
c_{33}	25.972	26.191
c_{12}	10.362	10.635
c_{13}	10.27	10.42
c_{44}	5.123	5.241
c_{66}	4.084	4.323
c_{14}	1.505	1.519
$\varepsilon_{11}/\varepsilon_0$	19.36	18.99
$\varepsilon_{33}/\varepsilon_0$	45.64	49.32
ρ	5.935	5.754

From Table 2 ones can see than LGG have higher piezolectric constants and less elastic constants than LGS.

Using Christoffel relations and Table 2 data characteristics of sound wave propagation at LGG crystals were calculated. On Fig 1 cross sections of the phase velocities of LGG by plates (001), (010) and (100) are shown compare with LGS values. From the picture ones can see that sound velocities of LGG less than LGS ones but for both crystals angular dependencies of all modes very similar. On the Fig2 are shown dependencies group wave direction from wave normal direction.

4. TEMPERATURE DEPENDENCES OF LGG SOUND VELOCITIES.

Temperature dependencies of sound velocities of various mode of LGG crystals were measured at temperature range 10- 70 °C. Some of them represented at Fig 3.

We can see from pictures that almost all modes have normal behaviour: decreasing of sound velocities with temperature growth. But velocities of two modes: slow shear waves along Y and X direction have abnormal increasing with temperature growth. Slow shear mode most important for BAW application and as for LGS abnormal behaviour indicate existance of rotated Y cuts with zero temperature coefficients.

5. CONCLUSION

$La_3Ga_5GeO_{14}$ (LGG) crystals could be useful for frequency control application. Above information also could useful for understanding nature of elastic anomalies at langasite family materials.

Work was partly support by the Russian Fund of Fundamental Investigation Grant N 99-2-17242

6. REFERENCES

[1]. R.C.Smithe, Material and resonator properties of langasite and langatate: a progress report Proc. 1998 IEEE Int. Frequency Control Symposium p.761-765

[2] J.Hornsteiner,E.Born,,G.Fisherauer,E.Riha Surfase acoustic wave sensors for high-temperature applications Proc. 1998 IEEE Int. Frequency Control Symposium p.615-620

[3] A.Bungo,C.Jian,K.Yamaguchi,Y.Sawada, S,Uda, Yu. Pisarevsky Analysis of Surface Acoustic wave properties of the rotated Y-cut langasite substrate. Jpn. J. Appl. Phys. Vol.38 (19990 p 3239-3243

[4]. I.M.Silvestrova, Yu.V.Pisarevsky, A.A.Kamiskii, B.V. Mill Elastic, piezoelectric and dielectric properties of crystals $La_3Ga_{5.5}Nb_{0.5}O_{14}$ Sov.Phys. Solid State vol.29 N5 1987

[5] Yu.V.Pisarevsky, P.A.Senyushenkov, B.V.Mill, N.A.Moiseeva Elastic, piezoelectric and dielectric properties of crystals $La_3Ga_{5.5}Ta_{0.5}O_{14}$ Proc. 1998 IEEE Int. Frequency control Symposium p.742-747

[6] A.A.Kaminskii, E.L.Belokoneva, B.V.Mill, Yu.V.Pisarevsky, S.A.Sarkisov, I.M.Silvestrova, A.V.Butashin and G.G.Khodzhabagyan. Pure and Nd3+doped $Ca_3Ga_2Ge_4O_{14}$ and $Sr_3Ga_2Ge_4O_{14}$ single crystals, their structure , optical, spectral luminescense, electromechanical properties and stimulated emission. Phys Stat. Sol (a.) v86, 1984, p345-362

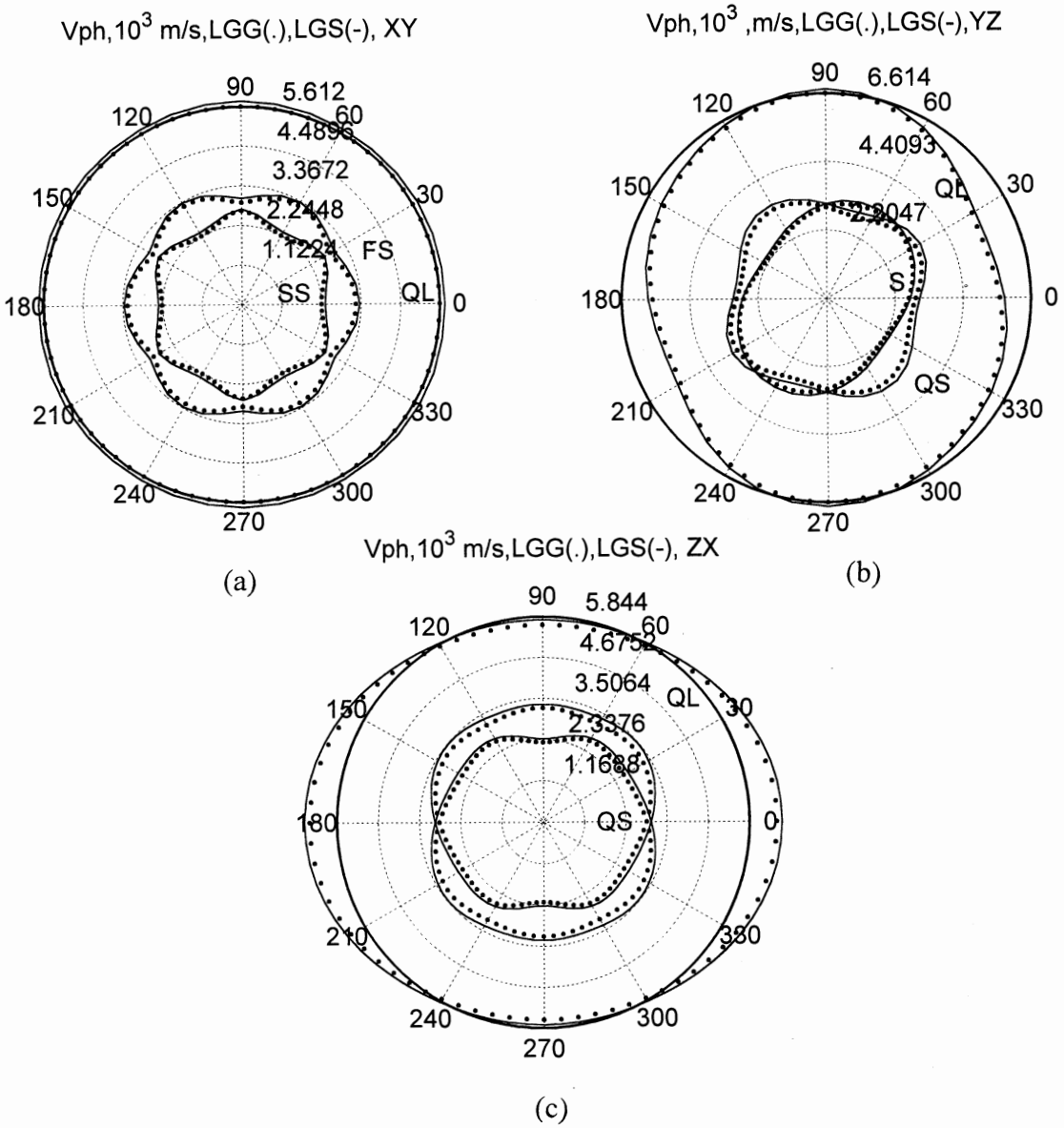


Fig.1 Cross sections of the phase velocities surfaces in $La_3Ga_5Ge\ O_{14}$ (…) and $La_3Ga_5SiO_{14}$ (-) crystals by the (a) (001), (b) (100) and (c) (010) planes: Q-quasi, L-longitudinal, S-shear

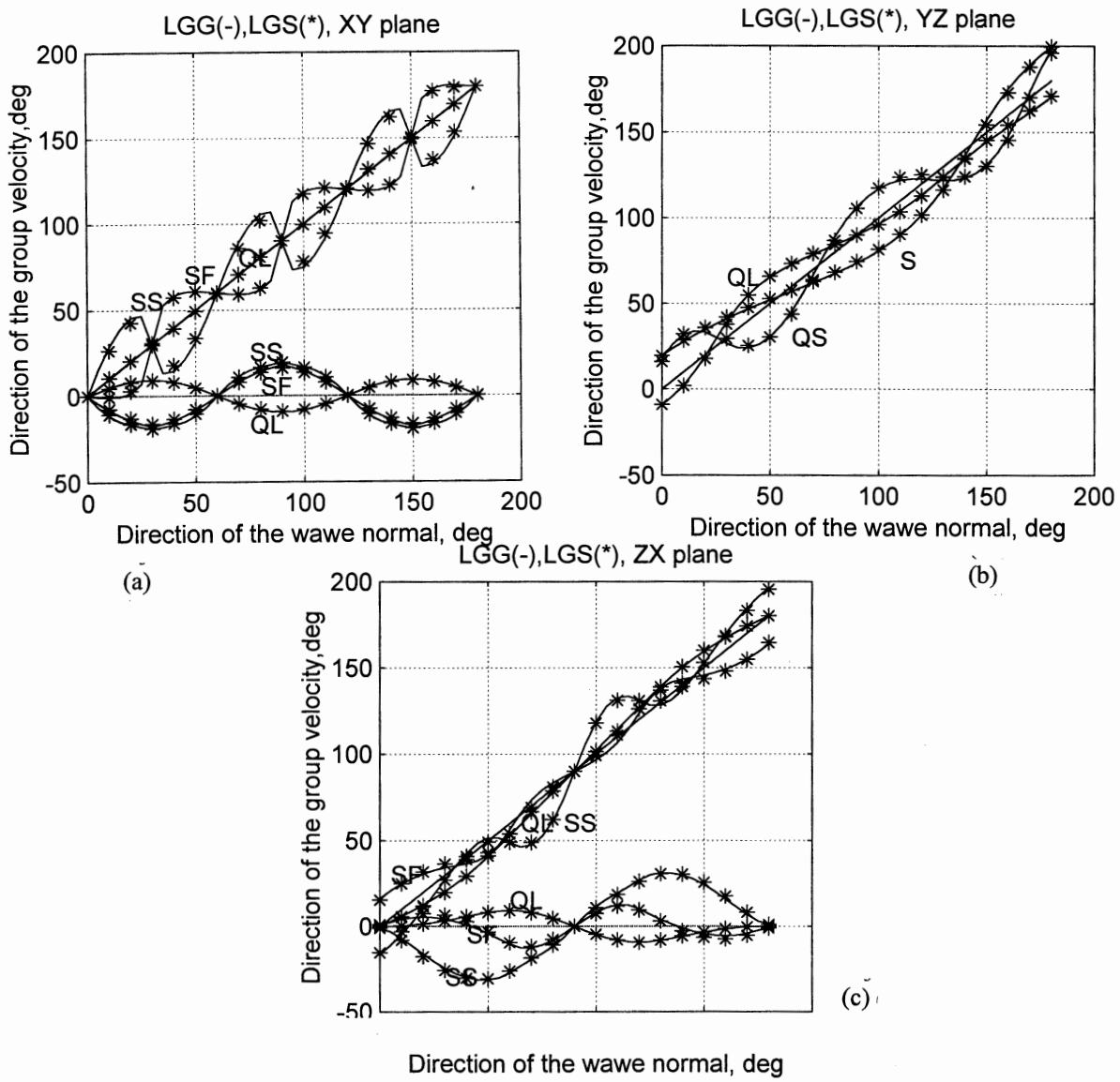


Fig.2 . Direction of acoustic beam propagation (group velocities) as a function of the direction of the wave normal for $La_3Ga_5Ge\,O_{14}$ (-) and $La_3Ga_5SiO_{14}$ (*) crystals , (a) (001), (b) (100) and (c) (010) planes: Q-quasi, L-longitudinal, S-shear waves.

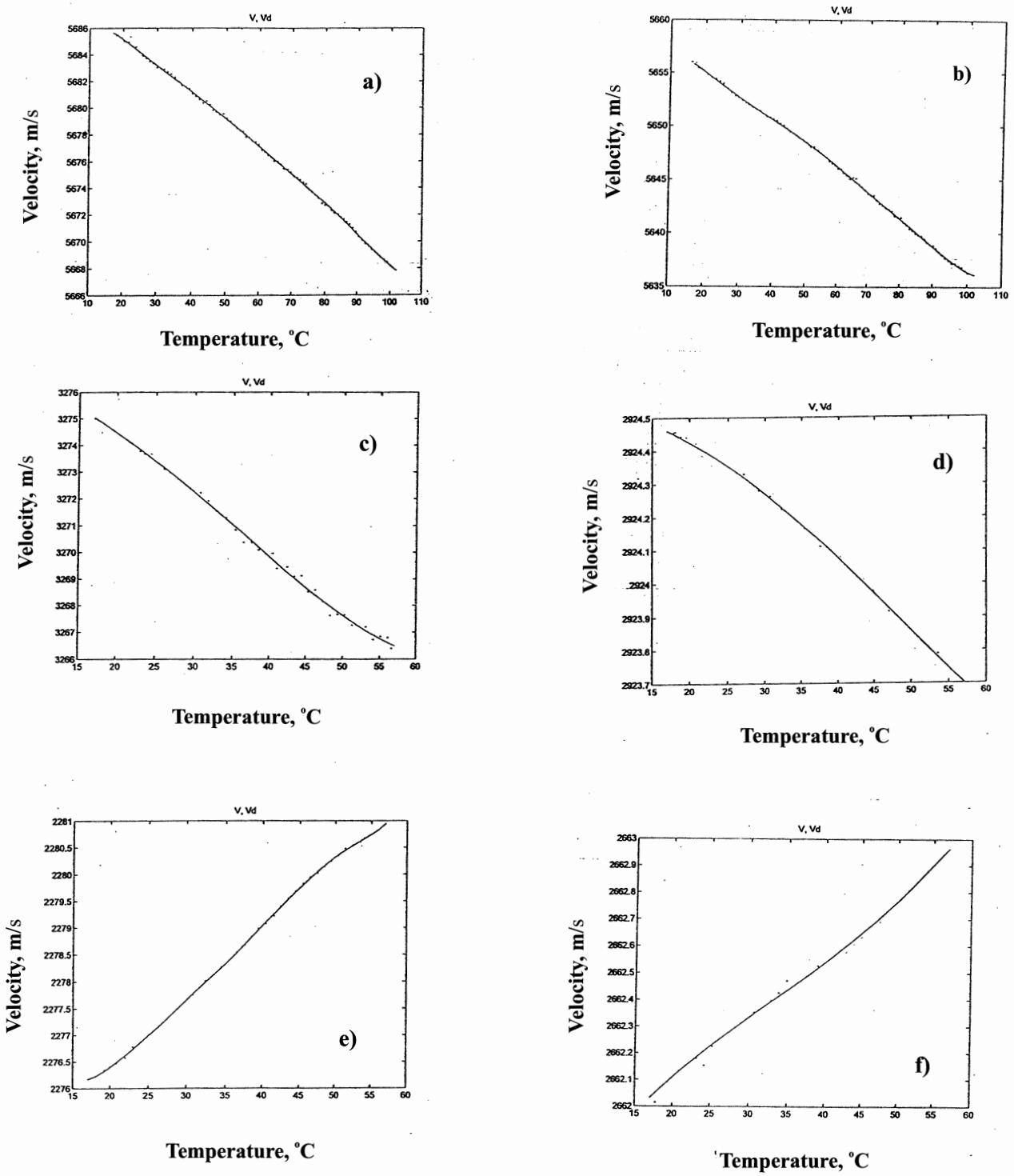


Fig. 3 - Temperature dependencies of the phase velocities of $La_3Ga_5GeO_{14}$ crystals a) longitudinal along X, b) longitudinal along Y, c) fast shear along X, d) fast shear along Y, e) slow shear along X, f) slow shear along Y

THE MASS-LOADING INFLUENCE ON THE ELECTRICAL PARAMETERS OF LANGASITE RESONATORS

I.Mateescu, G.Johnson*, A.Manea, I.Boerasu

National Institute of Materials Physics, PO Box MG-7, 76900 Bucharest-Magurele, Romania
Phone: 401-493 0047; Fax: 401-493 0267; E-mail: imate@ alpha1.infim.ro
*Sawyer Research Products, 35400 Lakeland Blvd., Eastlake, Ohio 44095, USA

ABSTRACT

Based on the Tiersten's theory of the monolithic structures [1], a theoretical calculation of the coupling coefficient of quartz and langasite monolithic two-pole filters as a function of mass-loading was performed [2,3]. The study revealed that the mass-loading influence on coupling coefficient of langasite filters is much smaller than in the case of quartz filters.

This conclusion is supported by the new experimental results of mass-loading effect on langasite electrical resonator parameters which are presented in this paper. The Y-cut plan-parallel polished Sawyer langasite blanks with 14mm diameter, 5 MHz frequency and AT-cut quartz resonators having the same parameters were deposited by vacuum evaporation of the Ag and Au electrodes.

Using the Ballato's transmission-line analogs of thickness-shear vibrating plates [4,5], the effective mass-loading, effective coupling coefficient, motional inductance and quality factor as a function of harmonic order were computed. The analysis of the results shows that the mass-loading effect on characteristics of langasite resonators is less important than in the case of quartz resonators.

1. INTRODUCTION

Langasite crystal is one of the promising material for applications in acoustoelectronic and piezoelectric devices. The larger coupling coefficient of langasite in comparison with quartz allow to get filters with larger bandwidth and to use higher ranks overtones. Also, the main characteristics of the langasite crystals compare favourably with those of quartz: the absence of phase transitions up to the melting point and low acoustic wave propagation losses [6,7,8]. Y-cut langasite plates, which vibrates in thickness-shear mode, exhibit a good thermal stability, low values of the equivalent motional inductance and series resistance and high electromechanical coupling coefficient. The size of the LGS devices are smaller by 30% than that of quartz devices [9,10,11,12].

Theoretical and experimental investigations of the mass-loading effect on plan-parallel and plan-convex AT and SC-cut quartz resonators were performed [13,14,15,16]. In this study was used the Ballato's transmission-line analogs of the thickness-shear resonators with a non-uniform distribution of vibratory motion over the electroded area of the crystal. As suggested in the papers [15,16] the non-uniform distribution of motion is associated with the effects at interface electrode-piezoelectric substrate. Using the Tiersten's analysis of trapped-energy resonators vibrating in coupled thickness-shear and thickness-twist modes [17], a correction of the mass-loading and coupling coefficient was performed in [18]. Thus, we can suppose, that non-uniform distribution of motion could be due to two factors: 1) the coupling of the thickness-shear and thickness-twist modes; 2) stress related effects at electrode-substrate interface. The experimental results regarding mass-loading effects on quartz resonators characteristics pointed out that the harmonic dependence of the motional inductance for large thickness of electrodes, small electrode diameters and high frequencies can be ascribed to stress effects associated to the deposited metallic electrodes on the piezoelectric plates, while in the case of small electrode thickness, large electrode diameter and low frequencies the inductance behaviour could be explained by the coupling of thickness-shear with thickness-twist modes.

Based on Tiersten's analysis an accurate calculation of the coupling coefficient dependence on geometrical dimensions and mass-loading for quartz and langasite monolithic structures was performed [2,3]. We have studied the effect of mass-loading on the coupling coefficient of langasite and quartz monolithic filters and we have found that the effect on langasite filters is lower than on quartz filters.

In this paper are compared the experimentally results of the mass-loading effects on Y-cut langasite resonators characteristics with those on AT-cut quartz resonators. The analysis of the results revealed that the maximum variation of the effective mass-loading, effective coupling coefficient and inductance as a function on harmonics for various electrode diameters and thicknesses is significantly lower for langasite resonators than for quartz resonators. We can ascribe this behaviour to lower interfacial stress in the case of Y-cut langasite resonators than for AT-cut quartz resonators. This new property of langasite allows to conclude that langasite is more advantageous than quartz for applications in piezoelectric devices.

2. EXPERIMENTAL

For experiments were used Sawyer Y-cut plan-parallel polished langasite and AT-cut quartz plates with 14 mm diameter and 5 MHz resonant frequency. The quartz blanks were processed with the usual technology. As-cut langasite wafers were lapped with 3 μm alumina powder. Between lapping and polishing steps, langasite wafers were chemically etched in a $2HNO_3$:CH_3COOH solution for removal of disturbed layer due to mechanical treatment. Polishing process used slurry of colloidal alkaline suspension of silica gel (QUSO) on a Politex pad material [19].

Ag and Au electrode layers with 75, 125, 200 nm thicknesses and with 4.6 and 7 mm electrode diameters were deposited on quartz and langasite blanks The electrode deposition was made by thermal evaporation in vacuum using a JEOL - JEE 4X installation.

The resonance and antiresonance frequencies and series resistances of the fundamental, 3rd, 5th and 7th overtones of the free and electroded plates were measured after every electrode layer. On the basis of the relations for the transmission-line equivalent electrical circuit of the piezoelectric plate resonator which vibrates in the thickness-shear mode, the effective mass-loading, piezoelectric coupling coefficient, motional inductance, motional capacitance and quality factor were calculated. Mass-loading effect could be evidenced by the change of these parameters on harmonic order.

3. RESULTS AND DISCUSSION

Based on our previous theoretical results [2,3] obtained for coupling coefficient of quartz and langasite monolithic filters, we expected to have a smaller influence of mass-loading on electrical characteristics of langasite resonators than on those of quartz resonators.

We analysed the harmonic dependence of the most significant resonator parameters: effective mass-loading, motional inductance and quality factor. Because the behaviour of the effective coupling coefficient and motional inductance are opposite and strictly correlated, we will discuss in detail the harmonic dependence of inductance for various electrode parameters (material, thickness, diameter).

To compare the influence of the electrodes on quartz and langasite resonator parameters, eight plates from each type of crystal having identical diameters (14 mm) and resonant frequencies (5MHz), were deposited Ag and Au electrodes with 7 and 4.6 mm diameters and thickness of 75, 125, 200 nm.

The values of the effective mass-loading, motional inductance and quality factor, averaged over eight resonators, were calculated in every case.

3.1 Effective mass-loading μ_{eff}

Figure 1 shows the effective mass-loading variation with harmonic order for quartz and langasite resonators deposited with Ag electrodes having 7 mm diameter and 75, 125 and 200 nm thicknesses.

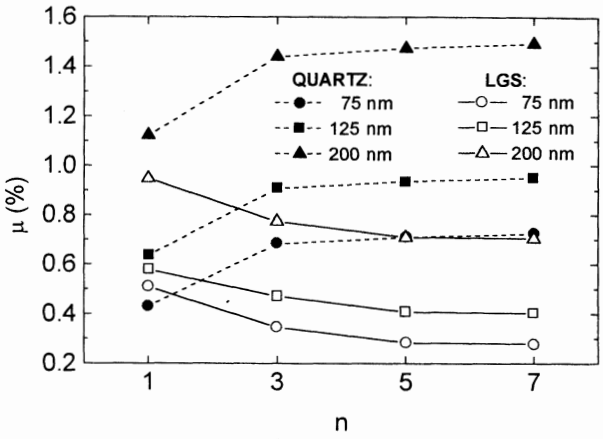


Fig.1Harmonic dependence of effective mass-loading for quartz and langasite resonators.

One observes that the effective mass-loading of quartz resonators is different from that of the langasite resonators: while for quartz resonators the effective mass-loading increases with harmonics, in the case of langasite resonators effective mass-loading decreases. The variation of effective mass-loading for various electrode thicknesses is smaller for langasite resonators than for quartz resonators.

3.2 Motional inductance L

The behaviour of the motional inductance is the most significant for comparison of mass-loading influence on quartz and langasite resonators. The inductance L was plotted versus harmonic order for quartz and langasite resonators (figure 2) with the same electrode parameters (Ag electrodes, 7 mm diameter and 75, 125, 200 nm thicknesses).

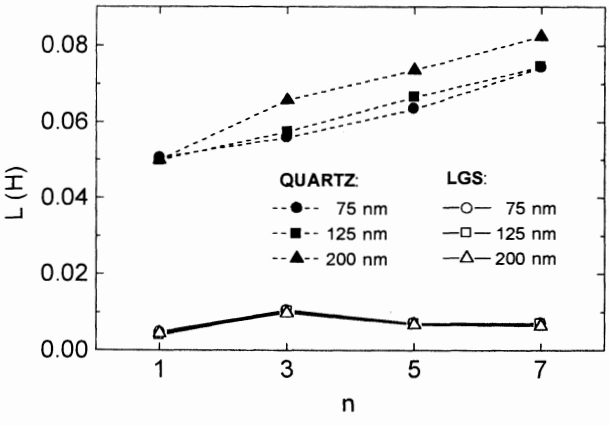


Fig.2 Inductance variation with harmonic order for quartz and LGS resonators

By analysing the figures we observe :

- the inductance of quartz resonators increases with harmonics for all electrode thicknesses, while the inductance of langasite resonators is almost constant;
- the change of electrode thickness determines a significant variation of inductance in the case of quartz resonators, but almost no change for langasite resonators.

Taking into account the previous results related to mass-loading effect [14,15,16] the change of inductance with harmonics was ascribed to the internal stress at interface electrode-piezoelectric substrate. We can conclude that the Y-cut langasite develops smaller stresses at interface than AT-cut quartz resonators having a similar behaviour with SC-cut quartz resonators. Consequently, the mass-loading influence on langasite resonators is lower than on quartz resonators due to the specific material constants of the two crystals and could be considered a new Y cut langasite property.

Figure 3 shows the harmonic dependence of motional inductance for quartz and langasite resonators with Au deposited electrodes having 100 nm thickness and 4.6 and 7 mm diameters.

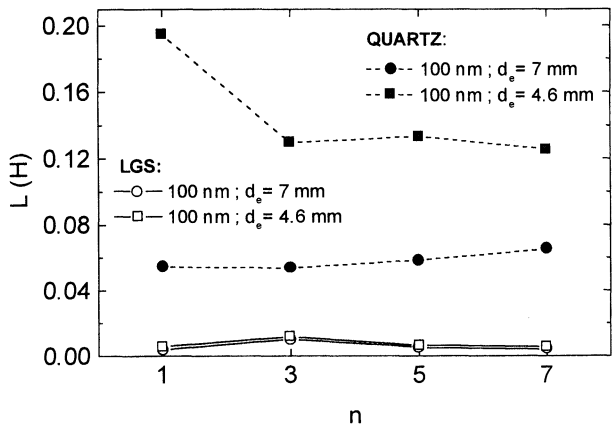


Fig.3. Harmonic dependence of inductance on electrode diameter.

The same conclusion (smaller mass-loading effect for langasite) can be drawn by comparing the quartz and langasite behaviour from figure 3. While the inductance of quartz resonators increases almost 2.5 times when the electrode diameter decreases from 7mm to 4.6mm, the inductance of langasite resonators presents a very small change with electrode diameter.

3.3 Quality factor Q

The quality factor values of langasite and quartz resonators deposited with Ag electrodes (7mm diameter, 100, 125, 150 nm thickness) are presented in figure 4.

Analysing this figure one observes that the normalized Q values measured for all electrode thicknesses are higher in the case of langasite (especially on third harmonic) than for quartz resonators. For both type of resonators the optimum electrode thickness which determines the maximum quality factor is 125 nm.

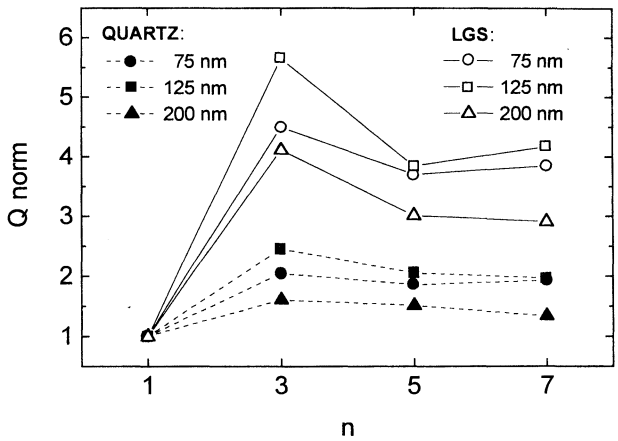


Fig.4 The quality factor dependence on harmonics for various electrode thickness

To see the quality factor dependence on electrode diameter for the same electrode thickness, in figure 5 is presented the normalized Q factor versus harmonics order for quartz and langasite resonators with 100nm electrode thickness and 4.6 and 7 mm electrode diameters.

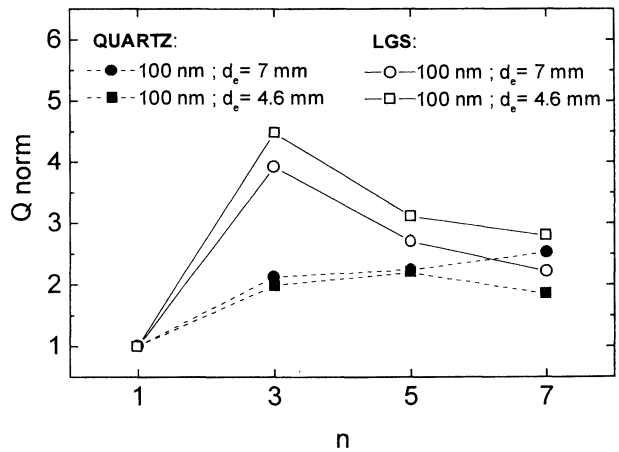


Fig.5 Quality factor variation with electrode diameter for quartz and langasite resonators.

The most important differences between the behaviour of quartz and langasite resonators are:

- for quartz resonators the quality factor increases with the electrode diameter change from 7 to 4.6 mm, while for langasite resonators Q factor decreases when the diameter increases;
- the increase of electrode diameter determines in the case of quartz resonators a shift of the maximum quality factor towards smaller harmonics, but the maximum Q factor of langasite resonators is situated at 3rd harmonic.

4. CONCLUSIONS

The experimental results obtained in this study allow to conclude that the mass-loading influence on langasite resonators characteristics is much smaller than in the case of quartz resonators. This behaviour is a feature of langasite, which has been pointed out for the first time in our theoretical calculation of coupling coefficient of monolithic structure as a function of mass-loading [2,3].

The analysis of the two types of resonators revealed the following differences between them:

- effective mass-loading increases with harmonic order for AT-cut quartz resonators and decreases for Y-cut langasite resonators;
- maximum variation of the effective mass-loading and inductance harmonic dependence for various materials, electrode thicknesses and diameters is significantly lower in the case of Y-cut langasite than for AT-cut quartz resonators;
- the quality factor of quartz resonators increases with decrease of electrode diameter and thickness; Q factor of langasite resonators decreases when the electrode diameter and thickness decrease;
- maximum Q factor values of the quartz resonators shift to lower harmonics with decrease of electrode diameter and thickness, while the langasite quality factor is situated at 3rd overtones.

Based on our previous results [13,14,15,16], according to which the inductance change with mass-loading is due to the internal stresses at interface electrode-piezoelectric substrate, we could ascribe the specific behaviour of inductance of Y-cut langasite resonators to a much smaller interfacial stresses in comparison with AT-cut quartz resonators.

The general conclusion related to mass-loading influence on langasite devices (resonators and monolithic filters) is that, due to the specific material constants of langasite crystal, the langasite devices are superior to quartz devices.

The present results could contribute at the optimization of the high performance devices.

5. ACKNOWLEDGEMENTS

The authors want to express their gratitude to Dr. Kelley Scott from Sawyer Research Products for the very good langasite samples.

Thanks are due to Dr. Mihai Popescu for comments and suggestions.

The authors acknowledge the help of F. Sava with the computer processing.

6. REFERENCES

[1] H.F.Tiersten, "Linear piezoelectric plate vibrations", New York, Plenum Press, 1969.

[2] I.Mateescu, G.Pop, A.Manea, M.Lazarescu, C.Ghita, "Investigation on behaviour of coupling coefficient of monolithic structure", in Proceedings of the European Frequency and Time Forum, Poland, 1998, pp.438-442.

[3] I.Mateescu, G.Pop, C.Ghita, "A new theoretical approach of coupling coefficient for quartz and langasite monolithic structures", in Proceedings of the IEEE Int. Frequency Control Symposium, USA, 1998, pp.271-277.

[4] A.Ballato, "Transmission-line analogs for stacked piezoelectric crystal devices", in Proceedings of the 26th Annual Symposium of Frequency Control, USA, 1972, pp.86-91.

[5] J.Kosinski, A.Ballato, S.Mallikarjun, "Mass loading measurements of quartz crystal plates", in Proceedings of the 43rd Annual Symposium of Frequency Control, USA, 1989, pp.365-377.

[6] A.Gotalskaya, D.Drezin, V.Bezdelkin, V.Stassevich "Peculiarities of technology, physical properties and applications of new piezoelectric material Langasite ($La_3Ga_5SiO_{14}$)", in Proceedings of the IEEE Int. Frequency Control Symposium, USA, 1993, pp.339-344

[7] I. Silvestrova , V. Bezdelkin V, P. Senyushenkov, Yu.Pisarevsky "Present stage of $La_3Ga_5SiO_{14}$-research", in Proceedings of the IEEE Int. Frequency Control Symposium, USA, pp. 348-350

[8] Yu. Pisarevsky, P. Senyushenkov, N.A.Moiseeva, "Elastic, piezoelectric, dielectric properties of LGT single crystals", in Proceedings of the IEEE Int. Frequency Control Symposium, USA, pp.742-747.

[9] G.Mansfeld , "Langasite as material for piezoelectric devices", in Proceedings of the 11th European Frequency and Time Forum, Poland, 1998, pp.350-357

[10] R.C. Smythe, "Material and resonator properties of langasite and langatite:a progress report", in Proceedings of the IEEE Int. Frequency Control Symposium, USA, 1998, pp.761-765.

[11] R.C. Smythe, R.C.Helmbold G.E. Hague, K.Snow, " Langasite, langanite and langatate resonators : recent results", in Proceedings of the Joint Meeting of European Frequency and Time Forum & IEEE Int. Frequency Control Symposium, France, 1999, pp.816-820.

[12] B.Chai, J.L.Lefaucheur, Y.Y.Ji, H.Qiu, "Growth and evaluation of large size LGS, LGT and LGN single crystals", in Proceedings of the IEEE Int. Frequency Control Symposium, USA, 1998, pp.748-760.

[13] I. Mateescu, E. Candet, "Non-uniform distribution of motion influence on the effective mass-loading in AT-cut quartz resonators", in Proceedings of the IEEE

Int. Frequency Control Symposium, USA, 1992, pp.561-566.

[14] I. Mateescu, I.V. Mateescu,"Complex mass-loading effects in AT-cut quartz crystal resonators", in Proceedings of the 7th European Frequency and Time Forum, Switzerland, 1993, pp. 63-66.

[15] I. Mateescu, J. Zelenka, I. V. Mateescu, "The contribution to the non-uniform distribution of motion effects obtained for the SC-cut quartz plates", in Proceedings of the 50th IEEE Int. Frequency Control Symposium, USA, 1996, pp. 405-411.

[16] I.Mateescu, I.V.Mateescu, G.Pop, "Comparison of mass-loading effects on the characteristics of AT and SC-cut resonators", in Proceedings of the Int. Piezoelectric Conference, Poland, 1996, pp. 14-19

[17] H.F.Tiersten, "Analysis of trapped-energy resonators operating in overtones of coupled thickness-shear and thickness-twist", Journal of the Acoustical Society of America, vol. 59, pp. 879-888, 1976.

[18] J.Kosinski, A.Ballato, I.Mateescu, I.V.Mateescu, "Inclusion of non-uniform distribution of motion effects in the transmission-line analog of the piezoelectric plate resonator: theory and experiment", in Proceedings of the IEEE Int. Frequency Control Symposium, USA, 1994, pp. 229-235.

[19] S. Laffey, M. Hendrickson, J. R. Vig, "Polished and etching langasite and quartz crystals", in Proceedings of the IEEE Int. Frequency Control Symposium, SUA, 1994, pp.245-250.

METHOD OF THE IMPROVEMENT OF THE SAW SENSOR SENSITIVITY

C. Kopycki, J. Filipiak
Institute of Applied Physics, Military University of Technology,
2, Kaliskiego St., 00-908 Warsaw, Poland

The main structural component of the sensor based on surface acoustic waves (SAW) is a delay line. It consists of two interdigital transducers (IDT). One of them generates a surface wave, which propagates in the surface region of the piezoelectric substratum. The wave reaches the receiving transducer after the delay time. The principle of the SAW sensor's operation is based on measuring the changes of SAW's delay time. The above changes are produced by influence of measured quantity on the piezoelectric surface region between IDT. The sensitivity of the sensor mentioned above is proportional to the changes of the SAW delay time. Mentioned values are up to few μs, therefore they are very small and difficult to measure. The change of the delay time is proportional to the length of the wave propagation path. Therefore, when the region between interdigital transducers is longer, changes of a delay time are bigger, so a sensitivity of the SAW sensor is higher. The distance between IDTs is limited by the length of the piezoelectric substrate. Moreover, long substrates put up the cost of the sensor and increase its dimensions.

A long wave propagation path can be obtained by using multiple transit signals (MTS) in a delay line. MTS are generated, in a SAW delay line, during reflection of a surface wave caused by an electrical regeneration. Transit signals' theoretical analysis was conducted. It allowed to maximise an amplitude of MTS and to obtain signals separated in time. Theoretical results were put up for experimental verification. Special measuring system, allowing to measure changes of delay time of reflected signals (MTS) between transducers as a function of changes state of the surface region, was designed and realised. The result of the measurement shows that the SAW sensor sensitivity is proportional to the number pathways of the wave between transducers. Presented method is a simple way to increasing (up to a few times) a sensitivity of the SAW sensor for measurements all non-electrical quantities.

Corresponding author:
Cezary Kopycki,
Military University of Technology,
Institute of Applied Physics,
2, Kaliskiego St., 00-908 Warsaw, Poland
tel. (+48 22) 685 48 51, fax (+48 22) 685 91 09
e-mail ck@wic.wat.waw.pl

THE SAW VIBRATION SENSOR

C. Kopycki, J. Filipiak
Institute of Applied Physics, Military University of Technology,
2, Kaliskiego St., 00-908 Warsaw, Poland

The principle of operation of the SAW vibration sensor is based on measuring the changes of SAW's delay time. The above changes are produced by influence of vibration on the piezoelectric surface region between interdigital transducers. The practical approach to the SAW vibration sensor requires investigation of the following subjects:

a) construction of the sensor,
b) way of the SAW delay-time measurement,
c) recording of the measurements.

During the above investigation the analysis of the piezoelectric-beam vibration was conducted. Theoretical results were put up for experimental verification. The breaking stress was determined for the investigated substrates. The described steps of the sensor realisation allow to determine dimensions of the substratum, and to select an aggregated mass (mounted on the free end of the beam). This part of the research allows to choose piezoelectric material basing on its mechanical properties.

The next task concerning the SAW vibration sensor is realisation of the SAW delay line. It was designed, and thereafter made using photolithography method. Further on the way of SAW delay-time measuring was developed. It allows to prepare a measuring system. Both methods: based on the SAW generator and the phase detector were analysed. During the selection the following criteria were taken into consideration: sensitivity and a resolution of the sensor, and simplicity of the measuring system. Finally, measurements was realised by the measurement of the frequency of the generator with delay line in its phase-locked feedback loop. The frequency of the SAW generator is in the range from a few tenth up to several hundred MHz. The range depends on the required parameters and the device's construction. The changes of the frequency caused by the measured acceleration are about 10 to 90 kHz so they are small in comparison to the frequency of the generator. The second generator with the constant frequency, nearly the same as the frequency of the main one, must be used and only measuring of the beat frequency allows us the precise measurement of the selected quantity.

The selected technical data of the SAW vibration sensor, which was realised are as follows:

a) measurement range: ±2,5g (g≈10 $\frac{m}{s^2}$), frequency range (0÷20)Hz,
b) sensitivity: 8,5 $\frac{kHz}{g}$, voltage sensitivity 180 $\frac{mV}{g}$

Mentioned above parameters of the sensor allows to refer it as a high sensitivity.

Corresponding author: Cezary Kopycki,
Military University of Technology,
Institute of Applied Physics,
2, Kaliskiego St., 00-908 Warsaw, Poland
tel. (+48 22) 685 48 51, fax (+48 22) 685 91 09
e-mail ck@wic.wat.waw.pl

MINIMIZATION OF AN AUTODYNE SENSOR'S ERROR, STIPULATED BY THE ANTENNA'S HEATING.

E. SAFONOVA
Kazan State Technical University, address: Kazan-420141, Kul-Gali st., 1-149, Tatarstan, Russia; e-mail: kate@cg.ru , Tel: +7(8432)-637-687; Fax: +7(8432) 766737 or 762758

ABSTRACT

The present paper is a continuation of the author's works on creation of the microwave self-excited oscillating sensors of high-temperature technological processes [1], mainly processes of burning. Sensor's circuit includes a flush-mounted antenna, emitting into monitoring medium. Then its admittance is transformed into a measuring signal - frequency of generation. Temperature error is stipulated by two reasons: natural frequency instability of self-excited oscillator (SEO) and antenna's influence during its non-stationary heating. Investigation of the last one is a subject of the present paper.

The research, conducted in dispersing approximation, includes:

- An approximation of four components of antenna's complex conductivity and transmission line temperature's dependencies,
- An analytical description of complex reflectivity from antenna $\dot{\Gamma}(T) = \mathrm{Re}\,\Gamma(T) + \mathrm{Im}\,\Gamma(T)$, which influence to the frequency of generation is investigated earlier in [1],
- Linearization of the found dependencies and searching of minimums $\mathrm{Re}\,\Gamma(T)$ and $\mathrm{Im}\,\Gamma(T)$.

It is shown, that under the random combination of electric and constructive antenna's parameters , such minimums are absent. Modifications of constructions, where an antenna's parts are made of an electrotechnical materials with various temperature coefficients (TC), are suggested. The relations between TC, providing the constancy of $\mathrm{Re}\,\Gamma(T)$ and $\mathrm{Im}\,\Gamma(T)$ in a temperature range (that is the average quadrate of temperature fluctuations), are found. This range is limited by the linear approximation.

The nonlinearity account of $\dot{\Gamma}(T)$ is effected by recovery of second members of the Tailor's series, discarded under the linearization and let to determine the boundaries of the specified range. An average quadrate of frequency's temperature fluctuations of the sensor and measurement error do not exceed an allowed significance within the limits of these boundaries.

The suggested technique is considered in more detail with reference to the ring slit antenna [1]. Here, the choice of TC provides variations of the slit width, more essential, than in traditional construction. Such way compensates not only a slit's capacities growth, but also an antenna's conductivity of emission.

Because the antenna's contribution to the frequency in sensors of the given type is dominant, the obtained result allows to improve stability more than on the order.

1. INTRODUCTION

. Head part of one of the sensor's models with an antenna is shown on fig.1.

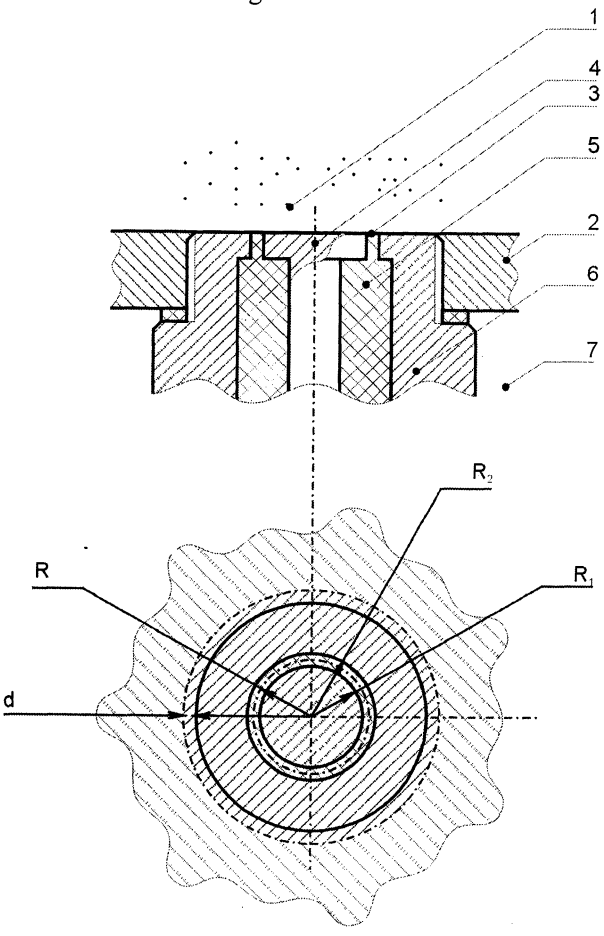


fig.1 Antenna's construction

Sensor is set up in a reactor's wall 2 with an operating body 1 being monitored, which has temperature T_1. The particular realization of an antenna is a ring slit 3, cut in short-circuited end of coaxial transmission line (4,5,6). The oscillator with a coaxial resonator (it isn't shown on fig.1) is in an external medium with temperature $T_0 << T_1$. Sensor's dynamics is described in [1].

The measuring signal is formed due to generation frequency dependence on an antenna's parameters - conductivity $\dot{y}_A$ and reflectivity $\Gamma_A(\dot{y}_A)$.

In the working position antenna is heated up to temperature $T_A < T_1$. During the propagation in a sensor's body the temperature waves are attenuating. The application of dielectric with low temperature diffusivity for main thermoleading details manufacturing promoted this. Antenna is required to be metallic, herefrom its influence to the frequency instability and measurement error is dominant. Fast response to the temperature variations is characteristic for antenna, while the heating of the sensor is rather long.

2. PROBLEM DESCRIPTION

Stabilization of reflectivity Γ_A is necessary for reaching the purpose, pointed in a title. Authors [2] have used here the distributed in engineering effect, based on temperature coefficients (TC) difference of two details, forming an internal and external edges of a slit. In our case their results could not be applied directly, because the antenna is principally displaced and $d\dot{\Gamma} \,/\, d\dot{y}$, $d\dot{\Gamma} \,/\, dT$ are big. However the mathematical statement of a problem coincides. In the specific temperatures range ΔT the reflectivity should satisfy to condition

$$\left|d\dot{\Gamma}/dT\right| \le \delta = D/\Delta T \quad , \qquad (1)$$

where the permissible error D is connected to frequency instability by a rather simple ratio [1]. With allowance that $\dot{\Gamma}_A = |\Gamma_A| \exp(j\varphi_A)$ the condition (1) is divided into two parts: for module and for phase. In the consequent computations it seems convenient to use a little bit changed statement:

$$\frac{d\,\mathrm{Re}\Gamma(T)}{dT} \le \delta_1; \quad \frac{d\,\mathrm{Im}\dot{\Gamma}(T)}{dT} \le \delta_2 \qquad (2)$$

The further material of the paper is devoted to determination of the ratios between antenna's conductivity components and transmission line, sizes of details and TC of materials, under which the conditions (2) are observed.

3. APPROXIMATION

Graph-analytical dependencies of conductivities g_A and ωC_A are obtained in work [2], the internal R_1 and external R_2 radiuses of slit act as arguments, referred to a wavelength. To say more exactly, arguments are: for g_A - is $(R_1 + R_2)/\lambda$, for C_A - is (R_2/R_1). Under the realization of graphic dependencies approximation, it seems convenient to use their half-sum r - average radius and dimension d - slit's width, also referred to the wavelength λ.

General view of approximating functions is:

$$g_A(T) = A \cdot [1 - d(T)/\rho] \cdot [r(T) - r_n]^2 \quad ,$$

$$C_A(T) = B \cdot R_1 / \left[1 + \tfrac{d(T)}{r(T)}\right]^{\alpha} \quad , \qquad (3)$$

where A, B are connected with an average magnitudes g_A and C_A under the nominal temperature, $r_n \sim 0{,}05\lambda$, $\rho \sim 0{,}01\lambda$ $(\rho > d)$, $\alpha \sim 1{,}5...2$.

With allowance of well-known dependence of sizes from temperature and under the condition of $K_T \Delta T = K_{1,2} \Delta T << 1$, where $K_{1,2}$ - is TC of details 4 and 6, let's transform the ratio (3) to Tailor series. Then in a linear approximation we receive:

$$g_A(T) \approx g_A \left\{1 + r\left[\frac{K_2 + K_1}{r - r_n} + \frac{K_2 - K_1}{d - \rho}\right] T\right\}$$

$$C_A(T) \approx C_A\{1 + K_1 T - \alpha[\frac{d}{r} \bullet$$

$$\bullet (K_2 + K_1) - (K_2 - K_1)] \cdot T\},$$

correspondingly the magnitudes of derivatives (when $d/r << 1$) is the following:

$$\frac{dg_A(T)}{dT} = r \cdot \left[\frac{K_2 + K_1}{r - r_n} + \frac{K_2 - K_1}{d - \rho}\right] \cdot g_A \quad ,$$

$(\rho > d,\ r > r_n)$

$$\frac{dC_A(T)}{dT} = [K_1(1 + \alpha) - \alpha K_2] \cdot C_A . \qquad (4)$$

Wave conductivity g and internal capacity C_r of excitation line mainly depend on permittivity. A dielectric 5, chosen by the recommendation [7], has TC two orders less than $K_{1,2}$. Therefore further we take into account only average magnitudes of g and C_r; and their contribution into fluctuation - is only qualitative.

4. BASIC EQUATIONS

Let's present reflectivity as a function of temperature:

$$\dot{\Gamma}_A(T) = \frac{g_A(T) - g + jb(T)}{g_A(T) + g + jb(T)} .$$

By replacing $g_a(T) = g_A + \Delta g_A(T)$ and $b = \omega C(T) = \omega[C + \Delta C(T)]$, $g(T) = g + \Delta g(T)$ with allowance for increment's smallness we have:

$$\dot{\Gamma}_A(T) \approx \dot{\Gamma}_A[1 + d_1 + d_2] \ , \qquad (5)$$

where $d_{1,2} = \dfrac{\Delta g_A \mp \Delta g + j\omega\Delta C}{g_A \mp g + j\omega\Delta C}$, symbol T as an increments argument hereinafter is omitted. Set of real and imaginary parts and differentiating, we shall receive:

$$\frac{d\,\mathrm{Re}\Gamma_A(T)}{dT} = R_1 \cdot \frac{dg_A}{dT} + R_2 \cdot \frac{dg}{dT} + \\ + R_3 \cdot \left(\frac{dC_A}{dT} + \frac{dC_r}{dT} \right)$$

(6a)

$$\omega \frac{d\,\mathrm{Im}\Gamma_A(T)}{dT} = -R_3 \cdot \frac{dg_A}{dT} + \omega R_2 \cdot \frac{dg}{dT} + \\ + \omega^2 R_1 \cdot \left(\frac{dC_A}{dT} + \frac{dC_r}{dT} \right)$$

(6b)

Magnitudes $R_{1,2,3}$ we shall reveal below. By virtue of the mentioned smallness of $\dfrac{dg}{dT}$ and $\dfrac{dC_r}{dT}$ let's present the corresponding members in both expressions by small constant magnitude μ (just not to lose it), saving this sign irrespective of a multiplying it on other magnitudes.

Substituting (4) in (6), we have

$$\delta_3 - \left|\mathrm{Y}_1 K_1 + \mathrm{Y}_2 K_2 + \mu_1\right| > 0 \ , \\ \delta_4 - \left|\mathrm{Y}_3 K_1 + \mathrm{Y}_4 K_2 + \mu_2\right| > 0 \ , \qquad (7) \\ \mathrm{Y}_{1,2} = R_1 \nu_{2,1} \mp R_3 (\xi_{1,2} + \alpha),$$

Where $\mathrm{Y}_{3,4} = -R_3 \nu_{1,2} \mp \omega^2 R_1 (\xi_{1,2} + \alpha)$,

$$\nu_{1,2} = \left[\frac{r}{r - r_n} \pm \frac{r}{d - \rho} \right]$$

$\xi_{1,2}$ - Kronecker delta ($\xi_1 = 0$, $\xi_2 = 1$),

$R_1 = [(g_A - g) + b]/M_1 - [(g_A + g) + b/M_2]$

$R_3 = b(1/M_1 - 1/M_2)\ \delta_{3,4} = \delta_{1,2}/\sqrt{M_2}$,

$M_{1,2} = (g_A \mp g)^2 + b^2$, dimension of inequalities (Ohm/degree)

Solving the system of inequalities, by taking for definiteness K_1 as argument, we shall receive an area of the permissible magnitude K_2, corresponding to the posed problem shown on fig.2:

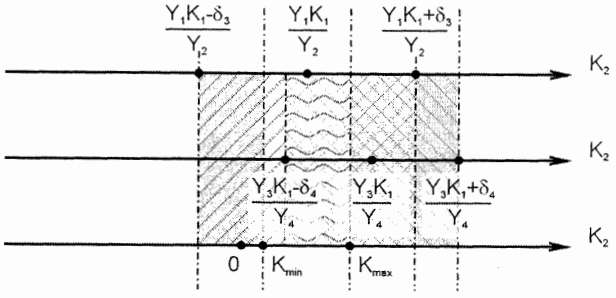


fig.2 Graphic solution of the inequalities system (7)

$$\frac{\mathrm{Y}_3 K_1 - \delta_4 + \mu_2}{\mathrm{Y}_4} \le K_2 \le \frac{\mathrm{Y}_3 K_1 + \delta_4 + \mu_2}{\mathrm{Y}_4}$$

$$\frac{\mathrm{Y}_1 K_1 - \delta_3 + \mu_1}{\mathrm{Y}_1} \le K_2 \le \frac{\mathrm{Y}_1 K_1 + \delta_3 + \mu_1}{\mathrm{Y}_2} \ ,$$

$$K_{\min} < K_{1,2} > K_{\max} \sim 10^{-5} \qquad (8).$$

The third inequality reflects the known physical fact. Preliminary estimations showed that the magnitudes $\mu_{1,2}$ do not exert the practical influence.

5. SPECTRAL DENSITY

The heat exchange between monitored operating body and chamber's wall happens by a radiation and convection. Last one predominates [3]. The amplitudes of temperature oscillations are connected by a ratio [4]: $T_A = T_1/\sqrt{1 + q/H}$, where $q = \sqrt{\omega/2a}$, ω - angular frequency, a -temperature diffusivity factor, $H = \alpha_T/\eta$, α_T -thermoreturn coefficient, η - heat conduction. For metals on frequency F>10Hz the ratio is $q/H > 1$.

The research level of spectral density (SD) $S_T(F)$ cannot be considered as sufficient, especially in a high-frequency (F>10Hz) range. Our point of view is based on temperature with pressure and acoustic occurrence correlation, which are investigated better. On fig.3 curves 1 and 2 correspond to SD of $S_1(F)$ and $S_A(F)$. The form of SD envelope of reflectivity $S_\Gamma(F)$ repeats $S_A(F)$. The curves 3 and 4 in this case reflect two situations: non-observance and tend to observance of conditions (8).

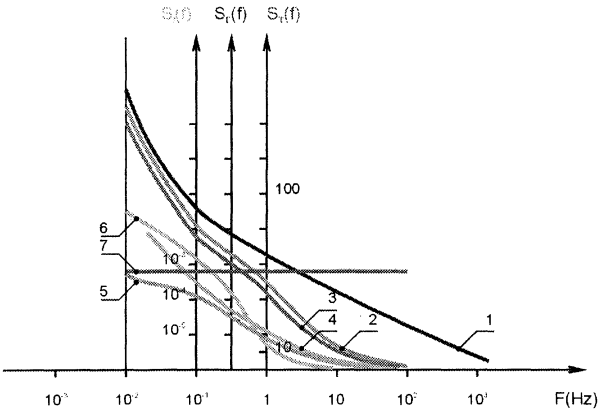


fig.3 Spectral densities

At last the curves 5 and 6 are the SD envelopes of frequency fluctuations, generated by an antenna (5) and resonator of the sensor (6). They are represented in relative units. In the last case, as it is shown in [5], the damping of high-frequency temperature waves is essentially more severe. In case of equality of frequency fluctuations on dispersion (this is comply with optimum designing norms), the contribution of an antenna to a high-frequency part of total SD $S_f(F) = S_5(F) + S_6(F)$ is more essential. Between that, the interest to the microwave sensors is connected just with this part of a spectrum.

Permissible level of frequency fluctuations, proportional to δ^2, is arbitrarily shown by horizontal line on fig. 2. It is seen that the suppression of low-frequency fluctuation processes is problematic.

6. NUMERICAL METHODS AND MINIMIZATION.

Minimization of frequency fluctuations now consists in magnification of the area of three inequalities intersection, down to full overlapping. To be exact, this will be a location of two first inequalities in the boundaries of third, which causes a technical solvability of the problem. For us these boundaries are physical constants.

This operation is made by choice of magnitudes $Y_{1..4}$. For this purpose three independent parameters can be used only: r, d and g. Other magnitudes or their functions (conductivity g_A and b), or constants (approximating factors r_n, ρ, α and implicitly present digit 2).

The following equality corresponds to the optimum situation:

$$Y_1 Y_4 = Y_2 Y_3 \qquad (9),$$

under which fulfillment the temperature fluctuations of a reflectivity in an accepted approximation are equal to zero (if not to take into account influence of g and C_r fluctuations being displayed by μ). Now the boundaries of two inequalities are symmetric relative to shareable point $K_2 Y_1 / Y_2 = K_2 Y_3 / Y_4$. Permissible errors while are different ($\delta_3 \neq \delta_4$), therefore observance of the other condition is desirable:

$$\delta_3 / Y_2 = \delta_4 / Y_4 \text{ or } \delta_3 Y_4 = \delta_4 Y_2, \qquad (10)$$

then also boundaries of inequalities will coincide. That means, the consequent selection of metals with TC $K_{1,2}$ will minimize errors of $\mathrm{Re}\dot{\Gamma}$ and $\mathrm{Im}\dot{\Gamma}$ at the same extent.

Narrow interval of TC magnitudes of electrotechnical metals (Kmin, Kmax) eliminates reaching of these conditions. Additional restrictions are: high transmission factor of working function preservation by the sensor: [1], for what $\omega C_A > g_A$, but g_A is rather big (not less than a few mSm). The last one ensures principally necessary antenna emission.

The analytical expressions obtained are hard-reviewed. The numerical analysis for global extremum searching is necessary. Here we shall limit by an evaluation of parameters and recommendations for improving frequency thermostability of concrete sensors, investigated experimentally before.

The sensor of permittivity of an ionized gas (flame). Parameters:

$g_A \approx (5...8) mSm$, $g = 20 mSm$,

$b = \omega C_A = (20...40) mSm$, $r = 0{,}075$

($R_2 = 5{,}6 mm$), $d = (1{,}5...3)10^{-3}$ (slit width is 0,3 $\pm$ 0,1 mm), $\lambda = 75 mm$ (f=4Hhz). It is seen that sensors differ only by the slit width.

Calculation has shown that the members of equality (10) differs in 1,3...1,6 times, and of equality (11) - little more. That means construction and electrical parameters are not optimal. Under the permissible instability δ_f and $\Delta T = 20^\circ$ the area of inequalities intersection is small, though the metals are selected on the ratio in [7].

7. DISCUSSION

Let's explain the physical processes, which take place here. The essential feature of antenna, that slit is non-resonance and its length is $L = 2\pi R < \lambda/2$ (or $r < 1/4\pi$). Under the heating the length is "getting closer" to resonance (albeit not much), and g_A is growing. At the same time the slit is widening and C_A is reducing. But this reduction is not enough for achievement the constancy of Γ_A (or, backward, too much) under the equal TC ($K_1 = K_2$). It is necessary

to have $K_1 < K_2$ or $K_1 > K_2$. Under the unsuccessful choice of base electrical parameters, i.e. average magnitudes of g_A and C_A, it could happen, that the necessary (from two first inequalities (7)) value $K_1 < 0$. Than the compensation is possible under the condition of electric magnitudes- conductivities variations. The suggested compensation method (as [ ]) in any case is effective in a limited temperature range, the reason is: nonlinearity of dependencies of $g_A(T)$, $C_A(T)$ and $\Gamma_A(T)$. At the same time there's no need to magnify this influence, because the maximum magnitude of KT under the fluctuations smaller than 10^{-2}, and in the interesting frequency range (more than few Hz) $<10^{-3}$. Corresponding nonlinear errors have the order $(KT)^2 < (10^{-4}...10^{-6})$. The other reason of error is multiplicative: dependence of conversion factor $d\omega/d\varepsilon$ (ε - monitored parameter) from g_A and C_A, therefore from T and $K_{1,2}$. In this case this dependence is bigger, than without the compensation. At last, the third error: number of electrotechnical metals (application of which is economically reasonable) is not much. This error could become the main one. Indeed, under the requested [ ], replacement of ruthenium ($K_1 = 9 \cdot 10^{-6}$), which is for sure more expensive than gold, by titanium ($K = 8{,}2 \cdot 10^{-6}$), applied to mentioned above examples, will give an error $10^{-3}...10^{-4}$.

8. CONCLUSIONS

- Bimetallic antenna application facilitates stability of an autodyne sensor frequency, subjected to the heating, and allows to reduce the monitoring error.
- Physical mechanism of stabilization is specified by finding the proportion between active and reactive components of antenna's admittance, which depend differently (and nonlinearly) from ring slit radius and width.
- TC values of two metals are diverse here, unlike in the matched antenna [5]. Their optimum values correlate with antenna's conductances and transmission line. Last interrelation could be considered nontrivial. Latter influences on the sensor's operating characteristics [1]. That is why the optimum could be obtained not always.
- Mutual compensation of temperaturedependent parameters, specifying high frequency stability, is obtained in a limited range of temperature fluctuations. This is explained by the referred nonlinearity. Fast fluctuations, which have smaller amplitudes, are being suppressed more effectively. That again proves the sensor's efficiency in monitoring of the fast processes.

Special thanks to A.Avksentiev for consulting.

REFERENCE

1) V.Boloznev, E.Safonova, "Diode microwave active oscillator with a multi-link resonator", Proc. International Symposium on Acoustoelectronics, Frequency control and Signal generation, pp.139-143, Russia, June 1998.

2) A.Avksentiev, N.Vorobiev, N.Stahova "Proektirovanie slabonapravlennih nevistupauschih antenn" Kazan 1984

3) F.Kreith, W.Z.Black, "Basic heat transfer": N.Y. etc., 1980

4) A.V. Luikov " Heat conduction theory"[in Russian], Moscow: High school, 1967, p.600

5) V.Boloznev, E.Safonova, "Microwave oscillator as a sensor of electrophysics flow parameters", Proc. 13th EFTF, Aprile 1999pp.1062-1066, Besancon, France

6) V. Boloznev, A.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 October 20, 1991

7) A.Avksentiev, N.Vorobiev and others, SU Inventors Certificate № 1628111, BI, № 6, 1991

SILICON STRUCTURES MICROMACHINED IN TMAH AND NAOH SOLUTIONS : CHARACTERIZATION OF THE ANISOTROPY AND COMPUTER AIDED DESIGN OF STRUCTURES

C.R. TELLIER, A. CHARBONNIERAS, C. HODEBOURG and T.G. LEBLOIS

Laboratoire de Chronométrie, Electronique et Piézoélectricité
Ecole Nationale Supérieure de Mécanique et des Microtechniques
26 Chemin de l'Epitaphe, 25030 Besançon Cedex, France
Tél. 33 (0) 3.81.40.28.30 – Fax 33 (0) 3.81.88.57.14 – e.mail : Colette.Tellier@ens2m.fr

ABSTRACT

TMAH and NaOH etchants are used to micromachine beams on various (hhl) silicon planes. An experimental and comparative study of the changes in the shape and in the width of beams with the cut and with the direction of alignment is undertaken. In addition for TMAH theoretical shapes are derived from the CAO tool TENSOSIM and the effect of misalignment is theoretically investigated.

1. INTRODUCTION

Resonant structures composed of a seismic mass suspended by bridges or of vibrating beams can be fabricated in silicon plates by micromachining [1,2]. Moreover similar structures and cantilever loaded by a mesa serve also of mechanical structures for piezoresistive or capacitive silicon sensors such as accelerometers [3]. Up to now KOH or EDP based etchants [1,2,4,5] are usually used to fabricate silicon structures essentially in {100} and {110} wafers. Recently interest in TMAH [6] and NaOH [5] etchants is revived. In particular TMAH solutions are fully compatible with IC fabrication processes. These etchants are highly anisotropic and as a result width of beams and of cantilevers and shapes of structures depend on the orientation of the silicon plate and of the direction of alignments of the mask pattern. It should be also outlined that it may be of interest to fabricate structures in non-conventional plates such as {hhl} plates. Consequently it becomes necessary to characterize the anisotropy of the chemical attack of silicon plates in TMAH and NaOH solutions.

For this purpose, in this paper a study of the anisotropy is made for two solutions : TMAH/Water (25% in weight) and NaOH (37% in weight). Experimental results on shapes of beams produced on various (hhl) plates are reported with a special attention on technical problems caused by the anisotropic chemical attack such as formation of limiting facets, lateral underetch and development of sharp profiles. A comparison of the two etchants is undertaken from a practical point of view (width w of beams, shape of cross-section). Finally the simulation tool TENSOSIM previously derived by C.R. TELLIER [4,6] is used to obtain theoretical shapes of beams etched in TMAH and to outline effects due to a misalignment of the mask pattern or to small changes of the wafer orientation.

2. EXPERIMENTALS

2.1. Experimental details

The starting materials were two sides polished singly or doubly-rotated silicon plates whose orientation are listed in table 1. A 1.3 µm thick masking oxide was grown at 1100°C in steam. The beams were fabricated using standard bulk micromachining techniques. A side of plates was patterned into a serie of beams with various alignment angles Ψ_A. Figure 1 gives the mask pattern and the corresponding designation for the various beams with initial width w_o= 600 µm.

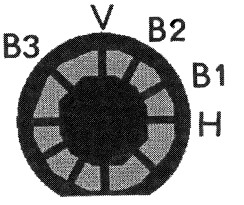


Figure 1 : The mask pattern

The horizontal beam (H) was theoretically aligned along a direction indicated in table 1. Patterned plates were etched in TMAH (25% in wt) and in NaOH (37% in wt) solutions maintened at a constant temperature of 70°C and 80°C respectively. The total durations of etching, for TMAH and for NaOH, were chosen to avoid the complete opening of beams. In this condition only cavities are formed and it remains possible to study the geometrical features of the etched bottom surfaces. Etched surfaces and general shape of beams were examined by Scanning Electron Microscopy. Width w of etched beams were measured at ± 5 µm using an optical microscope. In addition depths of etch Δd were evaluated from profilometry studies.

Plate	(100)	(110)	(112)	(441)
Δd ,TMAH	50	104	100	120
Δd ,NaOH	120	190	110	120
<D>	<110>	$<D_{\{111\}}>$	<110>	<110>

Table 1: Depth of etch and direction <D> of alignment

2.2 Experimental results

2.2.1 Etched surfaces

Even if in practice beams are open it remains of interest to compare briefly for the two etchants the geometrical features of etched surface because degraded surfaces are frequently associated with the formation of terraced

edges [4]. Figures 2 and 3 give SEM images of the bottom surfaces etched in TMAH and NaOH solutions respectively. We observe that {100} surface excepted the NaOH etchant causes a marked degradation of surfaces with the formation of terraced surfaces (on (112) and (441) planes) or of sharp hillocks on (110) plane. In particular on deeply etched (110) plane ($\Delta d \approx 190$ µm) the peak to valley height R_{max} as evaluated from profilometry traces exceeds 20 µm.

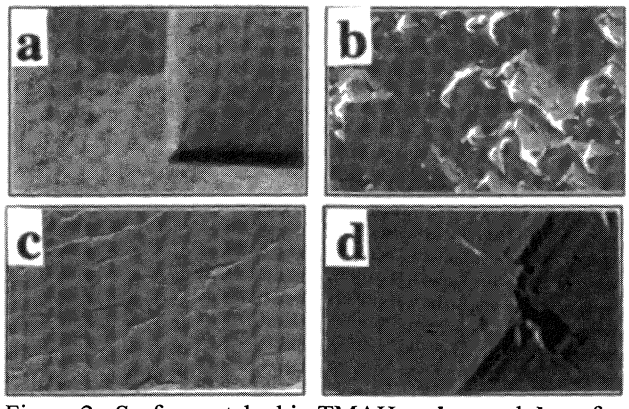


Figure 2: Surfaces etched in TMAH, **a** ,**b**, **c** and **d** are for (100), (112), (441) and (110) surfaces respectively.

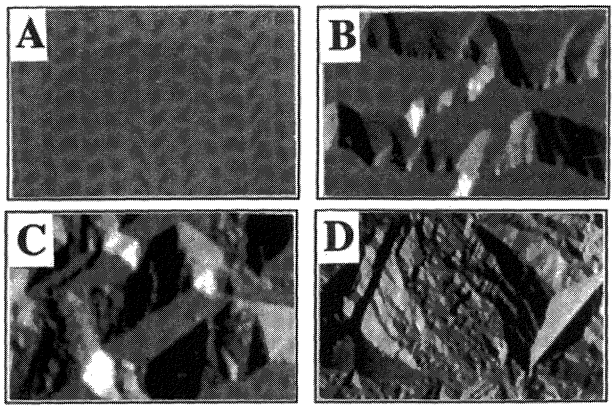


Figure 3: Surfaces etched in NaOH, **A** ,**B**, **C** and **D** are for (100), (112), (441) and (110) surfaces

2.2.2 Width of beams.

The final width w of beams formed on the differently oriented silicon plate are reported in table 2. To make the comparison between TMAH and NaOH etchants more easier the reduced decrement in width, $\Delta^* = \Delta w/\Delta d$, is also indicated in this table using the experimental values of the depth of etch, Δd, collected in table 1.

Examination of tables 1 and 2 reveals several features :

i) Turning attention to planes (100) and (111) for which the ratio Δd (110)/Δd (100) are crudely similar for the two etchants we observe that for most of beams changes in anisotropy are less rapid for NaOH than for TMAH when orientation of plate changes from (100) to (111). A behaviour which is not strictly verified for (hhl) planes.

ii) For most of the cuts investigated here the evolution of the ratio $\Delta w/\Delta d$ with the angle of alignment Ψ_A is not so marked for NaOH solution than for TMAH etchant. Only the (112) plate departs from this behavior which is particularly accentuated for (100) and (110) plates.

iii) For plates (100), (110) and (112) horizontal beams (H, $\Psi_A = 0°$) exhibit small decrement Δw in width as expected for beams aligned in theory along the intersection <D> of a {111} plane with the reference plate. The fact that beam H on (441) plane behaves differently indicates a misalignment with respect to the <110> intersecting direction.

2.3.2.General shapes of etched beams

SEM images of built-in structures are displayed in figures 4 to 9. A comparative examination of the general shape of beams etched in (110) plates with the two etchants (Figs. 4 and 5) allows us to draw some remarks:

i) Far from corners the cross-sectional profile of a beam is determined by the direction of alignment because we are concerned with a 2D etching process [4]. So rather planar facets (Figs 4 a, b,c) seem to bound beams in the case of TMAH etchant. On the contrary for the NaOH solution edges appear to be more terraced or terminated by terraced shoulders (Fig. 5C). This observation is also valid for beam H which is expected to be bounded only by vertical {111} planes.

ii) For the two etchants {111} planes which dissolve very slowly [4-6] participate to the concave corner undercutting and as a consequence top contours of

TMAH									
Beam	Ψ_A(°)	(100)		(110)		(112)		(441)	
		w	$\Delta w/\Delta d$	w	$\Delta w/\Delta d$	w	$\Delta w/\Delta d$	w	$\Delta w/\Delta d$
H	0	590	0.2	585	.14	560	.4	370	1.91
B1	30	410	3.8	385	2.07	380	2.2	390	1.75
B2	60	410	3.8	540	.58	490	1.1	420	1.5
V	90	590	.2	400	1.93	390	2.1	390	1.75
B3	135	510	1.8	410	1.83	435	1.65	485	.96
NaOH									
H	0	580	.17	560	.21	570	.27	440	1.33
B1	30	315	2.37	380	1.15	310	2.63	455	1.2
B2	60	320	2.33	515	.45	525	0.68	480	1
V	90	580	.17	420	.95	390	1.9	410	1.58
B3	135	345	2.12	320	1.45	355	2.22	580	0.17

Table 2 : width w and reduced width $\Delta^* = \Delta w/\Delta d$ for the various beams etched in (hhl) planes.

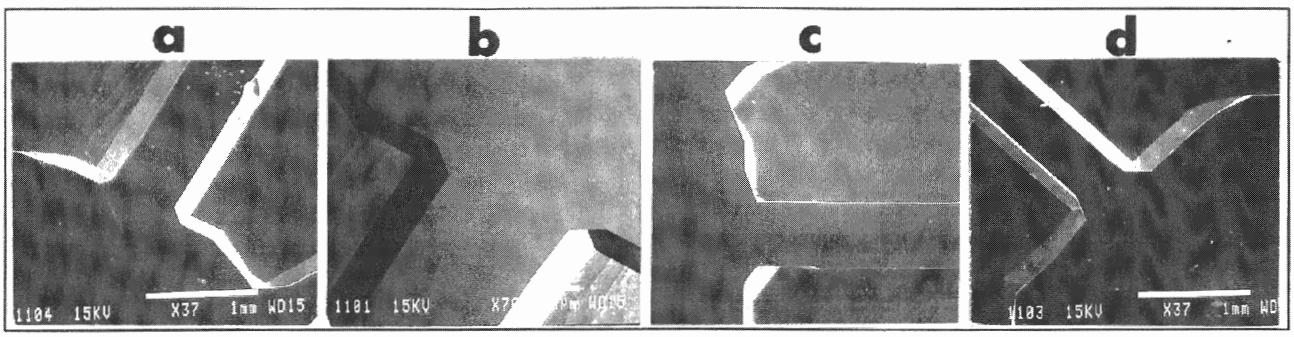


Figure 4 : SEM images of built-in beams B2 (**a,b**) , H (**c**) and B3 (**d**) etched in a (110) plane by TMAH etchant

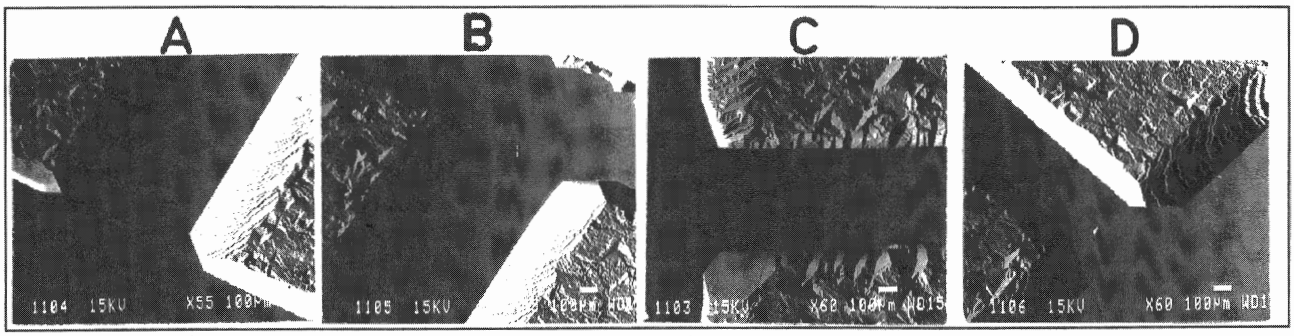


Figure 5 : SEM images of built-in beams B2 (**A,B**) , H (**C**) and B3 (**D**) etched in a (110) plane by NaOH etchant

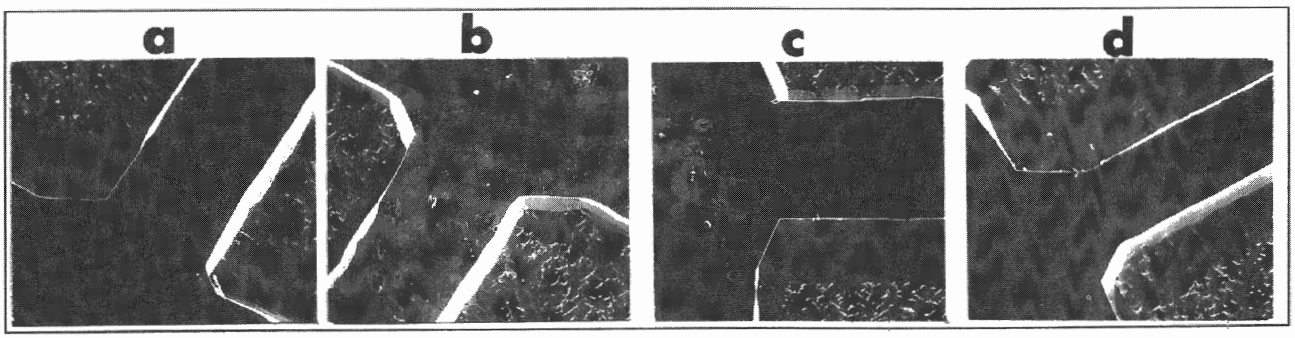


Figure 6 : SEM images of built-in beams B2 (**a,b**) , H (**c**) and B1 (**d**) etched in a (112) plane by TMAH etchant

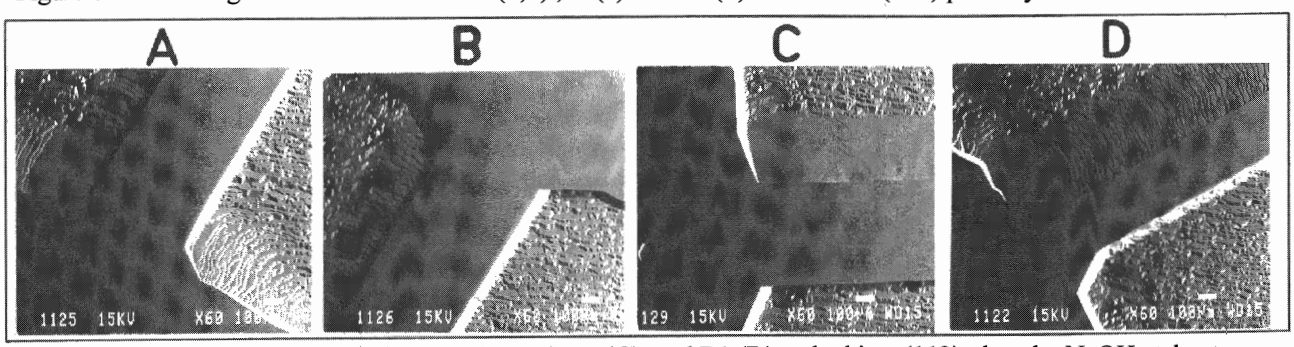


Figure 7 : SEM images of built-in beams B2 (**A,B**) , H (**C**) and B1 (**D**) etched in a (112) plane by NaOH etchant

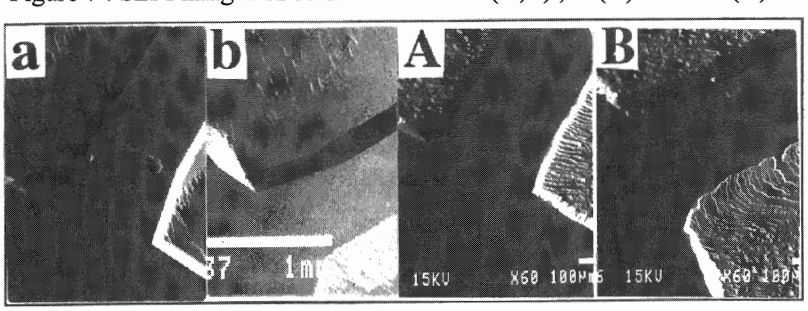


Figure 8: (441) Beams B2 (**a,A**) and B1 (**b, B**) on (441) plane etched in TMAH (**a,b**) and NaOH (**A,B**)

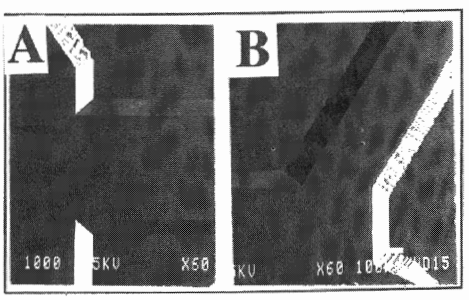


Figure 9: Beams H (**A**) and B1 (**B**) etched in a (100) plane by NaOH

corresponding concave corners look similar.
These remarks apply also for (112) and (441) plates with in particular :
i) Terracing present on beams shoulders (Figs 7A, B, D and 8 A, B) for NaOH in contrast with TMAH behavior (Figs 6 and 8a, b)
ii) Rather identical top contours at corners.
But extension of (111) facets at corner is less important for beams etched in NaOH than for those fabricated with TMAH (compare Fig. 7D with Fig. 6d for example).If we turn attention to (100) plates (Fig. 9) we observe that for the two etchants the development of planar {111} facets prevails at concave corners. But here again edges of beams are terraced (Figs. 9 A, B) in the case of NaOH solution.
The experimental study allows us to conclude that the anisotropy of the NaOH solution differs markedly from that related to TMAH 25%. Owing to the degradation of {110} and {hhl} surfaces etched in NaOH and to the simultaneous terracing of boundering planes the TMAH etchant seems at first sight more convenient than NaOH 35% solution. But NaOH can constitute an alternative etchant for the micromachining of structures in {100} planes.

3. THEORETICAL STUDY

3.1. 3D etching shapes

Tentatives are made to derive theoretical shapes for beams etched in TMAH using the simulation tool "TENSOSIM" [4,6] which presents the advantage to furnish 3D etching shapes for any cut and for any direction of alignment of beams. The simulator "TENSOSIM" is based on the tensorial model of the anisotropic dissolution [4] which predicts that for the beam structures studied here :
i) the 3D etching at concave corners built-in structures is governed by maxima maximora of the dissolution slowness. Thus {111} facets are in priority formed at corners as observed on SEM images
ii) The 2D etching shapes (cross-sections of beams) are composed of "facets" associated with maxima in the cross-sectional polar diagram of L. So etched shapes for cross-sections may be very sensitive to the alignment angle Ψ_A.
For convenience 3D etching shapes are represented as constant-level contour diagrams. Lateral underetchings U_L beneath the mask and mean slopes of edge beams are extracted from theoretical 2D cross-sections. Let us concentrate our attention to the (110) plane for which the two sides of beams satisfy the two-fold symmetry about the perpendicular [110] direction as given evidence by theoretical shapes of Figures 10 a,b and 11 related to beams H, B2 and B3. In practice table 2 reveals that $U_L \approx 20$ µm for the beam H indicating a misalignement $\Delta\Psi_A$ of few degree even if theoretical shape of Figure 10 is in accord with the SEM image of Fig. 4c. Figure 11 illustrates the effect of misalignments $\Delta\Psi_A = \pm 2.5°$ on the theoretical beam B3. Here the general shape and inclination of limiting facets are not markedly affected by $\Delta\Psi_A$ in the contrary of the top contour of corners. But for some directions of alignment both inclinations concave undercutting and width of beams may be significantly modified (see section 3.2).

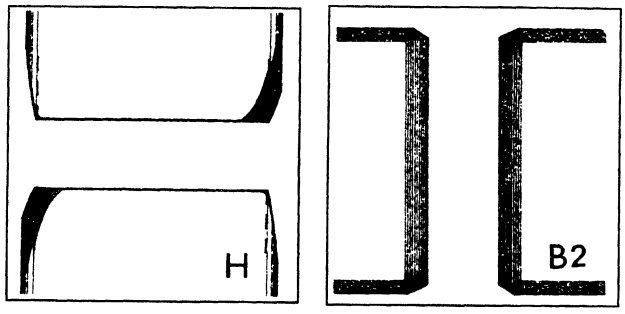


Figure 10 : Theoretical H and B2 beams on (110) plane

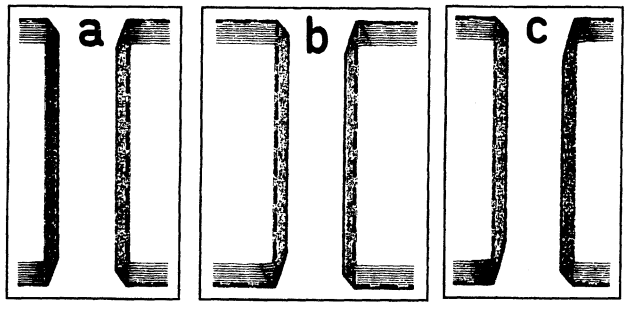


Figure 11 : Beam B3 on (110) plane , effect of misalignment. a, b, c are for $\Delta\Psi_A = +2.5°$, 0° and –2.5°

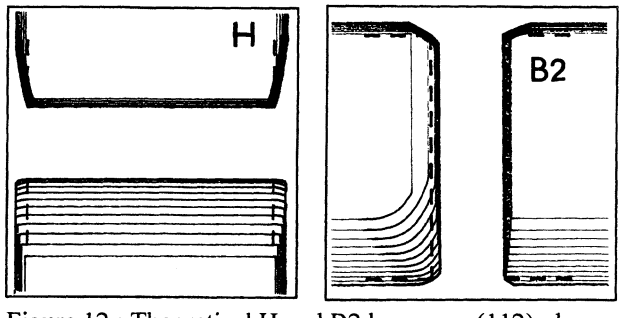


Figure 12 : Theoretical H and B2 beams on (112) plane

Simulations of theoretical shapes for beams H and B2 (Fig. 12) on (112) plane agree also with experimental shapes of Figures 6c and 6a,b. We thus infer that the database is rather appropriate and consequently we can use the simulator to study effects of misalignment $\Delta\Psi_A$ or of small deviations $\Delta\varphi_o$ and $\Delta\theta_o$ on the angles of cut (as defined from the IEEE standard of piezelectricity) and on the width of beams.

3.2. Effects of misalignment and of disorientation

First we focus on (110) beams and specially on beams with large decrement Δw in width (Figs. 13 and 14). A misalignment $\Delta\Psi_A = \pm 2.5°$ on beam B3 (Fig. 13) modifies essentially Δw. Since for (110) plane misaligments leave the cross-sectional profiles of beam symmetrical (see Fig. 13) modifications induced by misalignments $\Delta\Psi_A = \pm 2{,}5°$ are represented on the same figure (Fig. 14b) for the vertical beam V. For beam V small disorientations $\Delta\varphi_o = \pm 2°$, $\Delta\theta_o = \pm 2°$ (Figs. 14d and 14c) destroy the symmetry and change significantly both the cross-sectional shape and Δw. In

fact whatever is the starting plane disorientations and misalignments cause for specific alignments marked changes in the width w. Moreover $\Delta\varphi_o$ and/or $\Delta\theta_o$ affect more and less the depth of etch of Δd. In tables 3 and 4 are collected theoretical values of $\Delta w/\Delta d$ and Δd for some beams etched on (110) and (112) planes in the case of moderate misalignments and disorientations. Deviations of $\Delta w/\Delta d$ (in % with respect to theoretical orientation) are also indicated. Examination of these tables reveals that final beams are more sensitive to misalignment $\Delta\Psi_A$ than to small changes in the angles of cut. For misalignments of ± 2,5° departures for of $\Delta w/\Delta d$ can reach 25% for (110) planes and exceed 25% for (112) planes. Moreover more marked deviations are expected for misalignment of the horizontal beam.

	Beam B2			
	$\Delta\Psi_A$=+2.5°	$\Delta\Psi_A$= −2.5°	$\Delta\varphi_o$ = ± 2°	$\Delta\theta_o$ = ± 2°
Δ*	.133	.164	.151	.143
%	8	15.5	6.3	≈1
	Beam V			
	$\Delta\Psi_A$=+2.5°	$\Delta\Psi_A$= −2.5°	$\Delta\varphi_o$ = ± 2°	$\Delta\theta_o$ = ± 2°
Δ*	1.7	2.25	2.16	2.15
%	25.5	1	5.3	5
Δd	98	98	98	99

Table 3 : Effect of misalignment and disorientation on the reduced width of beams Δ*for the (110) plane.

CONCLUSION

Beams are micromachined on various {hhl} planes using NaOH 37% and TMAH 25% solutions. The results show that surfaces and edges bounding beams are less degraded with the TMAH etchant. Moreover etching shapes and width of beams depend on the direction of alignment. So it is shown that the simulator TENSOSIM constitutes a convenient CAO tool for evaluating changes due to misalignment or to small deviations in the angles of cut.

REFERENCES

[1] R.A. Buser, N.F. de Rooij, "Very high Q-factor Resonators in Monocrystalline Silicon", Sensors and Actuators, A21-A23 (1990) 323-327

[2] E. Stemme, G. stemme, "A Balanced Resonant Pressure Sensors", Sensors and Actuators, A21-A23 (1990) 336-341

[3] G.A. MacDonald, "A Review of Low cost Accelerometer for Vehicle Dynamics", Sensors and Actuators, A21-A23 (1990) 308-311

[4] C.R. Tellier, "Anisotropic etching of silicon crystals in KOH solutions. Part 3", J. Mater. Sci., 33 (1998) 117-131

[5] I. Zubel, I. Barycka "Silicon anisotropic etching in alkaline solutions", Sensors and Actuators, 70(1999) 81-97

[6] A.R. Charbonnieras, C.R. Tellier, "Characterization of the Anisotropic Chemical Attack of {hk0} Silicon Plates in a TMAH Solution. Determination of a Database", Sensors and Actuators, 77(1999) 81-97

	Beam B2			
	$\Delta\Psi_A$=+2.5°	$\Delta\Psi_A$= −2.5°	$\Delta\varphi_o$ = + 2°	$\Delta\theta_o$ = + 2°
Δ*	.11	.272	.127	.182
%	32	68	21	12.4
	Beam B1			
	$\Delta\Psi_A$=+2.5°	$\Delta\Psi_A$= −2.5°	$\Delta\varphi_o$ = + 2°	$\Delta\theta$ = + 2°
Δ *	1.14	.72	1.69	1.02
%	27	20	79	13.4
Δd	100	100	103	109

Table 4 : Effect of misalignment and disorientation on the reduced width of beams Δ* and Δd (in μm). The case of a (112) plane.

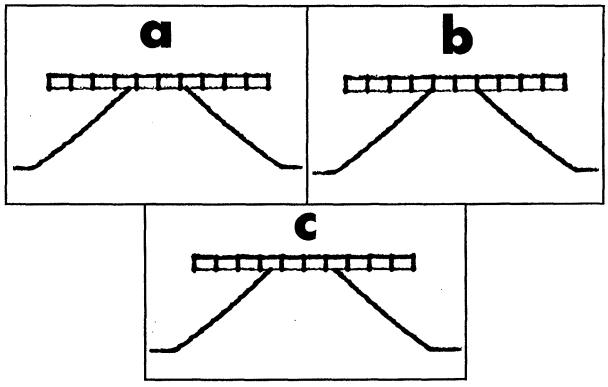


Figure 13: Cross-sections of the beam B3 on the (110) plane **a, b, c** are respectively for $\Delta\Psi_A$= 0°,$\Delta\Psi_A$=+2.5° and $\Delta\Psi_A$= −2.5°

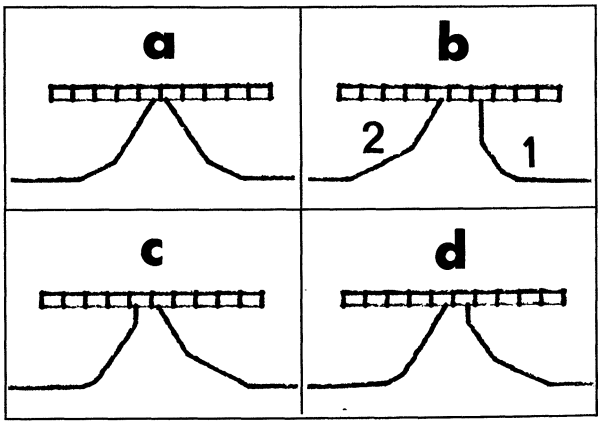


Figure 14: Cross-sections of the beam V on the (110) plane **b, c** and **d** are respectively for **b**: $\Delta\Psi_A$=+2.5° (1) and $\Delta\Psi_A$= −2.5° (2), **c** : $\Delta\varphi_o$ = −2°, **d** : $\Delta\theta_o$ = −2°. **a** is for a perfectly oriented beam

ESTIMATION OF A NON-LINEAR ELEMENT INFLUENCE ON PHASE NOISE IN CRYSTAL OSCILLATOR.

D. V. Bogomolov

JSC "Radiophysics", 10, Geroev Panfilovtsev st., O. Box 1, 123363 Moscow, RUSSIA

Fax.: 095 496 8790, tel.: 095 494 9573, e-mail: dvb@atom.ru

ABSTRACT

The numerical analysis of the phase noise dependencies upon an active device operation mode in the precision crystal oscillator is carried out. The relations between a coefficient of amplitude to phase noise conversion and a ratio of the device linear mode realm to non-linear mode realm are evaluated. It is shown, that the maintenance of operation mode which provide both said symmetric amplitude limitation of formed oscillation is the necessary condition for oscillator phase noise minimization.

1. TASK SETTING

The purposes of phase noise minimization in precision quartz oscillator demand the detailed analysis of all possible mechanisms of their appearance [1]. The device with the non-linear amplitude characteristic is an unavoidable part of any oscillator, therefore estimation of influence of it operation mode on a phase of noise level is an actual task. The digital mathematical simulation of oscillations is widely used to investigate noise in generators [2] and to evaluate influences of the circuit parameters on the frequency stability [3]. Importance of these research methods consists also in permitting to observe dependence on one parameter change within the interconnected parameters set. It is usually impossible in the natural experiments and measurements. Numerical simulation allows rather simple to set and to change the non-linear realm of active device operation mode. This paper discusses the dependencies of amplitude to phase noise conversion coefficient calculated by means of a self-oscillating system dynamical model. The ratio of time when the active device operates in linear mode to whole oscillation period is a main parameter which influence was researched. The second parameter which influence on phase noise level was investigated is the symmetry of the oscillation amplitude limitation. It can be tuned by changing a transistor initial current or by the variable loads in emitter and collector circuits. This factor influence on a whole level of the oscillator phase noise and was suggested that it affects through changing of the harmonics components, which level influences on a value of the device complex transfer function slope.

The considered model of the quartz oscillator takes into account the following parameters: nonlinear and linear behavior of the amplification device, non-linear dependencies of transistor base junction capacity upon signal level, transistor active device inertia and signal delay. The parameters of the precision crystal oscillator diagram are described in [4].

2. MODEL DESCRIPTION

The considered quartz generator can be represented by the following self-oscillating system. Such system contains: the active devise with nonlinear amplitude characteristic, in a feed-back of which a second order filter is included. The active devise has inertia. It can be represented as an integrating circuit time constant τ_α and as a time delay parameter τ_d. Thus the Laplace image of transistor current transmission coefficient $\alpha(s)$ is determined by following expression:

$$\alpha(s) = \alpha e^{-s\tau_d} /(1 + s\tau_\alpha),$$

$\tau_d \approx 0{,}2\ t_f$,
$\tau_\alpha \approx 0{,}8\ t_f$,
where t_f - time of a fly through the transistor base zone. The time constant of the integrating circuit can be expressed as $\tau_\alpha = C_b R_b$, it allows to include in a model non-linear dependence of

base-emitter transistor junction capacitance by following expression:

$$C_b(U)=\begin{cases} 0{,}5C_0U & when \quad U>0 \\ C_0/\sqrt{1-U/\Delta_\varphi} & when \quad U\le 0 \end{cases}$$

where $\Delta\varphi$ - height of a transition junction potential barrier. Such self-oscillating system is described by the following system of the differential equations:

$$\frac{d^2x}{dt^2}+\frac{\omega_0}{Q}\frac{dx}{dt}+\omega_0^2x=\omega_0Q\frac{d}{dt}F(y)$$

$$\tau_\alpha(y)\frac{dy}{dt}+y+n(t)=R_b x(t-\tau_d) \quad , (1)$$

where $x = x(t)$ - current of a quartz resonator at a sequential resonance, $y = y(t)$ - managing voltage on the emitter-base junction, Q and ω_0- loaded quality and resonance frequency of a quartz resonator correspondingly, $F(y)$- amplitude characteristic of the active device.

The transistor connection scheme with common base allows operating in circuit with zero power supply collector voltage. So such scheme has low value of a signal amplitude limitation in collector at a positive half-wave of an oscillation. The transistor collector junction becomes opened if the voltage exceeds value $\Delta\varphi$. The same voltage affects on emitter – base junction at a negative half-wave of an oscillation. So the scheme operate in the realm where non-linear behavior happens a main pate of oscillation period. The similarity of emitter and collector junction voltage –current characteristics allows to write a following expression for derivative of the function $F(y)$:

$$\frac{d}{dt}F(y)=\begin{cases} \frac{s_0}{\Delta_\varphi}\exp(y/\Delta_\varphi) & if \quad y<U_{SD} \\ s_0 & if \quad U_{SD}<y<U_{SUP} \\ \frac{s_0}{\Delta_\varphi}\exp((U_{max}-y)/\Delta_\varphi) & if \quad U_{SUP}<y \end{cases}$$

,(2)

where s_0 - a slope of the transistor transfer function in a linear amplification mode, which takes a voltage range on the transistor basis from U_{SD} up to U_{SUP}, U_{max} - maximum of signal amplitude value in the collector circuit. Model does not include values of passive components of the generator scheme, but by changing included in the formula (2) parameters it is possible to simulate the most various modes of the oscillator operation.

3. SOLUTION AND CALIBRATION METHODS

The set of equations (1) was solved with the help of numerical integration procedure. The fourth order Rung Kutte method was used. The choice of a method is caused by it high stability and small accumulation of a calculation errors.

To achieve correspondence between the calculation signal and the signal at the researched oscillator output some precaution had been made. The special calibration procedure of calculations accuracy was applied, it allows to calculate oscillations when dissipative term has high Q-factor and therefore has the very small coefficient. The oscillation time was calculated when the dissipative term coefficient was equaled to zero. If the oscillation (with selected integration time discrete) decreases faster than at Q oscillation periods, a neutralized term (like a positive feed-back) was added to the dissipative coefficient. It permits to compensate fading caused by excessive stability of a numerical integration algorithm. Such trick has allowed to achieve the reasonable compromise between computing expenditures and calculations accuracy. The usually demanded accuracy was achieved with calculation of 40 signal samples values on each oscillation period.

The processing of an obtained sequence of the signal samples and calculation of phase deviation dependence in time was produced with the help of algorithm described in [5]. The added error at this stage of the calculations did not exceed 10^{-4} rd.

4. CALCULATION RESULTS

The obtained functions for the amplitude to phase noise conversion coefficient k_{aph} are presented on a figure 1. The graphics show dependencies of the conversion coefficient upon ratio of realm in which transistor works in a linear mode $U_{lin}=U_{SUP} - U_{SD}$ to the whole signal amplitude realm - U_{max}. The solid curve belongs to a variant, when the function F'(y) has a central symmetry relatively the center of an oscillation amplitudes range. The dashed curve belongs to dependence at which, not symmetry of boundaries of a linear mode from U_{SD} to U_{SUP} equals 5 % from value of U_{max} and the doted curve corresponds to 10 % to difference of the linear site boundaries.

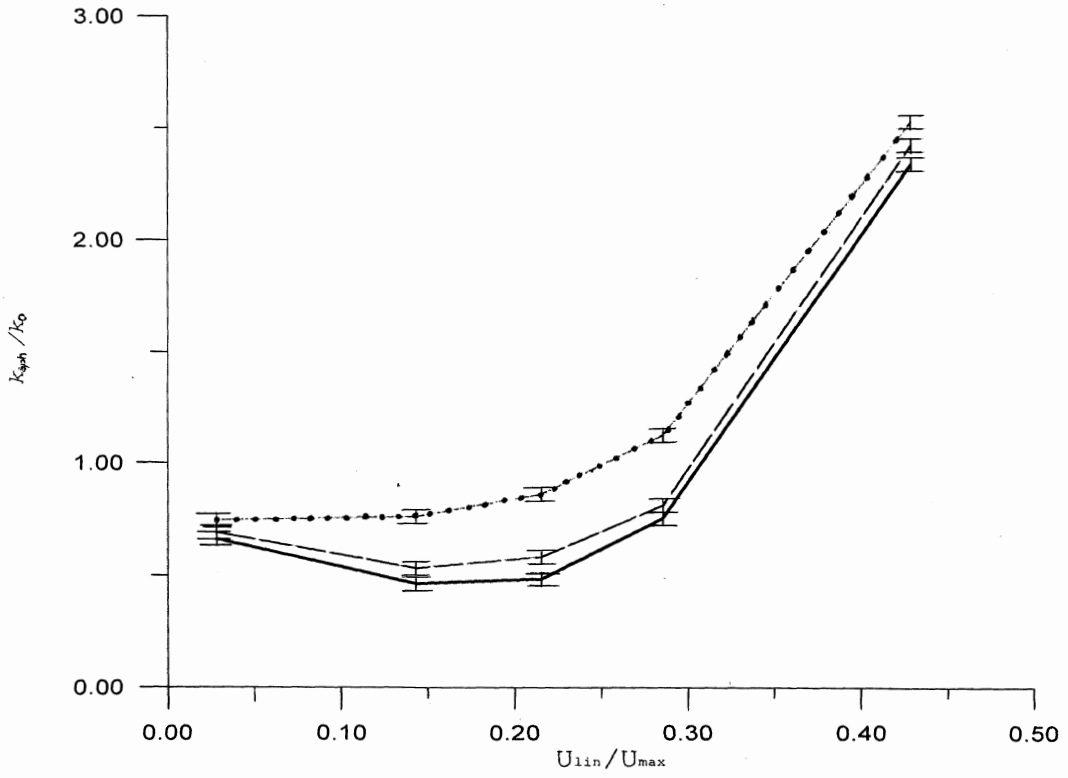


Fig. 1 The amplitude to phase noise conversion coefficient upon ratio of linear realm to whole operation realm.

The value of k_0 is equaled to 10^{-2} rd/V. The digital simulation experiment have showed that phase noise level was in 4 times lower than phase noise corresponding to k_0 in the case when capacitance of the base-emitter transistor junction didn't depend upon voltage. It confirms the suggestion that non-linearity of reactive components is the main source of amplitude to noise conversion. The influence of the amplitude noise on oscillation phase is decreased in non-linear operation mode because device is almost closed and disconnected. So the noise transfer function increases in linear operation mode. Some increase in k_{aph} dependence at a very small linear realm is possible to explain by added phase noise due to increase of amplification coefficient in a feed-back loop for obtaining generation in numerical model.

5. CONCLUSIONS

The obtained results show that the operation mode of the active device at which double side symmetric amplitude limitation is provided allows to minimize coefficient of the amplitude to phase noise conversion. The numerical experiments gives one more explanation of a low phase noise behavior of precision crystal oscillator described in [5], which special feature is prevailing of the non-linear mode in the oscillator active device operation.

6. REFERENCES

[1] R.J. Besson, M. Mourey, S. Gallion,... "10 MHz hyperstable quartz oscillator performances" in Proc. of the 1999 EFTF-IEEE IFCS Joint Meeting, pp. 326-330.

[2] A.V Kalianov "Oscillations in non-linear systems with delay and inertia", Journal of Communications Technology and Electronics, Vol. 40, N 2, pp. 123-126.

[3] 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.

[4] A.E. Silaev, D.V. Bogomolov, "Low noise ovenized quartz oscillator", in Proc. of the 1998 IEEE IFCS, pp. 3249-352.

[5] D.V. Bogomolov, D.C Ochkov, "Phase noise evaluation digital algorithm for precision oscillations", in Proc. of the 1999 EFTF-IEEE IFCS Joint Meeting, pp. 1117 - 1120.

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

THERMAL ADJUSTMENT FOR DOUBLE OVEN CRYSTAL OSCILLATOR WITH LARGE RANGE OF TURNOVER POINT

Laurent Couteleau – Marc Henrotte
Tekelec Temex Components
2, rue Robert Keller
10150 Pont Sainte Marie – France
e-mail : laurent.couteleau@temex.fr

ABSTRACT

CQE works for several years on double oven technology. This work completes the paper presented last year with some practical aspect[1]. This paper describes the thermal adjustment needed in function of the turnover point of the crystal. The span of the turnover point was 15°C around the nominal value. We will discuss about the non-linearity of the function for this adjustment. We will present the inner oven function which is mostly linear, and the second oven one which diverges in function of the turnover point. We also give a possible explanation of this non-linearity. This study gives us a solution to improve the thermal adjustment of our oven crystal oscillator. This work improves our ability for mass production of double oven. We present statistical results acquired during this experiment. We use this principle for other type of oven crystal oscillator in order to reduce the setting phase.

1. INTRODUCTION

Production of DOCXO increases for several years. These types of OCXO require high performance crystal resonators. One of the major characteristic for these crystals is the turnover point of the thermal response. The crystal manufacturers process shows a gaussian repartition for the value of the turnover point. Then, to increase the yield of the resonator, we use a large range of turnover point. This large range of turnover point introduces a large needs of operation during the setting phase. We will first present a functional scheme of DOCXO regulation and explain the traditional setting phase. Increasing the range of the turnover point induces an increasing number of the operation thermal cycling. We have tried to use a statistical method to help us to do the setting phase. We also give an explanation of the results obtained during this experiment.

2. GENERAL DESCRIPTION OF THERMAL REGULATION USING IN DOCXO

As everybody knows resonance frequency of quartz resonator is highly dependent on temperature. Temperature variations involve, in general, variations of oscillators' frequency. It seems to be interesting to choose a resonator which has a low thermal sensitivity in order to make up for drawback. But this process is not always sufficient so a correction device of thermal effects must be associated to the oscillator.

Two methods can be distinguished depending on whether the correction action acts on cause or on effect :

- Action on effect : it consists of putting right the temperature effect on frequency, it is the thermal balancing (TCXO : Temperature Compensated Crystal Oscillator)
- Action on cause : it allows to protect the crystal resonator from temperature variations, it is a thermostatic control (OCXO : Oven Controlled Crystal Oscillator)

In view of metrological oscillators' requirements, nothing short of OCXO's allow to obtain a good frequency stability. OCXO's are equipped with a resonator which has a discharged overvoltage the highest as possible so as to have insignificant repercussions from ageing or random fluctuations of electronics components which make up the oscillator on the resonator fundamental frequency. A part or the totality of the crystal oscillator is set in an electronic oven exact and very steady. In high stability OCXO, frequency drift due to temperature variation is minimized by using a double oven which immunizes the crystal oscillator from environmental temperature variation. This method enables the crystal resonator to operate at an almost constant temperature, leading to minimum frequency deviation (few part in 10^{-10}, 10^{-11}) regardless of the ambient temperature range.

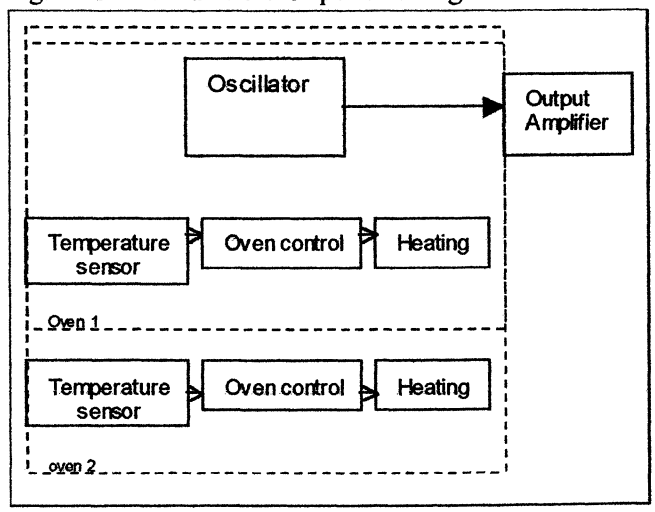


Fig 1 Functional scheme of a DOCXO

To reduce temperature fluctuations' effects of thermostat used, the temperature of the Oven 1 is adjusted on the temperature of resonator's turnover point. Resonator's functional temperature corresponds to a point of horizontal tangent of the frequency-temperature relation. The temperature of the Oven 1 is chosen a few degree above the wanted limit temperature of functioning of the OCXO in order to taken into account inner overheating.
The frequency's thermic stability of a such oscillator is obtained :

- by using an SC cut
- because of precision of setting of oven at a temperature about the resonator's turnover point (thermical offset)
- because of the oven's thermic stability.

These demands set up control process with low tolerances so expensive. This cost is to be divided between resonator's fabrication and electronic control.

The external oven is adjusted in such a way that internal control is only a low regulation. It involves that the oven 1 sees only temperature variation of a few degrees. Near the crystal, temperature variation is amply under the degree.

3. SETTING PHASE USING A NARROW RANGE FOR TURNOVER POINT

The setting phase of such a regulation consists first to adjust correctly internal control. Two method are used :

- temperature of oven 1 is adjusted on turnover point's temperature of the crystal. It is the ideal case where electronics parameters are constant on the temperature range.
- Temperature of oven 1 is adjusted a few degrees above or under turnover point's temperature of the crystal. In this case, oscillator's electronics parameters change over temperature range. Frequency variations due to electronic is compensated thanks to a variation of resonance frequency of the crystal in such a way that the oscillator gives a frequency as stable as possible in the temperature range.

The setting of the oven 2 is done so that consumption's variation of the internal oven is the lowest on OCXO's functional range.
Product's ideas man has a couple (Toven1, Toven2) which will serve to him of beginning point in optimisation process of the final choice of (Toven1, Toven2). Then by averaging, the best adjustment is established in order to have the best adjustment phase for the OCXO's production.

$$TOVEN1 = < Toven1 >.$$
$$TOVEN2 = < Toven2 >$$

When the range of crystal's turnover point becomes too much wide in comparison with the precision wanted, this adjustment is slightly changed while specifying the different offsets in comparison with crystal's turnover point.

$$TOVEN1 = T_{turnover} + Offset1$$
$$TOVEN2 = T_{turnover} + Offset2$$

Considering performances demanded for majority of OCXO's, it is this type of adjustment that is used. In fact this type of adjustment functions as long as the frequency precision is around a few part per billion. Moreover, to have such a precision, crystals are designed with a turnover point between a range of about five degrees. When frequency precision becomes lower and specification about crystal's turnover point becomes wider, we can notice that adjustment which contained a single offset is not enough efficient.

4. PROBLEMS INDUCED BY USING A LARGE TURNOVER POINT

For double oven crystal oscillator the usual frequency stability versus temperature range is contained between $2.\ 10^{-9}$ and 2.10^{-10} ; sometimes it could be 2.10^{-11}. Moreover the crystal turnover point is distributed as gaussian law around the specified value. The standard deviation is about 10 degrees. The specification asked for crystal used in double oven are within a few degree. All crystals with a turnover point outside this range are rejected. The yield for crystal production in these conditions is only around 50%. This is due to the narrow range of turnover point asked for the double oven oscillator. This yield has an important aspect for the price of the crystal. Also for the capacity of production. As the volume increases, and the price decreases, the crystal resonator cost is an important issue. One possibility would be the use of a wide range of turnover point. The benefit realised at one point cost of crystal resonator won't be overcharged by the cost of an increasing setting phase. That's exactly the purpose of this document explaining the way we used to work for double oven with crystal presenting large range of turnover point. The range used was more than 15 degrees.
The goals of our work was to achieve a frequency stability of 2.10-10 over a temperature range of 0°C to 70°C using crystal resonator with turnover point contained in the range 75°C to 90°C. This for hundred pieces, not anymore prototypes but serial production.
Rapidly we notice a real divergence of the value obtained and the usual setting

$$TOVEN1 = T_{turnover} + Offset1$$
$$TOVEN2 = T_{turnover} + Offset2$$

The way used is a successive approximation used on the two ovens without initial value. The interaction of the two ovens weren't taken into account.

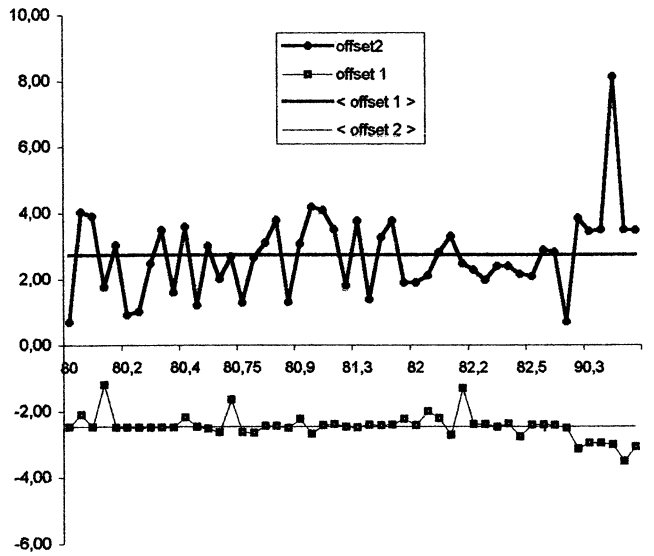


Fig 2 – Zero order approximation vs turnover point

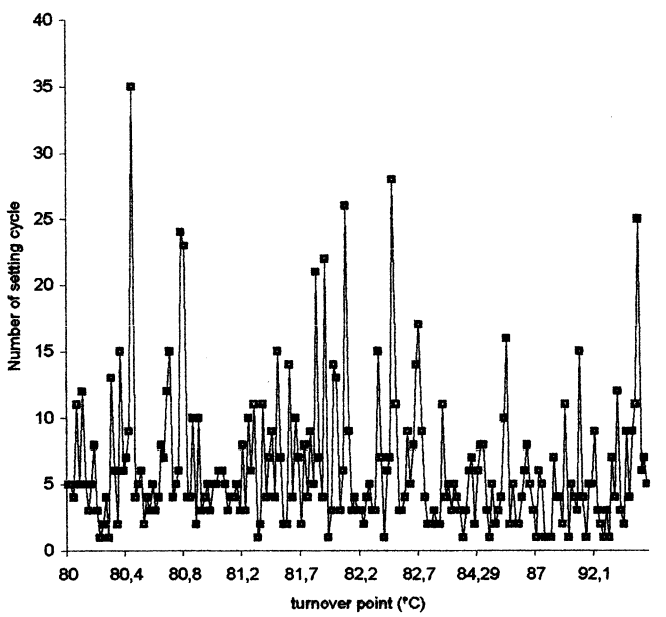


Fig 3 – Number of thermal cycle vs turnover point

After plotting the first data, we give the main idea to improve the setting phase. This was the use of a mathematical interpolation of data. This interpolation would be polynomial. In fact the traditional setting with narrow turnover was a constant offset, that is not more than a polynomial of degree zero. We have used at the beginning first degree polynomials.

5. SETTING PHASE USING A LINEAR METHOD

In fact we use a linear function to calculate the two value of the oven.

$$TOVEN1 = A1.\ T_{turnover} + B1$$
$$TOVEN2 = A2.\ T_{turnover} + B2$$

The coefficient of this function were calculated on the first value founded by the traditional methods.

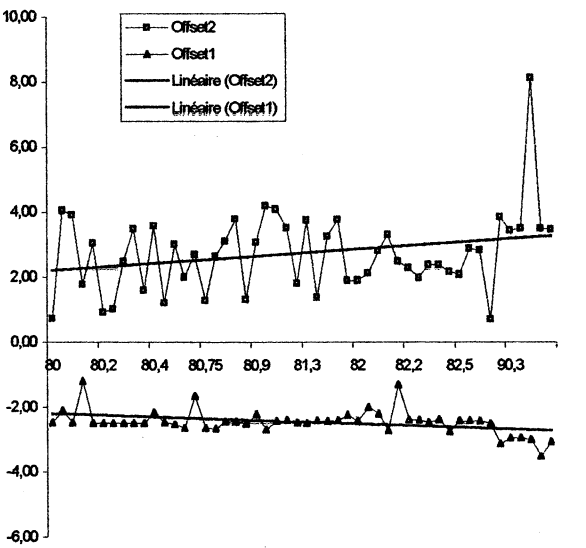


Fig 4 – One order approximation vs turnover point

This method was applied for about 50 pieces. By using the first value we obtain better results. In reality we gain of number of measurement done. Nevertheless we had difficulties with the linear function. We complete the plotting with the results of the new 50 pieces. At this point the deviation obtained by the linear function was too far away from the plot. Then we introduce more degree in the polynomial function. After calculation we obtain the best results with a third order polynomial. High order introduces artefact at the end of the series, low value and high value. Those extreme values are the worst case for the setting.

6. FURTHER IMPROVEMENTS WITH A NON-LINEAR LAW

We use a third order function to calculate the initial value for the temperature of the two oven. By this way we obtain more than 50% of the device within the specification with only one more adjustment.

$$TOVEN1 = A1.\ T_t^3 + B1.T_t^2 + C1\ T_t + D1$$
$$TOVEN1 = A2.\ T_t^3 + B2.T_t^2 + C2\ T_t + D2$$

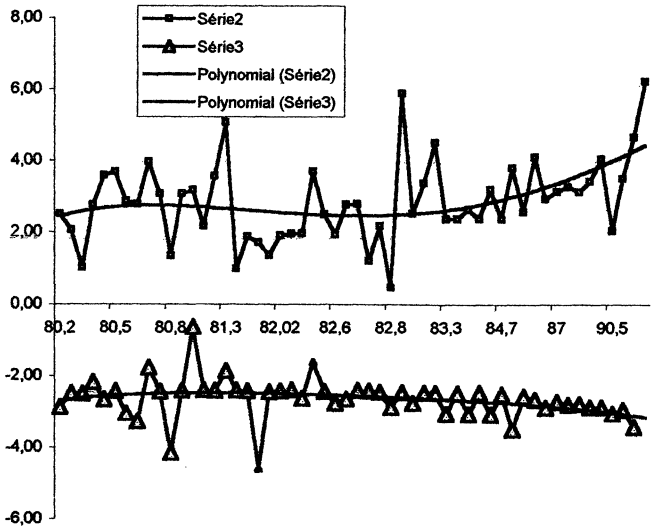


Fig 5 – Third order approximation vs turnover point

7. RESULTS

The divergence shows by the statistical analysis could probably be explained by the thermal modification inside the OCXO. We try to find out which parameter could change with a variation of the value of the temperature setting. Our conclusion is that the thermal resistivity of one material in the thermal path changes from the heating element to the crystal resonator. We check the thermal behaviour of all the materials used inside the OCXO [2]. None of them shows significant variation change of thermal resistivity.
The thermal is composed by ceramic, glass, but also a small portion of air. The inner oven is completely isolated by air. When we check the thermal behaviour of air with respect to the element give by [2]. we find a significant change of the thermal resistivity of the air. As the thermal resistivity of the air change the isotherm surrounding the inner oven will change. The variation shows a decrease of resistivity with an higher temperature. If we compare these results to an electrical case the isopotentiel extend more over. In our case the thermal resistivity decrease will extend the isotherm at the same temperature. In the seem time the thermal loss increase. As these losses increase the temperature setting needs to increase to maintain the same temperature at the crystal resonator. The extension of the isotherm at higher temperature will affect the ouside temperature setting. The second oven need to be colder than at low temperature.

8. CONCLUSION

This work highlight the support of statistical methods in OCXO production. The visualization of results give us a new approach for understanding the fundamentals of this types of OCXO. The physical phenomena could be explain with a rather simple approach. The goals which was to reduce the overall production cost is reach, we are able to use a wide range of turnover point with minimum setting phase. This methodology has been extend to other OCXO product.

9. REFERENCES

[1] F. Vignaud and M. Lancien, "Thermal regulation of Double Oven Controlled Crystal Oscillator, in Proceeding of the Joint Meeting , 1999, pp...

[2] W.M. Rohsenow, J.P. Hartnett, Y.I. Cho, "Handbook of heat transfer", McGraw-Hill, 1998.

FREQUENCY SPECTRUM OF LOKALIZED THICKNESS-SHEAR VIBRATIONS OF THE PIEZOELECTRIC RESONATOR

*S.S. Nedorezov, E. A. Ganenko, T. V. Emeliyanova, **Yu. S. Shmaliy, ***Kr. Weiss

*"Sichron" Center, 4 Skrypnyka Street, Kharkiv, 310057, Ukraine, stne@sichr.kharkov.ua
**Electronics Dept., Guanajuato Univ., S/N 36730, Salamanca, Gto, Mexico, shmaliy@salamanca.ugto.mx
***Piezoelectricity Dept., ITR, 11 Ratuszowa, 03-450 Warszawa, Poland, kweiss@org.pl

ABSTRACT

The paper addresses the new method for evaluation of the vibration frequencies of the resonator employing thickness-shear vibrations with the convex piezoplate. The method is based on the use of the approximating expression for calculation of the frequency spectrum, evaluation of its parameters proceeding from the obtained experimental data, and following account of the electrode boundary influence into the resonator frequency spectrum.

1. INTRODUCTION

Modern oven controlled crystal oscillators (OCXO) of precise frequency are now at that stage of development when further progress in the frequency instability by traditional methods is achieved by making rather noticeable efforts [1] because their long-term frequency noise is limited for the first hand by the destabilizing factors (temperature, aging, mechanical vibrations, accelerations, etc.) influencing crystal resonator. In principle, recently proposed the modulational method [2] allows substantial progress in this direction being based on use of a crystal resonator anharmonic modes as sensors of the environment [3]. However, the modulation results in the non-linear spectral products whose presence in the bandwidth of the sensor anharmonic mode may yield undesired splashes in the oscillator phase spectra. This phenomenon may be prevented if the natural frequencies of the anharmonic vibrations of a crystal resonator are accurately determined at the early stage.

In the report we consider in detail the methodology of the natural frequencies calculation of the localized thickness-shear vibrations in a frame of the advanced method in application to the crystal resonator of AT-cut with convex piezoelectric element. We obtain correspondent results and comparisons with experimental data [9]. The efficiency of the advanced method is explained by possibility to get the limited accuracy obtained by the effective equations. As the major practical sense of the studies we note that any shape of the crystal plate and electrodes may be accurately considered in a frame of this approach. We present the final expressions for the most interested shapes of a resonator accounting correspondent boundary conditions and piezoelectric anisotropy.

2. QUAZI-EQUIVIDISTANT FREQUENCY SPECTRUM

The frequency spectrum of natural vibrations of a convex piezoelectric element is defined in frame of the piezoelectric theory [10] as

$$\omega_{kln}^2 = \sum_{\alpha,\beta=0}^{\infty} \Omega_{\alpha\beta n} k^{\alpha} l^{\beta} + W_{kl}(n) \qquad (1)$$

where

$$\Omega_{\alpha\beta n} = \sum_{p=0}^{\infty} \Omega_{\alpha\beta n}^{(p)} R^{-\frac{p+\alpha+\beta}{2}}, \qquad (2)$$

$k, \ell = 0, 1, 2, \ldots, n = 1, 2, 3, \ldots, R$ is radius of curvature of a piezoelectric element (Fig.1).

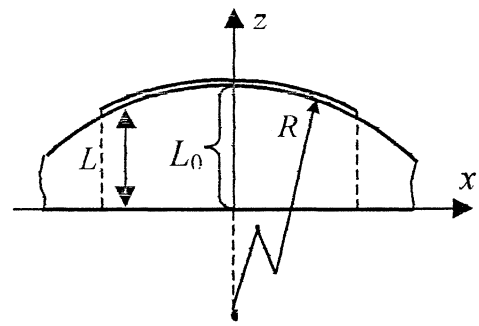


Figure 1. The intersection of the convex piezoelectric

Here the coefficients $\Omega_{\alpha\beta n}^{(p)}$ are expressed in terms of piezoelectric constants [10], and $W_{kl}(n)$ are the matrix elements of operator **W** which describes the influence produced by the electrode boundary on the frequency spectrum. The following holds true

$$\omega_{kln}^2 - \omega_{00n}^2 << \Delta, \qquad (3)$$

where Δ is a potential jump at the boundary of the electrode, and zero boundary condition for the vibrations may be taken [6] as an operator $\hat{\mathbf{W}}$. As a result, for $W_{kl}(n)$ we may get the explicit expressions presented as a function of ρ parameter which is defined by the vibrations damping length (an electrode radius ratio) for (3). An accuracy of the coefficient $\Omega_{\alpha\beta n}$ definition (2) in (1) is significantly limited by the approximate model of vibrations of a piezoplate with a electrodes. At the same time, the explicit expression (1) for the frequency ω_{kln} allows consideration with

high accuracy the coefficients $\Omega_{\alpha\beta n}$ as parameters derived from the experimentally obtained several frequencies. Below we consider the efficiency of such approach with respect to AT-cut quartz crystal resonator being based on the experimental data [9]. The "C" mode spectrum has been considered in assumption that the vibration modes k, $l = 0, 1, \ldots$ and $n = 1, 2, 3, \ldots$ correspond to that k, l, $n = 1, 2, 3, \ldots$ of [9].

With $\rho >> 1$, the influence produced by the electrode boundary can be neglected and in the quadratic approximation from (1) we get

$$\omega_{klm}^2 \cong \Omega_{00n} + k\Omega_{10n} + l\Omega_{01n} + k^2\Omega_{20n} + kl\Omega_{11n} + l^2\Omega_{02n} \quad (4)$$

The first three terms in (4) are obtained by the linear approximation that brings to the equidistant vibration spectrum, and the rest nonequidistance ones are the non-linear products. Once simulated, we have finally seen, the approach is valid as we got small disagreement of the predicted frequencies and experimental data of 10^{-3}–10^{-2}. It follows also that for some frequencies the linear and quadratic approximations give just the same error. To improve the convergence of series (1) and to get better approximation the partial summation has been done. As a result, we obtained

$$\omega_{k01}^2 \cong \frac{\omega_{001}^2}{1 - a_{10}k}, \quad (5)$$

where constant a_{10} is defined through the experimental frequency ω_{101}. Formula (5) gives more accurate estimates with the error of about 10^{-4} –10^{-3}. To get the same accuracy we get the following approximation

$$\omega_{0l1}^2 \cong \frac{\omega_{001}^2}{1 - \dfrac{a_{01}l}{1 + a_{02}l}}, \quad (6)$$

where the constants a_{01} and a_{02} are also given in Table 2 being defined by the method of least squares. For the rest frequencies we get

$$\omega_{kl1}^2 \cong \omega_{k01}^2 - \omega_{001}^2 + \omega_{0l1}^2 + kl\Omega_{111}, \quad (7)$$

where ω_{k01} and ω_{0l1} are derived from formulas (5) and (6), and constant Ω_{111} is obtained by data. In comparison with quadratic approximation (4) only the five parameters had been required in the case of (7) though we got more high accuracy.

3. INFLUENCE PRODUCED BY THE ELECTRODE BOUNDARY

In the frequency domain (3) closed to the fundamental frequency of ω_n the influence produced by the electrode boundary is analyzed being based on the equation

$$-d^{\alpha\beta}\frac{\partial^2 A(x_1,x_2)}{\partial x_\alpha \partial x_\beta} + \left[\omega_n^2 + a - b_\alpha x_\alpha + c(x_1^2 + x_2^2)\right]A(x_1,x_2), \quad (8)$$
$$= \omega^2 A(x_1,x_2)$$

where ω is vibration frequency defined by (8), with zero boundary conditions for the function $A(x_1,x_2)$ at the boundary of Γ –electrode

$$A(x_1,x_2)\big|_\Gamma = 0. \quad (9)$$

In equation (8), tensor $d^{\alpha\beta}$ and the constants a, b_α, and c account anisotropy and thickness $L(x_1,x_2)$ of the convex piezoelement and depend on the number n of the mechanical harmonic. In the axes x and y of a tensor $d^{\alpha\beta}$ the equation (8) takes the form

$$-d_1\frac{\partial^2 A(x,y)}{\partial^2 x} - d_2\frac{\partial^2 A(x,y)}{\partial^2 y} + \left[\tilde{\omega}_n^2 + c(x-x_0)^2 + c(y-y_0)^2\right], \quad (10)$$
$$\times A(x,y) = \omega^2 A(x,y)$$

where d_1 and d_2 are piezoelement axes, with the boundary condition $A(\vec{x})\big|_\Gamma = 0$, where $\tilde{\omega}_n^2 = \omega_n^2 + a - \dfrac{b_x^2 + b_y^2}{4c}$, $x_0 = \dfrac{b_x}{2c}$, and $y_0 = \dfrac{b_y}{2c}$. It follows from (10) that the center of natural frequencies is located at the point $O^1(x_0,y_0)$ and is shifted for the electrode center $O(0,0)$ that is pointed by vector $\vec{b}(x_0,y_0)$ (Fig .3). To estimate the boundary influence let us set an operator

$$\hat{B} = d_1\frac{\partial^2}{\partial x^2} + d_2\frac{\partial^2}{\partial y^2} \quad (11)$$

and consider the differential expression

$$U\hat{B}V - V\hat{B}U = \nabla\vec{P}(U,V), \quad (12)$$

where

$$\vec{P}(U,V) = d_1\left(U\frac{\partial V}{\partial x} - V\frac{\partial U}{\partial x}\right)\vec{i} + d_2\left(U\frac{\partial V}{\partial y} - V\frac{\partial U}{\partial y}\right)\vec{j}, \quad (13)$$

$V \equiv A(\vec{x})$ is a solution of the boundary-value problem (10), $U \equiv G(\vec{x},\vec{x}')$ is a Green function satisfying the equation

$$\left[\hat{B} - \omega^2 - \tilde{\omega}_n^2 - c(\vec{x} - \vec{x}_0)^2\right]G(\vec{x},\vec{x}') = -4\pi\delta(\vec{x} - \vec{x}'), \quad (14)$$

where $\delta(\vec{x} - \vec{x}')$ is delta function. We can obtain from (10) and (14) for $A(\vec{x})$ the following integral relationship

$$A(\vec{x}') = \frac{1}{4\pi}\iint_S\left[G(\vec{x},\vec{x}')\hat{B}A(\vec{x}) - A(\vec{x})\hat{B}G(\vec{x},\vec{x}')\right]d^2\vec{x}, \quad (15)$$

where S is a projection of an electrode onto the plane xOy. Based on the Green formula

$$\iint_S(U\hat{B}V - V\hat{B}U)d^2\vec{x} = \oint_\Gamma \vec{N}\vec{P}dl \quad (16)$$

and the boundary condition (9) we reduce the expression (15) into the form of

$$A(\vec{x}') = \frac{1}{4\pi}\oint_\Gamma G(\vec{x},\vec{x}')\left(d_1\frac{\partial A}{\partial x}\vec{i} + d_2\frac{\partial A}{\partial y}\vec{j}\right)\vec{N}dl, \quad (17)$$

where Γ is a contour which outlines the electrode bound $\vec{N}$ that is a normal to the contour Γ. The Green function $G(\vec{x},\vec{x}')$ is defined from (16) and (17) by the expansion

$$G(\vec{x},\vec{x}') = 4\pi\sum_{k,l}\frac{\Psi_{kl}(\vec{x}-\vec{x}_0)\Psi_{kl}(\vec{x}'-\vec{x}_0)}{\omega^2-\overline{\omega}^2_{kln}}, \quad (18)$$

where $\Psi_{kl}(\vec{x}) = \Psi_l(x)\Psi_k(y)$ and $\Psi_l(x) = \frac{1}{\sqrt[4]{4\pi\lambda^2}\sqrt{l!}} e^{-\frac{1}{2}\left(\frac{x}{\lambda}\right)^2} H_l\left(\frac{x}{\lambda}\right)$ are fundamental functions of the harmonic oscillator, $\lambda = \lambda(n)$, $\lambda^{-1} = \sqrt[4]{\frac{c}{d}}$, $d = d_1, d_2$, $H_l(x)$ is a Hermite polynomial. Here $\overline{\omega}^2_{kln}$ is a frequency spectrum of the vibrations localized between the electrode if we do not consider the influence of the boundary

$$\overline{\omega}^2_{kln} = \widetilde{\omega}^2_n + \Omega_{10n}\left(k+\frac{1}{2}\right) + \Omega_{01n}\left(l+\frac{1}{2}\right), \quad (19)$$

where k, l = 0, 1, 2, ... The solution of (10) may be presented in the form of

$$A(\vec{x}) = \sum_{k,l} A_{kl}\Psi_{kl}(\vec{x}). \quad (20)$$

Substituting (18) and (20) into (17) we obtain the equations system for $A_{k,l}$:

$$A_{k,l} = \sum_{k',l'}\frac{W_{kl,k'l'}}{\omega^2-\omega^2_{k'l'n}}A_{k'l'}, \quad (21)$$

where

$$W_{kl,k'l'} = \oint_\Gamma \Psi_{kl}(\vec{x}-\vec{x}_0)\left(d_1\frac{\partial}{\partial x}\vec{i} + d_2\frac{\partial}{\partial y}\vec{j}\right)\Psi_{k'l'}(\vec{x}-\vec{x}_0)\vec{N}dl. \quad (22)$$

We solve (21) with (22) in a frame of the disturbance theory that and get for (19) in a general case

$$\omega^2_{kln} = \overline{\omega}^2_{kln} + W_{kl}(n), \quad (23)$$

where $W_{kl}(n) \equiv W_{kl,kl}$ describes in the general case the influence of electrode boundary in the vicinity of the frequency ω_n. While studied we accounted that electrodes used for excitation of both anharmonic and fundamental modes may have, in principle, different shapes (Fig.2).

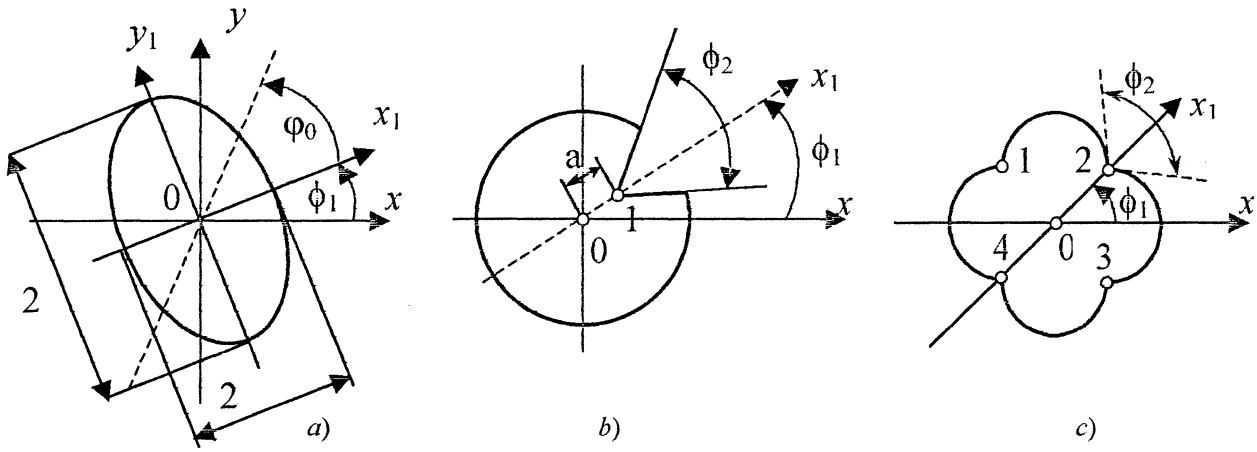


Figure 2. Shapes of the electrodes limited by the "smooth" contours or contours with the "break"

3.1. Ellipsoid shape of electrodes

Let us take that the symmetry axes x_1 and y_1 of an electrode of an ellipsoid shape with dimensions $2a$ and $2b$ are turned at an angle ϕ_1 relatively the axes x, y of a tensor $d^{\alpha\beta}$ (Fig.2*a*). In this case, the additional term $W_{kl}(n)$ in (23) is defined by the expression

$$W_{kl}(n) = -\frac{\sqrt{cd_1}\,2^{k+l+1}}{\pi k!\,l!\sqrt{\varepsilon}}\int_0^\pi \exp\left[-\rho^2_{x,\xi}(a;b)-\varepsilon\rho^2_{y,\eta}(a;b)\right]\left\{-\varepsilon\rho_x(b;a)h_l(\rho_{x,\xi}(a;b))\right.$$
$$\times\left[h_{l-1}(\rho_{x,\xi}(a;b))+\frac{l}{2}h_{l-1}(\rho_{x,\xi}(a;b))\right]h_k^2\left(\sqrt{\varepsilon}\rho_{y,\eta}(a;b)\right)-\frac{1}{\sqrt{\varepsilon}}\rho_y(b;a)h_l^2(\rho_{x,\xi}(a;b))$$
$$\left.\times h_k\left(\sqrt{\varepsilon}\rho_{y,\eta}(a;b)\right)\left[h_{k+1}\left(\sqrt{\varepsilon}\rho_{y,\eta}(a;b)\right)+\frac{k}{2}h_{k-1}\left(\sqrt{\varepsilon}\rho_{y,\eta}(a;b)\right)\right]\right\}d\varphi \quad (24)$$

where $\rho_x(a;b) = \sqrt[4]{\frac{c}{d_1}}\left(a\cos\varphi\sin\phi_1 - b\sin\varphi\cos\phi_1\right)$,

$\rho_y(a;b) = \sqrt[4]{\frac{c}{d_1}}\left(a\cos\varphi\sin\phi_1 + b\sin\varphi\cos\phi_1\right)$,

$\rho_{x,\xi} = \rho_x(a;b) - \xi$, $\rho_{y,\eta} = \rho_y(a;b) - \eta$, $\xi = x_0\sqrt[4]{\frac{c}{d_1}}$,

$\eta = y_0\sqrt[4]{\frac{c}{d_1}}$, $\varepsilon = \frac{\Omega_{01n}}{\Omega_{10n}} < 1$, and $h_l(x) \equiv 2^{-l}H_l(x)$, ρ, ξ, and η are non-dimensional parameters. We see that amendment (24) depends on the disorientation angle ϕ_1 of the symmetry axes of an electrode and tensor $d^{\alpha\beta}$ and the coordinates center $O^1(x_0, y_0)$. Its value is stipulated by those vibrations, which are localized along a direction $\varphi_0 + \beta(\varphi_0;\xi,\eta)$ in the area of $\Delta\varphi \approx 1/\rho$ wide. Assume, the shift between $O^1(x_0, y_0)$ and $O(x_0, y_0)$ does not exceed the reference length $\lambda = \sqrt[4]{d_1/c}$ and the smallest size of an electrode exceeds λ significantly. If

the transverse dimensions $2a$ and $2b$ are such that $\varepsilon < \frac{a^2}{b^2} \le 1$ then we get for expression (24)

$$W_{kl}(n) \cong \frac{\sqrt{cd_1}\,2^{k+l+1}}{\pi k!l!} e^{-\varepsilon\rho^2} \left\{ \frac{1}{\varepsilon} h_k\left(\sqrt{\varepsilon}\rho\right)\left[h_{k+1}\left(\sqrt{\varepsilon}\rho\right) + \frac{l}{2} h_{k-1}\left(\sqrt{\varepsilon}\rho\right)\right] \right.$$
$$\left. \times \Phi_l^{(0)}(1-U) + U\frac{1}{\sqrt{\varepsilon}\rho} h_k^2\left(\sqrt{\varepsilon}\rho\right)\Phi_l^{(1)}(1-U) \right\} \tag{25}$$

where $\Phi_l^{(0)}(1-U) = \int_{-\infty}^{\infty} e^{-(1-U)x^2} h_l^2(x)dx$,

$$\Phi_l^{(1)}(1-U) = \int_{-\infty}^{\infty} e^{-(1-U)x^2} h_l(x)\left[h_{l+1}(x) + \frac{l}{2} h_{l-1}(x)\right]dx,$$

$$\rho = \begin{cases} \frac{b}{\lambda} & \text{if } \phi_1 = 0;\pi \\ \frac{a}{\lambda} & \text{if } \phi_1 = \frac{\pi}{2};\frac{3\pi}{2} \end{cases}, \text{ and } U = \begin{cases} b^2\frac{\varepsilon}{a} & \text{if } \phi_1 = 0;\pi \\ a^2\frac{\varepsilon}{b} & \text{if } \phi_1 = \frac{\pi}{2};\frac{3\pi}{2} \end{cases}$$

In the particular case, electrodes have the circular shape with radius r_0 and the following expression is valid for the amendment $W_{kl}(n)$

$$W_{kl}(n) \cong \frac{\sqrt{cd_1}\,2^{k+l+1}}{\pi k!l!} \left\{ \frac{1}{\varepsilon} h_k\left(\sqrt{\varepsilon}\rho\right)\left[h_{k+1}\left(\sqrt{\varepsilon}\rho\right) + \frac{l}{2} h_{l-1}\left(\sqrt{\varepsilon}\rho\right)\right], \right. \tag{26}$$
$$\left. \Phi_l^{(0)}(1-\varepsilon) + \frac{\sqrt{\varepsilon}}{\rho} h_k^2\left(\sqrt{\varepsilon}\rho\right)\Phi_l^{(1)}(1-\varepsilon) \right\} e^{-\varepsilon\rho^2}$$

where $\rho = r_0/\lambda$. If the electrode dimensions anisotropy dominates over the anisotropy of tensor $d^{\alpha\beta}$, i.e. $\frac{a^2}{b^2} < \varepsilon$ and $\rho = \frac{a}{\lambda}$ then the add $W_{kl}(n)$ is estimated by the expression

$$W_{kl}(n) \cong \begin{cases} \frac{\sqrt{cd_1}\,2^{k+l+1}}{\pi k!l!} e^{-\rho^2} \left\{ \frac{a^2}{b^2\varepsilon^2\rho} h_l^2(\rho)\Phi_k^{(1)}\left(\frac{a^2}{\varepsilon b^2}\right) + h_l(\rho)\left[h_{l+1}(\rho) \right.\right. \\ \left.\left. + \frac{l}{2} h_{l-1}(\rho)\right]\Phi_k^{(0)}\left(\frac{a^2}{\varepsilon b^2}\right)\right\}, \quad \text{if } \phi_1 = 0;\ \pi \\ \frac{\sqrt{cd_1}\,2^{k+l+1}}{\pi k!l!} e^{-\varepsilon\rho^2} \left\{ \frac{h_k\left(\sqrt{\varepsilon}\rho\right)}{\varepsilon}\left[h_{k+1}\left(\sqrt{\varepsilon}\rho\right) + \frac{k}{2} h_{k-1}\left(\sqrt{\varepsilon}\rho\right)\right]\Phi_l^{(0)}\left(\frac{a^2\varepsilon}{b^2}\right) \right. \\ \left. + \frac{a^2\sqrt{\varepsilon}}{b^2\rho} h_k^2\left(\sqrt{\varepsilon}\rho\right)\Phi_l^{(1)}\left(\frac{a^2\varepsilon}{b^2}\right)\right\}, \quad \text{if } \phi_1 = \frac{\pi}{2};\ \frac{3\pi}{2} \end{cases} \tag{27}$$

3.2. Complicated electrodes

Let us consider resonator with the most complicated electrodes (Fig.2*b*, *c*). Assume an electrode shape has a "special" point (point 1 in Fig.2*b* and points 1, 2, 3, and 4 in Fig.2*c*). There the breaking sector is tangent to the contour and an angle $\phi_2 < \pi$. Such a shape with the only segment of an angle ϕ_2 yields

$$W_{kl}(n) \cong \begin{cases} \frac{\sqrt{cd_1}\,2^{k+l+1}}{\pi k!l!} e^{-\rho^2} h_l(\rho)\left[h_{l+1}(\rho) + \frac{l}{2} h_{l-1}(\rho)\right]\Phi_k^{(+)}\left(\frac{2\rho}{\sqrt{\varepsilon}} ctg\frac{\phi_2}{2}\right), \\ \qquad \text{if } \phi_1 = 0;\pi \\ \frac{\sqrt{cd_1}\,2^{k+l+1}}{\pi k!l!\varepsilon} e^{-\varepsilon\rho^2} h_k\left(\sqrt{\varepsilon}\rho\right)\left[h_{k+1}\left(\sqrt{\varepsilon}\rho\right) + \frac{k}{2} h_{k-1}\left(\sqrt{\varepsilon}\rho\right)\right]\Phi_l^{(-)}\left(2\varepsilon\rho ctg\frac{\phi_2}{2}\right), \\ \qquad \text{if } \phi_1 = \frac{\pi}{2};\frac{3\pi}{2} \end{cases} \tag{28}$$

where $\Phi_k^{(\pm)}(z) = \int_0^{+\infty} \exp(-zU) h_k^2(\pm x)dx$.

The electrodes may also be done of a another shape, namely, the rectangular, square, hexagons and so on. If a contour Γ has a flat section at a distance a from the center, then we must use the following add

$$W_{kl}(n) \cong \begin{cases} \frac{\sqrt{cd_1}\,2^l}{\sqrt{\pi}l!} e^{-\rho^2}\varepsilon h_l(\rho)\left[h_{l+1}(\rho) + \frac{l}{2} h_{l-1}(\rho)\right], & \text{if } \phi_1 = 0;\pi \\ \frac{\sqrt{cd_1}\,2^k}{\sqrt{\pi}k!\varepsilon} e^{-\varepsilon\rho^2} h_k\left(\sqrt{\varepsilon}\rho\right)\left[h_{k+1}\left(\sqrt{\varepsilon}\rho\right) + \frac{k}{2} h_{k-1}\left(\sqrt{\varepsilon}\rho\right)\right], & \text{if } \phi_1 = \frac{\pi}{2};\frac{3\pi}{2} \end{cases} \tag{29}$$

where $\rho = a/\lambda$.

3.3. Studies of the amendment $W_{kl}(n)$

It follows from the formulas (25)–(29) that for the circular-shape electrode only the amendment $W_{kl}(n)$ does not depend on angle ϕ_1. Let us investigate the influence of the disorientation angle ϕ_1 on the $W_{kl}(n)$ behavior. In first, we introduce the coefficient of anisotropy $K_{kl}(n) \ge 1$

$$K_{kl}(n) = \frac{W_{kl}(n)\big|_{\alpha}}{W_{kl}(n)\big|_{\alpha=0}}, \tag{30}$$

which maximal value $K_{00}(n)$ is given by

$$K_{00}(n) \cong \begin{cases} \frac{a^2}{b\sqrt{a^2 - \varepsilon b^2}} \exp\left(\varepsilon\frac{b^2 - a^2}{\lambda^2}\right), & \text{if } \varepsilon < \frac{a^2}{b^2} \le 1, \quad \text{(a)} \\ \frac{1}{\sqrt{\varepsilon}} \exp\left[(1-\varepsilon)\rho^2\right] & \text{if } \frac{a^2}{b^2} < \varepsilon, \quad \text{(b)} \\ \frac{1}{\varepsilon^2} \exp\left[(1-\varepsilon)\rho^2\right] & \text{(c)} \\ \frac{1}{\varepsilon\sqrt{\varepsilon}} \exp\left[(1-\varepsilon)\rho^2\right] & \text{(d)} \end{cases} \tag{31}$$

and depends on the piezoelectric parameters and electrodes shape. For the ellipsoid shape of electrodes we get that (*a*), (*b*), and (c), (d) in (31), correspond to the shape with broken boundary area. Let us add that the coefficient (31) for (30) strongly depends on the angle ϕ_1. Thus, the electrode shape influences on the frequency spectrum via the amendment $W_{kl}(n)$. Here the major factor is a disorientation angle ϕ_1. If the electrode has a circular shape then $W_{kl}(n)$ does not depend on ϕ_1, i.e. the coefficient $K_{kl}(n)$ equals 1. In contrary, possible cracks in the electrode (break of its segment) or chips (flat area of the electrode contour) increase that dependence. Appreciable "splash" of the amendment will occur in a case of the break direction coin-

cides to major axis of tensor $d^{\alpha\beta}$ or if direction of the normal $\vec{N}$ coincides to the flat area and the major tensor axis $d^{\alpha\beta}$. So, through the analysis of the tensor $d^{\alpha\beta}$ and orientation of its major axes we come to the accurate frequency spectrum definition of the convex piezoelectric element employing thickness-shear vibrations.

4. CONCLUSION

We have shown in the paper that the advanced approach for accurate frequency spectra definition of a quartz crystal resonator employing thickness-shear vibrations with electrodes of an arbitrary shape is efficient within a frame of the linear piezoelectric theory. Numerous calculations showed that in any case the predicted frequency values fit experimental data much more good in comparison with that obtained by the traditional linear theory approach. It allows offering the method for high accuracy definition of the crystal resonator frequencies being based on the minimal number of parameters some of those are obtained experimentally. To justify it, we brought the example of the frequency spectra calculation of natural vibrations of an AT-cut quartz crystal resonator based on the experimental data and predicted values obtained by Hertl, *et al* in a frame of the linear theory. We note with this that in order to get more accurate estimates and expend the analyzing spectra width within an advanced approach, one may take into consideration more large number of the parameters.

The most important and practically useful result consists expressions for the amendments $W_{kl}(n)$ correspond to the electrodes with the most preferable shapes, which may be used for anharmonic sensors excitation with required energy jointly with the fundamental mode by modulational method [2]. We have brought those for the electrodes of the circular, ellipsoid, rectangular, square and some another ones and show with this that the center of energy localization does not coincide with the center of symmetry of the convex electrode, thereat the similar oscillations are localized in the definite direction. The favored direction of localization is conditioned by many factors, namely, by anisotropy of piezoelectric, electrode shape, orientation of the symmetry axes of electrodes and tensor $d^{\alpha\beta}$. Thus, from the another hand, through the resonator spectra studies we have possibility to get the additional information about piezoelectric element.

5. REFERENCES

[1] J. R. Vig and F. L. Walls, "Fundamental limits on the frequency instabilities of quartz crystal oscillators," in *Proc. IEEE 48th Freq. Contr. Symp.*, 1994, pp.506-523.

[2] Yu. S. Shmaliy, "The modulational method of quartz crystal oscillator frequency stabilization," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. 45, pp.1476-1484, Nov. 1998.

[3] Yu. S. Shmaliy, A. Ph. Kurochka, E. G. Sokolinskiy, A. V. Marienko, "Quartz crystal oscillator with effective aging rate compensation", in *Proc. 12th European Freq. Time Forum*, Warsaw, Poland, 1998, pp.234-239.

[4] "An American National Standard. IEEE Standard on Piezoelectricity," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol.43, pp.717-772, Sept. 1996.

[5] J. Zelenka, Piezoelectric resonators and their applications, Amsterdam; New York: Elsevier, 1986. 301p.
S. S. Nedorezov, "Localized vibrations of piezoelectric resonators," *Izvestija VUZov USSR, Radiophyzika*, vol.33, no. 12, pp.1417-1422, Dec. 1990.

[6] D. S. Stevens and H. F. Tiersten, "Analysis of doubly-rotated controured quartz crystal resonators," in *Proc. IEEE 39th Freq. Contr. Symp.*, 1985, pp.436-446.

[7] R. C. Peach, "A variational method for the design of trapped energy resonators," in *Proc. IEEE 39th Freq. Contr. Symp.*, 1985, pp.392-399.

[8] S. Hertl, E. Bones, L. Winner, and M. Schmid, "Inverstigation of quartz crystal thickness shear and twist modes using a new noninterferometric laser speckle measurement method," in *Proc. IEEE 39th Freq. Contr. Symp.*, 1985, pp.535-542.

[9] S. S. Nedorezov and E. A. Ganenko, "Frequency spectrum of the piezoelectric resonators," *Dopovidi NAN Ukraine*, no. 10, pp.31-35, Oct. 1996.

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

HYPERSTABLE QUARTZ OSCILLATORS FOR SPACE USE

R.J. BESSON*, M. MOUREY*, S. GALLIOU*, F. MARIONNET*, P. GUILLEMOT**

***Ecole Nationale Supérieure de Mécanique et des Microtechniques**
26 Chemin de l'Epitaphe
25030 BESANCON CEDEX – FRANCE

****Centre National d'Etudes Spatiales**
18 avenue Edouard Belin
31404 TOULOUSE CEDEX 4 – France

Abstract

New hyperstable quartz oscillators have been introduced at the Joint Meeting in Besançon (April 1999). They exhibited stabilities down to $5x10^{-14}$ together with drift down to $5x10^{-12}$/day and resistance under severe environments.

However, those oscillators had to be modified for space qualification and miniaturized from 650 cm^3 down to less than 250 cm^3. In particular, the oscillators had to be constructed for use under high vacuum. In other words, best results should be obtained under vacuum though running the oscillator under atmospheric pressure should also be possible (then the thermal behavior could be shightly degraded).

In this paper, the oscillator construction is first presented together with reasons for technical choices (in fact, some construction details are very important to improve final results under qualification tests). Also presented are the most important qualification tests for space use, in particular phase noise and Allan Variance, sensitivity to environmental parameters (temperature, magnetic field, vibrations...).

In conclusion, the results are presented and discussed in view of future use under space conditions.

I - INTRODUCTION

Recently, quartz oscillators with short term stabilities clearly under $1x10^{-13}$ were needed for the cold atom space project (Project "Horloge Atomique par Refroidissement d'Atomes en Orbite" said **PHARAO Project**) and interesting results were obtained [1]. All published measurements of short term stabilities **were made outside the laboratory** in particular in Jet Propulsion Laboratory in Pasadena, CA (needless to say measurements of $\sigma_y(\tau)$ down to $5x10^{-14}$ or less on a 10 MHz quartz oscillator are not trivial). All fundamental concepts have been previously described, see [1] and references quoted then. Of course, mini-BVA resonators for space applications have been used [2]. The goal of this poster paper was to show progress in construction of **Identification and Qualification Models** (IQM) and deliver practical results on those.

II – CONSTRUCTION OF IDENTIFICATION AND QUALIFICATION MODELS

The hyperstable bva quartz oscillator for space applications can be seen on next general view.

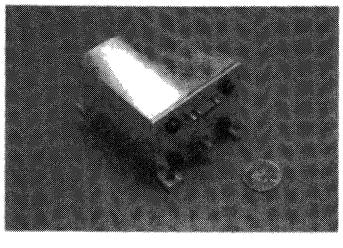


Fig. 1 - Hyperstable bva quartz oscillator for space applications - General view

Inside view of Fig. 2 shows various parts glued for vibration purposes. Internal oscillator block is about 48 cm^3.

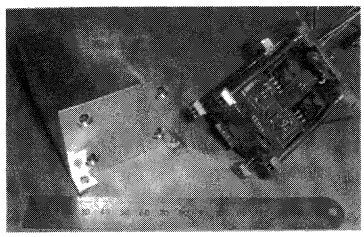


Fig 2 - Hyperstable oscillator (identification and qualification model). Inside view shows parts glued for vibration purposes. Internal oscillator block is about 48 cm³

Fig. 3 shows internal block opened with two internal mumetal half-shells, oven resonator and 4 stainless steel fixations. Those fixations have been carefully calculated to allow a given thermal loss flux under vacuum.

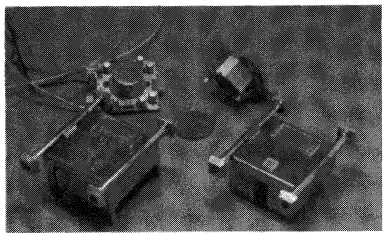


Fig. 3 - Internal block (48 cm³) opened showing the 2 internal half mumetal shells, oven, resonator and 4 stainless steel fixations.

Fig. 4 shows the internal oscillator with the two half-shells assembled.

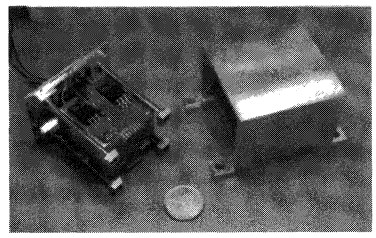


Fig. 4 - Internal oscillator with the 2 half shells assembled

III – OVEN DESIGN AND THERMAL DATA

Internal oscillator (48 cm^3) is located inside 2 mumetal half-shells. True thermal regulation (≈80°C) is achieved for resonator and oscillator circuits ; thermal regulation also drives heating elements located on external side of the two half-shells.

Hyperstable oscillator will be operated between 25°C and 35°C with temperature oscillation (± 1°C in 90 minutes) around working point (in vacuum or air).

Frequency variations all causes included must be less than $2x10^{-11}$.

Fig. 5 and Fig. 6 show concumption tests respectively under atmospheric pressure and under vacuum.

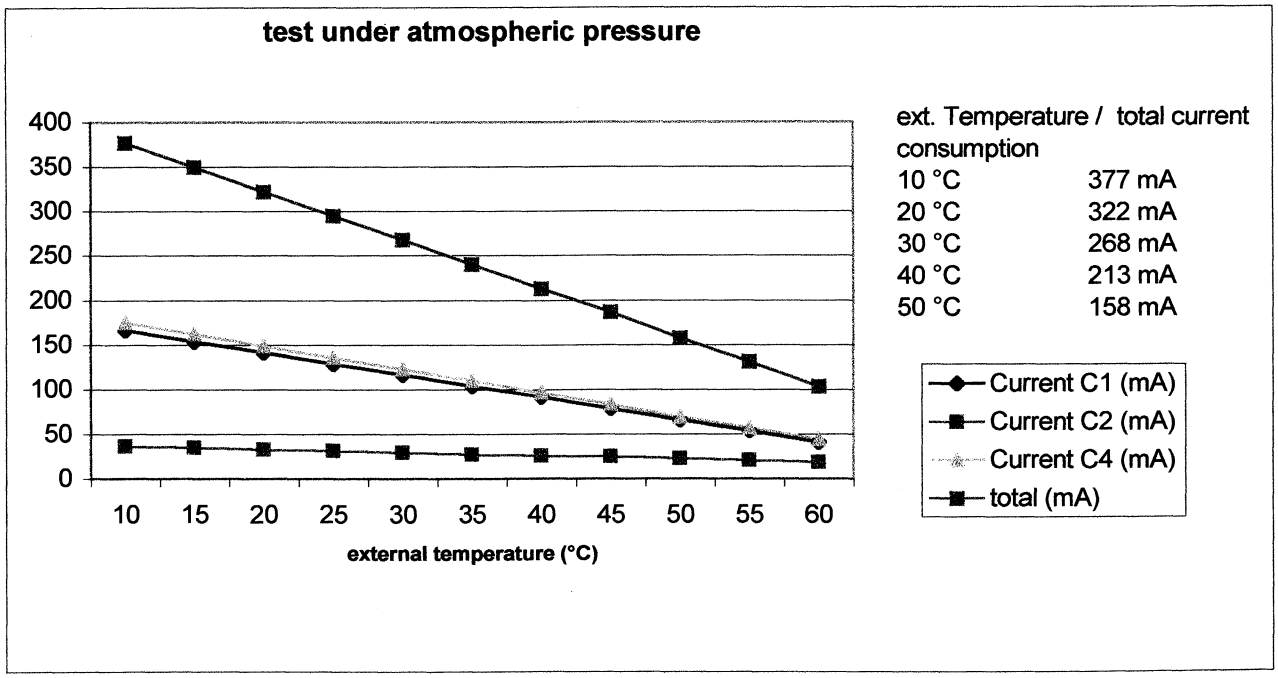


Fig. 5 - current consumption of the only oven (oscillator alone : 20mA) under 15V

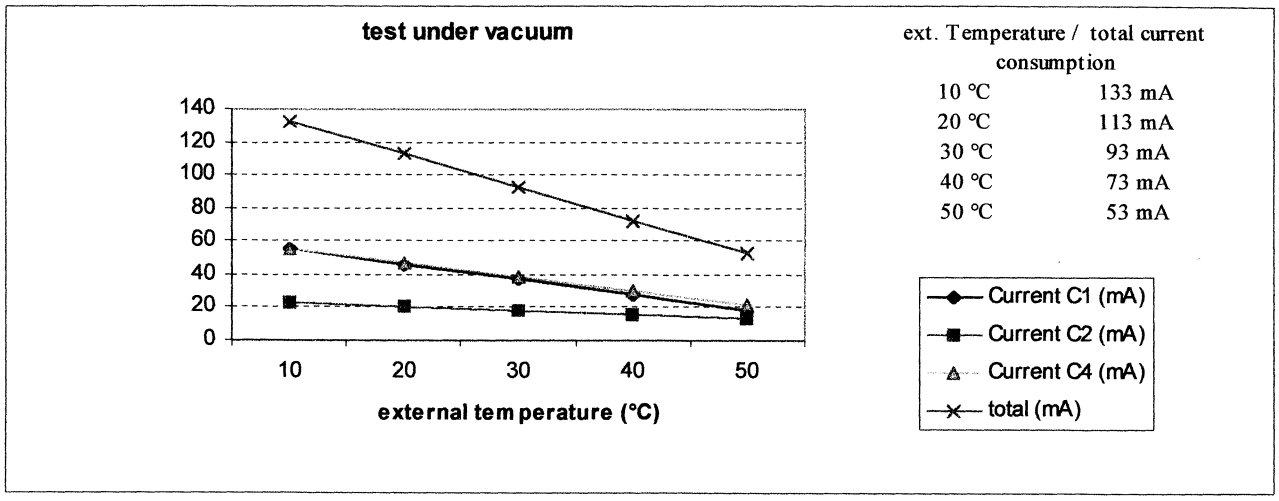


Fig. 6 - current consumption of the only oven (oscillator alone : 20mA) under 15V

IV – HYPERSTABLE OSCILLATOR UNDER VIBRATION TESTS

Two identification and qualification models have been extensively tested under vibrations (Fig. 7).

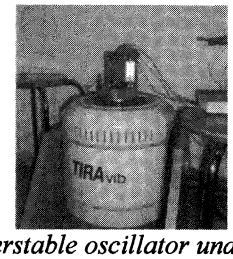
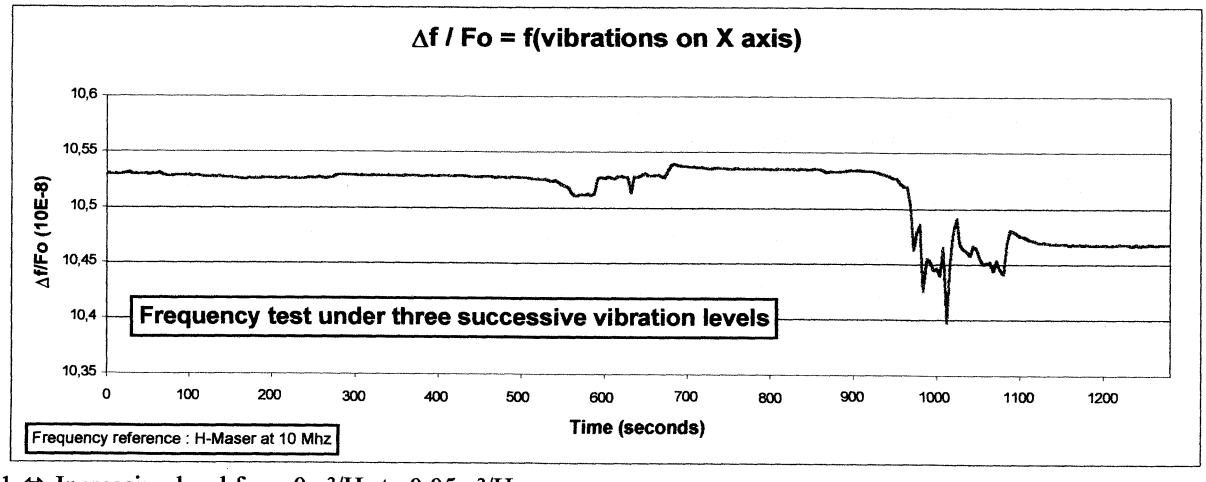


Fig. 7 - Hyperstable oscillator under vibration tests. (The oscillator, previously located very close to its actual place does not exhibit any magnetic field influence).

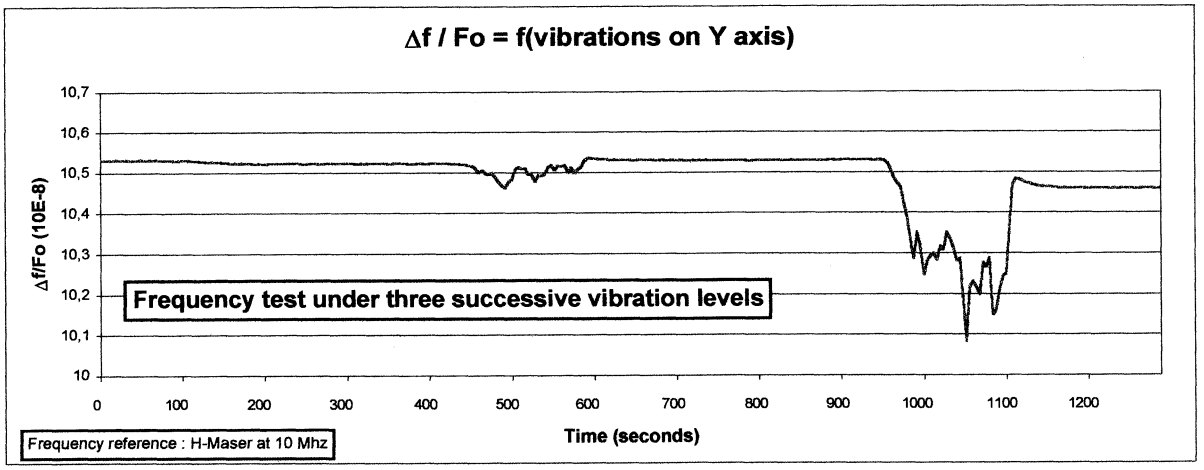
The first thing to do was to check that results obtained were not partially caused by some magnetic field influence. But the oscillator located very close to its actual place did not exhibit any magnetic field influence (more exactly influence was under our system resolution). Then, frequency tests under three successive vibration levels were performed with continuous frequency recording. This is unusual since frequency is generally measured before and after vibrations. Results appear on Fig. 8, Fig. 9 and Fig. 10.

One oscillator exhibited an important frequency permament deviation ($2x10^{-8}$) after the test but this turned out to be connected with deficient soldering of one connection and could easily be fixed.



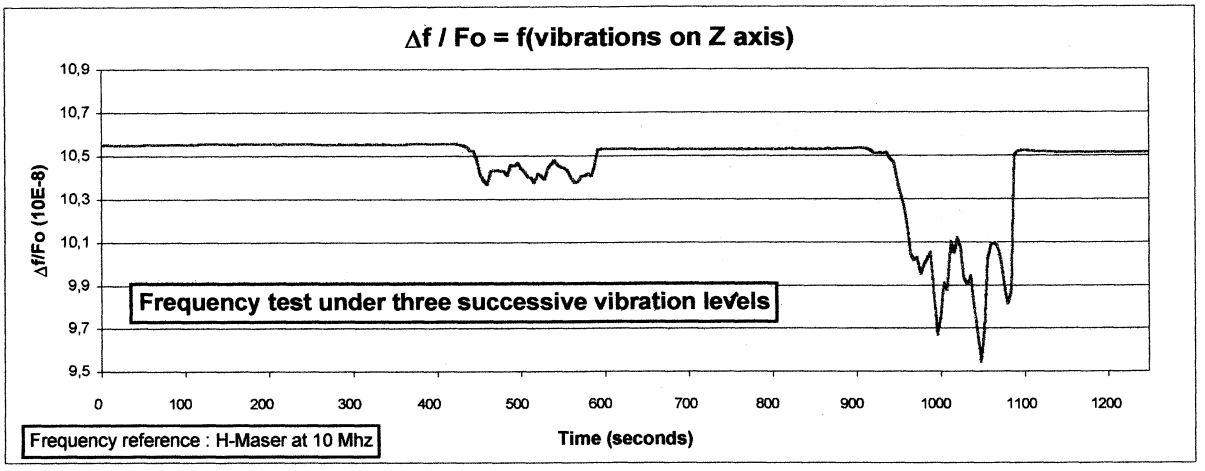
1 ⇔ Increasing level from 0 g²/Hz to 0,05 g²/Hz
2 ⇔ Constant level at 0,05 g²/Hz during 2 minutes, then null level
3 ⇔ Increasing level from 0 g²/Hz to 0,2 g²/Hz
4 ⇔ Constant level at 0,2 g²/Hz during 2 minutes, then null level
5 ⇔ Increasing level from 0 g²/Hz to 0,5 g²/Hz
6 ⇔ Constant level at 0,5 g²/Hz during 2 minutes, then null level

Fig. 8



1 ⇔ Increasing level from 0 g²/Hz to 0,05 g²/Hz
2 ⇔ Constant level at 0,05 g²/Hz during 2 minutes, then null level
3 ⇔ Increasing level from 0 g²/Hz to 0,2 g²/Hz
4 ⇔ Constant level at 0,2 g²/Hz during 2 minutes, then null level
5 ⇔ Increasing level from 0 g²/Hz to 0,5 g²/Hz
6 ⇔ Constant level at 0,5 g²/Hz during 2 minutes, then null level

Fig. 9



1 ⇔ Increasing level from 0 g²/Hz to 0,05 g²/Hz
2 ⇔ Constant level at 0,05 g²/Hz during 2 minutes, then null level
3 ⇔ Increasing level from 0 g²/Hz to 0,2 g²/Hz
4 ⇔ Constant level at 0,2 g²/Hz during 2 minutes, then null level
5 ⇔ Increasing level from 0 g²/Hz to 0,5 g²/Hz
6 ⇔ Constant level at 0,5 g²/Hz during 2 minutes, then null level

Fig. 10

V – CONCLUSION

One important test to be performed was to check short term stabilities. Results have been obtained in our facility against an hydrogen maser specially constructed for us by Observatoire de Neuchâtel (Switzerland). Four oscillators have been tested and results were always clearly under $1x10^{-13}$ between 1s and 100s. Results completely confirmed previous measurements [1]. Extensive tests will now be performed on final qualification models.

Then, **principal results are summarized as follows :**

- Hyperstable 10 MHz oscillators (using bva space qualified resonators 3rd overtone SC-cut resonators) have been modified in view of space applications.

➢ Performances already demonstrated (Joint Meeting 1999) are maintained under vacuum and space conditions.

➢ Principal technical choices justify as follows:

- Greatly improved stiffness allows no resonance below 2000 Hz and permits non destructive 0.5 g^2/Hz vibration levels.
- Four stainless steel fixations connect internal oscillator block to external shell and allow precise thermal exchanges.
- Almost no electrical wire connections.
- two mumetal half-shells allow insensitivity to magnetic field.

➢ Principal results and oscillator characteristics:

- Low volume and weight (227 cm^3 for 423 g)
- $\sigma_y(\tau)$ always < $1x10^{-13}$ (1s < τ < 100 s)
  σ_y(10s),
  typical: $7x10^{-14}$,
  lowest: $5x10^{-14}$,
  aging (typical): $5x10^{-12}$/day
- $\Delta f/f_o < 2x10^{-11}$ all causes included
- Excellent performances under vibrations.

REFERENCES

[1] R.J. BESSON, M. MOUREY, S. GALLIOU, F.MARIONNET and al
"10 MHz hyperstable quartz oscillators performances"
Joint Meeting EFTF-FCS, Besançon, April 1999
Proceedings pp. 326-330, 1999.

[2] J.J. BOY, R.J. BESSON, N. FRANQUET, D. THIEBAUT, R. PETIT, Ph. BAVIERE, M. BRUNET
"Evaluation of mini-BVA resonators for space applications: test program and results"
10th European Frequency and Time Forum, Brighton, March 1996
Proceedings pp. 31-36, 1996.

EVIDENCE FOR NON-EXISTENCE OF SURFACE ACOUSTIC WAVES IN EXCEPTIONAL CASES IN PIEZOELECTRIC CRYSTALS

V. G. Mozhaev* and M. Weihnacht**
*Faculty of Physics, Moscow State University, 117234 Moscow, Russia;
**Institut fuer Festkoerper- und Werkstofforschung, Helmholtzstrasse 20, 01069 Dresden, Germany

Surface-acoustic-wave (SAW) devices find extensive application in modern communications. Piezoelectric crystals are used as substrate materials for creating such devices. However, contrary to a non-piezoelectric case, piezoelectricity may result in the prohibition of SAW propagation in some exceptional cases. The first explicitly mentioned example of the geometry where no surface wave solution could be found corresponds to rotated *Y*-cut *X*-propagating berlinite [1]. Similar cases have been recently found for langasite and quartz [2]. On the other hand, the absence of SAW solution in a computational search for any specific geometry is not a guarantee of true non-existence of SAWs because it can be related only to numerical difficulties of computations near the degeneracy points.

To find more convincing evidence for non-existence of SAWs, we study in detail a fine structure of the phase velocity branches for normal and leaky SAWs near the degeneracy points of exceptional-wave type in rotated *Y*-cut piezoelectric trigonal crystals, with SAW propagation in the *X*-direction. It is found that the influence of piezoelectricity in the case of free surface of the crystal results in the splitting of exceptional waves which occur in langasite at the cut-rotation angles 139° and 150°. Such two generalized exceptional waves in crystals of 32 class possess only lateral piezoelectric activity, that is, they are accompanied by only near-surface electrical fields due to non-zero normal component of the electrical displacement while the electrical field in the bulk of crystal tends to zero with distance from the surface. The range of angles between the values corresponding to the splitted exceptional wave lies in the region 137.1° - 152.7° where according to Ref. [2] SAWs do not exist. We have found a prolongation of the SAW branch between the mentioned angles, that is, in the prohibited region. This continuity of the velocity curve is in agreement with an expected polynomial form of a secular equation for SAWs. On the other hand, the prolongation of the curve is identified as a non-physical solution because it includes the partial waves increasing away from the surface. Thus, this fine structure of the velocity curves is a convincing evidence of the non-existence of SAWs in the specific geometry under consideration. Similar fine structure has been found for leaky SAWs in quartz. The surprising thing is that there is a slight deviation of the leaky SAW degeneracy points and exceptional wave velocities in the case of free surface.

1. R. M. O'Connell and P. H. Carr, High piezoelectric coupling temperature compensated cuts of berlinite ($AlPO_4$) for SAW applications, IEEE Trans. **SU-24**, 376 (1977).
2. T. Sato, M. Murota, and Y. Shimizu, Characteristics of Rayleigh and leaky surface acoustic wave propagating on a $La_3Ga_5SiO_{14}$, Jap. J. Appl. Phys. Part 1 **37**, 2914 (1998).

HIGH STABILITY 10 MHz OCXO
NEW DESIGN IMPROVING PERFORMANCE vs SIZE AND COST RATIOS

E. GIRARDET, P. CAPPELLE A.R.ELECTRONIQUE F-25048 BESANÇON
B. WOLCOFF A.R.ELECTRONIQUE F-78500 SARTROUVILLE

14F, Rue de Franche Comté – 25048 BESANCON CEDEX
Tél. : (00.33.3) 81.88.66.77 – Fax : (00.33.3) 81.88.33.55
Email : are@ar-e.com

1- ABSTRACT

During the last few years, different papers have been presented, showing either 10 MHz ultra stable oscillators, stability class of a few 10^{-11}, short term of some 10^{-13}, or miniaturized, SMD versions, with a class of stability of a few 10^{-8}.

A.R.Electronique has investigated in an alternative way which takes place between these products and consists in providing high stability, very low phase noise in low size enclosures, combined with a very efficient performance versus cost ratio.

This concept is based on the use of A.R.Electronique designed, SC cut resonators (HC-53/U or TO-8, 4 points enclosure), associated with a very low noise electronic. The thermal regulation is done with a single oven, low size. This principle has been derivated in two versions :

The miniature JUMBOSTAR-FB family offers a class of stability of a few 10^{-9}, short term lower than 8.10^{-13} / 1s, low phase noise, in a 27 x 36 x 16 mm package.

The PULSAR-SL family reaches short term stability better than 5.10^{-13}/ 1s, very low phase noise, associated with temperature stability of a few 10^{-10}, in a very low profile case 41 x 51 x 20 mm.

The same concept has been used for ruggerized versions devoted to airborne applications (civil / military), with low g sensitivity resonators. These oscillators give very fast warm-up and low phase noise under vibration.
The paper will breafly describe the electronic, thermal and mechanical designs.

The main electrical characteristics and frequency stability will be given and placed in regard with other actual products.

Special developements and characterizations for particular applications will be also presented.

2- INTRODUCTION

Placed beetwen the miniature, SMD or not, OCXO's with limited characteristics and very high stability, high volume, expensive OCXO's, the last generation of A.R.Electronique oscillators have been designed for the optimization of three criterias : stability-size-cost.

TABLE [1] : TECHNICAL CHARACTERISTICS

PARAMETER	SMD OCXO	STD OCXO JUMBOSTAR	JUMBOSTAR-FB	PULSAR-SL	USO (GROUND)
Freq. vs Temperature (-20,+70)°C	+/-1.10-8 +/-5.10-8	+/-1.10-8 +/-2.10-8	+/-2.10-9 +/-1.10-8	+/-2.10-10 +/-1.10-9	+/-5.10-11 +/-5.10-10
Aging / month	+/-5.10-9 +/-1.10-8	+/-5.10-9 +/-1.10-8	+/-3.10-9 +/-1.10-8	+/-2.10-9 +/-5.10-9	+/-6.10-10 +/-4.10-9
Retrace (24 hours turn off, 60 min. after turn on)	N.S.	+/-3.10-9	+/-3.10-9	+/-2.10-9	+/-2.10-9
Phase noise L(F) (dBc / Hz) 1 Hz 10 Hz 100 Hz	N.S.	-100 -130 -145	-105 -133 -155	-110 -138 -156	-110 -138 -157
S.T.S. (Allan Variance) 1s 10s	N.S. N.S.		8.10-13 1.10-12	5.10-13 7.10-13	N.S. 5.10-13
Voltage sensitivity (Vcc+/-5%) Load sensitivity (50 Ohms +/- 10%)	N.S.	+/-2.10-9 +/-5.10-9	+/-5.10-10 +/-5.10-10	+/-5.10-11 +/-5.10-11	+/-5.10-11 +/-5.10-11
Frequency start-up (+25°C) +/- 1.10-8	N.S.	3 min	3 min	3 min	15 min
Power consumption (turn on / steady state)	5W / 1W	3W / 1W	3W / 1,3W	6,6W / 1,7W	8W / 2,5W
Size (mm) Volume (cm3)	25 x 22 x 14 8	27 x 36 x 16 15	27 x 36 x 16 15	41 x 51 x 20 40	60 x 67 x 40 160
Avionic versions		YES	YES	YES	

They are destinated for the new applications in the telecommunication fields, digital broadcast, TTNC, GPS receivers, instrumentation, where the continuous evolution of performances (transmission rate speed, precision of localization...) requires the use of high stability frequency sources, together with miniature size (low profile) and reasonable price.

This last oscillator's design has been declined in two families offering complementary classes of stability (JUMBOSTAR-FB, PULSAR-SL). Specific versions for hardened environmental conditions (civil and military avionic) have been also developped. This concept was also characterized, as preliminary investigation, under specific space environment (vacuum).

2- CHARACTERISTICS REQUIREMENTS

In order to precise the technical context of this study, the table [1] shows the main electrical and mechanical characteristics of these oscillators, in comparison with the related existing families.

3- DESIGN – TECHNICAL CONCEPT

The technical requirements have directly involved some technological choices concerning the crystal resonator, the oscillator's design, as well as the mechanical structure and the thermal control.

3.1- Quartz resonator :

The 10 MHz nominal frequency is well adapted for the majority of the expected needs. Based on SC cut crystals, 3rd overtone, this frequency range allows to reach a best compromise for stability rate, size and manufacturing costs.
The geometrical blanks design is optimized in order to reduce the complex lapping phases.
The housing in HC-53/U or low profile, T08 enclosure, 4 points mounting is well adapted for the reduction of the oven volum, and consequently of the whole oscillator's size.
The crystal manufacturing process includes stabilization cycles as well as computer characterized pre-aging tests, according with adapted high temperature profiles that have been optimized during this study.
The levels of achieved characteristics for this type of crystal are shown in the following chapters.
These measurements have been performed on the whole oscillators.
The 10 MHz range resonators can be used within a larger frequency scale, through a multiplication or a division of the oscillator's frequency. However, it will induce a phase noise degradation on the final frequency (20 log N factor for multiplication, noise contribution due to the logic IC's for division).

For very special requirements, the need of a direct, lower frequency resonator can be planed (typically 5 MHz, SC cut, 3rd O.T.). The fundamental concepts of these oscillators can be used, with some restrictions due to the crystal enclosure size (minimum oscillator's height of 25mm).

3.2- Electronic circuitry :

The high level of expected characteristics for phase noise, short term stability (Allan Variance) requires the use of low noise components with optimized bias points. This approach eliminates the use of full integrated, ASIC oscillators.
The oscillator is a colpitts types slightly modified. The peripheral funtions (insulation and amplifier stages, voltage regulators) have been specially designed, in order to reduce both phase noise and frequency stability contributions : supply voltage sensitivity, load sensitivity, as well as long term stability, thermal behavior.

3.3- Thermo-mechanical structure :

The mechanical design is involved by the expecting size and costs requirements. The internal structure is centered around the oven, which is optimized for the crystal resonator's thermal regulation. For the JUMBOSTAR-FB family, the sensitive electronic functions are pseudo-temperature regulated. This regulation is more precise for the PULSAR-SL versions, where the size criterions are not so severe. A special effort has been made for minimizing the number of the mechanical components.
These mechanical structures have been also declined for hardened versions, adapted to avionic requirements (larger temperature ranges, vibrations, schok). In that case, the optimization of the mechanical behavior under vibration modifies the thermal insulation of the oven, which degrades the consumption. The start-up speed, which is an important requirement for these products, will require a significative increasement of the oven power after turn-on.

3.4- Electronic temperature control :

The final thermal regulation of the oven is function of the mechanical structure as well as the associated electronic circuitry. The single oven solution has been choosen : a double oven concept is greatly unfavourable for global size, adjustment processes and reliability.
The control circuit is fully integrated (ASIC), except the temperature sensor and the power heating elements. Beyond the obtention of high regulation rates for low temperature slopes (<1°C/min.), the thermal design has taken into account the response to fast temperature steps : these kinds of OCXO's are normally working under very low temperature gradient. However, they have to keep their stability under accidental disturbances (typically, delocked PLL after the opening of a rack system).

For the derived avionic versions, the internal temperature gradient effect (ITGE) has to be characterized. It corresponds to the overall frequency shift due to the variation of the OCXO's geometrical orientation.

The table [2] shows the intrinsec thermal behavior of the different oscillator's families.
These characterizations are made by measurement of the b mode frequency of the SC cut resonators (linear slope of –30ppm/°C).
For evalutation, some probe tests have been realized under vaccum (pressure < 10-5 torr), for space applications.

Table 2 : THERMAL REGULATION

OCXO			TEMP.REGULATION (ext. ΔT = 100°C)	
Type	Volume (cm^3)	Height (mm)	Quartz	Oscillator
JUMBO-FB	15	16	<0,2°C	<3°C
PULSAR-SL	40	20	<0,1°C	<1°C
JUMBO-S	15	16	<0,2°C	<5°C
PULSAR-EM	52	25	<0,1°C	<2°C

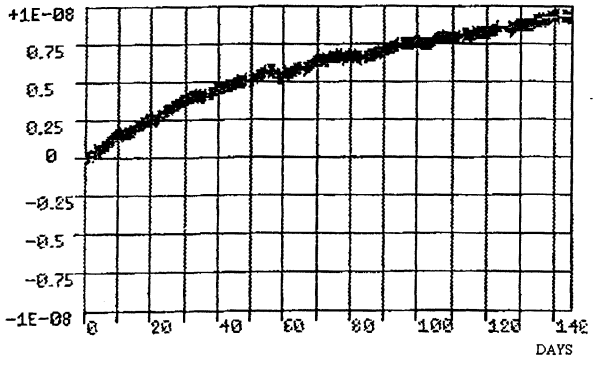
4- MEASUREMENT RESULTS

The most significative measurements are now shown and discussed. For general datas, please refer to table [1].

4.1- Long term stability

The aging is measured 2 times per day over 20 to 90 days (automatic bench). Fig [1] is a typical curve.

Fig [1] : PULSAR-SL 10 MHz - TYPICAL AGING



Statistical results on a batch of 100p are shown on fig [2] (after 21 days) and [3] (after 60 days).

Fig [2] : PULSAR-SL 10 MHz - AGING/MONTH AFTER 21 DAYS

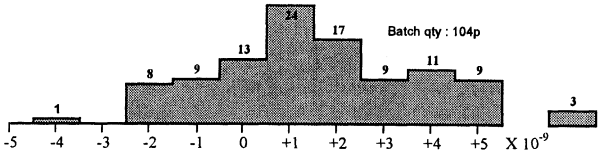
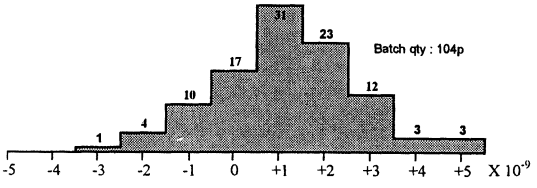


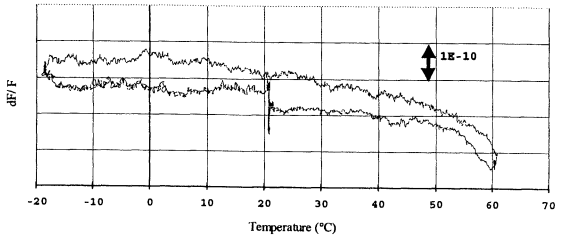
Fig [3] : PULSAR-SL 10 MHz - AGING/MONTH AFTER 60 DAYS



4.2- Frequency vs temperature

For very precise requirements, this characteristic is performed with a complete temperature cycle (+25, -20, +60, +25°C, 0,5°C/min, 400 min) (fig [4]).

Fig [4] : PULSAR-SL 10 MHz - FREQUENCY vs TEMP.



The similar test have been applied as preliminary experimentation under vacuum, without any change with the standard version (fig [5]).

Fig [5] : PULSAR-SL 10 MHz - FREQUENCY vs TEMP. (under vacuum)

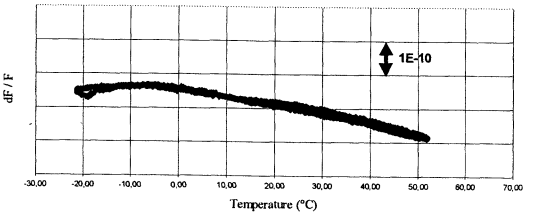


Fig [6] shows the frequency response to a temperature step of +/-35°C (temp. slope of 1°C/sec.).

Fig [6] : PULSAR-SL - FCY vs TEMPERATURE STEP

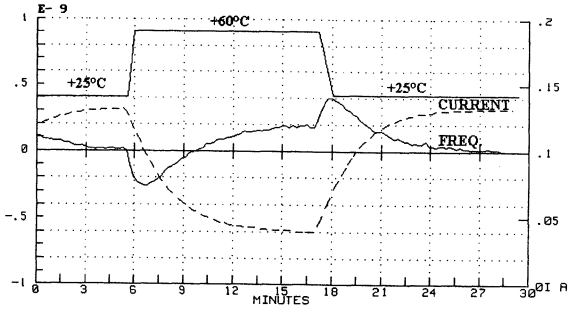
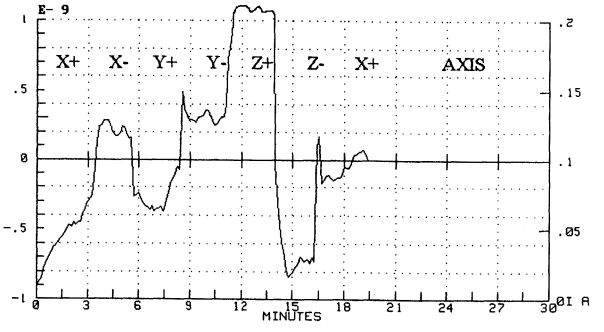


Fig (7] characterizes the I.T.G.E. test. Both temperature and g sensitivity effects are simultaneously taken into account.

Fig [7] : JUMBOSTAR-FB 10 MHz - I.T.G.E.



4.3- Start-up

The thermal design combined with the intrinsec SC cut characteristics offer a very short stabilization time for both frequency and consumption (fig [8], [9]).

Fig [8] : JUMBOSTAR-FB 10 MHz - START-UP (+25°C)

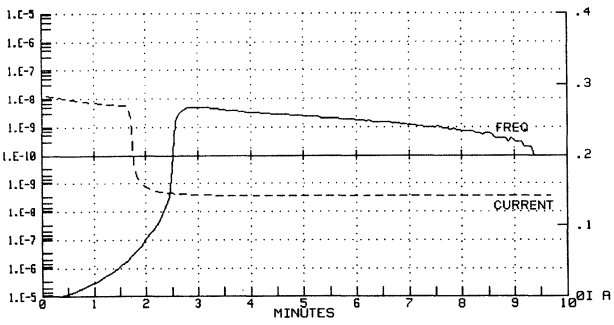
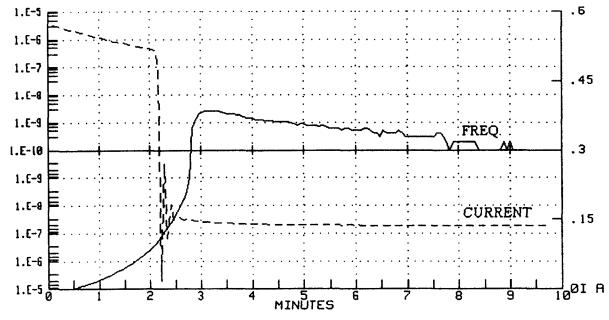


Fig [9] : PULSAR-SL 10 MHz - START-UP (+25°C)



4.4- Retrace

The retrace behavior is mainly dependant of the quartz resonator, as well as the test profile. Fig [10], [11] give statistical results on 20 pieces, 40 and 400 min after turn on (24 hours turn off at +25°C).

Fig [10] : PULSAR-SL 10 MHz - RETRACE (T0 + 40min)

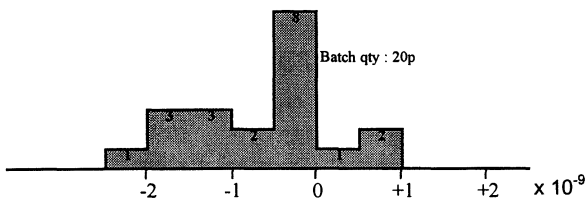
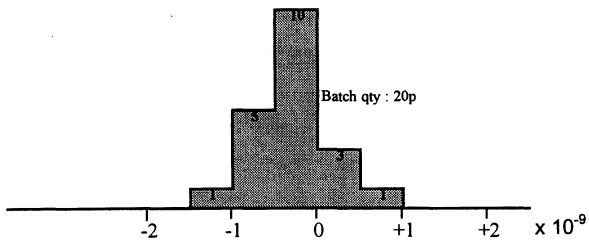


Fig [11] : PULSAR-SL 10 MHz - RETRACE (T0 + 400min)

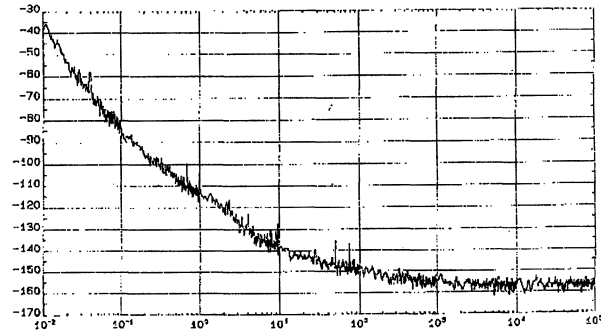


(frequency measurements refered to frequency before turn off)

4.5- Phase noise

This parameter has been particulary optimized during the study. Fig [12] shows a typical curve (PULSAR-SL).

Fig [12] : PULSAR-SL 10 MHz - PHASE NOISE (dBc/Hz)

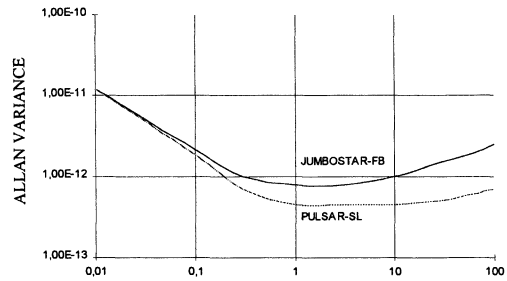


The JUMBOSTAR-FB is slightly degraded (see table [1]), because of some electronic simplifications due to the size.

4.6- Short term stability

The Allan Variance is measured in the time domain by comparison with a reference USO (fig [13]).

Fig [13] : ALLAN VARIANCE



This measurement can be also derivated from the phase noise curve, after frequency versus time domain conversion.

5- CONCLUSION

Initially designed for specific applications, these families of compact, high stability, low noise oscillators are now completely qualified. Their modular structures are well adapted for complementary developements where size and high performance criteria are critical.

REFERENCES

[1] B.Wolcoff, E.Girardet
High stability, miniature OCXO's for severe environmental conditions.
EFTF – IEEE, 1999

[2] J.Rutman, J.Groslambert
Characterization and measurement of frequency stability.
EFTF – 1990

[3] Identification and measurement of different types of noises present in a frequency source.
F.Vernotte, J.Groslambert, J.J.Gagnepain,
EFTF - 1990

SOME INVESTIGATIONS OF THERMOSENSITIVE QUARTZ RESONATORS AT CRYOGENIC TEMPERATURES

V.Georgiev, L.Spassov, L.Vergov, R.Velcheva
Institute of Solid State Physics, Bulgarian Academy of Sciences
72 Tzarigradsko Chaussee blvd., 1784 Sofia, Bulgaria

Introduction

It is well known that thermosensitive quartz resonators (TQR) can be used as high sensitive and high precise temperature sensor for temperatures from -80°C to +250°C (190K to 520K) [1]. Some publications [2,3] show the interest to implement such kind of resonators as temperature sensors for cryogenic temperatures. This paper represents some new results of our investigation of TQR in the temperature interval from 4.2K to 300K. The goal of this research was to test the possibility of using this kind of resonators as high sensitive temperature sensors for cryogenic temperatures. More over, we experiment with TQR prepared and tested to measure temperatures from 220K to 520K. So we hope to extend this interval and to cover the temperatures from 4.2K to 520K with one sensor.

1. Experiment

The TQR are prepared of synthetic quartz with Q-facto over 2.10^6 on yxbl/10°54'/ 11°06'-cut plates. The resonators are hermetically sealed using thermoresistive welding in standard capsules of HC-18U type (dimensions 13.5□11□4.6 mm), filled wit helium (fig.1).

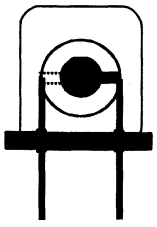


Fig.1 Quartz resonator

The basic problems we solved when testing TQR at cryogenic temperatures were:

- Assembling of proper equipment allowing to obtain temperatures in the temperature interval from 4.2K to 300K;
- Design of quartz generator with TQR outside of the electronic circuit;
- Measurement of the frequency and the temperature of the TQR simultaneously;
- Saving of the measured data for further observation and analysis.

The experimental equipment used is schematically shown on fig.2. It has two versions. One for measurement of temperatures from 77K to 300K where the coolant is liquid nitrogen and an other for temperatures from 4.2K to 300K using liquid helium. The liquid gases are saved in special containers. The temperature at the top of the container is room temperature (300K) and lightly decreases to 77K (4.2K) at the liquid surface. So to put the resonators at different temperatures it is necessary to move them vertically inside the container. For this purpose a special stock is designed. A copper shield thermally isolates its head where two resonators and a calibrated temperature sensor having good thermal contact are installed.

The TQR are connected by means of co-ordinated coaxial cables to special quartz generators (QG) and excited on 3-th overtone at 26.5 MHz. The connection between the QG and the TQR is fine balanced. The output of the QG represents a sinusoidal signal. The instability of the TQR within 10s at constant temperature is lass than $\pm 7.5 \times 10^{-9}$.

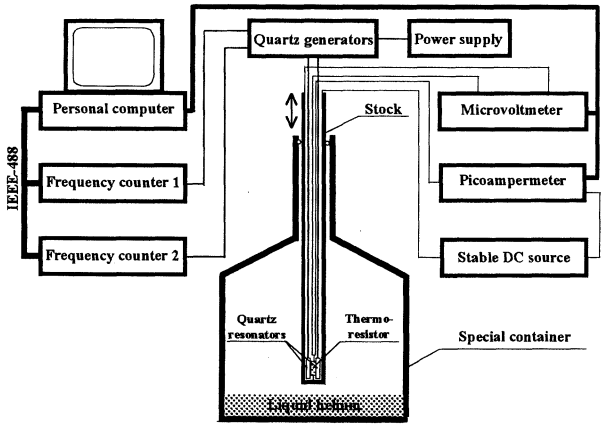


Fig.2 Experimental equipment

The output signals of the QG are measured using high precise frequency counters supplied with IEEE-488 interface. To compute the temperature we use a calibrated thermoresistor which currant and voltage are measured by picoampermeter and microvoltmeter respectively.

The data measured is send to the personal computer (PC) every second. The PC computes the temperature on the base of the calibration data and the measured currant and voltage data, saves the frequency and temperature data in files on a hard disk and shows their values as graphics on the screen. The graphics represent the course of the temperature and the frequency in chosen time interval, 10 minutes usually.

Having this equipment the experiment is carried out as follows. First the stock with the mounted thermoresistor and resonators in its head is placed at the top of the container. In a short time the temperature of the head obtains the temperature of the gas at this place. Then the stock is moved down. Again for some minutes the resonators and the thermoresistor obtain the temperature of the surrounding gas. So with slowly movements and waitings the head of the stock reaches the liquid gas. Then this procedure is repeated by movement in the reverse direction. During this time the temperature and the frequency are measured simultaneously and data is saved on computer's disk.

More then 100 resonators are tested at cryogenic temperatures. Some of them are measured two or three times below to nitrogen and helium temperatures.

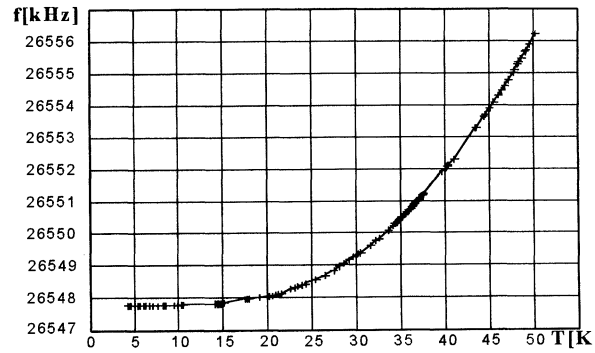
For the analyses of this big amount of experimental data a computer program is designed. Some of its functions are listed below.

1. Display of the experimental data. This function restores the graphics as they are during the experiment and allows observing the pictures in details. So some fine instabilities of the resonators behaviour were found.
2. Creation of temperature-frequency characteristics (TFC). This is the main function of the analyses. TFC can be prepared for different temperature intervals for different resonators.
3. Approximation of the data and the TFC using some mathematical functions. Because of the low sensitivity of the thermoresistors we use, in some temperature intervals the temperature data is measured inaccurate. To smooth the TFC-graphics we approximate temperature data using exponent function or other functions. The approximation of the TFC allows finding the temperature sensitivity of the resonator.
4. Elimination of wrong data. During the experiment casual apparatus errors input wrong frequency or temperature data. When analysing the data using functions 1 and 2 it is possible to remove such kind of errors.
5. Hardcopy of the graphics. This is a very useful function when showing and archiving data. The so prepared pictures can be inserted in text documents.

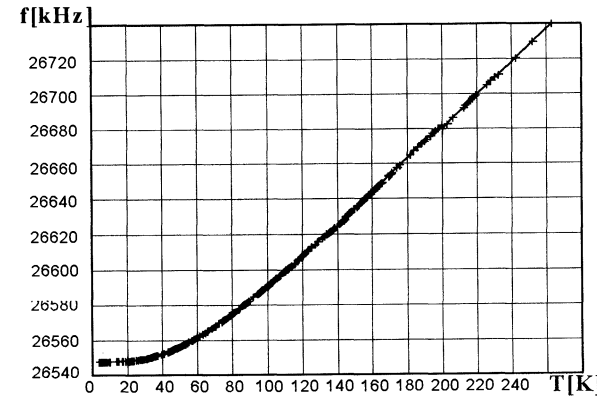
2. Results

The experience of more than two years of experiments and the analyses of the measured data can be generalised as follows.

1. In the temperature interval from 4.2K to 70K the TFC of all resonators are non-linear monotonous



a)



b)

Fig.3. Typical temperature frequency characteristic

increased functions (fig.3a). In this interval the sensitivity increases with the increase of the temperature starting with a value of about 2Hz/K at 4.2K and reaches up to 600Hz/K at 70K. Over 70K the TFC has linear character and the sensitivity reaches about 1000Hz/K. Typical TFC for the temperature interval from 4.2K to 300K is shown on fig.3b. The TFC in all temperature intervals are good approximated with polynomial of 3-th degree. Other measurements and analyses [3] show that the resonators keep their sensitivity and the linearity of their TFC to temperatures of about 500K.

2. For some of the resonators in particular temperature intervals (from 90K to 110K and from 120K to 135K) different disturbances and instabilities were registered:
   - increase of the noise level (fig.4a);
   - fluctuation of the frequency at some temperatures (fig.4b);
   - frequency jumping (fig.4c);
   - stopping of the resonator vibration (fig.4d).

It is interesting to note that in most of the cases of instability when putting the resonator again at the temperature of disturbance the TFC is normal, i.e the disturbance disappears. This is possible when the time between the two consequent measurements is within a few days. When this time is more than a week the instability can be observed again.

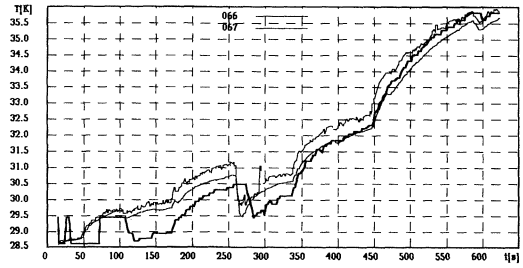


Fig. 4a Noise

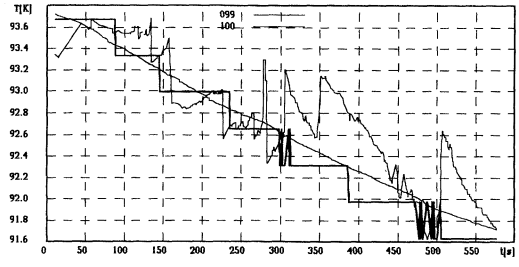


Fig. 4b Fluctuation of the frequency

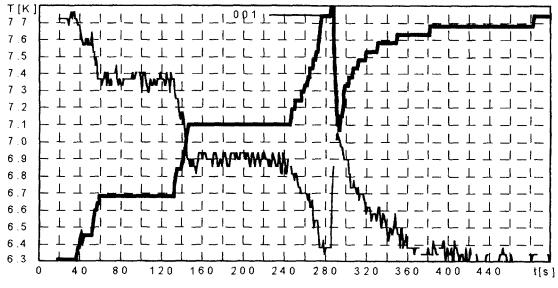


Fig. 5 Reverse TFC

3. Discussion

The most possible reasons of these instabilities are:

1. The influence of the resonator's holders.

Because of the different temperature coefficients of quartz and holder's metal it is possible the presence of tension between the holder and the crystal. The same situation is possible at the interface between the quartz plate and the resonator's electrodes.

2. The excitation of other mode and type of vibrations at some temperatures.

The change of the elastic constants and the dimensions of the quartz plate with the change of the temperature may cause excitation of other mode and type of vibrations. The extension of the temperature interval from 4.2K to 450K can give birth of occurrenceof so called "activity dips" phenomenon in the range of operation. Some modeling and preliminary experiments of such effect were calculated at high temperature (~ 450K) [4], but there is now no study available at cryogenic temperature. Such study is very important for the fundamental understanding of vibrating systems and the technical development of new cryogenic sensors.

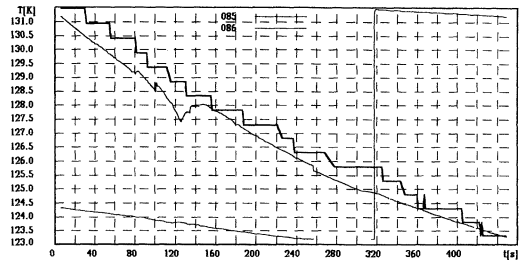


Fig. 4c Frequency jumping

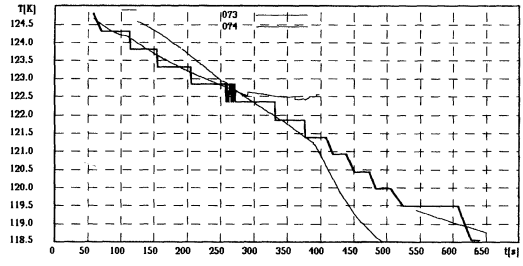


Fig. 4d `Stopping`

An interesting phenomenon is observed at temperatures under 10K (fig.5). Some of the resonators (no more than 7-8%) have reverse TFC at these temperatures (with the increase of the temperature the frequency decreases and vice versa). Such results are observed and described by other authors [5], but up to now no conclusive proof.

4. Conclusions

This type of TQR can be used as temperature sensor for cryogenic temperatures. For temperatures over 20K its sensitivity permits to achieve accuracy better than 0.002K. Before the calibration of the resonators it is necessary to test them for instabilities. The selection of the QG and the precise measurement of the resonator's frequency are very important too.

References

1. Quartz thermometer 2804A, Hewlett Packard catalog

2. E.Benes, M.Grösch, W.Burger, M.Schmid, "Sensors based on piezoelectric resonators", Sensors and Actuators A, 48(1995), pp.1-21.

3. L.Spassov, V.Georgiev, L.Vergov, N.Vladimirova, V.Drobin, N.Agapov, "Thermosensitive quartz resonators at cryogenic temperatures", Sensors and Actuators A, 62(1997), pp.484-487.

4. B.Dulmet, "Current Trends of Finite Elements Method Applied to Piezoelectric Resonators"

5. J.J.Suter, "Piezoelectric properties of quartz crystal resonators below 10K", Cryogenics 1990, Vol.30, June, pp. 547-548

MULTIPORT TECHNIQUE IN TRANSMISSION MEASUREMENTS OF QUATRZ RESONATORS

Aleksander Lisowiec, Marek Wójcicki, Tele and Radio Research Institute, Warsaw, Poland
(work sponsored by grant nr 8T11B02814 from State Committee for Scientific Study)

A. Lisowiec, Ph. D., Eng., M. Wójcicki, Ph. D., Eng., Tele and Radio Research Institute, Piezoelectric Department
Ratuszowa 11, 03-450 Warszawa, Poland; phone/fax: (+48 22)6196460; e-mail: alisow@itr.org.pl

ABSTRACT

In the paper a multi-port technique approach to the piezoelectric resonators basic parameters measurement by transmission method has been presented. The circuit for measurement of S_{21} coefficient of scattering matrix of two port devices (such as quartz resonator connected to test fixture) has been designed in the form of switched five-port. Analogously with classical six-port reflectometer, the switched five-port realizes various superpositions of the waves incident on and transmitted by two-port DUT. Proceeded by suitable calibration procedure, the S_{21} coefficient of DUT can be calculated from measurement of the power of these waves. The switched five-port for S_{21} measurement has been realized in lumped elements technique for frequency range 1 – 500 MHz. The mathematical analysis of measurement accuracy of five-port has been carried out. The parameters of selected types of quartz resonators have been measured with the use of this five-port and the results were compared with those obtained by other measurement techniques.

1. INTRODUCTION

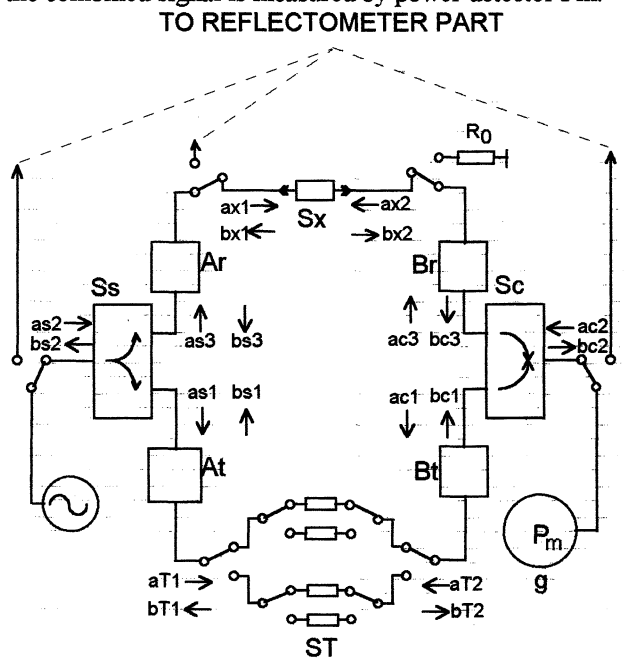
IEC recommends the quartz resonator basic parameters measurement by automatic network analyzer with automatic error correction [3]. The resonator is placed in special test fixture enabling the calibration of measurement path by connecting precision impedance standards in place of the resonator. Quartz resonator is generally a symmetrical device, and does not posses a well defined connection plane (such plane is arbitrarily defined 2 mm from resonator enclosure). In 50 Ohm systems there exists a scatter field around resonator leads that can not be calibrated out. The solution to this problem in the form of special test fixture has been presented during 1999 Joint Meeting EFTF-IEEE IFCS ([6], pages 449-452).

An alternative to heterodyne network analyzer technique for precise impedance measurements is multiport technique [1]. In this technique there is no down-converting of measurement signal to IF frequency but instead various superpositions of waves incident, reflected and transmitted by measured device are arranged and its S-matrix is calculated from measured power of these waves. The six-port reflectometer for 1-500 MHz bandwidth realized in the form of equivalent switched four-port reflectometer in lumped element technique was developed several years age in ITR and presented during 9th EFTF in Besancon ([4], pages 458-461) and 12th EFTF in Warsaw ([5] pages 280-285). Two such reflectometers can be arranged into complete network analyzer. The switched nature of the reflectometers makes the calibration process extremely difficult. At ITR another solution has been proposed. A separate circuit has been developed for S_{21} coefficient measurement and special commutation circuit arrangement makes the manual switching between reflectometer and S_{21}-meter unnecessary.

2. S_{21}-METER CIRCUIT

2.1 Practical realization

The block diagram of S_{21}-meter is presented in fig.1. The measurement signal is divided between one branch consisting of series connection of attenuator Ar, measurement adapter Sx and attenuator Br, and second branch made from series connection of attenuator At, switched transmittance ST and attenuator Bt. The signals from two branches are combined in Sc and the power of the combined signal is measured by power detector Pm.



$$Ss = \begin{bmatrix} G_s & C_s & I_s \\ C_s & G_s & C_s \\ I_s & C_s & G_s \end{bmatrix} \quad Sc = \begin{bmatrix} G_c & C_c & I_c \\ C_c & G_c & C_c \\ I_c & C_c & G_c \end{bmatrix}$$

$$Sx = \begin{bmatrix} Sx_{11} & Sx_{12} \\ Sx_{21} & Sx_{22} \end{bmatrix} \quad ST = \begin{bmatrix} ST_{11} & ST_{12} \\ ST_{21} & ST_{22} \end{bmatrix}$$

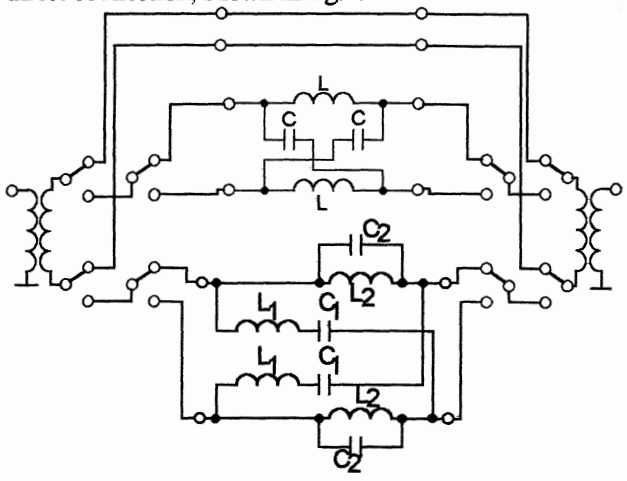
$$Ar = \begin{bmatrix} 0 & ar \\ ar & 0 \end{bmatrix} \quad At = \begin{bmatrix} 0 & at \\ at & 0 \end{bmatrix} \quad Br = \begin{bmatrix} 0 & br \\ br & 0 \end{bmatrix} \quad Bt = \begin{bmatrix} 0 & bt \\ bt & 0 \end{bmatrix}$$

Fig.1. Block diagram of S_{21} meter and S-matrices of its elements.

Assuming infinite isolation and perfect match of splitter Ss and combiner Sc (i.e. $I_s=I_c=G_s=G_c=0$), the amplitude of wave incident on power detector can be written as:

$$bc2 = as2 \cdot C_c \cdot C_s \cdot (at \cdot bt \cdot ST_{21} + ar \cdot br \cdot Sx_{21}) \qquad (1)$$

Equation (1) has the form analogous to equation describing six-port reflectometer [1] with S_{21} substituting Γ and ST_i substituting q-points. The same criteria for choosing ST_i [1] and the same calibration procedure can thus be applied to S_{21}-meter as to classical six-port reflectometer. The transmittances have been realized as all-pass ladder networks of first-order, second order and direct connection, shown in fig.2.



L=0.8 μH, C=300 pF, L_1=0.16 μH, C_1=1.6 nF, L_2=4 μH, C_2=64 pF

Fig.2. Circuit realization of switched transmittances.

The ladder form has been chosen because if offers greater flexibility in phase characteristics and doesn't need mutual transformers for practical realization (however balun transformers at the input and output are needed). Fig.3 shows the phase difference between the three all-pass networks. In the frequency range 4÷150 MHz the minimum phase difference is 90 degrees.

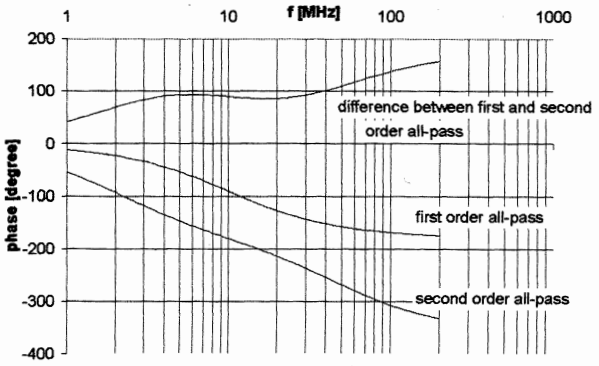


Fig.3. Phase difference between switched transmittances of fig.2.

When the assumption of infinite isolation and perfect match of splitter Ss and combiner Sc are dropped, the expression describing the amplitude of the wave incident on power detector is much more complicated than equation (1) and has the form

$$bc2 = \frac{a1 + b1 \cdot Sx_{21} + c1 \cdot Sx_{21}^2}{a2 + b2 \cdot Sx_{21} + c2 \cdot Sx_{21}^2}, \qquad (2)$$

with non-zero coefficients c1 and c2 of Sx_{21}^2 in numerator and denominator. As such it is not quite analogous to the equation describing the amplitude of waves incident on power detector in classical six-port reflectometer. However, the coefficients c1 and c2 are each sums of products of Is, Ic, Gs, Gc, at, bt, ar, br and are "small" compared to b1 and b2. The appearance of $c1Sx_{21}^2$ and $c2Sx_{21}^2$ terms can be treated as error in the power measurement. The values of Is, Ic, Gs, Gc have been measured and substituted (together with at=bt=ar=br=0.16) into equation (2). The coefficients c1 and c2 have been calculated as sums of absolute values (the relative phase of Is, Ic, Gs, Gc is not known). The upper bound on relative difference of the power of waves incident on power detector calculated by equation (1) and (2) has been estimated as 0.5%. This difference can be made arbitrarily small by increasing the attenuation coefficients of At and Bt but this is limited by power detector sensitivity and noise.

2.2 Calibration and accuracy analysis

The S_{21}-meter has been calibrated in the 1 – 150 MHz frequency range by the Engen method [2]. The circuit realizing various transmittances connected in place of Sx has been realized as series connection of switched attenuator and all-pass network identical to the one presented in fig.2. The calculated calibration constants were then used for accuracy analysis in accordance with the algorithm presented in flow chart form in fig.4.

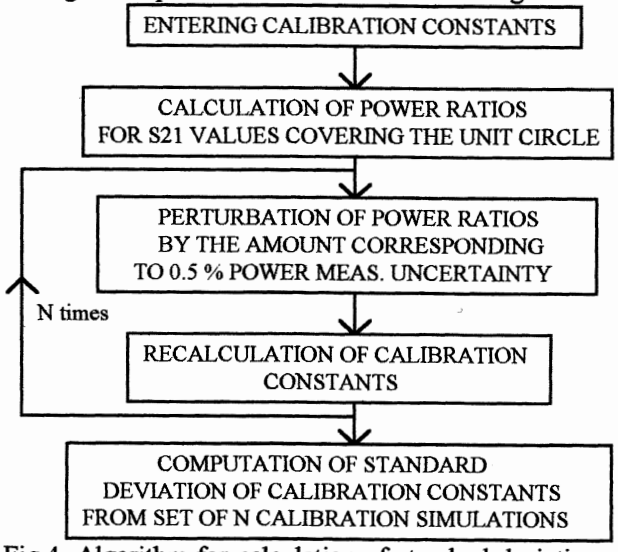


Fig.4. Algorithm for calculation of standard deviation of calibration constants.

The uncertainties of the real and imaginary part of S_{21} were calculated from equation (3) and (4)

$$\sigma Re(S_{21}) = \sqrt{\sum_{i=1}^{11}\left(\frac{\partial ReS_{21}}{\partial ki}\sigma ki\right)^2 + \ldots + \sum_{i=1}^{3}\left(\frac{\partial ReS_{21}}{\partial p_i}\sigma p_i\right)^2} \quad (3)$$

$$\sigma Im(S_{21}) = \sqrt{\sum_{i=1}^{11}\left(\frac{\partial ImS_{21}}{\partial ki}\sigma ki\right)^2 + \ldots + \sum_{i=1}^{3}\left(\frac{\partial ImS_{21}}{\partial p_i}\sigma p_i\right)^2} \quad (4)$$

where σki is the uncertainty (standard deviation) of calibration constant ki and ReS_{21}, ImS_{21} are real and imaginary parts of S_{21} as functions of calibration constants ki and power ratios p_i.

The accuracy analysis that has been carried out takes into account only the error caused by non-perfect nature of splitter Ss and combiner Sc. However, very good linearity and excellent resolution of the power meter [4,5] make the above mentioned error dominant.

The calculated measurement uncertainties (standard deviation) of the S_{21} absolute value and phase are presented in fig.5 and. fig.6.

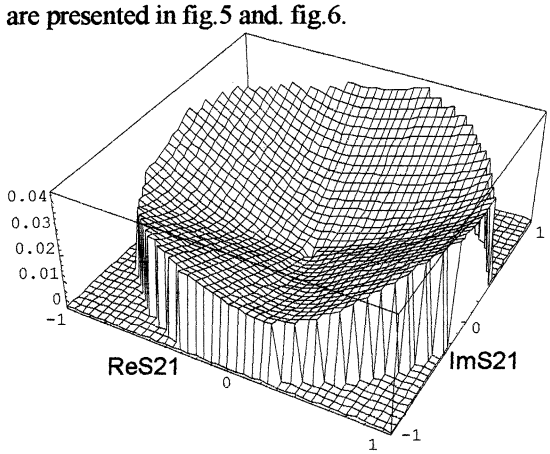


Fig.5. Uncertainty (standard deviation) of S_{21} absolute value.

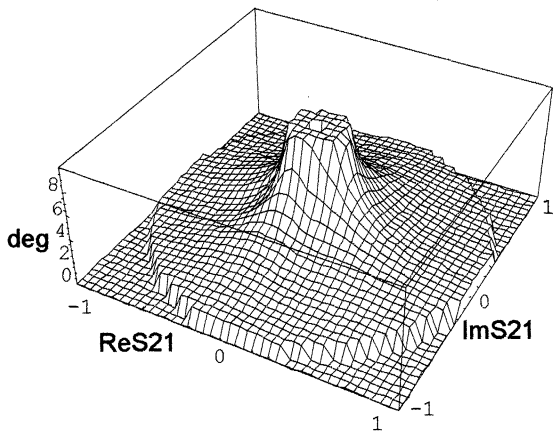


Fig.6. Uncertainty (standard deviation) of S_{21} phase

3. MEASUREMENT RESULTS

3.1 S_{21} measurement results

The S_{21} coefficient of measurement adapter with selected resonators has been measured by the switched multi-port S_{21} meter presented above and by HP8753D network analyzer. The results are presented in fig 7 to fig.9.

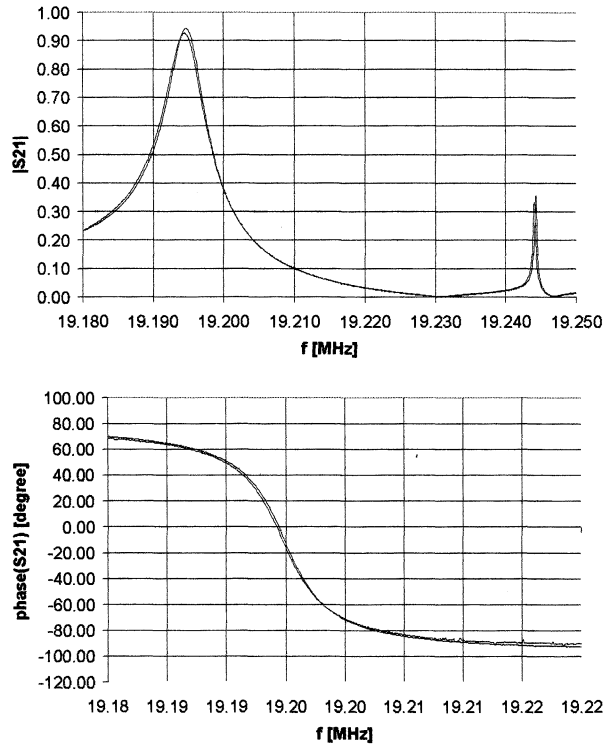


Fig.7. Magnitude and phase of S_{21} coefficient of measurement adapter with 19.193 MHz rezonator

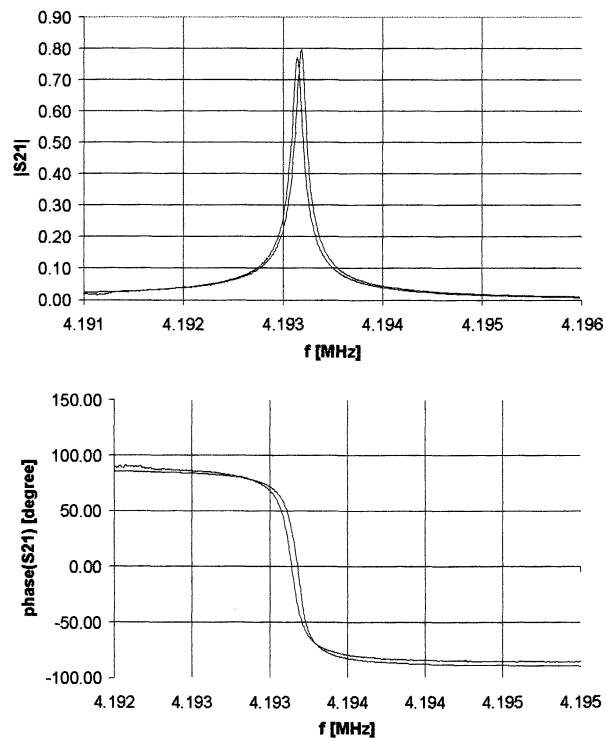


Fig.8. Magnitude and phase of S_{21} coefficient of measurement adapter with 4.193 MHz rezonator

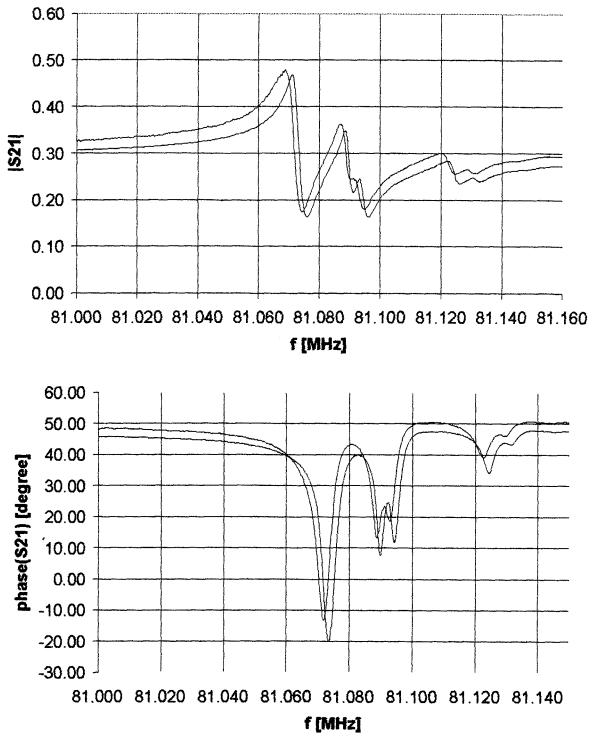


Fig.9. Magnitude and phase of S_{21} coefficient of measurement adapter with 81.07 MHz rezonator

3.2Resonator parameters preliminary measurement results

The complete circuit for S_{11} and S_{21} coefficients measurement is presented in fig. 10.

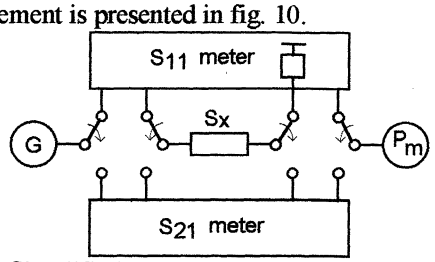


Fig. 10. Simplified diagram of multiport network analyzer

In the case of quartz resonator measurement the problem is to deembed the resonator scattering matrix from measurement adapter scattering matrix. The necessary equations have been presented at the Joint Meeting EFTF-IEEE IFCS ([6], pages 449-452). It has been assumed that the resonator is a symmetrical device with four-element electrical equivalent circuit. The resonator admittance Y has been calculated from:

$$R_0 Y = \frac{2S_{12}}{(1+S_{11})(1+S_{22}) - S_{21}S_{12}}, \qquad (5)$$

and the circle fitting method [3] has been used to extract the equivalent circuit parameters.

The parameters of three resonators of nominal frequencies 4.193 MHz, 19.193 MHz and 81.077 MHz have been measured with the use of specially designed measurement adapter [6] by the multiport circuit from fig. 10 and the HP8753D network analyzer. The results are presented in table 1.

Table 1.

Parameter	Circle fitting, multiport	Circle fitting, HP8753D
fs [MHz]	4.193156	4.193190
	19.194144	19.194800
	81.069934	81.069800
R1 [Ohm]	28	25
	6	6.9
	169	200
L1 [mH]	200	199
	3	3.04
	8	
C1 [fF]	7.18	7.2
	26.9	20
	0.48	
Q	195000	205000
	51947	60000
	20000	
C0 [pF]	3	3
	3.1	3.8
	6.8	6.1

Difference in fs resonant frequency has been caused by frequency shift of signal sources used by multiport network analyzer and HP8753D.

4. CONCLUSION

The main advantage of the presented multiport network analyzer is its simplicity. Although it has been calibrated in the frequency range 1÷200 MHz, increasing the number of switched transmittances STi could broaden its bandwidth. The intrinsic error of the S_{21} measurement method is of the order comparable to the S_{11} measurement error caused by power detector nonlinearity and noise in well designed multiport reflectometers.

The quartz resonator measurement results presented above should be compared with those obtained in APC-3.5 (or APC-7) test fixture.

5. REFERENCES

[1] G. F. Engen, "The six-port reflectometer: An alternative network analyzer", IEEE Trans. Microwave Theory and Tech., 1977

[2] G. F. Engen "Calibrating the six-port reflectometer by means of sliding terminations", IEEE Trans. Microwave Theory and Tech., MTT-26, 1978

[3] IEC Publication 444-5, 1994

[4] A. Lisowiec, M.Wójcicki, "Six-port technique in reference measurements of piezoelectric resonator parameters", Proc. of 9th EFTF Besancon, 1995

[5] A. Lisowiec, M. Wójcicki, " Aspects of calibration of a six-port circuit for precise measurements of piezoelectric resonator parameters", Proc. of 12th EFTF Warsaw, 1998

[6] A. Lisowiec, M. Wójcicki, "A new approach to the problem of transmission measurements of quartz resonators", Proc. of IEEE IFC and 13th EFTF Joint Meeting, Becanson, 1999

THEORETICAL AND EXPERIMENTAL INVESTIGATION OF SPECIFIC COMPONENTS IN OUTPUT SIGNAL OF DIGITALLY COMPENSATED CRYSTAL OSCILLATORS

Anatoly V. Kosykh, Alexander N. Lepetaev, Andrew A. Roy
Omsk State Technical University, Russia

ABSTRACT

For ensuring of high temperature stability of frequency under tight restrictions on dimensions, low power consumption and setting time of output frequency with assigned precision the thermocompensated crystal oscillators (TCXOs) are widely used. The best stability is achieved with the use of digital thermocompensated oscillators (DTCXOs). However, the use of digital methods worsens the spectrum of output oscillations.
In the present work the problems of influence of digital blocks on output spectrum of DTCXO with frequency-type thermosensors are considered. The results of theoretical and experimental researches are compared. The recommendations for account of parameters of compensation circuit from the point of view of output spectrum improvement are given.

INTRODUCTION

The requirements to frequency stability of reference oscillation sources are constantly increased. Simultaneously with this the requirements to their miniaturization are increased too. When it is necessary to ensure frequency stability within the limits of $3\cdot10^{-6}...3\cdot10^{-8}$, and the dimensions and power consumption should be minimum, TCXOs are used. The widely used oscillators with analog thermocompensation have low power consumption and small dimensions, but their frequency stability in the temperature range is usually not better than $5\cdot10^{-7}$ ... $1\cdot10^{-6}$. For obtaining of higher stability the crystal oscillators with digital thermocompensation (DTCXOs) are used. In [1, 2, 3] the oscillators with frequency stability of $3\cdot10^{-8}$ in the temperature range (-40 ... + 85) °C are described. However, except the temperature stability of frequency the main parameter for any oscillator is the cleanness of output signal spectrum. The digital shaping of compensatory effect inevitably results in the appearance of specific component in the output oscillation spectrum, however this phenomenon is investigated unsufficiently and there are no recommendations for spectrum improvement. The purpose of the present work is to fill this lack by theoretical and experimental researches of digital thermoconpensation process.

1. THE MAIN INFORMATION ABOUT THERMOCOMPENSATION PROCESS

In DTCXOs with the best stability (including those described in [1-3]) frequency-type thermosensors are used, namely thernosensitive crystal resonators or thermosensitive modes of dual-mode resonators. The block diagram of compensatory effect shaping in such oscillators is given in Fig. 1.
The compensatory effect can be both of analog, and digital form. In the first case the control code through the digital-to-analog converter (DAC) passes to the controlled oscillator, but in the second case this code passes to the frequency synthesizer.

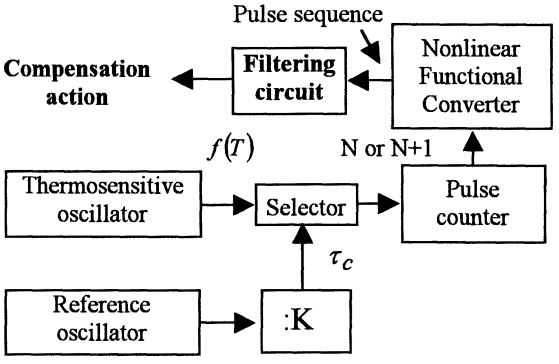


Fig. 1. Structure of compensation action forming

Let's consider the important ratio between the required frequency stability, steepness of thermosensor curve, time of impulse account and the pulsing amplitude of compensatory effect.
If thermosensor steepness is equal to S_{sensor}, operating temperature range is ΔT, and the word length of the counter is n, for definition of unambiguous measurement the inequality should be executed

$$\tau_c \le 2^n / S_{sensor}\,\Delta T \qquad (1)$$

Here τ_c - time of impulse account by the counter which influences greatly on output signal spectrum.

The word length of the counter n influences on digitization of temperature measurement and, corollary, on a potentially attainable accuracy of thermocompensation. In the process of step approximation of compensatory function the component of digitization error of thermosensor signal is determined as

$$\delta f_1 = \alpha_f \cdot \Delta T / 2^n \qquad (2)$$

where α_f is steepness of frequency vs. temperature curve of reference resonator in the given temperature range. For AT- cut the maximum value α_f in operating temperature range has the value of the order $1\cdot10^{-6}/°C$. For piecewise linear approximation of compensatory function thermocompensation error at the expense of digitization of thermosensor signal is determined as [4]

$$\delta f = \frac{\Delta T}{2^{m+2}} \cdot a_2(T) \qquad (3)$$

Where a_2 is factor of the second degree of frequency vs. temperature curve expansion of reference resonator in the set in the given temperature range, T – current temperature.

The expressions (2) and (3) allow to select the necessary value of n, ensuring allowable thermocompensation error.

The compensatory effect has a discrete nature, as its value can be changed only at the moments of termination of impulse account by the counter of impulses from the frequency-type thermosensor, and the argument of function of nonlinear transformation can have only discrete values. The discrete nature of compensatory effect results in the extension of its spectrum, that leads to deterioration of output signal spectrum dependent on the modulating action spectrum and modulation index. Proceeding from this, we shall conduct a research of spectrum of compensatory action.

The average amount of impulses with period of $\tau(T)$, which appeared in the counter from the thermosensor during τ_c, is determined as $\overline{N} = \tau_c / \tau(T)$. Generally, this ratio has fractional values, while the counter can record only integer amount of impulses (fronts). It results in the fact that under noncorrelatednesses of being accounted fronts of impulses and account interval either N (where $N = ent\{\overline{N}\}$), or N+1 impulses can be fixed in the counter. The probability P ($P \in [0,1]$) of N+1 impulses realization is determined by a fractional part of the ratio $\tau_c / \tau(T)$. During long time of observation the equality is executed

$$\overline{N} = N(P-1) + (N+1) \cdot P$$

As to each value of the counter (argument of compensatory effect) the own value of compensatory codes is corresponds, the average value of compensatory effect can be determined as

$$\overline{K} = K_N \cdot (P-1) + K_{N+1} \cdot P , \qquad (4)$$

where K_N and K_{N+1} - values of compensatory codes corresponding to the conditions of the counter N and N+1 accordingly.

The expression (4) corresponds to a piecewise linear interpolation of compensatory function according to Lagrange, that allows on the basis of averaging of compensatory sequence to realize the elementary interpolator [4] and to use the expression (3) for determination of word length of the counter.

Under the certain temperatures (on the boundaries of temperature ranges of the width $\Delta\Theta$) the ratio $\tau_c / \tau(T)$ accepts the integer values, and the compensatory effect is constant in time. Under the other temperature values of the environment the compensatory effect will be changed in steps, that results in the appearance of specific components in the spectrum.

2. COMPUTATION OF COMPENSATORY EFFECT SPECTRUM

At the output of nonlinear functional converter the impulse sequence generally is acyclic with wide spectrum. This sequence can be considered as sampling of periodic signal of unit amplitude with period T and on-off time ratio 1/P, which is made with period t. Let's present the sample and hold as the sum of products of shifted in time each from other on the magnitude d signals shown on Fig. 2, and $\delta \to 0$.

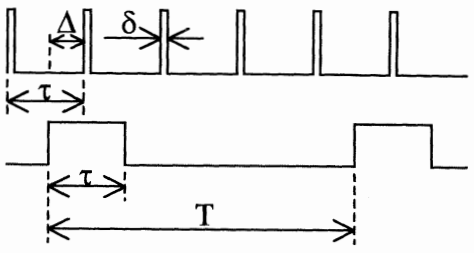


Fig. 2

As both signals in this Figure are periodic with period T, we shall take the advantage of the formula for presentation of periodic signals [1]:

$$L[f(t)] = \frac{1}{1-e^{-S\cdot T}} \cdot \int_0^T f(t) \cdot e^{-S\cdot t} dt$$

For logic signals of single amplitude and duration ? this expression will be transformed to a kind:

$$L[f(t)] = \frac{1-e^{-S\cdot\eta}}{(1-e^{-S\cdot T})\cdot S} \qquad (4)$$

Let's consider that the zero instant of time corresponds to the beginning of the second signal, and the first signal is shifted concerning the second one on the magnitude D (initial shift of signals). Then, according to (4), the Laplace transform of the upper signal in Fig.3 will be equal:

$$L1(S) = \frac{(1-e^{-S\cdot\delta})\cdot e^{-S\cdot\Delta}}{S\cdot(1-e^{-S\cdot\tau})} \qquad (5)$$

For the lower signal in the same Figure the Laplace transform is equal:

$$L2(S) = \frac{1-e^{-S\cdot\tau}}{S\cdot(1-e^{-S\cdot T})} = \frac{1-e^{-S\cdot\tau}}{S\cdot(1-e^{-S\cdot\tau/P})} \qquad (6)$$

The image of product of both signals is expressed by the formula:

$$M(S) = \frac{1}{i\cdot 2\pi} \cdot \int_{\lambda-i\infty}^{\lambda+i\infty} L1(Z)\cdot L2(S-Z)dZ \qquad (7)$$

The sum of shifted products (7) can be calculated on the basis of Laplace transformations and formula for the sum of geometrical progression:

$$W(S) = \sum_{k=0}^{\tau/\delta-1} M(S) e^{-k\cdot\delta\cdot S} = M(S)\frac{1-e^{S\cdot(\delta-\tau)}}{1-e^{-\delta\cdot S}} \qquad (8)$$

Substituting in (8) expressions (5..7) and passing further to the limit of $\delta \to 0$, we get the following formula for the signal at the output of thermocompensatoin system U (S) (9):

$$U(S)=\lim_{\delta\to 0}\left[\frac{1-e^{S\cdot(\delta-\tau)}}{1-e^{-\delta\cdot S}}\cdot\frac{1}{i\cdot 2\pi}\cdot\rightarrow\right.$$

$$\left.\rightarrow\int_{\lambda-i\infty}^{\lambda+i\infty}\frac{(1-e^{-Z\cdot\delta})\cdot e^{-Z\cdot\Delta}}{Z\cdot(1-e^{-Z\cdot\tau})}\cdot\frac{1-e^{-(S-Z)\cdot\tau}}{(S-Z)\cdot(1-e^{-(S-Z)\cdot T})}dZ\right]=$$

$$=\frac{1-e^{-S\cdot\tau}}{S\cdot i\cdot 2\pi}\cdot\int_{\lambda-i\infty}^{\lambda+i\infty}\frac{(1-e^{-(S-Z)\cdot\tau})\cdot e^{-Z\cdot\Delta}}{(1-e^{-Z\cdot\tau})\cdot(S-Z)\cdot(1-e^{-(S-Z)\cdot T})}dZ$$

For determination of signal spectrum we shall consider that the variable S in this formula accepts the values lying on a conjugate axis of a complex plane. Then the parameter l in the formulas (7) and (9) should be more than zero.

The integrand in expression (9) has, at least, a double zero when Z =-1, therefore for calculation of an integral (9) it is possible to close a path of integration by an arc of a circle of an infinite radius lying in the left half-plane. But then this integral can be calculated through residues of an integrand (10):

$$U(S)=\frac{1-e^{-S\cdot\tau}}{S}\cdot\sum_{Z_i}\mathrm{Res}\left[\frac{(1-e^{-(S-Z)\tau})e^{-Z\cdot\Delta}}{(1-e^{-Z\cdot\tau})(S-Z)(1-e^{-(S-Z)T})}\right]\quad(10)$$

Using the rules of calculation of residues and taking into account, that the pole of the second order when Z=S is removable, the expression (10) after some transformations can be presented as follows (11):

$$U(S)=\frac{P\cdot e^{-S\cdot\Delta}}{S}+\frac{1-e^{-S\cdot\tau}}{S}\cdot\rightarrow$$

$$\rightarrow\cdot\left[\sum_{n=-\infty}^{\infty}\left[\frac{(1-e^{-S\cdot\tau})\cdot e^{-j2\pi\cdot n\cdot\Delta/\tau}}{(S\cdot\tau-j2\pi\cdot n)\cdot(1-e^{-(S\cdot\tau-j2\pi\cdot n)/P})}\right]+\rightarrow\right.\quad(11)$$

$$\left.\rightarrow+\sum_{n=-\infty,n\neq 0}^{\infty}\left[\frac{(1-e^{-j2\pi\cdot n\cdot P})\cdot e^{-(S-j2\pi\cdot n/T)\cdot\Delta}}{(1-e^{-(S\cdot\tau-j2\pi\cdot n\cdot P)})\cdot j2\pi\cdot n}\right]\right]$$

Transferring from Laplace transform to the components of signal spectrum in the expression (11) it is necessary to substitute S by jω, and to multiply the expression by 2 for those components of spectrum, where ω≠0. However in practical accounts it is not convenient to use only imaginary values of S because when the values of angular frequencies correspond to the radicals of denominators, in the expression (11) the indeterminacy of an kind 0/0 are received which are necessary to be uncovered. Besides it is necessary to take into account, that during spectrum measurement by an actual device having nonzero band of the equivalent filter, the measured spectrum can considerably differ in its appearance from the theoretical one because the actual device determines average weighted value of signal spectral components in a filter passband. To make the calculation of signal spectrum like the kind that is received in the actual measuring instrument, it is necessary to refuse the use of only imaginary values of S in the expression (11). Let's consider that S =ρ+jω. The parameter r corresponds to a half of the width of an oscillatory outline passband: $\rho = 0.5\cdot\omega_0/Q$

Taking into consideration all mentioned above the expression for amplitudes of spectral components we shall present as follows:

$$U(\omega_i)=\left|(1-e^{-\tau\cdot S_i})\cdot\rightarrow\right.$$

$$\rightarrow\cdot\left[\sum_{n=-M}^{M}\left[\frac{(1-e^{-\tau\cdot S_i})\cdot e^{-j2\pi\cdot n\cdot\psi}}{(\tau\cdot S_i-j2\pi\cdot n)\cdot(1-e^{-(\tau\cdot S_i-j2\pi\cdot n)/P})}\right]+\rightarrow\right.$$

$$\rightarrow+\sum_{n=-M,n\neq 0}^{M}\left[\frac{(1-e^{-j2\pi\cdot n\cdot P})\cdot e^{-(\tau\cdot S_i-j2\pi\cdot n\cdot P)\cdot\psi}}{(1-e^{-(\tau\cdot S_i-j2\pi\cdot n\cdot P)})\cdot j2\pi\cdot n}\right]\left.\right]+\rightarrow$$

$$\left.\rightarrow+P\cdot e^{-\psi\cdot\tau\cdot S_i}\right]\cdot\frac{\rho_i}{S_i}\quad(12)$$

Here ω_i - the frequency set, in which the spectrum calculation is made, $S_i=\rho_i j\omega_i$, $\psi=\Delta/\tau$ - normalized shift of the first signal with reference to the second one ($\psi\in]0,1[$), M - limits of summation.

During spectrum calculation in narrow frequency band it is possible to use constant value for magnitude ρ, that will correspond to a constant width of an equivalent oscillatory outline passband. If the calculation need to be made in broad frequency band, it is necessary to use a criterion of persistence of a quality-factor. In this case magnitude ρ will depend on current value of frequency ω. If it is necessary to make calculation of a spectrum in N frequency points uniformly placed in logarithmic scale in the range from ω_l , (lower frequency of the range) up to ω_h , (upper frequency of the range), for choice of a parameter ρ it is possible to use the following inequality: $\rho_i \geq \omega_i\cdot(m-1)/(m+1)$ where $\omega_{i+1}=\omega_i\cdot m$, $m=\sqrt[N]{\omega_h/\omega_l}$.

In expression (12), in difference from (11), the calculation of the sums in finite limits is used.

The spectrum computation (15) for special cases (with periodic sequence of impulses) have shown that if the values of a limit M is of the order 60...100 the calculation errors should be less than 1 dB. The value of parameter y (normalized delay), as expected practically does not influence the results of spectrum account. However it is impossible to select the value P equal 0 or 1. It is connected with the fact that impulse signals (fig. 3) when P = 1/k2 are multiplied in the field of fronts, where their approximate representation has the greatest error. If to select values ψ in a range of values from 0.1 up to 0.9, the results practically do not differ from each other. It is necessary, however to monitor, that the magnitude of expression (1/P-ψ) does not accept the integer values, because in this case multiplication of signals in the field of fronts will also happen. It is possible to recommend making several accounts with different values of the normalized delay y to be convinced, that the identical results are received.

3.COMPUTATION OF OUTPUT OSCILLATION SPECTRUM

At small modulation indexes the spectrum components of output signal of controlled oscillator are directly proportional to magnitude of a modulation index (with a constant of proportionality of 0.5). Therefore for the final computation of an output spectrum at the supposition of a small modulation index the formula (13) may be used:

$$U_{out}(\omega_i) = U(\omega_i) \cdot K_F(\omega_i) \cdot \frac{\pi \cdot F_d}{\omega_i}, \qquad (13)$$

Where ω_i - value of angular frequencies, in which one is calculated the components of a output signal spectrum, $U(\omega_i)$ - spectral components of compensation voltage , calculated by the formula (12) , F_d - frequency of a deviation (in experiment it was equal to 30 Hz, that is typical for statistical piecewise linear interpolation [4]), K_F - transmission factor of the filter (in this case $K_F = 1 / (1+j\omega \cdot 0.00075)$).

In Fig.3 and Fig.4 the modulation indexes under various values of probability P (various relative position inside interpolation interval) are indicated.

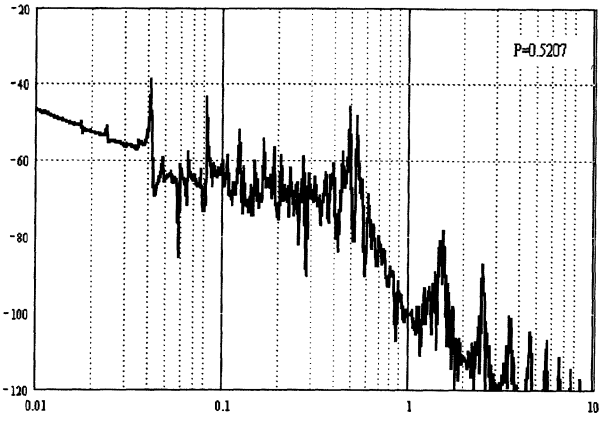


Fig.3. The modulation indexes when p=0.5207

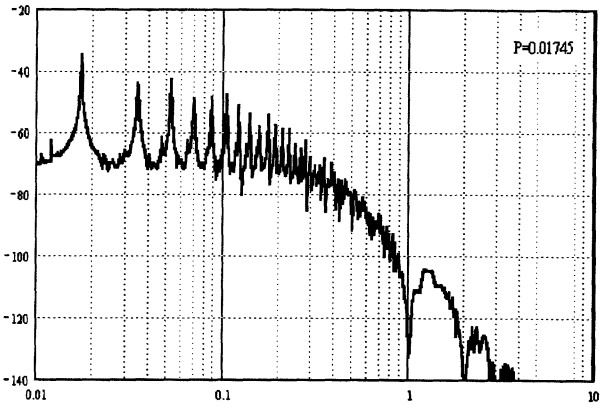


Fig.4. The modulation index when p=0.0017

Hereinafter this signal is subjected to filtration, and the amplitudes of maximum components essentially decrease. While selecting the type of filter it is necessary to remember that the increase of its time constant results in the appearance of dynamic thermocompensation error, and this effect should be taken into account.

You can see on Fig.3 and Fig.4, that the form of a spectrum essentially depends on parameter ?, which can change in the range of temperature. Therefore, the spectrum of compensatory effect in the range of temperatures will be nonconstant. The user usually is not interested in instantaneous spectrum value, but the limits in which the spectrum is under external actions, are of great interest for him. We had calculated the envelope of possible realizations of compensatory effect spectrum with modification of a parameter ? from 0 up to 1. The plot the envelope of spectrum is indicated in Fig5. The form of this plot can be approximated by the expression (14):

$$M(f) = \alpha_f \cdot \Delta T \cdot \tau_c \cdot K_F(f) \cdot \left[\frac{\sin(\pi \cdot f \cdot \tau_c)}{\pi \cdot f \cdot \tau_c}\right]^2$$

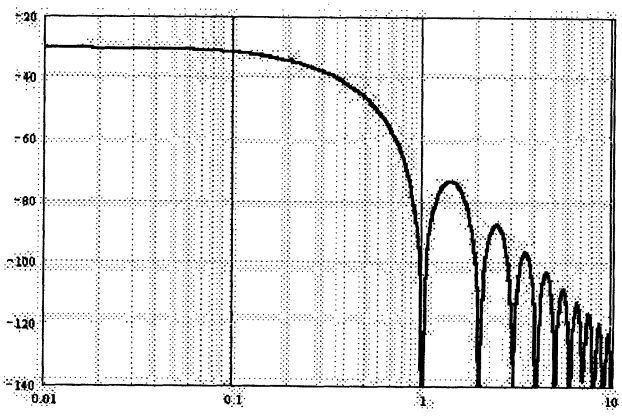


Fig.5 The envelope of compensatory voltage spectrum

From this expression it is visible, that close to the carrier the amplitude of spectral components is determined by the ratio $\alpha_f \cdot \Delta\Theta \cdot \tau_c = \alpha_f / S_{sensor}$

Obtained in (13) the expression allows to calculate the spectrum of compensatory action (which has pulsing character) with concrete parameters of thermocompensation system. Its form will be influenced to a great extent by thermosensor steepness S, range of operating temperatures ΔT, word length of the counter of impulses n and required stability δf.

The pulsing control signal will cause frequency modulation of the output signal. And, each of spectral components of control signal can cause the emergence of infinite number of components in the output spectrum (depending on modulation index). If modulation index is small, the form of FM oscillation spectrum coincides with the form of AM-oscillation with the precision of a factor. In Fig.6 is the plot of counted values of modulation index for actual values of

range of reorganization of the controlled oscillator and the spectrum of control action. It is visible, that the value of modulation index with all values of modulating frequencies is less than 1. Therefore, the expression (14) can be used for calculation of an output spectrum.

In Fig.7 and Fig.8 the spectrum obtained experimentally (in the unit which imitates the work of actual oscillator) are indicated. The measurements were made by spectrum analyzer EUROPTEST PN 9000, which has a threshold of its own noise of -160 dB.

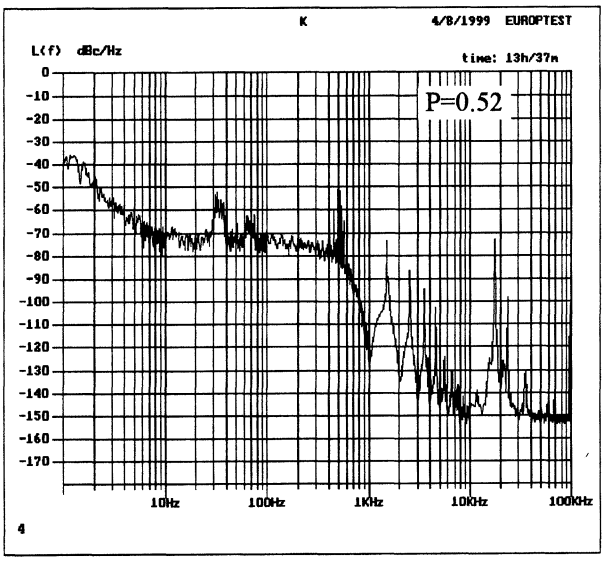


Fig.7. Experimental spectrum of imitative DTCXO

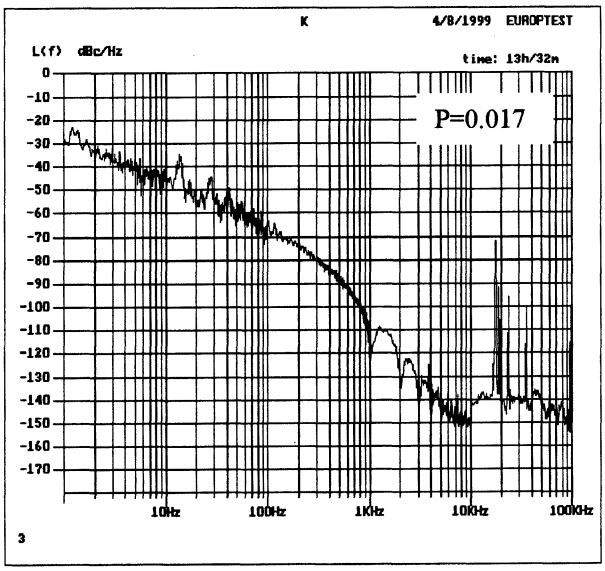


Fig.8. Experimental spectrum of imitative DTCXO.

From the comparison of fig. 7,8 and fig.3,4 it is visible that coincidence of experimental and counted values is rather close. Some differences are explained by influence of the controlled oscillator that has own noise (in the simulating unit the oscillator of a low class w used) on frequencies close to the carrier, a interference of the display's frequencies. It means, th the expressions (15) (16) can be really used f simulation of output signal spectrum of any DTCX using the compensatory circuit of Fig. 1.

CONCLUSION

On the basis of the produced researches it is possible make the following conclusions:

1. The spectrum of DTCXO output signal, having t structural diagram of Fig. 1, contains spurio components. The instantaneous shape of a spectru depends on ambient temperature. Enveloping of possible values of these spectrums is featured expression (14).
2. Near to carrier the amplitude of spurious compone of a spectrum is defined by the ration of a F-T cur slope of a reference resonator and of F-T curve slope a temperature sensor. Therefore, it is desirable to u temperature sensors with the greatest value of a slop From this point of view the usage as a temperatu sensor of B -mode a SC-cut [3,4] is more preferenti than usage of a difference of 1-st and 3-rd harmonics a C - mode C [1,2] .
3. The obtained formulas model a spectrum of DTCX output enough precisely, that allows to use them f optimization of DTCXO structure.

REFERENCES

[1]. M. Bloch, M. Meirs, J. Ho. The Microcomputer- compensated Crystal Oscillator. (MCXO). Proc. 43-r A.F.C.S., 1989, pp. 16 - 19.

[2]. Benjaminson A., Stalling. S. C. A Microcompute compensated Crystal Oscillator Using a Dual-Mode Resonator. Proc. 43-rd A.F.C.S., 1989, pp. 20 – 26.

[3]. A. Kosykh, I. Abramson, V. Bagaev. Dual-mode Crystal Oscillators With Resonators Excited on A an C Modes. Proc. 48-th A.F.C.S., 1994, pp. 578 – 586.

[4]. A. Kosykh, V. Bagaev, B. Ionov, A. Lepetaev, Zavialov. The new method of statistic piecewise-line interpolation and its application to DTCXO creation. Proc. 47-th A.F.C.S. pp. 687 – 697.

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

LOAD-RESONANCE MEASUREMENT ACCURACY OF CERAMIC SURFACE MOUNT QUARTZ CRYSTALS

M. van Herwijnen, F.K.C Sohre

SaRonix B.V., P.O.Box 70, 7000 AB Doetinchem, Netherlands.
T: +31 314 320 648; F: +31 314 320 636; E: mvherwijnen @ saronix.nl

ABSTRACT

Surface-mount quartz crystals and oscillators face a growing demand in compact electronic devices, extending into higher-accuracy applications. This calls for updating of the measurement techniques. Recent IEC proposals [4-5] adapt IEC standard 60444 [1-3] to include SMD quartz crystals. The present paper proposes a modified measurement set-up, derived from IEC-60444 but especially developed for ceramic surface-mount quartz crystals with a metal lid.
Important in this set-up is the grounding of the lid and the opportunity to apply a twin load capacitance. Metal-type crystals like e.g. HC49 are nearly symmetric regarding their connections, and they are frequently applied without connection to ground. Ceramic crystals with a metal lid (like e.g. NKS7) are often applied with the metal lid connected to ground, in order to minimize electromagnetic radiation. This makes the ceramic crystals intrinsically asymmetric.

In the regular measuring set-up, interchanging the crystal connections as well as using the set-up with or without grounding of the crystal lid yields significantly different measurement results for the static capacitance C_0 and the load-resonance frequency. These are caused by stray and holder capacitances. The value determined for C_0 becomes smaller by grounding the lid. As a consequence, the pulling range and pulling sensitivity change also. This can have a bearing to the behaviour of a grounded ceramic crystal applied in a VCXO-type oscillator.
This paper quantifies and explains the asymmetric behaviour for measurements with load capacitance and the influence on load-resonance frequency determination. With a one-sided load capacitance, measuring with grounded lid yields a lower load-resonance frequency than with ungrounded lid. Interchanging the connections yields also different values for the load-resonance frequency.
With twin load capacitance (equal on both sides of the crystal) the difference between grounded and ungrounded lid measurements will be minimized. Also, the asymmetric behaviour with respect to the connections is eliminated.

In series-resonance measurement, interchanging the connections of an NKS7-type crystal has practically no influence on the equivalent electrical parameters or on the series-resonance frequency.
In load-resonance measurement, it can be seen that the setup with twin load capacitance is less sensitive to stray and holder capacitances than the setup with the one-sided load capacitance.

INTRODUCTION

The methods for the passive measurement for quartz crystal units are described in the IEC 60444- series publications. For leaded crystals, the well-known pi-network with blade contacts is used. Although grounding of the enclosure is already mentioned in IEC 60444-1[1], it is often neglected in practice, i.e. measurements of the resonance frequency f_r and load resonance frequency f_L are usually performed with ungrounded enclosure.
With the advent of SMD-type crystals, the pi-network had to be modified for the adaptation of non-leaded crystal units. In a recent IEC-proposal [5] details of an SMD-adapter have been described, consisting of an adapter plate for f_r- and equivalent parameter measurement and a second plate equipped with a load capacitor for f_L-measurement.
In order to suppress electromagnetic interference, modern SMD-type crystals are built with a ceramic carrier that is covered by a metal lid. The carrier has four connecting pads. Two of them are internally connected with the lid and they are used for grounding the device.
Due to this construction the distribution of static capacitances within the crystal has become asymmetric. As a consequence, load-resonance frequency will depend on the orientation of the crystal with respect to the load capacitor when the crystal is grounded.

In the following section experimental data will be given for three fundamental crystals in a 5x7mm ceramic enclosure with a metal lid (NKS7). The nominal frequencies of the three crystals are 16.384 MHz, 20.833 MHz and 26.8 MHz.

2. EXPERIMENTAL RESULTS

2.1 Distribution of static capacitances of the crystal

The static capacitance of the crystal can be split into three contributions:

- the static capacitance between the electrodes of the crystal resonator: C_{0el}
- the capacitances between the metal lid and lower and upper electrode, respectively, of the crystal resonator: C_{0h1} and C_{0h2}.

The values can be determined from a measurement with a capacitance meter. To this end, the following capacitances have been measured at 1MHz (see Figure 1):

- C_I: capacitance between both resonator connections.
- C_{II}, C_{III}: capacitances between the resonator pads and metal lid.

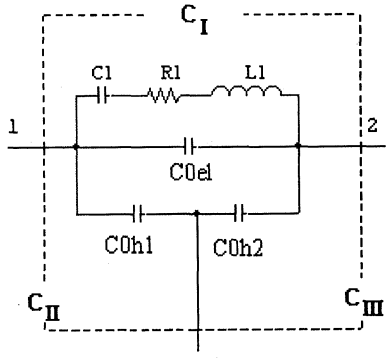


Figure 1

Then C_{0el}, C_{0h1} and C_{0h2} can be determined by a least squares fit from the equations:

$$C_I = C_{0el} + \left(\frac{1}{C_{0h1}} + \frac{1}{C_{0h2}} \right)^{-1} \quad (1)$$

$$C_{II} = C_{0h1} + \left(\frac{1}{C_{0el}} + \frac{1}{C_{0h2}} \right)^{-1} \quad (2)$$

$$C_{III} = C_{0h2} + \left(\frac{1}{C_{0el}} + \frac{1}{C_{0h1}} \right)^{-1} \quad (3)$$

Average values obtained by the least squares fit for the three crystals are given in Table 1.

Table 1: Calculated values for C_{0el}, C_{0h1} and C_{0h2}.

Fnom [MHz]	**C_{0el}** [pF]	**C_{0h1}** [pF]	**C_{0h2}** [pF]
16.384	2.58	1.12	1.41
20.833	3.23	1.09	1.39
26.800	3.73	1.10	1.40

2.2 Cross-talk capacitance

For the measurement of leadless crystal units the pi-network had to be equipped with contact springs. Due to them the cross talk capacitance has been increased considerably. According to IEC 60444-1 [1], the cross talk capacitance C_c can be determined from the cross-talk attenuation A_c with open pi-network:

$$C_c = \frac{10^{-\frac{A_c}{20}}}{25\omega} \quad (4)$$

We found for the SMD pi-network: C_c = 2.15 pF. This value is taken into account during the normalization.

2.3 Results of the measurement of resonance frequency and equivalent parameters

Resonance frequency and equivalent electrical parameters have been measured for the three crystals in two configurations: 1) crystal ungrounded;
2) crystal grounded.

Results are given in the following tables.

Table 2: f_r and equivalent parameters for 16.384 MHz.

GND Y/N	**C_0** [pF]	**C_1** [fF]	**L_1** [mH]	**R_1** [Ω]	**f_r** [kHz]
N	3.46	12.93	7.31	11.7	16377.636
Y	2.81	12.90	7.32	11.3	16377.658

Table 3: f_r and equivalent parameters for 20.833 MHz..

GND Y/N	**C_0** [pF]	**C_1** [fF]	**L_1** [mH]	**R_1** [Ω]	**f_r** [kHz]
N	4.09	15.47	3.78	7.6	20824.468
Y	3.45	15.51	3.77	7.6	20824.469

Table 4: f_r and equivalent parameters for 26.8 MHz.

GND Y/N	**C_0** [pF]	**C_1** [fF]	**L_1** [mH]	**R_1** [Ω]	**f_r** [kHz]
N	4.59	17.11	2.07	6.7	26785.239
Y	3.93	17.12	2.06	6.8	26785.248

Due to grounding, the effective C_0 for the three crystals is lowered respectively by 0.65pF, 0.64pF and 0.66pF. The change of the other parameters is negligible.

2.4 Results of f_L-measurement

Load resonance frequency has been measured for the three crystal frequencies with C_L=10pF and C_L=5pF, respectively, in the following six configurations:

A. Load capacitance C_L consisting of two capacitors placed on both sides of the crystal (C_{L1}=C_{L2}=20pF for C_L=10 pF; C_{L1}=C_{L2}=10pF for C_L=5 pF).
   1) ungrounded (used as reference); 2) grounded
B. one-sided capacitor of 10 pF, connected to crystal connection pad 2 (located close to C_{0h2})
   1) ungrounded; 2) grounded
C. one-sided capacitor of 10 pF, connected to crystal connection pad 1 (located close to C_{0h1})
   1) ungrounded; 2) grounded

Results are given in Tables 5-10.

Table 5: f_L-measurement (C_L=10pF) for 16.384MHz

GND Y/N	**C_{L1}** [pF]	**C_{L2}** [pF]	**f_L** [kHz]	**$(f_L-f_{ref})/f_{ref}$** [ppm]
A1) N	20	20	16384.997*	0.0
A2) Y	20	20	16385.008	0.7
B1) N	-	10	16384.691	-18.7
B2) Y	-	10	16384.414	-35.6
C1) N	10	-	16384.709	-17.6
C2) Y	10	-	16384.491	-30.9

* taken as reference

Table 6: f_L-measurement (C_L=5pF) for 16.384MHz.

GND Y/N	**C_{L1}** [pF]	**C_{L2}** [pF]	**f_L** [kHz]	**$(f_L-f_{ref})/f_{ref}$** [ppm]	**Pull. rng** [ppm] **5...10pF**
A1) N	10	10	16389.028*	0.0	246.0
A2) Y	10	10	16389.034	0.4	245.7
B1) N	-	5	16388.324	-43.0	221.7
B2) Y	-	5	16387.672	-82.7	198.9
C1) N	5	-	16388.366	-40.4	223.2
C2) Y	5	-	16387.891	-69.4	207.5

* taken as reference

Table 7: f_L-measurement (C_L=10pF) for 20.833MHz.

GND Y/N	C_{L1} [pF]	C_{L2} [pF]	f_L [kHz]	$(f_L-f_{ref})/f_{ref}$ [ppm]
A1) N	20	20	20835.012*	0.0
A2) Y	20	20	20835.012	0.0
B1) N	-	10	20834.589	-20.3
B2) Y	-	10	20834.199	-39.0
C1) N	10	-	20834.630	-18.3
C2) Y	10	-	20834.338	-32.4

* taken as reference

Table 8: f_L-measurement (C_L=5pF) for 20.833 MHz.

GND Y/N	C_{L1} [pF]	C_{L2} [pF]	f_L [kHz]	$(f_L-f_{ref})/f_{ref}$ [ppm]	**Pull. rng [ppm] 5...10pF**
A1) N	10	10	20840.327*	0.0	255.1
A2) Y	10	10	20840.320	-0.3	254.8
B1) N	-	5	20839.418	-43.6	231.8
B2) Y	-	5	20838.558	-84.9	209.2
C1) N	5	-	20839.499	-39.7	233.7
C2) Y	5	-	20838.901	-68.4	219.0

* taken as reference

Table 9: f_L-measurement (C_L=10pF) for 26.8 MHz.

GND Y/N	C_{L1} [pF]	C_{L2} [pF]	f_L [kHz]	$(f_L-f_{ref})/f_{ref}$ [ppm]
A1) N	20	20	26799.570*	0.0
A2) Y	20	20	26799.575	0.2
B1) N	-	10	26799.030	-20.2
B2) Y	-	10	26798.497	-40.0
C1) N	10	-	26799.069	-18.7
C2) Y	10	-	26798.711	-32.1

* taken as reference

Table 10: f_L-measurement (CL=5pF) for 26.8 MHz.

GND Y/N	C_{L1} [pF]	C_{L2} [pF]	f_L [kHz]	$(f_L-f_{ref})/f_{ref}$ [ppm]	**Pull. rng [ppm] 5...10pF**
A1) N	10	10	26806.387*	0.0	254.4
A2) Y	10	10	26806.366	-0.8	253.4
B1) N	-	5	26805.232	-43.1	231.4
B2) Y	-	5	26804.144	-83.7	210.7
C1) N	5	-	26805.329	-39.5	233.6
C2) Y	5	-	26804.656	-64.6	221.8

* taken as reference

From these results, the following conclusions can be drawn:

- For the three crystals the frequency deviations are comparable in size.
- If a symmetrical, twin load capacitance is used the frequency difference between ungrounded and grounded measurement is practically negligible.
- In the case of a one-sided load capacitance with the crystal not grounded large deviations occur (~ -19 ppm for C_L=10 pF; ~ - 41 ppm for C_L=5pF) with respect to the case of twin load capacitance. There is also a slight dependence on the orientation of the crystal (~ 2 ppm for 10pF; ~ 4 ppm for 5pF).
- If the crystal is grounded even larger deviations occur (~ -31 .. -40 ppm for C_L=10pF; ~ -64 .. -85 ppm for C_L=5 pF) compared to the twin load capacitance situation. The dependence on the orientation of the crystal becomes substantial (~ 5 ppm for C_L=10 pF; ~13 ppm for C_L=5 pF).
- The pulling range is evidently affected by the measuring configuration (change of ~ 18%, see last columns of Tables 6, 8 and 10).

3. RESULT OF PSPICE SIMULATIONS

To understand the frequency deviations, PSPICE simulations have been performed using the experimental equivalent parameters of the crystals.
For the determination of load-resonance frequency of the crystal in the pi-network , phase-zero frequency has been calculated with PSPICE using a frequency resolution of 1 Hz. All six configurations (A1, A2, B1, B2, C1, C2) have been simulated for the three crystals and both values of C_L (10 pF resp. 5 pF). To explain the frequency deviations, configurations A3, B3 and C3 are simulated. Explanation follows in the next section.

In order to obtain a better agreement with the measured data, we had to take the stray capacitances of the contacts for the capacitors and the crystal into account. The circuit used for simulation is given in Figure 2.

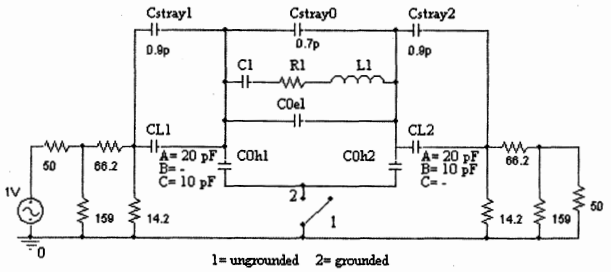


Figure 2: Simulation circuit. DC-paths are omitted for simplicity.

The simulation results for the three crystals are presented in Tables 11, 12 and 13. In the last column the experimental values are shown for comparison. The parameter sets used for the circuit are given in the headings of the tables.

Evidently, there remain discrepancies between theory and experiment. These are due to the finite measurement accuracy for the electrical parameters of the crystal and to the simplifications made for the stray effects. But the order of magnitude and the trend of the values are described correctly.

4. DISCUSSION

The difference regarding C_0 in the equivalent parameter measurement can be explained in the following manner: If the crystal is not grounded the series combination of C_{0h1} and C_{0h2} is in parallel to C_{0el}, i.e. $(1/C_{0h1} + 1/C_{0h2})^{-1}$ is added. For the 16.384 MHz example this amounts to 0.62pF (in agreement with the experimental difference).

If the crystal is grounded C_{0h1} and C_{0h2} are practically shorted by the low impedance of the pi-network (12.5 Ω on each side). This is low compared to the impedance of the holder capacitances (about 7 kΩ). Therefore the contributions of C_{0h1} and C_{0h2} approximately vanish in the grounded measurement.
Similar arguments can be used for the grounded measurement with Cload. Again, the pi-network can be considered approximately as a short compared to the capacitive impedances.

Table 11: Results of PSpice simulations for the 16.384 MHz crystal. We used: C_1=12.93fF; L_1=7.30364mH (chosen to match the experimental f_r); R_1=11.7Ω; C_{0el}=2.58 pF; C_{0h1}=1.12pF; C_{0h2}=1.41pF; C_{stray0}= 0.7pF; C_{stray1}= 0.9pF; C_{stray2}= 0.9pF.

GND Y/N	$C_{L1}/C_{0h1}/C_{0h2}/C_{L2}$ [pF]	f_L [kHz] Sim.	df/f [ppm] Sim.	df/f [ppm] Meas.
	$C_L=\infty$			
N	- / 1.12/ 1.41/ -	16377.636	-450.1	-449.3
Y	- / 1.12/ 1.41/ -	16377.635	-450.2	-447.9
	C_L=10 pF			
A1) N	20/ 1.12/ 1.41/ 20	16385.011*	0.0	0.0
A2) Y	20/ 1.12/ 1.41/ 20	16385.006	-0.3	0.7
A3) N	21.12/ - / - / 21.41	16385.006	-0.3	
B1) N	- / 1.12/ 1.41/ 10	16384.787	-13.7	-18.7
B2) Y	- / 1.12/ 1.41/ 10	16384.425	-35.8	-35.6
B3) N	- / - / - / 11.41	16384.425	-35.8	
C1) N	10/ 1.12/ 1.41/ -	16384.787	-13.7	-17.6
C2) Y	10/ 1.12 / 1.41/ -	16384.554	-27.9	-30.9
C3) N	11.12/ - / - / -	16384.554	-27.9	
	C_L=5 pF			
A1) N	10/ 1.12/ 1.41/ 10	16388.951*	0.0	0.0
A2) Y	10/ 1.12/ 1.41/ 10	16388.942	-0.6	0.4
A3) N	11.12/ - / - / 11.41	16388.942	-0.6	
B1) N	- / 1.12/ 1.41/ 5	16388.432	-31.7	-43.0
B2) Y	- / 1.12/ 1.41/ 5	16387.631	-80.5	-82.8
B3) N	- / - / - / 6.41	16387.631	-80.5	
C1) N	5/ 1.12/ 1.41/ -	16388.432	-31.7	-40.4
C2) Y	5/ 1.12/ 1.41/ -	16387.912	-63.4	-69.4
C3) N	6.12/ - / - / -	16387.912	-63.4	

* taken as reference

This means that C_{L1} and C_{0h1} resp. C_{L2} and C_{0h2} are connected in parallel to each other when the crystal is grounded resulting in an effective capacitance of $C_{L1eff}=C_{L1}+C_{0h1}$ resp. $C_{L2eff}=C_{L2}+C_{0h2}$.

Then, the effective load capacitance C_{Leff} for the three configurations with grounded crystal is:

$$\text{A2)}\quad C_{Leff}=\left(\frac{1}{C_{L1}+C_{0h1}}+\frac{1}{C_{L2}+C_{0h2}}\right)^{-1} \qquad (5)$$

$$\text{B2)}\quad C_{Leff}=C_{L2}+C_{0h2} \qquad (6)$$

$$\text{C2)}\quad C_{Leff}=C_{L1}+C_{0h1} \qquad (7)$$

Equation 5 (case A2) can be simplified if everything is symmetric ($C_{L1}=C_{L2}=2\,C_L$ and $C_{0h1}=C_{0h2}=C_{0h}$).

Then $C_{Leff}=(2/(2\,C_L+C_{0h}))^{-1}=C_L+C_{0h}/2$.
Since the contribution of the holder capacitances is only $C_{0h}/2$, this explains the increase of load resonance frequency in the case of twin load capacitance. (Incidentally, the same holds for the contribution of C_{stray1} and C_{stray2} (see Figure 2)).
On the other hand, the influence of asymmetric holder capacitances in the case of grounded measurement with one-sided Cload is clarified by equations 6 and 7.

Table 12: Results of PSpice simulations for the 20.833 MHz crystal. We used: C_1=15.51fF; L_1=3.766mH (chosen to match the experimental f_r); R_1=7.6Ω; C_{0el}=3.23pF; C_{0h1}=1.09pF; C_{0h2}=1.39pF; C_{stray0}=0.7pF; C_{stray1}= 0.9pF; C_{stray2}= 0.9pF.

GND Y/N	$C_{L1}/C_{0h1}/C_{0h2}/C_{L2}$ [pF]	F_L [kHz] Sim.	df/f [ppm] Sim.	df/f [ppm] Meas.
	$C_L=\infty$			
N	- / 1.09/ 1.39/ -	20824.477	-516.9	-506.1
Y	- / 1.09/ 1.39/ -	20824.475	-517.0	-506.0
	C_L=10 pF			
A1) N	20/ 1.09/ 1.39/ 20	20835.247*	0.0	0.0
A2) Y	20/ 1.09/ 1.39/ 20	20835.239	-0.4	0.0
A3) N	21.09/ - / - / 21.39	20835.239	-0.4	
B1) N	- / 1.09/ 1.39/ 10	20834.933	-15.1	-20.3
B2) Y	- / 1.09/ 1.39/ 10	20834.430	-39.2	-39.0
B3) N	- / - / - / 11.39	20834.430	-39.2	
C1) N	10/ 1.09/ 1.39/ -	20834.933	-15.1	-18.3
C2) Y	10/ 1.09 / 1.39/ -	20834.617	-30.2	-32.4
C3) N	11.09/ - / - / -	20834.617	-30.2	
	C_L=5 pF			
A1) N	10/ 1.09/ 1.39/ 10	20840.635*	0.0	0.0
A2) Y	10/ 1.09/ 1.39/ 10	20840.620	-0.7	-0.3
A3) N	11.09/ - / - / 11.39	20840.620	-0.7	
B1) N	- / 1.09/ 1.39/ 5	20839.938	-33.4	-43.6
B2) Y	- / 1.09/ 1.39/ 5	20838.864	-85.0	-84.9
B3) N	- / - / - / 6.39	20838.864	-85.0	
C1) N	5/ 1.09/ 1.39/ -	20839.938	-33.4	-39.7
C2) Y	5/ 1.09/ 1.39/ -	20839.259	-66.0	-68.4
C3) N	6.09/ - / - / -	20839.259	-66.0	

* taken as reference

When the real capacitances are replaced by the effective load capacitances mentioned above and grounding is omitted (see rows A3, B3, C3 in Tables 11, 12 and 13) , PSPICE calculations give the same results as in the real case (for fundamental SMD-type crystals up to 27 MHz) This confirms the interpretation.

4.1. Impact on oscillators

The crystal parameters are vital to the design of oscillators, in particular for those having narrow tolerances without such compensation mechanisms as applied e.g. in TCXOs. New SMD designs of oscillators often use an open blank and oscillator IC packed in a ceramic enclosure, the metal lid of which is grounded to minimize electromagnetic radiation. The crystal structure is then equal to the oscillator structure apart

from the IC; but the oscillator is always connected to ground. The equivalent to the NKS7-crystal is a 5x7 mm clock oscillator or VCXO, see Figure 3 and 4.

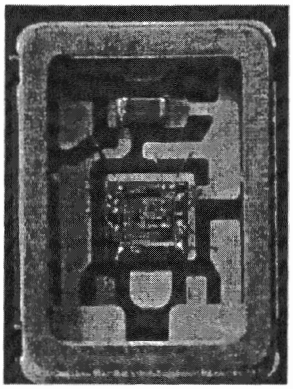


Figure 3: Clock oscillator without resonator.

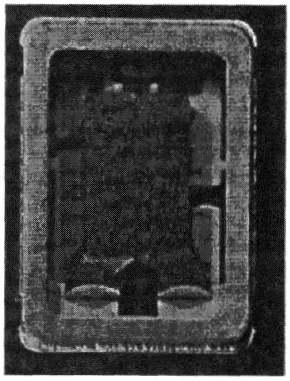


Figure 4: Clock oscillator with resonator.

Table 13: Results of PSpice simulations for the 26.8 MHz crystal. We used: C_1=17.12fF; L_1=2.06227mH (chosen to match the experimental f_r); R_1=6.8Ω; C_{0el}=3.73pF; C_{0h1}=1.10pF; C_{0h2}=1.40pF; C_{stray0}=0.7pF; C_{stray1}= 0.9pF; C_{stray2}= 0.9pF.

GND Y/N	$C_{L1}/C_{0h1}/C_{0h2}/C_{L2}$ [pF]	f_L [kHz] Sim.	df/f [ppm] Sim.	df/f [ppm] Meas.
	$C_L=\infty$			
N	- / 1.10/ 1.40/ -	26785.234	-552.1	-534.8
Y	- / 1.10/ 1.40/ -	26785.234	-552.1	-534.4
	C_L=10 pF			
A1) N	20/ 1.10/ 1.40/ 20	26800.031*	0.0	0.0
A2) Y	20/ 1.10/ 1.40/ 20	26800.011	-0.8	0.2
A3) N	21.10/ - / - / 21.40	26800.011	-0.8	
B1) N	- / 1.10/ 1.40/ 10	26799.618	-15.4	-20.2
B2) Y	- / 1.10/ 1.40/ 10	26798.935	-40.9	-40.0
B3) N	- / - / - / 11.40	26798.935	-40.9	
C1) N	10/ 1.10/ 1.40/ -	26799.618	-15.4	-18.7
C2) Y	10/ 1.10 / 1.40/ -	26799.187	-31.5	-32.1
C3) N	11.10/ - / - / -	26799.187	-31.5	
	C_L=5 pF			
A1) N	10/ 1.10/ 1.40/ 10	26807.075*	0.0	0.0
A2) Y	10/ 1.10/ 1.40/ 10	26807.056	-0.7	-0.8
A3) N	11.10/ - / - / 11.40	26807.056	-0.7	
B1) N	- / 1.10/ 1.40/ 5	26806.178	-33.5	-43.1
B2) Y	- / 1.10/ 1.40/ 5	26804.780	-85.6	-83.7
B3) N	- / - / - / 6.40	26804.780	-85.6	
C1) N	5/ 1.10/ 1.40/ -	26806.178	-33.5	-39.5
C2) Y	5/ 1.10/ 1.40/ -	26805.289	-66.6	-64.6
C3) N	6.10/ - / - / -	26805.289	-66.6	

* taken as reference

A typical total tolerance is 25 ppm, including calibration, drift due to reflow soldering, stability over temperature from -40 to +85 °C, load and supply variations, and aging. For VCXOs, an additional absolute pulling range is often defined as ±50 or ±100 ppm.

As the first stage in development of a new oscillator we normally manufacture an NKS7 crystal with the same blank frequency. The load-resonance frequency as well as the pulling range of crystals depend on the measurement configuration, see Tables 6, 8 and 10. This must be taken into account to obtain the right correlation between crystal and oscillator.

5. CONCLUSION

It has been shown that there is a considerable influence of the Cload configuration on load-resonance frequency if the crystal is grounded. Further, the asymmetry of holder capacitances plays a role if the measurement is performed with a one-sided load capacitance.

In order to avoid measuring differences between different places it is desirable to take those aspects into account in the standardization of load-resonance frequency measurement. The twin load capacitance configuration appears to be the better choice because the results do not depend on the orientation of the crystal and because it is less sensitive to stray capacitances.

Of course, correlation between frequency measurement according to IEC 60444 and working frequency in the application is just as important. Therefore, it is advisable to mark the pad 1 position of the SMD-crystal and take care of a fixed orientation in the oscillator circuit when placing the crystal.

ACKNOWLEDGEMENTS

The authors would like to thank E. Dekker and R. Lendering for their support, which was highly appreciated.

REFERENCES

[1] IEC Publ.60444-1: "Measurement of quartz crystal unit parameters by zero phase technique in a pi-network – Part 1: Basic method for the measurement of resonance frequency and resonance resistance of quartz crystal units by zero phase technique in a pi-network", 1986.

[2] IEC Publ. 60444-4: "Measurement of quartz crystal unit parameters by zero phase technique in a pi-network – Part 4: Method for the measurement of load resonance frequency f_L, load resonance resistance R_L and the calculation of other derived values of quartz crystal units, up to 30 MHz", 1988.

[3] IEC-Publ. 60444-5: "Measurement of quartz crystal unit parameters Part 5: Methods for the determination of equivalent electrical parameters using automatic network analyzer techniques and error correction", 1995.

[4] IEC-Proposal 49/433/NP: "Standard for the Test fixture of Surface Mount Quartz Crystal Units", 1999.

[5] IEC-Proposal 49/434/NP: "Standard for the Test-fixture with Load-capacitance of Surface Mount Quartz Crystal Units", 1999.

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

VERY LONG TIME SCALE AGING PERFORMANCE RESULTS OF QUARTZ CRYSTAL OSCILLATORS

Manish Vaish
MTI-Milliren Technologies, Inc.
Two New Pasture Road, Newburyport, MA 01950

ABSTRACT

Extended aging test results for time periods of up to 1900 days will be presented to provide an insight into the long-term drift characteristics of quartz crystal oscillators. The results will be discussed to show the effects of variations in environmental conditions and power on-off cycles.

1. INTRODUCTION

Very long time scale aging measurements have been made for several types of quartz crystal oscillators including AT and SC cut oven controlled crystal oscillators (OCXO) as well as temperature compensated crystal oscillators (TCXO.) Although every OCXO is aged in production, data is usually only collected for the period of time necessary to meet the aging rate specification. TCXOs are seldom aged at all due to the long test times required to determine the true aging performance. Studies on aging performance results for large groups of oscillators over extended periods of time are typically not available.

Many of the test results shown are of oscillators that are either overruns from production orders, units that did not meet a particular specification or exhibited abnormal behavior during the aging test measurement. It must be noted that the large number of data shown with frequency perturbations is not the norm and represents a small fraction of a percent of oscillators that do not show abnormalities. Nevertheless, data collected from these oscillators over extended periods of time exhibit interesting qualities that are worth noting. A number of oscillators were, in fact, intentionally left on the Aging System specifically to study the aging process and the data for these oscillators are also presented.

The data represents real world performance as it captures the cyclic variations in daily and seasonal changes as well as the effects of retrace due to power failures and scheduled systems maintenance. It is important to realize that the true aging performance of the quartz crystal oscillator is aliased by these factors.

2. METHOD

The data presented for each oscillator has been collected by averaging up to 20 samples of a 1 second gate interval approximately every 2 hours. Automated test measurements on all oscillators are made continuously via software control except when they are either loaded onto or unloaded from the Aging System each day. During this process, data collection is ceased for approximately 3 to 5 hours.

Interruptions may also be a result of routine maintenance work conducted on the Aging System or electrical power failures. In such cases, the oscillators may be powered off for several hours and data collection may be stopped for up to a few days.

All data presented have been reduced in size by removing every other data point due to limitations in the graphing software. The time scale has been normalized with the last data point taken on 15 February, 2000 representing day 0 on the x-axis labeled "Number of Days." As a result, all data points taken prior to this date are represented by a negative number of days from the normalized date. This is to facilitate locating events that are common to the data set of more than one unit and, hence, isolating them from being the behavior of a single oscillator.

Stray points have been removed from the results shown to allow observation of data of interest. Stray data points occur primarily due to wear in the Aging System test fixture sockets and RF switches from excessive use. This results in erroneous frequency readings due to occasional intermittent contact along the RF signal path.

3. AGING

Aging is the change in frequency of the oscillator over time under constant environmental and system-level conditions. Aging in quartz crystal oscillators is caused by changes in either the quartz crystal itself or the remaining components in the oscillator assembly.

Aging in quartz crystals is the result of a combination of several factors. Some of these factors may include the diffusion of impurities and the outgassing of the quartz crystal, its holder, the glass or ceramic base and the adhesive used to mount the quartz. It may also include metal migration from the electrodes into the quartz surface. These events involve an exchange in the mass of the quartz crystal which causes a change in its frequency.

Other factors that also contribute to the aging of quartz crystals include stress relief of the crystal mounts and microscopic holder leaks. While gross leaks in holders are known to cause a downward shift in the frequency, the effects of microscopic leaks on long term aging performance is not well understood.

Frequency drift due to changes in component values over its life may either directly affect the oscillating loop or the steady state function of the sustaining circuits such as voltage regulation, oven control and signal output stages.

The magnitude of aging related to components within the oscillating loop or the surrounding circuits is dependent on the reactance slope of the quartz crystal at the operating frequency. To better understand the effect of component value drift within the oscillating loop on the aging performance, we need to consider the simplified impedance block diagram shown in Figure 1.

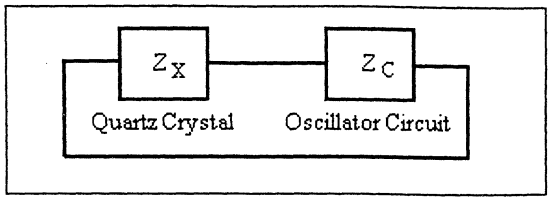


Figure 1

Under resonance,

$$Z_X = - Z_C \quad (1)$$

Where Z_X is the quartz crystal impedance and Z_C is the oscillator circuit impedance.

Since,

$$Z = R + jX \quad (2)$$

Substituting Equation (2) into (1) gives us,

$$R_X + jX_X = - R_C - jX_C \quad (3)$$

$$\Rightarrow \quad R_X = - R_C \quad (4)$$

and $$X_X = - X_C \quad (5)$$

Equation (4) represents the condition for sustained oscillation at a given amplitude and is independent of frequency. Whereas, Equation (5) represents the frequency of oscillation. The slope of the Reactance vs. Frequency curve at the operating point, dX/df, is sometimes referred to as the "pullability" of the quartz crystal.

Some typical values of dX/df in Ω/Hz at series resonance for different crystal cuts and overtones are shown in Table 1.

Frequency	AT Fund	AT 3rd	SC 3rd
5 MHz	-	7.1	110
10 MHz	0.2	1.4	14.5

Table 1

The reactance of the circuit components is responsible for the frequency of the oscillating loop. Suppose that the reactance of all components other than the quartz crystal changes by ΔX_C after several years. For resonance to occur under the new conditions, Equation (5) must still be satisfied.

Therefore,

$$X_X + \Delta X_X = -X_C - \Delta X_C \quad (6)$$

$(X_X + \Delta X_X)$ represents a new operating point on the Reactance vs. Frequency curve of the quartz crystal and a corresponding frequency. The change in frequency is given by,

$$\Delta f = \Delta X_X / (dX/df) \quad (7)$$

Therefore, if a 10μH inductor drifts 1% over the course of its life, using Equation (7), a 5MHz SC Cut 3rd Overtone oscillator would drift 0.029 Hz or 5.7×10^{-09} fractional frequency. Compared to this, a 10 MHz SC Cut 3rd Overtone oscillator would drift 0.433 Hz or 4.3×10^{-08} and a 10 MHz AT Cut 3rd Overtone oscillator would drift 4.488 Hz or 4.5×10^{-08}. It is important to note that an oscillator capable of excessive tuning (or "pullability") is also prone to greater aging rates.

The magnitude of aging related to components in the oven control circuit of an OCXO is dependent on the temperature co-efficient of the quartz crystal. Table 2 shows approximate values of the temperature co-efficient, C, at the turn-point temperature of AT and SC Cut quartz crystals in $°C^{-2}$.

AT Cut	2×10^{-7}
SC Cut	2×10^{-8}

Table 2

The fractional frequency, Δff, deviation due to a ΔT oven temperature change is given by,

$$\Delta ff = C\Delta T^2 \quad (8)$$

Therefore, a 0.2 °C change in the oven temperature over the life of an AT Cut OCXO operating at the turn-point, will result in a 4×10^{-9} fractional frequency change. As it can be seen, the effects of component value drift in the oven control circuit will usually be overshadowed by drift within the oscillating loop.

The aging performance of quartz crystal oscillators is a complex phenomenon that is the cumulative effect of many factors, only a few of which have been mentioned. The effects of some of these factors may tend to cancel one another while others may dominate the aging performance. Furthermore, the effects of these factors may also decay at different rates causing them to dominate at different times during the aging process.

4. FIELD OPERATING CONDITIONS

Under normal operating conditions an oscillator may be subject to a number of environmental changes such as temperature, humidity and atmospheric pressure fluctuations, as well as changes in system-level parameters such as power on-off cycles, supply voltage and tuning voltage instabilities. To determine the aging performance, it is important to distinguish and isolate the effects of these factors on the frequency of the oscillator.

Evident in the data presented are the effects of warm-up, retrace and thermal stability. They are discussed below to differentiate these factors from the aging process, which is only possible with large data sets spanning several years.

4.1 Warm-up

Warm-up is the change in frequency that occurs when the quartz crystal and the components are raised in temperature due to the application of power. While this is obvious in OCXOs, it is also present to a smaller extent in TCXOs and VCXOs (voltage controlled crystal oscillators) as a result of heat dissipation from the circuit

components. Warm-up in OCXOs, where it is most evident, is typically specified on the order of several minutes. The warm-up time is the time it takes for the frequency of the oscillator to reach within a specified frequency tolerance taken at a referenced time period of usually 1 hour from when the power is applied to the oscillator. However, it must be understood that warm-up does not cease after the specified warm-up time period and, therefore, must not be misinterpreted as aging. In fact, frequency changes due to the warm-up process may last up to several weeks. This is evident from the graph shown in Figure 8.

4.2 Retrace

Retrace is the shift in frequency observed by powering off the oscillator and then back on after some time. It is measured by taking the difference between the stabilized frequency, followed by a specified power off time period of usually 24 hours and the frequency measured after a power on time period of usually 1 to 2 hours conducted at a defined constant ambient temperature. Retrace is depicted in the graph in Figure 2 as a result of a power failure at –150 days and –47 days which lasted for approximately 2 to 4 hours.

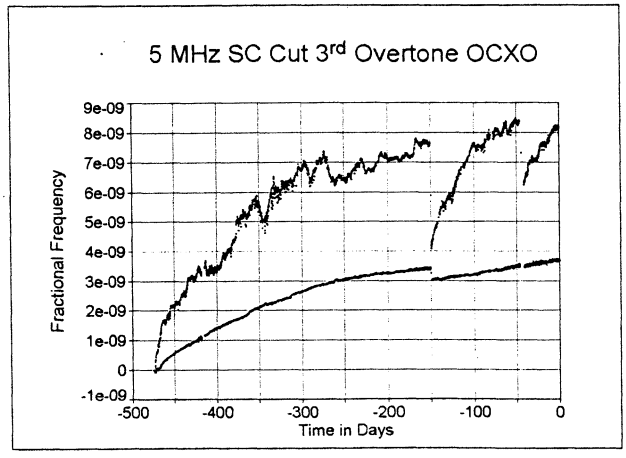


Figure 2

When operating for extended periods of time under static conditions, primarily temperature, the quartz crystal oscillator components tend to physically "anneal" themselves to that operating state. When the oscillator is powered off, the components begin the process of "annealing" themselves to the new state, the extent of which is dependent on the time period in this new state and the temperature difference between the two. When the oscillator is powered on again, the components physically change back to a state similar to right before when the power was turned off. However, they may not reach the exact state at which they originally existed resulting in retrace effects. Since, physical changes in a component causes a change in its electrical characteristics, this in turn will result in a frequency change. Components within the oscillating loop are the primary contributors to retrace and as with warm-up, OCXOs are much more susceptible to retrace due to the greater temperature difference between the power on and off states. In TCXOs, retrace is usually overshadowed by the larger temperature dependent variations and higher aging rates as seen in Figure 3.

Depending on the time period for which the oscillator was powered off and the type of product, it may take anywhere from a few hours to several weeks for the quartz crystal oscillator to recover from the effects of being powered off and return to its previous aging rate. In other instances, the oscillator may not exhibit the effects of retrace, for the same power failure event. In analyzing the data for this study, it was noted that the magnitude of retrace offset was typically proportional to the aging rate of that oscillator.

4.3 Thermal Stability

The effects of thermal stability can overshadow the true aging performance of an oscillator under normal operating conditions, especially when seasonal changes are considered. Quartz Crystal Oscillators with lower thermal stability such as TCXOs and VCXOs are much more susceptible to misinterpretation of the aging rate measurements than, say, precision double-oven OCXOs as seen in Figures 3 and 9.

5. AGING PERFORMANCE RESULTS

Aging test results for several types of oscillators is presented. Two recent power failures are prominently evident on almost all the data presented. A power failure on 30th December, 1999 at day –47 (not a Y2K related event!) lasted for approximately 3 to 4 hours. As a result, data collection was stopped for 4 days. Another power failure at day –150 lasted for nearly 2 hours. The entire Aging System was physically moved to the new manufacturing facility in an adjoining building in 1997. The resulting power interruption lasting over 6 hours is evident at day –837.

5.1 AT Cut Quartz Crystal Oscillator Results

Figure 3 depicts the aging performance of a 10 MHz AT Cut TCXO over 1799 days.

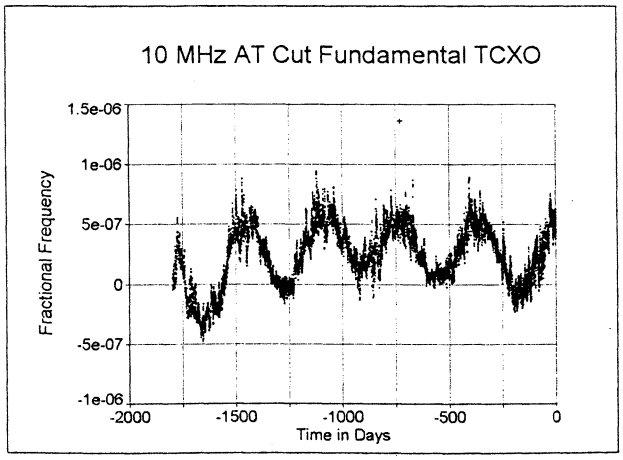


Figure 3

The daily variations in frequency due to temperature is clearly evident and expected. It is quite interesting to also note the cyclic behavior due to seasonal changes. It is important to realize that the yearly aging rate cannot be accurately determined by the first year of data collection alone, as the aging slopes during the transition from the Winter to Summer season and vice versa, are several magnitudes higher than the average yearly aging rate.

Similar aging performance was recorded for 3 other 10 MHz TCXOs over similar time periods and 2 TCXOs at 17.382812 MHz over 243 and 1119 days each.

In comparison, aging performance for 13 MHz Fundamental OCXOs are shown in Figure 4.

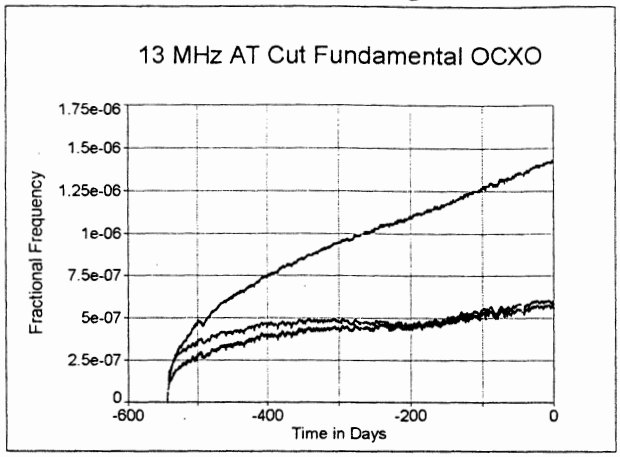


Figure 4

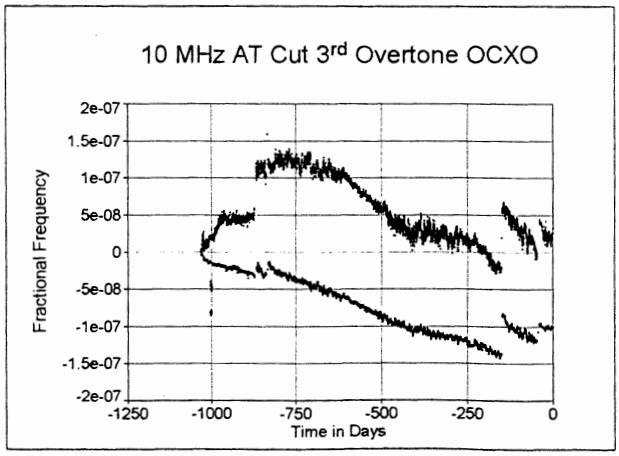


Figure 5

Aging results for 10 MHz 3rd Overtone quartz crystal oscillators are shown in Figure 5.

The aging performance of 2 OCXOs using 50 MHz 3rd Overtone quartz crystals are shown in Figure 6. The steps seen appear on both oscillators at exactly the same time period indicating some external perturbation.

100 MHz 3rd Overtone OCXO aging performance results are shown in Figure 7. Once again, the frequency shift occurs at exactly the same time period on both units with the exception of two frequency steps at –617 days and –433 days seen on one of the oscillators. The oscillators without the steps, is less prone to retrace from the power failures at –150 and –47 days.

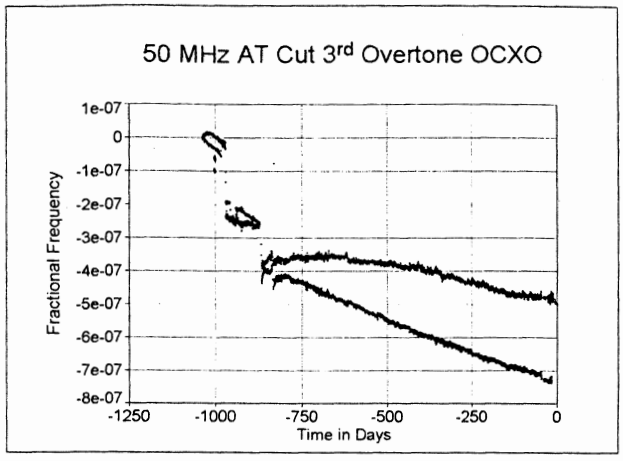


Figure 6

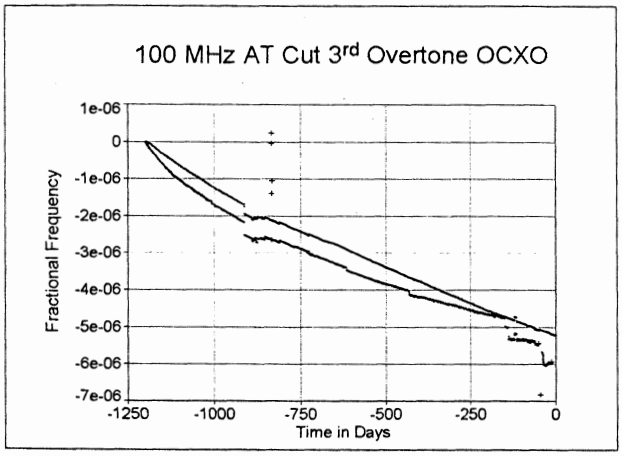


Figure 7

With the exception of the 100 MHz oscillators with much higher aging rates, the cyclic seasonal behavior can be noticed in each of the results presented.

5.2 SC Cut Quartz Crystal Oscillators Results

Aging results for 5 and 10 MHz SC Cut OCXOs are presented. Figure 8 shows a very correlated aging performance of 2 oscillators over a 1951 day time period. Seasonal cycles are also noticeable in these results.

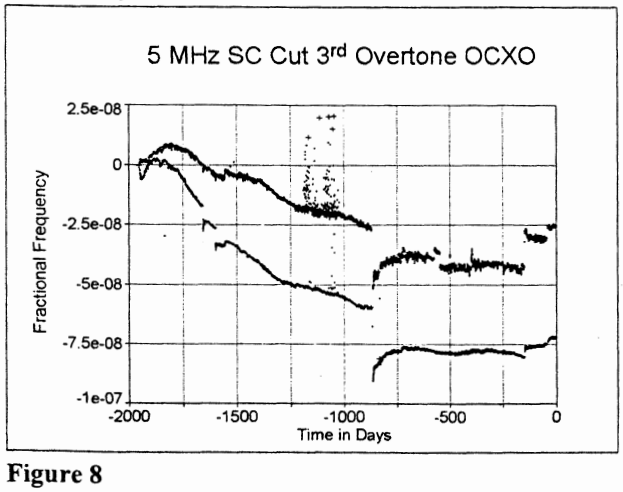


Figure 8

One oscillator exhibited a downward tail during the first few days of aging. Thereafter, it aged with a positive slope, followed by a steady negative slope with a diminishing rate. The gradual and smooth transition from either a positive to negative aging slope or vice versa which can occur over various time periods, is not an uncommon behavior that is perhaps more noticeable in oscillators with low aging rates.

Figure 9 shows additional performance results for 5 MHz 3rd Overtone OCXOs.

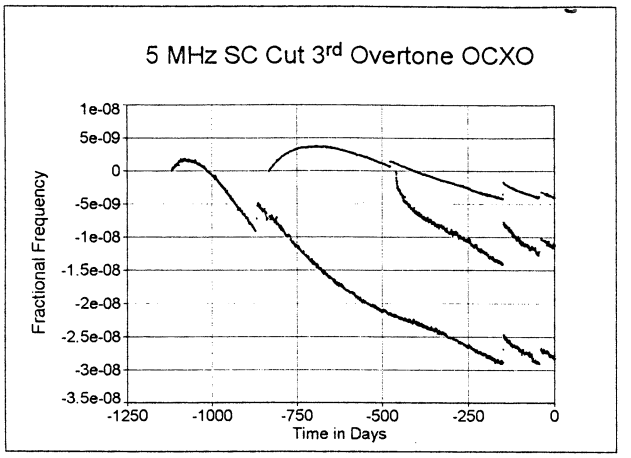


Figure 9

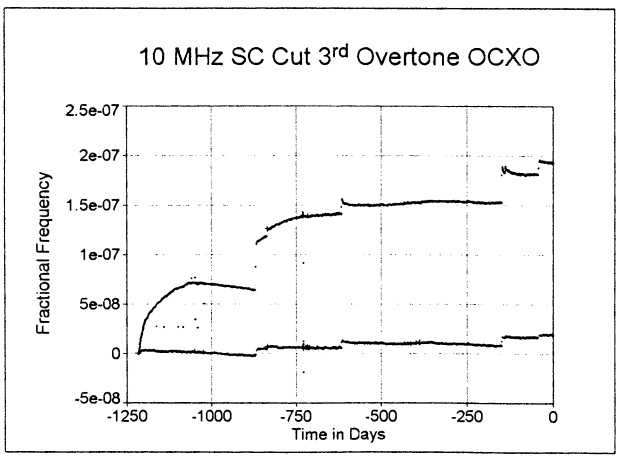


Figure 10

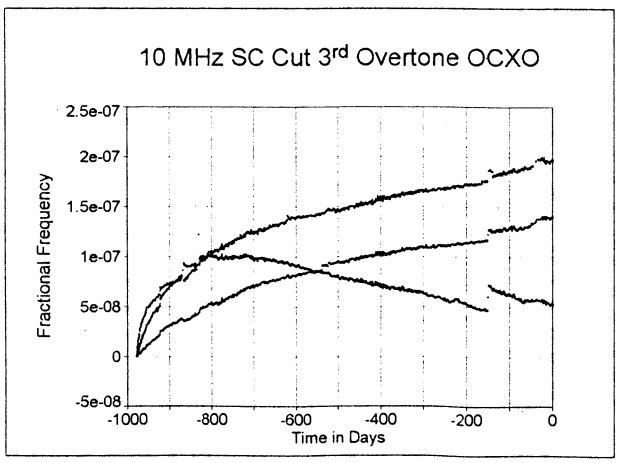


Figure 11

Aging results for 10 MHz 3rd Overtone OCXOs is shown in Figures 10 and 11. It is interesting to note that one of the oscillators exhibits two frequency states between which it jumps from –925 days to –872 days.

Figure 12 shows aging results for 10 MHz 5th Overtone OCXOs.

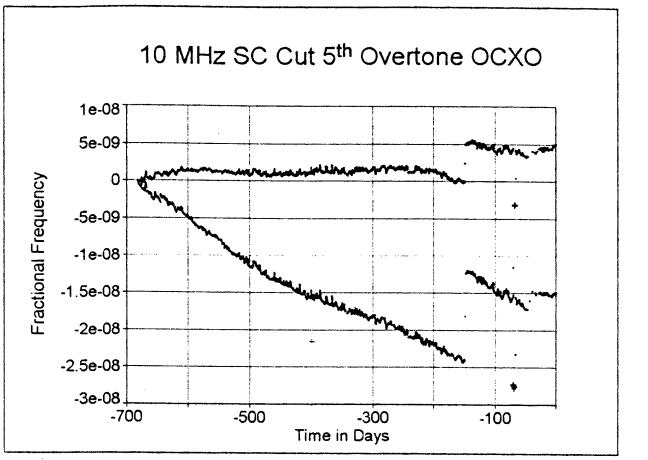


Figure 12

5.3 Experimental Results

An aging experiment was devised involving 32 OCXOs to study aging trends of 5 MHz SC Cut 3rd Overtone crystals.

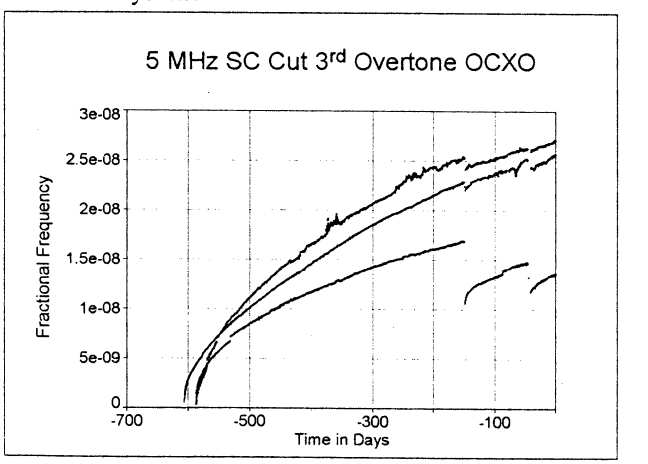


Figure 13

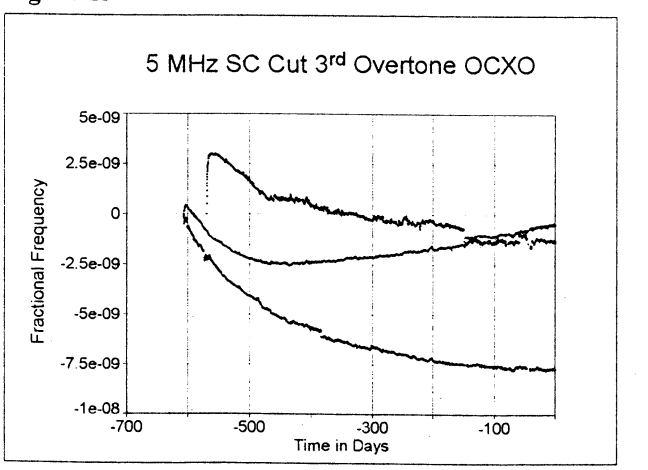


Figure 14

It was noticed that 13 oscillators exhibited a primarily positive aging slope while 19 oscillators exhibited a negative aging slope. Oscillators with positive aging slopes recorded an average of 1.5×10^{-08} total fractional frequency change over a period of approximately 600 days, or about 2.5×10^{-11}/day. Whereas, those with a negative slope recorded an average of 7.5×10^{-09} over the same period, or about 1.3×10^{-11}/day, a factor of 2 better. Representative results from the two groups are shown in Figures 13 and 14.

In analyzing the results of 226 oscillators for this study, the retrace effects for the 3 known power failures at day –837, –150 and –47 were recorded and are presented in Table 3 below.

Aging Slope Direction / Retrace Offset Direction	**# of Units**
Positive / Negative	25
Positive / Positive	29
Positive / Both	3
Positive / Negligible	5
Negative / Positive	93
Negative / Negative	24
Negative / Both	2
Negative / Negligible	6
Undetermined (Due to Data Quality)	39
Total	**226**

Table 3

It is interesting to note that, of the 171 oscillators with discernable retrace offset in one direction, 69% moved in the opposite direction to the aging slope.

5.4 Other Interesting Results

A 10 MHz AT Cut 3rd Overtone OCXO aging performance results due to a leak in the quartz crystal holder is shown in Figure 15.

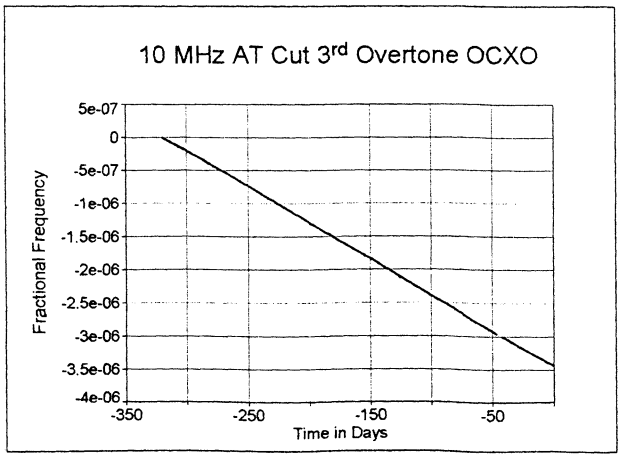


Figure 15

Figure 16 shows the result of a different 10 MHz AT 3rd Overtone OCXO with a precursor event occurring at day –254. This may be the result of the intersection of another oscillation mode aging at a different rate.

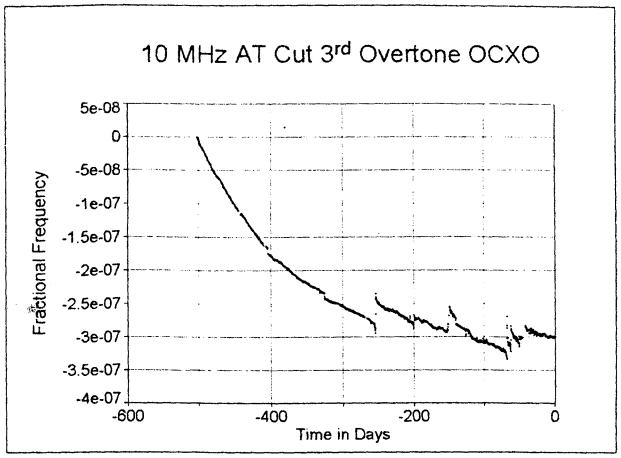


Figure 16

6. CONCLUSION

Aging is a complex process. The aging rate typically scales with the reactance slope of the quartz crystal, indicating some dependency on component performance.

It can sometimes be difficult to detect the aging performance of an oscillator without sufficient amount of data. Results heavily aliased by environmental effects can lead to misinterpretation for small data sets.

Retrace effects are common to all types of oscillators. However, the magnitude of retrace is somewhat independent of the type of oscillator.

For system level calculations, developing an error budget that includes all frequency dependent factors, primarily aging, is a good practice. This may be used to determine whether the application is able to tolerate the expected change in frequency of the quartz crystal oscillator over its life or if the available minimum tuning range is sufficient to offset this change.

7. REFERENCES

[1] J. Vig, Quartz Crystal Resonators and Oscillators For Frequency Control and Timing Applications A Tutorial, U.S. Army Communications - Electronics Command, January, 2000.

SIMULATION OF SAW FILTER BASED ON ZnO/DIAMOND/Si LAYERED STRUCTURE BY SIMULATOR USING COUPLING OF MODE THEORY AND INCLUDING VELOCITY DISPERSION

M. B. Assouar, O. Elmazria, R. Jiménez Riobóo, F. Sarry & P. Alnot,
Laboratoire de Physique des Milieux Ionisés & Applications CNRS UPRES A 7040
Université Henri Poincaré, Nancy I.
E-mail: omar.elmazria@lpmi.uhp-nancy.fr, Address: LPMIA, Université de Nancy I, Bd des Aiguillettes, 54506 Vandoeuvre les Nancy, France, Phone: (33/0) 3 83 91 25 75, fax: (33/0) 3 83 27 34 98

ABSTRACT

CVD diamond[1], due to there remarkable properties, offer much attraction in fabrication of layered structure SAW devices operating on the GHz range. The complexity of layered structures requires the establishment of precise simulation techniques to reduce the number of cycles of design-process-test necessary to perform and to improve the performance of SAW devices.

A simulator based on coupling of mode (COM) theory previously developed for modelling the bulk structure SAW devices, was modified to be adapted for layered structures. The frequency response of ZnO/diamond/Si SAW filter was calculated and the results was compared with experimental ones extracted from the literature[2]. A good agreement is obtained for the frequencies within and close to the pass-band of the filter. Outside of this pass-band, the experimental frequency response exhibits an asymmetry, which is not reproduced by simulation. This asymmetry is attributed to the dispersion, as a function of frequency, of SAW velocity (V_P) and electromechanical coupling coefficient (K^2) which can not be neglected in the case of the layered structure and particularly on the GHz range. In the original program developed for bulk structure, K^2 and V_P was assumed to be constant. To take into account the effect of dispersion, the program was modified by the introduction of dispersive model. The confrontation between the results obtained by simulation including the dispersive model and by experimental measurements shows a good agreement.

1. INTRODUCTION

The SAW devices which the principle is based on the propagation of surface acoustic wave, record a greatly evolution du to there various areas of applications: resonators, filters, sensors...This devices are typically made in piezoelectric materials such as quartz, lithium niobate, or lithium tantalate in which an acoustic wave can be launched by the application of a RF field to a surface interdigital transducer. The travelling wave can then be intercepted and converted back to an electric signal at a neighbouring IDT. The resonant frequency (f_0) of such structure is related to the spacing between the IDT fingers (λ/4) and the acoustic velocity on the substrate (V_P) by the formula : $f_0 = V_P / \lambda$. However it is very difficult to design such a filter in GHz frequencies, due to limitations through the photolithography process. These limitations are related to the wavelength of the acoustic vibration in solid medium, which is close to the micrometer range in the materials listed above.

CVD diamond is used as substrate material in order to obtain high frequency SAW devices. It's receiving a much attraction due to its remarkable characteristics of large SAW velocity. A variety of layered structures with a diamond film such as Zno/Diamond or AlN/Diamond (figure 1) are expected to be applied for high frequency SAW devices having large velocities[2].

In order to realise the layered structures SAW devices, an important number of cycles of design-process-test is necessary. This number can be greatly reduced by to proceeded of numerical simulation. This theoretical study based on the coupling of modes (COM) theory, enable the determination of frequency response of SAW device.

In this paper, the simulator[3] based on COM theory, previously developed for bulk substrate SAW devices, was modified to be adapted for layered structures.

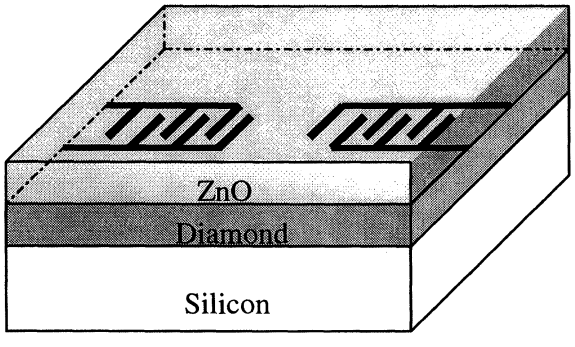


Figure 1: Schematic design of ZnO/Diamond/i layered structure

2. COM EQUATIONS AND ADAPTATION OF THE PROGRAM

The coupling of modes theory is based to the concept that the progressing wave and counter-progressing wave couple with each other in periodic structures[4]. This theory can express the device characteristics satisfactorily and is thus a most appropriate means of analysing SAW devices.

We considered a SAW mode amplitude propagating $A^{\pm}(x)$ in $\pm x$ direction. The COM equations[2] governing this propagation can be written as:

$$\frac{\partial A^{+}(x)}{\partial x} = -j\theta_u A^{+}(x) - j\kappa_{12} A^{-}(x) + j\zeta V$$

$$\frac{\partial A^{-}(x)}{\partial x} = j\kappa_{12} A^{+}(x) + j\theta_u A^{-}(x) + j\zeta V$$

$$\frac{\partial I(x)}{\partial x} = -4j\zeta A^{+}(x) - 4j\zeta A^{-}(x) + j\omega CV$$

where $\theta_u = \beta - \frac{2\pi}{P}$, β is the wavenumber, P is the periodic length of IDT, V is the applied voltage of the IDT, $I(x)$ is the current following the structure, κ_{12} is the mutual coupling coefficient, ζ is the transduction coefficient and C is capacitance of IDT.

The software based on this theory is used to calculaate a frequency response of different bulk substrate SAW devices such as filters and resonators.

In the first time, this program is validated by the simulation of frequency response of SAW filter bulk substrates (quartz). The simulation results were compared to experimental ones performed on device completely developed in our laboratory. It consists of two identical bidirectional interdigital Transducers etched from a thin film deposited on a ST cut quartz crystal by conventional lithography process. The IDTs wich have a 150 nm aluminium thickness, present a lines and spaces of 8µm and consequently the period of the transducer is λ=32 µm. This result in a centre frequency of approximately 98.4 MHz. Figure (2) shows that the experimental and simulated result exhibits a good agreement in the pass-band. Outside this band, loss divergence is observed probably due to the experimental set-up.

In the second time, the program will be adapted for layered structures in order to take into account the effect of influence of each layer on the propagation. For that, a sub program related to this structure is introduced.

Thanks to its highest SAW velocity among all materials, diamond combined with a piezoelectric thin film as ZnO, can provide advantage for the fabrication of high frequency SAW devices. In order to validate this program, we simulated the frequency response of the ZnO/Diamond/Si SAW filter and the result will be compared with experimental ones performed by Nakahata et al[2];

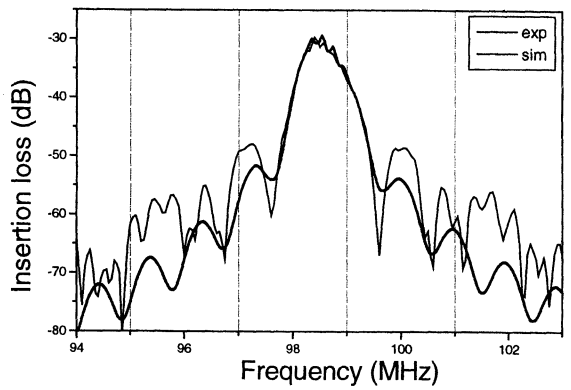


Figure 2: Frequency response of quartz SAW filter: experimental and simulated results

To calculate a propagation velocity (V_p) of surface acoustic wave in ZnO/Diamond/Si layered structure, we assumed that the thickness of diamond film used is large enough to neglected the influence of the silicon substrate. This velocity (V_p) is calculated by another software[5], which calculates the resulting spectrum of the light scattered by the ripple effect (surface corrugation) as obtained by the Brillouin spectroscopy. This software uses a Green's function formalism to calculate the corresponding scattering cross section. To simulate the surface mode spectrum it is necessary to introduce some elastic parameters in the software. In our case, we have combined the experimental values obtained for the elastic constants of an, from the elastic point of view, isotropic CVD-Diamond film[6] (our substrate, c_{11} = 1140.5 GPa and c_{44}=525.5 GPa) with the literature[7] values of ZnO (assuming an isotropisation of the elastic constants of the ZnO film, c_{11} = 210 GPa and c_{44} = 42 GPa). The obtained results are shown in figure 3. The lower curve represents the variation of the Rayleigh mode velocity (the relevant SAW mode) as function of the product kh (k corresponds to the acoustic wave vector and h is the thickness of the ZnO film, $k = \frac{2\pi}{\lambda}$). The COM parameters has been also calculated for this structure.

Considering the technological parameters used by Nakahata et al[2], kh = 0.5, the velocity of the first mode of surface acoustic wave (Rayleigh mode) deduced for the plot of figure (3) is $V_p = 10000\ m/s$; This value agrees well with the experimental one.

The frequency responses of ZnO/Diamond/Si simulated and measured are shown in figure (4). We can observe that the simulation agrees well with the measured results at the pass-band. Although a discrepancy is obtained at the stop-band

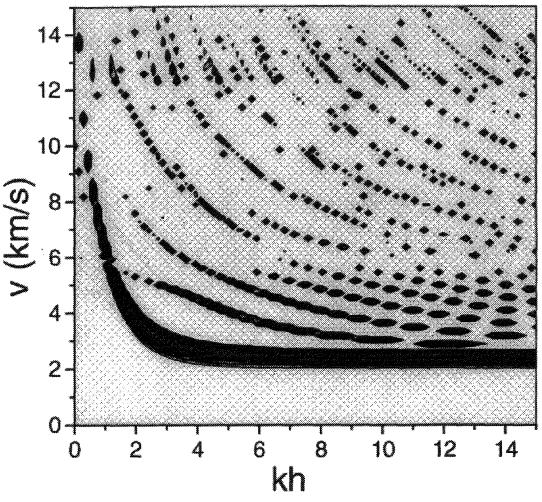


Figure 3: Velocity dispersion versus (kh) of ZnO/Diamond/Si SAW filter for different propagation modes of surface acoustic waves

The experimental frequency response exhibits an asymmetry, which is not reproduced by the simulation. This asymmetry can be attributed to the dispersion of SAW velocity (V_P) and electromechanical coupling coefficient (K^2) as a function of frequency. This dispersion can not be neglected in the layered structures because the various materials composing this structure determine the value of K^2 and V_P. In the original program developed for bulk substrates, these parameters were assumed to be constant.

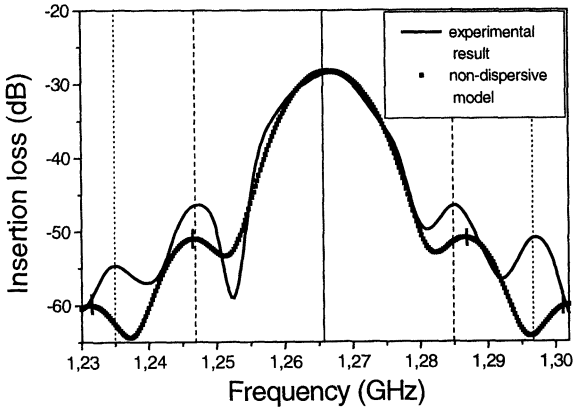


Figure 4: comparison between experimental and simulated frequency responses for ZnO/Diamond/Si SAW filter

3. Dispersion effect

The velocity dispersion in SAW layered structure was already studied by Hachigo et al[8-9] using Smith's equivalent circuit model. In their simulations the dispersion of electromechanical coupling coefficient as a function of frequency is neglected. In this study, the dispersion of V_P and K^2 as a function of frequency will be considered and their effects, on frequency response will be presented and compared to experimental results.

Velocity dispersion

Considering the non-dispersive model, the frequency response is symmetric around central frequency (f_0). The number of arches contained in a whole frequency response band is the same in the right side or in the left side of f_0. This is not the case of the experimental response where the number of arches is lower in the left side of f_0. This can be explained by a variation of acoustic velocity as a function of frequency. Figure 2 shows that the velocity varies with the product kh, knowing that is h constant and k is equal $2\pi / \lambda$. The asymmetry observed is easily explained with the variation of velocity shown in figure 2. In order to take into account this effect, a velocity dispersive model is introduced in the program. The frequency response obtained by this model is reported in figure 5 and the comparison with experimental one shows a good synchronisation. The discrepancy observed in the magnitude of signal is due to the dispersion of electromechanical coupling coefficient as a function of frequency.

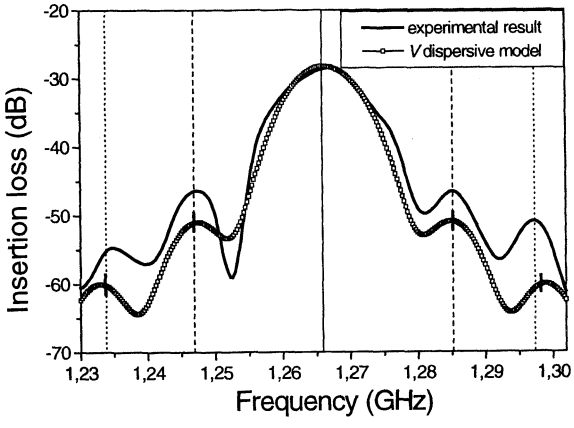


Figure 5: simulated result including velocity dispersion compared with experimental once

Electromechanical coupling coefficient dispersion

The electromechanical coupling coefficient is defined as $K^2 = \frac{V_0^2 - V_\infty^2}{V_0^2}$ Where:

V_0 and V_∞ are respectively open-circuited and short-circuited phase velocities of the surface acoustic wave. The K^2 coefficient is of crucial importance for estimating strength to surface acoustic waves on layered structures. The experimental results shows that the rejection of the frequency above f_0 is lower than the rejection of frequency below f_0. Knowing that the insertion loss varies with the K^2 coefficient, the asymmetry observed can be explained by its dispersion as a function of frequency. This dispersion can be attributed to the influence of various materials forming the SAW structure on the velocity and consequently on the K^2. Figure 6 shows the effect of K^2 dispersion on the frequency response of SAW filter by the comparison of dispersive and non-dispersive model.

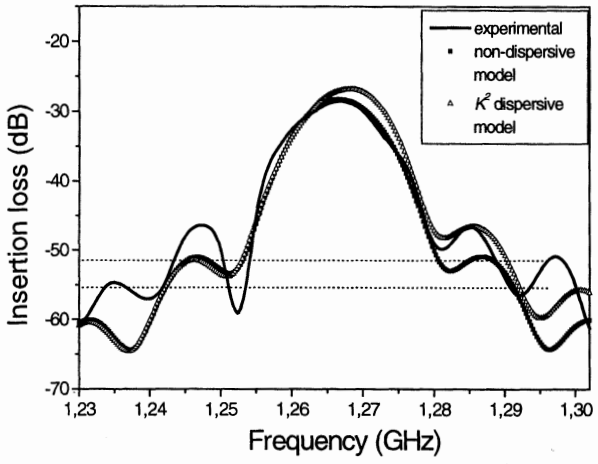


Figure 6: Simulated result including K^2 dispersion as a function of frequency compared with experimental once

4. Conclusion

A COM program originally developed for SAW bulk substrates was adapted for layered structures. The effect of dispersion on SAW velocity V_P and electromechanical coupling coefficient K^2 as a function of frequency was introduced in the program. Layered structure ZnO/Diamond/Si was simulated by the modification of the program with dispersive and non-dispersive models, and the results compared with the experimental one. The non-dispersive model agrees well with measurements but only in the pass-band. The velocity dispersive model enables to improve the frequency response in stop-band too in regard to synchronisation of arch. However, the magnitude of arch shows a discrepancy compared to experimental responses. This discrepancy is corrected by the introduction of K^2 dispersion effect. The combination of both V_P and K^2 dispersion effect is in progress.

Acknowledgements

We would like to thank Dr. K. Y. Hashimoto at Chiba University in Japan for provide the COM software and Dr. A. G. Every at the University of Witwatersrand in South Africa for the software to calculate the SAW velocity.

References

[1]- B. Dischler and C. Wild, "Low-Pressure Synthetic Diamond, Manufacturing and Applications", Springer (1998).

[2]- Nakahata, K. Higaki, A. Hachigo, S. Shikata, N. Fujimori, Y. Takahashi, T. Kajihara and Y. Yamamoto, "High Frequency Surface Acoustic Wave Filter Using ZnO/Diamond/Si Structure", Jpn. J. Appl. Phys. Vol 33, pp: 324-328, (1994).

[3]- K. Y. Hashimoto, "General-Purpose Simulator for Leaky Surface Acoustic Wave Devices Based on Coupling-of-Modes Theory", IEEE Ultrasonics symposium, pp: 117-122. (1996).

[4]- K. Nakamura, "A Simple equivalent circuit for interdigital transducers based on the coupled-mode approach", IEEE Transductions on Ultrasonics, Ferroelectrics, and Frequency Control, Vol 40, No 6, pp: 763-767, (1993).

[5]- A. G. Every, private communication, University of Witwatersrand (Johannesburg, South Africa).

[6]- J. K. Krüger, J. P. Embs, S. Lukas, U. Hartmann, C. J. Brierley, C. M. Beck, R. Jiménez, P. Alnot and O. Durand, " Spatial and Angle Distribution of Internal Stresses in Nano- and Micro-Structured Chemical Vapour Deposited Diamond as Revealed by Brillouin Spestroscopy", Journal of Applied Physics (in press)

[7]- D. Royer et E. Dieulesaint, "Ondes élastiques dans les solides", Massson (1996).

[8]- A. Hachigo and D. C. Malocha, "SAW Dispersive Equivalent Circuit for Diamond Layered Structures", IEEE Ultrasonics Symposium, (1996).

[9]- A. Hachigo and D. C. Malocha, "SAW Device Modeling Including Velocity Dispersion Based on ZnO/Diamond/Si Layered Structures", IEEE Tranductions on Ultrasonics, Ferroelectrics, and Frequency Control, Vol 45, No 3, pp: 660-666, (1998).

A MINIATURE PRECISION OVEN QUARTZ OSCILLATOR SETS NEW SIZE VS. PERFORMANCE STANDARD

Bryan T. Milliren
MTI-Milliren Technologies, Inc.
Two New Pasture Road, Newburyport, MA 01950

ABSTRACT

The next generation frequency references for high volume applications, such as base stations and telecom digital switches, require a very stable ovenized crystal oscillator. The 220 Series oscillator has been designed by MTI-Milliren Technologies, Inc. and through extensive testing and evaluation, the oscillator has proven to provide the performance attributes necessary to fulfill most, if not all, future system requirements.

1. INTRODUCTION

The 220 Series oscillator utilizes SC cut quartz resonators in recently developed low profile TO-8 holders, achieving a typical thermal stability of 5e-9 per 100°C at a low cost. Housed in a standard SMT or through-hole 16 pin ·dual in-line welded hermetic package, the product measures 24.77mm x 20.32mm x 12.70mm. The oscillator consumes 2.5W (typical) during warm-up and approximately 0.8W @ 25°C, steady state. Key parameters, such as thermal stability, phase noise, short term stability, supply voltage sensitivity, and aging are comparable and, in many cases better, than that obtained by traditional and much larger units currently available.

The 220 Series utilizes a new oscillator topology, which reduces the component count and improves reliability by a factor of approximately 2 times over traditional circuit designs. The entire assembly including oven control, heaters, voltage regulation, and oscillator circuit is accomplished with five active components.

The 220 Series has been designed to be a product that can be consistently manufactured in a high volume production environment. Test data characterizing the various key parameters will be shown.

2. PRODUCT DESIGN

The key design goals in developing the 220 Series oscillator have been:

- Thermal Stability < 1e-8 per 100C
- High Reliability, MTBF
- Low Power Consumption
- Very Fast Warm-up
- Standard Package, Reduced Size, Surface Mounting, High Integrity Hermetic Seal, Rugged for High Shock and Vibration Environments
- Manufacturability and Consistency
- Low cost, High Volume Production

A circuit topology was adopted to allow a significant reduction in component count compared to traditional approaches such as Colpitts, Pierce, etc. A MMIC amplifier gain block with the quartz resonator, mode selector, and matching network in the feedback path achieves the targeted goals. Proper choice of the mode selector and matching filter component values allow the quartz to operate around series resonance. Figure 1 shows the oscillator block diagram. The entire oscillator assembly, including buffer amplifiers, is ovenized for best performance.

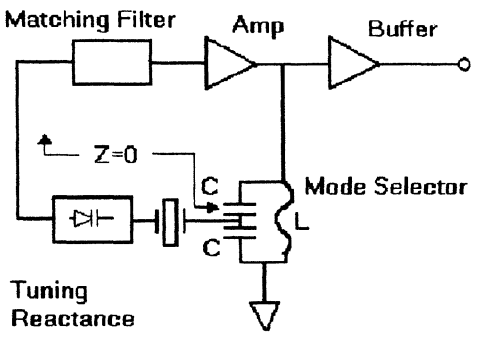


Figure 1

The 220 Series uses a common 16 pin DIL mechanical package which offers good hermetic seal characteristics and the ability to provide a surface mountable package. Figure 2 shows photo of both pin and surface mounting types compared to a US 5 cent coin.

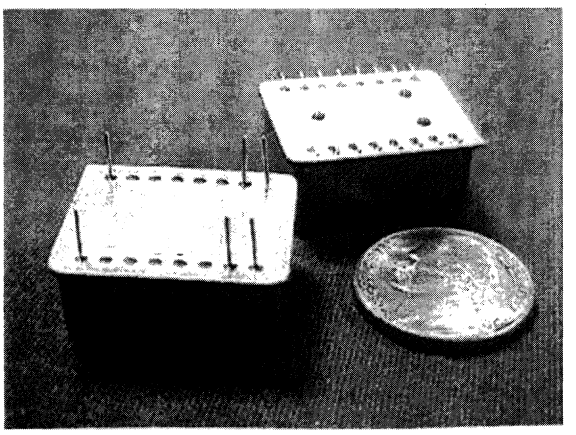


Figure 2

3. PERFORMANCE DATA

Data is shown for a number of important oscillator parameters. Sets of results were chosen to be representative for 2 common products, 5 and 10MHz units using SC cut quartz with 12V supply inputs and +9dBm sine outputs.

3.1 Phase Noise Results

Phase noise data is shown in Figure 3 and Figure 4 for 5 and 10MHz versions respectively. Results were obtained using an HP 3048 system for pairs of like units. Typical specification limits are shown as solid straight lines.

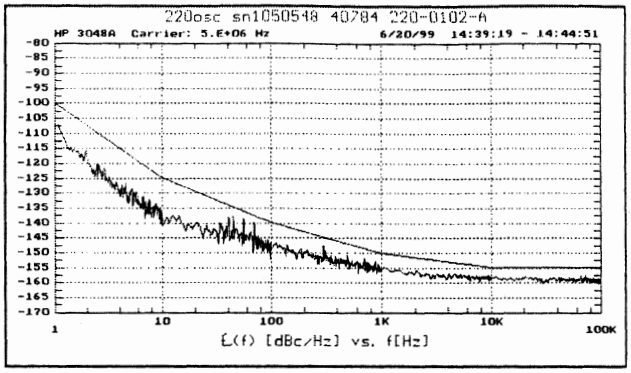


Figure 3

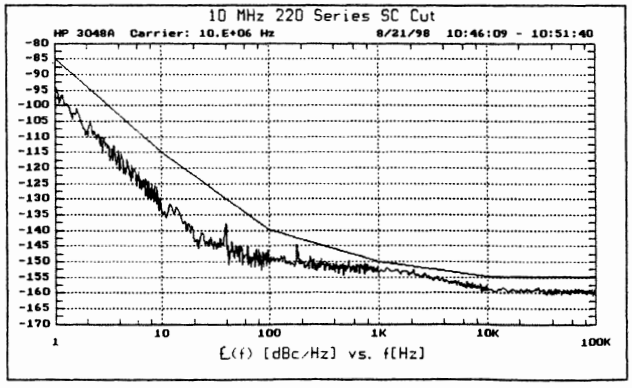


Figure 4

3.2 Output spectrum Results

Output spectrums are shown in Figures 5 and 6 for typical examples of 5 and 10MHz oscillators respectively. The output circuit contains a matching filter, which allows for low harmonic content as seen in the graphs below.

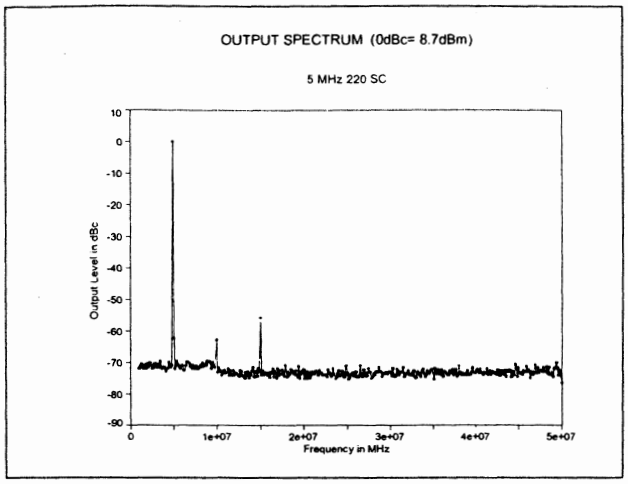


Figure 5

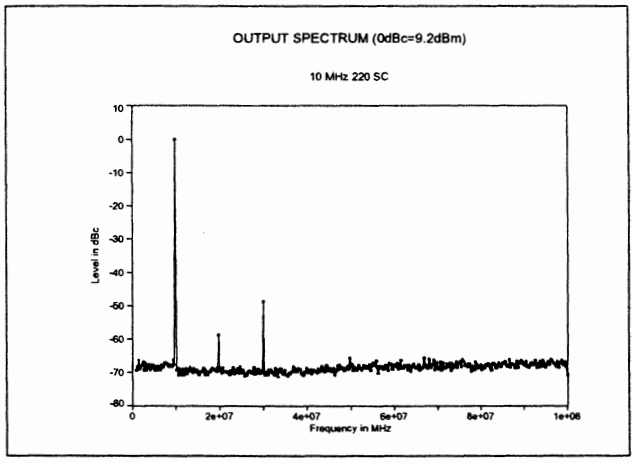


Figure 6

3.3 Thermal stability Results

The thermal stability is shown for both 5 and 10MHz units in Figures 7 and 10. Each stability graph is referenced to its respective "Temperature vs. Time" graph. This method of presentation allows the complete data set to be viewed in context to thermal transients and other phenomena, which might otherwise be missed in a simple Frequency vs. Temperature presentation. The Temperature vs. Time graph is shown for each frequency in Figures 8 and 11 respectively. The data for a 5MHz unit shows an up-down ramp with returns to 25C, while the 10MHz unit is a unidirectional ramp. Up-down ramps are most useful for qualification testing while single direction ramps are most used in production test. Each set of data also includes a Power vs. Time graph, see Figures 9 and 12 for 5 and 10MHz respectively.

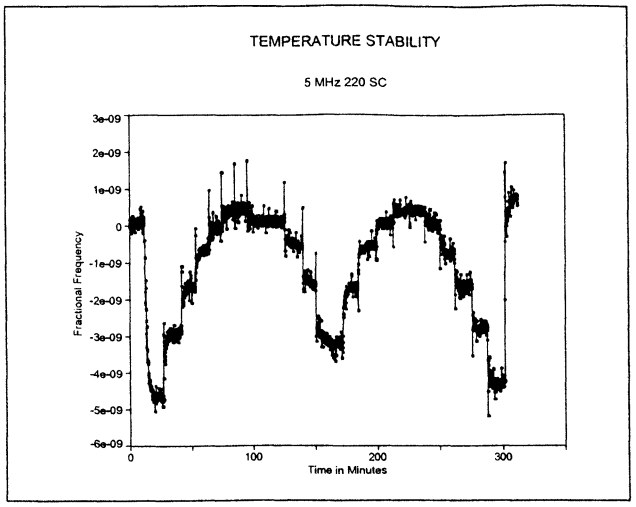


Figure 7

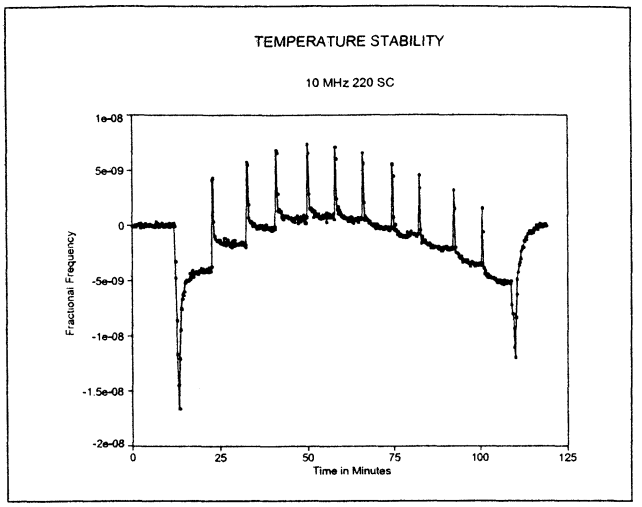


Figure 10

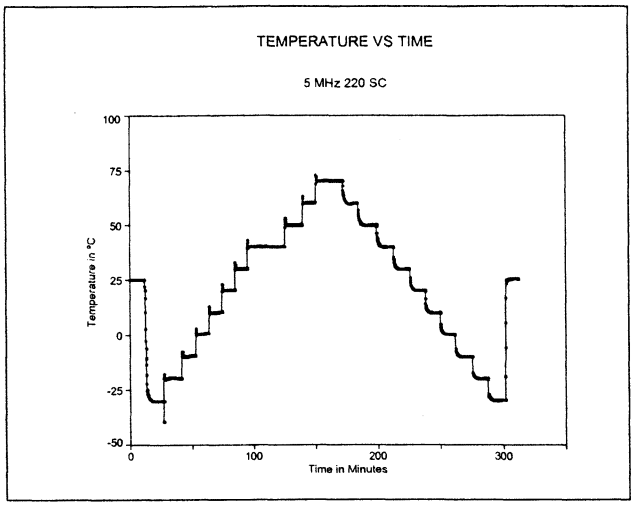


Figure 8

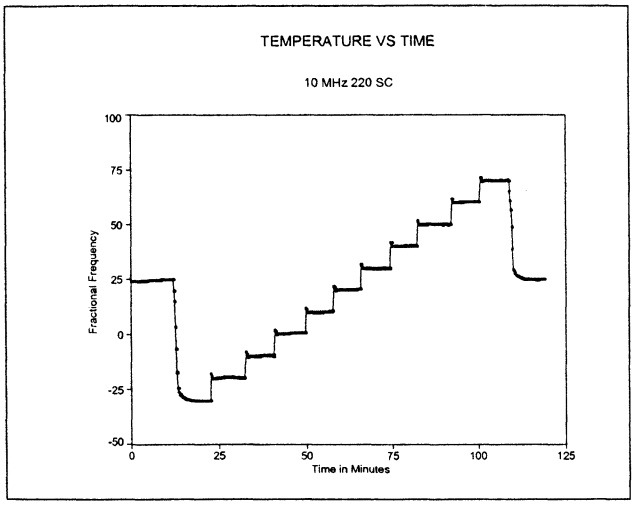


Figure 11

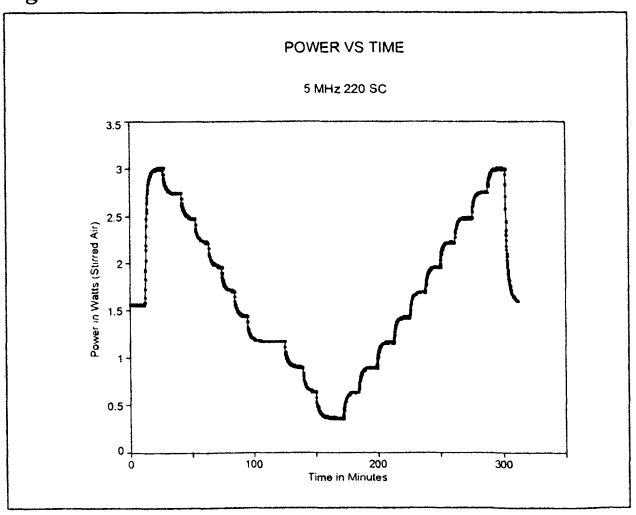


Figure 9

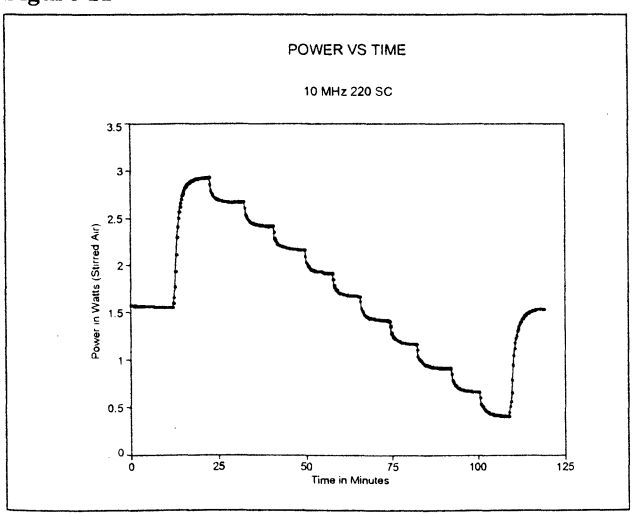


Figure 12

3.4 Warm-up Results

Warm-up data is shown in Figure 13 for a 5MHz SC 220 series oscillator. The measurements were made after an off period of greater than 2 hours. A graph of power vs. time is shown in Figure 14. The power consumption was measured in still air. The frequency is within a 2e-8 window in approximately 2 minutes at 25C. Although not shown, the frequency actually starts out at –20e-6 and quickly rises to the operating frequency. The small positive section of the graph represents an overshoot caused by a less than perfect SC crystal cut. Performance data for 10MHz product is very similar to 5MHz and thus no separate data is presented for this case.

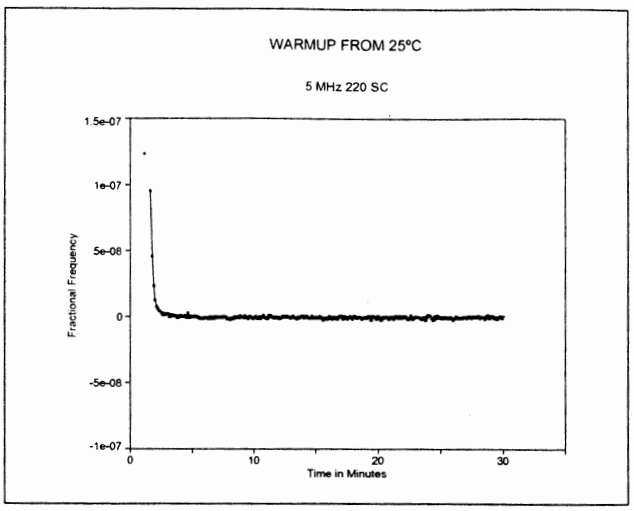


Figure 13

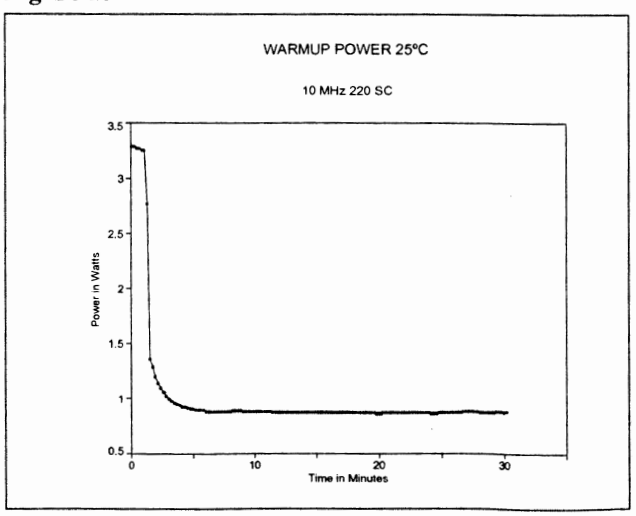


Figure 14

3.5 Tuning Function Results

Tuning curves are shown below for 5 and 10MHz 220 series oscillators in Figures 15 and 16 respectively. Note that the tuning rate of the 10MHz oscillators is approximately 4 times greater than the 5MHz product. The tuning range scales proportionally to the dX/dF of the quartz crystal and the available reactance swing of the tuning varactor diode. Note the non-linearity on the 5MHz plot. Many specifications omit the 0 to 0.5V range to avoid this area. The linearity represented by the graphs below is typically 10% or less.

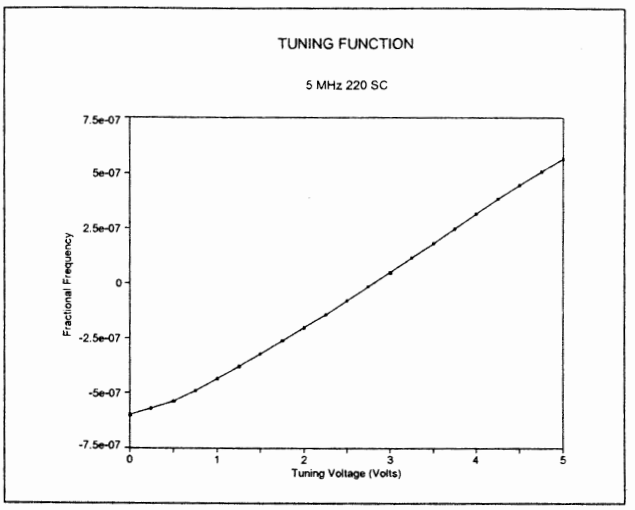


Figure 15

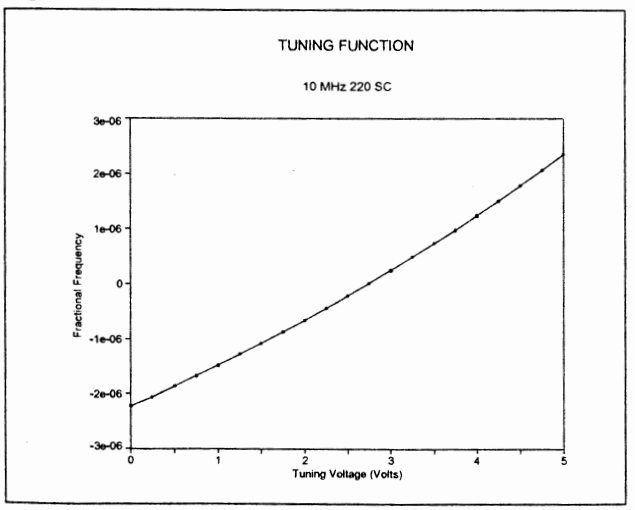


Figure 16

3.6 Supply Sensitivity Results

The supply sensitivity is characterized in a similar fashion to the thermal stability. The supply is varied over the operating extremes while the frequency is being recorded. Like the thermal stability measurements, the frequency and voltage are measured with respect to time, and then correlation is made between the time and voltage axis. As is seen in Figures 17 and 19, this method captures any transient effects, which may occur. Figure 18 shows the Voltage vs. Time function. Note that in each case for 5 and 10MHz the transient behavior is greater in magnitude than the static frequency offsets.

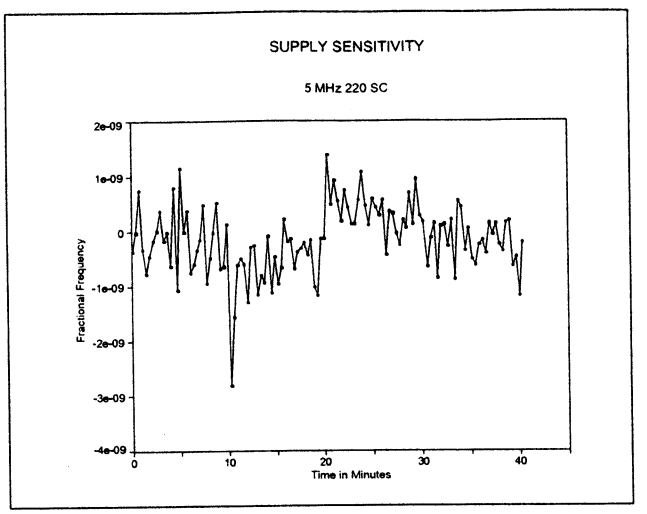


Figure 17

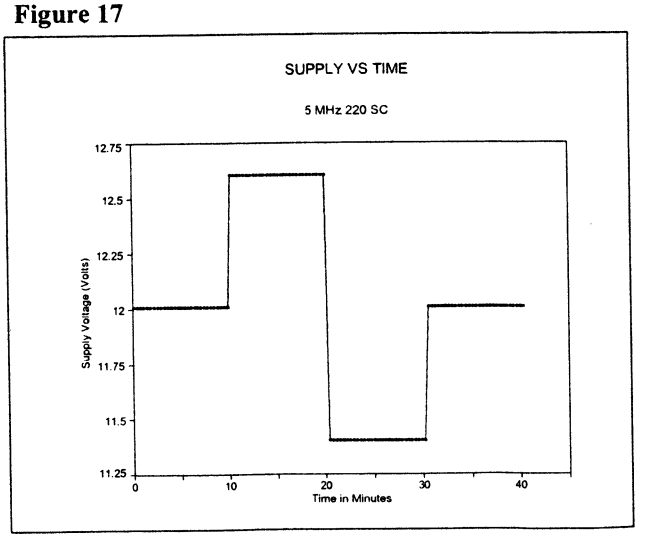


Figure 18

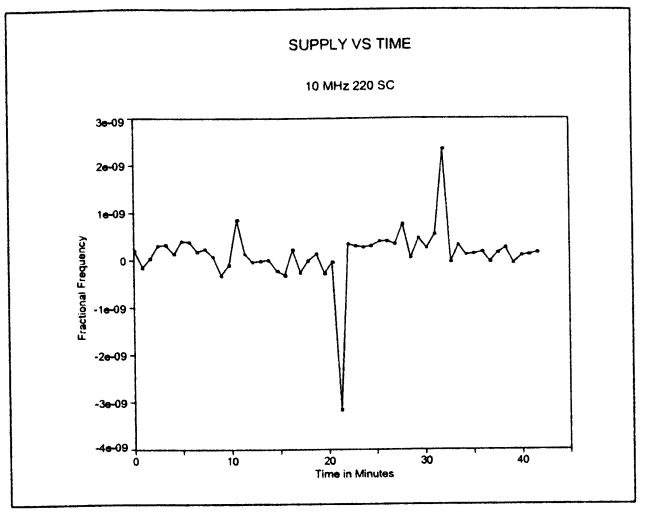


Figure 19

3.7 Aging Test Results

Aging test results for 5 and 10MHz 220 series oscillators are shown below in Figures 20 and 21. The average initial aging rates for 5MHz 220 series units are less than 2e-10/day over the first 30 days. 10MHz types have and average value for the same period of less than 5e-10/day.

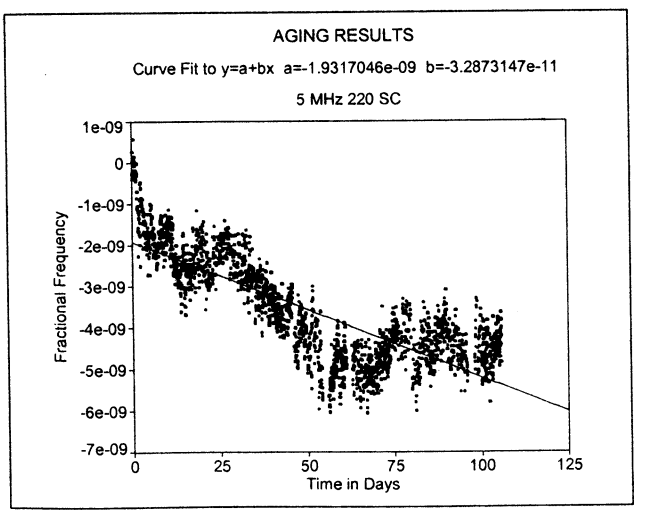


Figure 20

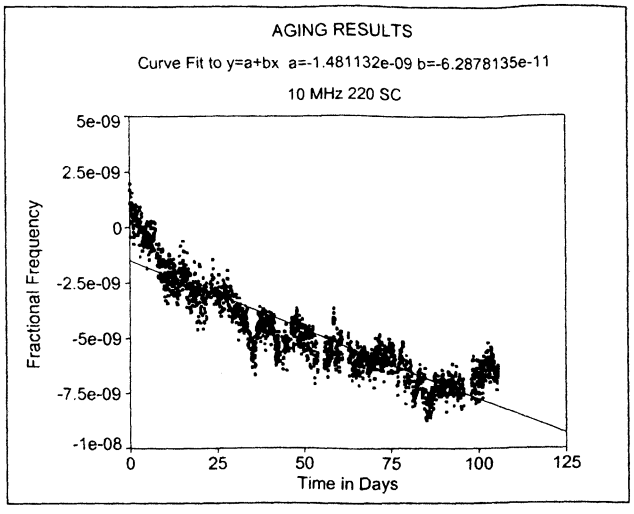


Figure 21

4. CONCLUSION

The data presented gives a brief overview of the 220 series product performance. Results collected over the past 2 years show that oscillators with a small physical size can equal or exceed performance obtained with much larger products of the past. The reduced complexity and lower component stresses help greatly in the goal of improving failure rates as well as reducing overall production costs.

5. REFERENCES

[1] J. Vig, Quartz Crystal Resonators and Oscillators For Frequency Control and Timing Applications A Tutorial, U.S. Army Communications - Electronics Command, January, 2000.

[2] M. M. Driscoll, Low Noise Crystal Oscillators using 50-Ohm, Modular Amplifier Sustaining Stages, Frequency Control Proceeding # 6-40-329.

THERMODYNAMIC ASPECT OF SHORT-TERM FREQUENCY STABILITY OF DIRECTLY HEATED RESONATORS

Igor Abramzom
Consultant, P. O. Box 7357, Omsk 644020, Russia
Roman Boroditsky
Valpey-Fisher Corporation, 75 South Street, Hopkinton, MA, 01748 USA

Abstract.

This paper is an attempt to study a problem of degradation of short-term stability (STS) of OCXO based on the directly heated resonators (DHR) due to variations in the heating power. The carried out research allowed development of the means for designing DHR with STS better than 5E-12 per 1 s, which is comparable to the best results attainable with the conventional technique of heating the crystal.

1. INTRODUCTION.

The directly heated resonator (DHR) technology provides significant reduction in warm-up time, size, and power consumption of OCXO thanks to employing thin film heater and thermistor directly deposited on the crystal surface [1]. Frequency stability of such devices, however, is essentially dependent on the temperature gradients over the crystal plate. The gradients are produced by power dissipated in the film heaters. Variations in the ambient temperature and/or supply voltage cause variations in the power dissipated by the heaters, which in turn leads to thermal gradients variation. This effect results in degradation of temperature and/or supply voltage frequency sensitivity of the OCXO. Periodic or random variations of the heating power caused by changes in external conditions or noise in thermocontroller circuitry translate into fluctuations of the DHR frequency through thermodynamic sensitivity of the crystal plate. That leads to degradation of STS and close-to-the-carrier phase noise performance of the OCXO employing such resonators.

A goal of the present paper is theoretical and experimental analyses of the influence of the thermodynamic effects on STS of the DHR, which would ultimately lead to the development of effective means of their prediction and improvement.

2. MODEL OF THE HEATING POWER FLUCTUATIONS

To understand an origin of the heating current fluctuation in the DHR lets consider its construction along with the thermocontroller circuit used to accurately control the crystal temperature. The thermocontroller circuit is shown schematically in Fig. 1. It consists of a thermo-sensitive bridge, an amplifier and a regulating transistor governing the current through the heater resistor deposited on the crystal plate. The DHR construction (Fig.2) contains inside the vacuum holder the crystal plate with the film heaters and the thermistor arranged on its surfaces. The thermocontroller circuitry can be partly or entirely located inside the resonator volume [3, 4]. Effective thermal insulation of the heated part of the DHR from environment is used to minimize the heating power consumption. For the described circuit the power through the film heaters in the steady state can be expressed as:

$$Ps = Ec\ Is\ Ka\ \Delta Rt / 4R1, \qquad (1)$$

where Ec - voltage on the thermo-sensitive bridge; Ka - the amplification coefficient of the circuit; ΔRt – the bridge misbalance providing the steady heating current Is, R1 – resistance of the bridge resistors.

The power dependence on the thermocontroller parameters and ambient temperature can be found as a derivation of (1):

$$dPs = Is^2 h\ dEc/Ec + Is\ Ka\ Ec\ \alpha\ dTc(4R1), \qquad (2)$$

where α- temperature coefficient of the thermistor resistance; dTc – temperature variation of the thermistor; $\alpha\ dTc = d\Delta Rt$.

As one can see from the expressions the power variations are dependent on the thermistor temperature and on the input voltage changes. The power variations decrease with reduction of steady state heating current and of the amplification coefficient. The input voltage changes are translated directly into the power fluctuations. So effective voltage regulator should be used with the thermo-sensitive bridge to eliminate influence of the voltage supply on the DHR frequency.

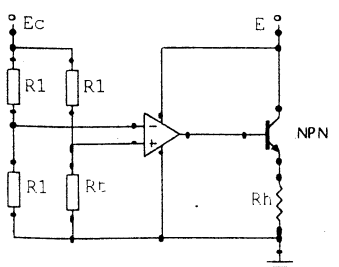


Fig.1.

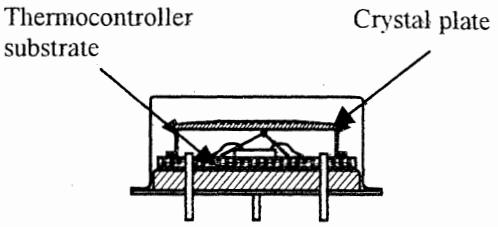


Fig. 2.

The temperature fluctuations in the thermistor result from environmental temperature fluctuations. Electrical network simulating the temperature variations of the crystal versus the ambient temperature deviations is shown in Fig. 3.

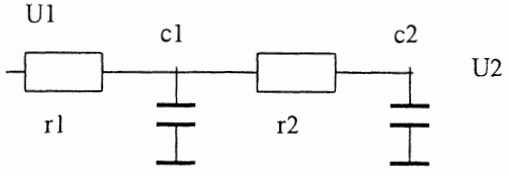


Fig.3.

In this model r1 – is an electrical equivalent of the thermal resistance between the holder and the environment, r2 – is an equivalent of the thermal resistance between the crystal and the holder, c1, c2 – are electrical equivalents of the thermal capacity of the holder and the crystal plate respectively; U1 – simulates ambient temperature variation $dTa(\Omega)$ with frequency Ω; U2- simulates the temperature variation of the crystal $d\mathrm{Tc}(\Omega)$.

From the model one can find the crystal temperature variation $d\mathrm{Tc}(\Omega)$ versus amplitude and frequency of ambient temperature $d\mathrm{Ta}(\Omega)$ changes:

$$d\mathrm{Tc}(\Omega)=d\mathrm{Ta}(\Omega)/\sqrt{(1+\Omega^2\tau1\tau2)^2+\Omega^2(\tau1+\tau2+\tau12)^2} \quad (3)$$

where $\tau1 = c1 * r1$; $\tau2 = c2 * r2$; $\tau12 = c1 * r2$.

Hence from expression (3) the crystal temperature fluctuations decline with the improvement of the effectiveness of the thermal insulation and with the increase of frequency of the temperature variations. Substitution of expression (3) in (2) allows determination of the heating power fluctuations versus ambient temperature variations, the thermocontroller parameters and the DHR thermal properties. Next stage of the work determines frequency changes induced by the power fluctuations in the crystal plate.

3. GENERAL EQUATION FOR STRESS-INDUCED FREQUENCY CHANGES OF THE CRYSTALS

It's known that temperature gradients in the crystal plate produce non-uniform mechanical stress field resulting in frequency shift via variation of the stiffness coefficients. Due to non-uniform distribution of the vibration amplitude over the plate the resulting frequency deviation of the crystal is a function of both fields interaction. To define the frequency shift of the crystal under the stress field let's consider the energy balance equation of the plate vibrating in the thickness-shear mode:

$$\mathrm{Kmax} = \mathrm{Pmax}, \quad (4)$$

Kmax is maximal kinetic energy while the vibration period:

$$\mathrm{Kmax} = \omega \rho \int u^2 (x,z)\sin^2(m\pi y/2h)ds$$

and the maximal potential energy:

$$\mathrm{Pmax}=(m^2/4h^2) \int q\, u^2 (x,z) \cos^2(m\pi y/2h)dv$$

Here u (x, z) – is a distribution of the thickness-shear mode displacements over the crystal plate; ρ - density of quartz, m - overtone number of the vibrating mode, q – effective coefficient of elasticity for the vibrating mode under the stresses; 2h – thickness of the plate.

Substituting expressions for Kmax, Pmax in (3) and taking that $q = qo + dq(x,z)$, and $\omega = \omega o + \Delta\omega$ (where qo – the effective coefficient of elasticity in absence of the stress field; $dq(x,z)$ – variation of the elasticity coefficient under the stress field $\sigma(x,z)$, ωo-resonance frequency of the crystal in the absence of the stress field, $\Delta\omega$ - the frequency change due to the stress field, we get the balance equation for the crystal vibrating under the stress field:

$$\rho(\omega o+\Delta\omega)^2 \int A^2 (x,z)ds = (m/2h)^2 \int (qo + dq(x, z))\, A^2(x, z)\, ds,$$

where A(x,z) – normalized vibrations amplitude:

$$A(x,z)=\exp(-x^2/a^2-z^2/b^2),$$

where a, b – active area sizes along x and z axis of the crystal plate.

Since $(\omega o+\Delta\omega)^2 \approx \omega o^2 + 2\Delta\omega\omega o$, and $\omega o = (\pi m/2h)^2(qo/\rho)$, expression (4) can be written as:

$$\rho(\omega o^2 + 2\Delta\omega\, \omega o) \int A^2(x,z)\, ds = (\pi m/2h)^2 \int dq(x,z)A^2(x,z)ds + \omega o \int A^2(x,z)ds$$

Solving the equation for the frequency shift $\Delta\omega$ yields:

$$\Delta\omega = (\pi\, m/2h\omega o)^2 \int dq(x,z)A^2(x,z)\, ds \,/\, (2\rho \int A^2(x,z)\, ds$$

Dividing the equation by ωo we come to the expression for fractional frequency changes:

$$\Delta\omega/\omega o = \int dq(x,z)\, A^2(x,z)ds \,/\, (2qo \int A^2(x,z)\, ds) \quad (5)$$

The crystal plate can be imagined as a composition of small areas ds with coordinates x, z being under the uniform stresses of $\sigma(x,z)$ value (within ds) which produce the frequency change $\partial\omega/\omega o$ due to the elasticity coefficient variation:

$$\partial\omega/\omega o = \partial q/2qo = K\sigma(x,z)\,\sigma(x,z),$$

where $K\sigma(x,z)$ - coefficient equaled to the frequency change produced by the uniform stresses of $\sigma(x, z)$ along axis x or z. Then equation (5) can be written as:

$$\Delta\omega/\omega = \int K\sigma(x,z)\,\sigma(x,z)\,A^2(x,z)\,ds \,/\, \int A^2(x,z)\,ds \qquad (6)$$

Obtained equation allows calculation of the crystal frequency changes produced by arbitrary field of mechanical stresses $\sigma(x,z)$ which in its turn can be a result of different factors imposed on the crystal, such as: non-uniform plate heating field, acceleration, mounting stress relaxation, etc.

In case of the circular crystal plate under the symmetrical (about the plate center) stresses field equation (6) can be written as

$$\Delta\omega/\omega = K\sigma \int \sigma(r)\,A^2(r)\,dr \,/\, \int A^2(r)\,dr, \qquad (7)$$

where $K\sigma$ – integral frequency-stress coefficient dependent on the crystal cut only. For the case of non-uniformly heated plate the thermal stress field $\sigma(r)$ is a linear function Ψ of the thermal gradient field $\partial T/\partial r$. Thus, equation (6) can be written as:

$$\Delta\omega/\omega = \int K\sigma(x,z)\,\Psi(x,z)\,A^2(x,z)\,ds \,/\, \int A^2(x,z)\,ds \qquad (8)$$

For the similar boundary conditions and the same plate geometry following equation is valid:

$$\Psi 1(x,z)/\Psi 2(x,z) = (\partial T1/\partial x)/(\partial T2/\partial x), \qquad (9)$$

where $\partial T1/\partial x$, $\partial T2/\partial x$ – thermal gradients produced by some temperature fields T1(x) and T2(x).

4. THERMAL GRADIENTS IN THE PLATE AND THEIR INFLUENCE ON THE CRYSTAL FREQUENCY

Typical designs of the DHR crystal plate are shown in fig. 4. The crystal plates can be circular or rectangular with the film heaters deposited on the peripheral. In the rectangular plates the thermal power flows from the heaters into the central part of the plate along one axis only. So, a one-dimensional model can be used for determination of the thermal gradient pattern. The thermal gradient in the circular plate with sufficient accuracy can be estimated as a superposition of two one-dimension solutions in the square plate with thermal flows along **x** and **z** axes (fig. 4 c).

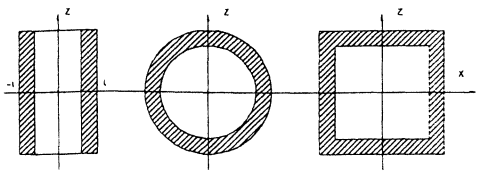


Fig.4.

At ambient temperature noise the power variation in the heater is a sum of periodical functions:

$$\Delta\mathbf{P} = \Sigma\,\Delta Poi \cos(\Omega i\, t),$$

where ΔPoi, Ωi–amplitude and frequency of i-component of the power variations.

For one-dimensional model the temperature field is described by the differential equations:

$$\partial\mathbf{T}/\partial x = R\,\Delta\mathbf{P},\quad -\partial\Delta\mathbf{P}/\partial x = j\Omega C\mathbf{T} \qquad (10)$$

at the boundary conditions:

$$\Delta\mathbf{P} = \tfrac{1}{2}\,\Delta Poi\cos(\Omega i\, t) \text{ at } x = l,$$
$$\Delta\mathbf{P} = -\tfrac{1}{2}\,\Delta Poi\cos(\Omega i\, t) \text{ at } x = -l, \qquad (11)$$

where **T**, $\Delta\mathbf{P}$ – complex temperature and the power variations respectively; R – thermal resistance of the plate equaled to $1/(hl\lambda x)$ (λx-thermal conductivity of quartz crystal along **x**-axis); C = cophl (co - thermal capacity of the plate).

Solution of system (10), (11) about the temperature gradient ($\partial\mathbf{T}/\partial x$) yields:

$$\partial T/\partial x = R\Sigma\Delta Poi\cos(\Omega it)|\sin(z1)/\sin(z2)|, \qquad (12)$$

where $z1 = \alpha(l-x) + j\beta(l-x)$, $z2 = \alpha l + j\beta l$,
$\alpha = \beta = \sqrt{\Omega\,co\rho/2\lambda}$

Calculation of the temperature gradients over a SC-cut plate with 10 mm width and 0.55 mm thickness in dependence on the fluctuation frequency Ω is shown in Fig. 5.

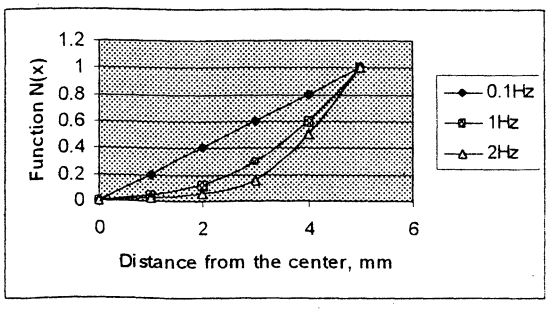


Fig. 5.

. As one can see the temperature gradients reach maximum at the plate edges and fall to zero in the plate center. The gradients decrease with raise of frequency Ω: for low $\Omega<0.1$ Hz the gradients are linearly distributed over the plate while for higher Ω the gradients decrease essentially in the plate center region.

Functions $\Psi(\partial T/\partial x)$ in (8) is the most difficult for determination as being dependent on a variety of the DHR parameters, i.e. the crystal plate cut and geometry, the heater configuration, as well as of the manufacturing imperfections. That makes theoretical definition of the function hardly practical.

To avoid calculation of Ψ we used experimental method of determination of thermodynamic sensitivity of the crystal plates based on measurement of "overshoot" of their frequency during warming-up. The "overshoot" phenomena is well studied at comparatively slow heating speeds - below than 1°C/s. We studied fast heating process at about 10°C/s for SC-cut and modified SC-cut (φ=23°25'; θ=34°) crystals (Fig. 6).

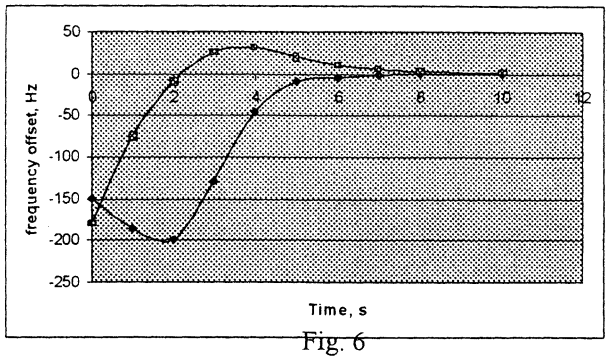


Fig. 6

Fig. 6

As it's seen form the curves sign and amplitude of the "overshoot" are different for these crystals. The "pure" SC-cut crystal "overshoot" is negative with 4-8 ppm amplitude while the modified SC-cut has positive sign with the range of about 2 to 5 ppm. Obviously, "zero-overshoot" cut must be oriented at about 23°00'.

Analysis of the warming-up process reveals that the "overshoot" is caused by thermal gradients between the plate center and its edges (where the thermistor is located) appeared by the moment when temperature of the thermistor reaches a preset value.

Process of heating the crystal plate at the start-up is governed by the differential equation of thermal conductivity:

$$\partial T^2/\partial x^2 = (\rho co/\lambda)\, \partial T/\partial t \qquad (13)$$

at the boundary conditions:

$$\text{at } x = -l \quad \partial T/\partial x = (\rho co/\lambda)\, Po/2,$$
$$\text{at } x = l \quad \partial T/\partial x = -(\rho co/\lambda)\, Po/2, \qquad (14)$$

where Po – start power in the heaters.

Solution of equation (13) for time $t>\rho col^2/\lambda$ (about 2 s for the 3rd overtone 10 MHz crystals) allows determination of the gradient values resulting in the frequency "overshoot":

$$(\partial T/\partial x)ov = Po\,(x/2\lambda hl^2)$$

Then the frequency "overshoot" can be found from (11):

$$(\Delta\omega/\omega)ov = \int K\sigma(x,z)\, \Psi(x, z)ov\, A^2(x, z)\, ds / \int A^2(x,z)\, ds \qquad (15)$$

On the other hand, frequency deviation due to heating power fluctuations can be expressed as:

$$(\Delta\omega/\omega)p = \int K\sigma(x,z)\, \Psi(x, z)p\, A^2(x, z)\, ds / \int A^2(x,z)\, ds, \qquad (16)$$

where $\Psi(x, z)p$ is the function of thermal gradients $(\partial T/\partial x)p$ produced by power fluctuations in the heaters and governed by equation (12).

Taking into account (9) equations (15), (16) along with expressions for $(\partial T/\partial x)p$ and $(\partial T/\partial x)ov$ allow determination of frequency deviations caused by the heating power variations:

$$(\Delta\omega/\omega)p = (\Delta\omega/\omega)ov\, (\Delta Ps/Po) \int l\, N(x)\, A^2(x,z)\, ds / \int x\, A^2(x,z)\, ds, \qquad (17)$$

where $N(x)= |sh(z1)/sh(z2)|$.

Function N(x) at small x (near a center of the plate) can be approximated by a linear function: N(x)=N(a)(x/a), where a – is a size of an active area of the plate. Integrating (17) we obtain simple expression for the frequency fluctuations versus the heating power variations:

$$(\Delta\omega/\omega)p = (\Delta\omega/\omega)ov\, (\Delta Ps/Po)(l/a)N(a) \qquad (18)$$

Substituting equations (2), (3) into (18) one can calculate the frequency fluctuations versus ambient temperature variations:

$$(\Delta\varpi/\varpi)t = (\Delta\varpi/\varpi)ov\, N(a)\, (l/a)\, Is\, Ka\, Ec\, \alpha\, \Delta Ta(\Omega)/ (\Omega^2\, \tau 1\, \tau 2\, 4R1\, Po) \qquad (19)$$

The crystal frequency deviation dependence on the temperature fluctuations frequency was calculated (table 1) for 10 MHz 3d overtone SC-cut DHR with the following properties: thermal insulation resistance 500 K/W; frequency "overshoot" $(\Delta\varpi/\varpi)ov$ = 3 ppm; $\tau 1$ = 70 s; Ka = 20; Ec = 8V; α=1KOhms/°C and ΔTa = 0.1°C).

From the data one can see that noticeable (above 1E-11) DHR frequency deviations take place only at slow ambient temperature variations – below 0.01 Hz. tests of OCXO based on different types of

DHR, however, show noticeable variation of their heating current (within 0.005 - 0.01 mA) resulting in considerable degradation of the STS. The most likely reason for the current fluctuations is noise in the thermocontroller circuitry components, primarily in the thermistor.

Table 1

Temperature fluctuation frequency, Ω, Hz	Crystal frequency deviation, Δf/f
0.01	1E-11
0.1	1E-13
1.0	1E-15

To calculate dependence of the DHR frequency on the heating current fluctuation expression (19) can be transformed by substituting $\Delta Psi = 2Isi\ Rh\ dIs$ and $Po = Io^2Rh$:

$$(\Delta f/f)i = 2(\Delta f/f)ov\ (l/a)\ N(a)\ dIs\ Is/Io^2 \qquad (20)$$

Equation (21) allows prediction of the DHR frequency fluctuations on the basis of the heating current fluctuation measurements regardless of its origin. It was applied for calculation of STS of different types of DHR. The results are shown in Table 2 in comparison with experimental data obtained for the OCXO built using different types of DHR. Denotation of the table:

A. Glass packaged AT-cut DHR, 5th overtone, 10 MHz

B Glass packaged DHR, modified SC-cut, 3rd overtone, 10 MHz

C. TO-8 packaged RT, modified SC-cut, 3rd overtone, 10 MHz.

Table 2.

Type of the DHR design	Steady state power in the heaters, mW	Start-up current/steady state current, mA	Overshoot, ppm	Current variations at 1Hz, mA	Frequency fluctuations, E-11	
					th e	te st
A	60	30/12	3.0	0.005	20	15
B	3.5	70/5	1.0	0.005	0.6	0.8
C	6.0	200/11	3.0	0.01	0.9	1.0

As one can see the theoretical results are close to the experimental data. For the OCXO based on the AT-cut DHR design the STS is 20-30 times worse as compared with that of the SC-cut DHR that results from significant thermodynamic sensitivity of the AT-cut crystals and considerable power dissipated in the heaters.

On the basis of above consideration following means of reduction in the DHR frequency fluctuations caused by the ambient temperature deviations and inherent noise in the thermocontroller can be proposed:

1. Reduction of power consumption of the DHR is the most effective way to improve STS. It effectively shields the crystal from transferring the ambient temperature variations. The result is proportional reduction in heating current fluctuation.
2. The translation of the crystal temperature deviation into the heating power fluctuations can be reduced by reduction of the thermocontroller amplification coefficient, the thermistor coefficient α and the thermo-sensitive bridge input voltage Ec. These methods however are limited by possible degradation of the DHR temperature stability.
4. Thermo-dynamic sensitivity of the crystal plate should be minimal, which demands application of SC-cut plates with optimized configuration of the heaters. Moreover the 5th overtone crystals are preferred over the 3d overtone ones while the later have better STS in comparison with the fundamental mode crystals.
5. Low-noise components of the thermocontroller circuitry should be used to minimize its inherent noise resulting in fluctuations of the heating current.

As concluded above, the most effective way to improve DHR STS is a reduction of the heating power by improvement of the thermal insulation. We've developed ultra-low consumption SC-cut DHR packaged in TO-8 vacuum holder. Its thermal insulation resistance is about 1000 W/K, which results in 5 mA steady state current and thermal power dissipated in the crystal of about 1 mW. The STS of such designs is measured at about 5E-12 per 1s, the Single Side Band Phase Noise Power Density at 1 Hz offset is below –90 dBc/Hz.

REFERENCE

1. B.Long, G.Weaver, «Quartz Crystal Oscillators with Direct Resonator Heating», Proc. of the 45 AFCS, 1991, pp. 384-392.
2. I.Abramzon, «OCXO Design Using Composite-Heating of the Crystal Resonator», IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, Vol.41, No.2, 1994, pp.284-287.
3. I.Abramzon, R. Boroditsky, D. Cocuzzi, «Miniature OCXOs Using DHR Technology», Proc. of 1997 IEEE AFCS, pp. 943-947.

APPROXIMATION AN ELECTROSTATIC END-EFFECT IN IDT TOPOLOGY USING DIPOLE-BASIS

YURI ABRAMOV
Soliton, Har Hatzofim 24/4, Holon 58492, Israel

WADIM DUNZOW
VTI Telefilter, Potsdamer Str. 18, D-14513 Teltow, Germany

Abstract

The electrostatic charge distribution in the dipole-basis is presented. The benefit properties of this basic solution are an electrostatic end-effect reduction and a possibility of using the transversal filters theory and the linear programming in a synthesis problem.

Introduction

The exact solution for a two-dimensional electrostatic field of a finite inter-digital transducer (IDT) [1,2] was modified for the effective numerical calculation in [3,4]. However, this solution is not suitable for using the transversal filters theory and the linear programming in the synthesis problem.

In order to use linear programming for the IDT topology calculation, the transform function of IDT was performed as multiplication of an element factor and an array factor [5,6,7,8,9,10]. The widespread interpretation of the weights of the array factor as an electrodes overlaps becomes incorrect for SAW filters with high requirements to a shape factor, a pass band, a delay variation and a stop-band attenuation, especially for tilted IDT topologies because of the uncompensated electrostatic end-effect [6,11,12]. The electrostatic field model with the given charges on electrodes of the infinite periodic grating [13] is more appropriate for the real IDT characteristics synthesis and prediction. In this paper the interpretation of the weights of the array factor as an electrostatic field displacement flux values through the gaps between the neighbor electrodes as well as the definition of the respective element factor are suggested. In this case the end-effect is approximated much better and calculated easier. The definitions of the voltages in the dipole-basis and of the element factor in frequency domain are presented.

Idea

Let us consider an infinite periodical electrode sequence, where electrodes are enumerated from $-\infty$ to $+\infty$. All the electrodes excluding the zeroth one and the first one are isolated. The zeroth and the first electrodes are connected to an external source of static electrical voltage U_o (**Fig. 1.**). So the zeroth and the first electrodes are charged and other electrodes keep zero charge. Let the zeroth electrode's charge be equal to **+1** and the first electrode's charge be equal to **–1**. Let us consider sum of the electrostatic charges on the electrodes numbered from $-\infty$ to n. This sum is equal to **+1** for $n=0$, and is equal to zero for $n\neq 0$. So, we consider this sum of the electrostatic charges as a basic function and this structure as a basic structure. This basic structure was described in [13]. Defining the IDT response through superposition of such basic structures characteristics, we approximate the field in the uncharged free space outside the IDT by the field of isolated electrodes. Despite the fact that there are no isolated electrodes outside the real IDT structure, this

approach approximates the electrostatic end-effect with high precision and at the same time allows to use the transversal filters theory and the linear programming in the synthesis problem.

The problem formalization

To formalize the electrostatic problem we admit the following assumptions.

- The infinite periodical sequence of the parallel electrodes is placed in (X, Z) - plane. The intervals occupied by electrodes along X-axis are (a_n,b_n). The electrode's width equals w, $w= b_n-a_n$, and the grating period equals p, $p= b_n-b_{n-1}$.
- The upper half-plane is the vacuum with the dielectric permittivity ε_o and the lower half-plane is a dielectric material with the effective dielectric permittivity ε_p.
- The electrodes are infinitely long in Z-direction, so the electrostatic problem becomes two-dimensional in (X, Y) - plane.
- The electrodes have zero thickness and zero electrical resistance.
- The electrodes are enumerated from -∞ to +∞.
- All the electrodes excluding the zeroth one and the first one are isolated.
- The electrostatic voltage applied between the zeroth and the first electrodes is such that the zeroth one is charged with **+1** and the first one is charged with **–1**.

We call the structure formalized with such assumptions as "dipole-basis" (**Figure-1**).

The problem is to find a function of the charge distribution density on the electrodes in the dipole-basis.

In other words, the problem is to find the two-dimensional electrostatic intensity function satisfying the boundary conditions along axis X and the associated potential satisfying Laplace's equation in the upper half-space and in the lower half-space. By Gauss's law the free charge distribution density is equal to the jump of the electrostatic field displacement discontinuity on the metal space. The free charge density must be zero in the gaps between the electrodes, the potential on each electrode must be uniform, and the induced charge on electrodes must satisfy the dipole-basis definition. I.e. the summary free charge must be equal to **+1** on the zeroth electrode, to **–1** on the first one and to zero on the others. The condition of zero summary free charge is satisfied for the whole dipole-basis.

Electrostatic solution.

The charge distribution density $\rho_o(x)$ satisfying the dipole-basis boundary conditions is written by the following equation:

$$\rho_o(x) = 0, \qquad x \in (b_m, a_{m+1}) \qquad (1a)$$

$$\rho_0(x) = \frac{(-1)^m}{p} \frac{(\varepsilon_0 + \varepsilon_p)\sqrt{2}\ \Gamma(x,\theta)}{\sqrt{Cos(\pi - \theta) - Cos\frac{2\pi x}{p}}}, \qquad x \in (a_m, b_m) \qquad (1b)$$

where :

m - is electrode number,

$\theta = \pi\eta$,

η is a metallization ratio, $\eta = w/p$,

w is an electrode width,

p is a grating period,

$$\Gamma(x,\theta) = \int_{-1}^{+1} \frac{Cos\frac{\pi t}{2} Cos\frac{\pi x t}{p}}{P_{\frac{t-1}{2}}(Cos\theta)} dt$$

$P_s(x)$ - is the Legendre function.

The associated potential $\varphi(m)$ on the m-th electrode in the dipole-basis structure is equal to:

$$\varphi(m) = -\int_0^1 \frac{P_{-s}(-Cos\theta)}{P_{-s}(Cos\theta)} Sin[(2m-1)\pi s]ds \quad (2)$$

The integral in (2) may be simplified for the metallization ratio $\eta = 0.5$, i.e.

$$\varphi(m) = -\frac{1}{\pi(m-\frac{1}{2})} \quad (2.1)$$

The voltage $U(m)$ in the gap between the m-th and the (m+1)-th electrodes in the dipole-basis structure is:

$$U(m) = -2\int_0^1 \frac{Sin(\pi s)Cos(2\pi ms)P_{-s}(-Cos\theta)}{P_{-s}(Cos\theta)} ds \quad (3)$$

The integral in (3) may be simplified for the metallization ratio $\eta = 0.5$, i.e.

$$U(m) = \frac{4}{\pi(4m^2-1)} \quad (3.1)$$

Element factor

Let now the applied voltage Uo between the zeroth and the first electrodes in the dipole-basis be oscillated with cycle frequency ω, and there is a piezoelectric substrate instead of dielectric substrate in the lower half space. Considering the oscillated electrical charges on electrodes as sources of elementary SAW with effective wave number κ, we obtain the normalized amplitude of the total SAW generated with the dipole-basis structure as Fourier-transform of the basic charge distribution $\rho_o(x)$.

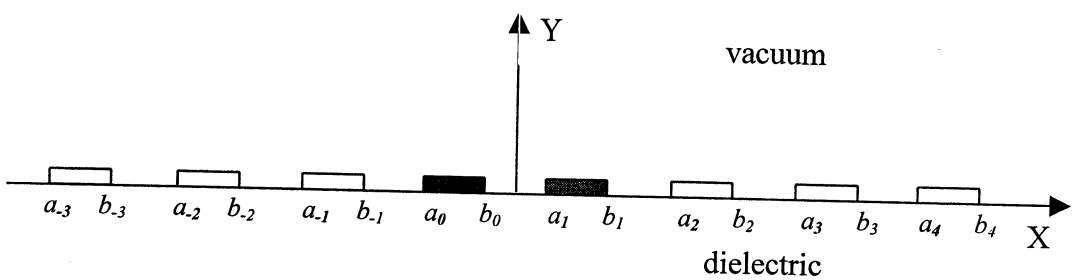
So, the dipole-basis element factor $E_0(\kappa)$ is defined as Fourier transform of the charge density $\rho_o(x)$ in κ-domain:

$$E_0(\kappa) = j\frac{2P_n(\cos\theta)Sin(\frac{p\kappa}{2})}{P_{n-s}(\cos\theta)} \quad (4)$$

where

n is integer, $\quad \frac{2\pi}{p}n \le \kappa \le \frac{2\pi}{p}(n+1)$

$S = \frac{p\kappa}{2\pi}$



Charge on electrode	0	0	+1	–1	0	0	0
$\sum_{m=-\infty}^{n} Charge(m)$	0	0	1	0	0	0	0

Figure 1. The dipole-basis structure.
The electrode with number zero is charged with **+1** and the electrode with number one is charged with **-1**. Other electrodes are isolated and keep zero charge. The sum of the electrostatic charges on the electrodes numbered from $-\infty$ to n has physical meaning of the electrostatic field displacement flux through the gap between n-th and $(n+1)$-th electrodes. This electrostatic field displacement flux through the n-th gap is equal to **+1** for $n=0$ and is equal to zero for $n \neq 0$.

Transform function

An arbitrary sequence $\{A(n)\}$ of an electrostatic field displacement flux values through the gaps between the n-th and the $(n+1)$-th electrodes is obtained by the superposition of the dipole-basis electrostatic fields shifted by np and weighted with $A(n)$. So, the charge distribution density for such an IDT is expressed by superposition:

$$\rho(x) = \sum_{n} A(n)\rho_0(x - np) \qquad (5)$$

The frequency spectrum of the SAW acoustic potential $\varphi_a(\omega)$ on the left port of the radiated transducer is expressed through the Fourier transform of the charge distribution density $\rho(x)$ as a multiplication of the element factor and the array factor [6], i.e.

$$\varphi_a(\omega) = j\Gamma_s E_0(\kappa) \sum_{n=1}^{N} A(n)\exp(-j\kappa np) \quad (6)$$

where

Γ_s is a constant, that depends on the piezoelectric material and cut; this factor is a measure of the piezoelectric coupling of the substrate material.

j is an imaginary unit;

ω is a cycle frequency;

κ is a wave number, $\kappa = \omega / V$;

V is an effective SAW velocity;

p is a grating period;

N is a amount of overlaps of electrodes in an IDT structure;

n is a number of grating period (n-th electrode and gap);

$E_0(\kappa)$ is an element factor;

$A(n)$ is an n-th weight of an array factor.

If we consider the SAW device consisting of two SAW transducers, one of which operates as a radiator and the other operates as a receiver, the device response is proportional to the multiplication of the characteristic SAW IDT responses $H_1(\omega)H_2(\omega)$. To analyze devices, it is convenient to define $H_1(\omega)$ and $H_2(\omega)$ so, that both of them would be written in the similar way $H(\omega)$ [6]. In our case the $H(\omega)$ is written using the expression (5) for the charge distribution density, i.e.

$$H(\omega) = \frac{\sqrt{W\omega\Gamma_s}}{U_t} E_0(\kappa) \sum_{n=1}^{N} A(n)\exp(-j\kappa np) \quad (7)$$

where

U_t is a voltage applied to the SAW transducer.

W is a width of the propagated SAW front.

This definition of the SAW IDT frequency response $H(\omega)$ is suitable for using the transversal filters theory and the linear programming for the IDT topology synthesis. In this case the electrostatic end-effect of an implemented IDT topology is negligibly small.

Conclusion

The dipole-basis for the regular electrode grating is presented. The simple expressions for the calculation of the associated potentials on electrodes and voltages on gaps between electrodes in the dipole-basis structure are found. The electrostatic problem for the regular electrode grating with an arbitrary sequence of an electrostatic field displacement flux values through the gaps between electrodes is obtained by the superposition of the dipole-basis electrostatic fields. The usage of the basic solution for the approximation of an electrostatic end-effect in a finite IDT topology allows to calculate charge distribution, voltages and frequency characteristics with high precision. At the same time the frequency response of the SAW IDT is

written in a transversal filter form, so this approach allows to use the transversal filters theory and the linear programming in the synthesis problem.

References

1. **L.L.Goryshnik, S.N.Kondratiev,** - Radiotekh. Electron. **8,** 1719 (1974) [English transl:Sov. Phys.-Radio Eng. Electron.,vol.19, No. **8,** pp.1719-1728 (1974)]
2. **L.L.Goryshnik, S.N.Kondratiev,** - Radiotekh. Electron., No. **1,** 151 (1978) [English transl:Sov. Phys.-Radio Eng. Electron., No. **1,** 151 (1978)]
3. **S.P.Stepura,** *Calculation of the SAW transducers electrical characteristics* – Electron. Techn. Ser.10, Microelectronic devices, issue 5(29)-1981,pp.53-56
4. **S.N.Kondratiev, J.A.Khabarov,** - Electron.Tech. **3(90),** 27 (1990)
5. **H.Engan,** *Excitation of elastic surface waves by spatial harmonics of interdigital transducers,* – IEEE Trans. Electron Devices, vol. ED-16, No. 12, pp. 1014-1017, (1969)
6. **D.Morgan,** *Surface-wave devices for signal processing, Chapter-4,* Elsevier, Amsterdam, 1985.
7. **S.Datta, B.J.Hunsinger, and D.C.Malocha,** *A generalized model for periodic transducers with arbitrary voltages* – IEEE Trans. Sonics and Ultrason., vol. SU-26, No.3, pp.235-242, (1979).
8. **S.Datta and B.J.Hunsinger,** *Element factor for periodic transducers* – IEEE Trans. **SU-27,** 42-44 (1980)
9. **R.C.Peach,** *A general approach to the electrostatic problem of the SAW IDT* – IEEE Trans. **SU-28,** 96-105 (1981)
10. **A.Rukhlenko,** *Charge Distribution and Capacitance Calculation for Generalized Periodic SAW Transducers Using Floquet's Technique* – IEEE Ultrason. Symp., 1994, pp.325-329
11. **C.S.Hartmann and B.G.Secrest,** *End effects in interdigital surface wave transducers* – IEEE Ultrason. Symposium, 1972, pp.413-416
12. **H.Skeie and A.Ronnekliev,** *Electrostatic neighbour and end effects in weighted surface wave transducers* – IEEE Ultrasonics Symposium, 1976, pp.540-542.
13. **Yu.M.Abramov, V.P.Dunzov, and V.S.Orlov,** *Electrostatic Field Model with the Given Charges on Electrodes of the Infinit Periodic Grating* – Theses of the Report of the IV School-Seminar "Acoustic Electronic Devices". USSR Ministry of Communication, Moscow-1991, p.24

COMBINED OPTIMIZATION FOR THE SAW FILTER DESIGN

YURI ABRAMOV
Soliton, Har Hatzofim 24/4, Holon 58492, Israel

Absrtact.

A two-dimensional synthesis of filter characteristics considering an external circuit influence (50Ohm load), when the insertion loss is one of the goal parameters, is presented. The problem is formalized as a non-linear programming problem.

Introduction.

The SAW design is considered as a triple problem: optimization, implementation and prediction. That means the implementation with the predictable characteristics is a criterion of the combined optimization procedure. So, the optimization procedure

- tries to reduce the amount of small overlaps with behavior that is hard to predict. Such effects as diffraction and distortions around apodisation gaps along electrodes might be ignored, when there are few small overlaps in the apodised inter-digital transducers (IDT).
- corrects the electrostatic end-effect, especially for the tilted topology, constructing the non-symmetric geometry of guard electrodes;
- chooses an appropriate (i.e. the most predictable and physically feasible) pair of functions for IDT-1 and IDT-2 for the goal frequency characteristics decomposition;
- takes into account a behavior of the IDT complex acoustic conductivity and an external circuit influence (for example, 50 Ohm load); the insertion loss is one of the goal parameters.

The apodisation functions and apertures of IDT-1 and IDT-2 are the output parameters of the optimization procedure. In order to ignore the SAW velocity dispersion, an appropriate electrode thickness is chosen.

Goal Function.

The goal function for the main signal of desired insertion loss is used for the combined optimization procedure. It takes into consideration the external circuit influence. The goal frequency spectrum $A(\omega)$ of the SAW filter loaded with external circuits, is written in form [1]:

$$A(\omega) = \frac{-H_{sc} \cdot Y_c}{(Y_1 + Y_{c1})(Y_2 + Y_{c2})} \qquad (1)$$

where

- the function H_{sc} is a short-circuit response defined as[1]

$$H_{sc} = H_1(\omega) H_2(\omega) S_{14} \exp(\omega T_{12}) \qquad (2)$$

$H_1(\ )$ and $H_2(\ )$ are the transform functions of the IDT-1 and IDT-2 respectively;
S_{14} is the multi-strip coupler (MSC) characteristics;
T_{12} is the SAW delay between IDT-1 and IDT-2;

- Functions Y_1 and Y_2 are the complex acoustical conductivity of the IDT-1 and IDT-2 respectively,

$$Y_1 = G_{a1}(\omega) + j(B_{a1}(\omega) + \omega C_1)$$

$$Y_2 = G_{a2}(\omega) + j(B_{a2}(\omega) + \omega C_2)$$

C_1 and C_2 are the static capacity of the IDT-1 and IDT-2 respectively;

- Functions Y_{c1} and Y_{c2} are the complex conductivity of the external circuits connected to IDT-1 and IDT-2 respectively.

Optimal Decomposition.

If the transition band and near stop band are the most critical areas to predict, we need to design both IDT characteristics having the same pass bands and the same transition bands. I.e., for example, to obtain the stop band of 15.00MHz on the level 50dB, both of IDT-1 & IDT-2's stop bands of 15.00MHz on the level 25dB should be built. In this case the transition area characteristics are the most predicted.

Topology Implementation.

It is assumed, that the goal topology structure meets the following conditions.

- The filter consists of two IDTs and MSC.
- Both of IDT are of apodised structures.
- Apodisation functions have few small overlaps, so the diffraction and distortions along electrodes might be ignored.
- In order to reduce RF leakage, the apodisation functions have a big number of long grounded electrodes near the MSC.
- The tilted topologies have the benefit property. They have a smoothed acoustical conductivity. However, this feature is not a criterion of the synthesis. The desired smoothing of the acoustical conductivity is achieved by enlarging of the electrode overlaps. Moreover, the criterion of smooth multiplication $(Y_1+Y_{c1})(Y_2+Y_{c2})$ is used in the design.
- If we consider the anti-symmetrical structures, the big number of long hot electrodes is placed on the opposite side of each IDT. It may lead to big parasitic charge distribution end-effect. In order to reduce this unwanted effect, the IDT topology function has a non-symmetric component that enlarges length of grounded electrodes in both sides of the IDT.

Figure-1 shows schematically the topology implementation of SAW filter built using MSC. The apodisation functions for IDT-1 and IDT-2 implementations are performed on Figure 2. Special apodisation for electrodes, which are coated with glue, is used in order to reduce an electrostatic end-effect.

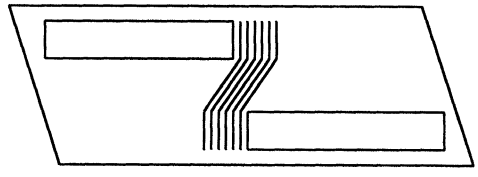


Figure 1. Schematic SAW filter topology designed using two apodised transducers and MSC.

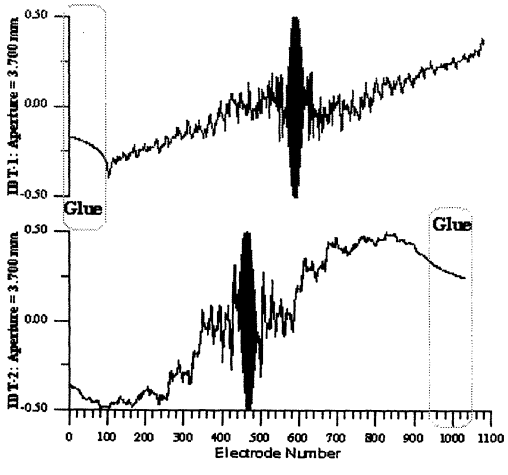


Figure 2. Apodisation functions for IDT-1 and IDT-2 implementation.

SAW velocity.

The additional problem of the development of the vestigial transition band filters, which are built with optimal decomposition, is the SAW velocity dispersion. Analysis model is in need of account a dependency of SAW velocity on electrical load (i.e. on apodisation function) and frequency. This effect becomes essential for

narrow transition zones and band pass more than 15% of the center frequency. This effect might be ignored when both of the apodisation functions have few small overlaps and thickness of electrodes is chosen so that the SAW velocity is not changed in the pass band frequency area.

Step by Step Algorithm.

The step-by-step algorithm may be separated conditionally on the stages of following meaning.

- The two-dimensional synthesis (or weighting with decomposed weighting function)[2] is used. This optimization is based on the criterion of the IDT characteristics predictability (i.e. optimal decomposition and topology implementation conditions).
- The two-dimensional synthesis for $(Y_1+Y_{c1})(Y_2+Y_{c2})$ is used (in particular case, $Y_{c1}=Y_{c2}=1/R$, $R=50$ Ohm).
- The procedure, that takes into consideration the external circuit influence (***50*** Ohm loading), where insertion loss also is one of goal parameters, finds optimal aperture and corrects the apodisation function.

Results.

One of the principal benefits of using of the combined optimization procedure is the fact, the desired frequency characteristics may be reached without additional tuning. And if, nevertheless, we would like to use external circuit tuning, the pulsation in the frequency characteristics band pass is not increased dramatically because the acoustic conductivity as well as multiplication $(Y_1+Y_{c1})(Y_2+Y_{c2})$ is smoothed.

The SAW filter with the relative pass-band ***B*(-3dB)/*Fo*** of about **14.5%** and insertion loss of **24.5dB** was accomplished using this technique. The frequency characteristics shape factor ***B*(-40dB)/*B*(-3)** equals **1.065**, peak-to-peak pass band pulsation is of **0.7dB** and delay variation is of **30nsec**, stop band attenuation is of **56dB**. The pulse characteristics is predicted down to **-60dB** level. (All the measurements were performed without additional circuits tuning.)

The calculated frequency characteristics of this SAW filter are performed on the figure-3, figure-4, figure-5, figure-6, figure-7. The group delay variation and the amplitude pulsation in pass band is shown on Figure-4. Big unwanted signal in the far zone from the pass band (Figure-5) is the result of the overlaps enlarging procedure. So, the overlaps enlarging

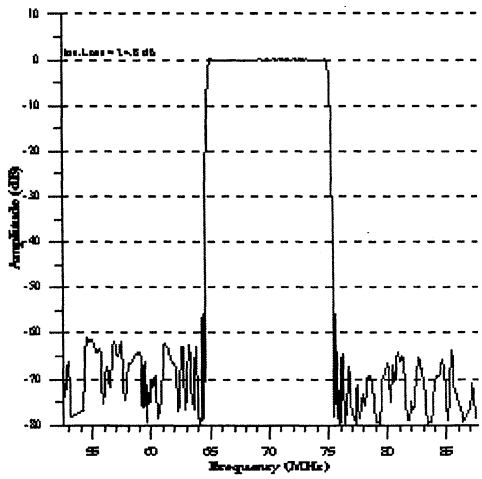


Figure 3. SAW Filter characteristics in the 50% area span.

is limited with specified requirements in the far frequency area. Detailed shape form is shown on Figure-6. Figure-7 demonstrates the result of the two-dimensional optimization of the multiplication $(Y_1+Y_{c1})(Y_2+Y_{c2})$, where the criterion of this optimization was the smoothing of the pulsation in the pass band. This is a principal feature of the combined optimization procedure.

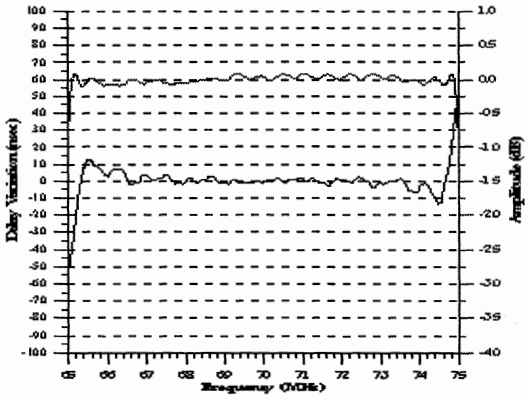


Figure 4. The delay variation and amplitude pulsation in the pass band.

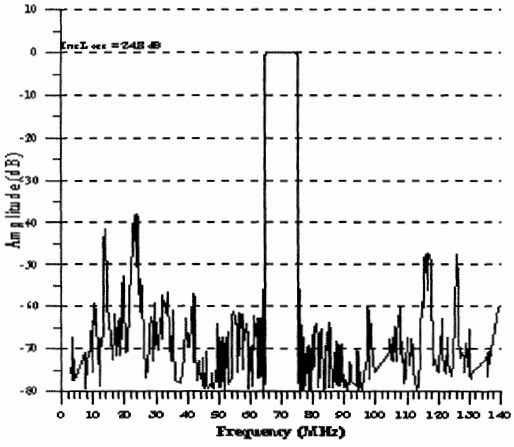


Figure 5. The SAW filter frequency characteristics in the 200% area span.

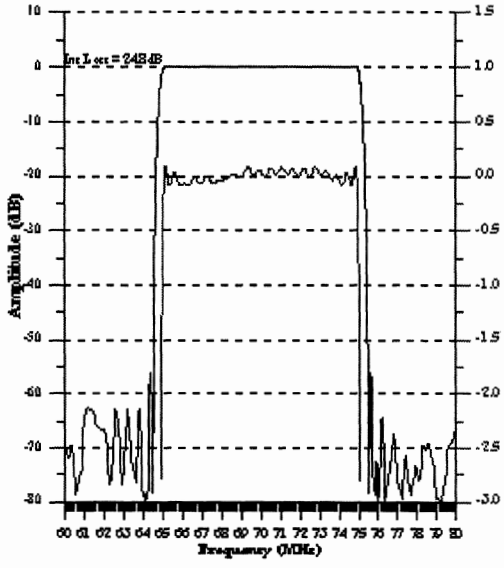


Figure 6. The SAW filter shape form.

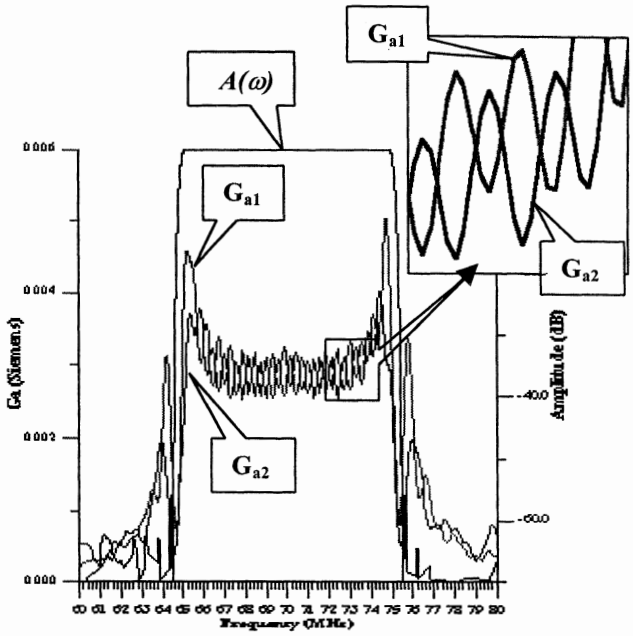


Figure 7. G_{a1} and G_{a2} are the real parts of acoustical conductivity for IDT-1 and IDT-2. $A(\omega)$ is the frequency spectrum of the SAW filter loaded with external resistance of 50 Ohm. Inter-behavior of G_{a1} and G_{a2} is the result of the two-dimensional optimization procedure using the band pass pulsation smoothing as a criterion to approach.

Conclusions.

This optimization procedure is a step-by-step algorithm of sequential approach to the goal function. Usually these calculations take about 500 hours (i.e. three weeks) running on PC based on the Intel Pentium-III 450MHz processor. So, this technique may be successfully used for commercial design of SAW filters with high requirements to the shape factor, pass band, delay variation and stop-band attenuation.

Acknowledges.

The author is grateful to Dr. B.Wall and V.Duntzov (TELEFILTER) for the productive cooperation.

References.

1. **D.P.Morgan,** *Surface-wave devices for signal processing*, Elsevier, Amsterdam, 1985.
2. **C.Ruppel, W.Ehrmann-Falkernau, H.R.Stocker and W.Mader,** *A Design for SAW Filters with Multi-Strip Couplers* – IEEE Ultrason. Symposium – 1984, pp.13-17.

NEW EQUIVALENT CIRCUIT MODEL ANALYSIS FOR SAW RESONATORS IN UHF RANGE

JUN YAMADA

Semiconductors & Integrated Circuits Group, Hitachi Ltd.,
6-2, Otemachi 2-Chome, Chiyoda-ku, Tokyo, 100-0004 Japan
Tel:+81-3-5201-5234, Fax:+81-3-3270-7827
E-mail:yamadaju@denshi.head.hitachi.co.jp

ABSTRACT

This paper describes a new equivalent circuit model analysis which holds stored energy effect in the transit angle for UHF SAW resonators, and re-examinations of circuit parameters of reflection factor and stored energy effect by experiments in UHF range. Then the new equivalent circuit model is verified by experiments.

1. INTRODUCTION

Many researches and developments on Surface Acoustic Wave devices (SAW devices) have been performed to achieve better frequency characteristics and miniaturization of electric equipments. Especially SAW resonators are now widely adopted as narrow bandpass filters, like ladder filters or IIDT filters for RF and IF bands in Mobile Telecommunication Systems.[1][2].

However many subjects have been remained for practical design of UHF SAW resonators, and they are mainly caused by their narrow electrode-width. The electrodes are generally fabricated thick to avoid increase of insertion loss, which is due to finger resistance[3]. And the conventional equivalent circuit model for VHF SAW resonators, which includes the stored energy effect as a shunt susceptance in the transmission line, has been found in such fabrication case that the model is not sufficient for experimental results in UHF range.

This paper describes a new equivalent circuit model analysis which holds stored energy effect in the transit angle for UHF SAW resonators, and re-examination of reflection factor and stored energy effect by experiments in UHF range. Then the new circuit model is verified by experiments.

2. NEW CIRCUIT MODEL

2.1 Admittance Matrix of SAW resonators

An electrode constitution for a two-port SAW resonators is shown in Figure 1. Two grating reflectors are located on the outer sides of the input and output interdigital transducers, and high resonant Q can be obtained due to confine surface wave energy at the intervals of these grating reflectors. Then resonator design has to consider the reflection factor in each electrode of gratings and transducer.

At first, total circuit model is studied, which is shown in Figure 2. A two-port admittance matrix (Y_{Gr}) of the grating is expressed as follows

$$(Y_{Gr}) = \frac{1}{B} \begin{pmatrix} D & -1 \\ -1 & A \end{pmatrix} \quad (1)$$

Here A, B and D are F parameters of the grating, which are obtained from cascade connections of each F matrix. (Fn) and N are F matrix of the n-th electrode and total electrode number in the grating respectively.

$$(F) = \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \prod_{n=1}^{N} (F_n) \quad (2)$$

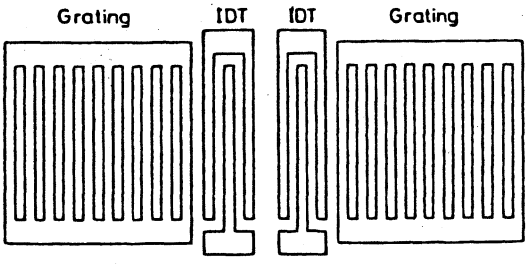


Fig. 1 Electrode constitution of a two-port SAW resonator

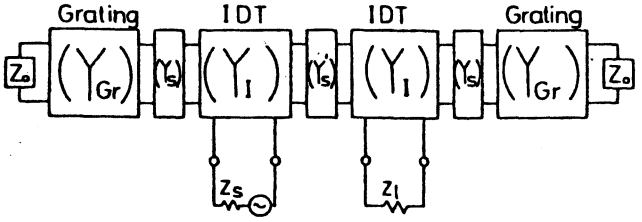


Fig. 2 Equivalent circuit model of a two-port SAW resonator

Admittance matrix (Ys) of free space between the transducer and the grating is expressed as follows.

$$(Y_s) = \frac{1}{jZ_o \sin\theta_s} \begin{pmatrix} \cos\theta_s & -1 \\ -1 & \cos\theta_s \end{pmatrix} \quad (3)$$

$$\theta_s = 2\pi f L_s / v_o \quad (4)$$

Here θs ,Ls and vo are the transit angle , the space distance and surface wave phase velocity in the free surface respectively.

Three-port admittance matrix (Y_I) of the interdigital transducer has been already obtained[4]. And transit angle θs has to be considered on propagation loss α [5], then it is replaced as follows.

$$\theta_o = \theta_s - j\alpha L_s \quad (5)$$

Besides, bulk wave conversion loss μ [6] is added to transit angle on the metallized surface.

$$\theta_m = \theta_o - j\mu (h/\lambda)^2 \quad (6)$$

Here, h and λ are electrode thickness and wavelength of surface acoustic wave respectively. And conductive loss caused by electrode resistance is considered as a lumped finger resistance Rf in the equivalent circuit model. Then the frequency characteristics S_{12} is obtained by the whole two-port admittance matrix (Y) of the resonator with source impedance Zs and load impedance Zl.

$$(Y) = \begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{pmatrix} \quad (7)$$

$$S_{12} = \frac{-2\sqrt{Re(Z_s)Re(Z_l)}\, y_{12}}{\Delta} \quad (8)$$

$$\Delta = (1+Z_s y_{11})(1+Z_l y_{22}) + Z_s Z_l y_{12} y_{21} \quad (9)$$

2.2 Stored Energy Effect

Reflection and stored energy factors are key design parameters inherent to SAW resonators. Relection is caused by characteristics impedance discontinuities of surface wave propagation path. Reflection factor ε is defined by free surface impedance Zo and metallized surface impedance Zm [7].

$$\varepsilon = Z_m / Z_o - 1 \quad (10)$$

And stored energy effect, caused temporarity at step discontinuities on the grating reflectors and transducers, brings about wave velocity decrease for fundamental operation. The factor B is defined by the resonant frequency fp and its frequency shift Δf [8].

$$B = \pi \cdot \Delta f / f_p \quad (11)$$

At the first stage, I calculated the frequency characteristics with the conventional equivalent circuit model [8] and the design parameters [9][10] in the VHF range, but it could not explain the experimental results well. Therefore these parameters are experimentally looked over again in the UHF range.

Many samples with various electrode thickness are examined. Then the parameters are determined by the following new equivalent circuit model, because in the conventional model I could not obtain suitable values of the parameters which satisfy stopband-width and stopband-frequency of frequency characteristics.

The new design parameters are shown in Figures 3 and 4. As to reflection factors both previous and new values agree well in the range of thin electrodes, but they differ greatly each other in the range of thick electrodes. The reason is considered that the previous values were calculated by neglecting higher order terms on reflection, moreover cross sections of examined samples with thick electrodes are trapezoidal.

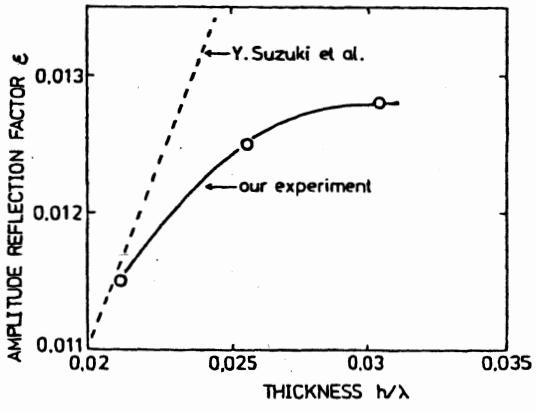


Fig. 3 Influence of electrode thickness on reflection

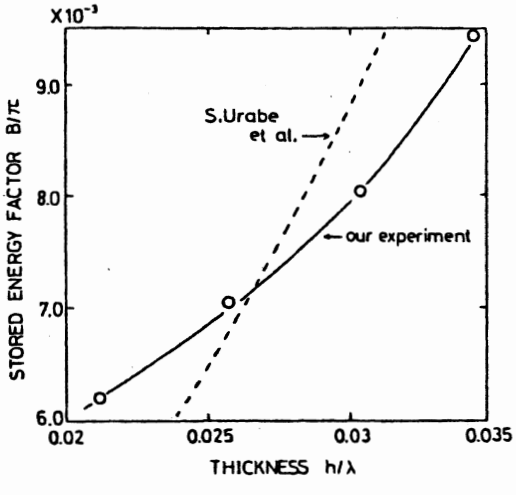


Fig. 4 Influence of electrode thickness on stored energy effect

Meanwhile on the stored energy factor, the new values relatively agree with the previous values after consideration of the measuring accuracy ($\pm 2\times10^{-3}$). This stored energy effect had been introduced before as a shunt susceptance B in the transmission line model [8], which is shown in Figure 5. This model has been available when the electrodes are relatively thin and the stored energy factor B is smaller than the reflection factor, then in such a case previou calculation explained experimental results well. But another reflection, which arises at the shunt susceptance, is added to reflection caused by characteristics impedance discontinuity when the electrodes are relatively thick in UHF SAW resonators.

So a new equivalent circuit model whose transit angle holds the store energy effect has been investigated, which are shown in Figure 6 for the grating reflectors and in Figure 7 for the interdigital transducers.

This model can avoid the reflection from the shunt susceptance and satisfies the resonant frequency shift and stopband characteristics, at that time the reflection from impedance discontinuities has been only considered.

The n-th F matrix of the new model (Fn) is expressed as follows.

$$(F_n) = \begin{pmatrix} \cos\theta'_m & jZ_m\sin\theta'_m \\ j\sin\theta'_m/Z_m & \cos\theta'_m \end{pmatrix} \cdot \begin{pmatrix} \cos\theta_o & jZ_o\sin\theta_o \\ j\sin\theta_o/Z_o & \cos\theta_o \end{pmatrix} \quad (12)$$

$$\theta_o = 2\pi f L / v_o \quad (13)$$

$$\theta'_m = \theta_m + B = 2\pi f L_m + B \quad (14)$$

Here, L and Lm are the length of free surface and electrode width. Besides the lumped finger resistance Rf is considered at the electrical ports in the interdigital transducers.

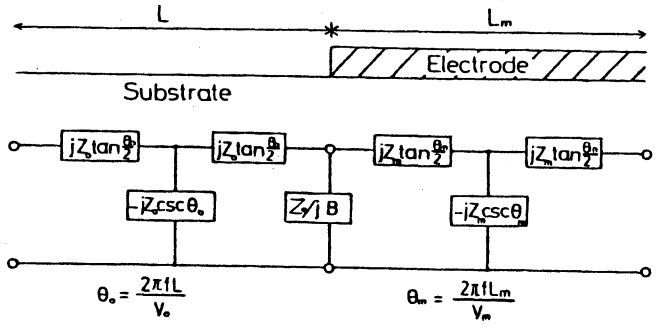


Fig. 5 Conventional circuit model
(grating reflector)

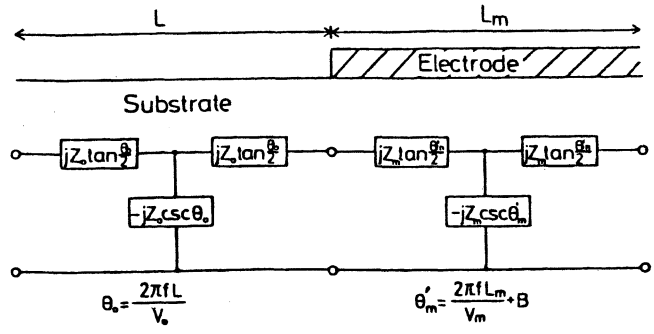


Fig. 6 New circuit model
(grating reflector)

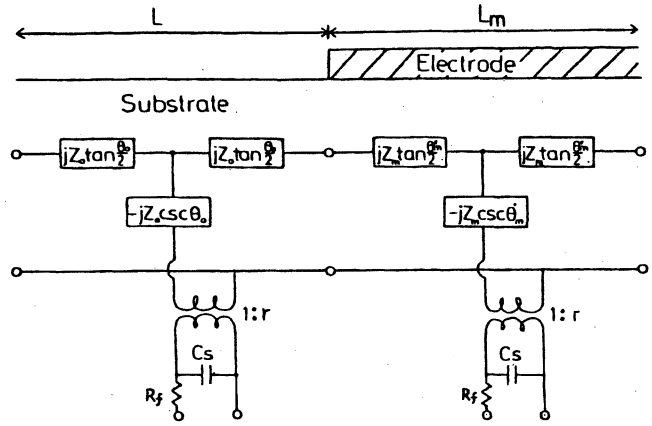


Fig. 7 New circuit model
(interdigital transducer)

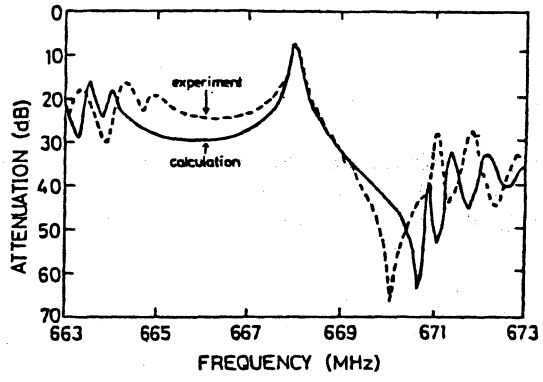
2.3 Verification of New Model

A UHF SAW resonator for the CATV 2nd local oscillator, whose resonant frequency is 668 MHz, is studied for verification of the new equivalent circuit model. The 60 pairs unapodized interdigital transducers and the 500 finger electrodes for grating reflectors are constituted on ST-X quartz. These electrodes are fabricated with thickness of 104 nm and metallization ratio of 45 %. The efficacy of the new circuit model is verified in Figure 8(a) and 8(b), which compare calucated and experimental values. The new model can explain not only the details of passband characteristics but resonant Q, stopband frequency and outband suppression.

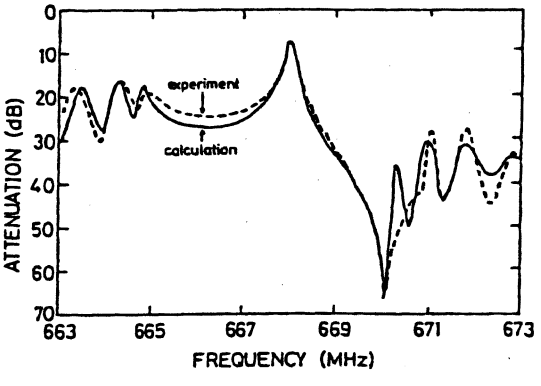
3. CONCLUSION

A new equivalent circuit model for UHF SAW resonators is studied. Reflection and stored energy factors are experimentally looked over again. And it is found that reflection factors in UHF range differ from reflection factors in VHF range, particular when the electrodes are thick.

The new equivalent circuit model, whose transit angle holds stored energy effect, has been proposed, and it is verified by experiments. The new model is adopted for the design of the UHF SAW resonator at 668 MHz in CATV 2nd local oscillator.



(a) with conventional circuit model



(b) with new circuit model

Fig. 8 Comparison between calculated and experimental values

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] J.Yamada, "Design and Fabrication Technology of Low Loss and High Frequency SAW Devices for Mobile Radio Phones", Proc. 1999 Joint Meeting EFTF and IEEE IFCS, pp.926-931, 1999.

[3] J.Yamada, A.Yuhara, T.Shiba and T.Toyama, "High Frequency SAW Filter Stretched Over Non-Overlapped Electrodes for Satellite Broadcast Receivers", IEEE Traans., Cons. Electr., CE-34, No.4, pp. 673-679, 1985

[4] W.R.Smith, H.M.Gerard, J.H.Collins, T.M.Reeder and H.J.Shaw, "Analysis of Interdigital Surface Wave Transducers by Use of an Equivalent Circuit Model", IEEE Trans., Microwave Theory & Tech., MTT-20, No.7, pp.458, 1972

[5] A.J.Slobodnik,Jr., AFCRL-72-0082, NTIS AD-742287, 1971

[6] P.S.Cross, W.R.Shreve and T.S.Tan, "Synchronous IDT SAW Resonator with Q above 10,000", IEEE 1979 Ultrasonics Symp. Proc., pp.824-829, 1979

[7] E.K.Sittig and G.A.Coquin, "Filters and Dispersive Delay Lines Using Repetively Mismatched Ultrasonic Transmission Lines", IEEE Trans., Sonics & Ultrason., SU-15, No.2, pp.111-119, 1968

[8] R.C.M.Li and J.Melngailis, "The Influence of Stored Energy at Step Discontinuities on the Behavior of Surface Wave Grating", IEEE Trans. Sonics & Ultrason., SU-22, No.3, pp.189-198, 1975

[9] Y.Suzuki, H.Shimizu, M.Takeuchi, K.Nakamura and A. Yamada, "Some Studies on SAW Resonators and Multiple-Mode Filters", IEEE 1976 Ultrasonics Symp. Proc., pp.297-302, 1976

[10] S.Urabe, Y.Koyamada and S.Yoshikawa, "Experimental Studies of SAW Metal Strip Reflectors", J.Elect. Comm. Eng. Japan, Vol.60, No.9, pp.875, 1977 (in Japanese)

SHOCK TESTS WITH A PHASE LOCKED CRYSTAL OSCILLATOR

Dr. Harri Eskelinen and Prof. Pekka Eskelinen
Lappeenranta University of Technology, P.O. BOX 20, FIN-53851 Lappeenranta, Finland
Phone: +358-5-6212442, email: harri.eskelinen@lut.fi

1. ABSTRACT

A simple 5 MHz phase locked oscillator has been used to evaluate the possibilities to reduce the effects of external shocks by mechanical absorbers. The acceleration maximum has been 160 m/s^2 and the initial shock duration 500 μs. Two different frequency shift types have been recorded. Alternatively the crystal shows a quasi-permanent frequency step of 5*10^{-8} lasting for about 400 ms or a sharp excursion of 2*10^{-7} however having a duration of 10 ms and followed by a decaying ripple below 10^{-9} level for 500 ms. The XOR-based phase lock circuit employed together with its loop filter tolerates a reference loss or frequency jump up to 130 ms and starts an erratic correction if the reference de-tuning caused by an external shock lasts more than 320 ms. Once started, the re-settling may take over 3 seconds. The combined phase lock electronics may under less fortunate conditions, if a longer frequency step occurs, turn thus totally unstable. A slight change in the mechanical arrangement e.g. by immersing the oscillator unit into a viscous fluid moves the natural frequencies or adjusts the damping factor reducing the tendency to long-lasting frequency shifts.

2. INTRODUCTION

Shocks and vibration are stated to be the worst enemy of precision oscillators [1] and are nowadays very obviously encountered in a portable environment such as in mobile phones and various military gear [2]. Typically a relative frequency change of e.g. 5 * 10^{-6} is specified or tolerated but very little has been documented about the specific correlation between applied acceleration and induced frequency shifts particularly for complete phase locked circuits.

The purpose of this evaluation was to test the performance obtainable with a simple unprotected oscillator system having both an internal crystal unit and analog phase lock circuitry and to find out if it is feasible to try and redesign the mechanics in such a way which prevents excessive stresses from getting to the quartz or other acceleration sensitive (microphonic) parts. Tuning of the phase locking time constants together with or solely by a an electronic refinement of the low-pass filter in the phase comparator circuit might be able to limit the effects of the problem but it obviously affects the spectral (electrical) performance of the unit in steady state conditions as well. The main applications possibly benefiting form these results include several vehicle-mounted synthesizers and various portable gadgets.

3. DIFFERENT LOAD CASES DUE TO MECHANICAL VIBRATIONS

Practical efforts to find an optimal mechanical construction or system to damp vibrations or even just to insulate the critical components properly in a vehicle environment is a complex problem. The exact values of the forces, which are transferred from road and tyres to the vehicle's fuselage, are difficult to determine exactly in different circumstances. Further when a vibration spectrum is collected by using practical measurements during a test period it is difficult to generalize these results to cover all the load cases. It is also important to notice that the mechanical properties of a phase locked crystal oscillator construction should be adjusted with the properties of the vehicle's spring suspension and shock absorber system. These problems will be double as difficult if vibration problems should be solved not only in vertical but also simultaneously in the horizontal plane. In practice, this means firstly a selection of encapsulating materials [3] (E.g. cast irons have extremely good damping properties if their allowed yielding stress is not the critical property) and secondly optimising procedure to find the most appropriate combination of spring constants and damping factors of the oscillator's mechanical construction.

However, if the surface properties of the terrain can be estimated and further presented in a mathematical form it is relatively easy to derive the relationship between the roads shape X_1 and the corresponding displacement or amplitude X_2 of the vehicles fuselage as shown in Eq. 3 [4]. In this most simplified example we have assumed that the surface shape of the road can be presented in a mathematical form presented in Eq. 2 (see also Fig. 1). If needed the complete solution could be derived from the basic differential equation presented in Eq.1. If these simplified assumptions are

acceptable the maximum force F_{sm} transmitted to the vehicle's fuselage can be calculated from Eq. 4.

$$m\ddot{x} + c\dot{x} + kx = m\omega^2 X_1 \cos \omega t \qquad \text{(Eq. 1)}$$

$$x_1 = X_1 \cos \frac{2\pi s}{L} \qquad \text{(Eq. 2)}$$

$$X_2 = X_1 \sqrt{\frac{1 + 4\zeta^2 r^2}{\left(1 - r^2\right)^2 + 4\zeta^2 r^2}} \qquad \text{(Eq. 3)}$$

$$F_{sm} = kr^2 X_2 = m\omega^2 X_2 \qquad \text{(Eq.4)}$$

where ζ is relative damping factor ($=c/2\sqrt{km}$) and the other symbols are as illustrated in Fig. 1.

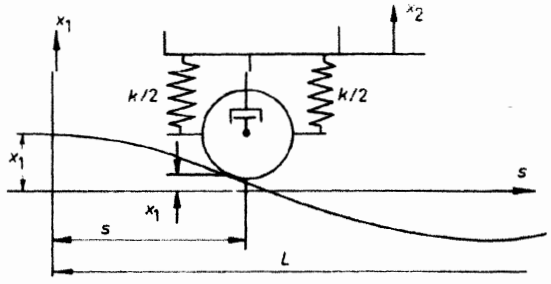


Fig. 1. One possible simple presentation of the road's surface shape and the mechanical modelling of the vehicle's spring suspension and shock absorber system.

In this particular case, if the force transferred from the vehicle's fuselage to the electronics is F_{sm} and the maximum effective force is F_0 we can present the relationships between these forces, corresponding amplitudes (X/X_a), relative damping factor (ζ) and frequency ratio r (showing the relative difference between the frequencies of the impacting force and the mass which is vibrating $r=\omega/\omega_n$) as illustrated in Fig. 2. These kinds of figures are useful when designing the mechanical construction and especially its dynamic behaviour.

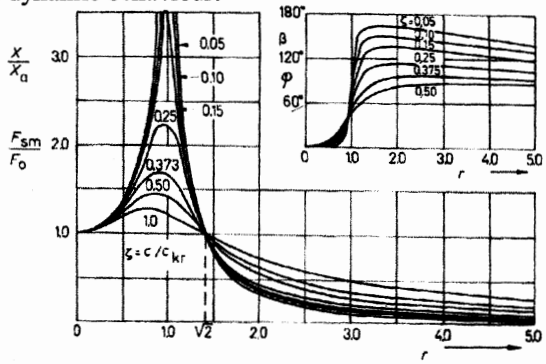


Fig. 2. Basic diagram for designing spring suspension and shock absorber systems [5].

4. UTILISATION OF COMPUTER AIDED VIRTUAL ENGINEERING (VE)

As justified above the exact mathematical solution of these kinds of practical vibration problems can be awkward and multi-level optimisation procedures are typically necessary. During this research we have found it more reasonable try to find the first acceptable mechanical solutions by using computer aided virtual engineering (VE).

Firstly we have 3D-modelled the encapsulating and secondly the adjustable spring mounting system and thirdly all the electrical components. Because the components are known exactly it is easy to include real mass properties into the model. The 3D-presentation of the final construction is shown in Fig. 3. If the 3D-model is produced in DXF- or even in IGES-format it is compatible in several software applications for dynamic analysis and further on for dynamic simulation.

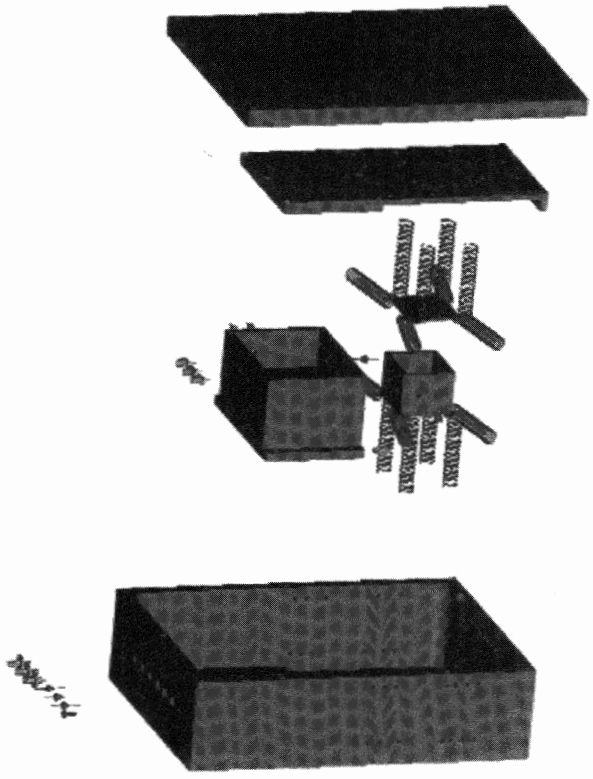


Fig. 3. During this research 3D-modelled construction which is compatible for VE simulations if it is produced in a proper file format. Electrical components expect the connectors are cut out from the model to illustrate the mechanical construction more clearly.

In our case we have decided to use standardized miniature springs which makes it much easier to ulilize virtual simulating on a computer screen because no modelling or calculations for the spring itself are needed. According to our opinion VE gives practical advantages in cases where several load types should be examined and the final mechanical construction is most probably a compromise of different variables.

5. SOME TEST ARRANGEMENTS

The basic quartz concept and the low-cost crystal unit were taken as such from [6]. A simple PLL circuit was connected between the crystal unit and a communications receiver FM demodulator which provided the stable reference through our cesium-based T&F transmitter. Due to historical reasons, 1 kHz was used as the comparison frequency. On the mobile platform a second, ovenized oscillator from Telequartz was used as a measuring reference for short term stability tests. Some of the laboratory installation is shown in Fig. 4 in which the high speed time interval analyzer is located near the left corner. The simple crystal oscillator - this time alone without any damping structures - is better visible in Fig. 5 with an acceleration transducer fastened above the cover. There was no way of reducing the commercial G-transducer's cable thickness, which caused a number of tedious re-runs during the measurements.

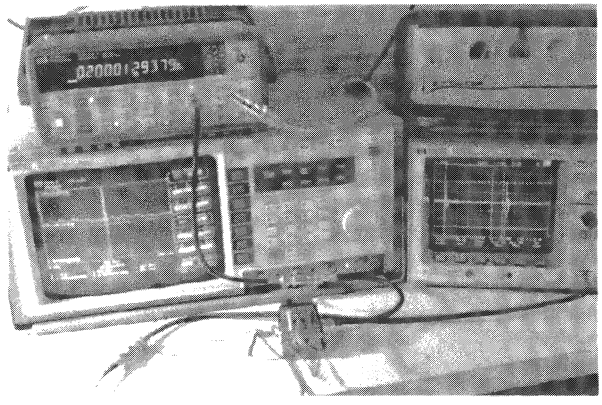


Fig. 4. The initial laboratory test arrangement includes the oscillator assembly (front), a time interval counter, a modulation domain analyzer (left) and a digital oscilloscope (right) which records the acceleration data.

Initially the phase lock time constants were extensively tested to obtain a background for the possible requirements set to the mobile crystal's vibration performance.

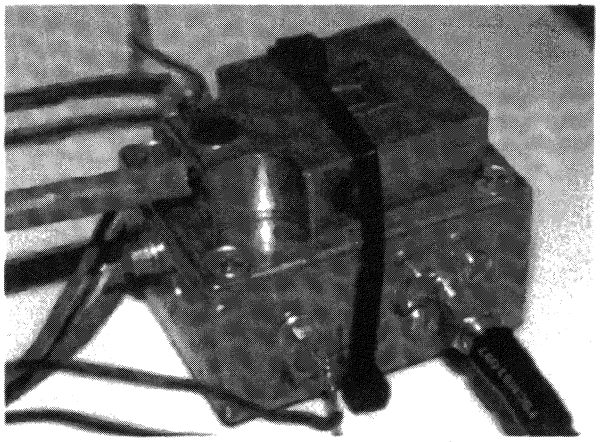


Fig. 5. The acceleration transducer is seen here mounted above the oscillator housing. The RF cable is in the lower right corner.

In an attempt to reduce noise and random fluctuations, the PLL low pass filter was designed to be quite slow which causes unavoidably a sluggish behaviour if one of its inputs is temporarily lost. This is illustrated in Fig. 6 where the experimental case was set up by disconnecting one of the inputs for 300 ms. As can be seen on the recording, the re-stabilization takes more than three seconds. Naturally the lock aquisition time depends on the relative accumulated error during the disturbed state and if the lock was completely lost for some reason, bringing the phase voltage to its extreme, the process was measured to last 16.8 seconds. Apparently every effort should be thus taken to prevent a prolonged loss of reference.

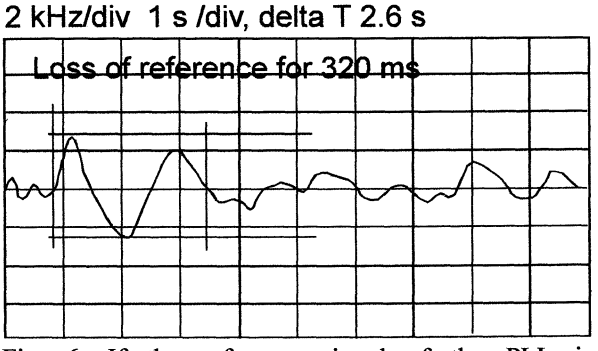


Fig. 6. If the reference signal of the PLL is momentarily lost or its frequency jumps, the circuit will loose lock. In our case a frequency deviation lasting about 300 ms is enough to cause a 3 s disturbance.

6. SHOCK TESTS

Two kinds of shock tests were carried out. First the performance of the crystal unit was measured without any absorbing mechanisms and later on extensive tests were carried out with the same electronics

supplemented by a number of alternative damping structures. The worst of the three major crystal orientations was selected for all tests in order not to show a "typical" performance but rather to give an estimate of worst case figures. However, the applied acceleration was kept well below the destructive force limit found earlier by breaking a couple of samples. Figure 7 shows the shock pattern used in our tests. It has a 500 μs wide first half cycle, which extends down to - 16 G.

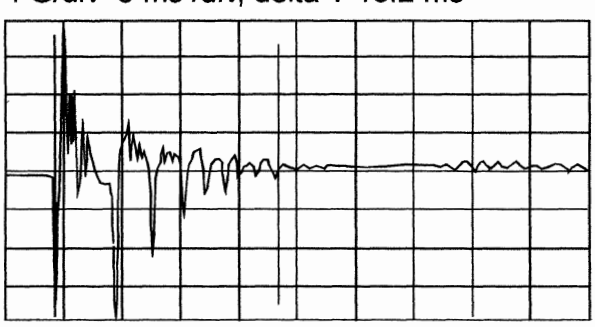


Fig. 7. When zoomed in, the details of the shock pattern observed in a typical vehicle installation reveal both a 100 - 400 Hz and a several kilohertz component. They decay in the test assembly within 20 ms.

Two different phenomena have been observed in the frequency characteristics of the basic oscillator. Alternatively the first impulse causes a rapid change in the oscillator frequency for about 10 ms as is illustrated in Fig. 8 or there may appear a more step-wise change which can last up to 400 ms as is plotted in Fig. 9. The relative frequency difference compared to the nominal value stays fairly repeatable, typically around 10^{-7}. Taking into account the previous evaluations of the PLL we easily note that the frequency step will induce a most severe disturbance.

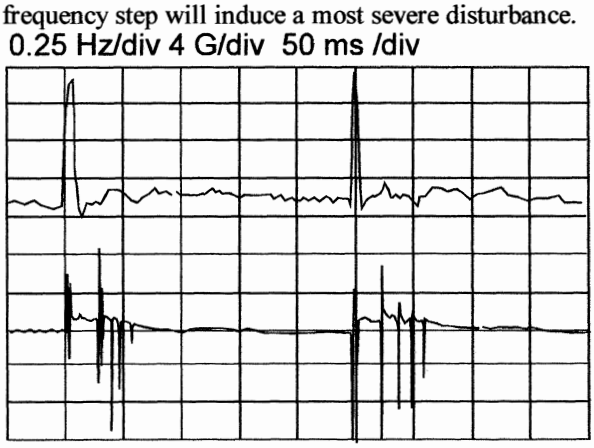


Fig. 8. This plot demonstrates the coupling between acceleration and reference oscillator frequency when the mechanical damping is disabled. The upper trace shows deviation from 5 MHz.

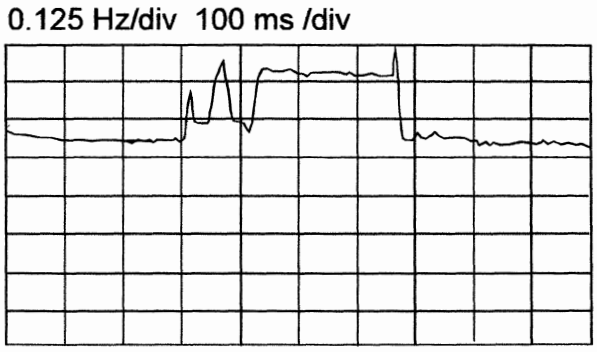


Fig. 9. Another type of primary oscillator behaviour was occasionally observed. Just one shock lasting only 500 μs can cause a stepwise change in frequency if there is no mechanical protection.

Several mechanical arrangements were tried ranging from miniature metal spirals and rubber tapes up to a viscous geel. As we were not so much interested in the possible spectral residuals caused by high frequency vibration but more concerned about the single-shot shocks, the dimensioning of the floating arrangement was done accordingly. Applying the same test impulse as during the preceding measurements we note in Fig. 10 a clear improvement in the frequency characteristics of the oscillator. The disturbed state lasts for about 120 ms. The respective situation, as seen at the PLL output is illustrated in Fig. 11.

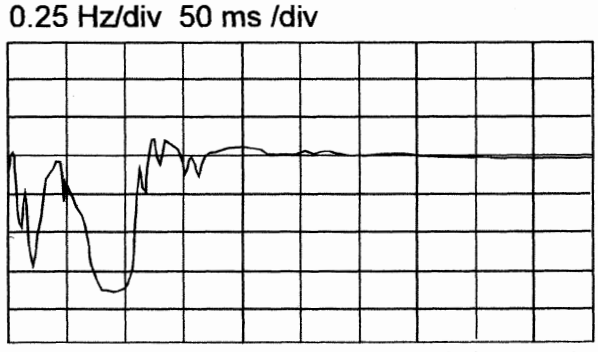


Fig. 10. The step-like frequency deviation of the crystal unit has disappeared after the introduction of a mechanical damping structure and thus the PLL no more gets disturbed for long periods of time.

As a supplement some general vibration tests were conducted, too. Quite apparently this type of stress is most often encountered e.g. in military wheeled vehicles going through rough terrain. The test signal was a 6 Hz sinewave with a maximum peak acceleration of 3 G - again along the most sensitive crystal axis. Figure 12 demonstrates the PLL output, which shows a nearly square nature and the recorded simultaneous platform acceleration.

25 mHz/div 50 ms /div

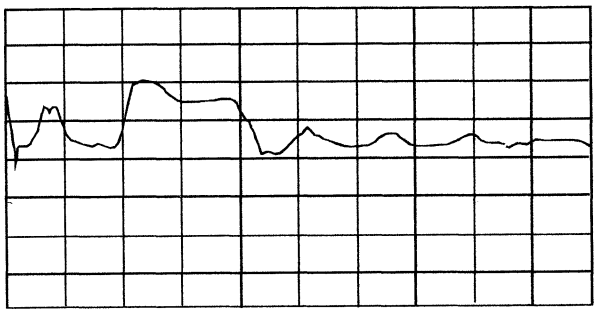


Fig. 11. The PLL output reacts with some delay to the crystal excursion but within 200 ms everything is over.

4 G/div 25 mHz/div 50 ms /div

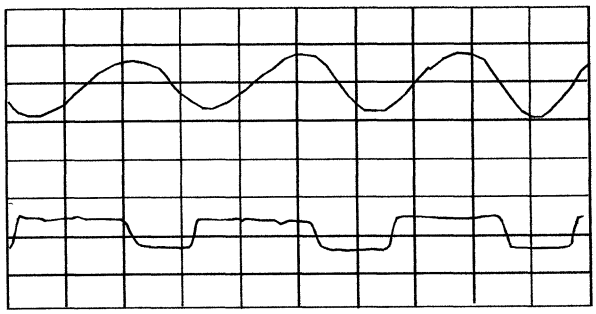


Fig. 12. The lower trace shows the PLL frequency as a function of time whereas the upper curve is the platform-mounted acceleration transducer output.

7. CONCLUDING REMARKS

It seems fairly reasonable to improve the performance of phase locked crystal oscillators by adding suitable mechanical shock damping structures. Without being involved in the spectral behaviour of the combined quartz unit - which would be the case if e.g. an intervention in the loop filter were made - we are able to push most of the slow vibrations caused by mechanical shocks outside the critical range of the mass system. The submerging approach gives thermal benefits, too, by stabilizing the spatial temperature gradients and giving some additional lag.

8. REFERENCES

[1.] Efratom, Precision Time and Frequency Handbook ,1996

[2.] Neubig B., Das grosse Quarzkochbuch, Franzis' Verlag, 1997

[4] Bishop, R. & Johnson D., The Mechanics of Vibration, Cambridge University, 1970

[3] Eskelinen, H.,"Suggestions for finding optimum materials and manufacturing methods for metallic microwave cavity resonators", Proc 13th EFTF'99, Besancon 1999

[5] Timoshenko, S., Vibration Problems in Engineering, Van Nostrand Co., Inc. New York 1966

[6.] Eskelinen P.,"A low cost ovenized crystal oscillator", Proc 9th EFTF'95, Besancon 1995

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

Frequency noise at a receiver and the Riemann zeta function

Michel Planat
Laboratoire de Physique et Métrologie du CNRS,
32 avenue de l'Observatoire, 25044 Besançon Cedex, France
e-mail: planat@lpmo.edu

Abstract

We recently reported on new experiments and a number theoretical understanding of the intermodulation spectrum in a superheterodyne receiver. Close to baseband a transition from white to 1/f power spectrum was observed. This was interpreted from a statistical analysis of the so-called Franel-Landau $\delta_i = |i/n - p_i/q_i|$, where the fraction p_i/q_i results from the finite resolution experiment and is the truncation of the continued fraction expansion of input oscillators frequencies; δ_i is the amplitude at mode i and measures the distance of rationals to the equally spaced graduation; and n is the number of resolved modes.

The Franel-Landau shift follows from the mode decomposition of the inverse zeta function $1/\zeta(s)$, with $\Re(s) > 1$. This can be explained physically if one looks at $\zeta(s)$ as the partition function of a "Riemann gas" with energies quantized on a logarithmic scale. $\zeta(s)$ already plays a role in the black body radiation where it defines the exponent of the temperature dependence in thermodynamic quantities. When properly generalized to the whole complex plane, the complex variable function $\zeta(s)$ shows a pole at $s = 1$, trivial zeros at negative even values of s (this is related to quantum noise) and, according to Riemann hypothesis, infinitely many zeros on the critical line $s = 1/2 + it$.

Looking at the inverse zeta function the non trivial zeros are infinitely many poles. In the high frequency range the power spectral density is found to be flat (white noise) at the critical line and shows the $1/f$ dependence close to it ($1/f$ noise). There are peaks in the low frequency range of the p.s.d. at $f_m = f_0 \log m$ with $m = p^l$, l integer and p prime. The first low frequency peaks are at 2, 3, 2^2, 5, 7, 2^3, 3^2, 11, 13.... This is related to the definition of total energy when $\zeta(s)$ is viewed as a partition function.

This confirms our earlier announcement of a connexion etween intermodulation dynamics and Riemann hypothesis. In addition to the understanding of low frequency noise the work will have applications in the field of digital modulation and coding.

1 Introduction

In his early study of electric noise J.B. Johnson[1] already found a large amount of voltage noise $S_v(f) \sim KV^2/f$ of the flicker type in addition to thermal noise $S_V(f) = 4kRT$. The $1/f$ noise has $K \sim \gamma/N$, with N the number of carriers in the sample and γ in the range 10^{-3} to 10^{-8}. This is generally attributed to different scattering mechanisms, by the crystal lattice or the impurities, leading to mobility fluctuations of the electrons. These findings point to **a nonlinear origin of the $1/f$ noise**. Fine structure revealing the interaction with bulk and surface phonons in several solid-state physical systems were found by Mihaila [2]. On the theoretical side a quantum electrodynamical theory was developed by P. Handel based on the infrared divergent coupling of the electrons to the electromagnetic field in the scattering process [4].

More recently the nonlinear approach has been revitalized in a series of new experiments performed at the LPMO [5],[6]. **The basic nonlinear electronic set-up is a Schottky diode mixer cascaded with a low pass filter**. The frequency of the beat signal is defined from a diophantine approximation of the frequency ratio of input oscillators. The amplitude results from the position of resolved fractions with respect to the equally spaced graduation. For the frequency fluctutions a transition from white frequency noise to $1/f$ frequency noise is observed close to baseband [6].

The distribution of energy among the modes is known in mathematics as the Franel-Landau (FL) conjecture and is equivalent to Riemann hypothesis (RH). RH relates on the distribution of zeros of the Riemann zeta function $\zeta(s)$ [3]. This function $\zeta(s) = \sum_{n\geq 1} n^{-s}$ with $\Re(s) > 1$ already plays a role in the physics of blackbody radiation laws [1]. The number of photons per unit volume is proportional to $\zeta(3)$ and the energy to $\zeta(4)$. Similarly in a Bose-Einstein condensate

[1] In black body radiation laws the Riemann zeta function reads $\zeta(s) = \frac{1}{\Gamma(s)} \int_0^\infty \frac{x^{s-1}}{e^x - 1}\, dx$ with $\sigma = \Re(s) > 1$ and $\Gamma(s) = \int_0^\infty u^{s-1} e^{-u} du$ is gamma function and $\Gamma(s) = (s-1)!$ if s is an integer.

the number of modes is proportional to $\zeta(3/2)$ and the energy to $\zeta(5/2)$. Thus the parameter in the zeta function also defines the exponent in the temperature dependence [7].

A first extension to lower than 1 values of $\Re(s)$ is required if one wants to model, as $\zeta(-3)$, the Casimir (vacuum) energy between two parallel conducting plates [8]. This is achieved thanks to the connexion of $\zeta(s)$, with s a relative integer, to Bernoulli numbers: they are defined from the coefficients in the algebraic expansion of the Planck factor $x/(e^x-1)$.

2 The inverse Riemann zeta function

One can looks at $\zeta(s)$ as a partition function (this was emphasized by B. Julia [9] in 1994) of some Riemann gas with energies $E_0 \log q$ (instead of $qh\nu$ in the conventional quantum harmonic oscillator). Since thermodynamical quantities such as the total energy are proportional to the inverse of the partition function, the zeros of $\zeta(s)$ should play a dominant role in the dynamics. To consider the problem in its full generality it is thus necessary to extend analytically $\zeta(s)$ to the full complex plane s [2].

The zeros of $\zeta(s)$ are of two types. They may be located on the negative real axis at

$$s = -2l, \quad l \text{ a positive integer} \tag{1}$$

and are called trivial. They are physically connected to the Fock space formulation of quantum mechanics [10].

All known non trivial zeros (they are infinitely many) are on the critical line

$$s = \frac{1}{2} + it \tag{2}$$

The original statement of RH is that all non trivial zeros are of the form (2).

2.1 The partition function

If one puts $E_q = E_0 \log q$, the "log" gas exhibits the partition function $Z = \sum_{q\geq 1} \exp(-\beta E_q) = \sum_{q\geq 1} \exp(-\beta E_0 \log q) = \sum_{q\geq 1} q^{-s} = \zeta(s)$ with the parameter $s = \beta E_0$. According to the classical formalism of thermodynamics the energy is $E = -\frac{\partial \log Z}{\partial \beta} = -E_0 \frac{\zeta'(s)}{\zeta(s)}$. Also

$$-\frac{\zeta'(s)}{\zeta(s)} = \sum_{q\geq 1} \frac{\Omega(q)}{q^s} \tag{3}$$

where $\Omega(q)$ is the van Mangolt's function, that is $\Omega(q) = \log p$, if $q = p^l$ (p prime and l integer) and $\Omega(q) = 0$ otherwise [3].

2.2 The Franel-Landau shift

The inverse zeta function is related to the Franel-Landau formulation of RH. One uses

$$\frac{1}{\zeta(s)} = \sum_{q\geq 1} \frac{\mu(q)}{q^s} \tag{4}$$

with $\mu(q)$ the Möbius function, that is $\mu(q) = 0$ if q is divisible by a square and $\mu(q) = (-1)^r$ otherwise, where r is the number of distinct factors in the prime decomposition of q.

According to Riemann hypothesis the domain of validity of equ(4) is $\Re(s) > 1/2$ [3].

One formulation of RH is $M(q_{\max}) = \sum_{q=1}^{q_{\max}} \mu(q) = \sum_{i=1}^{\Phi(q_{\max})} \cos(2\pi \frac{p_i}{q_i}) = 0(q_{\max}^{1/2+\varepsilon})$ for all $\varepsilon > 0$.

Here q is the denominator of an irreducible fraction and $q_{\max}$ relates to the resolution of the frequency mixing experiment mentioned at the beginning of this paper and $\Phi(q_{\max}) \sim \frac{3}{\pi^2} q_{\max}^2$ is the number of resolved fractions.

An equivalent formulation which also follows from equ(4) is in terms of "amplitudes" (or the so-called Franel-Landau shifts: see Fig. 1):

$$\delta_i = |\frac{i}{n} - \frac{p_i}{q_i}| \tag{5}$$

The statement equivalent to RH is $S(q_{\max}) = \sum_{i=1}^{\Phi(q_{\max})} \delta_i = 0(q_{\max}^{1/2+\varepsilon})$. In our previous publications [6] we have identified δ_i with the amplitude on the mode p_i/q_i corresponding to the nonlinear interaction between the frequencies $p_i f_0$ and $q_i f_1$ (f_0 and f_1 are respectively the frequencies of the RF signal and of the LO signal at the input of the mixer-filter element) leading to the beat signals

$$f_i^{p_i,q_i} = |p_i f_0 - q_i f_1| \tag{6}$$

Equs. (6),(5) describe in a compact way the intermodulation dynamics in the stationnary regime [3]

[2] One analytical continuation [3] of the Riemann zeta function is $\zeta(s) = \frac{e^{-i\pi s}\Gamma(1-s)}{2i\pi} \int_{\mathbf{C}} \frac{z^{s-1}}{e^z-1}\, dz$ where the path of integration $\mathbf{C}$ begins at $+\infty$, moves to the left down the positive real axis, circles the origin in the counterclockwise direction and returns up the positive real axis to $+\infty$. It is analytic at all points of the complex plane s except for a single pole at $s = 1$ with residue 1.

Since $\zeta(-n) = (-1)^n B_{n+1}/(n+1)$ with the B_n's the Bernoulli numbers, they are trivial zeros at $\zeta(-2l)$, l a positive integer. In addition Riemann hypothesis asserts that they are infinitely many non trivial zeros all located on the critical line $s = \frac{1}{2} + it$, with $t \sim 14.134$, $t \sim 21.002$, $t \sim 25.011$, ...

[3] If one rewrites equ. (6) by normalizing it with respect to the RF signal frequency f_0 we get

$$\mu^{p_i,q_i} = q_i|\nu - \frac{p_i}{q_i}| \tag{7}$$

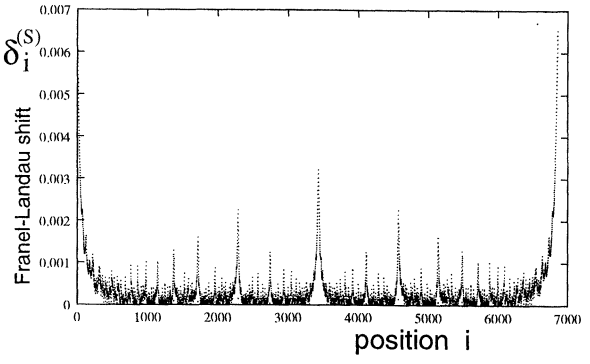


Figure 1: The Franel-Landau shift δ_i versus the position of the fraction p_i/q_i.

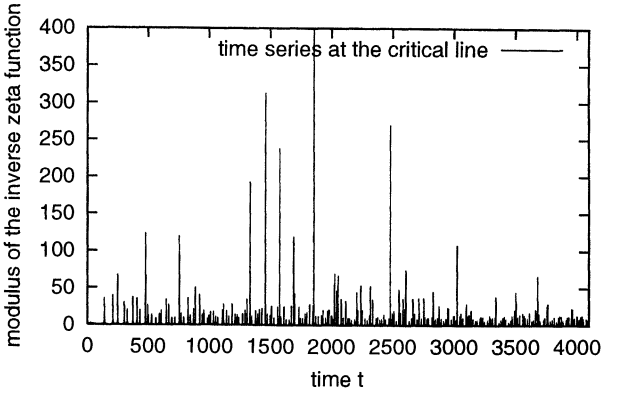


Figure 2: Time series of the modulus $Z^{-1}(t)$ at the critical line

Time fluctuations of the beat signal f_i were recorded [6]. **Using Allan deviation of the relative frequency $y = \delta f_i/\langle f_i \rangle$, we found $\sigma_y(\tau) = \tau^{-1/2}$ (that is white frequency noise) away from baseband and $\sigma_y(\tau) = C^{\text{te}}$ (that is $1/f$ frequency noise) close to baseband.**

This is related to the FL measure as $\sigma(q_{\max}) = S(q_{\max})/q_{\max} = 0(q_{\max}^{-1/2+\varepsilon})$, if one identifies the integration time τ with the denominator $q_{\max}$ of the bigger resolved fraction.

In the conventional approach all Farey fractions up to the denominator $q_{\max}$ are present. Close to baseband it is more reasonable to use the truncation at the partial quotient $a_{i+1} = a_{\max}$ in the continuous fraction expansion since from diophantine analysis we have $|\nu - \frac{p_i}{q_i}| \leq \frac{1}{a_{i+1}q_i^2}$, (with $\nu = f_1/f_0$ the frequency ratio of the RF and LO signals). In such a case the FL measure reads $\sigma(a_{\max}) = S(a_{\max})/a_{\max} = 0(1)$ from numerical experiments.

From this form it is expected that the device will product a rational approximation p_i/q_i of the real number $\nu = f_1/f_0$ with the error $\mu^{p_i,q_i} = f_i^{p_i,q_i}/f_0$. The approximation accounts for the finite resolution of the mixing and filtering process. It has been found that it is of the diophantine type (in contrast to a decimal approximation), that is corresponds to a truncation at some level i of the continued fraction expansion

$$\nu = [a_0; a_1, a_2, ...a_i...] = a_0 + 1/\{a_1 + 1/\{a_2 + ... + 1/\{a_i + ...\}\}\}; \quad (8)$$

and the a_i's are positive integers (called partial quotients) which are readily obtained from the formulas: $a_0 = [\nu], \alpha_0 = \{\nu\}$ and if $i \geq 1$, $a_i = [1/\alpha_{i-1}], \alpha_i = \{1/\alpha_{i-1}\}$, where $[\nu]$ denotes the integral part of ν and $\{\nu\} = \nu - [\nu]$ is its fractional part. Successive best approximants of ν, that are convergents p_i/q_i are obtained by truncating the expansion at some stage i. The best fit of the intermodulation experiments [6] has been obtained by truncating the CFE at the partial quotient $a_{\max}$ instead of $q_{\max}$ as would suggest the Franel-Landau formulation of the Riemann hypothesis. A full constructive account of fractions with a bounded partial quotient $a_{\max}$ can be found in [11].

3 $1/\zeta(s)$ and the transition from white to $1/f$ noise

The inverse zeta function may be decomposed in polar coordinates as

$$\frac{1}{\zeta(\frac{1}{2} + \eta + it)} = Z^{-1}(t)\exp(i\theta(t)) \quad (9)$$

where η is a shift parameter indicating the proximity to the critical line and t is physically the time. At the critical line there are the (Riemann-Siegel) approximating formulas [3]:

$$\Theta(t) = \frac{t}{2}\log\frac{t}{2\pi} - \frac{t}{2} - \frac{\pi}{8} + 0(t^{-1})$$
$$Z(t) = 2 \sum_{n<[(t/2\pi)^{1/2}]} n^{-1/2}\cos(\theta(t) - t\log n) + 0(t^{-1/4})$$

From now we consider the discrete Fourier transform (DFT) of waveforms as follows

$$S_{Z^{-1}}(f) = \sum_{t=0}^{\infty} Z^{-1}(t)\exp(2i\pi ft) \quad (10)$$

and similarly for the DFT of the phase $\theta(t)$. The following results are obtained by using the fast Fourier transform (FFT):

(a) $S_Z(f) = \text{Cte}$ at the critical line $\eta = 0$:
this is white noise, see Fig.2 (time series) and Fig.3 (FFT).

(b) $S_Z(f) = \text{Cte}/f$ close to the critical line $\eta \neq 0$
this is $1/f$ noise, see Fig.4 (time series) and Fig.5 (FFT).

In addition there are main low frequency peaks in the noise which are located at $f = f_0 \log q$, with $q = p^l$, p prime and l integer. The first peaks are at $q = 2, 3, 2^2, 5, 7, 2^3, 3^2, 11, 13 \ldots$. We remind that they are related from equ. (3) to the logarithmic derivative of

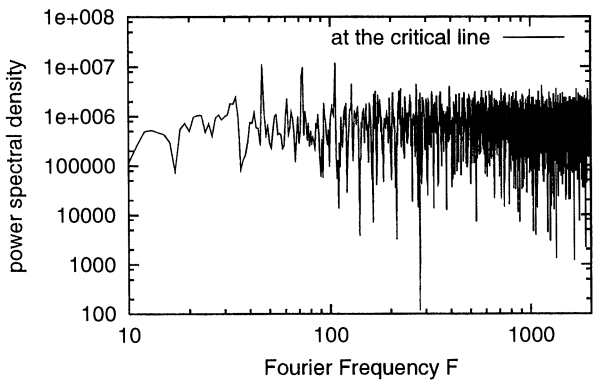


Figure 3: Power spectral density of $|1/\zeta(1/2+it)|$ (at the critical line). This looks like a white noise in the high frequency region.

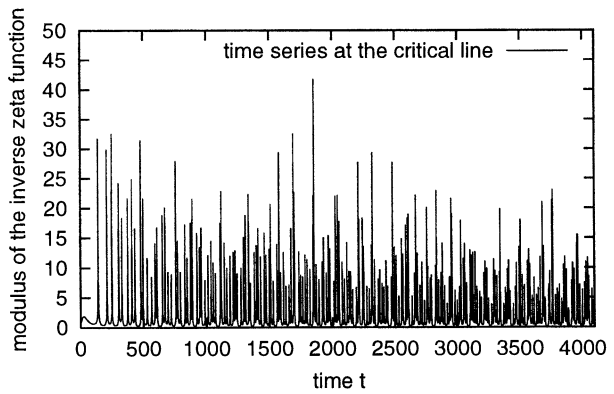


Figure 4: Time series of the modulus $Z^{-1}(t)$ close to the critical line at $\eta = 0.52$

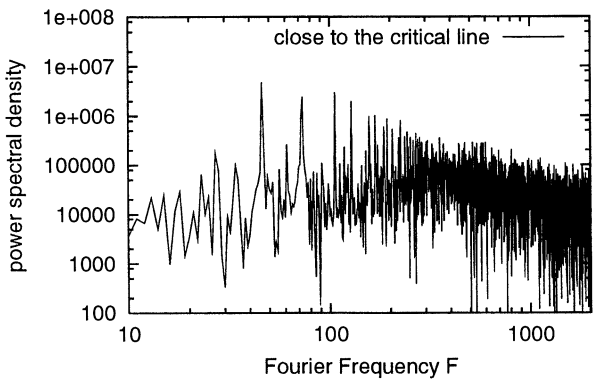


Figure 5: Power spectral density of $|1/\zeta(1/2+it+\eta)|$ with $\eta = 0.52$ (close to the critical line). This looks like a $1/f$ noise in the high frequency region. Main low frequency peaks are at $f = f_0 \log q$ with $q = p^l$, p prime and l integer

$\zeta(s)$ and to the Van Mangolt's function appearing in the expression for the total energy.

Similarly for the phase we get $S_\theta(f) = 0(f^{-2})$ at the critical line $\eta = 0$ and $S_\theta(f) = 0(f^{-3})$ away from it. Converting the phase spectrum into the frequency spectrum by using $S_{\dot{\theta}}(f) = f^2 S_\theta(f)$ we see that white frequency noise converts to $1/f$ frequency noise close to baseband.

The frequency analysis of $\frac{1}{\zeta(\frac{1}{2}+\eta+it)}$ using FFT thus confirms the statistical analysis which made use of the FL shift.

References

[1] J. B. Johnson, "Thermal agitation of electricity in conductors", *Physical Review* vol. 32, p. 97,1928.

[2] M. N. Mihaila, "Phonon fine structure in the $1/f$ noise of metals, semiconductors and semiconductor devices", *in Noise, Oscillators and Algebraic Randomness: NOAR*, edited by M. Planat, Springer Verlag, 2000

[3] M. H. Edwards, "Riemann's zeta function", Acad. Press, New York, 1974

[4] P. H. Handel, "The general nature of fundamental $1/f$ noise in oscillators and in the high technology domain", *in NOAR: see ref. [2].*

[5] S. Dos Santos and M. Planat "Arithmetical fractals in an electronic loop", *in Fractals and Beyond* , edited by M. M. Novak, World Scientific, Singapore pp. 297-306, 1998.

[6] M. Planat and C. Eckert " On the frequency and amplitude spectrum and the fluctuations at the output of a communication receiver", *IEEE Trans. on UFFC : Special Issue*, 2000.

[7] J. Kestin and J. R. Dorfman " A Course in Statistical Thermodynamics", Academic Press, New York, 1971.

[8] E. Elizalde, "Ten Physical Applications of Spectral Zeta Functions", Springer Lecture Notes in Physics, p. 111 , 1995.

[9] B. Julia, "Thermodynamical limit in number theory: Riemann-Beurling gases", *Physica A*, vol. 203 pp. 425-436, 1994.

[10] P.Cartier, "Mathemagics (A tribute to P. Euler and R. Feynman", *in NOAR: see ref. [2].*

[11] J. Cresson, "Geometry and dynamics of numbers under finite resolution", *in NOAR: see ref. [2].*

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

PM NOISE STUDIES OF QUARTZ CRYSTAL RESONATORS

F. STHAL, M. MOUREY, F. MARIONNET, S. GALLIOU

Laboratoire de Chronométrie, Electronique et Piézoélectricité
Ecole Nationale Supérieure de Mécanique et des Microtechniques
26, Chemin de l'Epitaphe - 25030 BESANÇON CEDEX – France
Phone: (33) 3 81 40 28 31, Fax: (33) 3 81 88 57 14, email: fsthal@ens2m.fr

ABSTRACT

Investigations of the phase modulation (PM) noise in quartz crystal resonators are presented. A study of the phase noise in 10-MHz BVA quartz crystal resonators is given. Two methods are investigated to classify the phase noise of the resonators. Comparison between these approaches and a model of frequency stability of ultra stable quartz oscillator is given.

1. INTRODUCTION

The first carrier suppression techniques to measure PM and amplitude noise (AM) noise were demonstrated by K.H. Sann[1] in 1968. Recently several teams have used these techniques along with some modifications and improvements[2-6].

In this paper, a study of the phase noise in 10-MHz BVA quartz crystal resonator pair is given. The measurements of the phase noise of the quartz crystal resonators are obtained with a crystal resonator tester[7]. This unit uses carrier suppression based on the bridge technique[8]. 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.

Measurements of phase noise of a quartz crystal resonator pair are given according to the dissipated power of the quartz crystals. Two methods are investigated to classify the phase noise of the resonators. Comparison between these approaches and a model of frequency stability of ultra stable quartz oscillator is given.

2. MEASUREMENT SYSTEM

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). The DUTs can be resistors (to measure the noise floor of the system) or crystal resonator pairs. The resonant 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. Because 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.

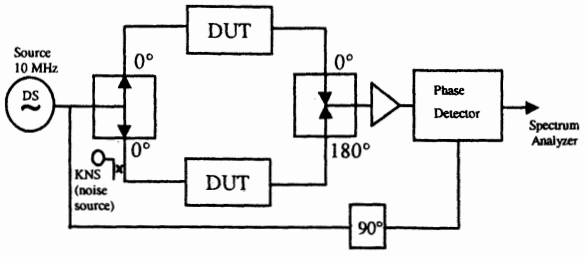


Fig. 1: Carrier suppression principle.

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 and is then turned off to measure the noise of the DUT. The noise of the DUT, as seen on the fast Fourier transform analyzer, is corrected using the calibration factor determined when the noise source was on. This calibration is done for each measured frequency and gives a very precise calibration of the system.

The struggle in using this system was to get enough carrier suppression to get a good measurement of the noise of the resonators without saturating the output amplifiers with the noise of the KNS. Because 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. MEASUREMENTS

Fig. 2 gives the power spectral density of the phase fluctuations (PSD) of DS, used in these measurements. It has a noise of -110 dBc/Hz at 1 Hz, -138 dBc/Hz at 10 Hz and a noise floor around -150 dBc/Hz.

The noise floor of the measurement system is given in Fig. 3. This measurement was obtained using two 91-Ω resistors. The measurement results are given in terms of the single sideband power spectral density of the phase fluctuations, $\mathcal{L}_\phi(f)$. We achieved a system

noise floor of approximately -149 dBc/Hz at 1 Hz and -153 dBc/Hz at 10 Hz.

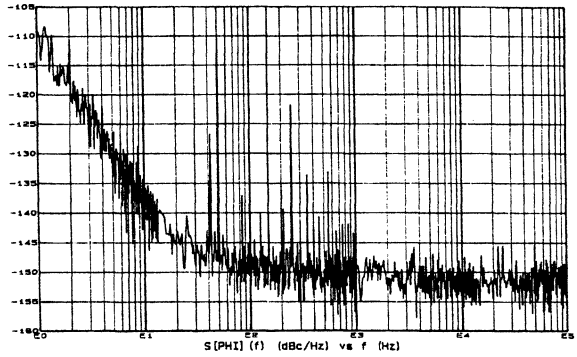


Fig. 2: Power spectral density of the phase fluctuations of the driving source.

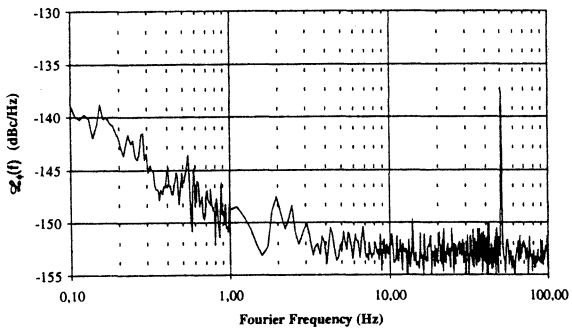


Fig. 3: Noise floor of the system.

Noise from two 10-MHz SC-cut BVA quartz crystals 5103 and 5123 is given in Fig. 4. We can observe the 1/f attenuation of the quartz crystal noise. The power applied to the quartz crystal is about 200 µW. If the noise of the DUT is above the calibration noise source, calibration is obtained with an unknown noise. This phenomenon appears below 0.2 Hz. The results are therefore not calibrated.

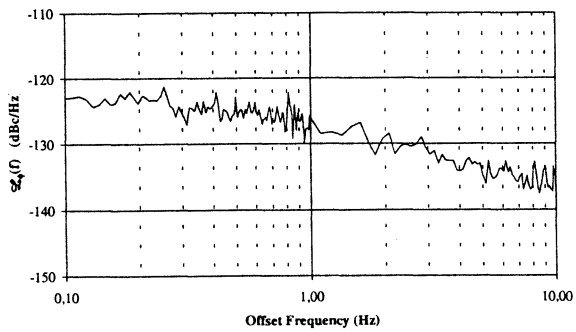


Fig. 4: Phase noise of the 10-MHz SC-cut BVA quartz crystal resonators.

Fig. 5 shows the single sideband power spectral density at 1 Hz following the power applied to the quartz crystals. We can observe a decrease of the phase noise when the power applied to the quartz crystals increase. These measurements are fit with a trend line using a 4th order polynomial approximation.

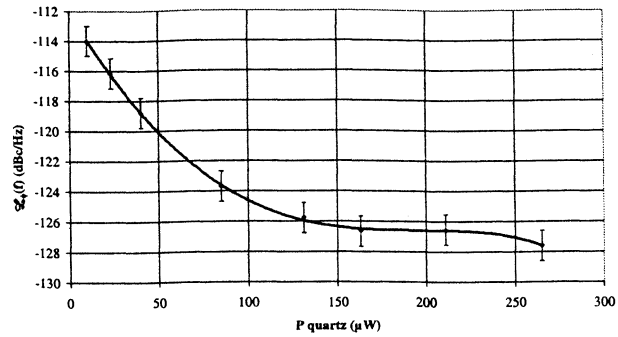


Fig. 5: Phase noise of the resonator versus the input power.

The generalization of this polynomial approximation requires further measurements with another crystal resonator pairs.

4. PHASE NOISE CHARACTERIZATION OF THE RESONATORS

4.1 Resonator in an oscillator

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 frequency fluctuations[9]:

$$\sigma_y^2(\tau) = \int_0^\infty S_y(f) \frac{2\sin^4 \pi f\tau}{(\pi f\tau)^2} df \qquad (1)$$

with f the Fourier frequency.

The standard Allan variance cannot be calculated from the phase noise of the resonator unless one knows the loaded quality factor of the resonator. Fig. 6 gives the loaded quality factor of the resonators Q_L obtained in the measurement system. The curve is obtained with the resonators 5123 and 5103.

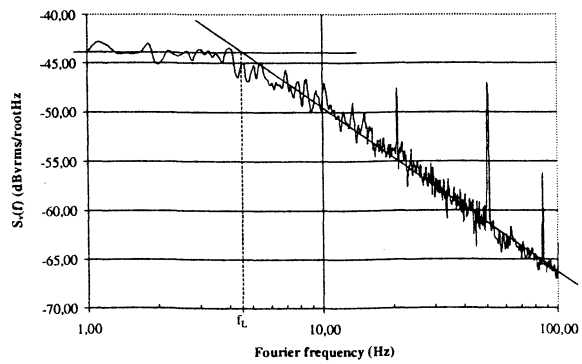


Fig. 6: Loaded quality factor Q_L of the resonators inserted in the PM measurement bench.

We evaluate the loaded quality factor with the relation[6]:

$$Q_L = \frac{\nu_0}{2 \cdot f_L} \qquad (2)$$

In this case, ν_0 = 10 MHz and $Q_L = 1.1 \cdot 10^6$.

The equations below show how $\sigma_y(\tau)$ is calculated. The resonator is considered as a low pass filter[10].

$$S_y(f) = \left(\frac{1}{2Q_L}\right)^2 S_\phi(f) \qquad (3)$$

Here, $S_\phi(f)$ is the measured phase noise of the resonator and $S_y(f)$ is the power spectral density of the frequency fluctuations of the resonator. The flicker phase in $S_\phi(f)$ gives flicker frequency in $S_y(f)$.

As expected $\sigma_y(\tau)$ is independent of τ for flicker frequency noise[11], $\sigma_y(\tau)$ is:

$$\sigma_y^2(\tau) = 2\ln(2) S_y(f = 1Hz) \qquad (4)$$

Thus the Allan variance is:

$$\sigma_y^2(\tau) = 2\ln(2)\left(\frac{1}{2Q_L}\right)^2 S_\phi(f = 1Hz) \qquad (5)$$

With both resonators 5123 and 5103, $S_\phi(1Hz)$ = -128 dBc/Hz. The same noise is assigned to the both measured resonators, thus $S_\phi(f)$ of one resonator is equal to the measured $\mathcal{L}_\phi(f)$. Standard deviation is computed at $2.13 \cdot 10^{-13}$ in the flicker floor.

The standard deviation calculated this way represents the standard deviation of an oscillator containing the test resonator in which the only source of flicker frequency noise is the test resonator. This oscillator has the same resonator loaded Q that the measurement system.

4.2 Resonator point of view

The Allan variance is computed from the Fourier frequency and from the power spectral density of the phase fluctuations following Eq. (1).

We consider that the resonator is submitted to a perfect sinus wave (Fig. 7).

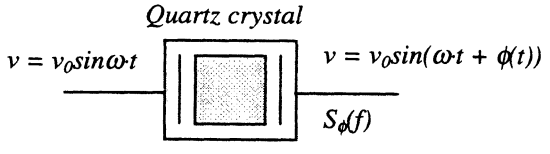


Fig. 7: Phase noise of the resonator.

In term of power spectral density of the phase fluctuations we have the following equation:

$$S_y(f) = \left(\frac{f}{\nu_0}\right)^2 S_\phi(f) \qquad (6)$$

Equation (1) becomes:

$$\sigma_y^2(\tau) = \int_0^\infty S_\phi(f) \frac{2\sin^4 \pi f\tau}{(\pi \nu_o \tau)^2} df \qquad (7)$$

Standard deviation is computed at $4 \cdot 10^{-14}$ at 1 s following the previous results (Fig. 4). The numerical integral calculus of the $\sigma_y(1s)$ is obtained with integral limits equal to 0.0001 Hz and 1000 Hz. The power spectral density of the phase fluctuations $S_\phi(f)$ which is used in this computation is a 1/f approximation of the previous results. It is an asymptotic curve including eight hundred points in a frequency decade and this particular value: $S_\phi(1Hz)$ = -128 dBc/Hz. In this case the resonators are qualified without an oscillator.

This point of view is more interesting because the resonator becomes independent of the oscillator where it will be inserted. This standard deviation can be considered as a characteristic of the resonator. This calculus depends on the upper frequency of the numerical integral because of the slope in f^{-1} of the phase noise. The same value of this upper limit of the integral must be used to compare the resonators each other.

4.3 Comparison with an oscillator approach

An improved model[12] of the frequency stability behavior of the oscillating loop is used to predict the phase noise performances of an oscillator which contains our test resonator. In this model, the transfer functions of the loop are entirely described. Fig. 8 shows the oscillating loop.

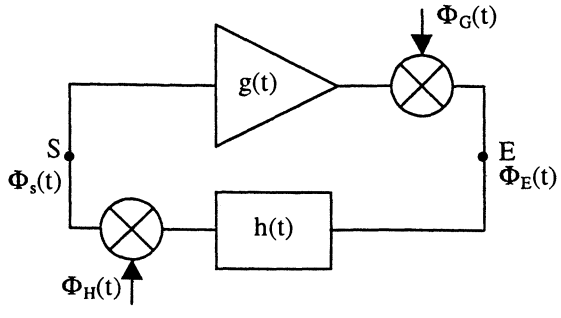


Fig. 8: Oscillating loop model.

g(t) denotes the impulse response of the loop amplifier (G(f) is its associated transfer function). h(t) denotes the impulse response of the quartz resonator (H(f) is its transfer function). $\phi_G(t)$ and $\phi_H(t)$ are the

amplifier and resonator phase noise respectively close to the oscillator frequency.

The transfer function H(f) which models the quartz resonator obviously depends on its loaded quality factor Q_L. The amplifier represented by G(f) has a quality factor rarely larger than 10^2. Thus the quality factor ratio H(f) and G(f) allows to simplify the model.

Considering that G(f) does not depend on f close to ν_0, the schematic loop in Fig. 8 is slightly transformed in Fig. 9 where $\phi_0(t)$ becomes the sum of equivalent resonator noise and amplifier noise close to the oscillator frequency ν_0. In our case we consider that the phase noise $\phi_0(t)$ is only due to the resonator.

$\overline{G}$ is the normalized transfer function.

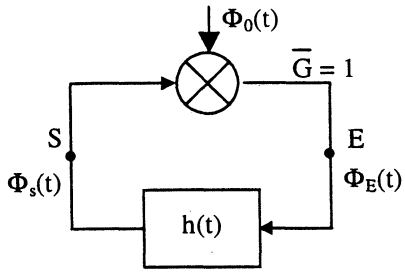


Fig. 9: Simplified oscillating loop.

The output power spectral density of the phase fluctuations of the oscillator can be expressed as follow:

$$S_{\phi osc} = S_{\phi_E}(f) = 10Log\left[10^{\frac{\mathcal{L}_{\phi_E}^{-}(f)}{10}} + 10^{\frac{\mathcal{L}_{\phi_E}^{+}(f)}{10}}\right] \quad (8)$$

in dBc/Hz, with

$$\mathcal{L}_{\phi_E}^{-}(f) = 10Log\left(\left|\frac{H(\nu_0 - f)}{H(\nu_0)} - 1\right|^{-2}\right) + \mathcal{L}_{\phi_0}(f) \quad (9)$$

$$\mathcal{L}_{\phi_E}^{+}(f) = 10Log\left(\left|\frac{H(\nu_0 + f)}{H(\nu_0)} - 1\right|^{-2}\right) + \mathcal{L}_{\phi_0}(f) \quad (10)$$

To take into account the spectrum in each side of ν_0, the computation uses single sideband power spectral densities $\mathcal{L}_{\phi_0}^{+}(f)$ and $\mathcal{L}_{\phi_0}^{-}(-f)$. These single sideband PSDs are given by the Eq. (11):

$$\mathcal{L}_{\phi_0}(f) = \mathcal{L}_{\phi_0}^{+}(f) = \mathcal{L}_{\phi_0}^{-}(-f) = \frac{S_{\phi_{res}}(f)}{2} \quad (11)$$

In accordance to the previous result (Fig. 4), the PSD of the resonator noise can be considered as follow:

$$S_{\phi_{res}}(f) = 10^{-12.8} \cdot f^{-1} + 10^{-17.3} f^{0} \quad (12)$$

$$\mathcal{L}_{\phi_0}(f) = 10 \cdot Log\left|10^{-12.8} \cdot f^{-1} + 10^{-17.3} f^{0}\right| - 3 \quad (13)$$

in dBc/Hz with $f > 0$

The f^0 term is considered as the thermal noise. The transfer function H(f) is given in Fig. 10, the resonator is loaded by a resistor. R_L is chosen in order to obtain the loaded quality factor of the resonator inserted on the oscillator.

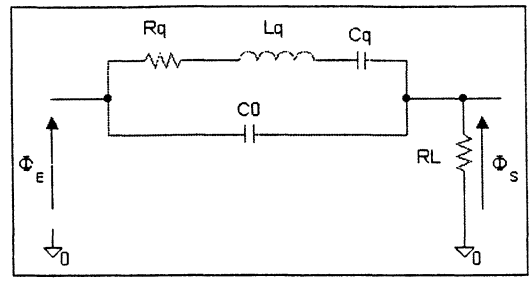


Fig. 10: H(f) transfer function.

Fig. 11 shows the PSD of an oscillator containing the test resonator in which the only source of flicker frequency noise is the test resonator with a loaded quality factor equal to $1.1 \cdot 10^6$.

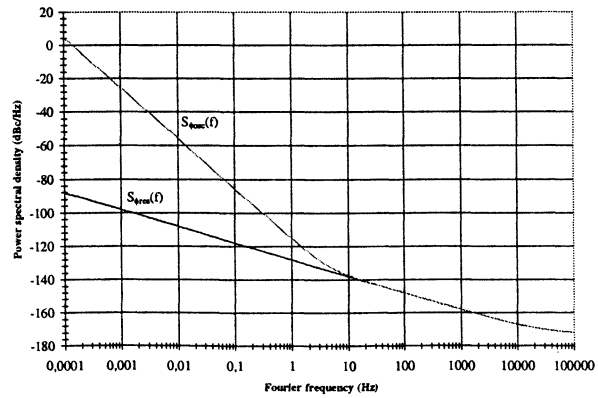


Fig. 11: PSD of the phase fluctuations of the oscillator containing the test resonator.

The performances of the oscillator will be affected by the noise of the resonator. With a given loaded Q factor, better oscillator cannot be obtained with these test resonators. Standard deviation is given using Eq. (7) for the resonator and the oscillator (Fig. 12). The maximum frequency for the integral in $S_\phi(f)$ is 1000 Hz.

We can see that a standard deviation for the resonator equal to $4 \cdot 10^{-14}$ is necessary to obtain a standard deviation for the oscillator equal to $2.13 \cdot 10^{-13}$.

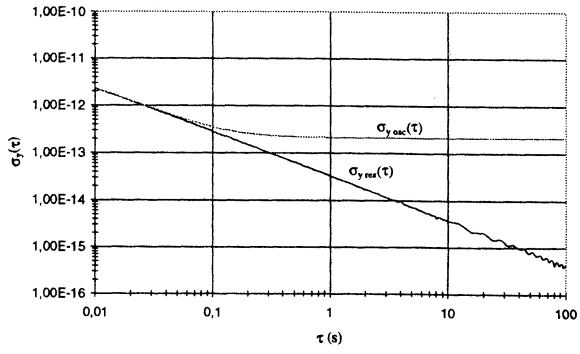


Fig. 12: Standard deviations of the resonator and the oscillator containing the test resonator.

In term of standard deviation, the measurement system presents a floor of the oscillator noise at 1 s equal to $2 \cdot 10^{-14}$ with a loaded $Q_L = 1 \cdot 10^6$ and a resonator floor equal to $8.2 \cdot 10^{-15}$ with a Fourier frequency maximum at 1000 Hz.

5. CONCLUSION

Phase noise of a quartz crystal pair is obtained. This phase noise is measured according to the dissipated power of the quartz crystal. The measured phase noise decreases when the power increases. We expect a noise floor around 200 µW.

In the time domain, the characterization of the resonator noise can be given following two ways. The first one, the standard deviation is calculated for an oscillator in which the only source of the noise is the test resonator. The second one allows to qualify the standard deviation of the resonator alone. The loaded Q is not necessary but the result depends on the upper frequency of the integral calculus. These both points of view seem to be complementary to interpret the phase noise measurement of a quartz crystal resonator.

We can predict the contribution in term of phase noise of an oscillator which contains a resonator. The best performances in term of phase noise of the oscillator are given by the numerical model. This model allows us to anticipate the future phase noise of the oscillator with a good precision.

6. 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 than 120 dB", in 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 interferometric PM and AM noise measurements", in Proc. 12th EFTF, Warsaw, Poland, pp. 286-291, March, 1998.

[6] E. Rubiola, V. Giordano, J. Groslambert, "Flicker noise measurement of HF quartz resonators", in Proc. Joint Meeting EFTF-IEEE, Besançon, France, pp. 1172-1175, April, 1999.

[7] F. Sthal, M. Mourey, F. Marionnet, W.F. Walls, "Phase modulation noise measurements of 10 MHz BVA using a phase bridge system with carrier suppression", in Proc. Joint Meeting EFTF-IEEE, Besançon, France, pp. 1176-1179, April, 1999.

[8] F.L. Walls, "Suppressed carrier based PM and AM noise measurement techniques", in Proc. 51st Ann. Freq. Cont. Symp., Orlando, FL, pp. 485-492, May, 1997.

[9] J. Rutman, "Characterization of phase and frequency instabilities in precision frequency source: Fifteen years of progress," Proc. IEEE, vol. 66, pp. 1048-1074, Sept., 1978.

[10] J. J. Gagnepain, "Fundamental noise studies of quartz crystal resonators," in Proc. 30th Annu. Freq. Contr. Symp. , pp. 84-91, June, 1976.

[11] J. A. Barnes, A. R. Chi, L. S. Cutler, D. J. Healey, D. B. Leeson, T. E. McGunigal, J. A. Mullen, W. L. Smith, R. L. Sydnor, R. F. C. Vessot and G. M. R. Winkler, "Characterization of frequency stability," IEEE Trans. Instrum. Meas., vol. IM-20, , pp. 105-120, May, 1971.

[12] M. Mourey, S. Galliou, R.J. Besson, "A phase noise model to improve the frequency stability of ultra stable oscillator", in Proc 51st Ann. Freq. Cont. Symp., Orlando, FL, pp. 502-508, May, 1997.

MODIFICATION OF THE RAW DATA SET FOR TIME EFFICIENT ADEV AND TDEV ASSESSMENT

Andrzej Dobrogowski, Michał 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 6652-293, fax ++48 61 6652-572

ABSTRACT

In the work the methods that produce good quality estimates of ADEV and TDEV in rather short calculation time are presented. The methods consist in the reduction of the amount of data used for calculation for long observation intervals. The processes of the data reduction consisting in the creation of a new data sequence that represents the old one are described. The results of ADEV and TDEV calculations made on the raw and modified data series are presented and compared.

1. INTRODUCTION

Allan deviation ADEV and time deviation TDEV are the parameters, which describe the quality of synchronization signals in the telecommunication network. The assessment of the parameter estimates is performed using the sequence of time error samples measured at some network interface. The amount of time error samples necessary for the reliable estimation and the complication of the estimator formulae result in rather very long time of the estimate calculation.

In the paper the methods of data reduction for the ADEV and TDEV calculation for long observation intervals are presented. Because of estimation accuracy reason the data reduction is only allowed for rather long observation intervals. The process of reduction is not a simple short cut of the measurement time. The short cut is unacceptable, because the telecommunication standards require the duration of measurement time twelve times longer than the longest observation interval taken into consideration. The reduction of data consists in the creation of a new data sequence that represents the old one. The methods of the creation of the new sequence based on the averaging, decimation and median searching of the data subset are presented in the Section 3 of the paper. The reduction of the data sequence used for estimate's calculation causes an obvious reduction in calculation time.

Presented methods of ADEV and TDEV estimates assessment were analyzed in the respect of the accuracy of calculated estimates. The modifications of the raw data set affect the quality of ADEV and TDEV estimates. The influence depends on the statistical behavior of the data. The results of ADEV and TDEV calculation made on the several time error sequences are presented in the Section 4. Some of the analyzed time error sequences were taken from the measurements of real several timing signals the others were obtained from the simulation process. The results of calculations performed on the modified data sets are compared with the results of calculations made on the raw data sets.

2. ADEV AND TDEV ESTIMATION

In order to calculate ADEV and TDEV time error (TE) is measured first. The results of measurement are taken with the interval τ_0 and stored in the data file. In the telecommunication standards [1] some conditions for the measurement of time error are specified. Time error shall be measured using an anti-aliasing filter with the cut-off frequency f_0. The minimum observation interval τ_{min} must be three times greater then maximum sampling time τ_0 and the anti-aliasing filter cut-off frequency f_0 must be equal $f_0=1/\tau_{min}$. The minimum measurement period T for time deviation calculation shall be twelve times longer then the observation interval τ.

In telecommunication standards the formulae for the estimators of the Allan deviation ADEV and the time deviation TDEV are defined:

$$\hat{ADEV}(\tau)=\sqrt{\frac{1}{2n^2\tau_0^2(N-2n)}\sum_{i=1}^{N-2n}(x_{i+2n}-2x_{i+n}+x_i)^2} \quad (1)$$

$$\hat{TDEV}(\tau)=\sqrt{\frac{1}{6n^2(N-3n+1)}\sum_{j=1}^{N-3n+1}\left[\sum_{i=j}^{j+n-1}(x_{i+2n}-2x_{i+n}+x_i)\right]^2} \quad (2)$$

where:

$\{x_i\}$ - sequence of N samples of time error function $x(t)$ taken with interval τ_0;

$\tau=n\tau_0$ - observation interval.

3. MODIFICATION OF DATA SET

In order to reduce the calculation time of ADEV and TDEV estimates for long observation intervals the modification of the time error set is proposed. The modification allows reducing the length of the time error sequence. As mentioned in the Section 1, the reduction is not the simple short cut of the measurement time. The reduction of data consists in the creation of a new data sequence that represents the old one. The three methods of the new data set creation are described below.

The reduction in the amount of data in each method for long observation intervals τ taken into consideration in

the parameter calculation is performed. The first step of the reduction is made when the observation interval is greater than $300\tau_0$. In this case the amount of data is ten times smaller than the amount of samples in the raw data set. The second step of the reduction is performed, when the observation interval exceeds $3000\tau_0$. The number of values in the new data set in this case is hundred times smaller than the number of samples in the raw data set. The next steps of the data reduction can be performed for the observation intervals greater than $30000\tau_0$ or $300000\tau_0$, if necessary.

The first method of the raw data set modification consists in the averaging of the time error samples [2]. The time error sequence is divided into the subsets of ten samples in the first step of data reduction. The new data set contains the average values of each subset. Similar operation is performed in the second step of reduction. The items in the modified data set are the average values of the subsets containing a hundred time error samples each. The operation of averaging can be performed in the next steps of the data reduction for the observation intervals greater than $30000\tau_0$ or $300000\tau_0$, if necessary.

The second method of the raw data set modification consists in the decimation of the time error samples. The new data set contains every tenth samples from the raw data sequence in the first step of the data reduction. In the second step the new data set contains every tenth samples from the data set created in the first step of the reduction.

The third method consists in the searching the median of the data subsets. The time error sequence is divided into the subsets of ten or hundred samples, depending on the reduction step, like in the first method. The modified data set contains of the median values found for each subset.

4. EXPERIMENT

The time of calculation and the accuracy of calculated estimates were analyzed in the experiment. Several time error sequences with different types were used in the experiment. Three sequences were obtained from the measurement process [3]. The first time error sequence (denoted as MSG, Fig. 1) was obtained from the measurement of the oscillator being the part of the timing signal measurement system. The GPS disciplined oscillator served as a reference. The second sequence (denoted as GPS, Fig. 2) resulted from the comparison of two different oscillators controlled by the GPS signals. The third sequence (denoted as DCF, Fig. 3) was obtained from the measurement of the oscillator controlled by the DCF signal with the GPS disciplined oscillator as a reference. The time error samples were taken with the sampling interval τ_0=1/30 s. The measurement period was 4000 s.

Other sequences were artificially created. These sequences represent the typical noises for the timing signals: white phase modulation (WPM), flicker phase modulation (FPM), white frequency modulation (FPM), flicker frequency modulation (FFM) and random walk frequency modulation (RWFM). The length of these sequences is the same as the length of the measured sequences.

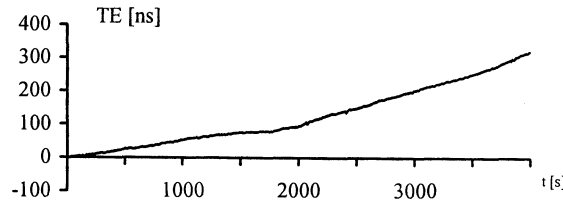


Fig. 1. Time error of the MSG signal measurement

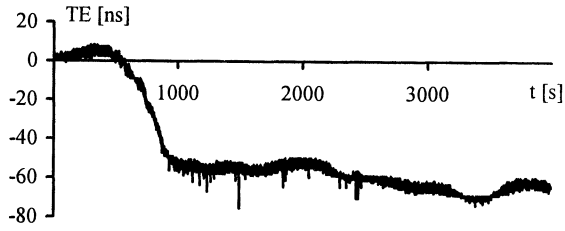


Fig. 2. Time error of the GPS signal measurement

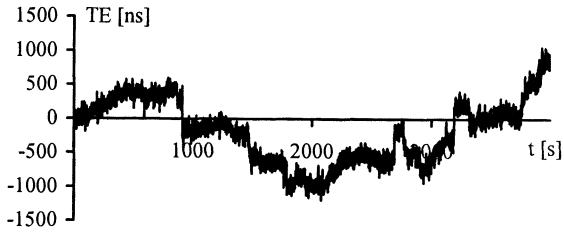


Fig. 3. Time error of the DCF signal measurement

In the experiment the 81 values of ADEV and TDEV estimate were computed. The observation interval varied from τ=0.1 s to 1000 s.

In the Table 1 and Table 2 the time of ADEV and TDEV calculation for the time error sequence measured during the time of 12000 s is presented. In the Table 1a and Table 1b the time of calculation for the whole range of 81 values of ADEV and TDEV and for the decade subranges of the whole range is presented. In the Table 2a and Table 2b the time of calculation for some chosen observation interval is given. The format of the time is as follows: minutes:seconds. The time of calculation using the modified data is the same for all modification methods. The PC computer with Pentium II 450 MHz processor was used for all calculations.

Table 1. Time of ADEV and TDEV calculation for the range of observation intervals

method	the range of observation intervals [s]				
	0.1-1000	0.1-1	1-10	10-100	100-1000
	ADEV				
direct	39:44.60	10:16.10	9:51.93	10:06.11	9:30.56
modified	21:12.57	-	-	0:59.05	0:05.49
	TDEV				
direct	53:31.16	14:09.53	13:13.40	13:31.03	12:37.20
modified	28:44.34	-	-	1:19.98	0:07.41

Table 2. Time of ADEV and TDEV calculation for some chosen observation intervals

method	observation intervals [s]				
	0.1	1	10	100	1000
	ADEV				
direct	29.33	29.38	29.99	30.16	25.54
modified	-	-	-	2.97	0.22
	TDEV				
direct	39.16	39.33	40.15	40.16	33.34
modified	-	-	-	4.01	0.33

The time of calculation with data modification is approximately ten times smaller for the observation intervals between 10 s and 100 s. It is about hundred times smaller for the intervals greater than 100 s. Thus the calculation time of the whole range of the parameter's estimate can be about two times smaller for the modified data.

The results of ADEV and TDEV calculations for the MSG time error sequence are presented in the Fig. 4 and Fig. 6. The relative error of the calculations with data modification is presented in the Fig. 5 and Fig 7.

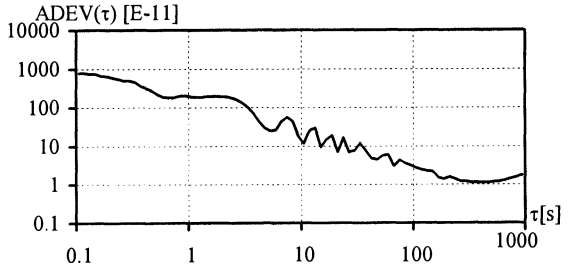


Fig. 4. ADEV for the MSG sequence

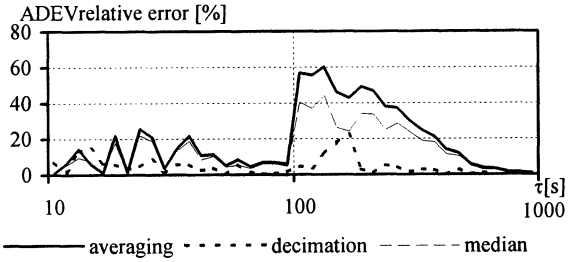


Fig. 5. ADEV relative error for the MSG sequence

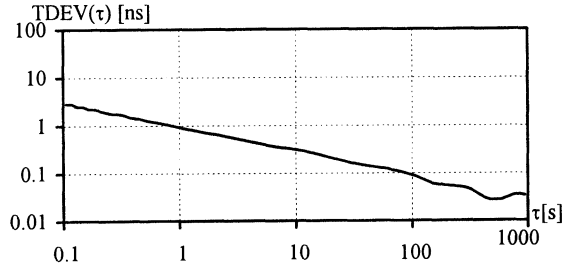


Fig. 6. TDEV for the MSG sequence

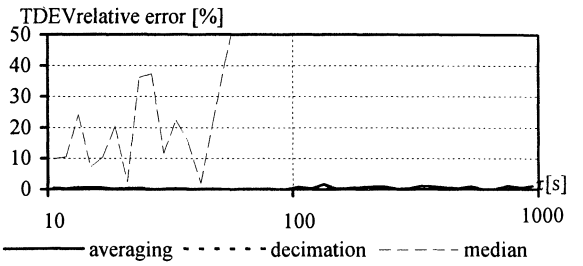


Fig. 7. TDEV relative error for the MSG sequence (the relative error for the decimated data exceeds 200%)

The relative error of the ADEV assessment for the MSG sequence is not greater than 60%. Satisfied results were obtained for the decimated data. The relative error in this case not exceeds 15% for the observation intervals smaller than 100 s. The error is about 22% for the observation interval of 200 s. Satisfied results of TDEV assessment for the MSG sequence were obtained for the averaged data. The relative error is about 1%. The calculations using other two modified data sets give the results with the error greater than 200%.

The results of ADEV and TDEV calculations for the GPS time error sequence are presented in the Fig. 8 and Fig. 10. The relative error of the assessment is presented in the Fig. 9 and Fig 11. The results of ADEV and TDEV calculations for the DCF time error sequence are presented in the Fig. 12 and Fig. 14. The relative error of the assessment is presented in the Fig. 13 and Fig 15.

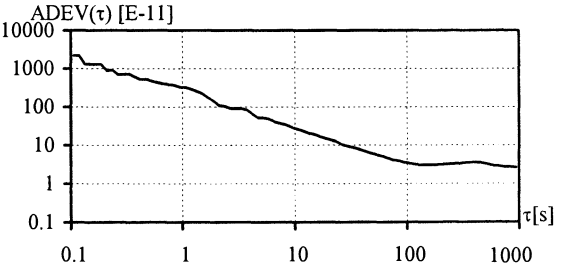


Fig. 8. ADEV for the GPS time error sequence

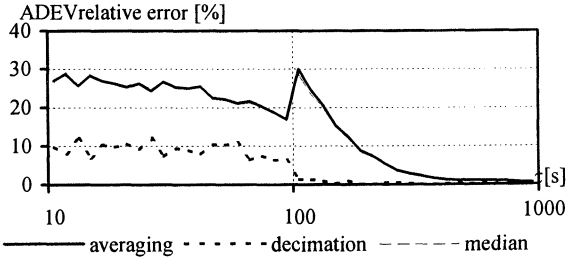


Fig. 9. ADEV relative error for the GPS sequence

The ADEV assessment results for the GPS and DCF sequence were acceptable for the decimated data. The relative error not exceeds 10% for the GPS sequence and 5% for the DCF sequence. Satisfied results of TDEV assessment were obtained for the averaged data and median data. The assessment relative error for the sequences was not greater than 5%.

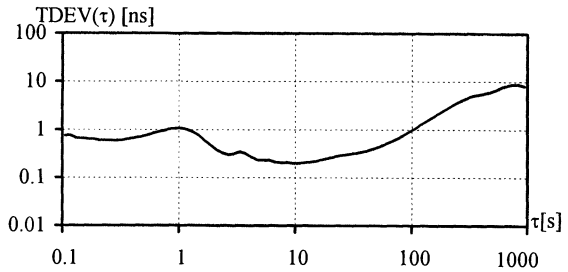


Fig. 10. TDEV for the GPS sequence

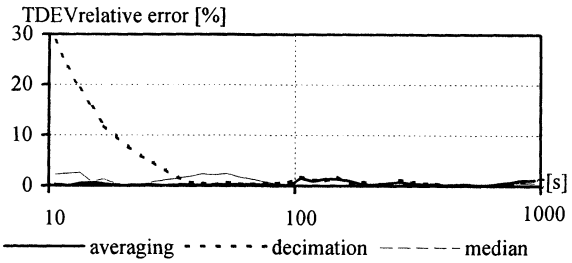


Fig. 11. TDEV relative error for the GPS sequence

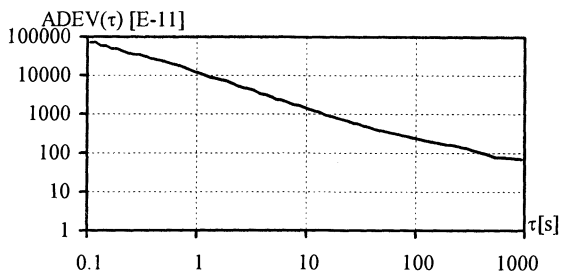


Fig. 12. ADEV for the DCF sequence

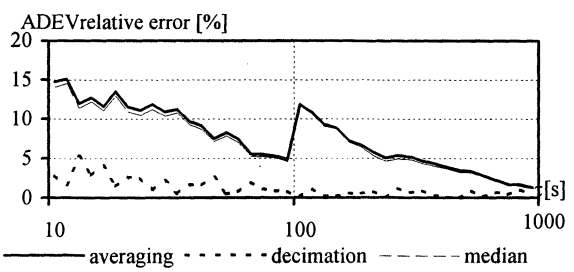


Fig. 13. ADEV relative error for the DCF sequence

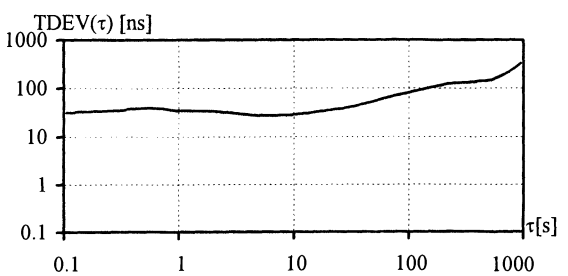


Fig. 14. TDEV for the DCF sequence

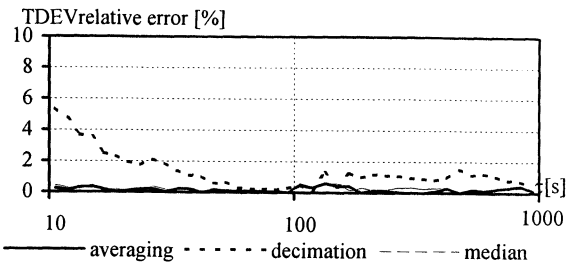


Fig. 15. TDEV relative error for the DCF sequence

The relative error of the ADEV and TDEV assessment performed for the WPM, FPM, WFM, FFM and RWFM sequences is presented in the Fig. 16-25.
The results of ADEV assessment for WPM and FPM sequences were satisfied using the decimated data. The relative error was not greater than 5%. The results of TDEV assessment for these sequences were acceptable using the data modification with averaging. The relative error not exceeds 1%. Satisfied results of ADEV and TDEV assessment were obtained for all modified data for the WFM, FFM and RWFM sequences. The ADEV assessment relative error was about 1%. The TDEV assessment relative error not exceeds 2%.

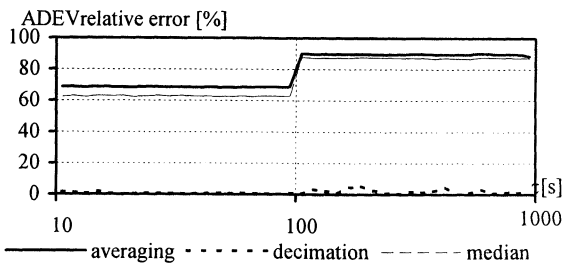


Fig. 16. ADEV relative error for the WPM sequence

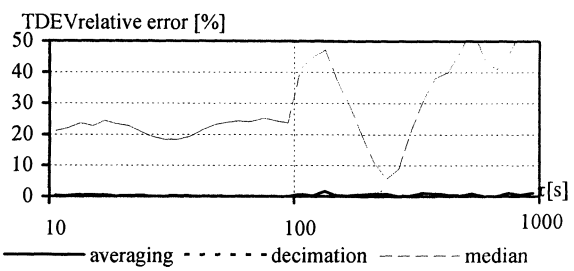


Fig. 17. TDEV relative error for the WPM sequence (the relative error for the decimated data exceeds 200%)

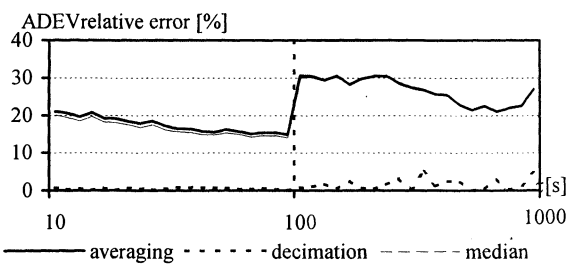


Fig. 18. ADEV relative error for the FPM sequence

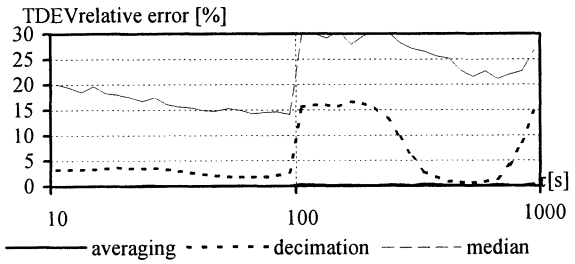


Fig. 19. TDEV relative error for the FPM sequence

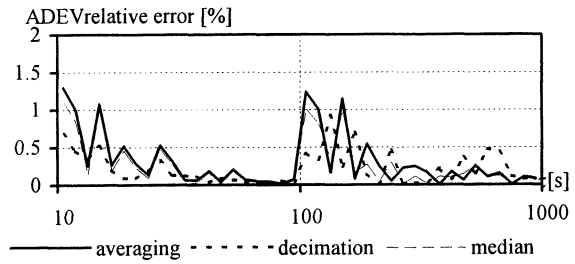


Fig. 20. ADEV relative error for the WFM sequence

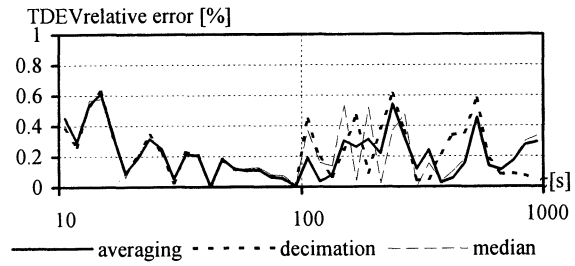


Fig. 21. TDEV relative error for the WFM sequence

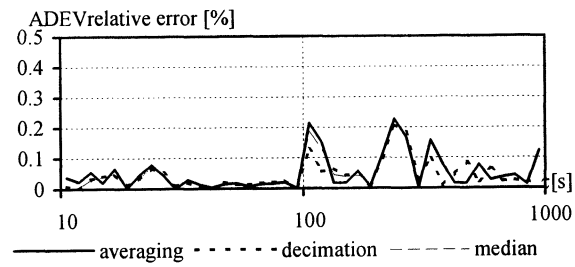


Fig. 22. ADEV relative error for the FFM sequence

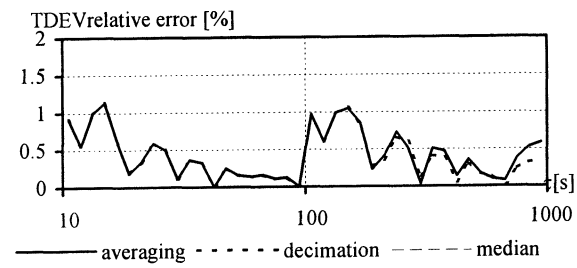


Fig. 23. TDEV relative error for the FFM sequence

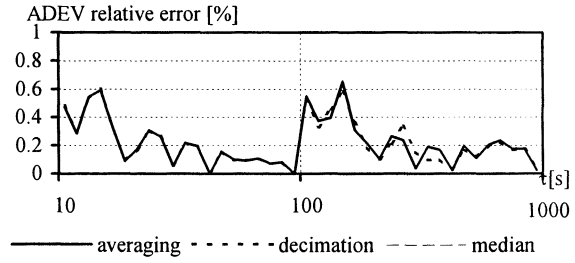


Fig. 24. ADEV relative error for the RWFM sequence

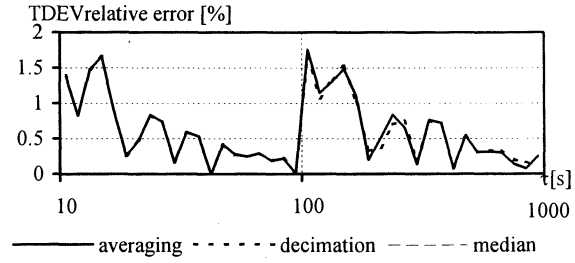


Fig. 25. TDEV relative error for the RWFM sequence

5. CONCLUSIONS

In order to shorten the time of finding the ADEV and TDEV estimate some modifications of the raw data are suggested. The modification results in some error of the parameter estimates. For some data sequences the error can be quite considerable. In general, it is not possible to point out the preferable procedure of the raw data modification. When we are interested in TDEV estimate it seems than the averaging is recommended. The decimation of the raw data can be recommended for the ADEV estimate assessment.

As an expected possibility of reducing the assessment error changing the number of the samples in the subset (for which the representative is established according to applied procedure) more often and with smaller step should be considered.

The methods proposed can be used for the ADEV or TDEV estimation, when the series of the parameter's values for the different observation intervals should be calculated. The time used for the data modification and single parameter value calculation can exceed the calculation time using direct method.

REFERENCES

[1] ETSI Draft ETS DE/TM-3017, "Generic requirements for synchronization networks", Helsinki 1995.

[2] M. Kasznia, "Some Approach to Computation of ADEV, TDEV and MTIE", in Proceedings of the 11th European Frequency and Time Forum, pp. 544-548, Neuchatel 4-6 March 1997.

[3] A. Dobrogowski, M. Kasznia, "Direct Approach to MTIE Calculation", in Proceedings of the Joint Meeting of the 13th European Frequency and Time Forum and 1999 IEEE International Frequency Control Symposium, Besancon 13-16 April 1999

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

THE REALIZATION OF A LOW NOISE PHASE COMPARATOR

Anca Niculescu
National Institute of Metrology, Sos. Vitan – Barzesti 11, Bucharest, Romania
Liviu Giurgiu, Octav Gheorghiu
National Institute of Laser, Plasma and Radiation Physics, PO Box MG 36,Bucharest, Romania

ABSTRACT

The appearance and the development of the new generation of frequency standards, based on atomic systems, with frequency stabilities of the order of $1x10^{-12}$ (τ =1 s) and better, imply the existence of the measurements systems able to measure frequency stabilities, at least, of the order of 1×10^{-14} (τ=1s).

This paper propose a method for the measurements of time domain stabilities of frequency and time standards based on the dual mixer time difference system.

1. INTRODUCTION

The measurements of the frequency stabilities could be done as well in the frequency domain as in the time domain. Time domain in contrast to frequency domain deals with the effects of averaging over different values of time, τ. Time domain characterization conventionally uses statistics computed from a discrete time series of average frequency samples obtained from a digital counter which is used to measure the beat note between the oscillator under test and a reference oscillator or the time difference between zero crossing of these beat supply by two identically channel. The conventional method used in time domain characterization is to heterodyne the original signal down to an arbitrary low beat frequency by mixing a reference signal with one or two other test signals. The low frequency beat signal reproduces faithfully the original frequency fluctuations of the test signals, which is typically measured with a digital counter. The commercial counter devices have "dead time" which destroys the opportunity to get to "true" time fluctuations. Also when making measurements between a pair of frequency standards or clocks is desirable to have less noise in the measurement system than the composite noise in the pair of standards being measured. The measurement system proposed in this paper has the advantage to satisfied these requirements.

2. DUAL MIXER TIME DIFFERENCE SYSTEM

The DMTD (Dual Mixer Time Difference) system proposed by D.W.Allan and H.Daams [1] is a system vary comfortable when we want to measure the frequency stability of the quasi-synchrone signals. The system described in this section open the door to making measurements of the frequency stability for sample times as short as a few milliseconds as well as for longer sample times and all without dead time. Unlike the system proposed by D.W.Allan and H. Daams we are made some modifications:

- We use two phase shifters with complementary action on the beat signals instead the high frequency phase shifter which could be o source of medium and long term drift ;
- We interposed a second order filter, with cut frequency f_c = 12 Hz, between the phase shifter and the zero crossing detector on each channel.

A block diagram of the dual mixer time difference system is shown in Figure 1. The signals from the Oscillators 1 and 2 go to the ports of a pair of double balanced mixers MX1 and MX2 (DBM166). The signal from the reference oscillator go to a power splitter, the two identically signals obtain are fed to the remaining other two ports of the pair of double balanced mixers. This common oscillator's frequency is offset by a desired amount, from the other two oscillators, producing the same beat frequencies with Oscillators 1 and 2. These beats will be out of phase by an amount proportional to the time difference between Oscillators 1 and 2 when running as clocks and will differ, in frequency, by an amount equal to the frequency difference between Oscillators 1 and 2.

Two phase shifters was inserted as illustrated in Figure 1 to adjust the phase so that the two beat frequencies are nominally in phase: this adjustment sets up the condition that the noise of the common oscillator tends to cancel when the time difference is determined – depending on the level and the type of noise as well as the sample time involved as described below.

The DMTD system has two identically channels. To simplify the explanation will be given only for one channel.

The beat signal go to the input of a preamplifier P with a gain Ap =23 dB. The amplified signal is fed to a phase shifter which adjust the phase by using a reference voltage V_{R1} (V_{R2}). By construction the reference voltage $V_{R2} = -V_{R1}$ which leads to $\varphi_2 = -\varphi_1$. After that the output signal go to a low pass filter and to zero crossing detector. The time interval between the two beat signals is measured with a time interval counter. The start port of a time interval counter is triggered with the zero crossing of one beat and the stop port with the zero crossing of the other beat. If the phase fluctuations of the reference oscillator are small during this time interval as compared to the phase fluctuations between Oscillators 1 and 2 over a full period of the nominal beat frequencies, the noise of the reference oscillator is insignificant in the measurement noise error budget, which means the noise of this oscillator can be worse than that of either Oscillators 1 and 2 and still not contribute significantly. The above condition will exist if the following equation is satisfied:

$$\Delta t \cdot \left\langle \sigma_y^2(2,\tau,\Delta t,f_c) \right\rangle^{1/2} << \tau \cdot \sigma_{y12}(\tau) \qquad (1)$$

The experimental verification confirmes the rejection of the noise of the local oscillator. In a real configuration the time difference δt between Oscillators 1 and 2 vary because the frequency differencies between oscillators under test. For short measure times ($\tau \sim$ few seconds) this supposition is confirmed but not in the case of long measurement time. In order to mend this problem we synchronized the reference oscillator with one of the test signal.

The electrical diagram of our DMTD is shown in Figure 2.

3. EXPERIMENTAL RESULTS

By using a low – frequency signal generator (Г3-110) as reference oscillator and by feeding the same signal from the atomic frequency standard HP 5071A, to the ports S1 an S2, we estimated the measurement noise introduced by the system itself. The low frequency signal generator was synchronized, all the time, with a signal of 10 MHz from the atomic frequency standard HP 5071A. The 1MHz signal from the signal generator was multiplied by 10 and fed to the port LO. To the ports S1 and S2 go the same signal, 10 MHz, from the same atomic frequency standard HP 5071A.

The time interval between two beat signals was measured with a time interval counter HP 53132A. A computer using an RS-232 interface and a one's own program collects the data. The subsequent process of the data was also made with a one's own program. The measurements are made with a nominal frequency of 10 MHz and the period of the beat was of 0.1 s, 0.2 s, 0.5 s and 1s.

We made the measurement in two configurations: using the operational amplifier OP27 and respective LT1028 as preamplifier P. The results are shown in Figure 3. The trendline for measurements made with OP27 (1) is:

$\sigma_y(\tau) = 3x10^{-12} \cdot \tau^{-0.89}$ and $R^2 = 0.978$.

The trendline for measurements made with LT1028 (2) is :

$\sigma_y(\tau) = 2x10^{-12} \cdot \tau^{-0.99}$ and $R^2 = 0.987$.

The trendline indicate an approximate dependence τ^{-1} for the noise floor measured with the low noise operational amplifier LT1028. This result confirms the necessity that a very good operational amplifier with small noise for low frequency must follow the double balance mixer.

Figure 4 shows the noise floor for the measurements made with the same nominal frequency (10 MHz) but with a beat period of 5 s. The symbols for the curves were maintain, 1(OP27) and 2 (LT1028). The tredline for both measurement points is :

$\sigma_y(\tau) = 3x10^{-12} \cdot \tau^{-0.96}$ and $R^2 = 0.999$.

This result suggests that the noise of the DMTD system be due fully to the type of the double balance mixer used.

4. CONCLUSIONS

The dual mixer time difference system allows, without doubt, better results. In future experiments we will use other mixers with less noise, especially for the flicker and phase noise. The dual mixer time difference system offer the greatest degree of flexibility in data analysis if the time difference can be measured with adequate precision.

5. REFERENCES

[1] D.W. Allan and H.Daams, "Picosecond time difference measurement system", Proc. 29th Annual Frequency Control Symp., pp. 404 – 411, May 1975

[2] R.Barillet, "Comparateur de phase ultra faible bruit pour les future etalons de frequence", 3e European Forum of Time and Frequency, March 1989

[3] Handbook Selection and Use of Precise Frequency and Time Systems, ITU, 1997

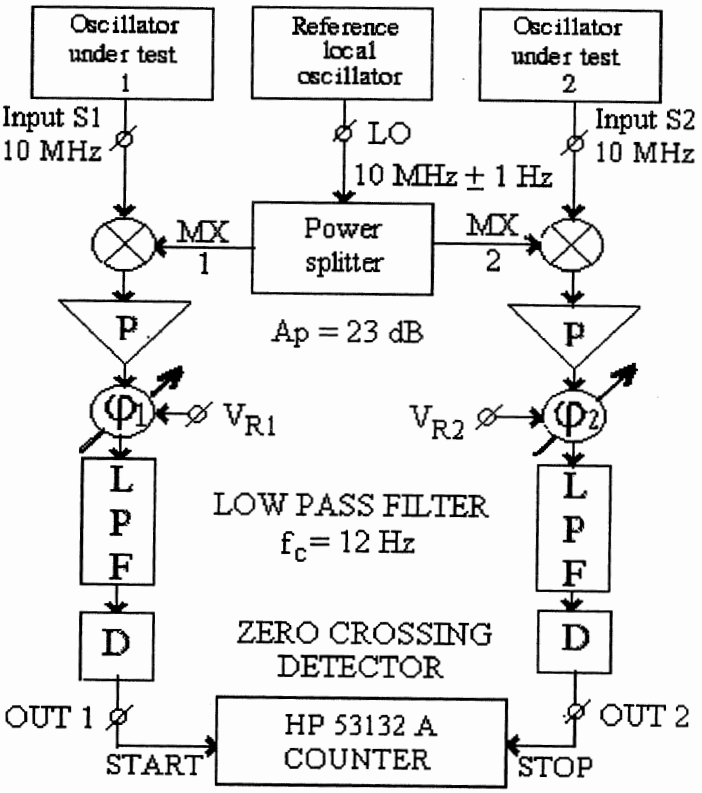


Figure 1. The block diagram of our dual mixer time systrm

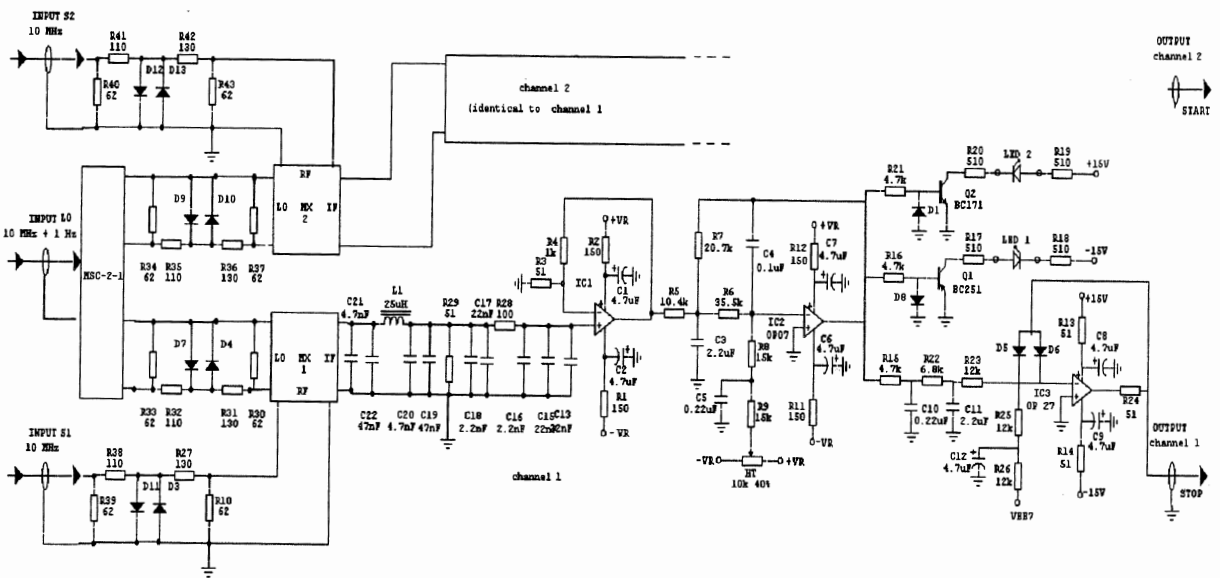


Figure 2. The electrical diagram of our dual mixer time system

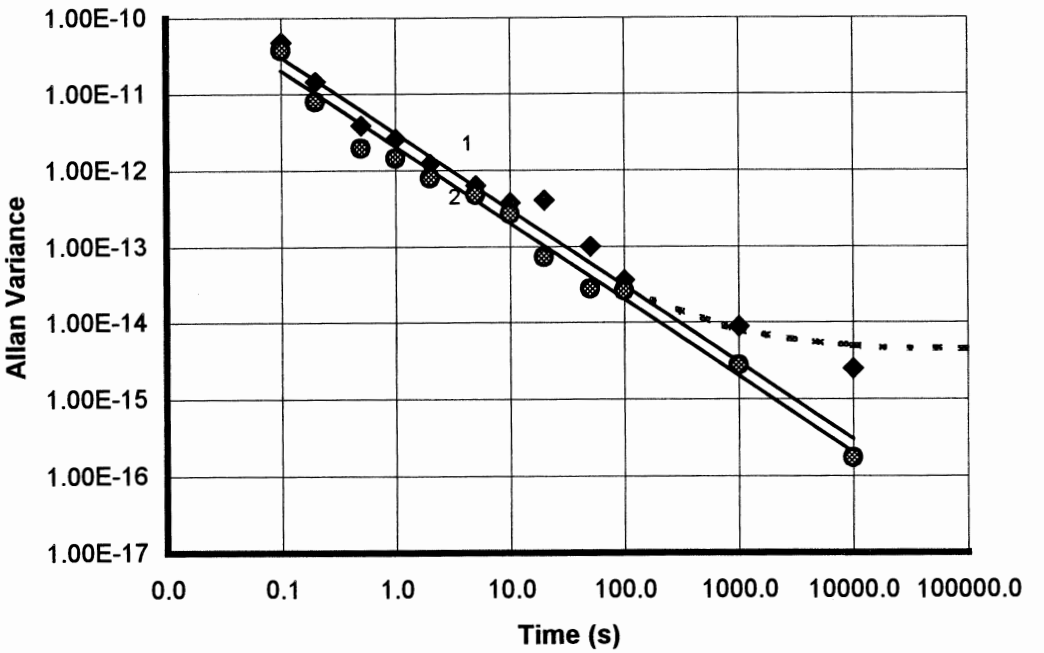


Figure 3. The stability foor measured at 10 MHz (T_B= 1 s)

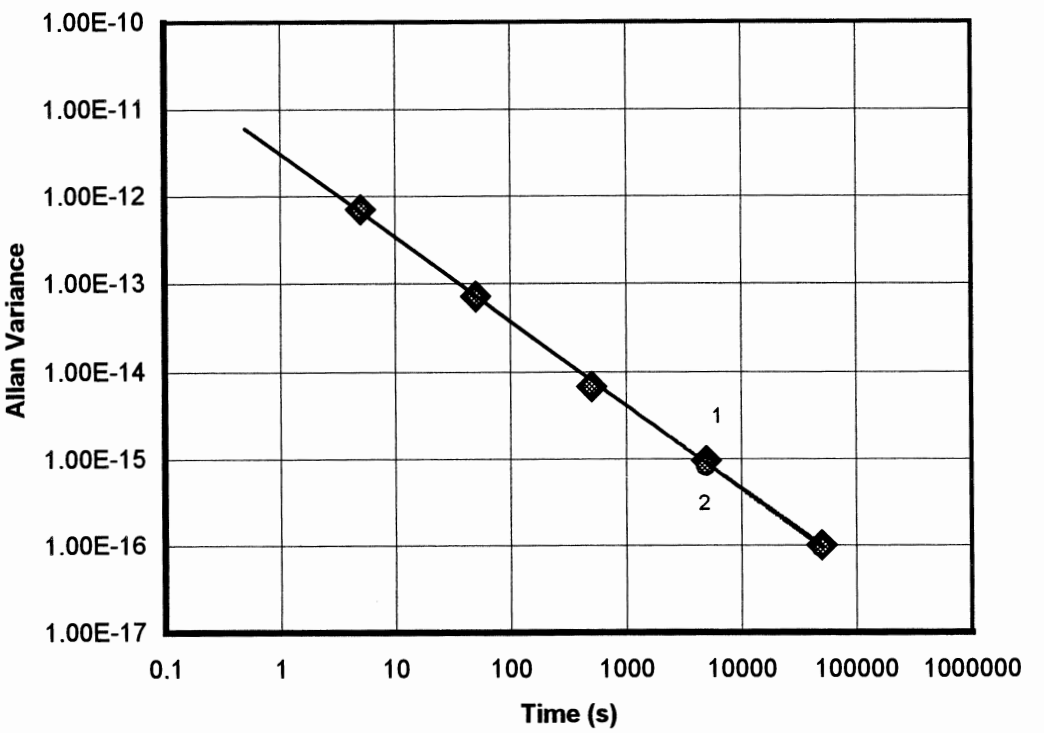


Figure 4. The stability floor measured at 10 MHz (T_B = 5 s)

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

EVALUATION OF TYPE A MEASUREMENT UNCERTAINTY FOR THE TIME ERROR

Ching-Haur Chang[1], Chia-Shu Liao, and Jiang-Lin Shi

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

A method of evaluating the Type A standard uncertainty for the time error, meeting the ISO's Guide, is presented in this paper. The evaluation of measurement uncertainty for the time error according to the standard remains so far an open issue. This is due to the fact that the time error is generally a function of time. Almost all oscillators display a superposition of deterministic and random variations in the time error. Any deterministic components must be removed prior to the computation of standard deviation. With a typical clock model, the random error is obtained with the removal of phase offset, frequency offset, and frequency drift. The resulting error is then used to compute the Type A standard uncertainty in this work. Using a high-performance cesium as a device under test, the uncertainty computed is 6.71×10^{-15} / hour, while that of computed from the Allan deviation is 6.66×10^{-15} / hour. The two results are almost the same.

1. INTRODUCTION

This paper presents a method for evaluating Type A standard uncertainty for the time error. Uncertainty, a parameter associated with the result of a measurement, which characterizes the dispersion of values that could reasonably be attributed to the measurand [1]. Specifying an uncertainty for a physical quantity of a measurement is important when reporting a result with its level of reliability. As an example, a national standard laboratory must provide source with specified uncertainty for traceability [2]. It is thus important to have guide to follow for testing and/or calibration laboratories. The guide specified in [1] only provides general rules for evaluating and expressing uncertainty rather than detailed, technology-specific instructions. For example, it does not take the time-varying characteristics into account. The ISO's Guide treats each measurand as an independent variable. However, a physical quantity of measurement is in general correlated with its adjacent values, especially from long-term point of view. The time error between two clocks is a good example. It is generally varying with time with values determined by a linear frequency offset and a linear frequency drift [3] and hence cannot be independent over time. This is why in timing community the expression of measurement uncertainty becomes difficult. Due to the importance of specifying uncertainty in timing laboratories [4], this needs to be solved in the near future.

[1] Corresponding author

A method of evaluating the Type A standard uncertainty for the time error is presented in this paper. This can be used as an extension of the ISO's Guide to this specific field. Using a high-performance cesium as a device under test, the uncertainty computed is 6.71×10^{-15}/hour, while that of computed from the Allan deviation is 6.66×10^{-15}/hour. The two results are almost the same.

2. TYPE A STANDARD UNCERTAINTY AND ITS EVALUATION

If the observation of a random variable is uncorrelated over time, the best estimate of the uncertainty is its standard deviation of the mean [1]. In this case, the expectation value μ_q of a quantity q is the average $\bar{q}$ of the n observations under the same conditions of measurement:

$$\bar{q} = \frac{1}{n}\sum_{k=1}^{n} q_k \tag{1}$$

The variance of the observations is given by

$$s^2(q_k) = \frac{1}{n-1}\sum_{k=1}^{n}(q_k - \bar{q})^2 \tag{2}$$

This quantity characterizes the variability of the values q_k, or more specifically, their dispersion about their mean $\bar{q}$. The variance of the mean is given by

$$s^2(\bar{q}) = \frac{s^2(q_k)}{n} \tag{3}$$

The quantity $s(\bar{q})$ is called the standard deviation of the mean. This quantifies how well $\bar{q}$ estimates the expectation μ_q of q, and may be used as a measure of the uncertainty of $\bar{q}$. Thus, for an input quantity X_i determined from n independent repeated observations $X_{i,k}$, the standard uncertainty $u(x_i)$ of its estimate $x_i = \bar{X}_i$ is $u(x_i) = s(\bar{X}_i)$, with $s^2(\bar{X}_i)$ calculated according to (3). For convenience, $u(x_i) = s(\bar{X}_i)$ is called a Type A standard uncertainty [1]. It is obviously from (3) that the quantity $u(x_i)$ is reducing as $\sqrt{n}$.

When the observations of a quantity are correlated, the mean and standard deviation of the mean may be inappropriate estimators. In this case, direct application of the statistics given by (1)-(3) cannot be appropriate for evaluating the uncertainty and, instead, special methods are required. The time error between two clocks can be described by [3]

$$x(t) = x_0 + y_0 t + \frac{1}{2}Dt^2 + \varepsilon(t) \tag{4}$$

where x_0 is the time offset, y_0 is the linear frequency offset, D is the linear frequency drift, and $\varepsilon(t)$ is the random deviation. It is clear from (4) that the measurements represented by $x(t)$ cannot be uncorrelated over time. Hence, special methods are required to treat measurements of this kind. Intuitively, the removal of x_0, y_0, and D in (4), i.e., the deterministic parts of the time error, may result in a series of uncorrelated, randomly varying measurements. Thus, the uncertainty of the measurement $x(t)$ is a result based on some statistical analyses of the random error $\varepsilon(t)$ as described in (4). Although it may not be useful in practice, since the parameters x_0, y_0, and D need to be estimated, this is the

intuitive method of expressing the uncertainty meeting the ISO's Guide.

The Allan variance [3] is a method of describing the dispersion about the estimator:

$$\sigma_y^2(\tau) = \frac{1}{2}\left\langle (\bar{y}_{k+1} - \bar{y}_k)^2 \right\rangle \qquad (5)$$

where $\bar{y}_i$ is the average frequency offset at time *i*. However, the use of Allan variance or the like only removes the components caused by the frequency offset [5], [6], [7], i.e., the first-order term in (4). Systematic components caused by the linear frequency drift cannot be filtered using similar approaches. Fortunately, many oscillators only exhibit frequency drift in some long-term measurements. It was stated in [4] that the uncertainty resulting from the long-term measurements can be ignored for simplicity. Due to this fact, by relating the Allan variance or the like to the statistical analysis mentioned above, a Type A standard uncertainty for the time error can also be determined from the Allan variance or the like. The relationship between the confidence interval and the Allan deviation can be found in [8]. It has been shown that the 68% confidence interval I_α or error bar for a Gaussian type of noise of a particular value $\sigma_y(\tau)$ obtained from a finite number of samples is

$$I_\alpha \approx \sigma_y(\tau)\kappa_\alpha n^{-1/2} \text{ for } n > 10 \qquad (6)$$

where n is the total number of data points in the estimate, α is the identification of the noise types, $\kappa_1 = \kappa_2 = 0.99$, $\kappa_0 = 0.87$, $\kappa_{-1} = 0.77$, and $\kappa_{-2} = 0.75$. Moreover, based on the definition of

$$\sigma_x^2(\tau) \equiv \tau^2 \sigma_y^2(\tau) \qquad (7)$$

the Type A standard uncertainty of the same dimension with the time error can be determined by $\sigma_x(\tau) = \tau\sigma_y(\tau)$.

3. ILLUSTRATIONS

Firstly, to see if the noises meet the desired property for evaluating the uncertainty, we compute the standard deviation of the mean for the five different types of noises: white PM (WPM), flicker PM (FPM), white FM (WFM), flicker FM (FFM), and random walk FM (RWFM). Fig. 1 shows the standard deviations of the mean for each type of the noises as a function of *n*, the number of samples used in the estimate. It can be seen that only the RWFM has values not reducing with $\sqrt{n}$. This tells that most of the noises do not violate the desired property in evaluating the uncertainty.

Secondly, the time error between the master clock derived in our lab and a high-performance cesium standard is used to demonstrate the consistency of the two methods mentioned in Section 2. Fig. 2 shows 650 points of the data. Fig. 3 shows the averaged frequency of the time error after the removal of the phase offset, frequency offset and the frequency drift. The standard deviation computed from this data is 1.718×10^{-13} / hour , and the standard uncertainty is 6.71×10^{-15} / hour . Fig. 4 shows the averaged frequency computed directly from the time error. The Allan deviation in this case is 1.714×10^{-13} / hour and $\alpha = 1$ is found by using Stable32[2]. Thus by using (6) with $\kappa_1 = 0.99$ and $n = 650$, the standard uncertainty is found to be 6.66×10^{-15} / hour . Note that the latter is close to the former.

[2] A software package developed by NIST.

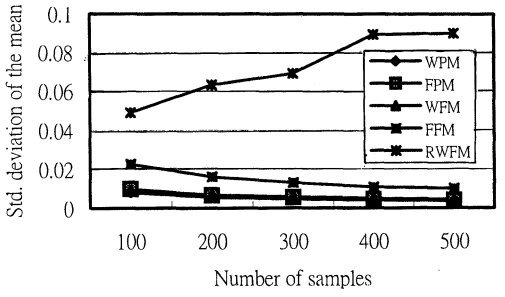


Fig. 1 Standard deviations of the mean for each types of noises.

Fig. 2 A set of time error for evaluating the uncertainty.

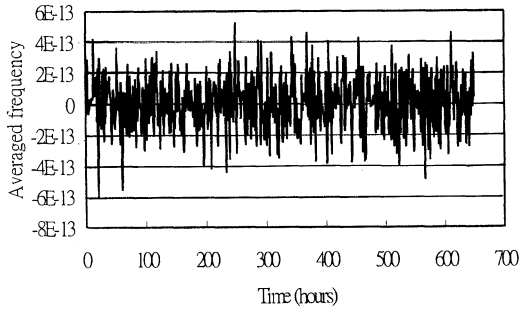


Fig. 3 Averaged frequency computed after the removal of the phase offset, frequency offset, and frequency drift.

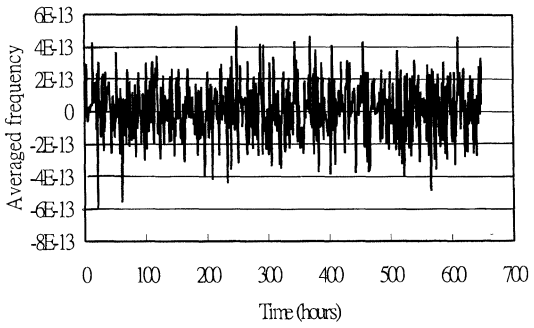


Fig. 4 Averaged frequency computed directly from the phase data.

4. CONCLUSIONS

A method of evaluating the Type A standard uncertainty for the time error is presented in this paper. With a removal of the deterministic components in the time error, the residual of the error is used to compute the Type A standard uncertainty. Using a high-performance cesium as a device under test, the uncertainty computed is 6.71×10^{-15} / hour , while that of computed from the Allan deviation is 6.66×10^{-15} / hour . The two results are almost the same.

5. REFERENCES

[1] BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML: "Guide to the expression of uncertainty in measurement," 1st Ed., ISO, 1993.

[2] CCTF, "Report of the CCTF Working Group on the expression of uncertainty in primary frequency standards," 1999.

[3] D.W. Allan, "Time and frequency (time-domain) characterization, estimation, and prediction of precision clocks and oscillators," *IEEE Trans. Ultrasonics, Ferroelectrics, and Frequency Control,*

vol. UFFC-34, no. 6, pp. 647-654, Nov. 1987.

[4] ISO/IEC 17025, "General requirements for the competence of testing and calibration laboratories," 1999.

[5] L.S. Cutler and C.L. Searle, "Some aspects of the theory and measurement of frequency fluctuations in frequency standards," *Proceedings of the IEEE*, Vol. 54, No. 2, pp. 136-154, Feb. 1966.

[6] P. Lesage and C. Audoin, "Characterization and measurement of time and frequency stability," *Ridio Science*, Vol. 14, No. 4, pp. 521-539, 1979.

[7] P. Lesage and T. Ayi, "Characterization of frequency stability: analysis of the modified Allan variance and properties of its estimate," *IEEE Trans. Instrumentation and Measurements*, Vol. IM-33, No. 4, pp. 332-336, Dec. 1984.

[8] P. Lesage and F. Audoin, "Characterization of frequency stability: uncertainty due to the finite number of measurements," *IEEE Trans. Instrumentation and Measurement*, Vol. IM-22, pp. 157-161, June 1973.

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

THE SUPPRESSION OF DC TUNING VOLTAGE NOISE IN PHASE NOISE MEASUREMENT OF VCO FREE RUNNING

ANTONY YAO

Center for Measurement Standards, ITRI, Taiwan

YAO-HUANG KAO

Institute of Communication Engineering, National Chiao Tung University, Taiwan

Antony Yao, Center for Measurement Standards, Bldg.16, 321 Kuang Fu Road, Section 2, Hsinchu, Taiwan

TEL:886-3-5732177, FAX:886-3-5724952, E-mail: 760561@itri.org.tw

ABSTRACT

The phase noise measurement of Voltage Controlled Oscillator(VCO) free running has become increasingly important to ensure the performance of some advanced instruments or mobile phones which use VCO as a signal generator. In the phase noise measurement of VCO free running, a DC tuning voltage is needed to bias the tuning varactor of VCO. The noise coupled with the DC voltage will degrade the performance of VCO in the phase noise measurement. To minimize the effect of noise, the battery source is often used to afford the DC voltage. However, it is not convenient to use a battery source to afford a continuous variation of DC voltage in the measurement. This paper presents a method for suppressing the noise coupled with the DC power supply output to measure the phase noise of VCO free running. We connected a resistor to the DC bias input port of VCO tuning varactor in series. Then we separately measured the phase noise of VCO free running which was biased by the battery source and the DC power supply. We found the phase noise of VCO free running biased by the DC power supply had been improved more than 10 dB due to connecting a resistor in series. In some offset frequency bands, a 20 dB improvement was seen.

1. INTRODUCTION

The DC power supply is the main power source to afford a DC voltage to bias the tuning varactor of VCO in the phase noise measurement of VCO free running. The output of DC voltage from the DC power supply is often coupled with the output of noise. The noise contains the harmonic components of power line frequency and a few spurs [1], [2]. The noise coupled with the output of DC voltage will make a degradation of performance in the phase noise measurement of VCO free running. To avoid the effect of noise, the battery is substituted for the DC power supply to afford a DC voltage. In the phase noise measurement of VCO free running, a few different values of DC voltage are needed to bias the tuning varactor of VCO for the different frequency output. It is not possible to use a battery to afford a few different values of DC voltage. The battery is not a convenient DC voltage source.

This paper presents a method to suppress the noise coupled with the output DC voltage of DC power supply in measuring the phase noise of VCO free running. We connected a resistor to DC bias input port of VCO tuning varactor in series. The phase noise of VCO free running biased by the DC power supply can be lowered sharply due to connecting a resistor in series.

2. EXPERIMENT AND RESULTS

In our experiments, we separately used a battery and a DC power supply to afford a 1.48 DC voltage to bias the tuning varactor of VCO. Fig. 1 shows the phase noise of VCO free running biased directly by the DC power supply. Fig. 2 shows the phase noise of VCO free running biased directly by a battery.

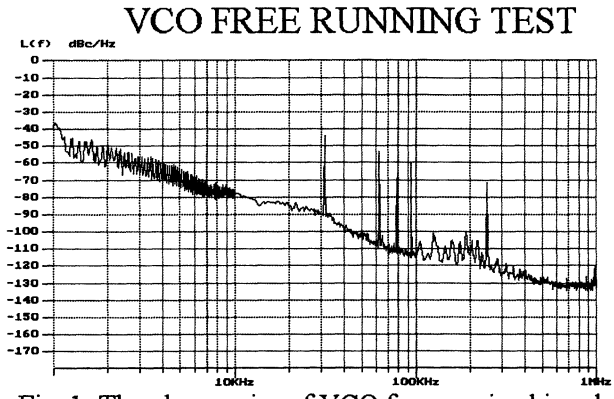


Fig. 1: The phase noise of VCO free running biased directly by the DC power supply.

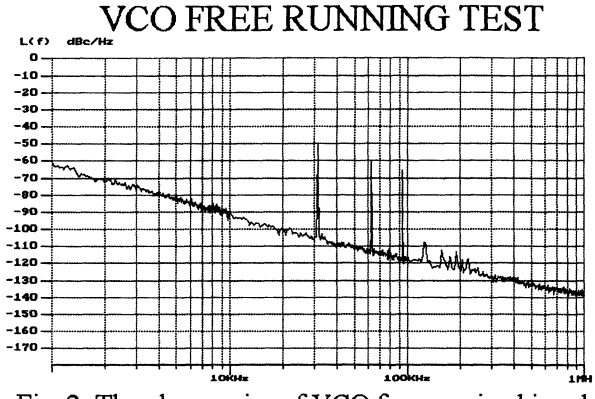


Fig. 2: The phase noise of VCO free running biased directly by a battery source.

To suppress the noise coupled with the DC voltage output of a DC power supply, a resistor was connected to the DC bias input port of VCO tuning varactor in series. The block diagram of connecting a resistor to the DC bias input port of VCO tuning varactor in series is shown in Fig. 3.

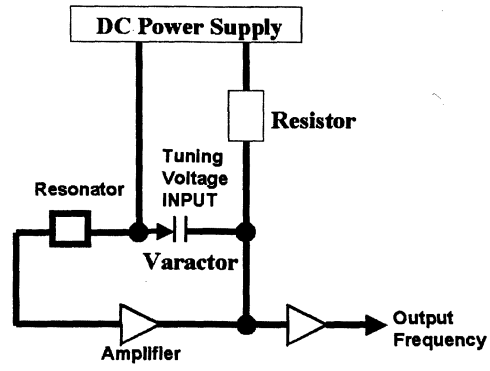


Fig. 3: The block diagram of connecting a resistor to the DC bias input port of VCO tuning varactor in series.

We separately used 10 ohm, 100 ohm, 1000 ohm and 100,000 ohm resistors in measurements. The measuring results of connecting the different resistor to the DC bias input port of VCO tuning varactor in series are shown in Fig. 4, Fig. 5, Fig. 6, and Fig. 7.

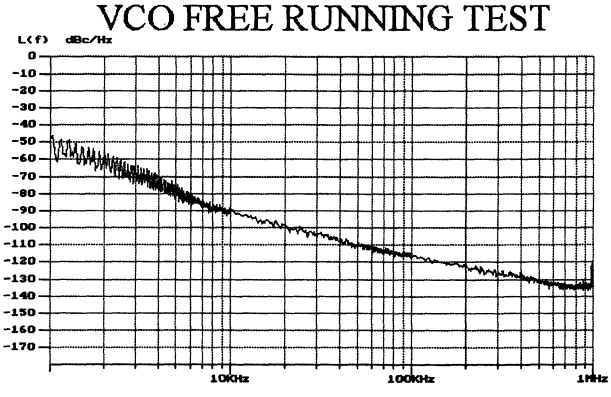


Fig. 4: The phase noise of VCO free running biased by a DC power supply with a 10 ohm resistor in series.

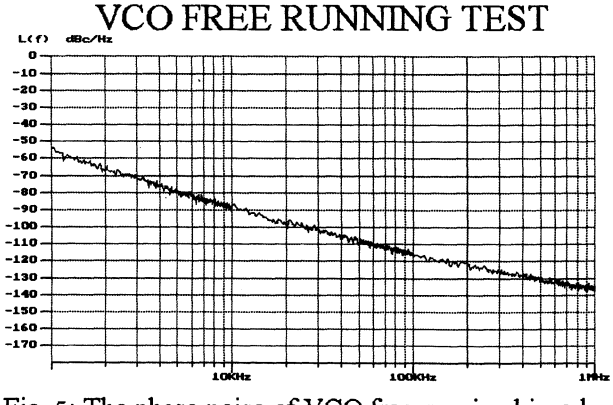


Fig. 5: The phase noise of VCO free running biased by a DC power supply with a 100 ohm resistor in series.

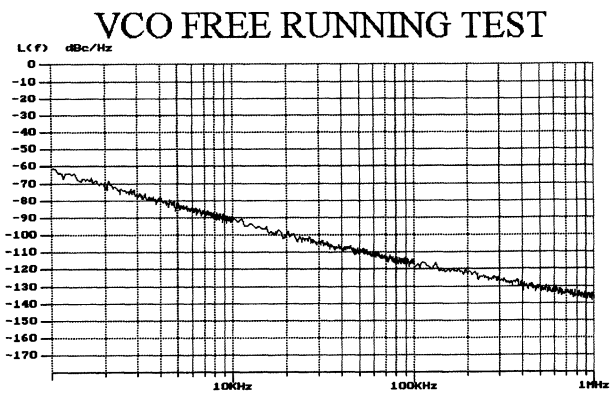


Fig. 6: The phase noise of VCO free running biased by a DC power supply with a 1000 ohm resistor in series.

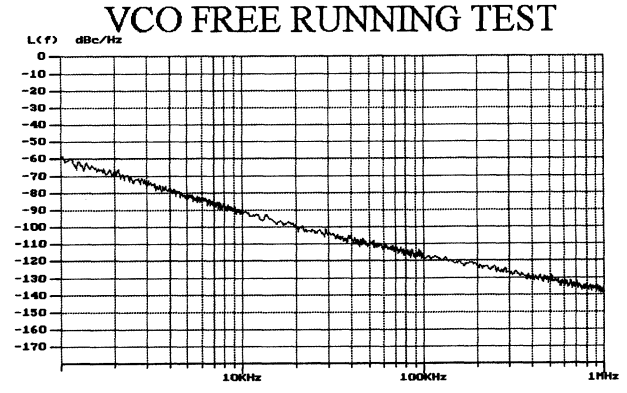


Fig. 7: The phase noise of VCO free running biased by a DC power supply with a 100,000 ohm resistor in series.

3. DISCUSSION

The degradation of noise coupled with the DC power supply output is obvious in the phase noise measurement of VCO free running shown in Fig. 1. In Fig. 1, it shows the sharp degradation caused by a few harmonic components of power line frequency between 1 kHz offset and 10 kHz offset. The degradation caused by the harmonic components of power line frequency is not appeared between 1 kHz offset and 10 kHz offset in Fig. 2 which shows the phase noise of VCO free running biased directly by a battery. The degradation caused by the noise coupled with the DC power supply output has been improved by connecting a 10 ohm resistor to the DC bias input port of VCO tuning varactor in series shown in Fig. 4. In Fig. 4, the degradation caused by the harmonic components of power line frequency still appears between 1 kHz offset and 10 kHz offset. The 10 ohm resistor has a suppressing effect on the noise of high frequency coupled with the DC power supply output between 10 kHz offset to 1 MHz offset. When an 100 ohm resistor was connected to the DC bias input port of VCO tuning varactor in series, the suppressing effect of the noise coupled with the DC power supply output is obvious in Fig. 5.

The phase noise level has 15 dB lowered at 1 kHz offset by connecting an 100 ohm resistor to the DC bias input port of VCO tuning varactor in series. We also separately connected an 1000 ohm and an 100,000 ohm resistor to the DC bias input port of VCO tuning varactor in series to measure the phase noise of VCO free running. The suppressing effects are similar between connecting an 1000 ohm resistor and an 100,000 ohm resistor shown in Fig. 6 and Fig. 7. The phase noise has sharply lowered by connecting a resistor, which resistance is more than 100 ohm in our experiments, to the DC bias input port of VCO tuning varactor in series.

4. CONCLUSIONS

It is an easy method to connecting a resistor to the DC bias input port of VCO tuning varactor in series. The suppressing effect of noise coupled with the DC power supply output is sharply obvious by connecting a resistor in series. When the resistance is more than 100 ohm in our experiments, the phase noise of VCO free running is lowered more than 10 dB. Some high frequency spurs are also suppressed by connecting a resistor in series. It has lowered the phase noise 20 dB at 1 kHz offset by connecting an 1000 ohm or an 100,000 ohm resistor in series in our experiments. From the experimental results, the noise coupled with the DC power supply output is not a problem in the phase noise measurement of VCO free running. The interference caused by the noise coupled with the DC power supply output can be moved by connecting a resistor in series. The DC power supply can afford any value of DC voltage without degradation of noise due to connecting a resistor in series for the phase noise

measurement of VCO free running.

5. REFERENCES

[1] Thomas E. Kiess, "Noise Filtering of DC Voltages in Both the Time and Frequency Domain", IEEE Transaction on Instrumentation and Measurement, VOL. 47, NO. 3, JUNE 1998.

[2] Martin Rowe, "Build a Low-Noise Voltage Source", Test & Measurement World, August 1998.

A 100 MHz TO 5.4 GHz FREQUENCY MULTIPLIER WITH LOW PHASE NOISE DEGRADATION

G. Galzerano, C. Svelto, C. Riccardi, and E. Bava

Dipartimento di Elettronica e Informazione – Politecnico di Milano, INFM, and CNR - CSTS
Piazza Leonardo da Vinci 32, 20133 Milano, Italy
E-mail: Gianluca.Galzerano@PoliMI.IT Phone: +39 02 2399 3601 Fax: +39 02 2399 3413

ABSTRACT

A 5.4 GHz frequency reference with low phase noise has been realized by means of a 100 MHz quartz oscillator followed by a frequency multiplier. This microwave frequency reference is intended for use as a low phase noise reference source for an optical frequency comb generator in the near infrared with a potential collapse frequency in excess of 10 THz.

1. INTRODUCTION

In the fields of time and frequency metrology, communications, and navigation systems very low phase-noise frequency references at around 10 GHz find useful applications.

In the case of frequency metrology it is extremely important to compare and measure frequency standards at different wavelengths from the ultra-violet to the microwaves [1]. Several non-linear techniques and devices are used to realize these frequency connections and they typically cover frequency differences up to some terahertz (*i. e.*, Metal-Insulator-Metal structures, optical-comb generators). With this last device to obtain such high-frequency difference measurements it is therefore necessary to realize high-order frequency-multiplication factors starting with an extremely pure oscillator (ultra-low phase-noise source) to be able to perform high-accuracy measurements [2, 3].

Quartz oscillators at 100 MHz (OCXO) with phase-noise levels of −170 dBc/Hz at 1 kHz from carrier (typical corner frequency value) are now available which could guarantee a coherent output at a few terahertz if a low phase noise frequency synthesizer is built.

In this paper a low noise frequency multiplier from 100 MHz to 5.4 GHz is described. Efficiency and phase noise measurements of the realized systems are also reported.

2. THE FREQUENCY MULTIPLIER CHAIN

In order to obtain the × 54 frequency multiplication factor, the realized circuit is composed by three different non-linear stages. From the scheme reported in Fig. 1 it is possible to recognize the different stages: a × 3 multiplier based on Schottky-diode clipper (100 MHz to 300 MHz) followed by a high-power linear amplifier (class A), a passive band-pass filter at 300 MHz, an active doubler using class C configuration (300 MHz to 600 MHz), a passive band-pass filter at 600 MHz, and a final stage composed by a step-recovery-diode (SDR) placed inside a microwave copper-resonator in order to select the output at 5.4 GHz.

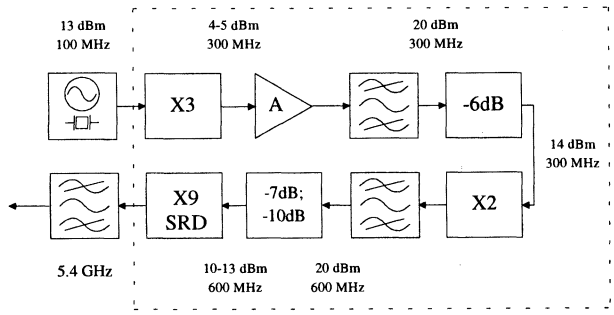


Figure 1. Block diagram of the frequency multiplier.

Starting with an input signal of 13 dBm at 100 MHz (typical output power level of a commercial low phase noise crystal oscillator), 12 dBm at a frequency of 5.4 GHz have been obtained at the resonator output. In order to completely characterize the whole system, intermediate conversion efficiencies have been measured. Particular attention was given to the design of the third harmonic stage (reported in Fig. 2) to reduce additive phase noise degradation [3]. According to [4] this circuit is designed to exhibit a low conversion from amplitude to phase modulation.

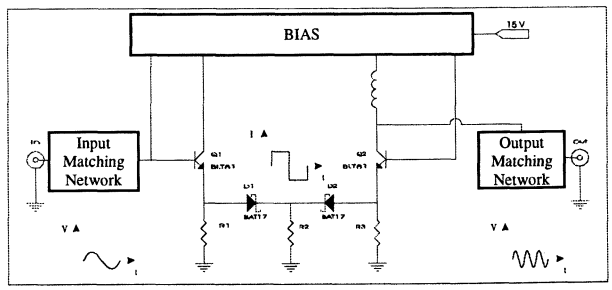


Figure 2. Schematic design of the third order multiplier.

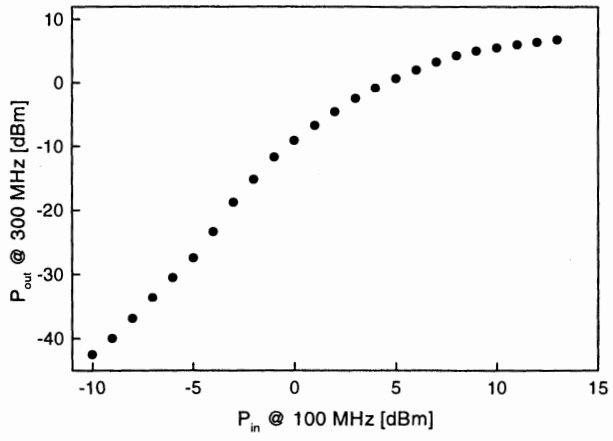


Figure 3. Output power at 300 MHz as a function of the 100-MHz input power.

The achieved third harmonic efficiency is reported in Fig. 3 as a function of the 100 MHz input power. In order to further reduce the spurious harmonic contributions, the 300 MHz output has been filtered using a fifth order pass-band Chebycheff network (insertion loss and bandwidth of 1.7 dB and 23 MHz, respectively).

The 300 MHz to 600 MHz conversion has been realized using a class C doubler (Fig. 4) followed by a fifth order 600 MHz pass-band Chebycheff filter (insertion loss and bandwidth of 2.3 dB and 65 MHz, respectively). An output power of 22 dBm at 600 MHz has been obtained using 14 dBm at 300 MHz.

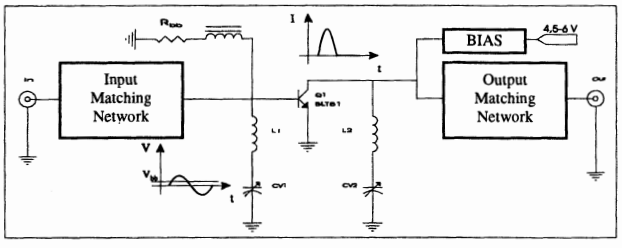


Figure 4. Schematic design of the class C doubler.

The last stage of the multiplier is based on a SRD placed inside a coaxial microwave cavity. The adopted scheme is reported in Fig. 5. The resonance frequency can be tuned around 5.4 GHz by means of a teflon screw.

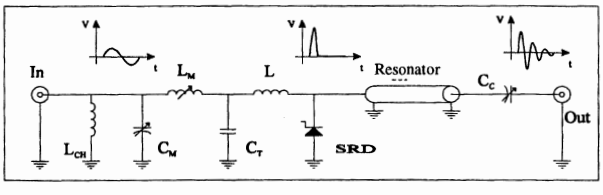


Figure 5. Schematic design of the step recovery diode multiplier.

The output power at 5.4 GHz for different values of the available power at 600 MHz is reported in Table 1.

P_{in} @ 600 MHz	P_{out} @ 5.4 GHz
16 dBm	4 dBm
19 dBm	6 dBm
20 dBm	11 dBm
22 dBm	12 dBm

Table 1. Output power at 5.4 GHz versus input power at 600 MHz.

3. PHASE NOISE MEASUREMENTS

In order to characterize the phase-noise degradation introduced by the multiplier, a second multiplier chain until 600 MHz has been realized. Using a double-balanced-mixer, the phase noise degradation due to the multipliers has been measured using the phase discriminator technique (Fig. 6) both at 300 MHz and 600 MHz.

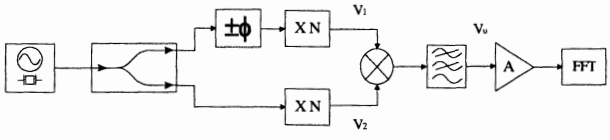


Figure 6. Phase noise measurement based on discriminator technique.

At the mixer inputs the two RF voltages can be represented by

$$V_1 = A_1(1+a_{01}+a_1)\cos\left[2\pi\nu_0 t+\phi_{01}+\phi_1\right]$$
$$V_2 = A_2(1+a_{02}+a_2)\cos\left[2\pi\nu_0 t+\phi_{02}+\phi_2+\theta\right] \quad (1)$$

where a_{0i} and ϕ_{0i} are the amplitude and phase noise processes at the multiplier output, induced by the common input signal, ϕ_i and a_i are the additive phase noise and amplitude noise, respectively, produced by the frequency multipliers and θ is the imposed phase difference between V_1 and V_2 in order to obtain an ideal phase quadrature between V_1 and V_2. Under usual conditions $\phi_i << 1$ and $\phi_{0i} << 1$, and the IF output signal of the mixer is

$$V_u \cong \frac{1}{2}K_m\ A_1A_2(1+a_{01}+\ a_{02}+a_1+\ a_2)\left[\phi_2(t)-\phi_1(t)+\vartheta_0\right]$$
$$= K_d\ (1+a_{01}+\ a_{02}+a_1+\ a_2)\left[\phi_2(t)-\phi_1(t)+\vartheta_0\right] \quad (2)$$

where K_m is the conversion loss of the adopted mixer, $\vartheta_0 = \theta+\frac{\pi}{2}$ is the quadrature residual error, and

$K_d = \frac{1}{2} K_m A_1 A_2 = \frac{\Delta V}{\Delta t} \frac{T}{2\pi}$ represents the phase noise discriminator sensitivity, which can be measured through the slope of the beat note of period T produced by two close input frequencies. In this way the power spectral density of the output voltage observed by a spectrum analyzer is mainly determined by

$$S_{V_u}(f) \cong K_d^2 S_{\phi_2 - \phi_1}(f) \quad (3)$$

where S_ϕ is the phase noise spectral density.
If the quadrature is not perfect and different AM/PM conversion effects are taken into account, the following additional contributions have to be considered

$$S_{unq}(f) = 4K_d^2 \theta_0^{\,2} \left[S_\alpha(f) + \tfrac{1}{2} S_{\alpha_0}(f) \right] \quad (4)$$

$$S_{uap}(f) = (K_1 - K_2)^2 \left[S_{\alpha_0}(f) \right] \quad (5)$$

where $S_\alpha(f)$ and $S_{\alpha 0}(f)$ are the additive and input amplitude spectral densities and K_i is the AM to PM conversion factor of a multiplier.

The adopted measurement set-up is reported in Fig. 7. By means of a synthesizer phase locked to a 10 MHz crystal oscillator the phase discrimination coefficient was measured at 300 MHz and 600 MHz, for a beat frequency of 1 kHz, as reported in Table 2.

f_0 [MHz]	$\Delta V / \Delta t$ [V/s]	K_d [dB]
300	1730	-11.2
600	1800	-10.9

Table 2. Phase discrimination slope sensitivity.

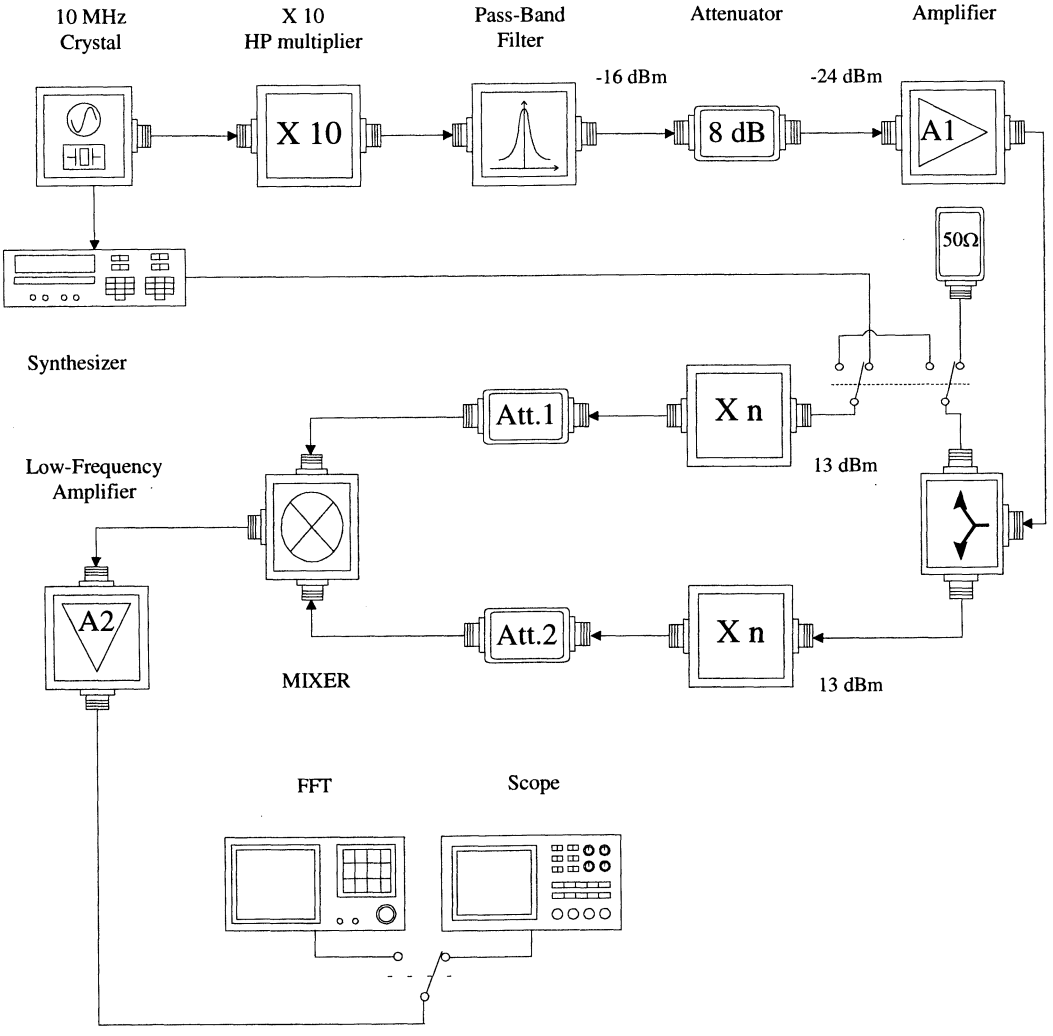


Figure 7. Phase-noise measurement set-up.

From the measurements at both 300 MHz and 600 MHz (Fig. 8), a -156 dBc/Hz at 100 MHz phase-noise pedestal has been measured at 1 kHz from the carrier. This white phase-noise level corresponds to a -121 dBc/Hz at 5.4 GHz. With this phase noise levels and with a typical bandwidth of 23 MHz for the band-pass filter at 300 MHz it is possible to reach a collapse frequency of 1.5 THz. The collapse frequency is related to the total phase noise power, ϕ^P, by the relation

$$\nu^{col} = n^{col}\nu_0 = \frac{\nu_0}{\sqrt{\phi^P(\nu_0)}} \tag{6}$$

where ν_0 is the central frequency of the oscillator, n^{col} is the collapse integer factor, and the phase noise is defined as

$$\phi^P(\nu) = \int_{f_0}^{\infty} S_\phi(\nu, f)\, df \ll 1 \tag{7}$$

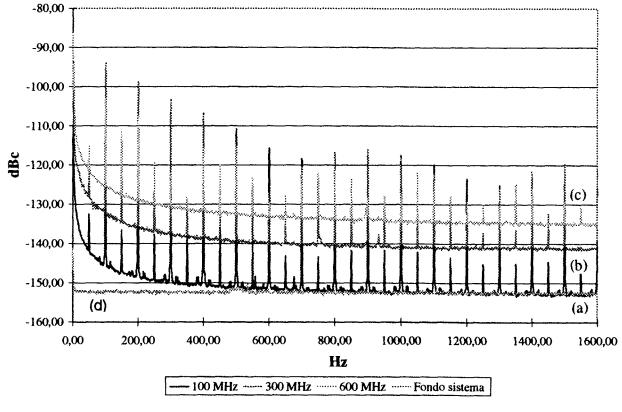


Figure 8. Phase-noise degradation measurements.

If a commercial high-quality factor 300 MHz band-pass filter is used ($Q \approx 1000$) the collapse frequency will be further increased near 13 THz.

4. CONCLUSIONS AND DEVELOPMENTS

Efficient $\times 54$ frequency multiplication from 100 MHz to 5.4 GHz has been obtained. From the characterization of phase noise degradation due to the 100 MHz to 600 MHz multiplier a collapse frequency in excess of 10 THz could be achieved.

In addition a novel design for a low-phase noise reference at frequencies of 9 GHz, 10 GHz and 11 GHz based on a low phase-noise Surface Acoustic Wave VCO at 1 GHz (RFM HO-4001) is now in progress. This more agile system consists of a linear power amplifier (Miteq AMF-4B-010020-20-33P), a step-recovery-diode (Herotek GC1000RC), and a tunable microwave cylindrical cavity. Moreover, by means of an harmonic mixer and of a 100 MHz low phase-noise crystal it is therefore possible to lock the VCO against the reference oscillator and reduce the VCO phase noise near the carrier.

ACKNOWLEDGMENTS

The authors gratefully acknowledge Dr. Giorgio Brida for his valuable technical support in phase noise measurements.

REFERENCES

[1] H. R. Telle, D. Meschede, and T. W. Hansh, Opt. Lett. 15, pp. 532-534, 1990

[2] F. L. Walls and A. De Marchi, "RF spectrum of a signal after frequency multiplication: measurements and comparison with simple calculation", *IEEE IM 24*, n. 3, 1975, pp. 210-217

[3] A. De Marchi, A. Godone, and E. Bava, "Phase-coherent synthesis and precision frequency measurements in the far infrared", Trans. IEEE on IM, vol. 30, n. 2, 1981, pp. 132-133.

[4] E. Bava, G. P. Bava, A. De Marchi, and A. Godone, "Measurement of static AM-PM conversion in frequency multipliers", IEEE Trans. IM, vol. 48, n. 1, 1977 pp. 33-38

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

Simultaneous amplitude and frequency noise analysis in Chua's circuit.

J.-M. Friedt, D. Gillet, M. Planat
LPMO/CNRS, 32, avenue de l'observatoire, 25044 Besançon Cedex
friedt@lpmo.univ-fcomte.fr

Abstract

Chua's circuit is a non-linear electronic circuit stripped to its basic elements: a parallel LC circuit connected though a resistor to a parallel CN circuit, N being a non-linear conductance. Chua's circuit displays a complex evolution: period doubling pattern from a periodic oscillation to a chaotic behavior as one of the parameters of the circuit increases. We have decided to study statistical properties of the "instantaneous" frequency and amplitude signals when the inductor parameter L evolves.

We performed a large number of simultaneous frequency and amplitude data acquisitions which were validated by plotting the bifurcation diagrams as well as the standard and the Allan deviation.

1 Introduction

We studied the simultaneous evolution of the frequency and amplitude noise characteristics in a chaotic electronic circuit. The advantage of studying an electronic circuit, in addition to the fact that such circuits can be used for data transmission, is that a large number of samples can easily be obtained and thus statistically significant values can be computed. Chua's circuit strips a chaotic electronic circuit to its basic elements, and isolates the bifurcation parameter to a single component. We will here briefly describe the electronic circuit and the three coupled differential equations describing its behavior, the experimental setup we have developed in order to obtain experimental data, the method used for validating the data we recorded and finally the signal processing techniques we applied to these data.

1.1 Description of Chua's circuit

Chua's circuit is composed of two parallel circuits LC_1 and C_2N (N being a negative conductance characterized by two slopes, cf figure 1) connected by a resistor R (figure 1).

Chua's circuit is thus described by three coupled differential equations as follows. They are obtained by describing the intensity at nodes X and Y in figure 1 and the voltage at node X:

$$\begin{cases} C_1 \frac{dv_{C_1}}{dt} = \frac{1}{R}(v_{C_2} - v_{C_1}) - g(v_{C_1}) \\ C_2 \frac{dv_{C_2}}{dt} = \frac{1}{R}(v_{C_1} - v_{C_2}) + i_L \\ L \frac{di_L}{dt} = -v_{C_2} - i_L R_L \end{cases}$$

where R_L is the internal resistance of the inductor. In our circuit $R_L \simeq 3.3\ \Omega$, $C_1 \simeq 20.3$ nF, $C_2 \simeq 1847$ nF and $R \simeq 1.46$ kΩ.

As opposed to the equations leading to Rössler or Lorentz attractors [4, pp. 77 and 131] in which the non-linearity is included as an explicit product between two of the differential equations variables, the only non linearity in this system comes from the equation including $g(v_{C_1})$ which describes the current to voltage response of the non-linear conductance N. A simulation of the current to voltage relation $I_N = g(V_N)$ is given in figure 1 (bottom) [5]:

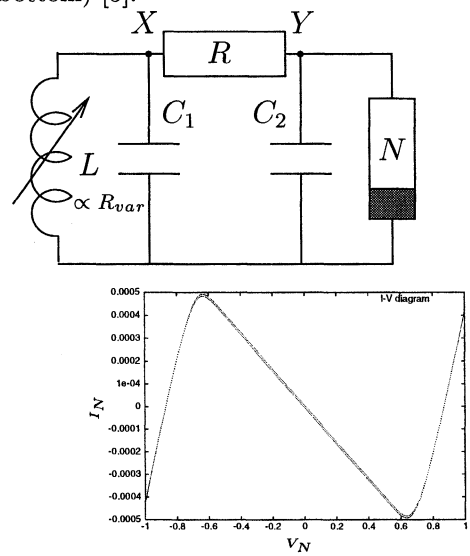


Fig 1: Schematic of Chua's circuit (top) and Spice simulation of the current to voltage characteristic of the non-linear element N (bottom).

We used a simulated inductor described by Weldon [1]: it is less voluminous and its value is defined by a variable resistor and thus can be computer controlled[1]

1.2 Data acquisition circuit

The three variables in Chua's circuit are the voltages at the two capacitors and the current running through the inductor. We have chosen for convenience to measure the voltage of one of the capacitors.

[1]The inductor's value L is related to that of the variable resistor R_{var}, which is the parameter shown in the figures of this article, from the relation $L = R_L \times (R_{var} + 30000) \times 0.1 \times 10^{-6}$ (30000 is the value of a resistor and 0.1×10^{-6} that of a capacitor).[1] With $0 < R_{var} < 24$ kΩ this leads to $L \sim 13.5$ mH (± 4 mH).

The parameter we have chosen to vary is the inductor's value L. Our circuit includes an operational-amplifier based simulated inductor whose value is defined by a computer controlled variable resistor. This parameter is thus much simpler to modify than the capacitors' values. The circuit also evolves more smoothly by shifting the L value rather than the R value. The behavior of the oscillating circuit does not seem to have been perturbed when replacing the mechanical variable resistors with electronic, computer controlled variable resistors (they are usually used for remote controlled audio systems). We have been careful in separating the various power supplies (for Chua's circuit, analog signal pre-processing card, computer and data acquisition cards) and to insert buffers where the probes were reading the voltages.

We first attempted to estimate the oscillation frequency of Chua's circuit in its first periodic mode by replacing the non linear element N by a linear negative resistor of value R_N ($R_N \simeq -1300\ \Omega$). By additionally assuming that the internal resistance of the inductor is null ($R_L = 0$) we obtain a variable frequency $f = \frac{1}{2\pi}\sqrt{\frac{\frac{C_1}{L} - \frac{1}{R_N^2}}{C_1(C_1+C_2)}}$, versus L (figure 5, bottom). Numerical application with the values given previously, and for a value of the parameter $L = 14$ mH, gives $f \simeq 760$ Hz, in close agreement with the observed values. Adding the term that was at first neglected makes the literal solution much more complex (after solving using the software Maple V) and does not change the numerical estimation of the frequency.

The electronic circuit used for shaping the analog signals includes a follower circuit, followed by a comparator (for frequency measurements) and a peak holder (for amplitude measurements) made of a germanium diode (so as to minimize the voltage loss), a high value capacitor and an analog switch (so as to be able to discharge the capacitor and the reset the amplitude measurement).

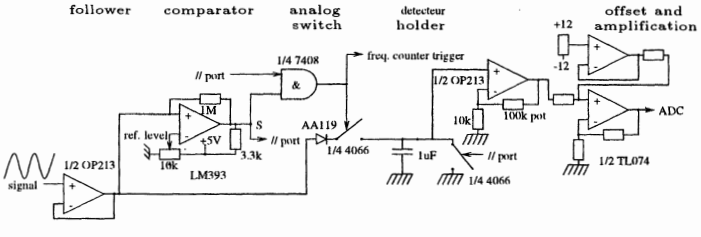


Fig 2: Schematic of the circuit used for measuring instantaneous frequency and the amplitude of the same period.

```
while S is high, wait
                    wait the end of the current period
set the ''AND'' gate triggering the frequency counter
                    get ready for a measurement
while S is low, wait
                    end of the low half-period
while S is high, wait
        end of the high half-period, measurement in progress
read amplitude and frequency measurements
                    data acquisition
capacitor reset
```

Measurement algorithm (this loop is executed for each new period to be measured)

We now understand that an acquisition only occurs when both the level of the voltage from Chua's circuit is higher than a given preset level (output S of the comparator) and the computer is ready for a new acquisition (second input of the "AND" gate set by the parallel port).

We have been able to verify, using a digital storagescope, that we indeed observed the expected signal: both amplitude (fig 3, top) and frequency (fig 3, middle) are measured on a single period. Amplitude must be read quickly, before the frequency, in order to avoid noise due to leakage currents coming from the analog switch (fig 3, bottom).

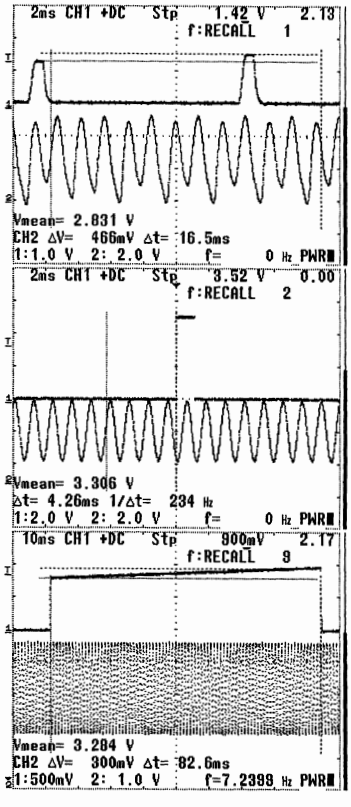


Fig 3: Storagescope outputs resulting from our data acquisition algorithm: top, the height of a pulse chosen randomly among the quasi-periodic signal is indeed proportional to the amplitude of this same period. Middle: the trigger pulse of the frequency

counter is indeed defined by the voltage going above a given preset value. Finally, bottom figure, the slow shift of the amplitude at the input of the analog to digital converter due to the leakage current leaving the analog switch. The conversion time during an actual acquisition is 70 times shorter than this pulse.

We found interesting that an experimental error lead to an optimized sweep of the value of the parameter L. As the computer controlled variable resistors we selected only allowed a choice of 64 resistance values, we decided to connect in series two variable resistors of different maximum values (10 and 50 kΩ) connected in parallel with fixed resistors so that the range of one of the variable resistors (10 kΩ) matches one step of the other variable resistor (50 kΩ). However, the composition law of resistors connected in parallel removes the linearity of adding the values of the two variable resistors connected in series. Hence, the composition of the successive values of the two variable resistors does not increase the resolution of each steps as expected, but gives a better distribution of the values taken by L than would have been obtained by connecting the two variable resistors in series, as shown figure 4.

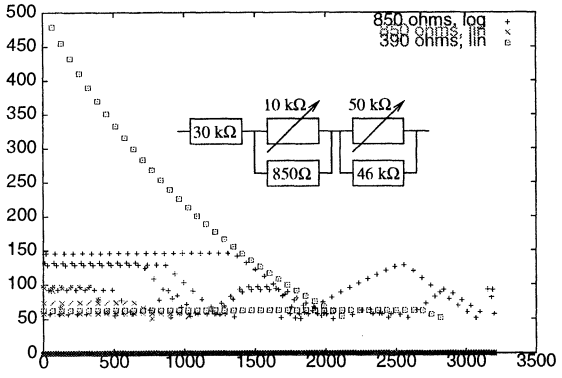


Fig 4: Difference between two successive values of the parameter L (as the sum of the resistance of the two variable resistors, each of which is in parallel with a constant resistor): comparison between various sweeping modes. The 50 kΩ variable resistor always sweeps a range from 0 to 39 kΩ in 50 linearly spaced steps (each of 780 Ω), while the 10 kΩ variable resistor sweeps either logarithmically (increases by 1, 2, 4, 8, 16 and 32 steps, resulting in $50 \times 7 = 350$ files) or linearly (10 kΩ increase in 64 steps, resulting in $50 \times 64 = 3200$ files). The curves shown on this figure result from varying the 10 kΩ variable resistor (lin or log) and the value of the resistance connected in parallel with this variable resistor (390 Ω – designed to fill the range between two steps of the 50 kΩ variable resistor – or 850 Ω – for which two successive ranges overlap and hence requiring some processing in order to put the files in the increasing values of R_{var}. We plotted here the resistance difference of R_{var} between two steps of the 10 kΩ variable resistor versus the index of the step.

2 Results

We graphically validated the data acquisition method (figure 5) by plotting the values taken by the amplitude and frequency time series versus the value of the L parameter. One can observe the first bifurcation, followed by the chaotic behavior. The widening of the frequency curve to the right of the graph (for large values of L) characterizes the transition from a Rössler-like attractor to a double scroll attractor, for which our measurements are harder to interpret (the separation of each period by the voltage crossing a reference level is no longer possible in this kind of oscillation mode as it was in the periodic regime).

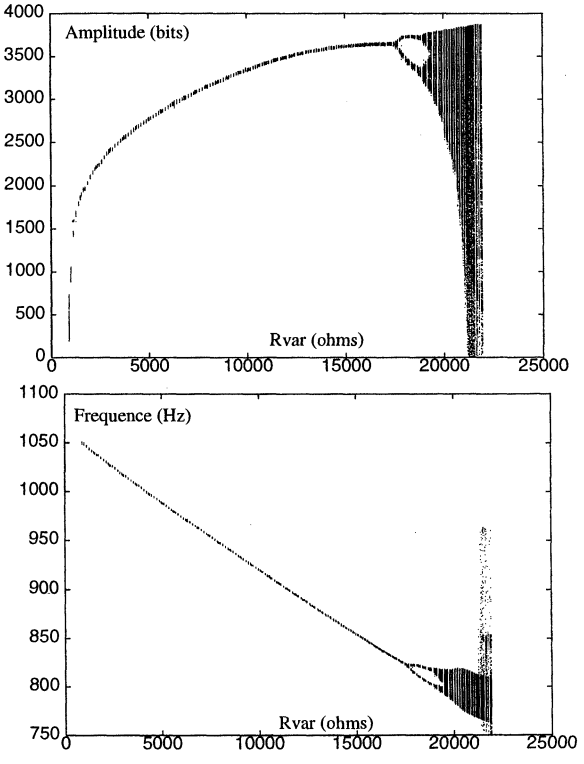


Fig 5: Bifurcation diagrams both in amplitude (top) and frequency (bottom). Each time series (plotted for a given abscissa corresponding to a value of the L parameter) includes 1000 points

2.1 Statistical results

As could be expected, the standard deviation quickly increases after the first bifurcation. This quick increase is partly due to our data analysis method as we do not separate various branches of the values taken by the time series (either in frequency or in amplitude) as such a distinction is no longer possible once the chaotic region is reached.

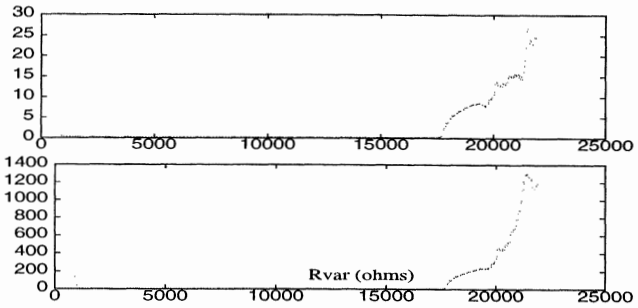


Fig 6: Evolution of the standard deviation with increasing values of the L parameter. Top: standard deviation of the frequency. Bottom: standard deviation of the amplitude.

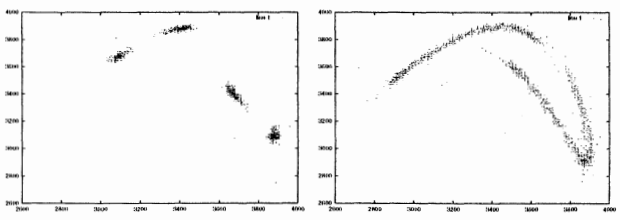
An automatic measurement of the slope of Allan variance [3], although quite inaccurate, always gives an average value of -0.5, whatever the oscillation mode. This result means that the frequency noise is a white noise. However, this result must be interpreted as showing that our data acquisition algorithm randomly chooses a period in the oscillating signal – depending on the communication time of the GPIB interface between the frequency counter and the computer – rather than the statistics of the signal itself. We will see later that different results are obtained when studying continuous time series (including *all* periods in a given time interval) rather than values chosen randomly in the oscillating signals (figure 8).

2.2 Sequential results

While processing the data obtained statistically, we have developed a frequency-counter able to measure frequencies around 1 kHz with a 15 mHz resolution, and most importantly to send the result to the computer in less than half a millisecond (duration of half a period), allowing the measurement of the properties of all the consecutive periods in a given time lapse.

Obtaining data on all the periods in a given time lapse shows clearly the acquisition noise due to a coupling of Chua's oscillator with the 50 Hz signal from the electric mains. Bifurcation diagrams obtained with this new data acquisition mode are noisier than those obtained previously.

Obtaining all the periods in a given time lapse was confirmed by drawing the return maps for various oscillation regimes (figure 7).



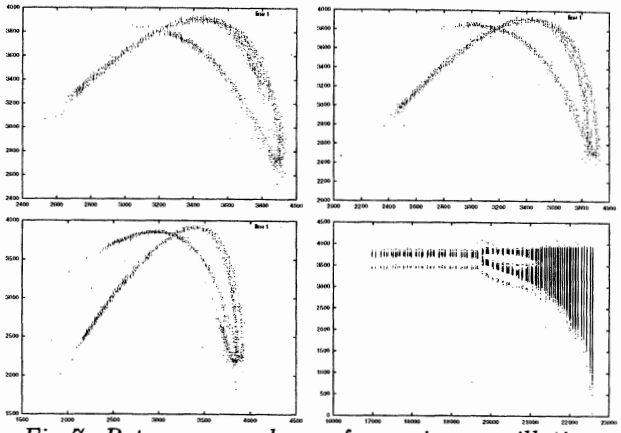


Fig 7: Return maps drawn for various oscillation regimes (from left to right and top to bottom, the R_{var} parameter is equal to 21254 Ω, 21484 Ω, 21711 Ω, 21933 Ω and 22151 Ω (which correspond respectively to values of 16.91 mH, 16.99 mH, 17.06 mH, 17.14 mH and 17.21 mH of L), and zoom on the bifurcation diagram's zone for which these return maps were shown

Allan variance is usually defined on frequency data. As we were not able to measure this variable quickly enough at the moment, we supposed as a first approximation that the behavior of the amplitude is similar to that of the frequency (as graphically shown by the bifurcation diagrams of fig. 5, although this hypothesis is still to be demonstrated). A first estimate of the slope of Allan variance versus the bifurcation parameter was thus computed on the sequential amplitude data (figure 8). We here expect, without being able to check this hypothesis, that the behavior of the amplitude noise is similar to that of the frequency noise.

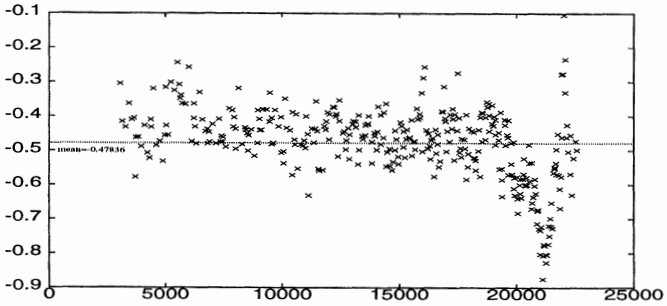


Fig 8: Evolution of the slope of Allan variance (computed on the experimental amplitude data) during the evolution from a periodic oscillation to a chaotic regime.

We observe a continuous evolution of Allan variance from a mean value of -0.5 (white frequency noise of thermal origin) to a slope of -1 when the circuit is in a chaotic regime. A slope of -1 of the Allan variance characterizes a noise spectrum of slope +1 or +2 (we

have not yet been able to distinguish between the two regimes), originating respectively from electronic noise in the oscillator or an external white noise [3].

3 Conclusion

We have successfully built a circuit for statistically measuring instantaneous frequency and amplitude of Chua's circuit. We checked the validity of our results by plotting the bifurcation diagrams, and analyzed the standard deviation of the time series.

We then modified the acquisition algorithm in order to measure all instantaneous amplitudes in a given time lapse. Such measurements enable the evaluation of Allan variance. We showed that its slope evolves from -0.5 to -1 during the evolution of the behavior from periodic to chaotic. We furthermore displayed the return maps of the amplitude for various values of the bifurcation parameter.

Our future work includes completing a fast frequency counter. It will be combined with the fast frequency acquisition algorithm for simultaneous amplitude and frequency measurements of all periods in a given time lapse.

Another future research idea is to study the ability of neural networks, or simpler derivated algorithms, to effectively predict the evolution of the measured parameters (frequency and amplitude) of the chaotic circuit.

References

[1] Thomas P. Weldon, *An inductorless double scroll chaotic circuit*, Am. J. Phys **58** (10), October 1990

[2] R.N. Madan, *Chua's circuit: a paradigm for chaos*, World Scientific Series in Non-linear Science (Series B, vol 1) (1993)

[3] François Vernotte, *Approche multivariance du bruit de fréquence des oscillateurs*, Ann. Télécommun. **51** (9-10) (1996)

[4] P. Bergé, Y. Pomeau, Ch. Vidal, *L'ordre dans le chaos*, Hermann (1984)

[5] A. Tamasevicius, G. Mykolaitis, A. Namajunas, *Double scroll in a simple '2D' chaotic oscillator*, Electronics Letters **32** (14), July 4th 1996

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

THREE-DIMENSIONAL OPTIMAL KALMAN FILTER FOR TIME ERROR, FREQUENCY, AND AGING ESTIMATES

Yu.S.Shmaliy*, A.V.Marienko, A.V.Savchuk**, Jörg Hahn, Rainer Krämer***

*Electronics Dept., Guanajuato Univ., S/N 36730, Salamanca, Gto, Mexico, shmaliy@salamanca.ugto.mx
**"Sichron" Center, 4 Skrypnyka Street, Kharkiv, 310057, Ukraine, sich@kharkov.ua
***German Aerospace Center, ICN, P.O. Box 1116, D-82230 Wessling, joerg.hahn@dlr.de

ABSTRACT

We present the results of optimal filtering of time error, frequency, and frequency aging of the crystal oscillator based on the Kalman filter matched with the oscillator time model by its three-dimensional algorithm, and on the reference timing signals provided by Motorola GPS UT+ Oncore receiver. We consider the oscillator timing model, the filter equations, and regression approach for definition of the coefficients. We compare the estimates provided with those obtained by direct measurement based on the rubidium oscillator.

1. INTRODUCTION

Fast and accurate optimal Kalman filtering of time error, frequency, and frequency offset of slaved sources (crystal and rubidium) is extremely important to create frequency and time standards looked after GPS timing signals. It is known that excellent accuracy 10^{-12} is obtained in practice by smoother approximately through 24 hours, and it is the reason why efforts are now underway to create the fast Kalman algorithm and obtain the same accuracy for the minimal processing time. Allan and Barnes have showed in [1] that the filtering effect strongly depends on a measuring time interval Δ that should be of 100...1000 sec for GPS timing signals. Later, series of reports had been devoted to the Kalman filters application especially for the phase, time errors, and frequency real-time estimates with prediction [2-5]. The papers are based on the Kalman fundamental approach [6] for the discrete-time optimal estimates with a white Gaussian noise and developed later by many authors in [7-8], for instance.

The report addresses the results of the three-dimensional Kalman algorithm to be used to estimate the synchronization time errors based on the timing signals of Motorola Oncore UT+ Receiver. While studied, we dealt with many measured and estimated functions of the time error, frequency, and aging related to the OCXO and rubidium standard. It allowed comparison and selection the appropriate optimal filter structure separately for each type of an oscillator and for measurement and synchronization tasks. As it was expected, different algorithms gave different filtering errors. Only when the frequency drift showed a quasi-stationary nature then it had not been important what type of the Kalman filter was used. On the contrary, in the non-stationary case, only an optimal filter provided sufficient accuracy being matched with an oscillator time model. Hence, the major point is the timing model of the steered source. One must expect with this that a rubidium source of frequency may be disciplined by GPS signals with relatively smaller fractional error with respect to that of a crystal oscillator because of its better prediction of the phase behavior in time.

2. MATHEMATICAL MODEL OF AN OSCILLATOR TIMING SIGNAL

Let us consider a generalized oscillator timing model [10]. In the ideal case, the total phase $\Phi_{id}(t)$ of an ideal timing signal is presented by

$$\Phi_{id}(t) = 2\pi\nu_{nom}t \qquad (1)$$

where ν_{nom} is called nominal frequency. Correspondingly, total instantaneous phase model of actual timing signals $\Phi(t)$ is modeled as:

$$\Phi(t) = \Phi_0 + 2\pi\nu_{nom}(1+y_0)t + \pi D\nu_{nom}t^2 + \varphi(t) \qquad (2)$$

where: Φ_0 is the initial phase offset, y_0 is the fractional frequency offset from the nominal value ν_{nom} (mainly due to finite frequency settability of the clock); D is the linear fractional frequency drift rate (basically representing oscillator aging effects); $\varphi(t)$ is the random phase deviation component. Based on the definition of time error and the above model (2), the following model for time error $x(t)$ results:

$$x(t) = x_0 + (y_0 - y_{0,ref})t + \frac{D - D_{ref}}{2}t^2 + \frac{\varphi(t) - \varphi_{ref}(t)}{2\pi\nu_{nom}}. \qquad (3)$$

Assuming that for the measurement of $x(t)$ the independent clock configuration applies and that the reference clock is properly chosen (i.e. all its degradation sources $y_{0,ref}$, D_{ref}, and $\varphi_{ref}(t)$ are negligible as compared to those of the clock under test), the $x(t)$ model reduces to:

$$x(t) = x_0 + y_0 t + \frac{D}{2}t^2 + \frac{\varphi(t)}{2\pi\nu_{nom}}. \qquad (4)$$

When the synchronized clock configuration applies and all slave clocks involved in the distribution of timing (including the clock under test) are operating in locked mode, $y_{0,\mathrm{ref}} = y_0$ and $D_{\mathrm{ref}} = D$ can be assumed; the $x(t)$ model then reduces to:

$$x(t) = x_0 + \frac{\varphi(t) - \varphi_{\mathrm{ref}}(t)}{2\pi\nu_{\mathrm{nom}}}. \quad (5)$$

We will use the model (4) to create correspondent Kalman filter.

3 DISCRETE-TIME KALMAN FILTER MODEL

The oscillator or clock discrete-time equations are given by observation and state vectors [9]

$$\boldsymbol{\xi}_\nu = \mathbf{H}_\nu \boldsymbol{\lambda}_\nu + \mathbf{u}_\nu + \mathbf{n}_{0\nu}, \quad (6)$$

$$\boldsymbol{\lambda}_\nu = \mathbf{A}_{\nu-1}\boldsymbol{\lambda}_{\nu-1} + \mathbf{n}_{\lambda\nu}, \quad (7)$$

where $\nu = 0, 1, 2, \ldots$ corresponds to discrete-time t_ν and measuring time interval $\Delta = t_\nu - t_{\nu-1}$, $\boldsymbol{\xi}_\nu = \boldsymbol{\xi}(t_\nu)$ is m-dimensional observation vector formed by the reference short-term noisy GPS timing signals and the oscillator, $\boldsymbol{\lambda}_\nu = \boldsymbol{\lambda}(t_\nu)$ is n-dimensional oscillator state vector (time, phase, frequency, aging, ...), $\mathbf{H}_\nu = \mathbf{H}(t_\nu)$ is $m \times n$ dimensional measurement matrix, $\mathbf{u}_\nu = \mathbf{u}(t_\nu)$ is m-dimensional vector contains the control signals, $\mathbf{A}_\nu = \mathbf{A}(t_\nu)$ is $n \times n$ dimensional state transition matrix, $\mathbf{n}_{0\nu} = \mathbf{n}_0(t_\nu)$ and $\mathbf{n}_{\lambda\nu} = \mathbf{n}_\lambda(t_\nu)$ are jointly independent vector white noises with zero expectations and covariance matrixes $\mathbf{V}_\nu = \mathbf{V}(t_\nu)$ and $\Psi_\nu = \Psi(t_\nu)$ are of $m \times m$ and $n \times n$ dimensions, respectively

$$\mathbf{V}_\nu = E\{\mathbf{n}_{0\nu}\mathbf{n}_{0\nu}^T\}, \quad (8)$$

$$\Psi_\nu = E\{\mathbf{n}_{\lambda\nu}\mathbf{n}_{\lambda\nu}^T\}. \quad (9)$$

As usually, they deal with single observation and estimate some states of an oscillator or clock. It means that if $m < n$ then one may use the following algorithm of linear Kalman filtering based on (1) and (2)

$$\hat{\boldsymbol{\lambda}}_\nu = \mathbf{A}_{\nu-1}\hat{\boldsymbol{\lambda}}_{\nu-1} + \mathbf{K}_\nu(\boldsymbol{\xi}_\nu - \mathbf{u}_\nu - \mathbf{H}_\nu\mathbf{A}_{\nu-1}\hat{\boldsymbol{\lambda}}_{\nu-1}), \quad (10)$$

$$\mathbf{R}_\nu = (\mathbf{I} - \mathbf{K}_\nu\mathbf{H}_\nu)\tilde{\mathbf{R}}_\nu, \quad (11)$$

$$\tilde{\mathbf{R}}_\nu = \mathbf{A}_{\nu-1}\mathbf{R}_{\nu-1}\mathbf{A}_{\nu-1}^T + \Psi_\nu, \quad (12)$$

$$\mathbf{K}_\nu = \tilde{\mathbf{R}}_\nu\mathbf{H}_\nu^T(\mathbf{H}_\nu\tilde{\mathbf{R}}_\nu\mathbf{H}_\nu^T + \mathbf{V}_\nu)^{-1}, \quad (13)$$

where $\hat{\boldsymbol{\lambda}}_\nu = \boldsymbol{\lambda}(t_\nu)$ is a vector of oscillator state estimates, $\mathbf{I}$ is unit matrix, and $\mathbf{R}_\nu = \mathbf{R}(t_\nu)$ is the error covariance matrix. Solution (10) is justified for a common case and may not be simplified, as a rule.

3.1. Regression Approach for Parameters Estimation

Major difficulties with the optimal Kalman algorithm design are caused by the problem of accurate definition of the necessary parameters in each step point in accordance with (8) and (9). The filtering error rises for small signal-to-noise ratio (SNR), and the possible way here is based on the regression estimates, knowing, that, strictly speaking, it is also the quasi optimal ones

$$y_i = \bar{\xi} + \frac{\sigma_\xi}{\sigma_t}\rho_{\xi t}(t_i - \bar{t}), \quad i = \overline{0,m} \quad (14)$$

where $\bar{\xi} = E\{\xi_i\}$ and $\bar{t} = E\{t_i\}$ are means of ξ_i and t_i, respectively, $\sigma_\xi^2 = E\{(\xi_i - \bar{\xi})^2\}$ and $\sigma_t^2 = E\{(t_i - \bar{t})^2\}$ are correspondent variances, and $\rho_{\xi t} = \dfrac{E\{(\xi_i - \bar{\xi})(t_i - \bar{t})\}}{\sigma_\xi\sigma_t}$ is correlation coefficient.

Procedure (14) takes first m points of the process to define the filter parameters by the following way. We put dawn $\lambda(0) \equiv \lambda_{\nu-1} = y_m$, then find out $d\xi_\nu = \xi_\nu - y_\nu$ and obtain

$$D_{0\nu} = E\{d\xi_\nu^2\}. \quad (15)$$

We take further the variance of the regression function (14) as $\sigma_y^2 = M\{(y_i - \bar{y})^2\}$ and get

$$D_{\lambda\nu} = D_{\lambda os} + \sigma_y^2, \quad (16)$$

where $D_{\lambda os}$ is expected variance of an oscillator frequency. Let us note that function $D_{\lambda\nu}$ depends on the regression function slope that results in the filter dynamic properties (inertia and inaccuracy). Finally, we take the initial value $R_{\nu-1} = D_{\lambda m}$.

3.2. Three Dimensional Kalman Algorithm

Based on the oscillator timing model (4), one may expect the best filtering effect as both the model and the filter dimension are matched. To show it, let us decompose the time error function x_ν into the series, and write

$$x_\nu = x_{\nu-1} + y_{\nu-1}\Delta + 0.5\alpha_{\nu-1}\Delta^2 + n_{x\nu}, \quad (17)$$

where $y_\nu = y_{\nu-1} + \alpha_{\nu-1}\Delta + n_{y\nu}$ is a frequency, and $\alpha_\nu = \alpha_{\nu-1} + n_{\alpha\nu}$ is a linear component of a frequency aging. This yields for (7)

$$\lambda_\nu = \begin{bmatrix} x_\nu \\ y_\nu \\ \alpha_\nu \end{bmatrix}, \; A_\nu = \begin{bmatrix} 1 & \Delta & \Delta^2/2 \\ 0 & 1 & \Delta \\ 0 & 0 & 1 \end{bmatrix}, \; n_{\lambda\nu} = \begin{bmatrix} n_{x\nu} \\ n_{y\nu} \\ \alpha\nu \end{bmatrix}, \quad (18)$$

and, respectively, for the observation equation (6)

$$\xi_\nu = [\xi_\nu], \; H_\nu = [1 \;\; 0 \;\; 0], \; n_{0\nu} = [n_{0\nu}]. \quad (19)$$

The Kalman filter equations are defined by (10)—(13), where for independent and non correlated noises we get

$$V_\nu = E\{n_{0\nu}n_{0\nu}^T\} = D_{0\nu} = S_{0\nu}\frac{1}{\Delta} = \frac{N_{0\nu}}{2\Delta} = \sigma_{o\nu}^2, \quad (20)$$

$$\Psi_v = E\{n_{\lambda v} n_{\lambda v}^T\} = \begin{bmatrix} \Psi_{11v} & \Psi_{12v} & \Psi_{13v} \\ \Psi_{21v} & \Psi_{22v} & \Psi_{23v} \\ \Psi_{31v} & \Psi_{32v} & \Psi_{33v} \end{bmatrix}, \quad (21)$$

where $\Psi_{11v} = S_{xv}\Delta + S_{yv}\Delta^3/3 + S_{\alpha v}\Delta^5/20$;

$\Psi_{21v} = \Psi_{12v} = S_{yv}\Delta^2/2 + S_{\alpha v}\Delta^4/8$;

$\Psi_{22v} = S_{yv}\Delta + S_{\alpha v}\Delta^3/3$; $\Psi_{33v} = S_{\alpha v}\Delta$;

$\Psi_{31v} = \Psi_{13v} = S_{\alpha v}\Delta^3/6$; $\Psi_{32v} = \Psi_{23v} = S_{\alpha v}\Delta^2/2$;

$S_{xv} = N_{xv}/2$, $S_{yv} = N_{yv}/2$ and $S_{\alpha v} = N_{\alpha v}/2$

are correspondent two-side spectral densities of the continuous noises depending on the time interval Δ in a general case. If to account an aging noise only then the matrix (21) yields

$$\Psi_v = \frac{N_{\alpha v}}{2}\Delta \begin{bmatrix} \Delta^4/20 & \Delta^3/8 & \Delta^2/6 \\ \Delta^3/8 & \Delta^2/3 & \Delta/2 \\ \Delta^2/6 & \Delta/2 & 1 \end{bmatrix} \cong \frac{N_{\alpha v}}{2}\Delta \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (22)$$

The prediction equations are obtained by the above-considered algorithm being based on the estimates of $\hat{x}_v$, $\hat{y}_v$, and $\hat{\alpha}_v$ as follows

$$\begin{aligned} \tilde{x}_{v+1} &= \hat{x}_v + \hat{y}_v\Delta + 0.5\hat{\alpha}_v\Delta^2, \\ \tilde{y}_{v+1} &= \hat{y}_v + \hat{\alpha}_v\Delta, \\ \tilde{\alpha}_{v+1} &= \hat{\alpha}_v. \end{aligned} \quad (23)$$

The equations (23), in bottom, are the working ones for the holdover operation mode of the slaved oscillator when GPS signals are not available [11].

4. EXPERIMENTAL RESULTS

We bring the results of the three-dimensional Kalman algorithm (10)—(13) use for the synchronization errors estimation. We consider the OCXO phase (time error) measured for $\Delta = 100$ sec based on the GPS timing signals of the Motorola ONCORE UT+ Receiver. Here digital sequence x_v (the oscillator phase reduced to the time interval error) is the initial time error data that is processed by the Kalman filter (10)—(13). As a result, we get the estimates $\hat{x}_v$, $\hat{y}_v$, and $\hat{\alpha}_v$. Frequency estimate $\hat{y}_v$ is compared also with its accurate magnitude obtained by direct measurement with the use of the reference time signal of 1sec of the rubidium standard in the same time scale.

Figure 1 shows the initial observation (time error) x_v jointly with its Kalman estimate $\hat{x}_v$ and expected error calculated as the difference $\varepsilon_v = x_v - \hat{x}_v$. Flowing from the physical sense, spectral density of the fluctuation of the aging speed within the noise matrix (22) is taken as $S_{\alpha 0} = 0.5N_{\alpha 0}\Delta = 4\cdot 10^{-22}$ Hz^2 for the process. With this, the expected root-mean-squire time deviation (TDEV) and maximal time interval error (MTIE) are estimated as $\sigma_\varepsilon = 282.7$ ns and $\Delta x_{\max} = 39680$ ns. Thus, in principle, we get possibility to reduce the time error by digital PLL to $\Delta x_{\max}/\sigma_\varepsilon \cong 140$ times.

Figure 2 exhibits the measured function of the relative frequency behavior of the same OCXO jointly with the Kalman estimate. If to take the measured data as the reference ones then the difference function may be taken as the error function, which TDEV equals $\sigma_{\varepsilon y} = 146\cdot 10^{-12}$ for the considered case. Correspondingly, TDEVs of the measured and estimated curves are $\sigma_{ym} = 325\cdot 10^{-12}$ and $\sigma_y = 286\cdot 10^{-12}$. Hence, the expected frequency control effect yields $a_y = \sigma_y/\sigma_{\varepsilon y} = 1.96$ and $a_{ym} = \sigma_{ym}/\sigma_{\varepsilon y} = 2.4$ with respect to the measured and estimated data correspondingly.

Finally, Figure 7 shows estimates obtained for the frequency aging. Mean value of α_v equals $\bar{\alpha}_v = E\{\alpha_v\} = -2.7\cdot 10^{-15}$/s upon the observation interval. TDEV for α_v is equal to $\sigma_\alpha = 3.4\cdot 10^{-14}$/c, and the ratio of the two statistical estimates is $\sigma_\alpha/\bar{\alpha}_v = 12.4$ that is caused by the extremely small SNR<<1.

5. CONCLUSION

We have considered the three-dimensional matched Kalman algorithm intended to estimate the time error, frequency, and frequency aging of the slave crystal oscillator based on the Motorola GPS Oncore UT+ Timing Receiver. General findings of the studies are following.

The best results of the parameter estimation are obtained if the filter is matched with the timing model of the oscillator. Thus, selection of the proper Kalman algorithms has to be done at the early stage of the filtering to get the best effect. We may note here to conclusion that three-dimensional Kalman algorithm is most efficient for rubidium oscillator and brings less good results for the crystal one as the aging function of the OCXO has the complex nature and is less predictable.

Three-dimensional matched Kalman filter exhibits small inertia (negligible dynamic error) and high accuracy jointly with prediction the time error for the holdover operation mode. Nevertheless, for our opinion, with respect to the crystal oscillators, the studies must be done to obtain the robust extended Kalman algorithm to reduce the errors caused by the frequency jumps and insufficient prediction of the aging. We plan to present the results of these studies in future.

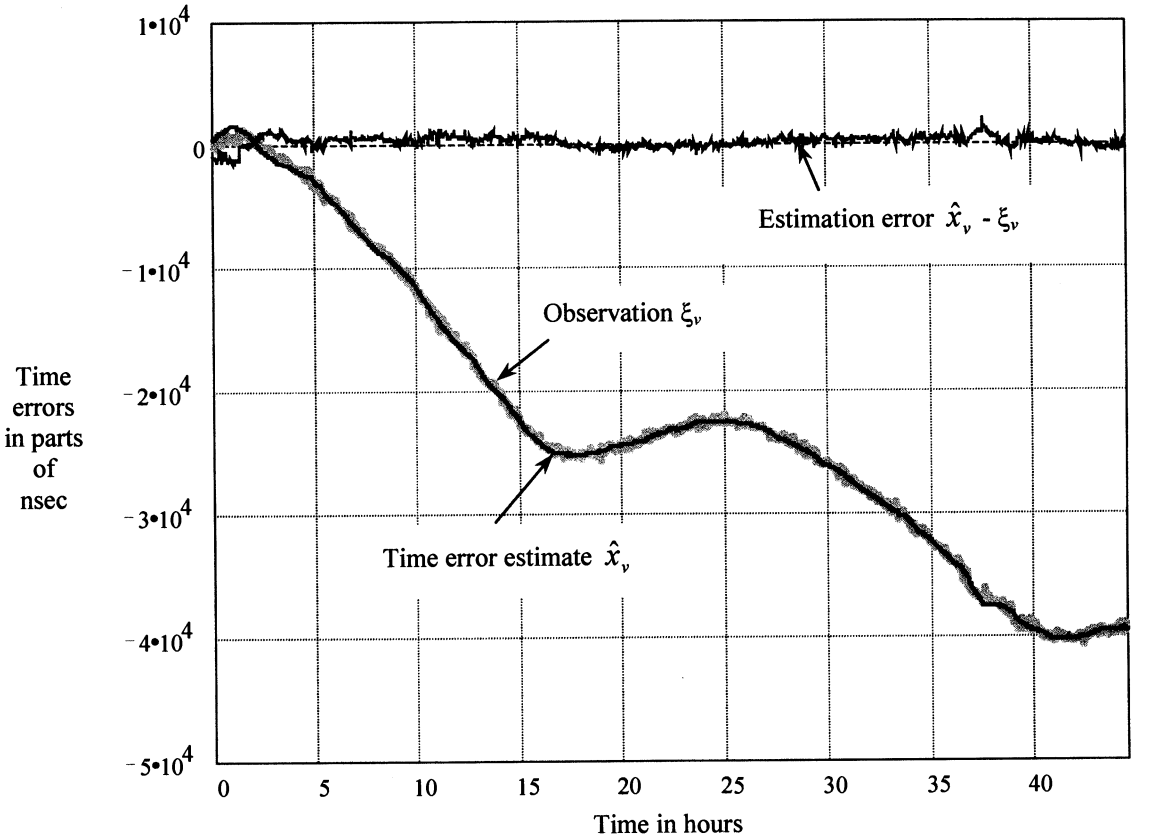


Figure 1. GPS-based three-state Kalman estimates of time errors

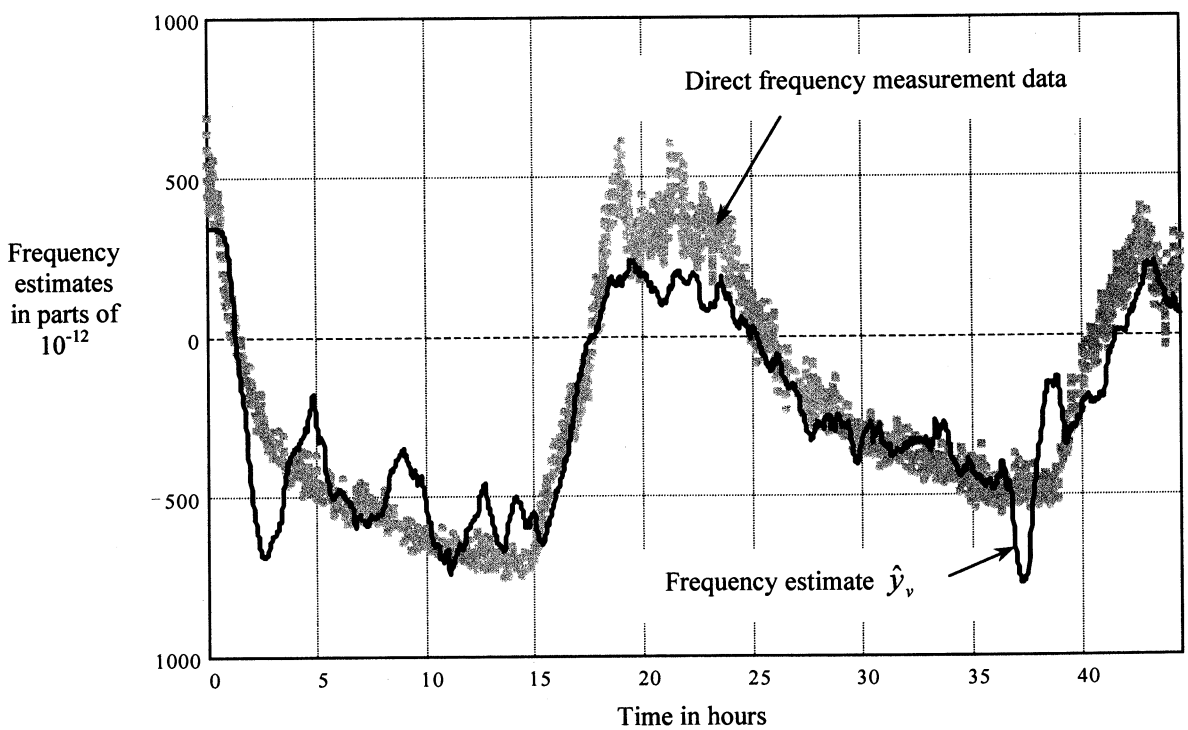


Figure 2. GPS-based three-state Kalman estimates of the frequency offset $\hat{y}_v$ and directly measured data

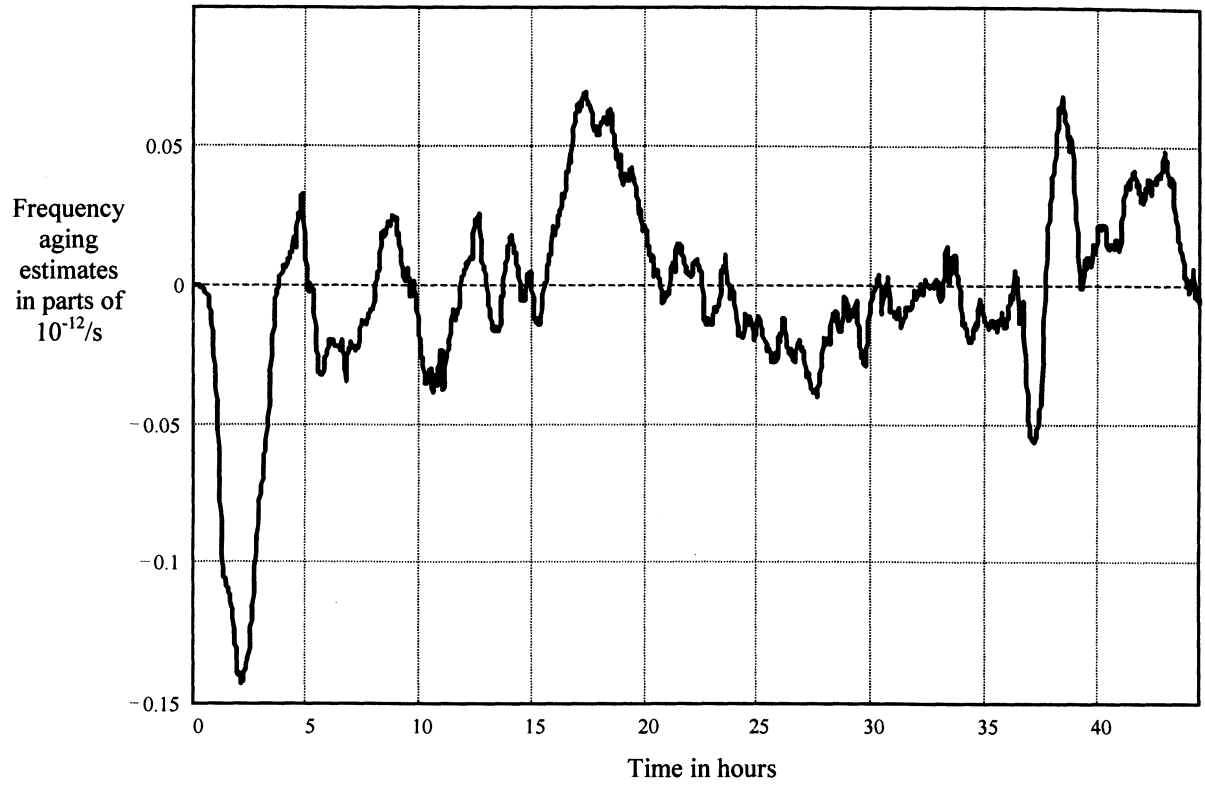


Figure 3. GPS-based three-state Kalman estimates of the frequency aging rate $\hat{\alpha}_v$

6. REFERENCES

[1] D.W. Allan and J.A. Barnes, "Optimal Time and Frequency Transfer Using GPS Signals", in *Proc. of the IEEE 36th AFCS*, pp.378-387, 1982.

[2] R.L. Filler and S.R. Stein, "Kalman Filter Analysis for Real Time Applications of Clocks and Oscillators", in *Proc. of 42nd AFCS*, pp.447-452, 1988.

[3] S.R. Stein, "Kalman Filter Analysis of Precision Clocks with Real-Time Parameter Estimation", in *Proc. of 43rd AFCS*, pp.232-236, 1989.

[4] W. Su, and R.L. Filler, "Application of Kalman Filtering Techniques to the Precision Clock with Non-Constant Aging", in *Proc. of 46th IEEE IFCS*, pp.231-237, 1992.

[5] W. Su and R.L. Filler, "A New Approach to Clock Modeling and Kalman Filter Time and Frequency Prediction", in *Proc. of 47th IEEE IFCS*, pp.331-334, 1993.

[6] R.E. Kalman, "A New Approach to Linear Filtering and Prediction Problems", *Trans. of the ASME, Ser.D, Journal of Basic Engr.*, Vol.82, pp.35-45, 1960.

[7] R.K. Mehra, "On the Identification of Variances and Adaptive Kalman Filtering", *IEEE Trans. on Automatic Control*, Vol.15, No.2, pp.175-184, 1970.

[8] B. Ekstrand, "Analytical Steady State Solution for a Kalman Tracking Filter", *IEEE Trans. on Aerospace and Electronic Systems*, Vol.19, No.6, pp.815-819, 1983.

[9] O.E.Rudnev, Yu.S.Shmaliy, *et al.* "Kalman filtering of a frequency instability based on Motorola Oncore UT GPS Timing signals," *Proc. of 13th EFTF and 1999 IEEE IFCS*, 1999.

[10] G.810. Definitions and terminology for synchronization networks, ITU-T, Geneva, Aug 1996.

[11] W. Su and R.L. Filler, "A New Approach to Clock Modeling and Kalman Filter Time and Frequency Prediction", in *Proc. of 47th IEEE IFCS*, pp.331-334, 1993.

[12] Yu. S. Shmaliy, A.V. Marienko, and A.V. Savchuk. "GPS-Based Optimal Kalman Estimation os Time Error, Frequency Offset, and Aging," *in Proc. of 31st PTTI*, 1999.

WIDEBAND SURFACE MOUNT HYBRID SAW MODULES WITH RFA FOR 146-174 MHZ MOBILE TRANSCEIVERS

S. A. Doberstein, A. V. Martynov, V. K. Razgonyaev
ONIIP, Maslennikov str., 231, Omsk, 644009, Russia, fax: 7-381-2-399391

ABSTRACT

This paper presents wideband surface mount hybrid modules with SAW filters for 146-174 mobile transceivers: the receiver modules with RF amplifier (RFA) having a fractional bandwidth of 3-7 %. These modules were mounted in the 16x7.3x2.6 mm SMD packages. The low-loss wideband three-transducers SAW filters using unidirectional IDTs on U-shaped MSCs on 128°YX, 64°YX, 41°YX $LiNbO_3$ were used. The modules were connected between a low impedance antenna and high impedance mixer in the receiver section of the mobile transceivers and used for the local oscillator frequency and image frequency suppressions. The modules contained two weighted SAW filters: the first filter with low input/output impedances, the second filter with high input/output impedances, connected across a low noise RFA on the one bipolar transistor. The 155, 165, 169 MHz samples of the surface mount hybrid SAW modules with RFA have shown an amplitude ripple of 1 dB within a 3-dB bandwidth of 4.65-11 MHz, 16 dB gain, suppression about 80 dB at ± 12 - ± 25 MHz offsets from the center frequency. A low noise and high intermodulation selectivity were provided with low power consumption. The new hybrid SAW modules have a wide bandwidth, small size and are compatible with SMT.

1. INTRODUCTION

Recently the surface mount hybrid SAW modules (functional devices combining the low-loss SAW filters, amplifiers, phase shifters in the SMD packages) were developed for simplifying and miniaturizing the 146-174 MHz mobile transceivers [1]. However these hybrid SAW modules have a narrow fractional bandwidth (below 2 %). The SAW modules with the wider passband are required for many applications of the 146-174 MHz mobile transceivers. This paper presents wideband surface mount hybrid modules with SAW filters: the receiver modules with RFA having a fractional bandwidth ($\Delta f/f_0$) of 3-7 %. The modules contained two wideband low-loss SAW filters connected across low noise RFA and were connected between an antenna and mixer in the receiver section (Fig. 1). These modules are used for the local oscillator frequency and image frequency suppressions.

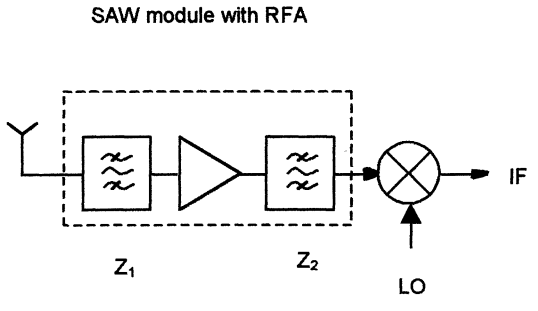


Fig. 1 Schematic layout of the SAW module with RFA

2. RFA FOR HYBRID MODULE

A feature of the RFA is that it is connected between a low impedance antenna and high impedance mixer. Currently the high impedance mixer (an input impedance of 100-200 Ω) is used for simplifying IF section of the 146-174 mobile transceivers. Also a gain of the front-end stage must be no more than 20 dB [1]. In the development of the RFA it is necessary to execute these requirements and also to provide a minimum mismatch with the SAW filters, to obtain a low noise. Besides it is necessary to use a comparatively simple circuit with the minimum parasitic couplings and with a possibility of the control of the RFA characteristics by the feedback elements and the regimes of the active elements. A previously developed wideband amplifier [1] on a common-emitter bipolar transistor 2T3132 with a

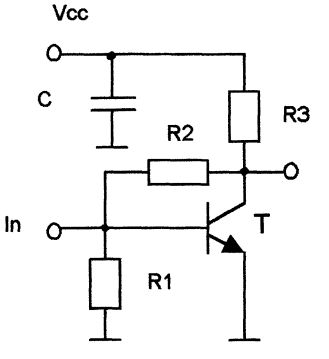


Fig. 2 Circuit diagram of RFA

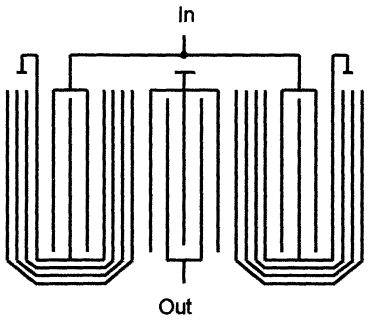


Fig. 3 Schematic layout of three-transducer SAW filter with U-shaped MSCs

cut off frequency of 7.5 GHz, collector capacitance of 0.7 pF, low noise factor, high gain under a low collector current (0.5-3 mA) was used for the hybrid module (Fig. 2). In this case a gain of 15-20 dB with low phase slope, noise factor of 1 dB, low current consumption of 1-3 mA and supply voltage of 5 V were achieved in the frequency range up to 800 MHz by choosing the collector current and the feedback elements of a transistor T. Under these conditions the amplifier input/output impedances were close to real values of 40-70 and 100-180 Ω, respectively.

3. SAW FILTERS FOR MODULE WITH RFA

For the module we used previously developed wideband three-transducer SAW filters with a insertion loss of 1 dB and fractional bandwidth of 3-8 % on 128°YX, 64°YX, 41°YX $LiNbO_3$ (Fig. 3) [2]. The low insertion loss is achieved by using unidirectional input interdigital transducers (IDTs) with U-shaped multistrip couplers (MSCs). Weighting of the central IDT provides sidelobe suppression in the frequency response. A construction does not require matching networks because it provides the specified real input/output impedances in the passband by self-matching when a static capacitance of the IDT is compensated by radiation susceptance. Moreover this construction allows to change the input/output impedances in sufficiently large limits by of a variable metallization ratio in the center IDT and by choosing the apertures of the IDTs (Fig. 3). Fig. 4 shows the frequency response of 165 MHz SAW filter on 41°YX $LiNbO_3$. All IDTs have an aperture of 76λ (λ is a SAW wavelength at a center frequency). The central IDT was with Hamming phase weighting and with an increased metallization ratio (k_m>0.5). This increased metallization ratio is used for an equalization of the input and output impedances in the filter [2]. At a center frequency of 165 MHz the input/output impedances of the filter are close to real magnitudes of 75 Ω. In a 75-Ω system the filter has

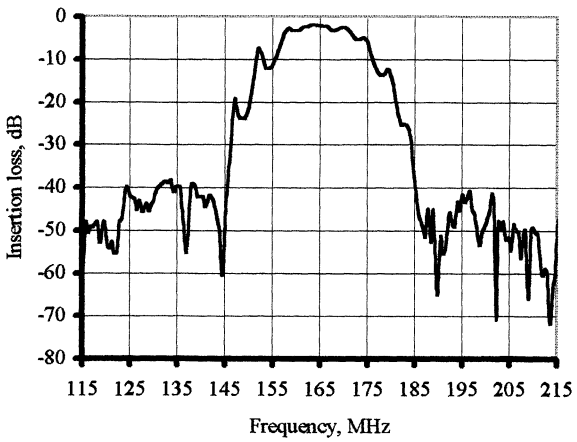


Fig. 4 Measured frequency response of the 165 MHz filter with 7 % fractional bandwidth on 41°YX

shown an insertion loss of 1.8 dB, 3-dB bandwidth of 11.2 MHz with a ripple of 0.3 dB, stopband attenuation of 35-40 dB at ± 25 MHz offset from the center frequency. Chip size was 3.8x3.8x0.7 mm. This filter is connected to the input SAW module with RFA. The input impedance of the filter increase up to a magnitude of 100 Ω, and the output impedance – up to a magnitude of 180 Ω with decreasing apertures of the IDTs to a magnitude of 57λ and k_m≤0.5 in the central IDT. The frequency response of such filter practically does not change and repeates a form showed in Fig. 4. Evidently it is appropriate to connect this high impedance filter (after prior replacing input by output) to the output of the RFA (Fig. 1) for matching with a high impedance mixer. Similarly a method can be shown for realization of the filters for the module with a fractional bandwidth of 3 and 4 %. For this purpose the three-transducer filters with U-shaped MSCs on 128°YX, 64°YX [2] can be used.

4. DEVELOPMENT OF WIDEBAND SURFACE MOUNT HYBRID SAW MODULES WITH RFA

The hybrid module contained two SAW filters: the first filter with low input/output impedances (Z_1), the second filter with high input/output impedances (Z_2), were connected across RFA (Fig. 1). An optimization of a SAW filter-amplifier-SAW filter system was provided using a computer simulation on a basis of an equivalent circuit model. The equivalent circuit of a filter- amplifier-filter system is shown in Fig. 5. Here [P_1], [P_3] are the mixed matrixes of the central IDTs for the first filter Z_1 and second filter Z_2 respectively; [P_2], [P_4] are the mixed matrixes of the IDTs with U-shaped MSCs

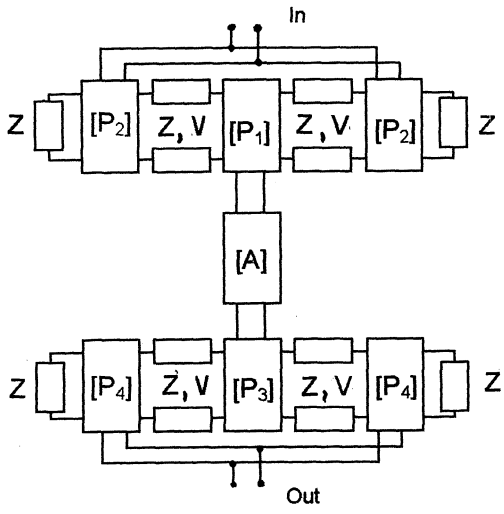


Fig. 5 Equivalent circuit of the wideband SAW module with RFA

for the first filter Z_1 and second filter Z_2 respectively; [A] – an A matrix of the amplifier; Z – characteristic impedance of a medium between IDTs, V- SAW velocity.

Much attention was given to achieving low ripple in the passband of the module. For this purpose the minimization of SAW filter and RFA mismatch was carried out. In this case the input/output impedances of RFA were fitted to the input/output impedances of the filters or conversely. The computer simulation demonstrated that this minimization is conveniently carried out with the first method by a careful selection of the feedback elements and regime of a transistor in RFA circuit. Fig. 6 shows the simulated normalized frequency response of 155 MHz filter-amplifier-filter system (prototype module) with in a 3 % fractional bandwidth with SAW filters on 128°YX. A plot 1 is the frequency response of the prototype module when the SAW filters are mismatched with RFA input/output. A plot 2 is the frequency response of the prototype module after a optimization of a RFA circuit when the SAW filters and RFA are minimal mismatched in the input/output impedances. As will be seen from Fig. 6 (plot 2) the passband ripple of the prototype module is not more than 0.2 dB. Fig. 7 shows the simulated normalized frequency response of the prototype module in the wide frequency band. The 100 dB stopband attenuation is obtained. Below the design of the hybrid SAW modules is shown. The modules were mounted in a 16x7.3x2.6 mm SMD packages with a12.2x4.8 mm mounting pad. The maximum depth of a paskage was 1.5 mm. So it is necessary to use a printed board with a thickness of 0.3 mm for RFA and special modernizated version of a transistor 2T3132 – A 871 with a package thickness and diameter of 0.7 and 2.15 mm respectively. This transistor is sticked on a 3.6x4.8 mm printed board with the evaporated resistors and SMD capacitor and is bonded by the gold wires. The module topology optimization was provided to achieve high stopband attenuation closed to a simulated magnitude of 100 dB (Fig. 7) by selection of a arrangement of SAW filters, RFA and ground bond wires in a SMD package. A combination of the wideband weighted SAW filters with low and high impedances has provided good matching with an antenna, input/output RFA, input of the mixer in the wide frequency range and allowed to obtain a high selectivity and to house optimally these components in the SMD package (Fig. 8). A photogragh of the SAW hybrid module in the 16x7.3x2.5 mm SMD package presented in Fig. 9. Fig. 10 shows a normalized frequency response of 165 MHz hybrid module the SAW filters on 41°YX. This module has shown an amplitude ripple of 0.8

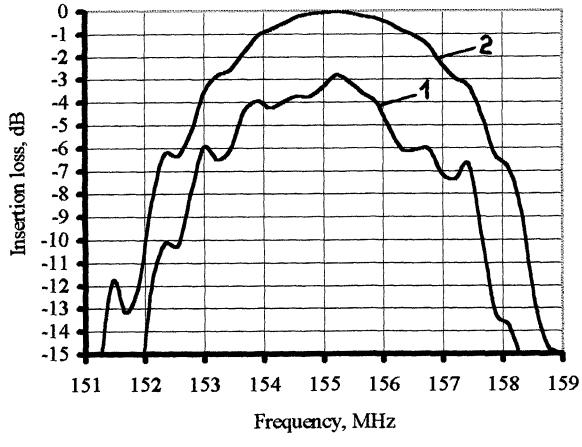


Fig. 6 Simulated frequency response of the SAW module within passband

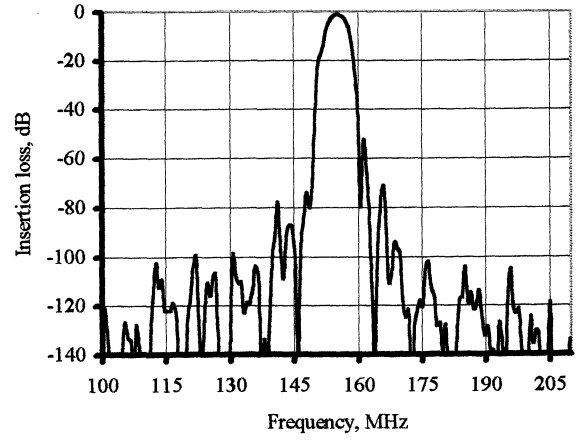


Fig. 7 Simulated frequency response of the SAW module in the wide frequency band

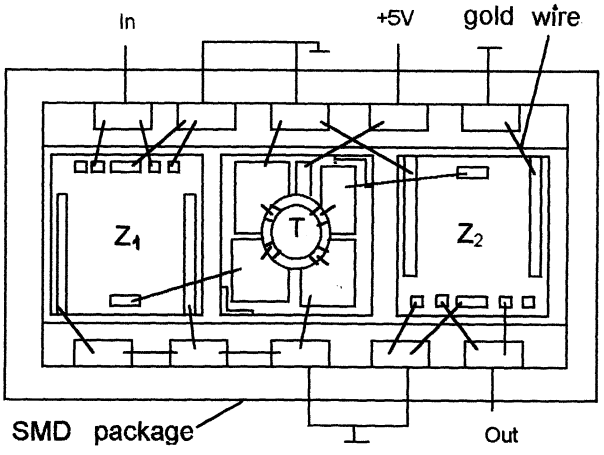


Fig. 8 Construction of the hybrid SAW module with RFA

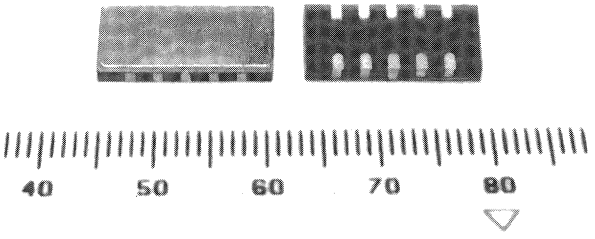


Fig. 9 Photograph of the surface mount hybrid SAW module with RFA

dB with in a 3-dB bandwiidth of 11 MHz ($\Delta f/f_0$=7 %), 16 dB gain, suppession of 80 dB at ± 25 MHz offset from the center frequency in a measurement system with an input load of 75 Ω and output load of 100 Ω. A noise factor of 2 dB and intermodulation selectivity of 70 dBμV were provided with a low current consumption of 2 mA and supply voltage of 5 V. An inadequate suppression (about 70 dB) in a low part of the response (131 MHz) and high part of the response (240 MHz) caused by Rayleigh SAW and bulk wave emitted by IDTs on 41°YX $LiNbO_3$ respectively. 64°YX $LiNbO_3$ with zero value of electromechanical-coupling coefficient for Rayleigh SAW allowes to eliminate one of these defectes. The normalized frequency response of 169 MHz hybrid module with the SAW filters on 64°YX is presented in Fig. 11. The module has provided an

amplitude ripple of 0.6 dB within a 3-dB bandwidth of 6.7 MHz ($\Delta f/f_0$=4 %), 16 dB gain, suppression of 80 dB at ± 21 MHz offset from the center frequency in a measurement system with an input load of 75 Ω and output load of 100 Ω. As will be seen from Fig. 11 the 80 dB suppression is provided in the entire low pait of the frequency response of the module. However the module has an inadequate suppression (about 70 dB) in high part of the response (243 MHz) as well as for module on 41°YX. This effect caused by a bulk wave emission. In this connection use of 128°YX $LiNbO_3$ minimized a bulk wave emission is very efficient. Fig. 12 shows the normalized frequency response of 155 MHz hybrid module with SAW filters on 128°YX. In a measurement system with an input load of 50 Ω and output load of 100 Ω the module has shown an amplitude ripple of 0.3 dB within a 3-

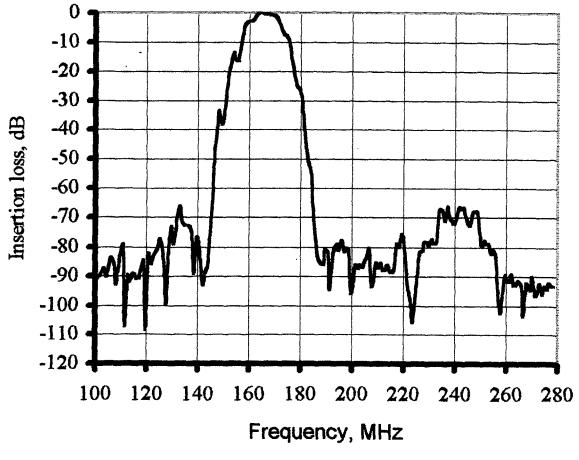


Fig. 10 Measured normalized frequency response of the 165 MHz SAW module with 7 % fractional bandwidth

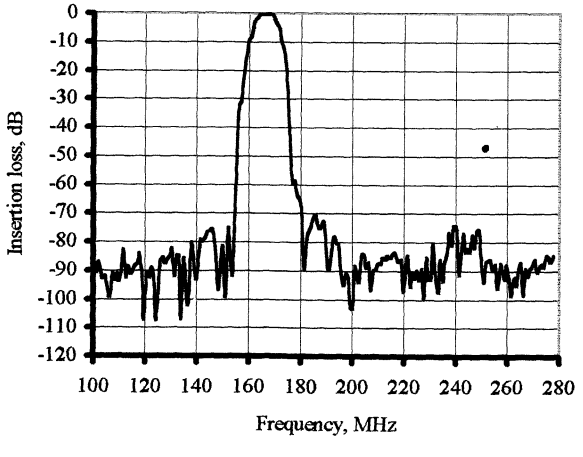


Fig. 11 Measured normalized frequency response of the 169 MHz SAW module with 4 % fractional bandwidth

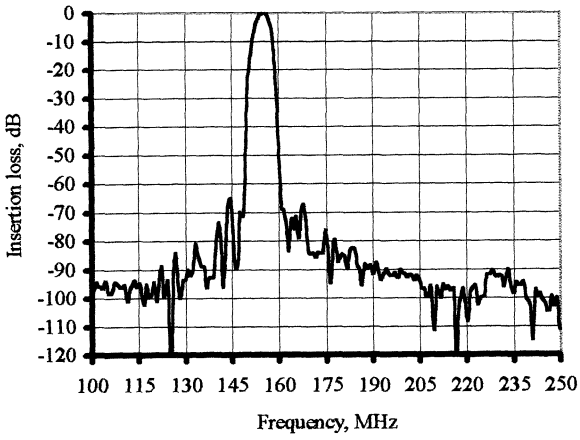


Fig. 12 Measured normalized frequency response of the 155 MHz SAW module with 3 % fractional bandwidth

dB bandwidth of 4.65 MHz ($\Delta f/f_0$=3 %), 17 dB gain, suppression of 80 dB at ± 12 MHz offset from the center frequency. As will be seen from Fig. 12 the parasitic peaks in the low and high parts of the stopband are absent.

5. CONCLUSION

We developed new surface mount hybrid SAW modules for 146-174 MHz mobile transceivers with wider bandwidth than the previous ones [1]. To extend the bandwidth the three-transducer SAW filters using unidirectional IDTs on U-shaped MSCs on 128°YX, 64°YX, 41°YX $LiNbO_3$ are used. These modules provided 16 dB gain within a fractional bandwidth of 3-7 %, stopband attenuation about 80 dB. We shall use them in the wideband front-end stages of the 146-174 MHz mobile transceivers for the local oscillator frequency and image frequency suppressions.

6. REFERENCES

[1] S. A. Doberstein, A. V. Martynov and V. A. Malyukhov, "Surface Mount Hybrid SAW Modules for Mobile Transceivers", in Proceedings of the 12th EFTF, 1998, pp. 472-476.

[2] S. A. Doberstein, V. A. Malyukhov and V. K. Razgonyaev, "Wideband Three-Transducer SAW Filters Using Unidirectional IDTs on U-spaped MSCs with Insertion Loss of 1 dB", presented at the 1999 IEEE Ultrasonics Symposium, Tahoe, USA, October 17-20, 1999.

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

GENERALIZATION OF THE TOTAL VARIANCE APPROACH TO THE DIFFERENT CLASSES OF STRUCTURE FUNCTIONS

F. Vernotte†
Observatoire de Besançon
41 bis, avenue de l'Observatoire
25010 Besançon Cedex, FRANCE

D. A. Howe‡
National Inst. of Standards and Tech.
325 Broadway
Boulder, Colorado, USA 80303

Abstract—**The Total variance approach has been developed for increasing the confidence of the estimation of the classical Allan variance (AVAR), particularly for large integration time values. This method is based on a procedure of extension of the original data sequence called the mirror-reflection which increases the equivalent degrees of freedom of each Allan variance estimate. Recently, we applied this approach to the Modified Allan variance (MVAR) and proved that, in this case, another procedure of extension of the data sequence should be used: the reflection-only extension.**

In this paper, we propose a criterion to select the most appropriate extension procedure for a given structure function (i.e. variance). This criterion is based on the sensitivity of a structure function to the linear and quadratic drifts, or, and this is equivalent, on the convergence of this structure function for the different types of noise.

This method and this criterion will be illustrated by its application to the "pulsar variance", which is insensitive to the quadratic and linear phase drifts, and converges from f^{-4} FM to f^{+2} FM.

1. INTRODUCTION

The estimation of the long term stability of oscillators (particularly very long term) is more difficult than the estimation of the short term stability. As a consequence, the random walk and flicker FM noise levels, which are likely to dominate at long term, are always less precisely estimated than the other noise levels. Nevertheless, the low frequency noise levels contain the statistical information of the future behavior of an oscillator, which is essential for extrapolating its performances [1].

Obviously, this lack of knowledge is due to the length of the time sequence which is often insufficient for insuring the preeminence of the low frequency noises. For ultra stable quartz oscillators, a data sequence of at least one day length must be used to distinguish the flicker FM from the white FM, and 5 to 10 days may be necessary for detecting the random walk FM. Moreover, the uncertainty of the estimates increases with the integration time. For example, with the Allan variance (AVAR) and considering a T-length time sequence, the longest integration time is $\tau = T/2$. Its corresponding estimate is chi-square distributed with only one degree of freedom. This means that the standard deviation of such an estimate is $\sqrt{2}$ times greater than its expectation. Furthermore, the distribution is negatively skewed with values twice as likely to be below the actual noise level than above. AVAR is also sensitive to linear frequency drift which must be removed, thus additionally suppressing the actual low frequency random noise levels. To be safe, the $T/2$ estimate is ignored, and the longest integration time is limited in practice to $T/4$ or less. Therefore, a data run of length 20 to 40 days must be used to provide sufficient confidence for detecting the presence of, say, random walk FM, let alone estimating its level. When long enough data runs become impractical to obtain, Total variance (Totvar, or its usually reported square-root Totdev) is recommended as an improved estimator of long term stability [2].

The goal of the concepts in this paper is to increase the confidence of long term frequency stability estimates without increasing the length of a data run for other classes of variances. Additionally, it is often important to distinguish white PM from flicker PM, unlike the Allan variance. To accomplish this, we apply the Total variance approach [3] to structure functions [4]. This is motivated by two issues.

The first issue involves defining a suitable variance from a combination of order and family of structure functions such that the variance is sensitive to all expected types of noises while it is insensitive to particular types of drift. This avoids the complication of removing drifts before applying the variance and consequently avoids suppression and underestimation of actual random noise level at long term.

The second issue involves retrieving the maximum information from the data run itself by using the Total variance approach. This approach is based upon a periodic extension of data sequences [5]. Obviously, this doesn't permit an increase of the integration time beyond a variance's normal upper limit, but the uncertainty interval of the corresponding long term estimates is significantly reduced by using the Total approach.

From our experience in constructing an improved estimator of the modified Allan variance (called Mod-Totvar), several types of data extensions can be considered [6]. This paper gives criteria for the extension type selection according to the variance which is obtained from the order and family of structure functions. Section 2 illustrates how the Total variance model is implemented on the modified Allan variance to obtain the best "Modified Total" variance. Section 3 defines relevant structure functions. Section 4 gives criteria for selecting a data extension type for these structure functions and ends with an example of the Total approach applied to the pulsar variance.

†This author is with the Laboratoire d'Astrophysique de l'Observatoire de Besançon (LAOB), UPRES-A CNRS 6091; E-mail: francois@obs-besancon.fr

‡This author is with the Time and Frequency Division; E-mail: dhowe@nist.gov.

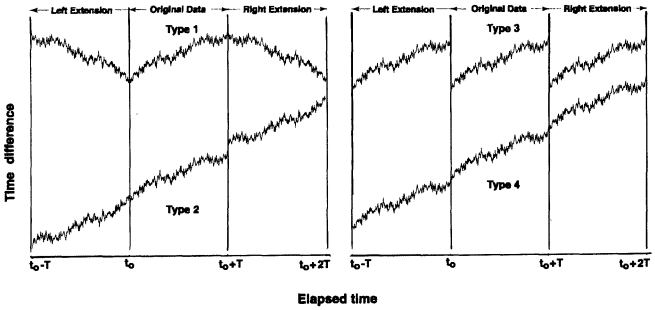


Fig. 1. The four types of data extensions.

2. THE TOTAL VARIANCE APPROACH

2.1 Bias and equivalent degrees of freedom

The Total variance approach involves periodically extending a data sequence beyond its normal measurement duration and in such a way that a particular time statistic is expected to have the same value with extended data as without. For those statistics which estimate components of broadband noise processes, the approach can significantly reduce the spread or uncertainty in the result.

We use two quantities to check the efficiency of the method:

- *the bias* defined as the percentage of error between the classical variance estimate and the Total variance estimate;
- *the equivalent degrees of freedom (edf)* defined, assuming a χ^2 distribution of the estimates, as

$$\text{edf} = \frac{2\,[\text{Mean(estimates)}]^2}{\text{Variance(estimates)}}. \qquad (1)$$

2.2 Types of data extension

Denoting T as the length of the calculation sequence (for example, $T = 2\tau$ for AVAR, $T = 3\tau$ for MVAR, ...), four types of extension have been used (see figure 1):

1. $2T$-periodic uninverted or even mirror-reflection,
2. $2T$-periodic sign-inverted or odd mirror-reflection,
3. T-periodic straight duplication,
4. T-periodic duplication with end-to-beginning connections.

Obviously, other extension types could be used, but these types summarize essential properties of interest.

Sequence extensions were originally tested with the Allan variance. The type 2 extension was found to be optimum and is used in defining Totvar [5],[7]. Since Totvar, like AVAR, doesn't distinguish white PM and flicker PM noises, the approach was generalized to the time variance (and time deviation) [8] and to the modified Allan variance [6]. These variances are specially designed for estimating the level of the phase modulation noise types (white PM, flicker PM, and random walk PM) as well as the frequency modulation noise types (flicker FM and random walk FM). In this case, the type 1 extension should be used in order to avoid a huge bias in the presence of high frequency (or PM) noises. This is because the type 2 extension modifies the mean of a sequence at each end-to-beginning connection, inducing a step in the subestimates and causing the overall mean of the Total approach to be biased very high (see figure 2 and ref. [6]).

$S_y(f)$	b	c	edf(τ_{max})	bias%
$h_{+2}f^{+2}$	1.9	2.1	3.6	-6%
$h_{+1}f^{+1}$	1.2	1.4	2.2	-17%
$h_0 f^0$	1.1	1.2	2.1	-27%
$h_{-1}f^{-1}$	0.85	0.50	2.05	-30%
$h_{-2}f^{-2}$	0.75	0.31	1.94	-31%

TABLE I
EDF MODEL DEFINED IN (2) AND BIAS FOR MOD-TOTVAR.

2.3 Modeling edf and bias

Howe and Greenhall defined an empirical model for edf of Totvar [5]:

$$\text{edf}\,[\text{Totvar}(\tau, T)] = b\frac{T}{\tau} - c. \qquad (2)$$

where the coefficients b and c were estimated for each type of noise from a Monte-Carlo method.

Using the same model (2) for Mod-Totvar, we obtained the coefficients b and c given in Table I, which also lists percentage bias of Mod-Totvar relative to classical MVAR for its range of noise types.

2.4 Practical implementation of Mod-Totvar

Let us consider a sequence of N time error data $\{x(t_i)\}$, with a sampling rate τ_0. Let us denote T the length of the total sequence: $T = N\tau_0$.

The integration time τ may be defined as $\tau = m\tau_0$ where m is an integer and $m \leq N$. The length of the calculation sequence of MVAR is 3τ, which will be called a *subsequence*, one of all possible consecutive sequences of 3τ length.

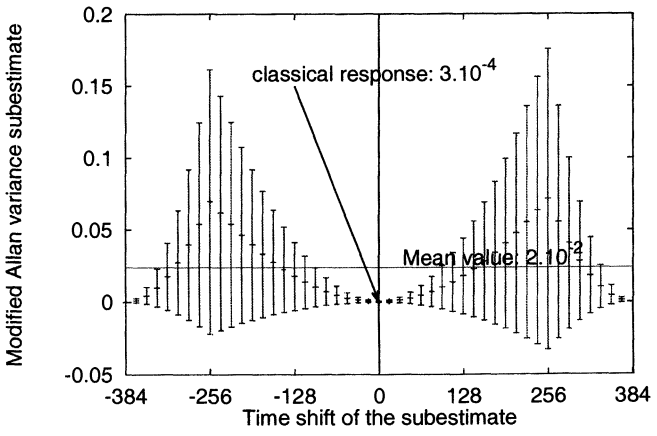


Fig. 2. Bias of the subestimates for a type 2 extension applied to a white PM noise and using MVAR. Shown are MVAR mean values and associated standard deviations (by the error bars) computed at $\pm$ time-shifts of $10\tau_0$, that is, shifts in the extended sequence of $0, \pm 10, \pm 20, \ldots \pm 384$. Each mean value is an average of 1000 estimates. The middle mean value is at a null-shift (0) which corresponds to the classical MVAR result.

In order to calculate the Mod-Totvar for one given value $\tau = m\tau_0$, one would:

1. extract all $3m$ data subsequences from the whole sequence $\{x(t_i)\}$;
2. remove their linear phase drift;
3. extend them at both ends by the type 1 even mirror-reflection to form $9m$ data subsequences;
4. calculate MVAR for each of these $9m$ subsequences;
5. average all these MVAR results.

2.4.1 2-τ and 3-τ extension

There are two cases for which MVAR can be computed. In case 1, MVAR subsequences can have a span of 2τ if three successive $\{x(t_i)\}$ values form a second-difference and consecutive second-differences spaced by τ_0 are subsequently averaged for integration time $\tau = m\tau_0$. Resulting values are then squared and averaged to compute MVAR. In case 2, MVAR subsequences can have a span of 3τ if three successive τ-averaged $\{x(t_i)\}$ values form a second-difference whose squared value is then averaged with all other possible squared second difference values. Since case 1 yields the same answer as case 2, the question arises, "Should we extend 2τ (case 1) or 3τ (case 2) subsequences?" Using simulation studies, we compared both cases, extending each subsequence to form a 2τ version and a 3τ version of Modified Total variance. The 2τ version had significantly more negative bias than the modest bias of the 3τ version shown in Table I. Moreover, the edf showed a reduction by 20% to 35% corresponding to f^{+2} FM to f^{-2} FM. These results show conclusively that the 3τ version is superior to the 2τ version, hence 3τ subsequences are used in all formulations of Modified Total variance.

2.4.2 Taking advantage of symmetries

Figure 3 shows an example of time-shifted MVAR means (and standard deviations of 1000 simulation trials) and exhibits 2 axes of symmetry located at time-shift $-384\tau_0$ and $+384\tau_0$ for a τ value equal to $256\tau_0$. This means that redundancies allow us to only calculate the subestimates for time-shifts contained between $-384\tau_0$ and $+384\tau_0$. The average of these subestimates is thus exactly the same as a complete calculation of Mod-Totvar, i.e. averaged from time-shifts $-768\tau_0$ to $+768\tau_0$.

Denoting s_k as the subestimate obtained for a time-shift equal to $k\tau_0$, it can be demonstrated that $s_{-3m+k} = s_{-k}$ and $s_{3m+k} = s_{3m-k}$.

Consequently,

- if m is even, we just have to calculate the subestimates from $s_{-3m+3m/2}$ to $s_{3m+3m/2}$ ($6m$ subestimates) ;
- if m is odd, we just have to calculate the subestimates from $s_{-3m+(3m+1)/2}$ to $s_{3m+(3m-1)/2}$ ($6m-1$ subestimates).

The periodicity of the subsequence is equal to $9m\tau_0$ (i.e. 9τ) but, thanks to the symmetries, we only have to consider $6m$ subestimates.

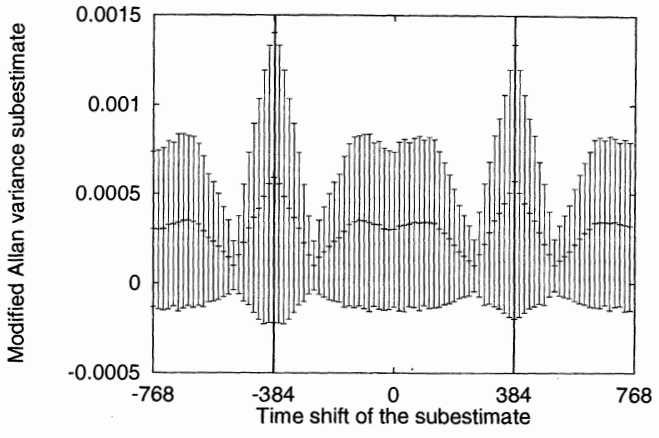


Fig. 3. Time-shifted MVAR mean and corresponding 1σ standard deviation of the mean (by the error bars) after extending a simulated subsequence by even reflection (type 1) for subsequence noise types white PM. The axis of symmetry are located at -384 and +384 for a τ value equal to $256\tau_0$.

3. THE STRUCTURE FUNCTIONS

The concept of structure functions is an extension of the variance approach to n^{th} difference operators [4].

The structure function $\sigma^2_{n,m}(\tau)$ is characterized by its order of difference n and its family m, expressing the "smoothness" (rectangular, linear, quadratic, ...shapes) of its sequence calculation (see figure 4).

The properties of a structure function $\sigma^2_{n,m}(\tau)$ may be summarized by:

- $\sigma^2_{n,m}(\tau)$ is insensitive to phase drifts up to t^n;
- $\sigma^2_{n,m}(\tau)$ converges for low frequency noises from f^{-2n} FM;
- $\sigma^2_{n,m}(\tau)$ converges for high frequency noises up to f^{+2m-2} FM.

It may be noticed that the convergence for low frequency noises and the insensitivity to phase drifts are linked. This property is known as the *moment condition* [9], [10]: it may be demonstrated that the insensitivity to a t^n frequency drift (t^{n+1} phase drift) yields the convergence up to f^{-2n-2} FM.

For example, we may choose the structure function $\sigma^2_{2,2}(\tau)$, called the pulsar variance [11], [12], because it converges for all types of noise and it is insensitive to the quadratic phase drifts.

4. CRITERIA FOR THE EXTENSION TYPE SELECTION

We checked

- 4 structure functions: the Allan variance ($\sigma^2_{1,1}(\tau)$), the modified Allan variance ($\sigma^2_{2,1}(\tau)$), the Picinbono variance ($\sigma^2_{1,2}(\tau)$), the Pulsar variance ($\sigma^2_{2,2}(\tau)$),
- with the 4 extension types described above,
- with or without removal of the linear or the quadratic phase drift (3 cases),

i.e. 48 different estimators.

We applied all these estimators to 100 simulated realizations (16384 data) of each of the 5 noise types (from f^{-2} FM to f^{+2} FM), i.e. 500 noise sequences.

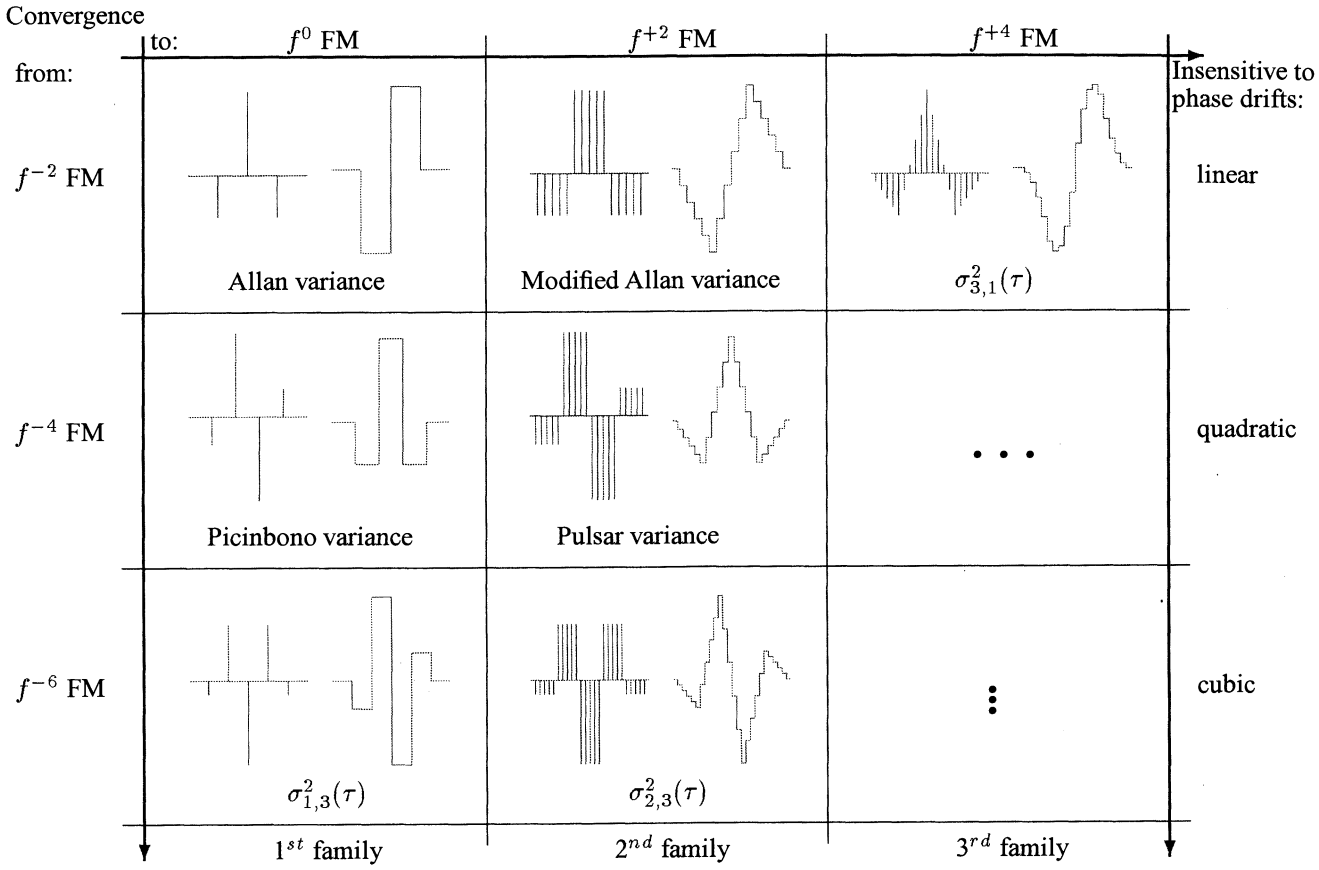


Fig. 4. The different classes of structure functions. The left figures in each box represent the calculation sequences for time error data and the right figures represent the calculation sequences for frequency deviation data. From top to bottom, the order of difference increases, yielding convergence for lower frequency noises and insensitivity to higher order polynomial drifts. From left to right, the different families yield convergence for increasingly higher frequency noises.

4.1 Results:

- extension types 1 and 2 are better than 3 and 4 (bias smaller and edf higher);
- extension type 1 is better than 2 from f^{-2} PM to white PM;
- as mentioned above, extension type 2 must be avoided for f^{+1} PM and f^{+2} PM;
- removing the quadratic phase drift decreases the edf.

4.2 Criteria:

1. since 1^{st} family structure functions don't converge for f^{+1} PM and f^{+2} PM, the type 2 extension should be used;
2. the type 1 extension must be used for the structure functions of the 2^{nd} family and higher;
3. since the 1^{st} order difference (and higher) structure functions are insensitive to linear phase drift, this drift must be removed over each subsequence of length 2τ (lower bias);
4. since the 2^{nd} order difference (and higher) structure functions are insensitive to quadratic phase drift, this drift must be removed over the whole sequence (lower bias) but not over each subsequence of length 2τ since this would decrease the edf (see last result above).

4.3 Examples:

Total variance (Totvar): extension type 2, no removal of drift;

Modified Total variance (Mod-Totvar): extension type 1, removal of the linear drift over each subsequence;

Total Picinbono variance: extension type 1, removal of the linear drift over each subsequence, removal of the quadratic drift over the whole sequence;

Total Pulsar variance: extension type 2, removal of the linear drift over each subsequence, removal of the quadratic drift over the whole sequence.

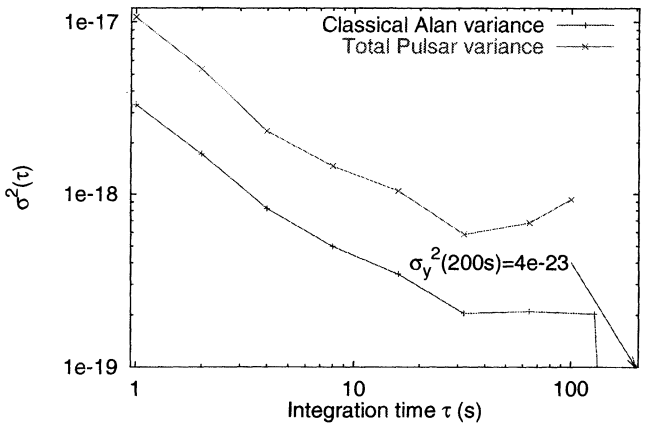


Fig. 5. Comparison of the classical Allan variance and the Total Pulsar variance. These curves were obtained from 400 frequency measurements of a quartz oscillator, with a sampling rate equal to 1s.

5. EXAMPLE

Figure 5 shows the results obtained with the classical Allan variance compared to the results obtained by the Total Pulsar variance applied to the same real sequence. Despite its lower integration time limit ($\tau_{max} = T/4$ for PVAR and $\tau_{max} = T/2$ for AVAR), only the Total Pulsar variance is able to show the positive slope for large τ values, corresponding to the detection of the random walk frequency noise. It may also be noticed that for $\tau = T/2$, AVAR gives a result 5000 times too low.

6. CONCLUSION

By using the Total approach on different variances, the equivalent degrees of freedom of the estimates at and near the longest averaging time $\tau = T/2$ (T is the total duration of the data sequence) increases by a factor of between 2 and 4 relative to the corresponding classical variance. The confidence interval over each variance estimate using the Total approach is then reduced from 70% to 50% relative to the classical variance estimate. Thus, the noise levels are more precisely determined (and this could be crucial, see for example ref. [1]) and the use of a Total variance over a sequence of duration T may be equivalent to the use of the corresponding classical variance over a sequence of duration $2T$.

REFERENCES

[1] F. Vernotte, J. Delporte, M. Brunet, and T. Tournier, "Estimation of uncertainties in time error extrapolation," in *Proc. of the 31th Annual PTTI Meeting*, 1999, p. in process.

[2] E.S. Ferre-Pikal and J.R. Vig et al., "IEEE 1139-1999: Standard definitions of physical quantities for fundamental frequency and time metrology—random instabilities," Tech. Rep., Institute of Electrical Engineers, Inc., New York, NY 10017, USA.

[3] D.A. Howe, "The total deviation approach to long-term characterization of frequency stability," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 2000, to be published.

[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] D. A. Howe and C. A. Greenhall, "Total variance: a progress report on a new frequency stability characterization," in *Proc. of the 29th Annual PTTI Meeting*, 1997, pp. 39–48.

[6] D. A. Howe and F. Vernotte, "Generalization of the total variance approach to the modified allan variance," in *Proc. of the 31th Annual PTTI Meeting*, 1999, p. in process.

[7] C. A. Greenhall, D. A. Howe, and D. B. Percival, "Total variance, an estimator of long-term frequency stability," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 46, no. 5, pp. 1183–1191, September 1999.

[8] M. A. Weiss and D. A. Howe, "Total TDEV," in *Proc. of IEEE International Frequency Control Symposium*, 1998, pp. 192–198.

[9] 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.

[10] 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.

[11] F. Vernotte, M. McHugh, and G. Zalamansky, "Cut-off frequencies and noise power law model of spectral density: Adaptation of the multi-variance method using the structure function approach," in *Proc. of the 9th European Frequency and Time Forum*, 1995, pp. 373–376.

[12] F. Vernotte, "Estimation of the power spectral density of phase: Comparison of three methods," in *Proc. of the Joint Meeting EFTF and IEEE IFCS*, 1999, pp. 1109–1112.

A STUDY OF CORRELATED NOISE WITHIN A CLOCK ENSEMBLE

John A Davis and Benedicte Rueff

Time and Frequency Section,
Centre for Electromagnetic and Time Metrology,
National Physical Laboratory,
Queens Road, Teddington, TW11 0LW.

ABSTRACT

A study has been undertaken, examining the treatment of correlated noise within a clock ensemble. The aim of this work is to enhance NPL's clock algorithm so as to optimise the clock weighting process in the presence of correlated noise. A detailed description of the underlying analysis is presented along with preliminary results using simulated clock data.

1. INTRODUCTION

NPL has recently been developing a clock algorithm to generate a composite timescale from NPL's ensemble of commercial caesium clocks and active Hydrogen Masers [1]. The aim of the work is to improve both the stability and the integrity of the UK National Time scale UTC(NPL).

NPL is particularly interested in generating a time scale with high medium term stability (τ = 1 day) against which to compare NPL's caesium fountain clock which is presently under construction [2]. Although in the process of purchasing a third active hydrogen maser, NPL presently operates only two masers. To achieve the required frequency stability from a composite timescale, NPL must include data from active hydrogen masers remote from NPL. There have been significant recent improvements in satellite time transfer techniques, in particular the development of geodetic GPS time transfer [3]. This has resulted in time transfers with sufficient precision to allow clock data from distant active hydrogen masers to be usefully incorporated into a time scale optimised for medium term stability. The first results have been reported in [4]. Many active hydrogen masers are routinely compared using geodetic GPS time transfer, including those in the GPS network processed by the University of Berne [5], and many International GPS Service (IGS) stations. A major concern when combining clock data from remote locations into a composite timescale is the effect of correlated noise. This may arise from the clocks and more likely from the time transfer systems providing the inter-comparisons.

In this paper, a study is performed on the effects of correlated noise within a clock ensemble. The aim of the work is to enable NPL to develop an algorithm that would generate an optimum composite time scale from an ensemble of clocks in the presence of correlated noise.

2. ANALYSIS OF CORRELATED CLOCK NOISE

Initial work in studying correlated clock noise has been reported by Tavella and Premoli [6,7]. Let $\mathbf{x}_i$ be a vector representing a sequence of m equally spaced phase / time records of a single clock i. The m by n matrix **X** where

$$\mathbf{X} = [\mathbf{x}_1..\mathbf{x}_n]$$

contains a complete set of phase records for the whole ensemble of n clocks. Each value of m represents a separate epoch. **R** is the n by n clock covariance matrix formed from the vectors $\mathbf{x}_i$, and so describes the noise distribution within the ensemble. The use of the classical variance in forming a clock covariance matrix **R** is uncommon, as most noise types found in time series analysis are non-stationary. Often a Allan variance, time variance or variance based on a forward error prediction is used.

Direct measurements of the absolute phase needed to determine the variance of individual clocks are not possible (we can only measure clock differences). Hence it is not possible to determine **R** directly. The variance of comparisons made between pairs of clocks may however be directly determined. The Tavella and Premoli analysis uses a reference clock against which all other clocks are compared [6,7]. A (n-1) by (n-1) measurement covariance matrix **S** is formed from these clock difference measurements. The reconstruction of the clock covariance matrix **R** from the measurement matrix **S** is a highly under-determined problem, and there may be an infinite number of possible solutions for **R**. A "best estimate" of **R** is obtained by finding a solution for **R** that minimises the "global correlation" [6,7].

NPL's clock algorithm, like the NIST AT1 algorithm on which it is based [8,9], does not use a reference clock. Each clock is compared against the composite timescale. Correlations between the individual clocks and the composite time scale are correctly treated by the algorithm, but correlations between individual clocks are not [6]. The algorithm applies weights **w** to each clock assuming that there is no correlation between individual clocks. The aim of this work is to find a series of weights ω that correctly minimises the variance of the composite timescale in the presence of correlated noise.

The Tavella and Premoli approach has been reformulated so that the measurement covariance matrix **S** is an n by n matrix formed from the difference between each clock and the composite timescale, rather than choosing an individual clock as a reference. This formulation has a major advantage that the symmetry between each clock is maintained. The composite timescale should be more stable than any contributing clock. The measured covariance matrix **S** will be dominated by the noise of the individual clocks and so show strong correlation with the clock covariance matrix **R**, and so form a good basis for estimating the individual elements of **R**.

The steps in NPL's analysis are as follows:

1) Compute a measurement covariance matrix **S** from (composite timescale - individual clock) values, using existing clock algorithm weightings to determine the composite.
2) Obtain a single (non-unique) solution for the clock covariance matrix **R**.
3) Determine the solution space for all possible solutions of the matrix **R**.
4) Use a suitable constraining function D to distinguish the realistic solutions from the unrealistic solutions and obtain a "most likely" estimate for **R**.
5) Determine the clock weights from the estimation of the clock covariance matrix **R** and then use these weights in the computation of the composite timescale.

Let $\mathbf{y_i}$ be a vector representing a sequence of m equally spaced records of the phase / time differences between the i th clock and the composite timescale. The m by n matrix **Y** where

$$\mathbf{Y} = [\mathbf{y_1}..\mathbf{y_n}]$$

containing a complete set of phase differences for the whole ensemble. The elements of $\mathbf{y_i}$ may be determined from actual measurements. The matrices **X** and **Y** are related by:

$$\mathbf{Y}^T = \mathbf{H}\mathbf{X}^T$$

where H is the n by n transformation matrix

$$\begin{pmatrix} 1-w_1 & -w_2 & .. & -w_n \\ -w_1 & 1-w_2 & & .. \\ .. & & & .. \\ .. & & & .. \\ -w_1 & .. & .. & 1-w_n \end{pmatrix}$$

and $\sum w_i = 1$. w_i is the weight of the i th clock assuming no correlation. The weights are easily obtained using NPL's existing clock algorithm. The covariance matrices **S** and **R** are related by [6,7]:

$$\mathbf{S} = \mathbf{H}\,\mathbf{R}\,\mathbf{H}^T$$

Because the sum of the column vectors of **H** is **0**, **H** is a singular matrix. There may be an infinite possible solutions for **R** given a matrix **S**. We wish to choose the solutions that is most likely to a good approximation to the clock covariance matrix **R**. This requires adding constraints to distinguish likely solutions from unrealistic solutions.

The covariance matrix **S** is symmetric and positive semi-definite [10]. Because of this it is possible to Cholesky factorize the matrix **S**. The matrix **R** may also be factorised as shown below (not necessarily Cholesky factors).

$$\mathbf{S} = \mathbf{U}\mathbf{U}^T \qquad \mathbf{R}=\mathbf{V}\mathbf{V}^T$$

The matrices **U** and **V** are related by:

$$\mathbf{U} = \mathbf{H}\mathbf{V}$$

Because H is singular, it does not have an inverse. However using a **QR** factorization of **H** [10], a pseudo inverse $(\mathbf{H})^+$ may be determined, from which a solution for **V** and hence **R** may be determined. This solution turns out to be unrealistic. The matrices **U** and **V** may be considered to be composed of a group of column vectors $\mathbf{u_i}$ and $\mathbf{v_i}$. If

$$\mathbf{v_i} = (\mathbf{H})^+ \mathbf{u_i}$$

is a particular solution then $(\mathbf{v_i} + \Upsilon_i\mathbf{q})$ is also a solution where [10].

$$\mathbf{Hq}=\mathbf{0}$$

In this example the vector **q** is the unity vector **1**. By adding the matrix $\mathbf{1}\Upsilon^T$ to the matrix **V** it is possible to generate all possible solutions for **R** simply by adjusting the elements of the vector Υ.

A simple iterative procedure has been established to determine the most suitable values for the vector Υ. and hence a physically realistic solution for the matrix R. Unrealistic solutions may be considered as those where there is strong correlation running through all of the data, or where large negative correlations are present between individual clocks.

The minimisation of the "global correlation" has been examined, both by applying the Tavella and Premoli function [6,7], and also by minimising C where:

$$C = \sum_{i \neq j} R_{ij}^{\ 2}$$

The approach being proposed in this paper is to exploit the similarity between the covariance matrices **S** and **R**, and minimise a function D that has been formed from elements of **S** and **R**. In the case of no correlation between clocks the covariance matrices **R** and **S** may be related as

$$R_{ij} = 0 \qquad S_{ii} = R_{ii}/(1-w_i)$$

$$S_{ij} = -w_j(1-w_i)\, R_{ii} \; -w_i(1-w_j)\, R_{jj}$$

Assuming that the correlated clock noise may be treated as a few small / medium amplitude R_{ij} terms. These R_{ij} terms will then add an additional values to the S_{ij} tems of S_{ij}* where:

$$S_{ij}^* = R_{ij}\,(1-w_i)(1-w_j)$$

By considering other relationships between the Matrices **R** and **S** and adjusting several free parameters a first estimate of a constructing function D was obtained where:

$$D = \sum_{i \neq j} 2(R_{ij}(1-w_i)(1-w_j)/2.5 - S_{ij} - w_iS_{ii} - w_jS_{jj})^2$$

$$+ \sum_{i=j} (R_{ii} - 1.5S_{ii}(1-K))^2$$

$$\text{where} \quad K = \sum_k ((S_{ik} + w_iS_{ii} + w_kS_{kk})/S_{ii}$$

Having obtained a suitable estimate of the covariance matrix **R** the next stage in the analysis is to determine an estimate of the clock weights ω in the presence of correlated noise. The variance of the clock ensemble is given by σ_T where:

$$\sigma_T = \omega^T \mathbf{R}\, \omega$$

ω is a vector containing the weights of individual clocks, and **R** is the clock covariance matrix. Weights need to be determined so that the variance of the clock ensemble σ_T is minimised. Non normalised weights ω' are initially determined where:

$$\omega' = (\mathbf{R})^{-1}\, \mathbf{1}$$

The i th normalised weight ω_i is given by:

$$\omega_i = \omega_i' / \sum \omega_i'$$

To create an iterative and adaptive algorithm the weights ω are regularly recalculated as the elements of the covariance matrix **S** change.

3. EFFECT OF CORRELATED NOISE ON NPL'S CLOCK ALGORITHM

NPL's clock algorithm will overweight correlated clocks. This is due to two effects:

Firstly, correlation between individual clocks will result in a larger than anticipated correlation between the individual clocks and the composite timescale. The noise of an individual clock then appears to be less than its true value when determined by comparisons against the composite.

Secondly, wholly correlated noise between two clocks A and B will add linearly so that the total σ_T is given by:

$$\sigma_T^{\ 2} = (\sigma_A + \sigma_B)^2$$

whereas non-correlated noise will add as:

$$\sigma_T^{\ 2} = \sigma_A^{\ 2} + \sigma_B^{\ 2}$$

Adding clocks with correlated noise therefore results in a larger contribution to the variance of the composite timescale $\sigma_T^{\ 2}$. Assuming that an accurate estimate can be made of the clock covariance matrix **R** the weights ω determined in the presence of correlated noise will be lower for clocks demonstrating strong correlations.

4. SOLUTIONS USING SIMULATED DATA.

Simulations of correlated time series may be easily generated. The actual value of the clock covariance matrix **R** may be determined in these simulations. This is not possible when performing studies using real data. A measurement matrix **S** is easily generated. Examples are shown below for a series of six clocks numbered sequentially from one to six:

$$\begin{pmatrix} 10.29 & 0.09 & 0.12 & 0.16 & 0.38 & 0.51 \\ 0.09 & 10.79 & 0.16 & -0.09 & -0.15 & -0.22 \\ 0.12 & 0.16 & 13.60 & 6.14 & 13.40 & 1.97 \\ 0.16 & -0.09 & 6.14 & 18.18 & 2.23 & 8.34 \\ 0.38 & -0.15 & 13.40 & 2.23 & 30.16 & 9.19 \\ 0.51 & -0.22 & 1.97 & 8.34 & 9.19 & 21.66 \end{pmatrix}$$

*Figure 1: Simulated clock covariance matrix **R** (arbitrary units).*

In this simulation there are only minimal correlations (off diagonal covariance terms) involving clocks one and two. There are significant correlations between clocks three to six.

The associated measurement matrix **S** may be formed from the data set knowing the weights **w** of the clocks determined assuming no correlations (Figure 2).

Similarities between the measurement covariance matrix **S** and the clock covariance matrix **R** may be observed. The variance (leading diagonal) terms are generally lower for the matrix **S**, due to correlations between the individual clocks and the composite timescale. Large covariance (off diagonal) terms in **R** map to lower but still significant covariance values in **S**. Small covariance (off diagonal) terms in **R** map to negative values in **S**. These effects are also due to correlations between the individual clocks and the composite timescale. These trends may be exploited in estimating the clock covariance matrix **R** from the measurement covariance matrix **S**.

$$\begin{pmatrix} 8.84 & -1.21 & -3.43 & -3.44 & -4.83 & -3.14 \\ -1.21 & 9.65 & -3.24 & -3.54 & -5.20 & -3.72 \\ -3.43 & -3.24 & 7.94 & 0.43 & 6.09 & -3.79 \\ -3.44 & -3.54 & 0.43 & 12.42 & -5.12 & 2.53 \\ -4.83 & -5.20 & 6.09 & -5.12 & 21.19 & 1.78 \\ -3.14 & -3.72 & -3.79 & 2.53 & 1.78 & 15.81 \end{pmatrix}$$

*Figure 2: Measurement matrix **S**, calculated from the clock covariance matrix **R** (arbitrary units).*

Several possible techniques for the reconstruction of the clock covariance matrix have been examined. Minimising the Tavella and Premoli function was not succesful.

$$\begin{pmatrix} 14.09 & 4.22 & -1.07 & 0.14 & -1.57 & -0.62 \\ 4.22 & 15.26 & -1.32 & 0.22 & -2.12 & -1.01 \\ -1.07 & -1.32 & 6.17 & 0.10 & 4.56 & -4.77 \\ 0.14 & 0.22 & 0.10 & 14.34 & -3.52 & 3.38 \\ -1.57 & -2.12 & 4.56 & -3.52 & 22.47 & 2.31 \\ -0.62 & -1.01 & -4.77 & 3.38 & 2.31 & 15.59 \end{pmatrix}$$

*Figure 3: Clock covariance matrix **R**, reconstructed by minimising the global correlation.*

Minimising the global correlation $\sum \mathbf{R}_{ij}{}^2{}_{i \neq j}$ also did not produce a good approximation for **R** (Figure 3). The variance of clock 3 is under-estimated, and whilst large covariance (off diagonal) terms are generally identified, they are heavily underestimated. False correlations may also be observed, for example between clocks one and two.

The reconstruction based on the use of the reconstruction function D was more successful. Strong covariance (off diagonal) terms are present at slightly below their true value. Further work is required to improve this reconstruction technique.

$$\begin{pmatrix} 9.76 & -0.47 & -1.06 & -1.05 & -0.47 & -0.35 \\ -0.47 & 10.17 & -1.07 & 1.35 & -1.05 & -0.42 \\ -1.06 & -1.07 & 11.74 & 4.90 & 11.87 & -0.85 \\ -1.05 & 1.35 & 4.90 & 16.26 & 0.67 & 7.48 \\ -0.47 & -1.05 & 11.87 & 0.67 & 28.96 & 8.69 \\ -0.35 & -0.42 & -0.85 & 7.48 & 8.69 & 21.86 \end{pmatrix}$$

*Figure 4: Clock covariance matrix **R**, determined by use of the reconstruction function D.*

$$\begin{bmatrix} 0.244 & 0.233 & 0.185 & 0.138 & 0.083 & 0.116 \end{bmatrix}$$

*Figure 5: Initial weights used in the computation of the measurement matrix **S**.*

$$\begin{bmatrix} 0.196 & 0.182 & 0.235 & 0.159 & 0.099 & 0.128 \end{bmatrix}$$

*Figure 6: Weights $\mathbf{w}^T$ calculated assuming no correlated noise using the measurement matrix **S** shown in figure 2.*

$$\begin{bmatrix} 0.316 & 0.308 & 0.239 & 0.049 & -0.040 & 0.127 \end{bmatrix}$$

Figure 7: Weights ω^T calculated assuming correlated noise using the reconstruction of the covariance matrix **R** shown in figure 1.

The weights $\mathbf{w}^T$, calculated from the measurement covariance matrix **S** assuming no correlated noise and weights ω^T calculated assuming correlated noise from the clock covariance matrix **R** (figure 1) are shown above. The weights ω^T of the clocks one and two which have low correlations with other clocks are increased at the expense of clocks three to six which display strong correlations.

$$[0.302 \quad 0.298 \quad 0.252 \quad 0.096 \quad -0.028 \quad 0.079]$$

Figure 8: Weights ω^T calculated assuming correlated noise using the reconstruction of the covariance matrix R shown in figure 4.

The weights ω^T calculated in figure 8 using the reconstruction of the clock covariance **R** obtained using the reconstruction function D was also encouraging. There were increases in the weights of clocks one and two and reduction of the weights of the other clocks when compared with determining the weights assuming no correlated noise.

5. CONCLUSION

A study has been performed on the treatment of correlated noise within a clock ensemble. An improved method of estimating the clock covariance matrix **R** and of weighting the clocks in the presence of correlated noise has been demonstrated. Further improvement is required before this method may be successfully used in a clock algorithm.

ACKNOWLEDGEMENTS

The authors thank, Peter Harris and Maurice Cox (Centre for Information Systems Engineering) for valuable advice given during the course of the work.

REFERENCES

[1] Rueff B R, Davis J A and Pardo E, "The development of a clock algorithm at NPL" EFTF 2000 (these proceedings).

[2] Whibberley P B, Henderson D, and Lea S N, "Development of a caesium fountain primary standard at the NPL", proc. 1999 joint meeting of EFTF and FCS, Besancon, France, pp 24-26, 13-16 April 1999.

[3] Dudle G, Overney F, Schieldknecht Th, Springer T, Prodt L, "Transatlantic time and frequency transfer by GPS carrier phase", proc. 1999 joint meeting of EFTF and FCS, Besancon, France, pp 243-246, 13-16 April.1999.

[4] Senior K, Matsakis D, "Removing clock jump discontinuities in GPS carrier phase time transfer" proc 31st PTTI-99, Dana Point California, Dec 9-11 1999.

[5] Schieldknecht T, Dach R, Springer T, "Recent results with the transatlantic GeTT campaign", proc 31st PTTI-99, Dana Point California, Dec 9-11 1999.

[6] Premoli A and Tavella P, "A revised three-cornered hat method for estimating frequency standard instability", IEEE Tran. Inst. Meas., **42** , 1, Feb. 1993.

[7] Tavella P and Premoli A, "Estimating the instabilities of N clocks by measuring differences of their readings", Metrologia 1993/4, **30,** 479-486.

[8] Weiss M A, Allan D W, and Peppler T, K, "A study of the NBS time scale algorithm", IEEE Tran. Inst. Meas., **38,** 2, Apr. 1989.

[9] Weiss M A and Weissert T P, "Sifting through nine years of NIST clock data with TA2", Metrologia 1994, **31,** 9-19.

[10] Gill E, Murray W, and Wright M, "Numerical linear algebra and optimisation", vol 1, Addison Wesley, chapter 4.9.2, and chapter 5.3.3.1.

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

A TEST SYSTEM FOR TIME TRANSFER UTILIZING OPTICAL FIBERS

Fang-Dar Chu*, Mingfu Li*, Chia-Shu Liao*, I-Yu Kuo+, Wen-Hung Tzeng+, Yung-Kuang Chen+

* National Standard Time and Frequency Laboratory
Telecommunication Labs., CHT Co., Ltd.
12 Lane 551, Sec. 5, Min-Tsu Rd.,
Yang-Mei 326, Taoyuan, Taiwan.
Tel : +886-3-4244973 E-mail : cfonda@ms.chttl.com.tw

+ Institute of Electro-Optical Engineering, National Sun Yat-Sen University
P. O. Box 59-83, Kaohsiung 804, Taiwan.
Tel : +886-7-5252000 Ext. 4455 E-mail : ykchen@mail.nsysu.edu.tw

ABSTRACT

The precise and accurate characteristic on time and frequency is extremely important in telecommunication and electric power systems. We have begun to investigate time transfer via optical fiber, and a test system has been built since 1998, either. One of the unused overhead bytes in each SONET/SDH frame is employed to transfer a time reference pulse. The test system was divided into three parts such as controller, timing board and transceiver. The transceiver is based on Odetics LIMO SONET/SDH OC-3 interface adapter to access the SONET overhead data. The controller is used to set and control the LIMO. The timing board is designed to transfer the time pulse generated by the atomic clock and to access the time pulse received via LIMO. The test system was configured for a loop-back test. A short-term stability of less than 10 ps has been achieved through short-distance (about 5m) single mode fiber. Another 25km loop-back test has been done in air-conditioned laboratory and the room temperature was maintained within 23 ±5 ℃. The propagation delay variation was about 1.4 ns.

1. INTRODUCTION

Precise time information is indispensable for time synchronization both in telecommunication and electric power systems. Most laboratories transfer time using measurements of the GPS time-code observed simultaneously at two different locations. Study of time and frequency transfer using optical fiber was started by Kihara [1], who reported on accurate time and frequency transfer method using bi-directional wavelength division multiplex (WDM) transmission. The National Institute of Standards and Technology (NIST) has built a two-way time transfer device which uses the unused bytes in the Synchronous Optical Network (SONET) / Synchronous Digital Hierarchy (SDH) overhead to effect the time transfer [2]. The hardware allows time transfer over short distance with stabilities less than 10 ps. In this work, the two-way time transfer using SDH/SONET OC-3 transceivers through conventional single mode fiber (SMF) is tested. The transmission distance considered ranges from several meters to tens of kilometers. The stability of the time transfer system configured in the Loop-back mode is better than 3×10^{-14} at measurement times longer than one day.

2. TWO-WAY TIME TRANSFER

Two-way time and frequency transfer [2] is generally used to compare two geographically separated clocks. The clocks at each end of a link that joins them transmit the time of the local clock and simultaneously receive the time of the remote clock. Each clock then measures the difference between the local clock and the remote clock. If the time difference data from the remote clock are different from that of the local clock, then the path delay effects can be removed, assuming that the path from the remote node to the local node is reciprocal with the path form the local node to the remote node. The time of the remote clock (clock B) relative to the local clock (clock A) can be written as [3]:

$$Time(A) - Time(B) = \frac{[TIC(A) - TIC(B)] + [T_x(A) - R_x(A)] - [T_x(B) - R_x(B)]}{2}$$

where

$TIC(i)$ = Time between transmitting and receiving 1 PPS at node i

$T_x(i)$ = Transmitting delay through hardware at node i

$R_x(i)$ = Receiving delay through hardware at node i

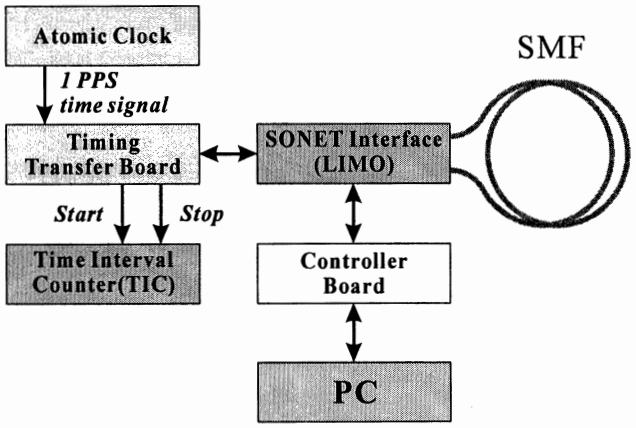


Figure 1. Block diagram of two-way time transfer system in loop-back mode

3. SYSTEM CONFIGURATION

Figure 1 shows the Loop-Back configuration of the two-way time transfer system. In this system, We utilize the unused overhead bytes in each SONET/SDH frame to transfer a reference pulse. The transmitted and returned pulses employ different overhead bytes. The system uses Odetics Inc. SONET/SDH OC-3 (155.52 Mbps) interface adapter, called LIMO, to access the SONET overhead data. We also design two devices, one is the controller board used to setup and control LIMO, the other is the timing transfer board provided to transfer a reference pulse generated by an atomic clock to LIMO and access the timing data from LIMO. The timing transfer board can also generate the pulses as the start and stop commands for the time interval counter.

4. RESULTS AND DISCUSSIONS

In the first experiment, the length of the used single mode fiber is 5 m. Because the fiber is too short, one can expect that the delay is mainly influenced by the hardware but not by the fiber path. The propagation delay is shown in Figure 2(a), and its variation is within 850 ps. The frequency stability [4], $\sigma_y(\tau)$, is plotted in Figure 2(b). The counter error due to the Time Interval Counter (Stanford Research Systems; model SR620) is also plotted. The experiment result shows that the frequency stability is better than 3×10^{-14} over a measurement period of one day.

In the second experiment, we use a single mode fiber with 25 km in length for the loop-back configuration. The whole system is set and placed within the laboratory. Under such a small room the temperature usually varies slightly. The propagation delay data is shown in Figure 3(a). The variation range of propagation delay is 1.3 ns, which is much larger than the former result in the first experiment. The thermal coefficient of the bare fiber on delay is 37 ps /℃/ km [5]. Thus, the delay variation will be increased when the fiber link distance increases. Fortunately, the variation due to the transmission path can be cancelled in the two-way configuration if the path is well symmetrical. The stability in the test is shown in Figure 3(b). The result exhibits frequency stability in the long-distance as good as that in the short-distance. Accurate time transfer requires the absolute magnitudes of the delay to be small; but accurate frequency transfer only requires the delay to be stable. These results show that our two-way time and frequency transfer system has a capability of frequency stability to 3×10^{-14}.

The time deviation performance $\sigma_x(\tau)$ is indicated in Figure 4. One can observe that the short-term stability is better than 10 ps for short-distance.

5. CONCLUSIONS

We operate the time transfer system in loop back mode, and demonstrate the frequency stability at 1 day of 3×10^{-14} approximately. These results show that the integral two-way time transfer system has a capability of transferring time and frequency accurately.

REFERENCE

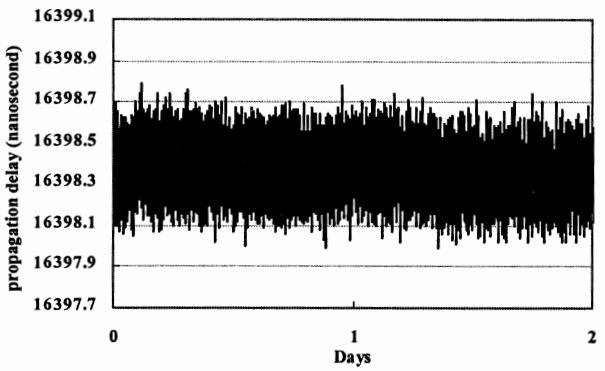
[1] A. Imaoka and M. Kihara., "Accurate time/frequency transfer method using bidirectional WDM transmission ," Electro. Lett., vol. 31, No. 16, pp. 1361-1362, August, 1995.

[2] S. R. Jefferts, et al., "Two-way time and frequency transfer using optical fibers," IEEE Transactions on Instrumentation and Measurement, vol. 46, No. 2, pp.209-211, April, 1997.

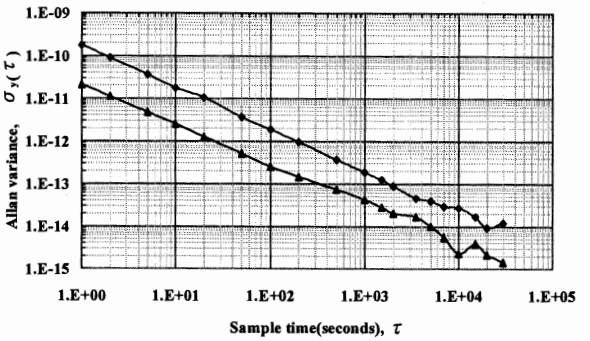
[3] C. Hackman, et al., "Common-clock two-way satellite time transfer experiments," Proc. 49th Annu. Symp. Frequency Control, pp. 275-281, 1995.

[4] D. W. Allan, et al., "A frequency-domain view of

time-domain characterization of clocks and time and frequency distribution systems," Proc. 45th Frequency Contr. Symp., pp.667-678, 1991.
[5] N. Shibata, et al., "Thermal characteristics of optical pulse transit time delay and fiber strain in a single-mode optical fiber cable," Appl. Opt., vol. 22, pp. 979-984, 1983.



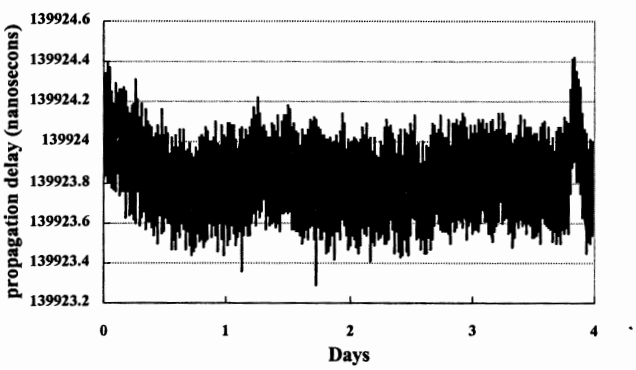
(a) Propagation delay

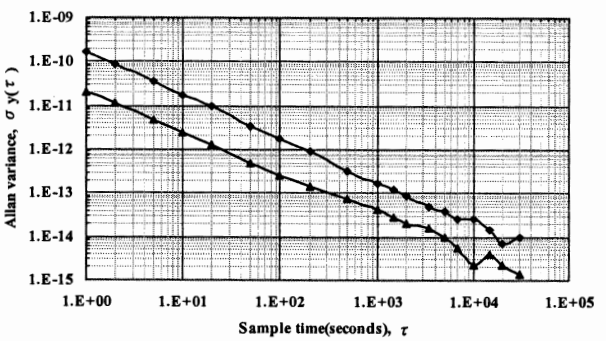


-●- loop-back test
-▲- counter error

(b) Frequency stability of loop-back test and the counter error

Figure 2. The loop-back test of two-way time transfer system with 5 m SMF link.





-●- loop-back test
-▲- counter error

(a) Propagation delay

(b) Frequency stability of loop-back test and the counter error

Figure 3. The loop-back test of two-way time transfer system with 25 km SMF link.

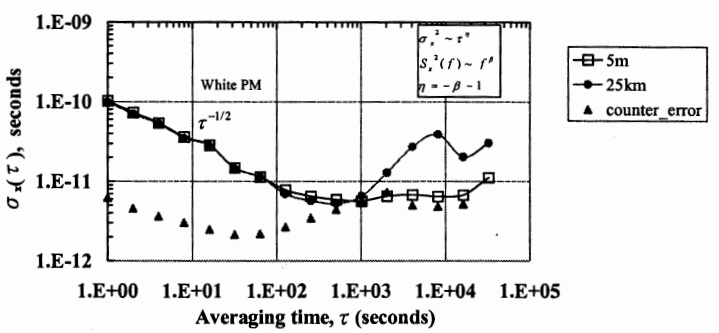


Figure 4. The time deviation performance of two-way time transfer system with SMF link.

FREQUENCY INTERLABORATORY COMPARISON BETWEEN ESTONIA AND FINLAND

K.Kalliomäki [1], T.Kiibus [3], T.Mansten [1], R.Rebane [2]

[1)] MIKES, Electricity and Time Metrology, Otakaari 7B, FIN-02150 Espoo, Finland
[2)] EKTA Cybernetics, Akadeemia tee 21, 12618 Tallinn, Estonia
[3)] AS METROSERT, Aru 10, 10317 Tallinn, Estonia

K. Kalliomäki, MIKES, Electricity and Time Metrology, Otakaari 7B, FIN-02150 Espoo, Finland
Tel. +358-40-510 4411, Fax +358-9-456 7029, email kalevi.kalliomaki@vtt.fi

ABSTRACT

A common view frequency comparison via Finnish TV broadcast signal was successfully accomplished between METROSERT and VTT Automation, Measurement Technology. The distance between laboratories is about 80 km over sea and both labs can receive Finnish TV1-program from Espoo. TV-synchronisation signals are locked to rubidium oscillator. Both direct frequency and phase comparison methods were utilised. Phase comparison method gives better results but suffers from unknown phase jumps requiring rather long comparison time. An inaccuracy of $1 \cdot 10^{-12}$ at best was obtained.

1. INTRODUCTION

Estonia is establishing its own accreditation system and its own national standards laboratories. METROSERT in Tallinn takes care of the most accurate calibrations in the field of electricity. METROSERT is accredited by Finnish accreditation system (FINAS). Accredited quantities are DC voltage and DC resistance, but METROSERT plans to add some new quantities to its scope, including frequency. A common view frequency comparison via Finnish TV broadcast signal was successfully accomplished between METROSERT and VTT Automation, Measurement Technology in September-October 1999.

MIKES VTT, Electricity and Time Metrology is the National Metrology Institute (NMI) of Finland. It was separated from VTT Automation at the end of the year 1999, after this comparison. MIKES maintains time and frequency by comparing four Cs standards to each other and to GPS receivers with a monitoring system /3/.

MIKES has a portable rubidium clock, which is used for frequency proficiency tests in Finland (/1/). Usually the clock is used for short distances (less than 1 hour) with a car and powered from the car's battery, because the accredited laboratories are situated in the vicinity of MIKES. The frequency uncertainty level of this kind of proficiency tests has been about $1 \cdot 10^{-11}$ in Finland. The distance between Helsinki and Tallinn is about 80 km. The connection is provided via ferries, see maps in Figs. 1 and 2.

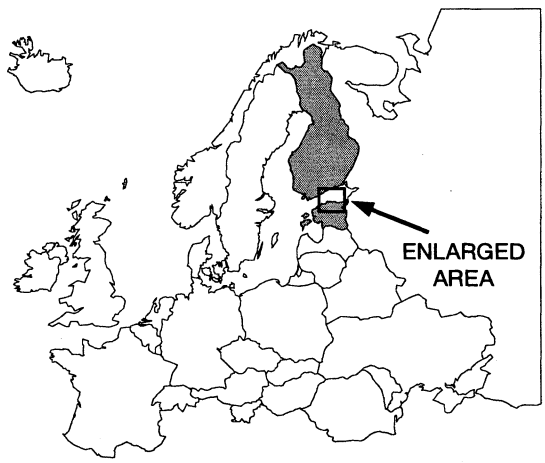


Fig. 1: A map showing Estonia and Finland

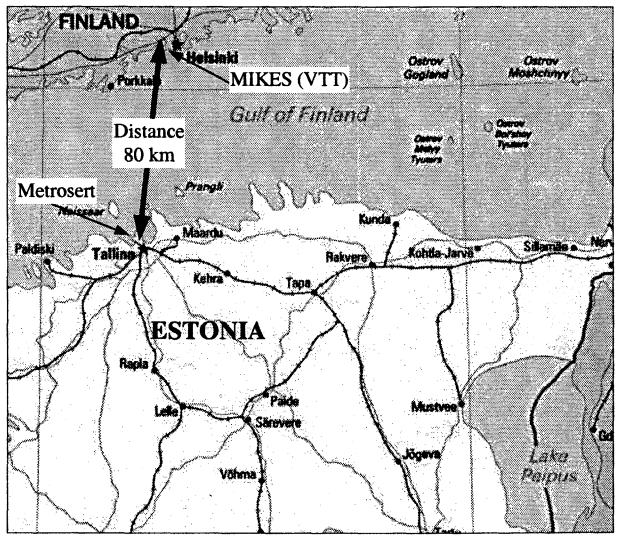


Fig. 2: A map showing locations of the co-operating laboratories

The expected uncertainty level for a one-day visit of the Rb-clock via a ship would be larger than $1 \cdot 10^{-11}$. Because of the relatively good equipment level of METROSERT (four Rb clocks, a high-resolution frequency comparator and a GPS receiver, manufactured by EKTA) a better level was desired (see Fig. 3).

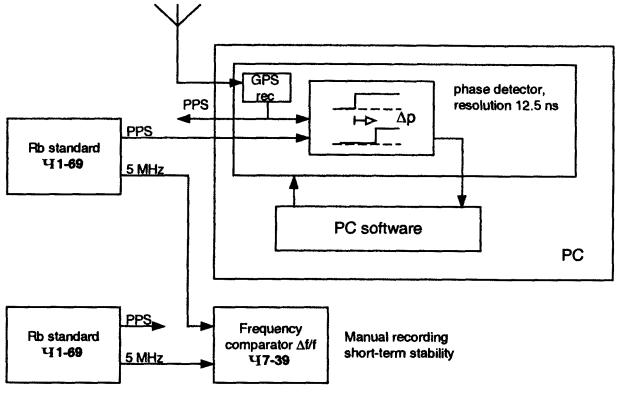


Fig. 3: The frequency traceability of METROSERT

The 15.625 kHz line frequency and other synch. signals of the TV transmissions of the Broadcasting company of Finland are based on a very stable Rb oscillator, monitored by the NMI. The line frequency has been the basis of frequency traceability in the accredited calibration laboratories in Finland. /2/. Up to the spring 1999 the overall quality of the line frequency was excellent. From that on the quality of the line frequency deteriorated due to the new digital links which contain picture memories and reproduce the line frequency after every link. The 24h average has been of good quality, but with shorter integration times there are slower (hours) and faster (minutes) variations, as can be seen from the Allan variance of the line frequency in Fig. 4.

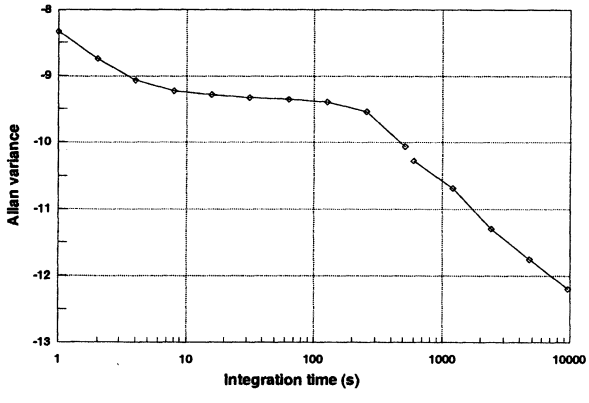


Fig. 4: The Allan FVAR of the Finnish TV line frequency

2. METHODS

To check the performance of the TV-comparison method, a short baseline (7.2km) test was done first in Finland and Allan FVAR was calculated from the phase record, Fig. 5. In this case both reference standards were Cs-clocks. Allan FVAR plot shows, that an uncertainty of $1 \cdot 10^{-11}$ can easily be attained in less than one hour and $1 \cdot 10^{-12}$ is possible within office hours.

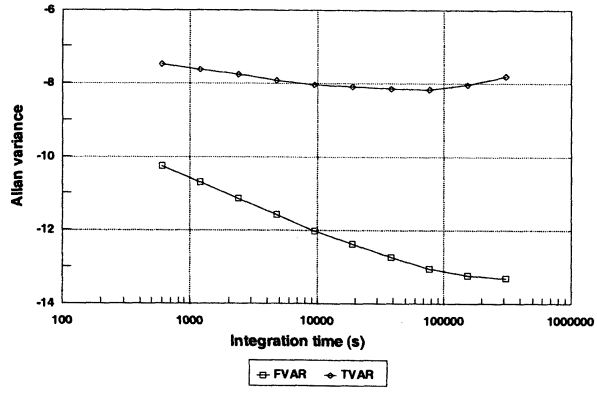


Fig. 5: The Allan variances of TV common view, baseline test, between two Cs-clocks

Because METROSERT did not have a phase measuring system for TV-synch. pulses, a direct frequency measurement was tried first. To accomplish this, a TV-frequency standard /2/ was sent to METROSERT, see Fig. 6. It contains a simple crystal oscillator and phase lock circuitry to lock into 15625 Hz line frequency of any TV-set.

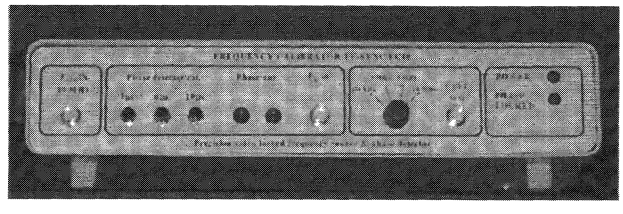


Fig. 6: The TV frequency standard of MIKES

METROSERT has a high resolution frequency comparator, ie. the dominating noise originates from TV-signal and above-mentioned phase-lock circuitry. As a first method 10 s readings were taken manually at 30 s intervals during one hour, see Fig. 7.

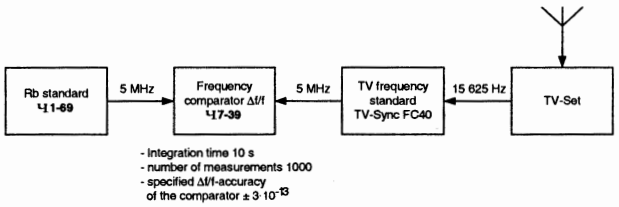


Fig. 7: The direct frequency comparison method

As a second method a long term phase measurement was carried out utilising the linear phase comparator included in the TV-frequency standard, see Fig. 8.

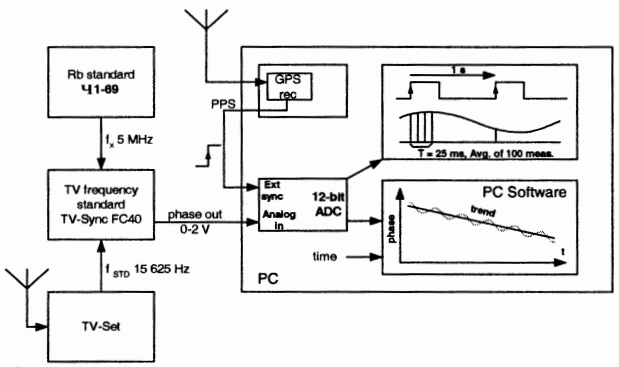


Fig. 8: The frequency comparison method via phase

MIKES monitors TV line frequency via a video recorder as a TV receiver, a frame sync. separator for video and a surveillance system (/3/) which measures all clocks, including 50 Hz frame pulses against the PPS pulses of the Cs main clock of MIKES.

3. RESULTS AND DISCUSSION

Fig. 9 shows the results of simultaneous frequency measurements at both ends. The common sinusoidal frequency variation of $3 \cdot 10^{-9}$ (pp) comes from the digital links explained earlier. Instability (SD) of each reading was $0.4 \cdot 10^{-9}$ and the final inaccuracy of the comparison was $3 \cdot 10^{-11}$ (SD). The result, df/f = $(1.6 \pm 3) \cdot 10^{-11}$, was not satisfactory, therefore we continued with phase measurements.

Raw phase plots at both ends, Helsinki and Tallinna are shown in Fig. 10. There are some non-coincident phase jumps we cannot explain. Probably they originate from TV-frequency standard which may lose phase-lock due to bad signal, noise or power line problems.

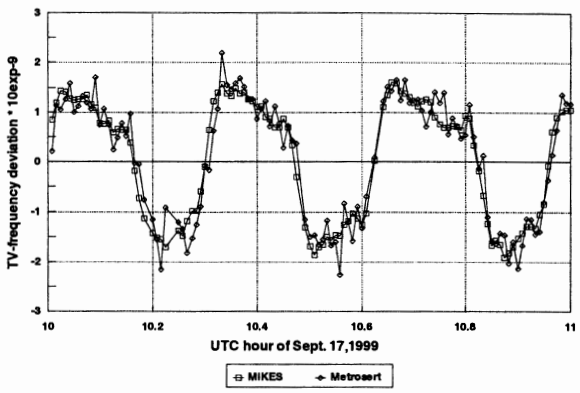


Fig. 9: The direct frequency comparison result

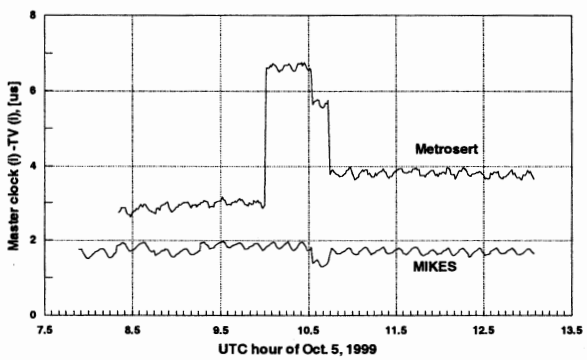


Fig. 10: The raw phase data at both ends

The phase plot is depicted in Fig. 11 after the elimination of obvious phase jumps. It contains a strange 700 ns_{pp} phase variation in the middle of comparison period. At the beginning of the period the linear behaviour turns into parabola due to warming-up phenomena (probably TV-set). When applying linear regression to the whole curve, a frequency difference seems to be $-20 \cdot 10^{-12}$ with the standard error of $1{,}3 \cdot 10^{-12}$. The observed frequency difference seems to decrease towards the end of comparison. Last two hours of the comparison period were quite free from phase jumps and above-mentioned thermal drifts. The best 2.5 hour piece of the phase difference gives a frequency difference of $-1{,}3 \cdot 10^{-12}$ with A-type uncertainty of $0{,}5 \cdot 10^{-12}$. The corresponding Allan FVAR is shown in Fig. 12.

At the time of frequency comparison there was no air conditioning in the METROSERT frequency laboratory. In addition, a window, looking towards south-east, was near the frequency standards. Therefore, the laboratory temperature may have varied several degrees during the comparison. The rubidium frequency standards and frequency comparator were installed in the open rack one

upon the other, a thermal design which may cause problems.

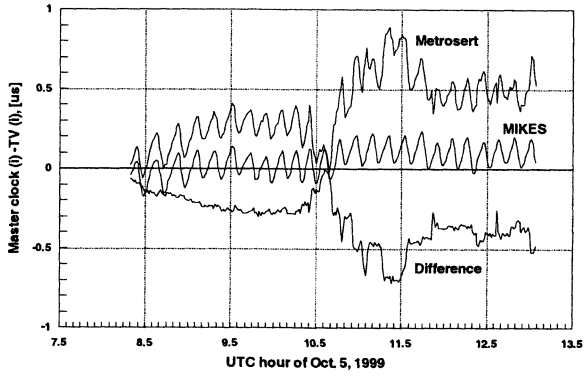


Fig. 11: The corrected phase data at both ends and their difference

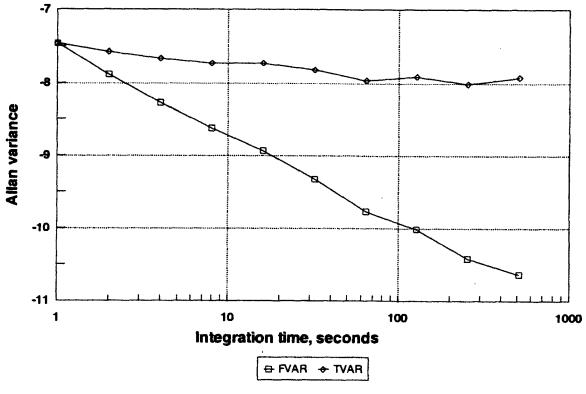


Fig. 12: The Allan FVAR and TVAR of the "best piece" phase difference

4. CONCLUSIONS

Phase comparison method is superior compared to direct frequency measurement. It suffers, however, from unknown phase jumps, which are not related to common signal source. In addition, environmental conditions were not in control during the comparison, leading to phase drifts, which makes the results difficult to decipher. It seems that the results were more stable towards the end of the five hour measurement period.

The frequency comparison was successful. METROSERT stated the frequency deviation of their reference clock. It was compared to the reference frequency of MIKES via TV common view comparison. The difference between the stated (METROSERT) and measured (MIKES) reference frequency deviation of METROSERT was clearly less than $1 \cdot 10^{-11}$ during the comparison. By utilising longer comparison times (up to several days) bad phase records could be rejected and an uncertainty of $1 \cdot 10^{-12}$ should be easily attainable by using the described method.

5. REFERENCES

/1/ The modification of a hp 105b frequency standard and its use as a portable frequency reference, Kalliomäki, Kalevi; Mansten, Tapio, proceedings of the 12th European Frequency and Time Forum. Warsaw, Poland, 10-12 March 1998, Tele & Radio Research Institute (1998), pp. 484-486

/2/ TV-frequency standard as a traceability source in accredited calibration laboratories, Kalliomäki, Kalevi; Mansten, Tapio, proceedings of the7th European frequency and time forum, Neuchatel, 16 - 18 March 1993, Switzerland. (1993) , pp. 383-385

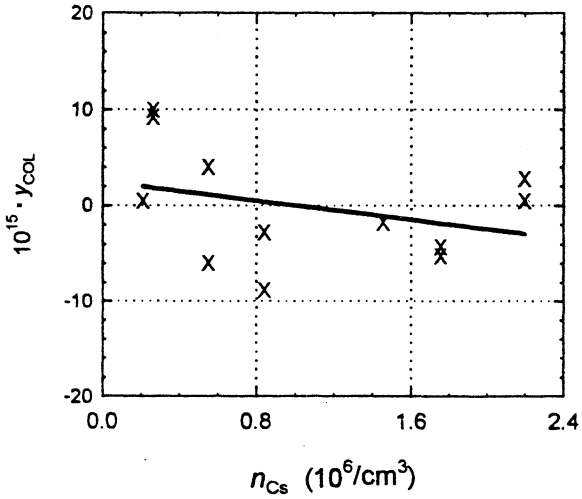
/3/ A Computerized Time and Frequency Standards Monitoring System, Kalliomäki, Kalevi; Mansten, Tapio, proceedings of the 4th European Frequency and Time Forum/EFTF 90, Neuchatel, Switzerland, 13 - 15 March 1990. (1990), pp. 177 - 179

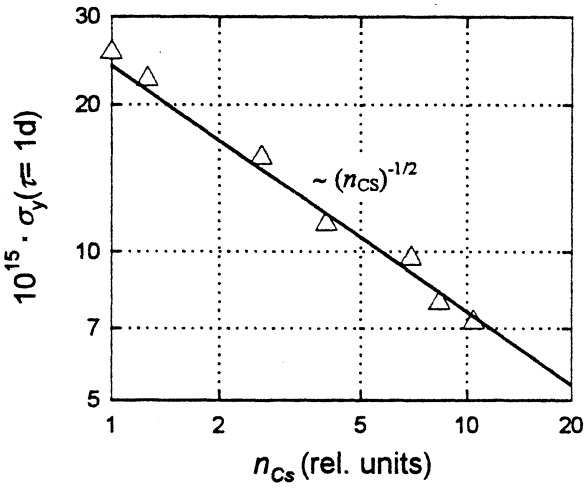
STUDIES ON COLLISIONAL FREQUENCY SHIFTS IN CAESIUM ATOMIC CLOCKS WITH A THERMAL BEAM AND MAGNETIC STATE SELECTION

R. AUGUSTIN[(1)], A. BAUCH[(1)], M. KAJITA[(2)]
[(1)] Physikalisch-Technische Bundesanstalt, Braunschweig, Germany
[(2)] Communications Research Laboratory, Tokyo, Japan

Currently, the collisional frequency shift in fountain type frequency standards is a matter of vivid interest. Collisions in a thermal gas and a thermal atomic beam have been discussed now and then. Only recently one of the authors (M.K.) pointed to the fact that under the experimental situation prevailing in an atomic clock with a thermal beam having a narrow axial atomic velocity distribution and very low transversal velocities, a shift might be expected which should even at moderate atomic densities in the beam be comparable in magnitude with the uncertainty of such standards.

In PTB's caesium frequency standard CSX the density of atoms contributing to the detector signal has been varied by about one order of magnitude, up to about $2 \cdot 10^6/cm^3$, in such a way that all known frequency shifts should have remained constant. Frequency measurements were performed between February and July 1999 with reference to PTB's primary clocks CS1 and CS2, and to UTC(NIST) and UTC(USNO).





In the figures, the average result of all frequency comparisons (X) is plotted as a function of the atomic density, and the measured relative frequency instability $\sigma_y(\tau=1d)$ (Δ), obtained from hydrogen maser comparisons. Each data point corresponds to typically 10 days of measurement. The results (Δ) prove shot-noise limited performance of the CSX for all atomic densities employed.

A linear fit to the frequency shift data (X) gives $\Delta y_{col}/\Delta n_{cs} = -2{,}5 \cdot 10^{-15}/(10^6$ atoms/cm^3) with $2{,}3 \cdot 10^{-15}$ as the 1σ - standard uncertainty of the slope. This result would have to be multiplied by (-4,8) to agree with M. Kajita's theoretical predictions (unpublished). The difference may be due to the modelling of the atomic density and the transversal velocity distribution in the theoretical calculations, which may be too simple to reflect the real situation in the frequency standard.

In the contribution we will explain details of the experiment, of the data analysis and of the theory which predicts the frequency shifts.

COORDINATION OF POLISH TIME AND FREQUENCY LABORATORIES AND CONSTRUCTION OF A NATIONAL ATOMIC TIME SCALE

J. NAWROCKI, *Astrogeodynamical Observatory, Borowiec, Poland*
J. SIEMICKI, *Central Office of Measures, Warsaw, Poland*
Z. RAU, *Institute of Telecommunications, Warsaw, Poland*
W. STRUS, *Polish Telecommunications Company, Warsaw, Poland*
A. STACHNIK, *Institute of Telecommunications, Warsaw, Poland*
J. AZOUBIB, *Bureau International des Poids et Mesures, Sèvres, France*
W. LEWANDOWSKI, *Bureau International des Poids et Mesures, Sèvres, France*

ABSTRACT

Several Polish governmental, scientific and commercial institutions are equipped with caesium frequency standards, which until recently were isolated and compared only occasionally one with another. The total number of standards approaches ten; most are type HP5071A. Development by the Astrogeodynamical Observatory (AOS), in cooperation with the Bureau International des Poids et Mesures (BIPM), of a cheap multi-channel time-transfer system (TTS-2) based on a Motorola GPS module has allowed these laboratories to quickly establish a system of time links. Most of the institutions are now linked through the TTS-2 to the Central Office of Measures (GUM) in Warsaw on an operational basis.

The TTS-2 comprises a Motorola VP Oncore multi-channel GPS C/A-code module, a dedicated personal computer, and software for common-view comparisons. The software uses raw pseudo-range measurements and rigorously follows standards developed by the CGGTTS (CCTF sub-group on GPS and GLONASS Time Transfer Standards). This new system enables the observation of several satellites simultaneously. Another of its advantages is that the data obtained are compatible with the earlier single-channel standard recommended by the CCTF. The multi-channel method improves the stability of the time transfer in comparison to the classical single-channel common-view mode. The results of such comparisons between several laboratories are presented.

Data from Polish frequency standards, compared by means of TTS-2s, are used to establish the Polish Atomic Time Scale TA(PL). Based on about ten caesium frequency standards, this will be much more stable than either of the existing scales generated by a single laboratory. When fully operational, this time scale will help to maintain two Polish realizations of UTC, UTC(GUM) and UTC(AOS), to within 100 ns of UTC.

ULTRA PRECISE TIME DISSEMINATION SYSTEM

C.LOPES, B.RIONDET

IN-SNEC a company of INTERTECHNIQUE GROUP

ABSTRACT

Last year at the Joint Meeting of the 13th European Frequency and Time Forum and the IEEE International Frequency Control Symposium in Besançon (FRANCE) 12-16 April, 1999, we presented an Ultra Precise Time Dissemination System reporting the possibility of performing time transfer between two areas separated by few kilometers with an accuracy of a few picoseconds. We demonstrated the ability to synchronize remote systems within 10 to 30 picoseconds over a temperature variation of 4°C. We have already performed the following improvements as announced in our previous paper:

The operating temperature range has now been extended from 15°C to 35°C.

The path variation range can be as high as several nanoseconds and the synchronisation process is maintained in the picosecond range.

INTRODUCTION

Our previous work has shown the architecture of the system.[5]

The figure 1 shows the setup used in order to validate the system.

We perform a transit time measurment between a master clock system and remote slaves sub-systems, after computation of the transit time measurement by the master

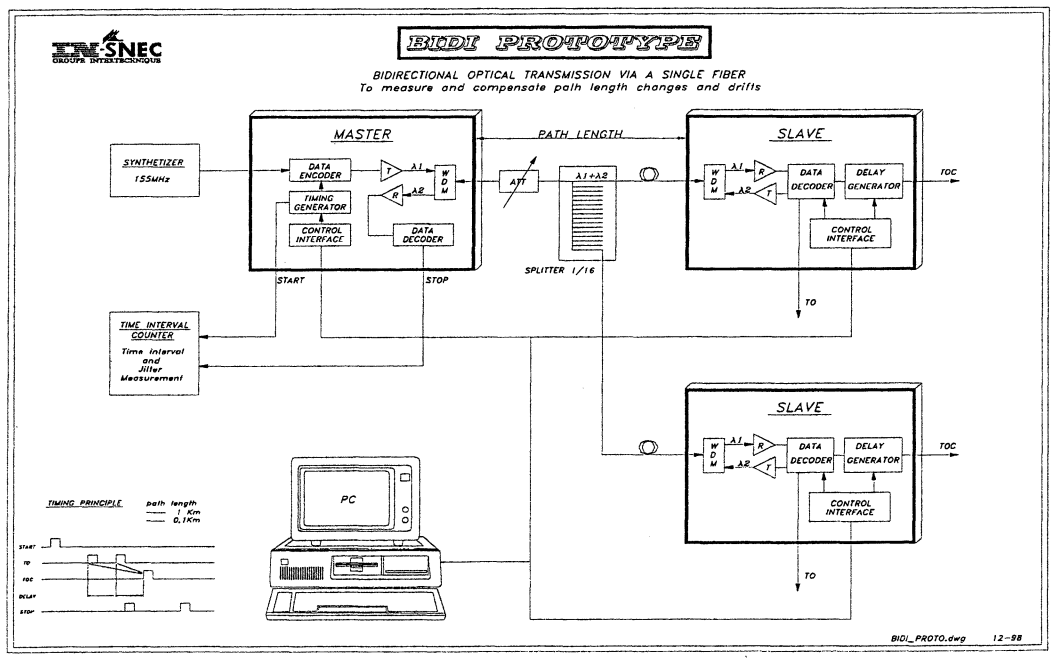


Figure1: Setup used to validate the system

computer, data are transfered to slave equipment to acco-

modate and compensate time differences and drifts.

To improve the results presented last year regarding temperature we had to work in two areas.

- Increasing the corrections efficiency of the delay generator over the 15 - 35 °C temperature range.
- Increasing the transit time measurement accuracy over the 15 - 35 °C temperature range.

SYSTEM IMPROVEMENT

Corrections of the delay generators

The simplified slave sub-system schematic is given by the figure 2:

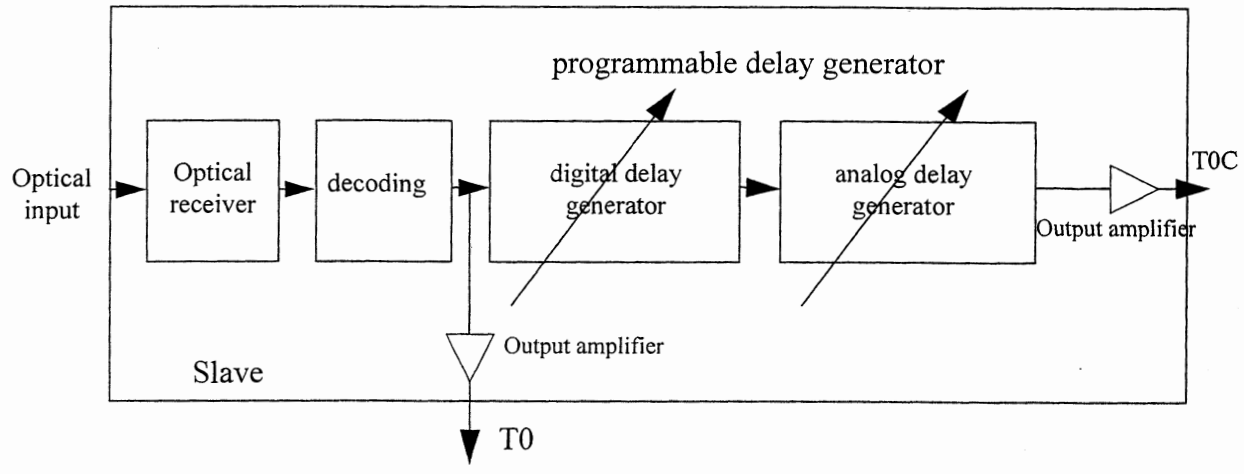


Figure 2: Simplified schematic of the slave sub-system.

The optical signal is demodulated, decoded and then transmitted to a programmable delay generator. The signal is also sent back through the optical link to the master.

The programmable delay generator is a two-stage delay generator. The first stage is digital and consist of a synchronous counter clocked by the 155.52 MHz clock signal providing a 6.4 ns resolution and a 100µs range.

The second stage is an analog interpolator providing a resolution of 400 femtoseconds and a delay range covering the 6.4 ns clock period.

The combination of these two stages allows to have a delay range of 100 µs and a resolution better than 1 picosecond.

The digital stage has no identified non-linearities.

The analog stage has some non-linearities over the period range of 7 ns.

The figure 3 shows the non-linearities of the analog interpolator over the 7 ns. We can observe that the worst

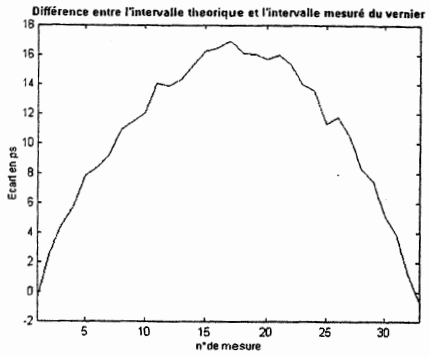


Figure 3: non-linearities of the analog interpolator

case of the non-linearities is in the middle of the delay range (about 3.5 ns) and it is 16 to 17 picoseconds.We can see that such a curve is easy to approximate by polynomial interpolation to implement the appropriate correction.

The figure 4 shows the residual non-linearities after the application of the correction over the delay range of the analog delay generator.

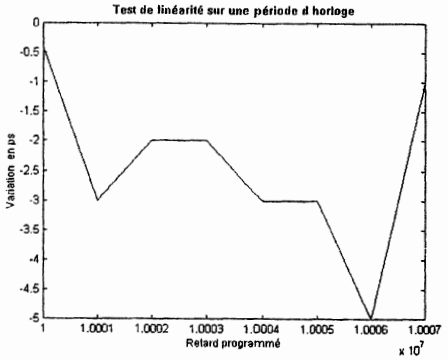


Figure 4: residual non-linearities after correction range: 0 to7ns

We represent the difference between the programmed delay and the measured delay at the output (T0, T0C) of the

We represent the difference between the programmed delay and the measured delay at the output (T0, T0C) of the delay generator for programmed delays from 10.001 µs to 10.007 µs by step of 1 ns.

We can appreciate that the residual non-linearities are on the order of 5 picoseconds peak to peak.

The figure 5 shows the residual non-linearities after the application of the correction over the delay range of 100 ns to 100 µs showing the residual non-linearities of the whole delay generator including both digital and analog stage.

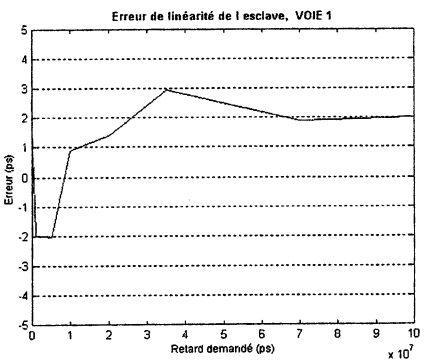


Figure 5: residual non-linearities after corection range: 100 ns to100 µs

We can appreciate that the residual non-linearities of the complete delay generator in the range of 100 ns to 100 µs is also as low as 5 picoseconds peak to peak.

We will discuss now about another source of errors in our timing system and especially due to the delay generator.

Even if we set the two stages of the delay generator to zero the effective delay will not be equal to zero, the delay generator due to transmission time in electronics of the two stage has a minimum delay.

This minimum delay is not constant over temperature.

The figure 6 shows the minimum delay variation over an ambient temperature varying from 15°C to 35°C.

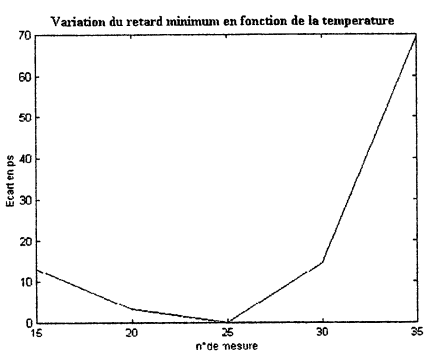


Figure 6: Minimum delay variations over temperature.

This variation of 70 ps of the minimum delay over the temperature range is also compensated by polynomial interpolation.

The figure 7 shows the difference between a programmed delay of 50 µs and the measured delay at the outputs (T0, T0C) of the delay generator for the ambient temperature varying from 15°C to 35°C.

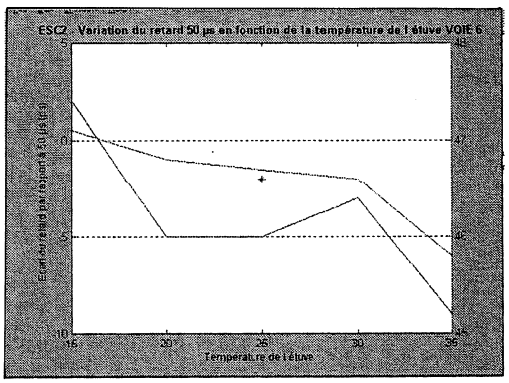


Figure 7: Minimum delay compensation over temperature.

The lower curve shows the errors between the programmed delay and the effective measured delay. We can see that the peak to peak variation is 10 picosecondes.

The upper curve show the temperature of the time interval meter.

We can see that the two curves area correlated, we believe that temperature on the time interval meter as a significant influence on the accuracy of the results when we are working below ten picosecondes.

Transit Time Measurement Accuracy

To perform an accurate synchronisation process between two remote equipment after having implemented an accurate delay generator, we should implement an accurate transit time measurement method.

The method already existed [1,2,3,4,5] but we worked to increase the accuracy of this method over the temperature range 15°C - 35°C.

The figure 8 shows the slave ambient temperature variation and its influence on the transit time measurement accuracy.

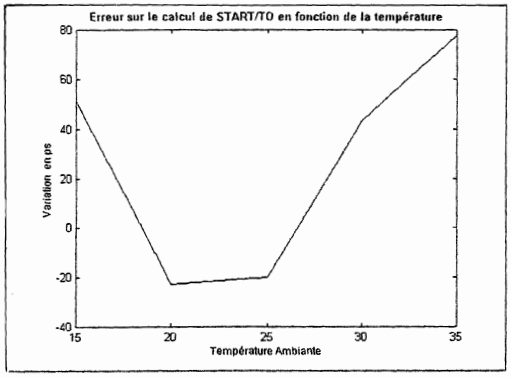


Figure 8: Temperature variation on the slave and its influence on transit time measurement accuracy.

We can see a peak to peak variation of 100 picoseconds.

The temperature on the master has the same influence on the the transit time measurement accuracy.

We should have a correction for the slaves temperature and the master temperature.

The figure 9 shows the variation of the delay programmed between two remote slaves when the ambiant temperature of one of the slave is varied from 15°C to 35°C.

We have programmed a delay of 50 µs between the outputs of the two slaves and we measured the difference between this two slaves and we represent the difference between the programmed and the measured delay.

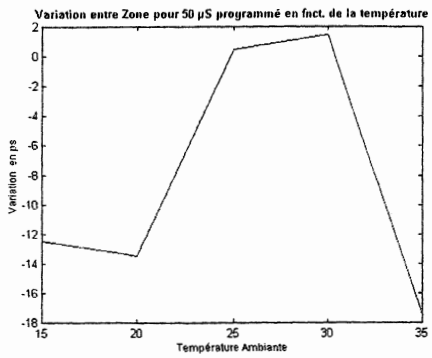


Figure9: variation of the delay programmed between two remote slaves versus temperature.

We can see that the we have a maximum variation of 15 picoseconds peak to peak between two remotes slaves.

The master and one of the slaves are at room temperature and the ambient temperature of the other slaves varies from 15°C to 35°C.

One of the slaves is linked to the master by 100 m of optical fiber and the other one is linked by 500m.

The 500 m of fiber optic is also exposed to the temperature variation; this represents a 2 ns transit time variation that is fully compensated by the system.

CONCLUSION

From last year we made a significant improvement in our system.

Last year we demonstrated that we were able to deliver over a wide area time epochs synchronous in a few picoseconds range implementing a suitable compensation for time propagation variation delay over 4°C temperature variation.

Now we have improved temperature corrections of the electronic sub-systems; we maintain the same timing performances (10 to 15 ps) over a temperature variation of 15 °C to 35 °C of the equipment.

The equipment are now in production and will be delivered to our customer in the next months.

The main challenge for the future is to increase the time performances, our goal is to have a synchronisation process between two remote equipment down to 5 picoseconds.

REFERENCES

[1] L. Primas R. Logan, G Lutes, "Application of ultra-stable fiber optic distribution system ," 43rd Annual Symposium on frequency control 1989 pp 202,211 June 89 IEEE.

[2] S.R. Jefferts, M Weiss, S. Dilla and T. Parkeer, «Two-way time and frequency transfer using optical fibers» Conference on precision electromagnetic measurements 17-21 June 96 Braunschweig, Germany pp. 520, 521 IEEE, ISBN 0 7803 3376 A.

[3] C. Lopes «Maquettage et validation d'une lisaison optique numérique bidirectionnelle faible gigue» June 1998 private communication.

[4] C. Lopes «Cahier de Résultats d'Essais de la maquette de liaison bidirectionnelle». November 1998 private communication

[5] C. Lopes,B. Riondet «Ultra Precise Time Dissemination System». Joint Meeting of the Frequency Control Symposium and the Eurpean Frequency and Time Forum April 1999, Besançon, France.

B.RIONDET/C.LOPES
IN-SNEC
5 Av. des Andes
91943 Les Ulis Cedex
FRANCE
tel: 33 1 69 82 79 70
fax: 33 1 69 07 39 50
e-mail: briondet@wanadoo.fr

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

THE DEVELOPMENT OF A CLOCK ALGORITHM AT NPL

Benedicte Rueff, John A. Davis and Eulogio Pardo

Time and Frequency Section
Centre for Electromagnetic and Time Metrology,
National Physical Laboratory,
Queens Road, Teddington Middlesex, TW11 0LW.

ABSTRACT

During the past year NPL has developed a clock algorithm to combine data from three active hydrogen masers and three commercial caesium clocks. The algorithm was initially based on the NIST AT1 algorithm but has been adapted to meet NPL's own requirements. It has been implemented to process firstly simulated clock data, then actual clock data. First results are presented in this paper.

1. INTRODUCTION

An ensemble time scale combining measurements from a group of clocks has many uses [1]. Firstly its increased integrity is useful when considering that NPL has to maintain the UK's national time scale : UTC(NPL). Many National Metrology Institutes (NMI) nowadays implement an ensemble time scale for generation of UTC and for internal laboratory use. The stability of NPL's existing timescale needs to be improved to achieve best possible comparisons with the new Caesium fountain presently under construction. The evaluation of NPL's new Caesium fountain will require the use of a highly stable reference time scale, which may be provided by a clock algorithm. Short term stability may be provided by a good Active Hydrogen Maser clock. Longer term stability will be provided by TAI. An ensemble time scale may be extremely good in the medium term, enhancing the performance of the individual Hydrogen Masers.

With these aims, the NPL has developed an algorithm to combine three masers and three commercial caesium clocks. The third maser is being purchased. At present the NIST(AT1) algorithm [2] has been used for several reasons. The long operational run of this algorithm at NIST is an excellent proof of the integrity and stability of the algorithm. The algorithm does not treat any single clock as the master. This is useful in the case where the reference clock fails, when a second clock may then automatically replace the bad clock. Finally, the algorithm can be made to operate in real-time or in a post-processed mode. The quality of the results may be improved by post processing.

In this paper, the principle of the algorithm NIST(AT1) and its implementation at NPL are described. Then, the problem of missing data (after a temporary clock disruption) is addressed. Results are presented using both simulated and real data. Finally, some improvements are discussed including use of the statistical tool called the variogram.

2. NIST(AT1) CLOCK ALGORITHM

The principle of the operation of a clock algorithm is to predict the future values in both phase and frequency for each clock, then determine the prediction error from which the weight of each clock is calculated. The ensemble time scale is calculated using these weighted values. In the case of the NIST algorithm, predictions are made of (Clock - ensemble) phase and frequency. For other algorithms a single clock is usually used as a reference.

Firstly an estimate of the phase difference $\hat{x}_i$ between the clock and the ensemble time scale is calculated using the previous value of the phase $x_i(t)$ and the frequency $y_i(t)$:

$$\hat{x}_i(t+\tau) = x_i(t) + y_i(t) \times \tau \,. \quad (1)$$

This result is used to determine the true value of (clock – ensemble) using every measured value x_{ij} between clocks, and a weight w_i estimated for each clock i with equation (10) :

$$x_j(t+\tau) = \sum_{i=1}^{n} w_i(\tau)_t \left[\hat{x}_i(t+\tau) - x_{ij}(t+\tau) \right]. \quad (2)$$

It allows us to calculate an estimate value of the frequency $\hat{y}_i$ for the clock i using the updated phase data (x_i = Clock - ensemble) and its previous value :

$$\hat{y}_i(t+\tau) = \frac{x_i(t+\tau) - x_i(t)}{\tau} \,. \quad (3)$$

At this stage one can improve the estimation of the frequency taking into account a known frequency drift D_i by calculating $\tilde{y}_i$:

$$\tilde{y}_i(t+\tau) = y_i(t) + \tau . D_i \,. \quad (4)$$

We can then calculate a value of the frequency y_i with a filter using a time constant m_i, at time $t+\tau$:

$$y_i(t+\tau) = \frac{1}{m_i+1}\left[\hat{y}_i(t+\tau) + m_i\,\tilde{y}_i(t+\tau)\right]. \quad (5)$$

The time constant of the frequency filter is calculated with the averaging time $\tau_{min,i}$ corresponding to the minimum value of the Allan deviation on a standard Allan deviation representation ($\sigma_y(\tau)$ against τ) of the clock. This gives an estimation of the level of white frequency noise in comparison with the level of frequency random walk of the clock.

$$m_i = \frac{1}{2}\left(-1+\left(\frac{1}{3}+\frac{4\tau_{\min,i}^2}{3\tau_0{}^2}\right)^{1/2}\right) \quad (6)$$

(τ_0 being the algorithm step size).

A tricky part of the process occurs when the relative weights (w_i) of each clock are estimated. This must come from an estimation of its prediction error ε_i. The forward prediction error in the extrapolation of (clock - ensemble) is used to estimate the performance and hence weight of each clock. It is calculated with the difference between the estimate phase of (clock - ensemble) ($\hat{x}_i$) and the true value x_i. If for example the error becomes very high because of an outlying bad point (it could come from a time step), the weight of the clock becomes very small. To avoid this kind of behaviour, the average value of the error estimator is filtered [3].

$$\varepsilon_i(\tau)_{t+\tau} = \left|\hat{x}_i(t+\tau) - x_i(t+\tau)\right| \times K_i(\tau)_t. \quad (7)$$

A correction factor (K_i) is introduced to take into account the correlation between the clock and the ensemble time scale. It is updated with equation (12).

Then an average value of the forward prediction error is filtered using a filter time constant N_τ. Since the noise characteristics of a clock may not be stationary, previous values are deweighted. The value of the filter time constant is chosen to represent the time during which one expects the white frequency noise level to be constant [3].

$$\left\langle \varepsilon_i^2(\tau)\right\rangle_{t+\tau} = \frac{1}{N_\tau+1}\left[\varepsilon_i^2(\tau)_{t+\tau} + N_\tau\left\langle \varepsilon_i^2(\tau)\right\rangle_t\right]. \quad (8)$$

The forward prediction error of the ensemble ε_x is calculated as :

$$\varepsilon_x^2(\tau)_{t+\tau} = \left(\sum_{i=1}^{N}\frac{1}{\varepsilon_i^2(\tau)_{t+\tau}}\right)^{-1} \quad (9)$$

N being the number of clocks.

So it allows us to determine the weight of each clock i :

$$w_i(\tau)_{t+\tau} = \frac{\varepsilon_x^2(\tau)_{t+\tau}}{\left\langle \varepsilon_i^2(\tau)\right\rangle_{t+\tau}} \quad (10)$$

with the condition

$$\sum_{i=1}^{N} w_i(\tau)_{t+\tau} = 1. \quad (11)$$

Then the correction factor is updated using the new value of the weight :

$$K_i(\tau)_{t+\tau} = \frac{1}{(1-w_i(\tau)_{t+\tau})}. \quad (12)$$

This equation order is used for the implementation. The algorithm cycle begins with the determination of the prediction of the phase for each (clock - ensemble) at time $t+\tau$ using previous phase value, the weight and the frequency calculated at time t ant the measurement value x_{ij} recording at time $t+\tau$. Then the (clock - ensemble) frequency at time $t+\tau$ is determined with the new phase value. Finally a forward prediction error is estimated giving a weight for each clock at $t+\tau$, which will be used for the calculation of the phase at time $t+2\tau$.

3. MISSING DATA PROBLEM

A major problem in the algorithm's operation occurs when data from clocks are not recorded during a known time duration. Some clocks are removed from the time scale and are then reinstated.

The solution required depends on the duration of the outage. There is a possibility to interpolate data if only a few data points are missing as long as the weights are not disturbed. If the gap is bigger, one has to estimate a new value of the frequency and phase data using the previous values obtained when the clock was operational. Setting the weight of this missing clock to zero does not disturb the remaining ensemble time scale. It allows us to calculate a new value of x_i in the absence of new data from the missing clock using equation (2). This new value of x_i allows us to determine a new estimation of the frequency y_i using equation (3). Then the returned clock can take part in the ensemble prediction error.

This process has been applied in the chart below. The clock weights and the resulting time scale are not disturbed.

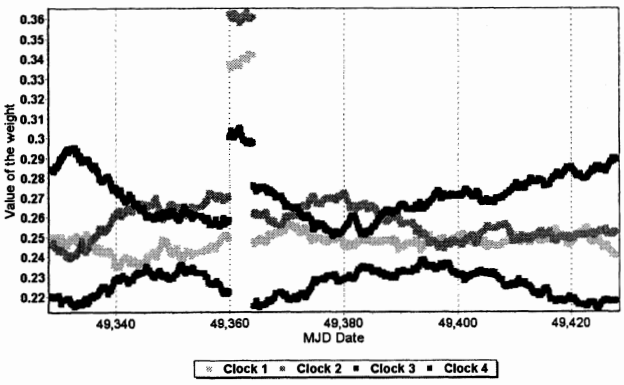


Figure 1 : Weight behaviour of the algorithm operating on four Caesium clocks (data supplied by USNO). Clock 4 is missing between MJD 49360 and 49364. The weight of the other clocks increases to keep a consistent ensemble time scale (through equation (10)).

4. RESULTS WITH SIMULATED AND REAL DATA

The first implementation of this algorithm was undertaken using simulated data. It is straightforward to generate standard noise types (white phase noise and white frequency noise, random walk, flicker frequency and linear frequency drift), and examine the results knowing the initial noise types. In this process, 5 clocks have been simulated using a combination of white phase noise and white frequency noise. The behaviour of the weight of each clock during the time series is shown in Figure 2.

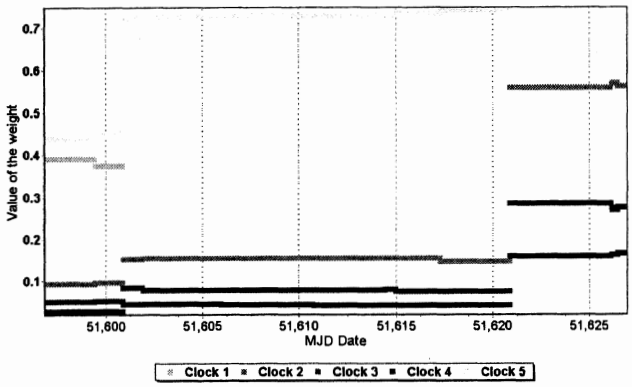


Figure 2 : Weights of five clocks during a 35 day simulation (τ_0 = 3600 s). Clocks 1 and 5 are removed at MJD 51600 and 51620. The weights of all other clocks rapidly readjust to stable values.

A very important consideration is the definition of initial conditions. The choice of initial values will dramatically affect the early stages of the algorithm operation.

Firstly, an initial frequency for each clock must be determined. This could be done by calculating a difference between two phase readings from the clock. The problem is to choose a suitable interval between these two readings. We can simply use the equation :

$$y_i(0) = \frac{x_{ij}(n\tau) - x_{ij}(0)}{n\tau} \qquad (13)$$

where x_{ij} is the measured phase value between the clock i and a clock chosen as a reference for this stage of the process. n must be chosen sufficiently large to characterise the true frequency of the clock.

Filter time constants (m_i, filter for the frequency and N_τ, filter for the prediction error) does not change during the process. So it is important to have some good estimates. For large values of N_τ there could be some transient effect in the behaviour of the prediction error due to initial conditions. But if N_τ is too low, the scatter of the prediction error is larger because the filter is not really efficient. So there is a compromise to find. This is shown in Figure 3.

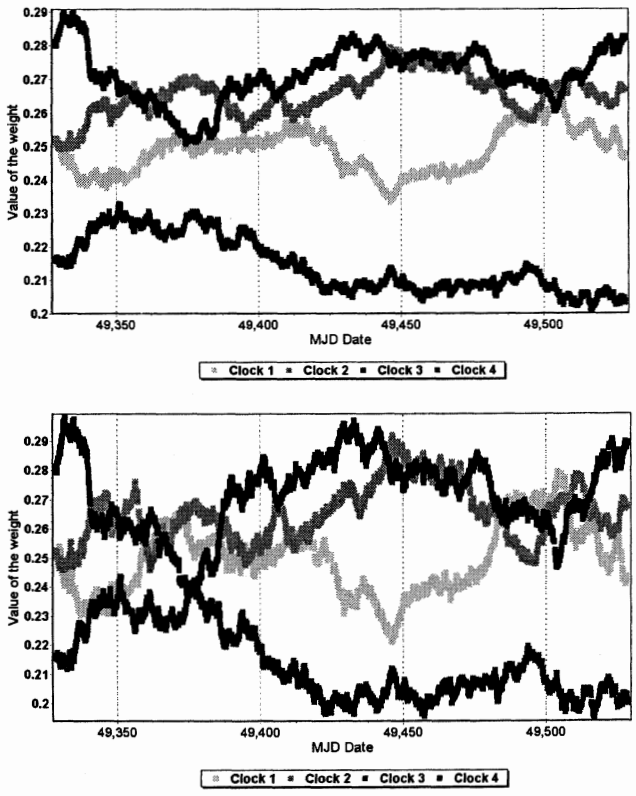


Figure 3 : Processing of real data (clock readings from USNO) showing weight behaviour using 2 different N_τ values (40 days for the top chart and 20 days for the bottom chart). The long term structure is about the same but some short term structures have been added in the bottom chart.

Figure 4 shows the stability of the clocks and the ensemble time scale through the Allan deviation representation. The three-cornered hat method is applied. However the problem of correlation between the ensemble time scale and individual clocks and between individual clocks has not been taken into account.

The short term stability of the ensemble time scale is better than any contributing clock. But in principle one can notice that in the long term, the stability of the ensemble time scale is worse than some of the contributing clocks. Figure 4 indicates that this may be happening. This is due to the fact that the prediction error is optimised for one step ahead averaging time. Prediction errors depend on levels of white frequency noise. But long term stability depends on levels of random walk frequency noise which may be different.

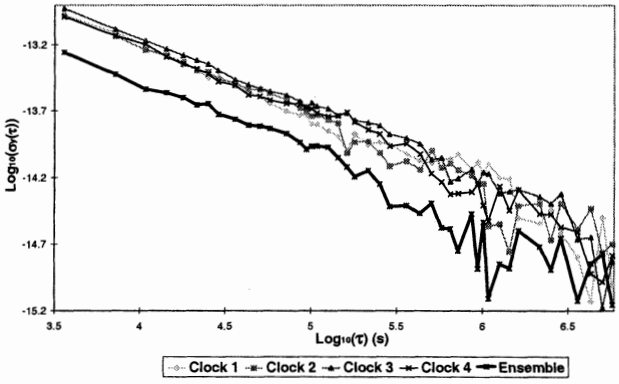


Figure 4 : Result of the processing of 4 Caesium clocks (data supplied by USNO) using the three-cornered hat method. The Allan representation for each clock and the ensemble time scale show a clear improvement in the stability of the ensemble time scale in comparison with any contributing clock.

5. IMPROVEMENTS

There are a few stages of the implementation that could be improved.

There is a possibility of improving the evaluation of several algorithm parameters. A statistical tool called the variogram [4] may allow us to estimate the level of noise of the clock frequency : white noise, random walk and possible linear frequency drift as shown in Figure 5. After applying these estimators in clock data, the time constant filter (as the frequency filter) can be replaced by a more suitable parameter that can change during the process. For example it could be the case of the frequency filter (m_i) that depends on the relative level of white noise and random walk.

The variogram has been applied using simulated frequency data as shown in Figure 5.

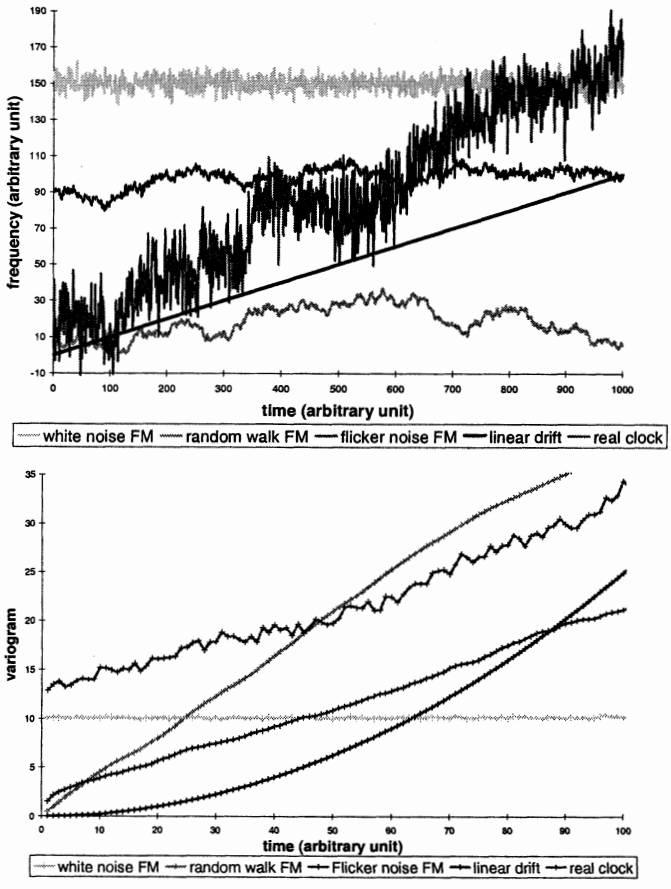


Figure 5 : Variogram applied in some simulated noise frequency data shown in the first plot. The Variogram and the fitted corresponding shape give estimations of the clock noise.

If a clock is very stable, the calculated weight is very high in comparison with the other clock's weight. It could be advantageous to put an upper value to prevent this clock contributing too heavily to the stability of the ensemble time scale. In this case, the continuity and integrity of the time scale is compromised if that clock was to fail.

The time scale can be calculated in real time. It allows NPL to have a new estimation of the ensemble time scale at every epoch chosen. It could be possible to implement the algorithm in a post-process mode. In this case, some better estimation of the clock stability against the ensemble time scale can allow us to improve the real-time processing.

The correlation between the clock and the ensemble time scale is taken into account in the calculation of the bias in the error estimates (K_i). But correlation among clocks may have a major influence on the result as well. That is why some work on clock correlation is being performed at NPL [5].

Finally, it could be possible to extend the laboratory time scale to a national time scale. In this case, some other good clocks may be integrated provided that there is a good time transfer link between the laboratory and the NPL. So after evaluating the stability of the clock and the noise generated from the time transfer link, it's possible to integrate it in the ensemble time scale.

6. CONCLUSION

The work on a ensemble time scale at NPL is in progress. The AT1 algorithm from NIST has been firstly implemented. Tests have been performed using both simulated and real data (supplied by USNO). Finally some improvements such as better estimation of the noise of each clock using the statistical tool, the variogram, will be implemented with the aim to meet NPL's requirements. These requirements are mainly to have a better comparison for the frequency standard (Caesium fountain) under construction at NPL. Moreover, a national time scale combining clocks, other than NPL clocks, could be more reliable for UTC(NPL) users.

ACKNOWLEDGMENTS

The authors wish to thank Peter Harris from CISE at NPL for his help on the variogram. We are also grateful to Demetrios Matsakis from USNO for supplying us some of the clocks data.

REFERENCES

[1] C. Thomas, P. Wolf, P. Tavella, "Time scales", Monographie 94/1 from BIPM, 1994.

[2] M. A. Weiss, D. W. Allan, T. K. Peppler, "A study of the NBS time scale", IEEE Tran. Ins. Meas., **38**, April 1989.

[3] M. A. Weiss, T. P. Weissert, "Sifting through nine years of NIST clock data with TA2", presented at the 7th EFTF, Neuchatel, 16-18 March 1993, pp. 199-210.

[4] N. Cressie, "Fitting variogram models by weighted least squares", Mathematical Geology, Vol. **17**, No. 5, 1985.

[5] J. A. Davis, B. Rueff, "A study of correlated noise within a clock ensemble", EFTF Torino, 2000.

TRUSTED TIME SERVICE FOR THE ITALIAN TIME STAMPING AUTHORITIES

F. Cordara, M. Mascarello, V. Pettiti, G. Vizio and R. Costa
Istituto Elettrotecnico Nazionale "Galileo Ferraris"
Strada delle Cacce, 91 – 10135 Torino, Italy
Tel +39 011 39191, Fax +39 011 346384, E-mail: cordara@tf.ien.it

ABSTRACT

The development of electronic commerce and of the secure electronic document exchange also requires that a trusted date and time information is included as a part of the digital signature of the documents exchanged.
To meet the requirements of the Time Stamping Authorities (TSA) that are going to be established in Italy, the IEN is implementing a trusted time distribution service for the remote certification of these authorities that must have their time sources traced to UTC(IEN) within 1 second.
The approach followed, is based on the accreditation for the time field of each TSA as a secondary centre in the Italian Calibration System (SIT), and the use of secure dial-up connections with the IEN Time and Frequency laboratory for a periodical certification of their date and time.
This paper describes the communication protocol that should be adopted and the equipment setup proposed, both at the IEN and at a TSA side.

1. INTRODUCTION

The great increase in the use of Internet that has been seen in last years, is beginning to change the society. Many every day operations in fact, are now possible using Internet and this is changing the way of living of the people. Already now, it is possible to buy whatever you want in every part of the world staying at home and using a personal computer, and e-mail is substituting the use of fax and telephone. Many people are already working at home exchanging their work through the net and even more so will do it in the future.
Nevertheless this new way of communicating has some problems to be solved yet, in fact the electronic exchange of documents is affected by security problems. As in the case of other ways of document exchange, it has to ensure ***integrity***, ***authenticity, confidentiality*** and ***non repudiation*** of the information. To comply with these requirements, it is essential if one wants to use Internet for e-commerce or electronic exchange of documents having legal value.
A way to grant security, liability and identity is to use the electronic signature within a Public Key Infrastructure (PKI); the advantage of public key cryptography consists in that it allows the confidential transmission of information in open networks where one doesn't know in advance or share secret key information.
Many working groups around the world are studying and proposing systems dedicated to the digital signature starting from the USA where it is already widely used [1]. Also in Europe the use of the digital signature is spreading over, and some studies and standardization projects have are undergoing [2,3].
The electronic signature has to grant the identity of contractors, the integrity and authenticity of the documents but also its legal value, so it is necessary to have a trusted time reference. A PKI is composed by some ***Certification Authorities*** that manage the cryptographic keys and by ***Time Stamping Authorities*** to provide the necessary trusted time reference.

2. PUBLIC KEY INFRASTUCTURES

The use and the exchange of electronic information in the global information society undoubtedly requires services to ensure the security of information and the identities of the participants in electronic events (transactions, administrative acts,...). These services are mainly provided by the PKIs. A PKI is an open public cryptographic system and addresses issues as important as:

- **Integrity of information**
- **Authentication**
- **Confidentiality**
- **Non-repudiation**

It is primarily composed by:

- Certification Authority (**CA**), a trusted third party that issues digital certificates to its subscribers, binding their identities to the key pairs they use to digitally sign electronic communications;

- Time Stamping Authority (**TSA**), the entity that manages the time used by the CA or by the users to sign the documents. It must keep a trusted time related to an internationally accepted reference of time such as UTC;

- Registration Authority (**RA**), an integration of the CA and manages the registration of the users, users identity verification, etc.

In a PKI scheme, every user owns two cryptographic keys, one called ***Public Key*** and the other called ***Private Key***. One has to keep jealously the private key, while the public key must be published so that everyone can read it. For this reason it is necessary a sort of list where all the public keys and the identity of their owners are published. This is made by the CA that ***certificates*** the correspondence between users and public keys.
The important property of the keys used is that they are asymmetric, in the sense that it is only possible to decode the message encrypted with the public key using the corresponding private key and vice-versa, but it is not possible to extract the private key (Fig.1).

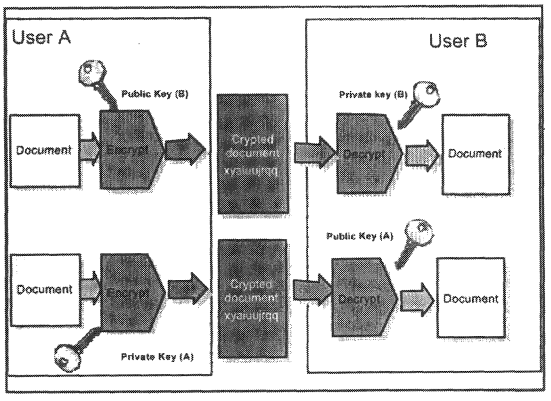


Figure 1 - Asymmetric keys criptography.

In the context of digital signatures (Fig. 2) a signer uses a ***"hash"*** function to create a compressed form of the message to be sent. This ***"message digest"*** is unique to that message and can be used subsequently to verify the authenticity of the document once received. Before sending the document electronically, the signer applies the private key to the message digest, thereby encrypting it and creating a secure digital signature. The document may then be sent, (encrypted with the receiver's public key) along with the digital signature.

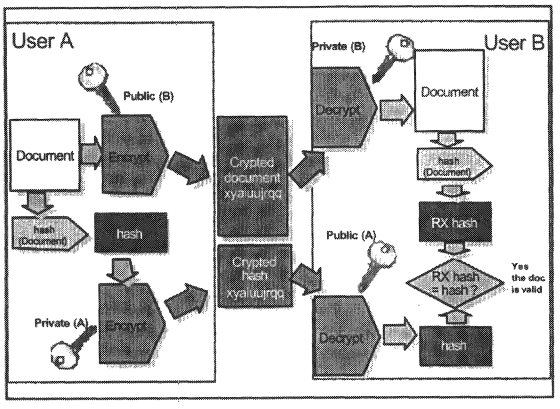


Figure 2 - Digital signature operational scheme.

Upon receipt, the digital signature can be decrypted with the signer's public key and the message digest can be used to verify the contents of the electronic document.
In this scenario, the only need is to assure that the document arrives to destination in a reliable way, but in the case of commercial transactions and administrative documents, it is necessary to add also a trusted time stamp to the electronic signature. To solve this problem the TSA has the duty to generate a trusted time signature and this time tag must be traceable to a legal reference of time.

3. IEN PROPOSAL FOR A TRUSTED TIME SERVICE

In Italy it's going to be developed a method to exchange electronic documents with the Public Administration, and a recent regulation [4] established the Italian CAs and stated that the TSA time reference must be within 1s from UTC(IEN), that is traceable to the international time scale UTC.
IEN already has some time dissemination services like the telephone time code (CTD), a radio dissemination service (RAI – SRC), the Internet Network Time Protocol (NTP) stratum 0 server, but none of them is compliant to the requirements of a trusted time reference for TSAs.
Therefore it has been studied a new dissemination service, to provide a trusted time reference to the TSAs and for the future activities that will use electronic exchange of information.
The approach proposed by IEN assumes that a TSA will have some sources of date and time synchronized e.g. by GPS, DCF77, RAI/SRC, IEN/CTD, to have the necessary redundancy, and will call some times a day the IEN trusted time service by a secure dial-up connection, to have a certification of its time information traceable to UTC as shown in Fig. 3.

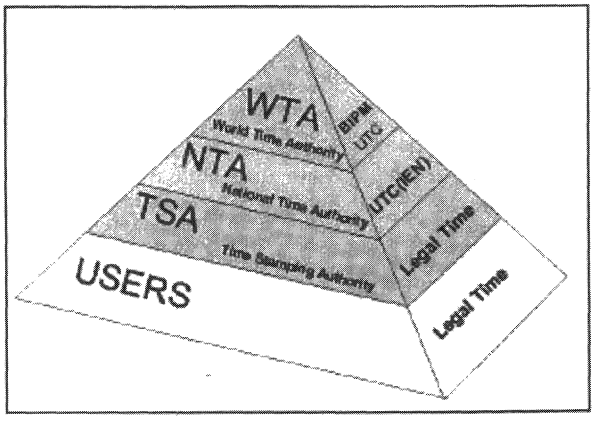


Figure 3 - Traceability of TSA time reference.

Operating according to this protocol, the TSA can also become a centre accreditated for time dissemination in the frame of the Italian Calibration Service (SIT), and

consequently have the international recognition by the European cooperation for Accreditation (EA). With the solution proposed, IEN will be also able to measure the rate of TSAs clocks and to provide a certification of their time error.
An additional feature can also permit the IEN to assess the compliance of the CA trusted time mark; this could be accomplished asking the TSA the signing of a document carrying the IEN time information, as can be done by regular user over the Internet connection.
In this scheme, the TSAs will have to implement a software similar to the NTP daemon to manage all the possible sources of time with suitable drivers then, using an additional software they will have to call the IEN Trusted Time Service some times per day to get the certifications of their time references. The same should be done in case of clocks failures The IEN Trusted Time Server (TTS) will be made of two identical systems as shown in Fig. 4, one serving as backup and will be connected to the UTC(IEN) via CTD time generators; it will be reachable by dedicated lines and cellular phone for redundancy.

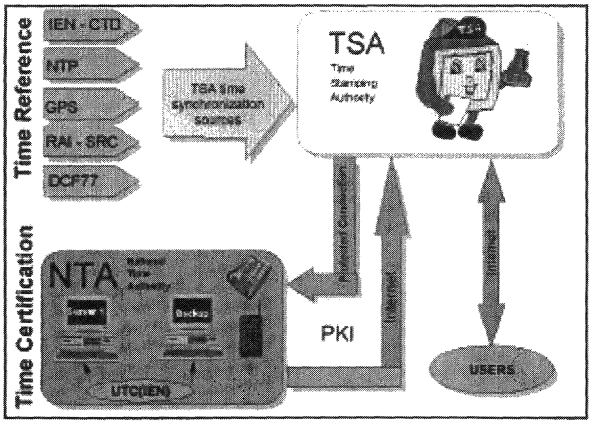


Figure 4 - IEN proposed Trusted Time Service.

The protocol of communication between a TSA and the IEN TTS is under development and will be sent to the AIPA (Authority for the Information exchange in the Public Administration) for the approval. In what follows is presented a draft version of the protocol.
The IEN trusted time server will use dial-up connections to overcome two main problems: delays estimation and security.
The first problem is due to the fact that, using Internet connection, it is very difficult to estimate link delays because of the asymmetry of network routing; any data packet in fact can follow different routes to reach its destination;
The second is even most relevant, because using Internet many hackers attacks are possible and there is a chance that some of them could also shut down the service.

Telephone connections have instead a higher security and, except for intercontinental calls, have substantially symmetrical transmission delays. It is therefore possible to evaluate the one way trip delay dividing by two the round trip delay.
The proposed data string for the communication of TSA date information to IEN, is showed in Fig. 5 and it is based on the format of the European Telephone Time Code as reported in the ITU-R TF.583.4 Recommendation

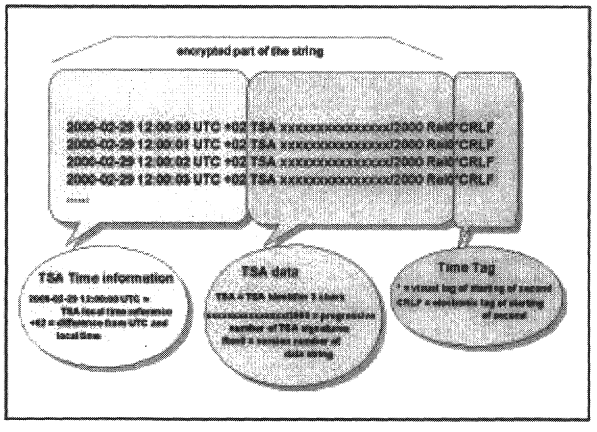


Figure 5 - Proposed TSA data string.

The communication protocol flowchart is showed in Fig. 6. The TSA shall call the IEN TTS at least once a day and after the modem connection the time certification begins following these steps:

A: Identity Verification:

- TSA sends its certificate to IEN,
- IEN sends its certificate to TSA,
- TSA generates a session symmetric key (eg. 3DES), signs it with its private key and encrypt it with the IEN public key,
- IEN receives and decrypts the session key and sends an acknowledge to the TSA.

- **B: Line delays calibration:**

- TSA goes in LOOP mode,
- IEN begins to send a special character to the TSA that sends it back immediately,
- IEN records the RTT (Round Trip Time) and after some measures (eg. 10) stops the process,

- **C: TSA time verification:**

- TSA begins to send its time according to the data string format,
- IEN makes a series of measurements (eg. 10) and computes every second the time difference with the resolution of the server clock (~50 ms) and records them,
- If UTC(IEN) – Time(TSA) is less than 1 second the IEN confirms the TSA time and certificates its validity.

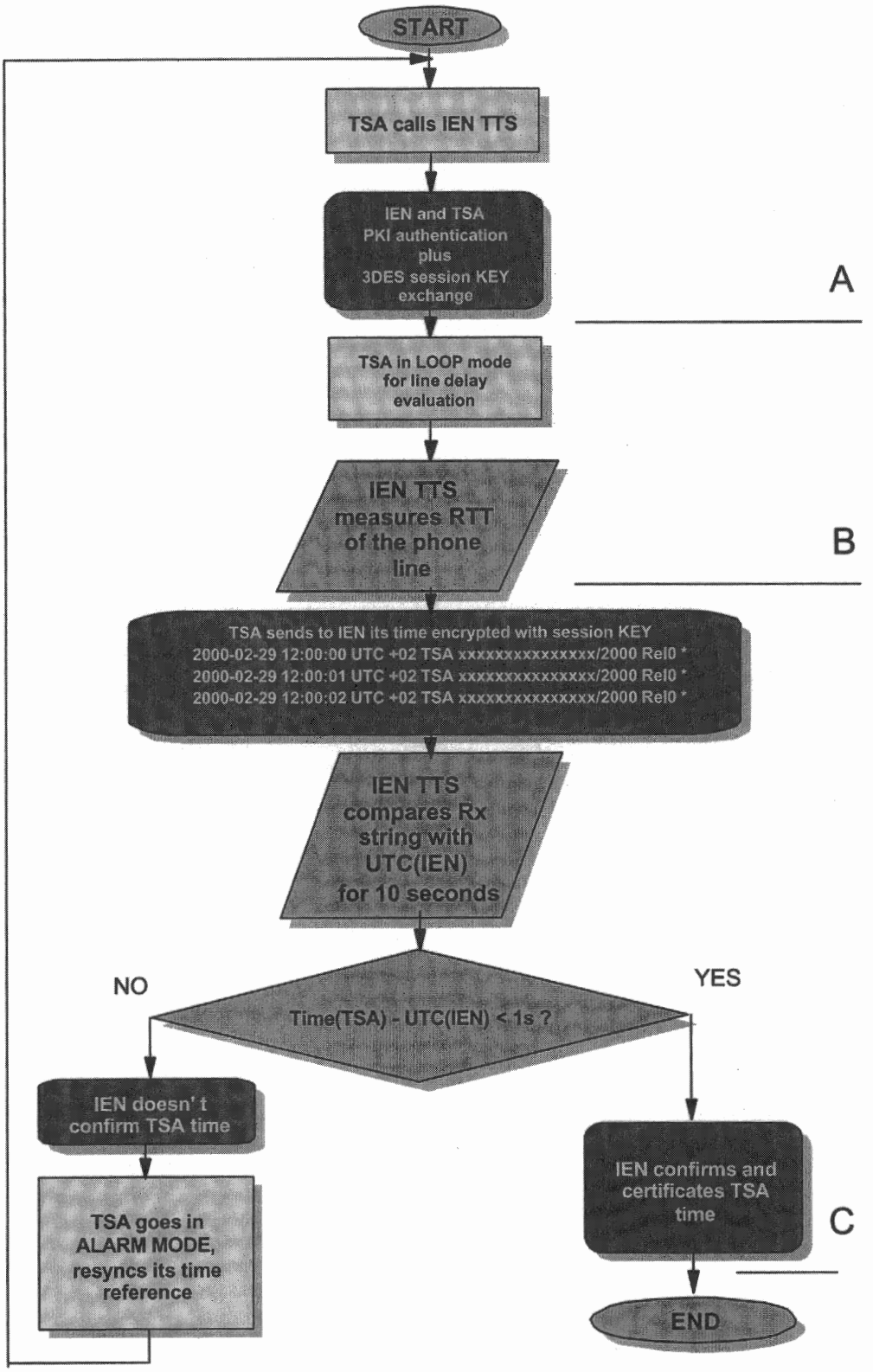


Figure 6 - TSA - IEN TTS communication protocol.

The data string used for the transmission of the time and date is derived from the CTD format and the data exchange will be made using a connection as if it was like an ASCII terminal program.
The proposed data string contains, apart from the date and UTC time, also the difference between the legal local time and UTC because in most countries the legal time differs from an integer number of hours from UTC. The previous information is followed by:

- an identifier of the TSA that is transmitting the string,
- twenty characters to indicate the actual number of time stamp performed,
- four characters where it is indicated the version of the data string for future releases.

At the end, like in the CTD code a character "*" is present followed by a carriage return (CR) and line feed (LF) to mark the starting of the time information indicated in the data string.
In the data exchange between TSA and IEN TTS, the first 57 characters will have to be encrypted with the session KEY while <*,CR,LF> characters will be transmitted in clear form. The encryption in fact adds bytes to the information, so the time tag must be in clear form to avoid added delays.
The total lenght of the data string is 60 characters, that could be transmitted in clear form using a 1200 bit/s modem (150 byte/s) but, using the encryption, a standard 9600 bit/s modem connection (1200 bytes/s) will be more suitable.
A possible cryptographic algorithm already widely used could be the 3DES one described in [6].

Some preliminary tests have been made at IEN, to check the possibility of using telephone lines and the computer time base to evaluate the telephone lines delay. In Fig.7 are presented the results of a measurement session of a 50km phone call round trip time (RTT) using a software developed at IEN. The two

lines represent the RTT measured with the internal PC

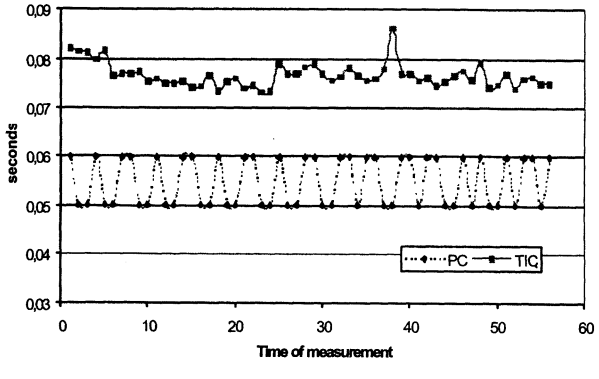


Figure 7 - Round Trip Time (50 km link).

timer and using an external counter. It can be seen that the bottom line shows the quantization effect of the limited resolution on the PC internal counter (~50ms). As regards the difference between the means values of the curves it is due to additional hardware and software delays of the two measurement set-ups.

4. CONCLUSIONS

A proposal for the trusted time dissemination service (TTS) that will serve the Italian companies involved in digital time stamping and e-commerce has been studied at IEN. In the next months it is planned to implement the solutions presented and to test the new service in the field.

5. REFERENCES

[1] A. Gidari, J.P. Morgan, P. Coie *"Survey of state electronic & digital signature legislative initiatives"* ILPF (Internet Law and Policy Forum) 12 September 1997.

[2] *"European Electronic Signature Standardization Initiative (EESSI)"* Final Report of he EESSI Expert Team 20 July 1999.

[3] Draft ETSI *"Electronic signature standardization for business transaction"* ES 201 733 v1.1.3 (1999-09).

[4] DPCM 8/2/1999 *"Regole Tecniche per la formazione, la trasmissione, la conservazione, la duplicazione, la riproduzione e la validazione , anche temporale, dei documenti informatici ai sensi dell'art. 3 comma 1, del decreto del Presidente della Repubblica 10 Novembre 1997 n.513"*.

[5] J. Levine *"Authenticating Time and Frequency signals"* Joint Meeting EFTF 13-16 April 1999 pp. 304-308.

[6] U.S. Department of Commerce / National Institute of Standards and Technology *"DATA ENCRYPTION STANDARD (DES)"* FIPS PUB 46-1999 October 25.

OPTICAL TIME TRANSFER USING AJISAI MIRRORS: STATUS PROGRESS

Leonardi Mauro

Alenia Spazio – space division, via Saccomuro, 24 - 00133 Roma, Italy
Tel +39 06 41512587; Fax +39 06 41294035 ; e-mail: m.leonardi@roma.alespazio.it

ABSTRACT

Present time transfer systems are based on microwave signals. The most popular remote clock synchronisation method is currently based on the Global Positioning System constellation (*GPS*). Its "*common view*" configuration guarantees accuracy of a few nanoseconds. Optical carriers could increase this timing accuracy for a number of reasons, such as the bandwidth or the nearly absence of ionospheric effects. In *1989* an ESA promoted experiment using a special geostationary satellite showed the potentiality of laser timing [1], [2]. Nowadays that spacecraft is no more available, but a Japanese one, named *Ajisai*, could offer the opportunity to use optical signals.

This satellite, of spherical shape, orbiting at *1500 km* of altitude, is fitted by *318* mirrors plus *120* arrays of retroreflectors. With a suitable laser emission scheme and with the detection of the light reflections from the mirrors, it is foreseen to realise both a *one* or a *two-way* optical link between remote ground laboratories.

This note is a progress report. The possibility to perform a time synchronisation using this peculiar satellite is examined. Preliminary steps are the redaction of precise satellite ephemeris, including also *attitude information*, *spin rate*, *spin axis orientation*, *mirror identification*, etc. Using the Sun light reflections as input signals, some data were collected and will be presented. *Ajisai common visibility* and *synchronisation windows* are determined considering some European laser ranging stations. Meanwhile, the link budget is evaluated for the same stations.

1. INTRODUCTION

In order to improve current accuracy, time and frequency community is investigating new remote clock comparison systems. Although the microwave domain is still of interest, in the last decades also the optical region has been analysis. With the *TWSTFT* (Two-Way Satellite Time and Frequency Transfer) microwave system, for instance, which is under test since some years, better values of accuracy respect to the GPS *common view* configuration can be reached [3], [4].

Synchronisation techniques in the visible range, on the other hand, seem to be the best candidate to guarantee measurement accuracy of sub nanosecond level with the smallest theoretical effort. The *T2L2* (Time Transfer by Laser Link) project is probably the most popular proposal currently under investigation [5], [6]. Its link scenario is based on an optical receiver/transmitter onboard a spacecraft, with the second station based on the Earth.

The hereafter considerations, however, refer to the possibility of performing laser time transfer between two remote ground clocks using the mirrors onboard the geodetic satellite *Ajisai* [7]. A complete passive *two-way* synchronisation can therefore be realised (Figure 1).

The optical solution, besides its high potentialities, presents also drawbacks and disadvantages that can extremely limit its applicability to synchronisation purposes, especially if clock comparisons are frequently and periodically needed. Furthermore, in case of optical applications performed inside the atmosphere, weather conditions, scattering, turbulence and refraction effects should also be taken into account.

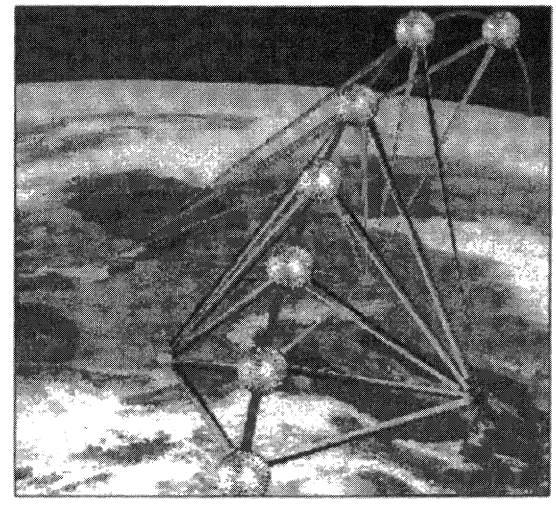


Figure 1. *A possible European two-way link scenario via a passive satellite (optical domain).*

In the specific case of the considered scenario, moreover, the links are feasible only when the mirrors are correctly oriented. Connections, in fact, can fail due to the spin stabilisation of the satellite. The consequence is that both the satellite orbit and attitude must be accurately investigated and predicted in order to determine exactly when and where the laser signals must be fired.

To analyse the synchronisation repeatability of the chosen technique, it is also necessary to examine the various passages of the satellite over a same site, in order to determine the *common visibility windows* (i.e. the time intervals during which two or more remote stations can view contemporaneously the satellite). During a single common visibility window, another subset of time intervals must be considered: the *synchronisation windows* (i.e. the intervals during which the mirrors are correctly oriented to establish the link).

Finally several comparisons, also during the same satellite passage, are mandatory to perform a statistical analysis to infer stability and uncertainty of the synchronisation link.

(Figure 2). The sphere is *1 m* in radii and is completely passive. Its surface is fitted only by *CCR* (corner cube retroreflectors) and mirrors having a curvature that represents a larger mirror sphere of *9 m* in radii [8]. The orbital plane is inclined of *~50°*, whilst the orbit shape is almost circular ($e_{average}$*=0.0011364*). The satellite revolution period is $T_{average}$*=115,716 min* (the orbit is *~1500 km* high).

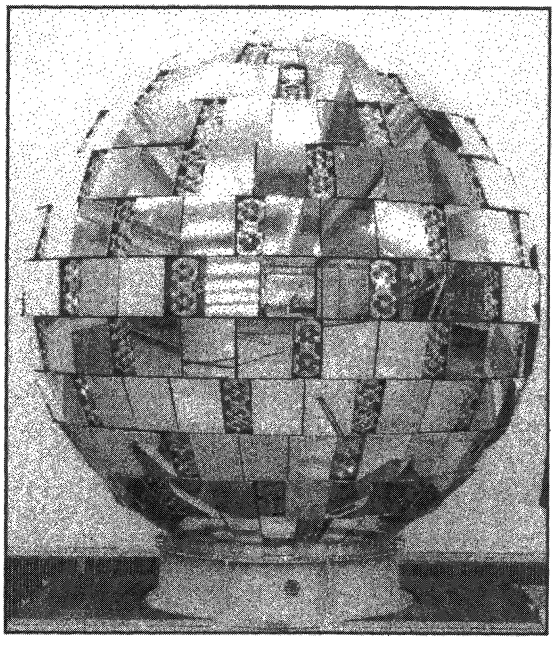


Figure 2. *EGP* (Experimental Geodetic Payload) satellite (*Ajisai* is its nickname).

To perform the proposed optical link, a prediction of the satellite motion is needed. Both gravitational and non gravitational effects must be considered and kept into account [9], [10]. In the case of *Ajisai*, the Earth oblateness is the dominant perturbation source and, for secular effects on *Ajisai* orbital plane, it is sufficient to consider only the J_2 coefficient. For this satellite, the main Earth gravitational perturbation causes the following effects [11]:

- *RAAN* (Right Ascension of the Ascending Node) precession (*-3.074°/solar day*).
- Argument of the perigee motion (*+2.547°/solar day*).

For accurate orbit prediction, it is obvious that J_2 cannot be the only factor to be considered. The basic step, however, in the satellite position determination is the availability of precise and update *orbital elements*. For geodetic satellites (as *Ajisai*) the *SLR* (Satellite Laser Ranging) world-wide tracking campaign permits to have very accurate *orbital elements* that can be used in the orbit prediction software. Very short term prediction, therefore, are sufficiently precise even in the case that only the main perturbation effects are taken into account.

A previous verification of the estimated *Ajisai* position was performed using Sun light reflections. Figure 3 shows the photo obtained at the astronomical observatory of Pino Torinese (I) where two adjacent sun flashes were detected (see arrows in figure). The stars are circular in shape due to the Earth motion compensation of the telescope. The exposition time was *1 min*, the telescope field of view was *10x10 arcsec*, and the CCD was composed by a matrix of *1152x1296 pixels*.

Because the position of the satellite and its relative velocity respect to the observer were known, by counting the pixels impressed by the two sun reflections it was possible to roughly estimate the time interval of the two flashes. Also the time separation between them was determined in a similar way. For the considered photo, it was obtained *11,7 ms* and *9,98 ms* for the flashes, whilst the measured separation time was *274,43 ms* [11].

To link successfully two remote stations, however, also the satellite attitude perturbation must be modelled. The long term effects are the *Ajisai* spin axis precession, its nutation and the spin rate decrement due to the *Eddy* current generation. Short term effects are also present and are currently under investigation. The attitude prediction is essential to know when the mirrors reach the right orientation to perform the link. Sun light can be useful also in this analysis.

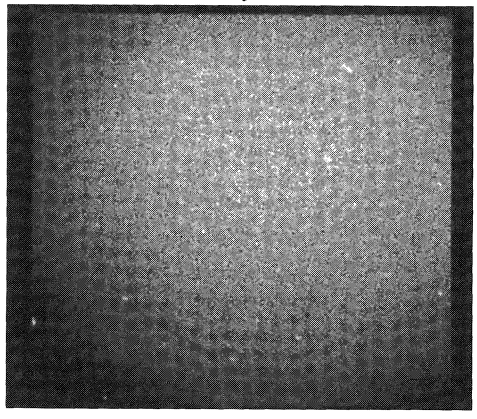


Figure 3. Sun reflections from the Ajisai mirrors detected at the Astronomical Observatory of Pino Torinese (I).

Detection of the Sun reflections from the *Ajisai* mirrors were executed. At each received flash was associated the detection time. In such a way the time interval between flashes is easily computed. If the time between two adjacent flashes generated by the same mirror is considered, the *Ajisai* rotation period can be evaluated. Figure 4 reports the various estimations obtained during a single *Ajisai* passage over the laser ranging station located in Cagliari (I). Note that the reported data in this figure represent the measured values without keeping into account the relative motion of *Ajisai* respect to the observer.

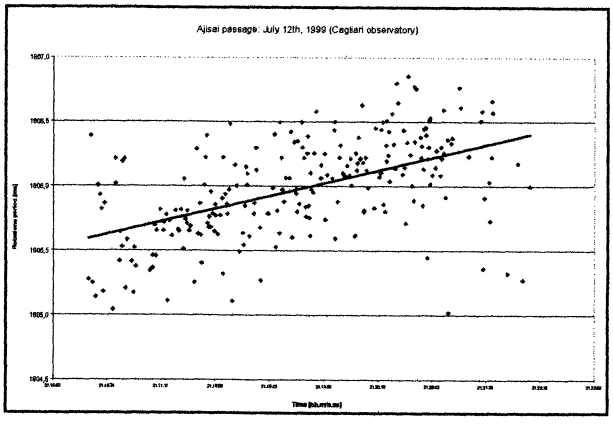


Figure 4. Raw data measured in Cagliari July 12th 1999.

Considering this effect, it was possible to determine that the *Ajisai* spin is clockwise for a celestial north pole observer and that the detected average rotation period was *1806.32 ms*.

Because the time separation between adjacent reflections depends on the *Ajisai* spin axis orientation, it is also possible to estimate such disposition. It is sufficient to compare the measured values with the time interval evaluations that arise from geometrical consideration. The satellite axis parallel to that of the Earth was the configuration that gave the lowest residuals.

The measured rotation period of *1806.32 ms* is compliant with the *Ajisai* spin rate secular decrement of *–0.00145 rotation per min/day* proposed by some Japanese researches [12], [13]. Figure 5 shows this secular decrement (dash line) and the spin estimations proposed by Kubo and Otsubo, together with the obtained value of *1806.32 ms* (*33.21 rotations per minute*).

Note that the detected short term decrement during the six minutes of the Cagliari acquisition is much higher: *-1.2535 rotation per min/day*. Also this value is compliant with the instant values measured by the Japanese colleagues [12].

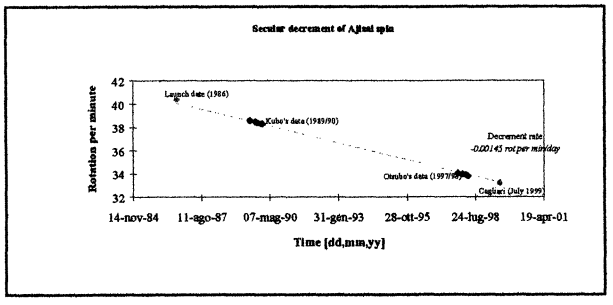


Figure 5. Ajisai spin rate secular decrement

A possible explanation of such a difference between the secular and the instant decrement can be found in the thermal transients suffered by *Ajisai* entering in (or exing out from) the Earth shadow (hyphotesis under investigation).

3. COMMON VISIBILITIES

Due to the Earth rotation, the *Ajisai* ground coverage areas are not constant in time. As a consequence also the common visibilities between two fixed stations change in time. Fortunately a particular time interval exists for each orbiting satellite, after which the coverage area repeat themselves. This satellite *periodicity* represents the time necessary to a chosen passage to repeat identically, i.e. the time interval necessary to the satellite ground tracks projected on the Earth surface, to return and to pass exactly over the previous tracks. In the case of *Ajisai*, it was determined a periodicity of [11]:

$$15^{solar\ day}\ 19^{h}\ 43^{m}\ 14^{s}.$$

This value keeps into account the J_2 Earth perturbation effect. During this period a number of *Ajisai* passages can be used to realise the synchronisation. Considering some specific European *SLR* stations it can be determined that from a single site there are from *68* to *82* passages during which *Ajisai* can be seen with a maximum elevation angle greater than *30°* [11]. Coupling two by two the considered sites, always in the obtained repetition time, there are from *65* to *75* common visibilities during which the satellite can be seen from the two sites with an elevation angle greater than *30°* contemporaneously [11].

These figures represent the useful *Ajisai* passages that were determined disregarding the distinction between night time and day. Although during daylight the acquisition of the return pulse is more critical than at night, current photodetectors are able to detect such signals. Furthermore, a selected passage will be detected during daylight or at night depending on the chosen reference orbit. However, remind that daylight acquisitions request for very good weather conditions and only the state of the art *SLR* stations can justify this assumption.

4. SYNCHRONISATION WINDOWS

The *Ajisai* common visibilities give the period during which a potential time transfer can be performed once the two remote stations are selected. To determine exactly when the synchronisation can be realised, it is necessary to find out the time windows during which the mirrors are correctly oriented.

Depending on which row of mirrors over *Ajisai* is considered, their width changes. The largest are those at the border with the satellite equator (*191.355 mm*). This value represent the average value between the two basis of the trapezoid which describes the mirror. The smallest ones in width are those covering the satellite polar zones (*80 mm*). Taking into account the average rotation period of *1806.32 ms* a solar flash (or consequently a synchronisation window) can last from *5.118 ms* to *12.222 ms* [11]. Note that the flash durations obtained with the pixel analysis of the Figure 3 are compliant with these values.

The obtained synchronisation intervals are simply average values. Besides the dependence from the satellite passage, which selects the mirror useful to the link, it is necessary to stress again that the *Ajisai* rotation period changes continuously due to the attitude perturbations. In the case of the detected passage, for instance, the time windows change from *12.2221 ms* (beginning of passage) to *12.2242 ms* (end of passage). However current laser signals are *100 ps* long, the oscillator delays are below the tens of microsecond and the additional atmospheric delays are of some tens of nanoseconds, therefore the windows duration oscillation is totally negligible.

The basic aspect concerning the optical link, however, is the a priori knowledge of when a mirror will assume the right orientation to permit the link. Because each single mirror position can be described respect to an ideal reference meridian, it is possible to associate a rotation phase to the mirrors. In such a way, the future position prediction can be performed once an initial disposition has been detected. In the case of the Cagliari acquisition, the instant, at which the first sun flash was detected, was taken as reference datum in order to predict the subsequent flashes. For the mirrors belongings at the same row, the prediction was acceptable (only few milliseconds of discrepancy). Although these results are encouraging, this discrepancy becomes unacceptable when the prediction is extended to the following *Ajisai* passage. Furthermore, problems arise when predicting the position of mirrors in different rows (discrepancy of tens of millisecond).

Such discrepancies are still under investigation and the main sources are probably connected to the attitude perturbation, the thermal transient and the not yet resolved displacement between the two rotation axes of the Earth and of *Ajisai*. Future acquisitions are scheduled in accordance with the satellite passages in order to deeply investigate these aspects.

5. LINK BUDGET

The considered link scenario implies the necessity to localise an object in the sky with angular dimension of less than a microradiant, in a telescope field of view of less than a hundred of microradiants. Furthermore the reflected signal is foreseen to be very weak although the transmitted powers are high. A link budget estimation is therefore essential to understand if the proposed link is feasible.

Besides the free space attenuation, the optical pulses during the forward and return propagations are perturbed by the atmosphere. Not only weather condition must be kept into account, also effects such ad *scattering* and *turbulence* degrade the transmitted beam divergence [11]. From the synchronisation point of view, the *refraction* effect must also be considered either by modelling the atmosphere (for instance with the *Marini & Murray* model) or by measuring the effects with two-colour acquisitions.

Furthermore, a number of different parameters must be considered in the budget estimation. Each one presents a high impact level on the received signal. The hereafter results must, therefore, be considered as examples of possible situations, but they do not represent a general situation. A case by case computation is mandatory. The received photon number at photodetector level (i.e. after the optical receiver but previous of the signal conversion), is given by [14]:

$$n_{rec,ph} = \left[\eta_T \left(E_T \frac{\lambda}{hc}\right) G_T\right] \cdot \left[\left(\frac{1}{4\pi R_1^2}\right) T_{A,1} T_{c,1}\right] \cdot \sigma_s \cdot \left[\left(\frac{1}{4\pi R_2^2}\right) T_{A,2} T_{c,2}\right] \cdot \eta_R A_R$$

The expression $4\pi R_i^2$ represents the sphere surface on which the optical power distributes at a distance R_i from an isotropic transmitter. The gain given by the telescope in focusing the beam is described by the factor G_T as in the microwave domain applications. The transmission efficiency η_T describes the signal attenuation inside the transmitter. The factors $T_{A,1}$ and $T_{c,1}$ represent respectively the *one-way* atmosphere transmissivity and the *one-way* cirrus trasmissivity from one station to the satellite. Analogously, factors $T_{A,2}$ and $T_{c,2}$ represent the atmosphere effects on the link from the satellite towards the second station. The transmitted energy is represented by E_T, the signal wavelength is λ, the Planck constant is h, the velocity of light is c, σ_s is the mirror cross section, η_R the receiver efficiency, A_R the collecting area of the receiver [14]. The received photon number, therefore, must be always associated to the considered weather conditions, to the apparatus parameters and to the sites position.
Also the used detector (PMT, APD, SPAD) play an important rule [15]. Not all the received photons that reach the photodetector, in fact, are really detected. The *quantum efficiency* of the receiver must also be taken into account. Depending on the detector, current quantum efficiencies belong to the range (*0.08 - 0.4*).

To estimate the received signal in a realistic situation two scenarios (within the same *Ajisai* passage) were considered. As reference *SLR* sites were chosen *Grasse* (F) and *San Fernando* (S). The two different satellite positions were:

- *Rising situation for Grasse (elevation angle=30°) and corresponding elevation angle of 36° for San Fernando (according to the selected Ajisai passage).*
- *Zenith situation for Grasse (elevation angle=86°) and corresponding elevation angle of 40° for San Fernando (according to the same selected Ajisai passage of the previous case).*

Note that, if another satellite passage is considered, the elevation angles related to *San Fernando* can change significantly even in the case that those for *Grasse* remains identical.

Nowadays *SLR* apparatus can be described as follows:

- *beam divergency* — *5 arcsec*
- *Transmission efficiency* — *0.6*
- *Receiver eff. (with interferometric filter)* — *0.3*
- *Transmitted energy* — *80 mJ*
- *Nominal wavelength* — *532 nm*
- *Transmitted photons per pulse (average)* — $3\ 10^{17}$
- *Telescope radius* — *30 cm*
- *Transmitted divergency (also the case 10 arcsec is considered)* — *5 arcsec*

The *Ajisai* mirrors are described by:

- *Mirror riflettivity* — *0.85*
- *Mirror cross section* — $286\ m^2$

To describe the atmosphere, the case $T_A=T_c=0.8$ for a zenith direction (elevation angle *90°*) was chosen. The corresponding values for the considered elevations were determined and reported in [11]. The second case of $T_A=T_c=0.6$ (zenith direction) was considered to emphasise the dependence from the atmospheric conditions. The situation $T_A=0.8$ represents a visibility at sea level of about *20 km*. $T_A=0.6$, instead, describes a visibility of *8 km*. The value $T_c=0.8$ corresponds to the attenuation caused by an average cirrus layer of *1.34 km* high.

By substituting these values in the previous formula, the expected received photons in the considered scenarios are reported in Table 1. As a consequence, when weather conditions are *acceptable*, current photodetectors are able to detect such signals. Remind that Table 1 values refer to a particular configuration. Real return signals can be even stronger due to better atmospheric conditions or different satellite positions or different stations with better performances than the one considered in this example.

Table 1 Average number of photon expected for the considered scenario.

Considered Ajisai positions	$T_A=T_c=0.8$ θ=5 arcsec	$T_A=T_c=0.8$ θ=10 arcsec	$T_A=T_c=0.6$ θ=5 arcsec
Satellite at zenith	26	6.5	4.6
Rising satellite	3	0.8	0.15

Analogous energy budget estimations were performed also for other European *SLR* stations. The obtain results were of the same order of magnitude of those reported in Table 1 and are not explicitly reported in the present note.

6. CONCLUSION

This progress report illustrates the current situation concerning the possibility to perform a remote *two-way* synchronisation via a passive satellite. The *Ajisai* attitude was determined considering the sun light reflections detected in *Torino* (*I*) and *Cagliari* (*I*). The satellite spin rate, its spin axis orientation and its rotation were estimated and verified. The repetition period, the common visibilities and the synchronisation windows were investigated. An energy budget was also calculated.

orientation and its rotation were estimated and verified. The repetition period, the common visibilities and the synchronisation windows were investigated. An energy budget was also calculated.

The average rotation period obtained by the analysis of sun reflections from the onboard mirrors was *1806.32 ms* (i.e. *33.21 rotations per minute*). This value is compliant with the spin secular decrement of *–0.00145 rotations per min/day* detected by some Japanese researchers [12], [13].

The obtained short term perturbation, instead, was *–1.2535 rotations per min/day*. Also this value is compliant with those obtained by the Japanese researchers during their acquisitions in *1989* and *1997* [12]. The corresponding increment of the rotation period, in the *6 min* of the observation, was *+0.284 ms* (i.e. *68.2 ms/day*), with a relative variation of $\Delta T/t=0.788 \cdot 10^{-6}$. A possible explanation concerning the difference between the secular and the short term perturbations, although not yet confirmed, is connected to thermal transients suffered by the satellite during the exit (or the entrance) from (or in) the Earth shadow.

The analysis and the corrections of the raw data also permitted to estimate that the *Ajisai* spin axis is parallel to that of the Earth and the rotation is clockwise for a north celestial pole observer.

The satellite attitude determination permitted the synchronisation windows evaluation. The obtained average values belong to the range (*6 ms*, *12 ms*). Analogous values were obtained analysing the photos realised at the Pino Torinese astronomical observatory (Italy).

Investigating the main *Ajisai* orbit perturbations, a repetition period of $15^{day}\ 19^{h}\ 43^{m}\ 14^{s}$ was obtained. During this period and according to the selected stations there are several *Ajisai* passages useful for a remote synchronisation. For the considered European *SLR* station there are from *65* to *75* common visibilities within the repetition period.

The expected number of photons are detectable, although they depend on a number of parameter. Considering state of the art systems and good weather conditions, it is foreseen to receive even a few hundreds of photons for each transmitted pulse.

From the synchronisation point of view, the accuracy obtained in the mirror position prediction is still not sufficient. A detailed investigation of the attitude determination is under progress. Currently the single mirror identification permitted the determination of a rotation phase, essential step in the attitude prediction. A small displacement between the Earth rotation axis orientation and the Ajisai one, is probably the main cause of the prediction errors. Further acquisition are planned during next months in accordance with the Ajisai passages.

In conclusion The optical link via Ajisai mirrors seems really feasible.

7. REFERENCES

[1] J. Gaignebet, S. Leschiutta, *"Time sincronisation using laser techniques"*.

[2] F. Beaumont, J. Gaignebet, W. Klepzinsky, S. Leschiutta, C. Veillet, *"LASSO, a satellite laser ranging and synchronisation experiment: Preliminary results"*.

[3] P. Tavella, M. Leonardi, *"The analysis of irregularly sampled TWSTT data"*, technical report n° 556, IEN "G. Ferraris" – Torino, November 1998

[4] P. Tavella, M. Leonardi, *"Noise characterisation of irregularly spaced data"*, in Proc. 12th European Frequency Time Forum, Warsaw, Poland, March 1998, pp. 209-214.

[5] E. Samain, P. Fridelance, *"The time transfer by laser link experiment on MIR (T2L2)"*, Workshop on New Generation of Space Clocks, Neuchatel, March 1997.

[6] E. Samain, P. Fridelance, *"Time transfer by laser link (T2L2) experiment on MIR"*, Metrologia, 1998, 35.

[7] H. Kunimori, F. Takahashi, A. Yamamoto, "*A system design for laser time synchronisation via geodetic satellite Ajisai*".

[8] M. Sasaki, H. Hashimoto "*Launch and observation program of the experimental geodetic satellite of Japan*", IEEE trans. on Geoscience and Remote Sensing, vol. GE-25, no. 5, Sept. 1987.

[9] G. Seeber, "*Satellite geodesy*", Walter de Gruyter, 1993.

[10] A. Milani, A. Nobili, P. Farinella, "*Non-gravitational perturbations and satellite geodesy*", Adam Hilger, 1987.

[11] M. Leonardi, "*Two-colour laser synchonisations*", Ph.D. thesis, Politecnico di Torino, February 2000.

[12] T. Otsubo, J. Amagai, H. Kunimori, "*Measuring Ajisai's spin motion*".

[13] Y. Kubo, "*The rotation of the geodetic satellite Ajisai around its center of mass*".

[14] J. J. Degnan, "*Millimeter accuracy satellite laser ranging: a review*", Contribution of Space Geodesy to Geodynamics: Technology; Geodynamic Series, Vol 25, 1993.

[15] L. Grunwaldt, H. Fischer, "*Field experince with various detector types at SLR- Station 7836 Potsdam*", Proc. 9th Intern. Workshop on Laser Ranging Instrumentation, Canberra, 1996.

TESTING THE INDEPENDENCE BETWEEN CLOCKS

FRANCISCO JAVIER GALINDO and JUAN PALACIO

Real Instituto y Observatorio de la Armada
C/CecilioPujazón S/N, 11.100 San Fernando, Cadiz, Spain
Tel. +34-956599286, Fax +34-956599366, jgalindo@roa.es, jpalacio@roa.es

ABSTRACT

A test based on the Chi-Squared distribution is described and applied to series of frequency deviations corresponding to several clocks contributing to an average time scale. This test allows to evaluate the hypothesis of independence usually supposed, helping even to detect a deficient operation in a clock.

Keywords: *Test of hypothesis of independence, Chi-Squared distribution, average time scale, contingency table, level of significance*

1. INTRODUCTION

When an atomic time scale is based on the weighted average of the contributing clocks, it is usually supposed that the clocks are independent. Many efforts are realised by the laboratories to avoid the influence between clocks, but always doubts exist about the consideration of independence.

A non-parametric procedure can be applied to test the hypothesis of independence between pairs of clocks. It is based on the comparison of the probabilistic distributions of the fractional frequencies deviations of two clocks measured versus a third reference one, more stable than the two ones under test. The observed fractional frequencies deviation of each clock are organised into classes and arranged in a two-dimensional array. For each cell of the table the observed probability is compared with the theoretical probability expected if clocks were independent: "The joint probability density of independent variables must be equal to the product of the two marginal probability densities".

The statistical test on the comparisons between observed and expected probabilities for every cell can be accomplished by using a χ^2 random variable.

This technique has been led to the practice on experimental data of frequency deviations measures of clocks maintained at the Real Instituto y Observatorio de la Armada "ROA" (referenced to UTC), coming out that correlation can be detected in some cases, particularly in case of aged clocks or "old generation" Cesium beam standards.

In the same way, some preliminary results were reached applying this method on rates series (referenced to TAI) of some contributing clocks to the TAI computation.

The presented method cans therefore helps in validating the assumption of clock independence.

2. THEORETICAL FOUNDATIONS

The test of independence, supported by the use and analysis of ***r* x *c* tables**, *r* by *c* tables or **contingency tables**, in which data are tallied into a two-way classification having *r* rows and *c* columns, is based on the sampling from a population and posterior classification of each item with respect to two (usually qualitative) categories, in several classes inside each criterion of classification. The rows represent levels or classes of criterion 1, and the columns represent the levels or classes of criterion 2.

For this particular application, each sample constitutes a two-dimensional random vector associated with certain instant t_0, obeying each component to the frequency deviations measures –for an integration time τ -, relative to a third more stable device (scale, clock, ...), of the i^{th} and j^{th} clocks respectively.

Once arranged the data into the contingency table, the Null hypothesis H_0 (independence of both criteria of classification) could be proved by comparison of how the frequency deviations are distributed into classes of the variable 1 (first criterion of classification) for each class of the variable 2 (second criterion of classification), or vice versa. If the statistical frequency distribution of the variable 1 is kept similar for each class of the variable 2, it is logical to think that both components of the two-dimensional random vector are independent. If on the contrary, distributions for each

Table 1. Test of independence, $\chi^2 > 13.28$ means dependence between pairs of clocks $(\alpha = 1\%)$.

	35 718	14 896	16 113	12 1223	14 1569	31 422	
35 583	4.63	4.21	11.47	6.56	0.79	27.88	35 583
	35 718	4.04	14.65	7.47	12.54	5.93	35 718
		14 896	2.87	4.10	8.61	27.03	14 896
			16 113	26.47	26.17	11.03	16 113
				12 1223	7.30	17.74	12 1223
					14 1569	21.97	14 1569
						31 422	

class of the variable 2 are different enough, it is reasonable to think that the hypothesis of independence is not completed (Alternative hypothesis).

The Null hypothesis could be expressed analytically as follows:

$$H_0: \quad p_{ij} = p_{i.} \cdot p_{.j} \tag{1}$$

Being p_{ij} the probability that an outcome belongs to the row i^{th} and column j^{th}, while $p_{i.}$ and $p_{.j}$ are the probabilities that an outcome belongs to the row i^{th} and column j^{th} respectively.

If the Null hypothesis is completed, the expected statistical frequency for any given cell e_{ij} may be obtained by multiplying the totals of the column and the row to which it belongs, and then dividing by the total number of samples n:

$$e_{ij} = \frac{p_{i.} \cdot p_{.j}}{n} \tag{2}$$

Now, the problem has been simplified to determine if the expected statistical frequencies represent to their corresponding observed o_{ij}. If agreement between both values exists, the deviations $o_{ij} - e_{ij}$ must be small. If on the contrary, the deviations are high, the independence hypothesis must be refused.

A reasonable indicator of the agreement between the observed and expected distributions will be:

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(o_{ij} - e_{ij})^2}{e_{ij}} \tag{3}$$

The Null hypothesis will be rejected if the value of this statistical exceeds that of $\chi^2_{1-\alpha}$ for $rc-1-[(r-1)+(c-1)] = (r-1)\cdot(c-1)$ degrees of freedom *dof*. Being α the level of significance, that is: the probability of committing a Type I error (probability that completing the null hypothesis, it has been rejected by no overcoming the test), and r and c the number of rows and columns of the contingency table (number of classes of the variables 1 and 2 respectively).

The level of significance should not be excessively small, because it would affect negatively to the probability that a Type II error is committed: probability that the Null hypothesis is accepted when the hypothesis of independence is false.

The values usually assigned to the coefficient α are: 0.05 and 0.01.

When the observed value χ^2 is such that $\chi^2 \geq \chi^2_{1-0.05}$, it is said that the observed value indicates a significant disagreement. If the boundary is established at 0.01, then it could say that the disagreement is highly significant.

3. PRACTICAL CONSIDERATIONS

So that this test type Chi Squared could apply with success, the following hypothesis should be verified:

- The n samples or random two-dimensional vectors implied in the test should be independent.
- The expected statistical frequency for each cell should be next to 5 or higher.
- The number of cells of the r x c table should be equal or higher than 4.

4. EXPERIMENTAL RESULTS

In a first experience, data of phase-time deviations (referenced to UTC –Annual Report of the BIPM Time Section, 1996 and 1997-) of seven Cesium beam standards belonging to the ROA, during the years 1996 and 1997, were taken. Starting from these data, frequency deviations were determined (time of integration τ =5d).

Since the predominant noise in Cesium beam standards is white of frequency type, the hypothesis of

Table 2.1. Test of independence applied to some USA HP 5071 hp Cs Standards, $\chi^2 > 6.63$ means dependence between pairs of clocks $(\alpha = 1\%)$.

	s/n 765 (AMC)	s/n 408 (NIST)	s/n 101 (USNO)	s/n 164 (USNO)	s/n 167 (USNO)	s/n 226 (USNO)	s/n 279 (USNO)	s/n 394 (USNO)	s/n 1097 (USNO)	
s/n 173 (AMC)	3.07	7.95	14.97	17.54	9.71	1.72	7.95	0.36	7.82	s/n 173
	s/n 765	0.65	1.63	0.88	5.31	0.17	3.28	0.11	1.32	s/n 765
		s/n 408	13.28	3.28	4.91	0.41	6.25	0.87	1.04	s/n 408
			s/n 101	9.03	9.24	0.26	13.28	0.01	3.71	s/n 101
				s/n 164	9.71	4.74	3.29	2.39	7.82	s/n 164
					s/n 167	4.59	4.91	1.65	1.73	s/n 167
						s/n 226	0.18	1.18	0.49	s/n 226
							s/n 279	0.88	1.04	s/n 279
								s/n 394	0.15	s/n 394
									s/n 1097	

Table 2.2. Test of independence applied to some Occidental Europe HP 5071 hp Cs Standards, $\chi^2 > 6.63$ means dependence between pairs of clocks $(\alpha = 1\%)$.

	s/n 172 (F)	s/n 198 (F)	s/n 385 (F)	s/n 536 (F)	s/n 609 (F)	s/n 770 (F)	s/n 781 (F)	s/n 583 (ROA)	s/n 718 (ROA)	
s/n 131 (F)	1.07	5.23	15.38	4.89	6.74	0.96	0.03	0.43	1.92	s/n 131
	s/n 172	1.73	4.57	4.81	3.38	0.24	1.73	0.11	0.46	s/n 172
		s/n 198	6.25	3.23	2.16	0.01	0.69	0.69	3.90	s/n 198
			s/n 385	3.23	5.23	0.69	0.69	0.01	3.90	s/n 385
				s/n 536	4.89	0.96	0.03	0.43	0.05	s/n 536
					s/n 609	0.03	0.96	0.03	0.34	s/n 609
						s/n 770	2.78	6.25	0.24	s/n 770
							s/n 781	2.78	0.24	s/n 781
								s/n 583	0.11	s/n 583
									s/n 718	

Table 2.3. Test of independence applied to some Central Europe HP 5071 hp Cs Standards, $\chi^2 > 6.63$ means dependence between pairs of clocks $(\alpha = 1\%)$.

	s/n 441 (GUM)	s/n 502 (GUM)	s/n 219 (IEN)	S/n 505 (IEN)	S/n 415 (PTB)	s/n 247 (TUG)	s/n 456 (VSL)	s/n 548 (VSL)	s/n 731 (VSL)	
s/n 771 (CH)	3.38	4.73	0.11	0.01	1.07	0.02	0.04	2.71	2.71	s/n 771
	s/n 441	12.75	0.03	1.96	0.04	0.05	2.56	5.00	5.00	s/n 441
		s/n 502	3.71	0.62	3.44	4.74	6.62	13.54	13.54	s/n 502
			s/n 219	8.36	9.64	14.31	4.17	9.37	9.37	s/n 219
				S/n 505	3.74	11.50	0.69	3.78	3.78	s/n 505
					S/n 415	7.00	2.56	9.42	9.42	s/n 415
						s/n 247	4.91	11.05	11.05	s/n 247
							s/n 456	5.25	5.25	s/n 456
								s/n 548	25.00	s/n 548
									s/n 731	

independence of samples will be assured. The other previous hypotheses are completed distributing the frequency deviations in three classes for each clock.

The number of *dof* is four, which added to the selected level of significance $\alpha = 0.01$, leads to $\chi^2_{0.99} = 13.28$.

Table 2.4. Test of independence applied to some Asia HP 5071 hp Cs Standards, $\chi^2 > 6.63$ means dependence between pairs of clocks $(\alpha = 1\%)$.

	s/n 715 (CRL)	s/n 732 (CRL)	s/n 908 (CRL)	s/n 224 (NRLM)	s/n 459 (NRLM)	s/n 523 (NRLM)	s/n 474 (TL)	s/n 1012 (TL)	
s/n 112 (CRL)	15.38	18.06	0.01	9.37	18.06	0.03	0.26	4.57	s/n 112
	s/n 715	21.30	0.96	15.23	14.55	1.99	0.07	7.00	s/n 715
		s/n 732	0.43	12.98	17.63	0.99	0.32	9.00	s/n 732
			s/n 908	0.12	0.03	2.16	0.26	0.24	s/n 908
				s/n 224	12.98	1.21	0.04	6.17	s/n 224
					s/n 459	0.99	0.32	9.00	s/n 459
						s/n 523	1.96	1.92	s/n 523
							s/n 474	2.93	s/n 474
								s/n 1012	

Table 2.5. Test of independence applied to some H Masers, $\chi^2 > 6.63$ means dependence between pairs of clocks $(\alpha = 1\%)$.

	s/n 5623 (KRIS)	s/n 204 (NIST)	s/n 1701 (NPL)	s/n 505 (PTB)	s/n 710 (USNO)	s/n 711 (USNO)	
s/n 713 (AMC)	4.89	18.06	17.63	4.89	5.94	9.64	s/n 713
	s/n 5623	9.64	4.89	0.04	0.99	0.03	s/n 5623
		s/n 204	18.06	2.16	3.71	6.25	s/n 204
			s/n 1701	4.89	5.94	9.64	s/n 1701
				s/n 505	7.35	11.78	s/n 505
					s/n 710	17.65	s/n 710
						s/n 711	

Table 1 displays the seven clocks arranged from high to low stability. Their type (two digits) and serial number in the type identify the clocks. The codes for the types are similar to those employees for the table 13A in the Annual Report of the BIPM Time Section. Each bin contains the obtained result χ^2 for each pair of considered clocks. Once applied the test, if dependence between clocks is deduced, the bin is pointed out with a grey background.

By a glance to this table, it seems obvious that the obtained results for the clocks serial number 113 and 422 are abnormally high. The first clock (an OSCILLOQUARTZ Mod.: 3200) failed definitively, stopping its contribution to the TAI on march 1998. The second clock (a HP5061B Opt. 4) is the older (eight years old) and the less stable of that laboratory, although it is still in operation. This is indicative that the behaviour has degenerated, due surely to the approaching to the limit of beam tube life.

In a second experience, data of mean clock rates relatives to TAI for one-month intervals from January 1998 to January 2000 were used. These series were obtained from the Annual Report of the BIPM Time Section (1998) or downloaded via ftp 62.161.69.5 user anonymous. The total number of complete data series obtained was 81, the data series have been grouped in accordance with criteria of model of clock and geographical vicinity into five tables.

Unfortunately, due to the optimisation time interval of the EAL was reduced to one month in January of 1998, the maximum number of available rate data $(\tau = 1\text{m})$ was 25. This is a reduced number in order to take out definitive conclusions, but they could give us a qualitative idea on possible dependence between pair of clocks contributing to TAI.

Although the flicker and random walk of frequency noise begins to be present, it has been considered that the hypothesis of independence between samples is still valid. We distributed the rates into two classes. The number of *dof* is one, and $\chi^2_{0.99} = 6.63$ for $\alpha = 0.01$. With great difficulty it could have completed the second previous hypothesis: minimum expected statistical frequency.

Nevertheless, tables 2.1 to 2.5 display the obtained preliminary results, showing those pair of clocks in which an effect of correlation is suspected (grey background). In these tables, their serial number and the laboratory of ownership identify the clocks.

It is emphasised that the obtained results are not definitive. In future works, when longer series of data are available, the results then obtained will be reliable enough.

5. CONCLUSIONS

A test based on the Chi-squared distribution, simple of applying and useful in order to detect correlative effects between pair of clocks, has been introduced.

Results obtained using experimental data of clocks maintained at the ROA warned of the faulty operation and predicted the death of a clock.

Also, initial results on clocks that contributing to the TAI were obtained, but these results are only of orientation. In a work later on, definitive results will be obtained to validate or not the usual supposed independence hypothesis.

6. REFERENCES

[1] H.J. Larson, Introduction to Probability Theory and Statistical Inference, John Wiley & Sons, Inc., 1982, pp. 541-551.

[2] I. Miller and J.E. Freund, Probability and Statistics for Engineers, Prentice-Hall, Inc., 1985, pp. 262-265.

[3] J.R. Taylor, An introduction to Error Analysis: The Study of Uncertainties in Physical Measurements, University Science Books, 1997, pp. 261-278.

OPTIMAL WEIGHTING OF CORRELATED CLOCKS IN THE DEFINITION OF AN ENSEMBLE TIME-SCALE

FRANCISCO JAVIER GALINDO
Real Instituto y Observatorio de la Armada
C/Cecilio Pujazón S/N, 11.100 San Fernando, Cádiz, Spain
Tel. +34-956599286, Fax +34-965599366, jgalindo@roa.es

PATRIZIA TAVELLA
Istituto Elettrotecnico Nazionale G. Ferraris
Strada delle Cacce 91, 10135 Torino, Italy
Tel. +39-11-3919235, Fax +39-11-3919259, tavella@tf.ien.it

ABSTRACT

The optimisation of the stability of an average time scale obtained with correlated clocks is considered, by deriving the optimal clocks weights and by examining, by means of simulated clocks, which is the impact of the use of those optimal weights.

1. INTRODUCTION

Correlation between clocks, mainly due to environmental conditions but not only, can often be detected in T&F laboratories. In the last years, a certain effort was devoted to the estimation of correlation besides the usual estimation of the clock variances [1-5]. Once a correlation is detected and estimated, the subsequent step is the possible use of such additional information. In particular, in the ensemble time algorithms, clocks are usually assumed as independent and each clock is weighted as inversely proportional to its variance. If a correlation exists, this weighting is no longer optimal for minimising the instability of the ensemble time. A different weighting has to be used which takes into account the variance of each clock and also the covariance, or cross-correlation, with any other clocks. The optimal weight in this case can be obtained by using the general theory of least squares. Also the common belief that with N clocks, the stability of the ensemble time is improved by a factor $1/\sqrt{N}$ is appropriate only if clocks are independent, if not the stability of the ensemble has a more general expression containing also clock covariances. In this paper, we present the matricial formulation of the least squares to calculate the optimal weights in case of correlated or uncorrelated clocks. Then, by the aid of simulated clocks, we examine the impact of this weighting procedure on the instability of the ensemble time when correlation is appropriately taken into account or in the case it is neglected. The problem of how a correlation can be estimated from real clock data is also addressed.

2. OPTIMAL WEIGHTS

2.1 General formulation

When we have at disposal N measures $\mathbf{Y}$ concerning Q unknowns $\mathbf{X}$ related by a known relationship, the measurement equation can be written in the matricial form

$$\mathbf{Y} = \mathbf{M} \cdot \mathbf{X} + \mathbf{e} \qquad (1)$$

where $\mathbf{Y}$ is the (Nx1) vector containing the measures, $\mathbf{X}$ is the (Qx1) vector containing the unknowns and the (NxQ) matrix $\mathbf{M}$ contains the relationships between $\mathbf{X}$ and $\mathbf{Y}$. The vector $\mathbf{e}$ contains the error of the measures $\mathbf{Y}$. The measures $\mathbf{Y}$ are supposed to be unbiased, therefore $E[\mathbf{e}] = 0$ and the covariance matrix of the measurement error is called $\mathbf{R}$:

$$\mathbf{R} = E[(\mathbf{Y} - E[\mathbf{Y}]) \cdot (\mathbf{Y} - E[\mathbf{Y}])^T] = E[\mathbf{e} \cdot \mathbf{e}^T] \qquad (2)$$

The term r_{ii} on the main diagonal is the variance of the i-th measurement, while the term off-diagonal r_{ij} $i \neq j$ is the covariance of the i-th and j-th measurement. The least square best estimator $\hat{\mathbf{X}}$ of the unknown $\mathbf{X}$ is given by the Gauss-Markov theorem [6]:

$$\hat{\mathbf{X}} = Cov[\hat{\mathbf{X}}] \cdot \mathbf{M}^{\mathbf{T}} \cdot \mathbf{R}^{-1} \cdot \mathbf{Y} \qquad (3)$$

where $Cov[\hat{\mathbf{X}}]$ is the covariance matrix of the estimates $\hat{\mathbf{X}}$ and it is given by:

$$Cov[\hat{\mathbf{X}}] = (\mathbf{M}^{\mathbf{T}} \cdot \mathbf{R}^{-1} \cdot \mathbf{M})^{-1} \qquad (4)$$

In case the observation $\mathbf{Y}$ are all independent and with the same variance σ^2, the matrix $\mathbf{R}$ is diagonal written as: $\mathbf{R} = \sigma^2 \cdot \mathbf{I}$, where $\mathbf{I}$ is the unit matrix. Then (3) and (4) are reduced to:

$$\hat{\mathbf{X}} = (\mathbf{M}^{\mathbf{T}} \cdot \mathbf{M})^{-1} \cdot \mathbf{M}^{\mathbf{T}} \cdot \mathbf{Y} \qquad (5)$$

$$Cov[\hat{\mathbf{X}}] = \sigma^2(\mathbf{M}^{\mathrm{T}} \cdot \mathbf{M})^{-1} \qquad (6)$$

2.2 Weighted average

Let's suppose to have at disposal N measures of clock readings h_i. They measure the same quantity, the "atomic time" TA (Q=1), therefore the model (1) may be written as:

$$\begin{vmatrix} h_1 \\ h_2 \\ . \\ h_N \end{vmatrix} = \begin{vmatrix} 1 \\ 1 \\ . \\ 1 \end{vmatrix} \cdot TA + \begin{vmatrix} e_1 \\ e_2 \\ . \\ e_N \end{vmatrix} \qquad (7)$$

The vector containing N elements equal to 1 is called **u.** By applying (3) and (4), we find that the best estimation of the TA is given by:

$$\hat{TA} = (\mathbf{u}^{\mathrm{T}} \cdot \mathbf{R}^{-1} \cdot \mathbf{u})^{-1} \cdot \mathbf{u}^{\mathrm{T}} \cdot \mathbf{R}^{-1} \cdot \mathbf{Y} \qquad (8)$$

and its covariance matrix reduces to one element that is the variance of TA

$$Cov[\hat{TA}] = \sigma^2_{TA} = (\mathbf{u}^{\mathrm{T}} \cdot \mathbf{R}^{-1} \cdot \mathbf{u})^{-1} \qquad (9)$$

Since the quantity $(\mathbf{u}^{\mathrm{T}} \cdot \mathbf{R}^{-1} \cdot \mathbf{u})$ is a scalar, (8) may be written as:

$$\hat{TA} = \frac{\mathbf{u}^{\mathrm{T}} \cdot \mathbf{R}^{-1}}{\mathbf{u}^{\mathrm{T}} \cdot \mathbf{R}^{-1} \cdot \mathbf{u}} \cdot \mathbf{Y} = \boldsymbol{\omega}^T \cdot \mathbf{Y} = \sum_{i=1}^{N} \omega_i h_i \qquad (10)$$

where the most common writing has been introduced with a vector $\boldsymbol{\omega}$ of weights. In the next section, an example will illustrate the content of these equations.

In this case, it should be pointed out that **R** is the Allan covariance matrix for the selected time interval τ to optimise, and therefore, the variance of the ensemble time TA in (9) is referenced to the same observation time.

An alternative derivation of the optimal weight is given by the more "traditional" procedure of minimisation of the ensemble time instability with appropriate constrains. In such a case, the $\hat{TA}$ is defined as:

$$\hat{TA} = \sum_{i=1}^{N} \omega_i h_i$$

and the weights ω_i are chosen in order to minimise the instability of TA. To this aim, the weights have to verify the following equation system:

$$\sum_{i=1}^{N} \omega_i \cdot (r_{ij} - r_{iN}) = 0, \quad j = 1,2,...,(N-1)$$

$$\sum_{i=1}^{N} \omega_i = 1$$

deduced by minimizing the expression of the variance of the ensemble time scale, i.e.

$$min\left(\sum_{i=1}^{N} \sum_{j=1}^{N} \omega_i \omega_j r_{ij} \right) = min\left(\sum_{i=1}^{N} \omega_i^2 r_{ii} + 2 \cdot \sum_{i<j} \omega_i \omega_j r_{ij} \right)$$

subjected to the constraint

$$\sum_{i=1}^{N} \omega_i = 1$$

The optimal weights, written in compact form, appear to be:

$$\boldsymbol{\omega} = \frac{\mathbf{R}^{-1} \cdot \mathbf{u}}{\mathbf{u}^T \cdot \mathbf{R}^{-1} \cdot \mathbf{u}} \qquad (11)$$

as introduced in (10). This writing is in agreement with another derivation recently presented [7].

Let's suppose that only the diagonal term of the matrix **R** are used in the weight definition, i.e. the covariance terms are neglected. Thus the complete matrix **R** is approximated by the diagonal $\mathbf{R_D}$. By the use of matricial algebra, it can be demonstrated that the new weighs are:

$$\boldsymbol{\omega}^* = \frac{\mathbf{R}_D^{-1} \cdot \mathbf{u}}{\mathbf{u}^T \cdot \mathbf{R}_D^{-1} \cdot \mathbf{u}} \qquad (12)$$

and the resulting TA has variance equal to:

$$\sigma^{*2}_{TA} = \frac{\mathbf{u}^T \cdot \mathbf{R}_D^{-1} \cdot \mathbf{R} \cdot \mathbf{R}_D^{-1} \cdot \mathbf{u}}{\left(\mathbf{u}^T \cdot \mathbf{R}_D^{-1} \cdot \mathbf{u}\right)^2} \qquad (13)$$

For any other weight definition, the variance of TA has a structure similar to (13).

2.2.1 Example of weighting

Let's suppose to have N independent clocks with variances σ_i^2. The matrix **R** is diagonal, written as:

$$\mathbf{R} = \begin{vmatrix} \sigma_1^2 & 0 & . & 0 \\ 0 & \sigma_2^2 & . & 0 \\ . & . & . & 0 \\ 0 & 0 & 0 & \sigma_N^2 \end{vmatrix} \qquad (14)$$

Eq. (10) and (9) becomes the commonly know definition of the ensemble time and its instability given by:

$$\hat{TA} = \frac{\sum_{i=1}^{N} \frac{1}{\sigma_i^2} \cdot h_i}{\sum \frac{1}{\sigma_i^2}} \qquad (15)$$

$$\sigma_{TA}^2 = \frac{1}{\sum \frac{1}{\sigma_i^2}} \qquad (16)$$

Let's now suppose to have N=2 clocks not independent and not with equal variance: the most general case. The covariance matrix **R** is written as:

$$\mathbf{R} = \begin{vmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{vmatrix} \qquad (17)$$

where $\sigma_{12} = \sigma_{21}$ by definition of covariance. In this case, (9) and (10) give:

$$T\hat{A} = \frac{\sigma_2^2 - \sigma_{12}}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}} h_1 + \frac{\sigma_1^2 - \sigma_{12}}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}} h_2 \qquad (18)$$

$$\sigma_{TA}^2 = \frac{\sigma_1^2 \sigma_2^2 - \sigma^2{}_{12}}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}} \qquad (19)$$

3 APPLICATION EXAMPLES

3.1 Case 1: Strong correlation between clocks

In Fig. 1, the stability of four simulated clocks is represented by dotted lines. The sampling time τ_0 is 1 day. Initially, the predominant noise is white FM. The four clocks have a predetermined amount of Allan deviation, which was chosen to be: $\sigma_y(2\,\text{days}) =$ $= 1,\ 2, 2, 15$ respectively. The estimates of Allan deviation have been determined by means of the overlapping estimation process.

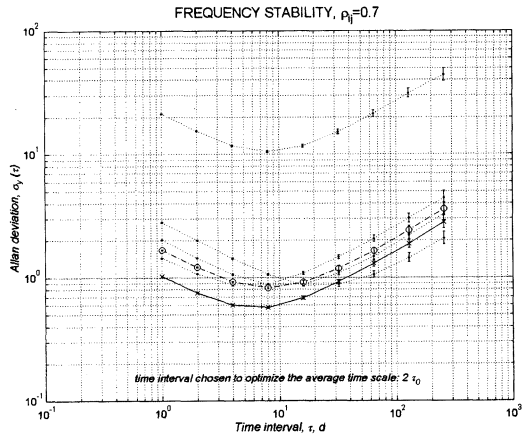


Figure 1

The clocks were simulated not independently but with a certain correlation that corresponds to a correlation coefficient between pairs of clocks equal to $\rho_{ij} = +0.7$.

The average time scale *TA* is computed with two different sets of weight. In the first case, clocks are supposed independent and, therefore, their covariance matrix **R** is approximated by $\mathbf{R_D}$ and the weights are given by (12). The resulting stability of *TA* is depicted (dashdot line with circles) in the same Fig.1

In a second step, the correlation between pairs of clocks is considered and the matrix **R** is fully used. The weights are defined by (11). The stability of the resulting *TA*, is in Fig.1 (solid line).

The obtained *TA* stabilities are in agreement with the expected values given in (13) and (9) respectively.

The selected time interval to optimize both average time scales is $\tau = 2$. Therefore matrix **R** is evaluated for this integration time.

The Fig. 2 shows the same four clocks and the corresponding average time scales when the optimization interval is $\tau = 64$ days. The dominant noise is, at this τ, random walk FM. The predetermined values of Allan deviation are $\sigma_y(64\,\text{days}) = 2,\ \sqrt{2}, 1,$ and 20 respectively.

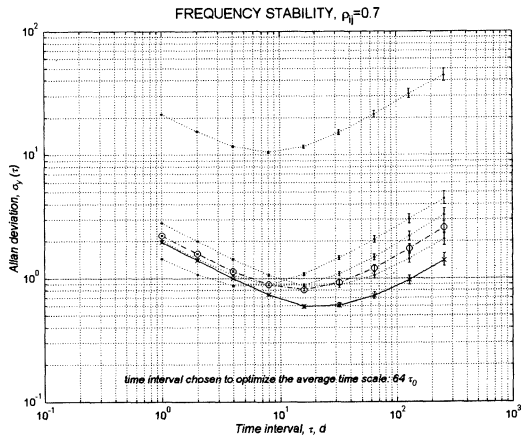


Figure 2

The cross correlation coefficient between pairs of data series maintains the value of $\rho_{ij} = +0.7$.

Some conclusions:

1. The average time scales obtained when independence among data series is supposed, is usually less stable than the one obtained by means of the full use of the covariance matrix **R** and therefore of the optimal weights.
2. If covariances are neglected, the average time scale can be less stable than some of the individual clocks.
3. When the weights are evaluated with their optimal definition (11), some of them may assume negative values. This is a typical situation also observed in the gain of the Kalman filter, where "weights" lose

their intuitive meaning due to the presence of correlation.

In this simulated case, the complete matrix $\mathbf{R}$ is known and therefore it is possible to evaluate its impact in the weight definition. In reality, the covariance matrix of individual clocks is not known because only comparison measures are possible. This is the so-called "*N* cornered hat" problem. Recently [3-5], some solutions were proposed for the *N* cornered and they are here examined with the aim of giving examples on how the matrix $\mathbf{R}$ can be estimated and used in the definition of the average time scale. Here we use the techniques developed by USNO [4] and ROA [5]. The matrix $\mathbf{R}$ is estimated by $\hat{\mathbf{R}}$ and it is used to define the weights. Again we consider the full $\hat{\mathbf{R}}$ or its approximation $\hat{\mathbf{R}}_D$.

According to the treatment leading to (12) and (13), the weights are then defined by:

$$\hat{\omega} = \frac{\hat{\mathbf{R}}^{-1} \cdot \mathbf{u}}{\mathbf{u}^T \cdot \hat{\mathbf{R}}^{-1} \cdot \mathbf{u}} \qquad (20)$$

$$\hat{\omega}^* = \frac{\hat{\mathbf{R}}_D^{-1} \cdot \mathbf{u}}{\mathbf{u}^T \cdot \hat{\mathbf{R}}_D^{-1} \cdot \mathbf{u}} \qquad (21)$$

and the stability of the average is given by:

$$\hat{\sigma}_{TA}^2 = \frac{\mathbf{u}^T \cdot \hat{\mathbf{R}}^{-1} \cdot \mathbf{R} \cdot \hat{\mathbf{R}}^{-1} \cdot \mathbf{u}}{\left(\mathbf{u}^T \cdot \hat{\mathbf{R}}^{-1} \cdot \mathbf{u}\right)^2} \qquad (22)$$

$$\hat{\sigma}^{*2}_{TA} = \frac{\mathbf{u}^T \cdot \hat{\mathbf{R}}_D^{-1} \cdot \mathbf{R} \cdot \hat{\mathbf{R}}_D^{-1} \cdot \mathbf{u}}{\left(\mathbf{u}^T \cdot \hat{\mathbf{R}}_D^{-1} \cdot \mathbf{u}\right)^2} \qquad (23)$$

Fig. 3 and 4 show the stability associated with the different time scales optimized for $\tau = 2$ and $\tau = 64$ days respectively. The full use of the estimated covariance matrix is represented by means of continuous lines, making use of the dotted lines when the only information considered is that of the diagonal of $\hat{\mathbf{R}}$.

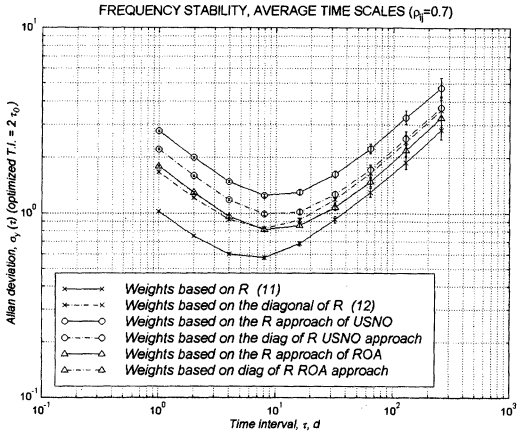


Figure 3

Some comments can be added:

1. When the covariance matrix $\mathbf{R}$ is estimated, results are worst than those reached using the true matrix $\mathbf{R}$. This is mostly justified by the fact that the *N* cornered techniques developed at USNO, as well as at ROA, minimize the quadratic sum of clock covariances and correlations respectively. This hypothesis is not valid now because a strong correlation was imposed in the simulation.
2. The ROA algorithm gives very small off-diagonal terms, therefore using the approximate $\hat{\mathbf{R}}_D$ or the complete $\hat{\mathbf{R}}$ doesn't make a substantial difference.
3. The ROA algorithm reaches a good estimate of the individual Allan variances (diagonal of $\mathbf{R}$), leading to an ensemble time scale very similar to the one obtained considering only the diagonal of $\mathbf{R}$.

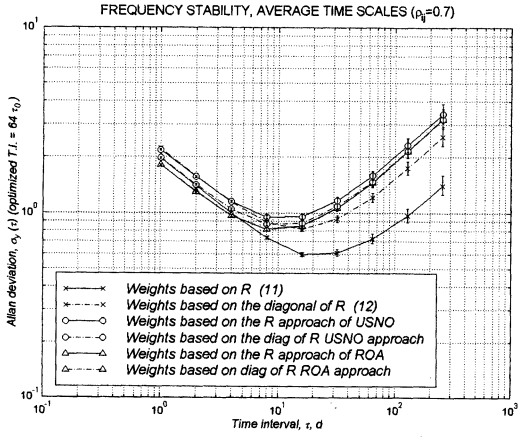


Figure 4

3.2 Case 2: Weak correlation between clocks

The same procedure was repeated with four simulated clocks with small correlation, $\rho_{ij} = 0.1$. The corresponding stabilities are showed in Fig.5 to 8.

The most important appearing feature is that, in this case of small correlation, there is no significant difference between the use of the full estimated covariance matrix, or only of the diagonal elements in the weight definition, as expected.

Main conclusions:

1. The best results are obtained using independent clocks. It is very important to assure, whenever it is possible, the independence of the measured clock deviations.
2. When independence between pair of clock measures is not assured, the covariance values should be estimated. This information must be used to calculate the weighs for optimizing the average time scale.
3. If the correlation between couple of clocks is low, the stability of the ensemble time is not dramatically affected by the approximation of the

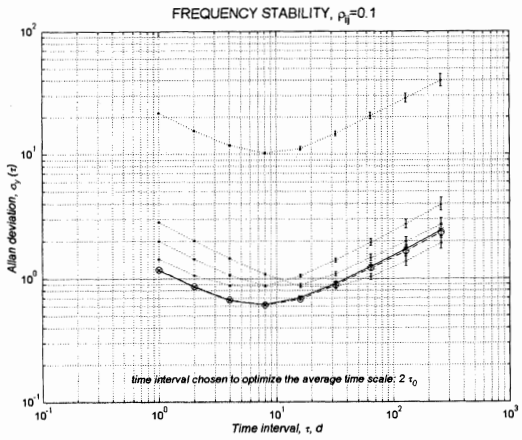


Figure 5

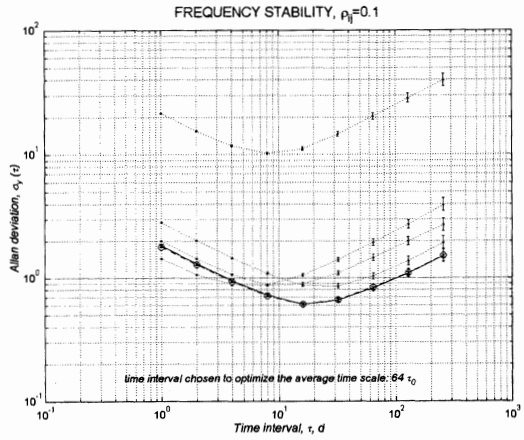


Figure 6

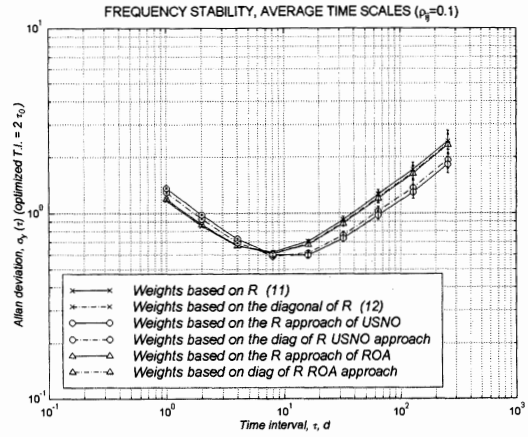


Figure 7

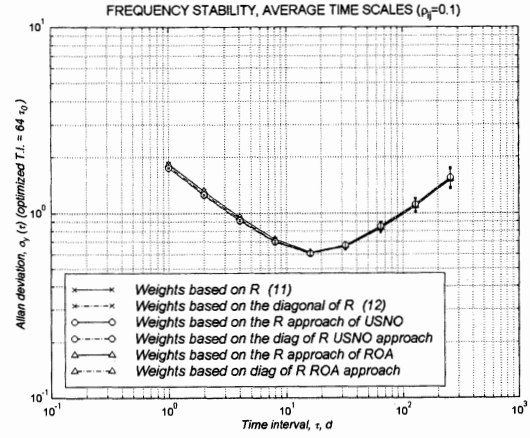


Figure 8

matrix $\hat{\mathbf{R}}$ with its diagonal form.

4 CONCLUSIONS

The optimal weighting of correlated clocks in the definition of an average time scale is obtained according to the least squares theory. The effect of the correlation in the definition of optimal weight is discussed with the aid of some simulated clocks showing that, in case of not negligible clock correlation, the stability of the average time scale is optimised only if the complete definition of weights, considering also covariance, is used.

ACKNOWLEDGEMENTS

The authors are thankful to W. Bich of IMGC and D. Matsakis of USNO for helpful discussions.

REFERENCES

[1] P. Tavella, C. Thomas, "Study of the correlations among the frequency changes of the contributing clocks to TAI," in Proc. 4th EFTF, Neuchatel, Swiss, 1990, pp. 527-541.

[2] A.Lepek, F. Walls, "Cross correlation analysis improves time domain measurements, Proc. IEEE Frequency Control Symp. 1993, pp. 312-321.

[3] P.Tavella, A.Premoli, "Estimating the instabilities of *N* clocks by measuring differences of their readings", Metrologia, 1994, 30, 5, pp. 479-486.

[4] F. Torcaso, C. R. Ekstrom, E. A. Burt, D.N. Matsakis, "Estimating the frequency stability and cross-correlations", in Proc. Precise Time and Time Interval, Washington, 1998, pp. 69-82.

[5] F.J.Galindo, J.Palacio, "Estimating the instabilities of *N* correlated clocks", to appear in Proc. Precise Time and Time Interval, Dana Point, CA, 1999.

[6] F.L. Lewis, Optimal estimation, New York: Wiley, 1986.

[7] C.R. Ekstrom, F. Torcaso, E.A. Burt, D.N. Matsakis, "Estimating the Stability of *N* correlated clocks with correlations", Joint meeting EFTF-IEEE IFCS, pp. 168-172.

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

USING PRECISE EPHEMERIDES FOR GLONASS P-CODE TIME TRANSFER

J. Azoubib[1], W. Lewandowski[1], J. Nawrocki[2],

[1]Bureau International des Poids et Mesures, Sèvres, France
[2]Astrogeodynamical Observatory, Borowiec, Poland

Abstract

GLONASS P-code has two main advantages for precision time synchronization. First, GLONASS P-code has a chip length that is 1/10th of GLONASS C/A-code chip length and about 1/5th of GPS C/A-code chip length. This allows GLONASS P-code pseudo-range measurements to be considerably more precise than comparable GPS or GLONASS C/A-code measurements. Second, GLONASS P-code is transmitted on both L1 and L2 frequencies, without Anti-Spoofing (AS) encryption. The absence of AS encryption allows GLONASS P-code measurements to be used for high-precision ionospheric measurements.

GLONASS data are subject to a receiver bias which may be different for each GLONASS frequency. The spread of these biases across satellites can reach 15 nanoseconds and therefore mask other noise sources. Based on the data available so far, GLONASS frequency biases appear to be a function of temperature and relate to specific receivers. But once calibrated with respect to a reference receiver, and provided that temperatures are maintained via laboratory air-conditioning together with a temperature stabilized antenna set-up, these values remain pretty constant and can therefore be compensated in the software.

In this paper we describe a test of long distance time transfer using GLONASS P-code multichannel common-view measurements between some US and some European time laboratories. At both sites temperature stabilized antennas were used, and GLONASS frequency biases were determined by means of a portable reference receiver. It is shown that the use of IGEX GLONASS precise ephemerides greatly improves the long distance GLONASS P-code time transfer: its performance is at least as good as that of GPS used with precise ephemerides.

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

SHORT-TERM STABILITY OF CRYSTAL OSCILLATORS IN COMMERCIAL GPS RECEIVERS

J.Mannermaa [1], K.Kalliomäki [2], T. Mansten [3], S.Turunen [1]

[1)] Nokia Mobile Phones, FIN-33100, Tampere, Finland
[2)] Univ. of Oulu, Dept. of Electrical Eng., FIN-90570 Oulu, Finland
[3)] MIKES, Electricity and Time Metrology, FIN-02150 Espoo, Finland

J.Mannermaa, Nokia Mobile Phones, FIN-33100, Tampere, Finland
Tel. +358-50-541 0801, Fax +358-10-505 6953, email jari.mannermaa@nokia.com

ABSTRACT

The short-term ($\leq$ 1 s) and long term (up to $3 \cdot 10^5$ s) stabilities of the crystal oscillators used in five different commercial GPS receivers were studied. A special set-up was designed and evaluated for the short-term measurements, otherwise a counter was used. Sampling intervals were from 1 ms to 10 minutes. A Cs clock was used as a time and frequency reference. Analyses of this study were carried out by using Allan variances (FVAR and TVAR), power spectra and noise estimation by standard deviation. The results show that Allan frequency variances are nearly constant at integration times between 10 ms and 100 s. After 100 s they begin to fall about 1.5 decades/decade of integration time, reaching $5 \cdot 10^{-14}$ at 3 day integration time. The respective Allan time variance above 1 s integration time is better than 100 ns. Power spectra point out that in the short-term case the noise floor of one of the receivers is lower than the others. Stability properties of the studied GPS receivers do not notably differ from each other. The results show that thermal design is important factor affecting the short-term stability and overall performance of the GPS receivers.

1. INTRODUCTION

The last decade has been the time of intensive studies of GPS receiver technologies all over the world. Scientific and economic opportunitities concerning the timing, positioning and navigation activities seem to create huge scenarios for the future. Due to the expanding range of GPS receiver applications with e.g. more stringent positional accuracy demands manufacturers have to concentrate more and more on the performance of internal clocks of GPS receivers. The R&T unit of Nokia Mobile Phones has also made these studies in co-operation with VTT [(1)] Automation, Measurement Technology. Tests were started in spring 1998, have a non-stop characteristic and reach for long term applications of GPS receivers, aiming to achieve comparable information and results of the overall performance of different GPS technologies. Some results of the long term studies are included in this article because they facilitate the understanding of the changes observed in the short-term vs. long term behaviour of in a receiver.

Our motivation and main interest in this study was the short-term stability of the crystal oscillators used in commercial GPS receivers down to 1 ms (the length of C/A code), especially their Allan variances with integration times below 1 s. Another reason for these researches was the lack of relevant knowledge of the short-term stability of the practical, cheap or modest price, commercial XOs and TCXOs (temperature compensated crystal oscillators). Our measurements covered integration times from 1 ms up to 300 000 s. A special set-up was designed for the measurements. To avoid loading the crystal, sensitive capacitive probes without galvanic contact were used. The noise bottom level for the set-up was determined, too. The oscillators were also tested by cooling them with short, local ventilation pulses while measuring the PPS pulse and the short-term crystal phase. The results were mostly omitted to keep the paper within the allowable size of the proceedings articles, and to be published later.

2. SET-UP

The used short-term measurement set-up is shown in Fig. 1. The probe signal was mixed down to about 100 kHz IF-signal using a stable local oscillator. IF-signal was fed to a phase detector together with another stable 100 kHz signal. The resulting phase signal was filtered with a 3 kHz low pass filter to a saw-tooth-shaped phase difference voltage, which was AD-converted, and after appropriate scaling Allan variances were calculated from it. The Allan frequency noise bottom level of the measurement system was determined to be about $5 \cdot 10^{-8}$ at 1 ms and drops in proportion to $\tau^{1.5}$ up to 0.5 seconds. Above 0.5 s the slope is slower reaching $1 \cdot 10^{-12}$ at 5 s. Above 5 s a fast time interval counter with Cs reference was used. The respective bottom level for short-term phase measurements is at best about 10 ps.

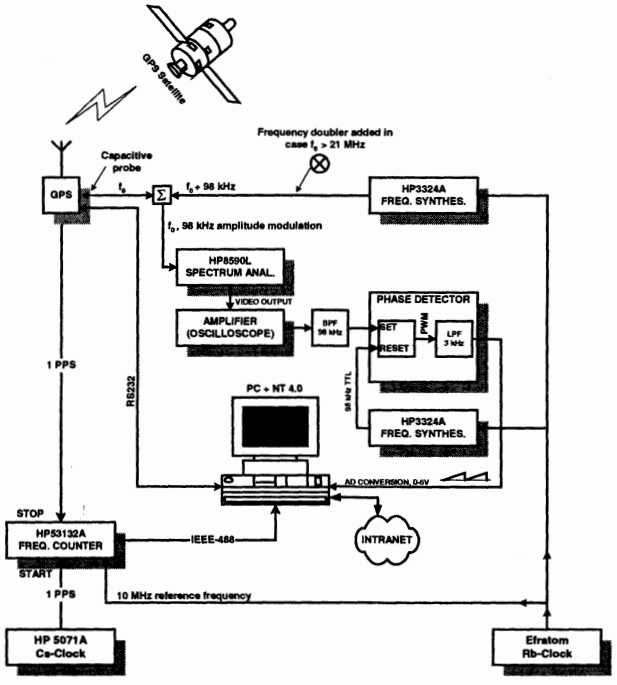


Fig. 1: Short-term measurement set-up

3. MEASURED GPS-RECEIVERS

The measured GPS receivers are given in Table I. Moreover, the Cs-monitored Motorola Oncore receiver of the NMI (National Metrology Institute) was used as a reference, see Fig. 3. The studied GPSs used their own antennae, which were installed on a flat roof. Because the antenna cables were short, a special installation housing had to be built inside, beneath the roof. All measurements were done by VTT at Espoo.

Table I: The measured GPS receivers.

Garmin	LP, chip set/OEM
Rockwell	chip set
Trimble	SK8, chip set
Ashtech	G8 chip set/OEM
SiRF	SiRFstar, LX, chip set/OEM

4. ANALYSIS METHODS

Sampling intervals were from 1 ms to 10 minutes. Due to a limited number of measurement channels in use, all GPS receivers were not measured at the same time. In this study short-term sampling intervals from 1 ms to 40 ms were applied to provide 8000 values for each time series.

Analyses of this study are focussed on some generally known parameters and functions e.g. on Allan variance (FVAR and TVAR), power spectrum and noise estimated by standard deviation of the PPSs. By FVAR and TVAR, the Allan frequency variance and Allan time variance is meant, respectively. Power spectra were computed by Maximum Entropy Method (MEM). The same methods were applied to time series of each sampling interval. Short-term data had to be preprocessed due to the fact that e.g. the output voltage of the phase detector was a (0 to 5 V) saw pulse train, and scaled into monotonous phase variable (in microseconds).

5. MEASUREMENTS AND RESULTS

In the beginning the noise bottom level components of the measurement set-up were measured and analyzed. The resulting FVAR and TVAR are in Figs. 2a, 2b. HP 5071 Cs specifications are added for reference.

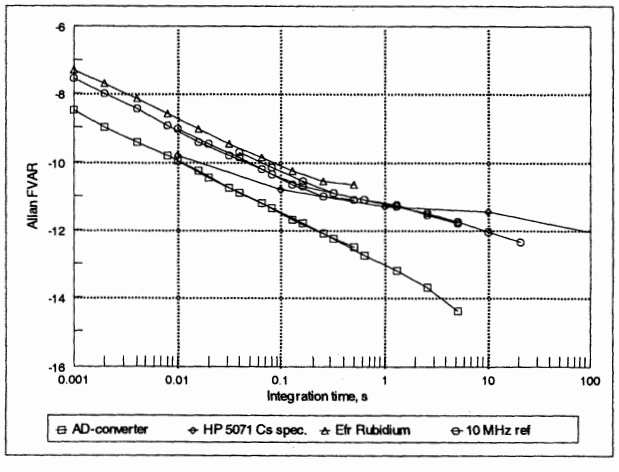


Fig. 2a: Allan FVAR bottom levels for fast Allan variance measurement set-up

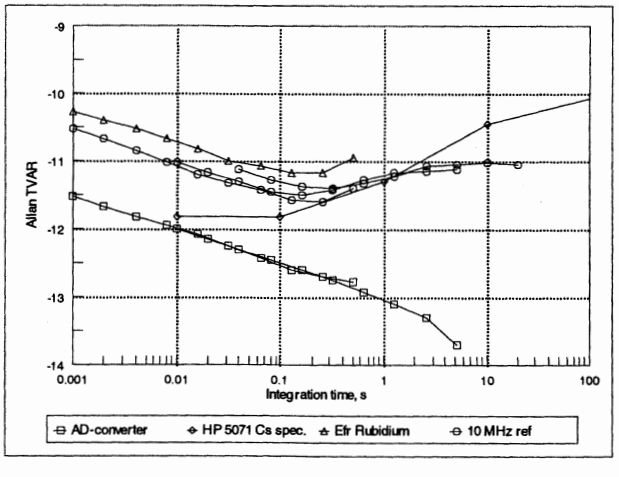


Fig. 2b: Allan TVAR bottom levels for fast Allan variance measurement set-up

The Allan FVAR of the Motorola Oncore GPS receiver is depicted in Fig. 3. No short-term measurements were done to it. The Allan FVAR of a SiRF GPS receiver is depicted in Fig. 4, including short-term results. The difference between the two short-term results is due to the ambient temperature, which causes parabolic phase variations to longer time series. The difference at 1 s is caused by the PPS phase correction, which is not visible to the short-term measurement.

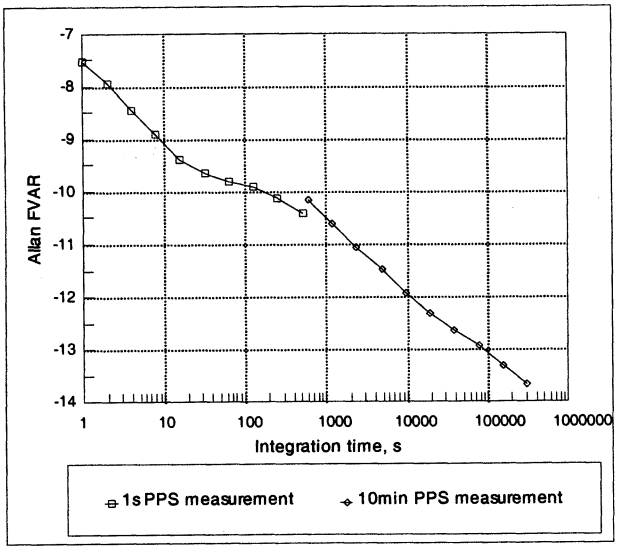


Fig. 3: Allan FVAR for Motorola Oncore GPS-receiver

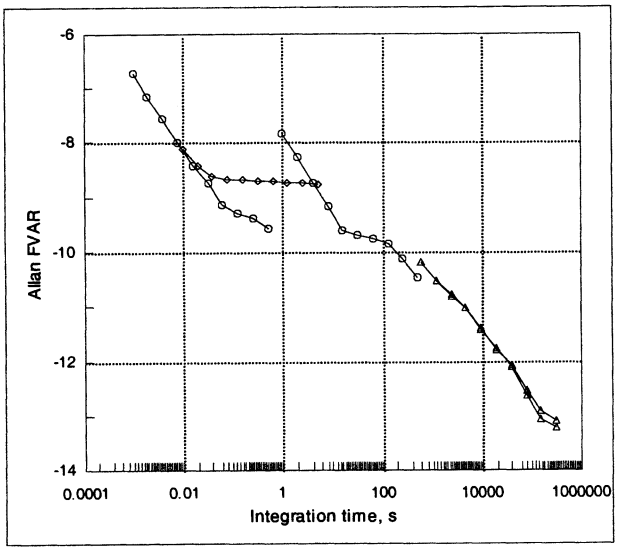


Fig. 4: Allan FVAR for SiRF GPS-receiver

The short-term and long term power spectra for the measured receivers are depicted in Figs. 5a and 5b, respectively. The results are calculated by maximum entropy method (MEM). M is the length of prediction error filter. The raw data for 5a is taken from short-term measurements (8000 points, 10 ms sampling), for 5b from 10 min interval measurements. The absolute origins (0 dB) of the vertical scales are not the same in Figs. 5a, 5b due to different averaging processes of the sampled data series.

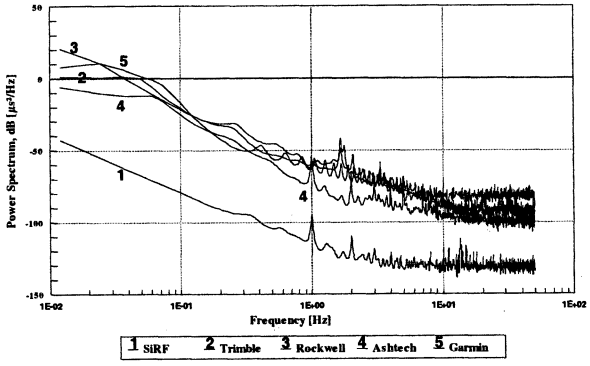


Fig. 5a: Power spectra of GPS receiver crystal phases; maximum entropy method (MEM); M = 500

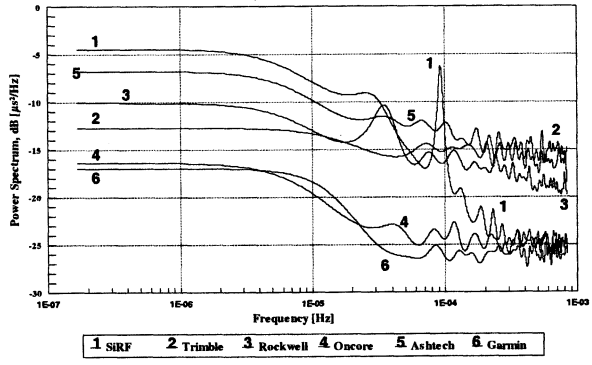


Fig. 5b: Power spectra of GPS receiver PPS pulses; maximum entropy method (MEM); M = 50

The Allan variances of all measured GPS receivers together with noise bottom levels and a few high-quality oscillator specifications from litterature are collected into Figs. 6a, 6b.

One oscillator was tested by cooling it with short, local ventilation pulses while measuring the PPS pulse and the crystal phase. The immediate relative frequency changes varied from $2 \cdot 10^{-7}$ to $1 \cdot 10^{-6}$. The frequency recovered with about 30 s time constant.

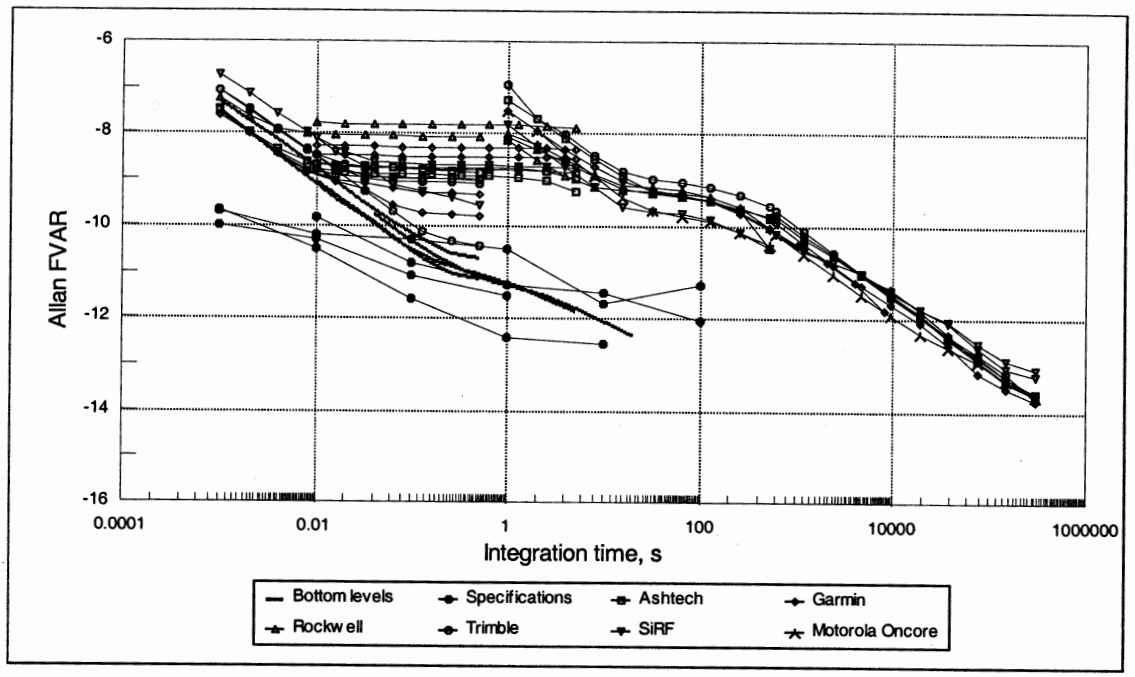


Fig. 6a: Noise bottom levels and all Allan FVARs

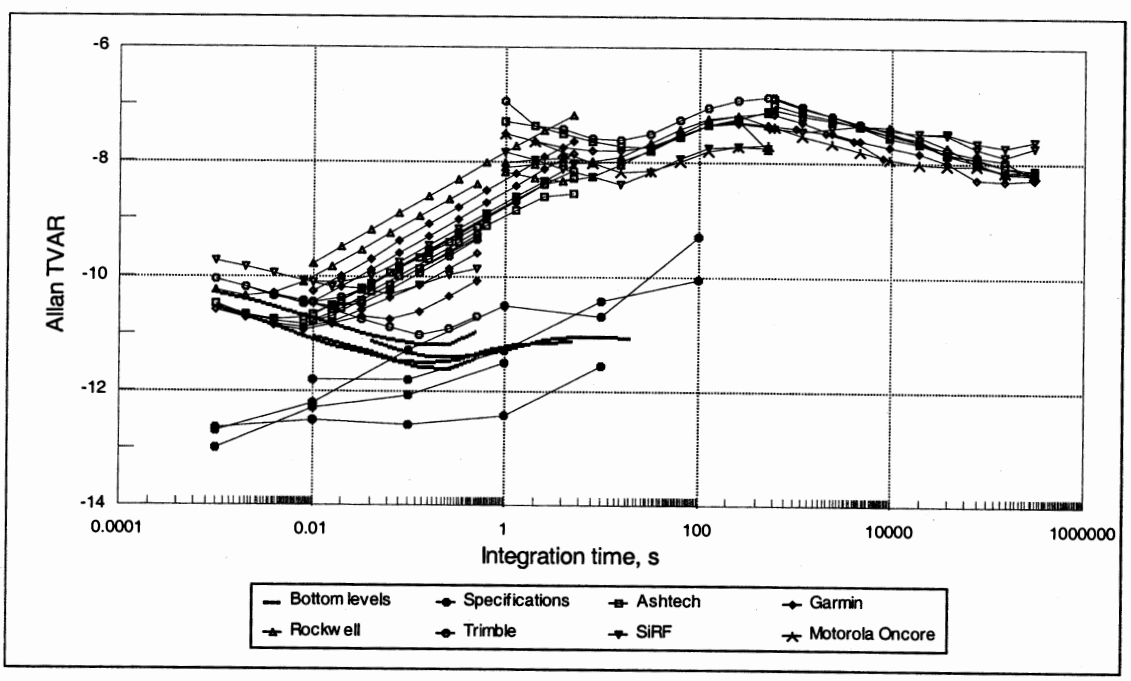


Fig. 6b: Noise bottom levels and all Allan TVARs

Moreover, noise powers were estimated by standard deviation (SD / variance) from randomly chosen samples (1 s and 10 s). The results agreed within a reasonable scale with the values achieved by Allan variance calculations (Figs. 6a, 6b).

6. CONCLUSIONS

The noise bottom level of the short-term set-up and five GPS receiver crystals were measured. The results show that all crystals stayed below the noise bottom level of

the set-up up to 5 ms integration time. Above 5 ms the observed Allan frequency variances remained near the value of $1 \cdot 10^{-9}$. The observed differences were probably caused by short term environmental changes, especially temperature. The observed behaviour at integration times <<1 s is mainly due to the crystal itself. In the locked mode above 1 s the behaviour is increasingly dominated by the control loop of the GPS receiver.

The Allan frequency variances (Figs. 6a and 6b) are nearly constant at integration times between 10 ms and 100 s. After 100 s they begin to fall about 1.5 decades/decade of integration time, reaching $5 \cdot 10^{-14}$ at 3 day integration time. The respective Allan time variance above 1 s integration time is better than 100 ns. Stability properties of the studied GPS receivers do not notably differ from each other. Their short-term stability is a couple of decades poorer than the specifications of OCXOs (oven controlled crystal oscillators) from litterature. Effective temperature compensation is probably hardly ever used in the crystals of GPS receivers.

One oscillator was cooled with short, local ventilation pulses while measuring the PPS pulse and the crystal phase. The immediate relative frequency changes varied from $2 \cdot 10^{-7}$ to $1 \cdot 10^{-6}$. The frequency recovered with about 30 s time constant. It did not necessarily return to the value it had before the cooling.

The results show that careful mechanical and electrical construction and thermal design are important factors affecting the short-term stability and overall performance of the GPS receivers.

---

(1) The metrological part of VTT Automation, Measurement Technology was separated and transferred to MIKES, Electricity and Time Metrology at the end of the year 1999.

7. REFERENCES

/1/ Mannermaa, J., Kalliomäki, K., Mansten, T., Turunen, S., "Timing Performance of Various GPS Receivers", Proceedings of the 13th European Frequency and Time Forum, pp. 287-290, Besançon, France, 1999.

/2/ Mannermaa, J., Kalliomäki, K., Mansten, "Long Term Stability of High Quality OCXOs", Proceedings of the 12th European Frequency and Time Forum, pp. 231-233, Warsaw, Poland, 1998.

/3/ Kalliomäki, K., Mansten, "SA and Control of Local Oscillator in GPS Receivers", Proceedings of the 5th European Frequency and Time Forum, pp. 211-213, Besançon, France, 1991.

OPERATIONAL STATUS OF THE IEN TWO-WAY SATELLITE SYNCHRONIZATION SYSTEM

M. Mascarello, F. Cordara, V. Pettiti, and P. Tavella
Istituto Elettrotecnico Nazionale "Galileo Ferraris"
Strada delle Cacce, 91 – 10135 Torino, Italy
Tel +39 011 39191, Fax +39 011 346384, E-mail: masca@tf.ien.it

ABSTRACT

To participate in the international synchronization network based on INTELSAT 706 satellite, a VSAT station designed for a two-way communication system operating in the Ku-band frequency range, has been completed at the IEN. A MITREX (Microwave Timing and Ranging Experiment) modem, modified for remote PC control, has been integrated in the measurement system together with a signal switching unit that will allow to monitor the station delays at each measurement session. The whole measurement set-up can be automatically controlled by a customized software that also provides a data file output in agreement with the ITU-R recommended standard.

A characterization of the noise associated with the measurement setup of the modem transmit and receive delays, by a suitable cross-connection of 3 modems and a statistical noise decoupling technique, was also performed.

The paper reports on what described above and gives some preliminary synchronization results.

1. INTRODUCTION

In 1997 IEN began to set-up a two-way station to participate in the international synchronisation network based on INTELSAT 706 satellite between European and American timing laboratories. Plenty of literature on this experiment and on the related activities

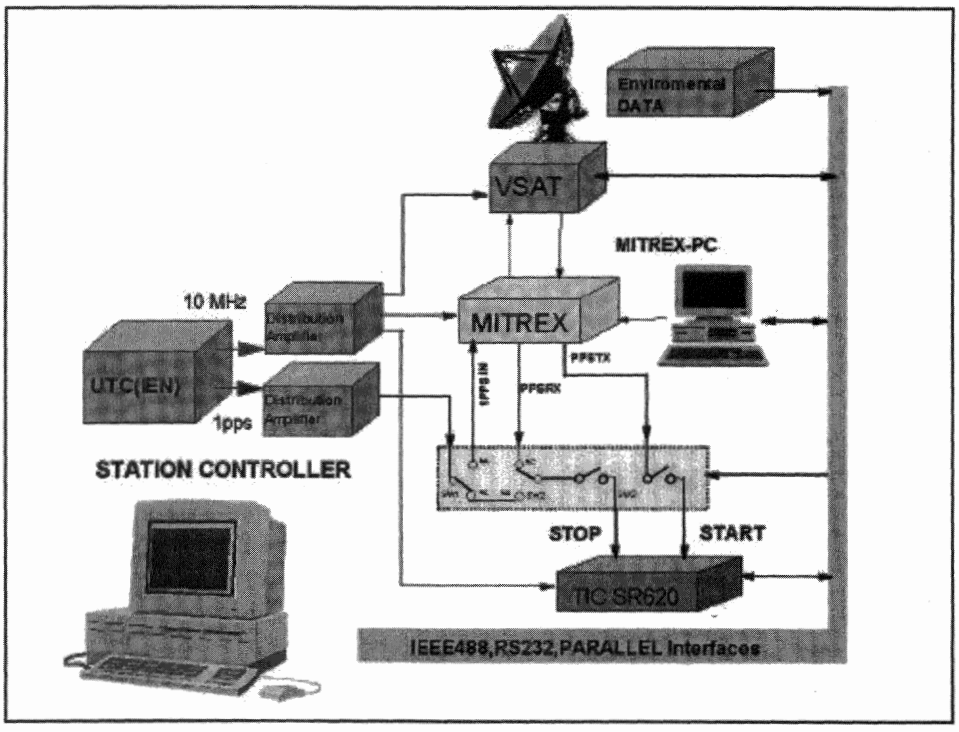


Figure 1 - IEN TWSTFT station set-up.

within the TWSTFT/CCTF Working Group is available in [1].
After some preliminary tests, the VSAT station had a failure so the work was stopped until the transmitter part was repaired. In the meantime contacts were made with the INS (Institute of Navigation University of Stuttgart) to modify the MITREX modem available at IEN to implement a remote PC control, allowing to automatize the measurement process.
At the beginning of 1999, all the parts were ready to operate, so the construction work was restarted and the TWSTFT station and the associated software were completed. During 1999, all the necessary tests to obtain the authorisations to transmit form the Italian Ministry of Communications, and the contacts with Telecom Italia, the Italian INTELSAT provider, were established and a completely new automatization and data analysis software were developed. A fully operative status of the station has been reached at the beginning of 2000.
In the followings are given detailed descriptions of the equipment set-up at IEN, of the software developed and are reported some results obtained in the stability characterisation of the instrumentation used. To test the performance of the IEN station, some measurement are being performed with the Technical University of Graz; some first experimental results are also reported.

2. THE IEN TWSTFT STATION

The IEN TWSTFT station setup is shown in Fig.1 and is composed by:

- SSE Technologies transceiver mod. ASAT1214/AB;
- 1,8m offset antenna Prodelin mod. 1194;
- MITREX modem mod 2500A-PC modified version;
- 386 PC using MsDOS for MITREX control;
- Stanford Research mod. SR620 Time Interval Counter (TIC);
- Pentium PC using Windows95 for the station control;
- Environmental data acquisition system;
- 10 MHz signal distribution amplifier;
- signals switching unit (PC controlled).

Apart from the VSAT station, all the equipment is placed in the IEN Time and Frequency laboratory where the temperature is stabilised within 2°C.
The VSAT station operates in the 14 / 12 GHz band and can be remotely controlled by the station main TW PC via a serial interface. Its oscillators are locked to an external 10 MHz reference and the operating frequency can be set with a 1kHz resolution. The transmitting power is adjustable by a programmable attenuator and the maximum output power is 4W.

The 1,8m offset antenna can be manually pointed and has a transmission gain of 45,2 dBi and a receiving one of 46,7 dBi.
The MITREX 2500A-PC modem is a modified version of the widely used MITREX modem allowing to have a remote PC control. This MITREX version, shown in Fig 2, modified by the INS doesn't have any keypad and display, but only a flat cable connection direct to a PC ISA card.

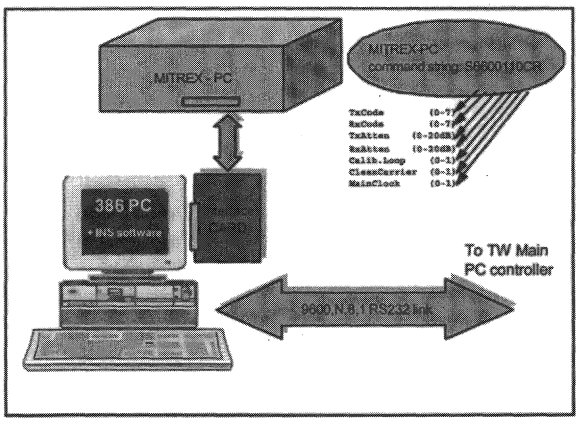


Figure 2 - MITREX-PC interfacing block diagram.

In this way, using a dedicated software written in Pascal and running under MsDOS [2], it is possible to have a remote control of the modem. This software permits also to interface the MITREX-PC to the TW main station controller by a serial port, thus controlling also the modem configuration changes needed during the measurement sessions. Nevertheless, to simplify the developed software, it has been necessary to dedicate a second 386 PC to interface the MITREX-PC to the main TW controller by a simple serial link.
The measurement system of the station includes a time interval counter SR620 interfaced to the TW main PC controller by an IEEE488 card.
All the operations of the TW station are controlled by a Pentium 100 PC running Microsoft Windows95 and the software has been developed with the Visual Basic™ programming language.
The PC internal clock is synchronised to UTC(IEN) by a serial connection to the IEN/CTD (telephone time code) generator to provide the necessary timing for all the operations during the measurement sessions.
The environmental data (external temperature, rel. humidity, pressure) are collected using the Time and Frequency Laboratory data acquisition system.
Also the MITREX modem and the SR620 TIC are locked to the UTC(IEN) 10 MHz reference through a distribution amplifier.
A switching unit controlled by the main TW PC via the parallel port, to manage all the signals involved in the

two-way measurements and future delays measurement satellite simulator, completes the set-up.

3. TWSTFT SOFTWARE

The TWSTFT network, Fig.3, is based on measurement sessions of two minutes, where couple of laboratories compare their clocks using a different code from the 8 available in the spread spectrum modems. In this way several couples of laboratories can make measurements at the same time.

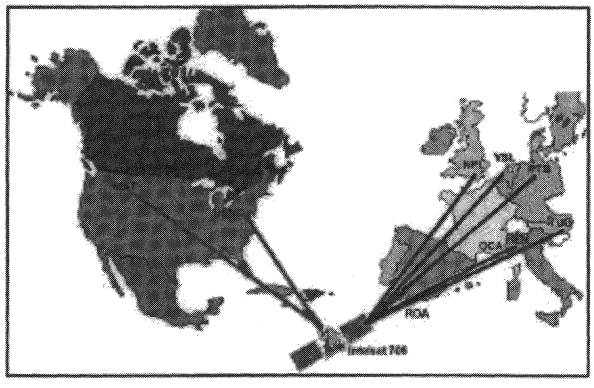


Figure 3 - TWSTFT INTELSAT 706 network.

To do that, it is necessary to program the TW station to operate the transceiver, to change the modem configuration every 2 minutes, to acquire the 1s time interval data from the counter and the environmental data of the session.
To have the possibility to make the suitable changes for the different kind of measurements, a TW software must be configurable for all necessary parameters and operations.
To respond to these requirements it has been developed a dedicated automatization software using Microsoft Visual Basic and running under Windows95. It has been choose because of the easy user interface, also if many timing and crashing problems due to the operating system had to be solved.
The TW software is similar to a command interpreter, in fact it reads a scheduling file where a sequential list of commands to be executed and the associated time is reported. All the configurations and scheduling files are in text form, so it's easy to change everything with a simple text editor. The main window of the software shows all the TW station parameters like operating frequencies, VSAT status, MITREX status and, in a monitor window, are reported the command executed and the results obtained. Some programmable flags permit to use the program in test mode and to change the days of measurement. At the end of every 2 minutes session the collected data are saved according to the ITU-R format.

Thanks to the signal switching unit commanded via the parallel port, the software can also measure the MITREX PPS(TX) – UTC(IEN) delay and report it in the data file.
The PC internal clock is synchronised to UTC(IEN) before starting the usual TW session at 14 UTC in order to execute the commands at the right time and assure the correct time tagging of the data acquired from the counter.
At the end of a TW session it is necessary to analyse the measured data, calculate the quadratic fit and generate a data report according to the ITU-R format.
For this purpose another software that reports in a graphic form the measured data, computes the residuals of the experimental data and the quadratic fit, eliminates the data outliers and generates the IEN TW file has been developed. Moreover, with another function it is possible to calculate the differences between the clocks of the laboratories. For this last computation it is necessary to exchange the TW result file with the other laboratory using e-mail or FTP. A future release of the software will permit to manage the FTP files exchanges as an internal function without using external software.
At the moment the analysis software uses the complete TW formula [3] to compute the clock differences but doesn't use all the data like ionospheric and or troposferic delay corrections yet.

4. EQUIPMENT CHARACTERIZATION

During the implementation of the station, the characterisation of the equipment used has been performed.
In the context of a degree thesis in 1997 [4], were investigated the instabilities of three electronic counters mod. SR620, of a 2500A MITREX modem and some distribution amplifiers.
As far as the TIC characterisation is concerned, in Fig 4 a typical Modified Allan Deviation of 120 measurements, 1s apart, of the delay of a coaxial cable is represented. It can be seen that the instability is $< 2 \cdot 10^{-11} \cdot \tau^{-3/2}$ mostly due to a white phase noise.

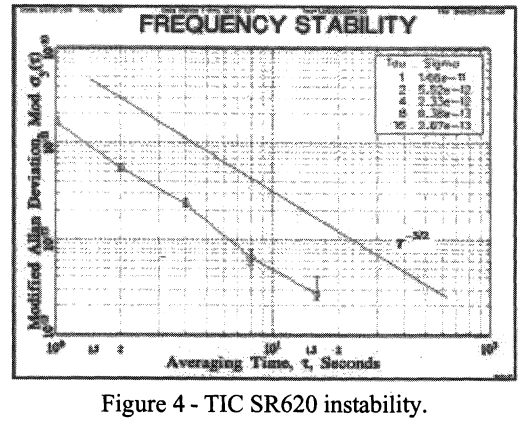


Figure 4 - TIC SR620 instability.

The investigations about the MITREX, concerned the instability of the PPS(TX) output versus UTC(IEN) delay for measuring times from 100 s to 2,5 days and the results are shown in Fig. 5. The effects of the temperature versus the 70 MHz station IF loop is reported in Fig.6; it can be seen that there is a correlation evaluated at level of – 0,1 ns/°C. The temperature effect is probably due to the outside part of the cable connections, since the temperature coefficient of the MITREX only was of the order of 0,02 ns/°C.
This fact confirms the necessity of calibrating the differential delays of the station at every TW session.
In 1999, the chance of having three MITREX 2500A modems available in the laboratory, allowed to measure the transmitting and receiving delays by suitable cross-connections between the modems. Since the measurement scheme didn't allow the estimation of the separate Tx and Rx delays, using a noise de-coupling technique similar to the known N-cornered hat method, the separate instabilities of the receiving and transmitting sections were evaluated only. The most interesting finding was that the receiving delay (TDEV[τ=1s] ~ 70ps) is more stable than the transmitting one (TDEV[τ =1s] ~ 150ps), probably because of the presence of a 70 MHz band pass filter in the transmitting chain.

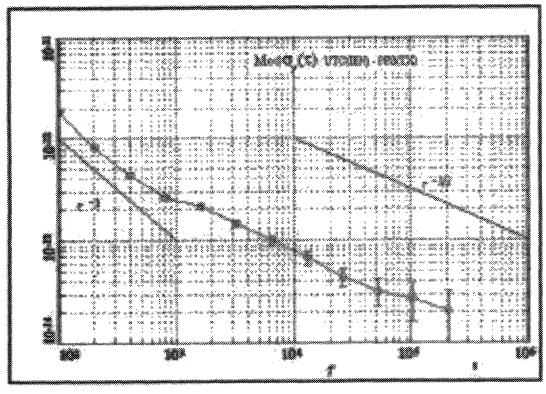


Figure 5 –MITREX time reference instability.

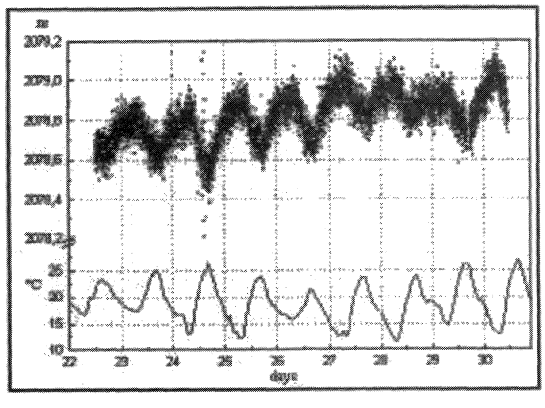


Figure 6 – TW station 70 MHz IF loop delay vs. outside temperature.

To characterise the whole station behaviour, some investigations were performed in February 2000 and, in Fig 7, it is reported the MDEV of the residuals of the ranging measurements respect to a quadratic fit, obtained changing the transmitted power. Practically it is showed, as expected, that the stability of the station is a function of the signal to noise ratio of the received signal.
As regards to the 10 MHz distribution amplifiers, the measured instability was 5 ÷ 10 times lower than that of the MITREX modems.

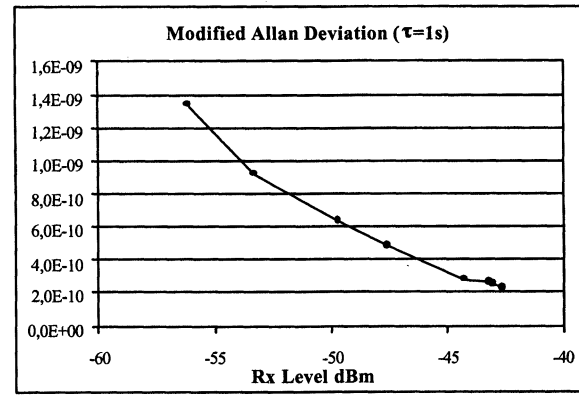


Figure 7 –TW station instability vs. S/N ratio.

5. PRELIMINARY RESULTS

Thanks to the collaboration of the Technical University of Graz, some measurements were made since November 1999 to test all the station set-up and to check and debug the software.
In Fig 8 are reported the first results of these comparisons between UTC(IEN) and UTC(TUG) using the TW method, the GPS-CV and the BIPM Circular T data. To compute the TW differences, the parameters of the quadratic regression found in the TW data files, have been entered in the ITU-R complete formula but, corrections were applied for *Earth Rot. Corr.* and *Ref. Delays* only. It can be seen that the agreement between the Circular T data and the GPS-CV computed at IEN is very close meanwhile there is a systematic difference versus the TW data due to the fact that no correction for the stations delays has been applied. A first estimate of its differential value has been performed considering three TW time differences between IEN and TUG in coincidence with the Circular T standard dates, namely MJD 51494 (1999 Nov. 12) , MJD 51534 (1999 Dec. 22) , MJD 51569 (2000 Jan. 26) and computing their mean value and the standard deviation that has been found of 262,9 ns ± 2,4 ns (1σ).
The TW results of Fig. 8 have been consequently corrected to obtain the coincidence with the Circular T data.

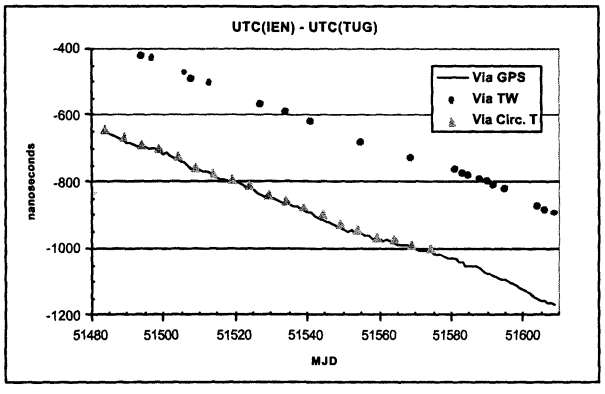


Figure 8 – UTC(IEN) vs. UTC(TUG) via GPS, TW and BIPM Circ. T (November 2, 1999 – March 6, 2000).

The same set of data with the time offset (-659,3 ns) and rates (-4,0 ns/d) between the two clocks removed, are shown in Fig. 9.
Finally, applying the same interpolation criteria used by BIPM in the monthly reports [6], the residuals of the differences between the TW and the Circular T data have been computed and a mean value of –0,3 ns with a standard deviation of 1,7 ns has been found.

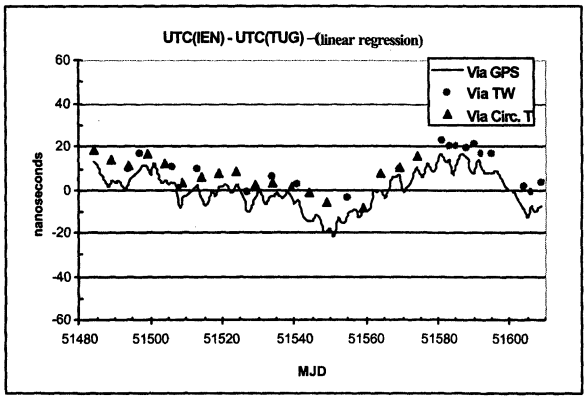


Figure 9 – UTC(IEN) vs. (UTC(TUG) corrected for the estimated TW differential delay.

6. CONCLUSIONS

The IEN TWSTFT station is now ready to be operative and it is to add in the near future the necessary hardware to monitor the differential station delays. Next April it is foreseable that the omologation and approval of the station will be finished, so that the IEN can join on a regular basis to the international TW measurement network.

7. ACKNOWDLEDGEMENTS

The authors wish to thank Dr. D. Kirchner and Mr. R. Robnik of TUG for the helpful cooperation and suggestions. The MITREX modems used have been provided by Ing. G. Tignanelli of the Istituto Superiore delle Comunicazioni e delle Tecnologie dell' Informazione – Roma in the frame of a scientific cooperation project.

8. REFERENCES

[1] Proceedings of PTTI (Precise Time and Time Interval Meeting) and EFTF (European Frequency and Time Forum) meetings.

[2] M. Thomas: "*MITREX CONTROL Ver. 10*" INS, April 1996.

[3] ITU-Recommendation TF1153, October 1996.

[4] M. Sturniolo, M. Raffa: *"Time scales comparison using the Two-Way method"* degree thesis in Telecommunication Engineering, Politecnico di Torino, December 1997, *"Confronto tra scale di tempo mediante tecnica a due vie"*.

[5] M. Mascarello, P. Tavella, V. Pettiti and F. Cordara: *"Evaluation of MITREX modems transmit and receive instability"* 31th Precise Time and Time Interval Meeting, Dana Point (CA), December 1999.

[6] 10 th BIPM TWSTFT Monthly Report, March 2000.

PROGRESS REPORT ON THE TRANSATLANTIC TIME TRANSFER BY GETT

G. Dudle, L. Prost
Swiss Federal Office of Metrology, Lindenweg 50, CH-3003 Berne-Wabern
gregor.dudle@eam.admin.ch

R. Dach, T. Springer, T. Schildknecht
Astronomical Institute of the University of Berne, Sidlerstrasse 5, CH-3012 Berne

ABSTRACT

We report on the progress of the transatlantic GeTT campaign for time and frequency transfer by GPS Carrier Phase (GPS CP). This experiment lasts already for almost two years and will probably be extended another year. The comparisons with other, independent time transfer methods, namely TWSTFT and data from Circular T, have been continued. The sine-variation of the difference between GeTT and TWSTFT reported last year could, however, not be confirmed. To overcome the day boundary problem, we describe an alternative data processing technique, involving overlapping batches of data.

1 INTRODUCTION

GPS Carrier Phase (GPS CP) has been presented several years ago as new technique for time and frequency transfer with considerable potential to improve the results of GPS Common View (GPS CV) [1]. Numerous groups are now studying various aspects of the technique such as local delays variations or short term stability over small and medium baselines [2, 3, 4]. Hardware issues are addressed and different software packages are used to process the data. The interest in GPS CP as time transfer method has also lead to the joint pilot project of the BIPM and the International GPS Service (IGS) [5].

The Astronomical Institute of the University of Berne (AIUB) and the Swiss Federal Office of Metrology (OFMET). AIUB and OFMET have been working on GPS CP time transfer for some years already and developed a technique under the name Geodetic Time Transfer (GeTT). Local delays and their variations were the main interest in former studies [6]. Another important aspect, especially for time transfer, is the long term behaviour of the method. For this purpose, OFMET and AIUB have initiated a GPS CP experiment between the Physikalisch Technische Bundesanstalt (PTB) in Braunschweig, Germany and the US Naval Observatory (USNO) in Washington DC. The campaign has been set up in summer 1998 and first results have been presented at the EFTF 1999 [8]. Here we report on the progress of the work and new insights after 20 months of almost uninterrupted operation.

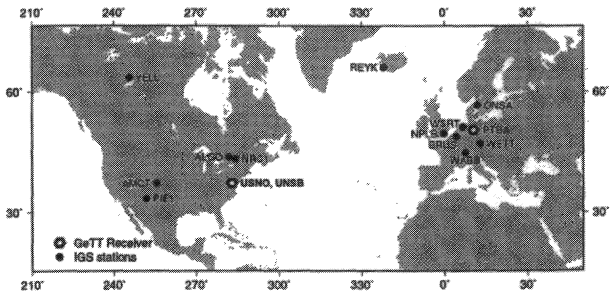


Figure 1: The locations of the stations forming the small subnet for the time transfer experiment

Section 2 briefly recalls the main aspects of a GeTT campaign. Updated results of time transfer are presented in section 3. In section 4, finally, we discuss the so called day boundary problem of GeTT.

2 GETT TERMINAL AND PROCESSING

For a GPS CP time transfer campaign a geodetic type receiver is installed on every participating site. For the GeTT experiment modified ASHTECH Z12 T receivers were used. The modification allows to replace the internal clock by an external reference (20 MHz signal and 1 PPS). These receivers are integrated in a dedicated time transfer terminal. The central part of the terminal is a thermally stabilized box in which the receiver and auxiliary electronic boards are placed to minimize the changes of the internal delays due to temperature variations. A commercial temperature control regulates the temperature to within 1 K. Another important part of the station is the data acquisition PC. Automated processes are scheduled to download the observations

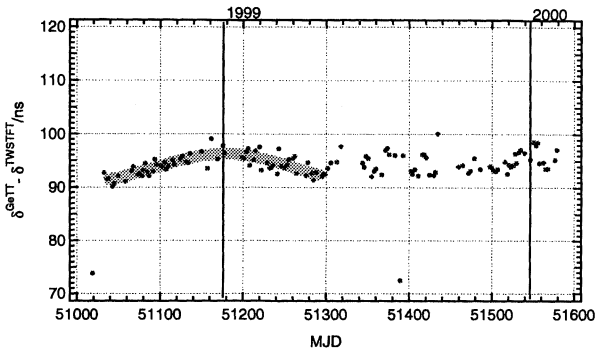


Figure 2: Comparison between time transfer by GPS CP and time transfer by TWSTFT between USNO and PTB.

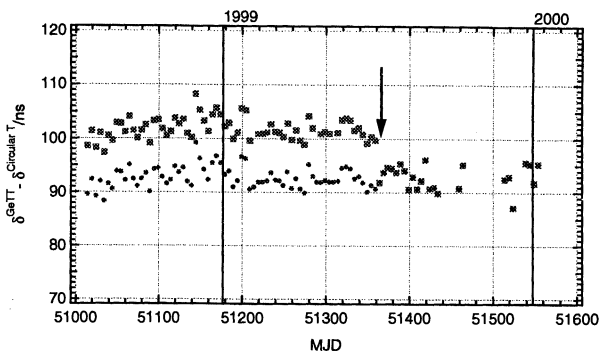


Figure 3: Comparison between time transfer by GPS CP and the Circular T published by BIPM. Squares show the raw data, for the circles the new ionospheric correction has been taken into account for all data points.

from the receiver every day and to send them by Internet to the processing center (see below). The PC is also logging various temperatures such as antenna-, receiver- or room-temperature. GeTT terminals are also equipped with a UPS to ensure continuous recording of the observations during power failures. Two terminals are involved in the transatlantic campaign. A third terminal is assembled and currently tested at OFMET. More details on the GeTT terminals can be found in [6].

The processing of the data is carried out at the AIUB which hosts the Center for Orbit Determination in Europe (CODE) for IGS. According to the IGS policy, data are sent by the observing stations to the data centers. Once a day all stations submit all observations by FTP in RINEX format (Receiver INdependent EXchange format). AIUB computes for each epoch the difference of the clock reading of a given station with regard to an arbitrarily chosen reference station, taking full advantage of IGS products, such as precise orbits or ionospheric delays. For most of the work presented here PTB was chosen as reference station. Those results are delivered in the RINEX-clock format [7]. Rapid products become available on the day following the observations, final CODE results are provided with a delay of 5 days.

For the time transfer experiment, a small subnet of IGS stations has been defined at the AIUB. The locations and the IGS labels are depicted in Fig. 1.

3 TIME TRANSFER

Early results of the transatlantic GeTT campaign have been published at last year's EFTF [8]. Here we concentrate on the long term character of the experiment. In particular, it is of interest, how the results compare to those of other, independent methods. As both, PTB and USNO, participate in Two Way Satellite Time and Frequency Transfer (TWSTFT) sessions on a regular basis (each week on Monday, Wednesday and Friday at 14:22 UTC), the results of GPS CP can be compared with those of the TWSTFT link.

In Fig. 2 we show $\Delta\delta = \delta^{\mathrm{GeTT}} - \delta^{\mathrm{TWSTFT}}$ vs. MJD where δ^i =UTC(USNO)−UTC(PTB) as measured by method i. Data outages stem from the absence of results from one or the other method. As far as GeTT is concerned, the most important gap of data was introduced around MJD 51420 when the firmware of the receivers had to be upgraded.

Note that there is no calibration of the differential station delay available for the two GeTT terminals. This explains the important offset between the two methods and, thus, that $\Delta\delta$ is not close to zero. We focus therefore not on the value of $\Delta\delta$ but on its stability throughout the 20 months of operation.

In Fig. 2 we can distinguish two different regimes for $\Delta\delta$. For the first half year of the campaign, the difference between the two methods seems to exhibit a seasonal variation. It is possible to fit a sine-wave through these data points having a period of 330 days and an amplitude of 2.7 ns. As already discussed in [8], the temperature dependent delay variations of GeTT are not sufficiently large to explain the order of magnitude of the sine-wave. Beyond MJD 51300 this behaviour is no longer observed, but at the same point the spread of the difference increases. At present, it remains unclear whether the sine variation can be ascribed to one of the methods and correlates with another observable or is simply noise. As far as GeTT is concerned, no hard- or software modification that could explain the suppression of the seasonal term occurred around MJD 51300.

With the help of a third method one could possibly identify the source of variations. As both data sets

are reduced to the local UTCs, we have tried to check the results with the Circular T data published by BIPM [9]. Fig. 3 shows the comparison between time transfer by GeTT and time transfer by Circular T (squares) over the same period than displayed in the previous figure.

Again, the unknown differential delay calibration for the GeTT-terminals gives rise to an important offset between the two results. The spread of the data is unfortunately too large to rule out a sine variation of the difference GPS CP - Circular T with an amplitude of 2.7 ns. On this basis, it is therefore not possible to incriminate or exclude GPS CP as source of the seasonal term. But another, unexpected step was detected by this difference. On MJD 51364 BIPM modified the ionospheric correction for the link between America and Europe. For the comparison PTB - USNO this adjustment amounts to 9 ns. As no change was introduced for the processing of the GeTT data, a step is clearly visible at this particular MJD. For Fig. 3 we have taken into account this correction also for comparisons before MJD 51364 (circles). The corrected data set shows no structure but is unfortunately too noisy to draw any conclusions about the 'sine term'.

No clear structure can be made out in this new set of data, but unfortunately the results are too noisy would hide a possible sine-dependence with time.

4 DAY BOUNDARY PROBLEM

For the processing of GPS data both, pseudorange and carrier phase observations are used. While the latter give the high resolution of the increment of the clock reading between two adjacent epochs, the former help mainly to determine the absolute level of the estimated clock parameters. In the Bernese Software a least square adjustment is performed for all parameters estimated, including those of the clocks. Continuous data analysis such as e.g. using the Kalman filter approach is, therefore, not possible. Every day has to be analyzed separately, independent from the preceding days. Depending on the noise of the pseudorange observations, the absolute level of the estimated clock parameters can vary slightly from one day to the next. This translates into steps at the day boundaries.

To investigate possible solution to this day boundary problem, we have processed two weeks of February 1999 in two different ways. First, as for the routine processing, batches were formed with all observations from 0 h to 24 h UTC of a given MJD. For the second processing, the batches were regrouped to contain all observations between 12 h of MJD and 12 h of MJD + 1. The results of both calculations are represented in Fig. 4. The lighter gray shows the results of the batches 12 h to 12 h, the 0 h to

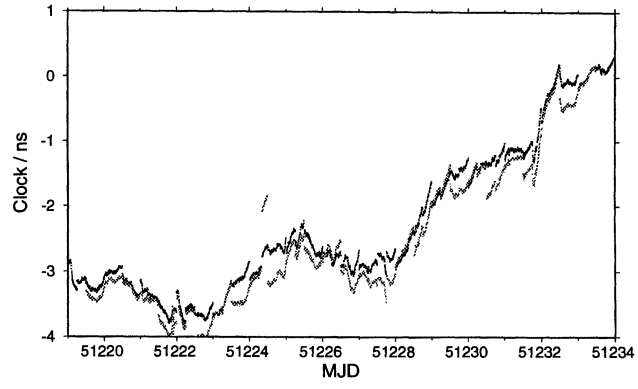


Figure 4: Processing with batches regrouping observations between 0 h and 24 h UTC and with batches grouping observations between 12 h of day MJD and 12 h of day MJD + 1

24 h solutions are given by the dark dots. Analyzing the same observations twice, regrouping being the only difference, one would expect similar results. In reality, small offsets are observed between the two solutions as the absolute value of the clock parameters are determined by the noisy pseudorange observations. For the chosen interval, the offsets remain well below the 1 ns limit.

We are currently testing to what extend a combination of the two solutions could reduce further the steps at the day boundaries.

5 CONCLUSION

Three units for time and frequency transfer by GPS CP—also known as GeTT terminals—are now operational. Two of these stations have been used successfully in a long term experiment lasting for more than 20 months. The campaign is still ongoing. The terminals are fully automated and require a minimum operator interaction.

Throughout the campaign, frequent comparisons between GPS CP on one side and TWSTFT on the other have been performed. The annual term in the difference of time transfer by GPS CP and time transfer by TWSTFT observed over the first year of operation could not be confirmed. The agreement between these two independent methods remains within a few ns.

The comparison with Circular T shows clearly the 9 ns step caused by the adoption of a new ionospheric correction scheme for the values published by BIPM.

To improve even further the quality of the results, especially at the day boundaries, a different approach for the data processing, involving overlapping time series is currently under investigation.

ACKNOWLEDGMENT

The authors wish to thank both institutes, PTB and USNO for accepting and maintaining the GeTT-terminals in their labs. In particular we are grateful to D. Matsakis and P. Hetzel and their co-workers for many helpful discussion and advice.

REFERENCES

[1] Th. Schildknecht *et al.*, "Towards sub-nanosecond GPS time transfer using geodetic processing technique," in *Proc. of the 4th EFTF, 13 - 15 March 1990, Neuchâtel, Switzerland*, 1990.

[2] C. Bruyninx *et al.* "Frequency Transfer using GPS : a comparative study of code and carrier phase analysis results," *Proceedings of the 30th PTTI Meeting, 1 - 3 December 1998, Reston, Virginia*, 1998.

[3] G. Petit *et al.* "Use of GPS ASHTECH Z12T receivers for accurate time and frequency comparisons," *IEEE Transaction on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 46, no. 4, pp. 941 - 949, 1999.

[4] G. Petit *et al.* "Differential calibration of ASHTECH Z12-T receivers for accurate time comparisons," *These proceedings*

[5] Updated informations can be found on the internet http://maia.usno.navy.mil/gpst

[6] F. Overney *et al.*, "GPS time transfer using geodetic receivers (GeTT): Results on a European Baseline," in *Proc. of the 12th EFTF, 1998, Warsow, Poland*, 1998.

[7] A description of the data format can be found at ftp://igscb.jpl.nasa.gov/igscb/data/format/rinex_clock.txt

[8] G. Dudle *et al.*, "Transatlantic time and frequency transfer by GPS carrier phase," in *Proc. of the 13th EFTF, 13 - 16 April 1999, Besançon, France*, 1999.

[9] Circular T data are available on BIPM's homepage ftp://62.161.69.5/pub/tai/publication/

Atomic Collimation with a Single Laser Pulse

Masatoshi Kajita* and Takao Morikawa
Communications Research Laboratory
4-2-1 Nukui-Kitamachi, Koganei, Tokyo 184-8795, Japan
*E-mail address: kajita@crl.go.jp

1. Introduction

There are two forces which laser light field exerts on atoms: the scattering force and the optical dipole force. The scattering force is given by the transfer of photon momentum when photons are absorbed or emitted by atoms. It is not a conservative force and has mainly been used for laser cooling. The combination of laser-Doppler cooling and polarization-gradient cooling can reduce atomic temperature to less than 10 μK.[1] With this combination of methods, however, the temperature cannot become lower than the photon recoil energy. To reach temperatures lower than the photon recoil energy, researchers have developed the velocity-selective coherent population trap (VSCPT)[2] and Raman cooling.[3] The scattering force is also used to accelerate the laser-cooled atoms.[4]

The optical dipole force is caused by the interaction between an induced atomic dipole moment and a light electric field. It makes a potential, called optical potential. The optical potential has been mainly used for atomic trapping. Using far-detuned lasers the heating effect of the spontaneous scattering can be reduced[5].

In this paper we describe a new method to reduce the velocity components transverse to the laser beam using the optical potential. Because the optical dipole force is a conservative force, it is not possible to reduce the phase space volume. However, it is possible to obtain an narrower velocity distribution at the expense of broadening the position distribution. If the laser is switched off when the broadening of the position distribution is maximum, a narrow distribution of velocities can be obtained. When a far-detuned laser is used, it is possible to obtain the sub-recoil temperature using a simple apparatus.

As an example, we have calculated the distribution of the velocity components of cesium atoms after irradiation by the collimating laser. It is shown that the velocity component transverse to the laser beam can be reduced to several mm/s.

2. Analysis with a One-Dimensional Model

The potential ($U(x,y,z)$) provided by the interaction between an induced atomic dipole moment and a light electric field (optical potential) is described by

$$U(x,y,z) = U_0\, I(x,y,z) \tag{1}$$

$$U_0 = -\alpha E_0^2/2\,, \tag{2}$$

where α is the atomic polarizability and E_0 is the amplitude of the light electric field at the center of beam. When the laser frequency is in the infrared (IR) region, α can be approximated by the static value (for a Cs atom, 59.6 A^3).[3] $I(x,y,z)$ denotes a power density distribution in which the laser light propagates in the z-direction and U_0 is the depth of the optical potential. Here, we discuss a case when $I(x,y,z)$ is given by

$$I(x,y,z) = \exp\left[-x^2/(\Delta x)^2\right]. \tag{3}$$

The equation of motion for the atom is then described by

$$\frac{d^2x}{dt^2} = -\omega^2 x \exp\left[-x^2/(\Delta x)^2\right] \approx -\omega^2 x \qquad (x << \Delta x)\,,$$

$$\omega^2 = -\frac{2U_0}{M(\Delta x)^2} = \frac{\alpha E_0^2}{M(\Delta x)^2}\,, \tag{4}$$

where M is the atomic mass.

Here we consider the case in which the laser is switched on at $0 < t < T$ and is switched off at $T > 0$. The temporal changes of x and v_x (velocity component in the x-direction) are described by

$$\begin{aligned} v_x(t) &= v_x(0)\cos\omega t - x(0)\,\omega \sin\omega t \qquad (0 < t \le T) \\ &= v_x(T) \qquad (t > T), \end{aligned} \tag{5}$$

$$\begin{aligned} x(t) &= x(0)\cos\omega t + \frac{v_0(0)}{\omega}\sin\omega t \qquad (0 < t \le T) \\ &= x(T) + v_x(T)(t - T) \qquad (t > T) \end{aligned} \tag{6}$$

At $T = \pi/2\omega$, $v_x(T)$ is $-x(0)\omega$. When $x(0)$ is zero, $v(T)$ becomes zero with any value of $v(0)$.

Here, we perform the numerical calculation of Eq. (4), with the following parameters:

$$\Delta x = 1.5 \text{ mm}, \quad T = 0.1\text{s} \quad (\omega = 2\pi \text{ x } 2.5 \text{ radian/s}). \tag{7}$$

Figure 1 shows $v_x(T)$ as a function of $v_x(0)$ for $x(0)$ = 0 mm and $\pm$0.5 mm,. When the optical potential is given by the harmonic potential, $v_x(T)$ (= $-x(0)\omega$) is 0 m/s when $x(0)$ = 0 mm and is $\mp$7.8 mm/s when $x(0)$ = $\pm$0.5 mm. Actually, this approximation is valid when $v_x(0)$ < 1 cm/s. When $v_x(0)$ is sufficiently small the atomic motion is harmonic oscillation at $0 < t \le T$.

Figure 2 shows the distribution of kinetic energy after the IR-laser pulse is applied under the conditions specified in Eq. (7). Here, we assume the following distributions of $v_x(0)$ and $x(0)$:

$$\rho_{v_x(0)} = \frac{1}{\delta v\sqrt{\pi}}\exp\left[-v_x(0)^2/(\delta v)^2\right]$$

δv = 1.8 cm/s (corresponds to 2.6 μK)

$$\rho_{x(0)} = \frac{1}{\delta x\sqrt{\pi}} \exp\left[-x(0)^2/(\delta x)^2\right]$$

$$\delta x = 0.25 \text{ mm.} \qquad (8)$$

The distribution of the kinetic energy is not in thermal equilibrium, but is approximately described by the temperature 110 nK. This corresponds to the approximate estimation by $\omega\delta x = 3.75$ mm/s (110 nK).

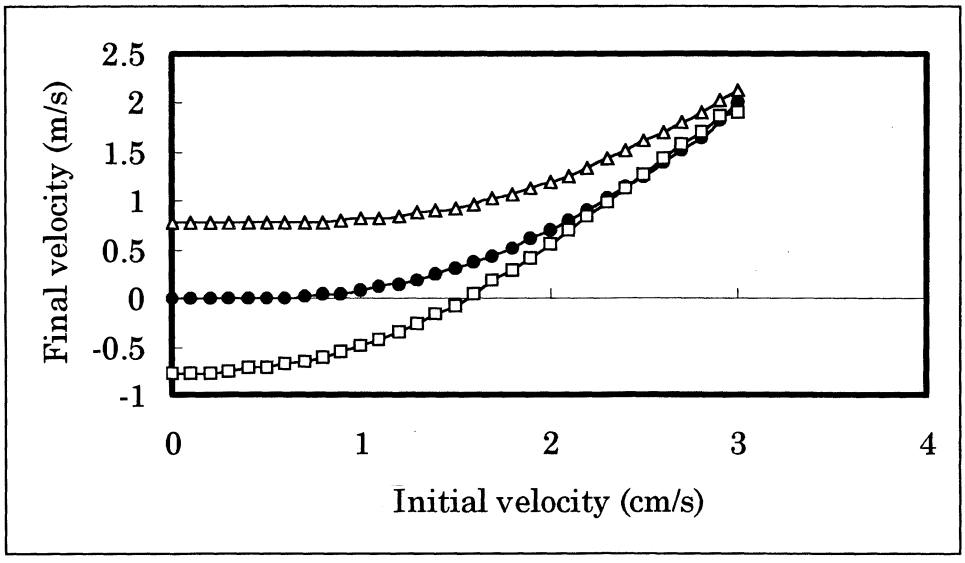


Fig. 1: Atomic velocity after the IR laser is applied under the conditions specified in Eq. (7). ●: $x(0) = 0$ mm □: $x(0) = 0.5$ mm △: $x(0) = -0.5$ mm.

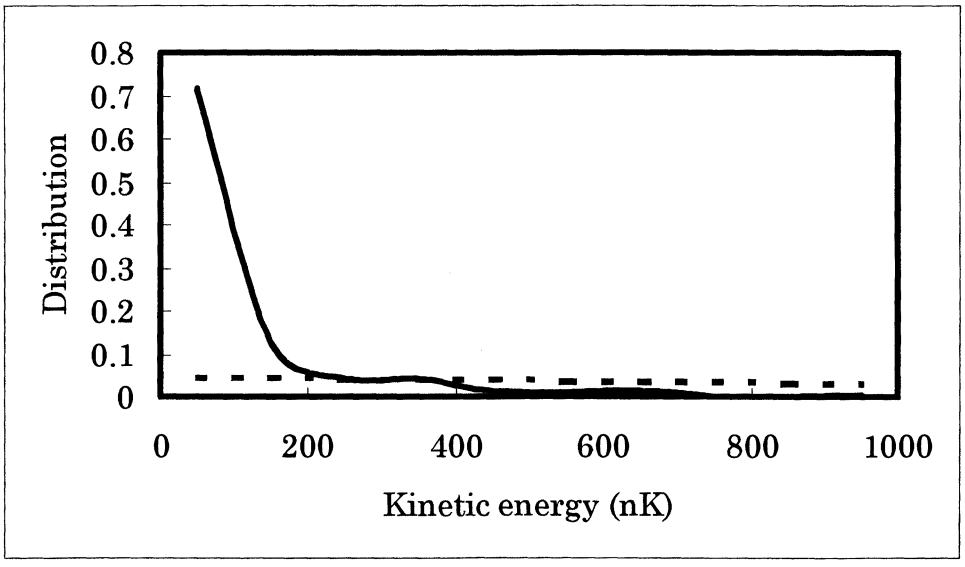


Fig. 2 : Solid line indicates the kinetic energy distribution after the IR laser is applied under the conditions specified in Eq. (7), with the one-dimentional model. Dotted line indicates the initial kinetic energy distribution (2.6 μK)

3. Analysis with a Two-Dimensional Model

It is convenient to describe the equation of motion with cylindrical coordinates (r,θ,z), when a two-dimensional model is used. Here we assume that the laser power density distribution is described as

$$I(r) = \exp\left[-r^2/(\Delta r)^2\right]$$

$$r^2 = x^2 + y^2 \tag{9}$$

When the centrifugal force is taken into account, Eq. (4) must be revised as

$$\frac{d^2r}{dt^2} = -\omega^2 r \exp\left[-\frac{r^2}{(\Delta r)^2}\right] + \frac{L^2}{r^3}.$$

$$\omega^2 = -\frac{2U_0}{M(\Delta r)^2} = \frac{\alpha E_0^2}{M(\Delta r)^2}$$

$$L = xv_y - yv_x \quad \text{(angular momentum).} \tag{10}$$

Note also that L cannot be changed by the optical potential field, which gives only the central force. The kinetic energy $K(t)$ is given by

$$K(t) = \frac{M}{2}\left[v_r(t)^2 + \left(\frac{L}{r(t)}\right)^2\right], \tag{11}$$

where v_r (= $(xv_x + yv_y)/r$) is the velocity component in the radial direction. Figure 3 shows the $K(T)$ distribution calculated with the following parameters:

$$\Delta r = 1.5 \text{ mm}, \qquad T = 0.1 \text{ s} \qquad (\omega = 2\pi \times 2.5 \text{ radian/s}). \tag{12}$$

The distributions of $v_{x,y}(0)$ and $r(0)$ are given by

$$\rho_{v_x(0),v_y(0)} = \frac{1}{\pi(\delta v)^2}\exp\left[-\frac{v_x(0)^2 + v_y(0)^2}{(\delta v)^2}\right]$$

$\delta v = 1.8$ cm/s (thermal equilibrium with 2.6 μK),

$$\rho_{r(0)} = \frac{2r(0)}{(\delta r)^2}\exp\left[-\frac{r(0)^2}{(\delta r)^2}\right]$$

$$\delta r = 0.25 \text{ mm}. \tag{13}$$

The distribution of $K(T)$ is roughly described by the thermal equilibrium with temperature 180 nK, which is lower than the photon recoil energy (200 nK). When a two-dimensional model is used, the kinetic energy is not reduced as much as when a one-dimensional model is used, because of the centrifugal force.

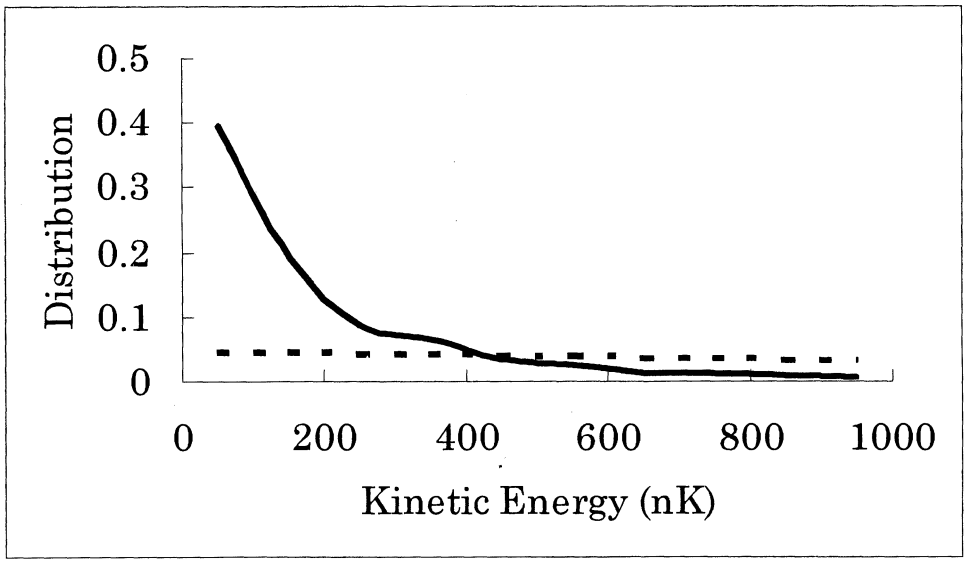


Fig. 3: Solid line indicates the kinetic energy distribution after the IR laser is applied under the conditions specified in Eq. (12) with the two-dimentional model. Dotted line indicates the initial kinetic energy distribution (2.6 μK).

4. Discussion

In this paper we describe a new optical method for obtaining a sub-recoil kinetic energy in one or two directions. VSCPT [2] and Raman cooling[3] have been used to get sub-recoil kinetic energies, but VSCPT can only be used for atoms with a certain quantum energy structure and Raman cooling requires a complex apparatus. Using a single IR-laser beam, we can obtain a sub-photon recoil kinetic energy with a simple apparatus. As the optical potential by the far-detuned laser does not depend on the quantum energy structure of atoms, this method can be used for any type of atom.

Equations (4), (5), and (6) are valid when U_0 is larger than the initial kinetic energy. To get a significant U_0, we require a large laser power. For example, the conditions specified in Eq. (12) ($U_0/k_B = 2.3$ μK) hold for cesium atoms when the laser power is 360 W. The laser line does not have to be a single laser line, as the optical potential does not depend on the laser line itself.

5. Application to the atomic fountain frequency standard

It is useful to apply this method to reduce the horizontal velocity component of the launched atoms in the atomic clock based on the atomic fountain[4]. After the launching, the diameter of the atomic cloud increases (at 3.6 cm/s when the atomic temperature is 2.6 μK). When the atomic cloud passes through the hole (diameter is typically 10 mm) of the microwave cavity at the second time (1 sec after the launching), the diameter is as large as 36 mm. Therefore only 8 % of the atoms can pass through the microwave cavity and fall

into the detection region. If the collimating laser is applied under the conditions specified in Eq. (12), the mean horizontal velocity component is reduced from 1.8 cm/s to 4 mm/s, while the diameter of the atomic cloud is expanded from 0.5 mm to 3 mm. The diameter of the atomic cloud is 11 mm when it passes through the microwave resonator; in this case more than 80% of launched atoms fall into the detection region and the signal-to-noise ratio of the spectrum is improved. There is no light frequency shift because the collimating pulse ends before atoms have their first interaction with the microwave field.

References

[1] J. Dalibard and C. Cohen-Tannoudji, "Laser cooling below the Doppler limit by polarization gradients: sinple theoretical models", J. Opt. Soc. Am. B, vol. 6, pp.2023-2045, November 1989.

[2] A. Aspect, E. Arimond, R. Kaiser, N. Vansteenkiste, and C. Cohen-Tannoudji, "Laser cooling below the one-photon recoil energy by the velocity selective coherent population trapping", Phys. Rev. Lett., vol.61, pp.826-828, August 1988.

[3] J. Reicel, F. Bardou, M. Ben Dahan, E. Peik, S. Rand, C. Salomon and C. Cohen-Tannnoudji, "Raman Cooling of Cesium below 3nK: New Approach Inspired by Levy Flight Statistics", Phys. Rev. Lett. vol.75, pp.4575-4578, December 1995.

[4] A. Clairon, P. Laurent, G. Santarelli, S. Ghezali, S. N. Lea and M. Bahoura, "A Cesium Fountain Frequency Standard", IEEE Trans. Ins. Meas., vol. 44, pp.128-131, April 1995.

[5] K. M. O'Hara, S. R. Granade, M. E. Gehm, T. A. Savard, S. Bali, C. Freed, and J. E. Thoma, "Ultrastable CO_2 Laser Trapping of Lithium Fermions", Phys. Rev. Lett., vol. 82, pp.4204-4207, May 1999.

DEVELOPMENT OF A DOUBLE–BULB RUBIDIUM MASER AND A COHERENT POPULATION TRAPPING RB CLOCK

D.F. PHILLIPS, E.M. MATTISON, R.F.C. VESSOT, and R.L. WALSWORTH
Harvard–Smithsonian Center for Astrophysics

Results will be presented from two new Rb atomic clocks under development: the double–bulb rubidium maser (DBRM) and a coherent population trapping (CPT) clock. The DBRM consists of connected quartz chambers: one for optical pumping of rubidium atoms and one where active maser oscillation can occur on the ^{87}Rb hyperfine transition. The bulbs are coated and have no buffer gas, allowing the Rb to effuse from the pumping cell to maser cell as has been recently observed. The novel design of the DBRM will effectively eliminate the optical pumping light shift and buffer gas pressure shift that limit conventional rubidium frequency standards. The DBRM may provide short term frequency stability superior to that of a room temperature hydrogen maser, but in a smaller and more robust unit. Current investigations include the effects of wall-coatings on the decoherence rates for the Rb atoms. The CPT clock places the ^{87}Rb atoms into a superposition of hyperfine states using a laser modulated at a subharmonic of the hyperfine frequency and coherently resonant with either the D1 or D2 transition on both hyperfine levels. The atoms are pumped into a dark state in which there are no light shifts. Unlike in conventional optical pumping, the atoms directly radiate microwave power to a resonant cavity from the dark state. Present studies are of a CPT clock employing coated quartz cells.

David Phillips
Center for Astrophysics
60 Garden St. MS-59
Cambridge, MA 02138

email: dphil@cfa.harvard.edu
phone: (617) 496-7828
fax: (617) 496-7690

ON THE LINESHAPE OF THE CPT DARK LINE AND MICROWAVE EMISSION

*Filippo Levi, *Aldo Godone, ^S. Micalizio and +J. Vanier
*Istituto Elettrotecnico Nazionale, Torino, Italy
^Politecnico di Torino, Torino, Italy
+Université de Montréal, Montréal, Canada

1 ABSTRACT

The line-shape of an atomic transition is one of its most important attributes for high resolution spectroscopy and if a non Lorentzian shape is found a deep study of the physics of the system is required. In the experiments done on a resonant Λ scheme applied to alkali atoms in a cell with buffer gas, significant modifications and asymmetries have been found on the line-shapes of the Dark Line and of the CPT maser emission profile, as a function of the lasers power and detuning. In this paper we analyze different experimental conditions and a tight comparison between theory and experiments is reported for the cases of Cesium and Rubidium in a buffer gas.

2 INTRODUCTION

In recent years the possibility of using atomic three level systems to realize a new generation of atomic frequency standards [1-2-3] or other measurement instruments such as magnetometers [4] has been proposed by several groups. In the practical implementation of a high resolution atomic reference all those effects that may lead to shifts or asymmetries in the line shape are of crucial importance. In this paper we analyze both theoretically and experimentally several effects that may affect the atomic response to a Λ excitation scheme.

The observation of a Λ transition in an alkali vapor cell with buffer gas is a straightforward experiment, in fact only a single laser modulated at half the hyperfine transition frequency is required in order to excite the Λ transition and to produce a coherent superposition of the ground atomic states (Coherent Population Trapping or CPT).

When the coherence is established two main effects appear: the Dark Line in the fluorescence spectrum (an Electromagnetic Induced Transparency effect) [5] and the coherent microwave emission due to the hyperfine coherence created in the ground state (CPT-Maser) [6].

The theory of these systems has been widely examined [7-8-9]. Here we will recall first some theoretical results in order to go into a more detailed theoretical analysis of the line-shape.

In particular the effects of the laser beam profile and of the asymmetries of the excitation process on the line-shape will be considered and compared with the experimental results.

3 THEORY

The evolution of the three level system of figure 1 may be analyzed in the ensemble averaged density matrix formalism. With reference to the theory developed in [8], we have:

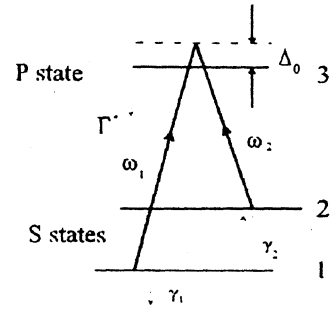


Figure 1 – Three level scheme representing the Λ transition in alkali atoms, typical values of the relaxation parameters are: $\Gamma^*/2\pi \approx$ 700 MHz and $\gamma_1 \approx \gamma_2 \approx 500\ s^{-1}$

$$\begin{cases} \rho_{11}+\rho_{22}\approx 1 \\ \dot{\rho}_{33}+\Gamma^*\rho_{33} = -\omega_{R1}\,\mathrm{Im}\,\delta_{13}-\omega_{R2}\,\mathrm{Im}\,\delta_{23} \\ \dot{\Delta}+\gamma_1\Delta = +\omega_{R2}\,\mathrm{Im}\,\delta_{23}-\omega_{R1}\,\mathrm{Im}\,\delta_{13} \\ \dot{\delta}_{12}+(\gamma_2+i\Omega_\mu)\delta_{12} = +i\tfrac{1}{2}\omega_{R1}\delta_{32}-i\tfrac{1}{2}\omega_{R2}\,\delta_{13} \\ \dot{\delta}_{13}+\left(\tfrac{1}{2}\Gamma^*+i\Delta_0\right)\delta_{13} = -i\tfrac{1}{4}\omega_{R1}(1-\Delta)-i\tfrac{1}{2}\omega_{R2}\,\delta_{12} \\ \dot{\delta}_{23}+\left(\tfrac{1}{2}\Gamma^*+i\Delta_0\right)\delta_{23} = -i\tfrac{1}{4}\omega_{R2}(1+\Delta)-i\tfrac{1}{2}\omega_{R1}\,\delta_{12}^* \end{cases} \quad (1)$$

where we have indicated with: Δ the population difference in the ground states ($\Delta=\rho_{22}-\rho_{11}$), ω_{R1} and ω_{R2} the Rabi frequencies of the optical transitions, Γ^* the lifetime of the excited state taking into account all the relaxation processes (spontaneous decays, collisions with the buffer gas, etc), γ_1 and γ_2 the population inversion and coherence relaxation rates of the ground state, Ω_μ the detuning between the laser radiation difference and the hyperfine transition frequency and with Δ_0 the detuning of the laser carrier from the optical line "center of gravity".

For the off diagonal terms of the density matrix (coherences) we have assumed solutions of the form:

$$\begin{cases} \rho_{12} = \delta_{12}\,e^{i(\omega_1-\omega_2)t} \\ \rho_{13} = \delta_{13}\,e^{i\omega_1 t} \\ \rho_{23} = \delta_{23}\,e^{i\omega_2 t} \end{cases} \quad (2)$$

To obtain the system (1) the following physical assumptions have been done:

1) the, saturation of optical transitions is small, $\omega_{R1}, \omega_{R2} \ll \Gamma^*$;

2) the Rabi frequency of the generated microwave is small compared to the atomic linewitdth and its effect on the system can be neglected [10];
3) the microwave detuning is small compared to the optical linewidth.
In [8] it was discussed in details the case where the optical coupling between the two ground state and the excited state are equal ($\omega_{R1}=\omega_{R2}=\omega_R$). Let's summarize the main results there reported with respect to the population in the excited state and the coherence in the ground state.
The optical fluorescence is directly proportional to the population in the excited state ρ_{33}; solving the system (1) in the stationary condition ($\partial/\partial t = 0$) we have:

$$\rho_{33} = \left(\frac{\omega_R}{\Gamma^*}\right)^2 \left(1 - \frac{\omega_R^2}{\Gamma^*}\frac{\Gamma'}{\Gamma'^2+\Omega_\mu^2}\right) \tag{3}$$

where $\Gamma' = \gamma_2 + \omega_R^2/\Gamma^*$.
It can be observed from (3) that a Dark Line appears in the fluorescence spectrum when the frequency difference of the two laser radiations equals the hyperfine splitting of the ground state. The line-shape of the Dark Line is given by:

$$\pi_{DL}(\Omega_\mu) = \omega_R^2/\Gamma^* \frac{\Gamma'}{\Gamma'^2+\Omega_\mu^2} \tag{4}$$

and the linewidth of the Dark Line is:

$$\Delta\omega_{1/2} = 2\Gamma' \tag{5}$$

The coherence term in the ground state is then given by:

$$\rho_{12} = -\frac{\omega_R^2}{2\Gamma^*}\frac{1}{\Gamma'+i\Omega_\mu} e^{i(\omega_1-\omega_2)t} \tag{6}$$

This coherence produces an oscillating magnetization as discussed in [8] at the angular frequency (ω_1-ω_2), that acts as source of a microwave field when the atomic ensemble is coupled to a microwave cavity; the power dissipated in the cavity P_{diss} and the emission profile of the CPT-maser turn out respectively:

$$P_{diss} = \frac{\frac{1}{2}\hbar\omega_{21} k N |2\delta_{12}|^2}{1+4Q_L^2(\Delta\omega_c/\omega_{21})^2} \tag{7}$$

$$\pi_M(\Omega_\mu) = \frac{(\omega_R^2/\Gamma^*)^2}{\Gamma'^2+\Omega_\mu^2} \tag{8}$$

In expression (7) N is the total number of atoms coupled by the Λ scheme, Q_L is the quality factor of the loaded microwave cavity, $\Delta\omega_c$ is the detuning between the cavity and the atomic line, k stands for the number of emitted microwave photons per atom per second; the expression of k is given by [8]:

$$k = \frac{\mu_0 \mu_{21}^2 \eta' Q_L N}{\hbar V_a} \tag{9}$$

where μ_0 is the vacuum permeability, μ_{21} is the magnetic dipole moment of the transition, η' is the filling factor of the cavity and V_a is the volume occupied by the atomic vapor interacting with the laser field. The linewidth of the maser emission profile is given by (5) also in this case.
According to theory both phenomena, Dark Line and maser profile, show Lorentzian shapes. We will discuss in the following paragraphs the experimental situation and the conditions which can lead to a different line-shape.
At low laser intensity the intensity of the transition instead is proportional to ω_R^2/Γ^* for the Dark Line, and to $(\omega_R^2/\Gamma^*)^2$ for the CPT maser; this difference is not surprising because of the different origin of the two signals: the Dark Line originates from an incoherent fluorescence, while the CPT Maser is generated by a coherent microwave emission.

4 EXPERIMENTAL SET UP

The experimental set-up is shown in figure 2, we have done experiments based on the D_1 and D_2 optical transition of ^{87}Rb and on the D_2 optical transition of ^{133}Cs.

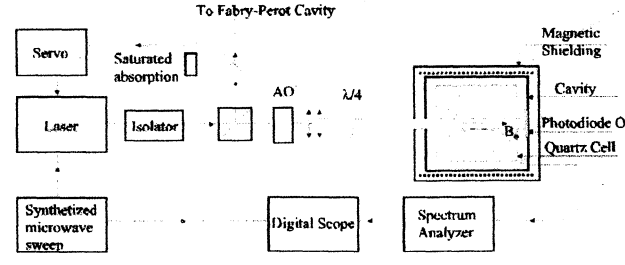


Figure 2 – Experimental set-up.

There are three main parts in the experimental set-up: the optical system, the physical package and the microwave generation and detection apparatus.
The optical part provides the radiations necessary to excite the Λ transition: it is realized starting from a single laser diode, frequency modulated at half the hyperfine frequency splitting of the ground state; in such a way, when the laser carrier is set to the line "center of gravity" of the transition, the first modulation sidebands correspond to the exciting radiations. The laser center frequency is locked to a saturated absorption signal obtained with an external absorption cell and its spectrum is monitored with a Fabry-Perot

cavity used as optical spectrum analyzer. The polarization of the beam is made circular with a λ/4 wave-plate before being sent to the microwave cavity, so the selection rules allow to excite only one of the possible Λ transitions; the laser intensity is approximately 200 μW/cm^2 for each sideband.
An acusto optic modulator is placed along the laser beam to act as an optical switch and to allow a precise frequency tuning with respect to the atomic optical reference.
The physical package consists of a sealed quartz cell inside a TE_{011} cylindrical microwave cavity that at its turn is contained inside a magnetic shield. The cell containing the alkali vapor and a buffer gas (N_2 at a pressure of 19 Torr with Cs and 32 Torr with Rb in our experiments) is thermally controlled; the operating temperature is around 45 °C when using Cs and around 60 °C when using rubidium. We operated with a cell volume of 3 cm^3 and a quality factor of the loaded cavity Q_L of 3 000 in the Cs case, and with a cell volume of 6 cm^3 and a Q_L of 10 000 in the Rb case.
A hole, covered with a layer of metal-mesh, is present on the cavity front end-cap allowing a laser beam diameter up to 2 cm inside the maser active region without destroying the cavity mode.
Finally a microwave synthesizer was used to modulate the laser carrier and a high sensitivity spectrum analyzer was used as heterodyne detector.
For the observation of the Dark Line, the experimental set-up was basically the same, but the detection was done with a photodiode by observing the laser transmission through the atomic vapor cell and the microwave cavity was absent.

5 LASER GAUSSIAN PROFILE EFFECT ON THE LINE-SHAPE

It is interesting to analyze how the shapes of the dark line and of the CPT-maser are modified by the laser profile used in the excitation process; these kind of effects were also studied in [10]. We assume in the following that the laser radiation has a Gaussian profile (TEM_{00} mode); then the pumping rate ω^2_R/Γ^* is a function of the radial position across the laser beam, we assume moreover to deal with an optically thin medium and that the light absorption along the z axis is negligible.
We can define the laser beam profile utilizing cylindrical coordinates as:

$$\omega_R^2(\rho,\varphi,z) = \omega_{R0}^2 e^{-\rho^2/\rho_0^2} \qquad \forall\varphi,\ \forall z \tag{10}$$

and, assuming no spatial variation along z the line-shape can be obtained by a double integration of the profiles expressed by (4) and (8) over the cylindrical coordinates ρ and φ.
The new profiles are then obtained in the two cases as:

$$\Pi_M(\Omega_\mu) = \frac{\int_0^{2\pi} d\varphi \int_0^{\infty} \rho d\rho \pi_M(\Omega_\mu)}{\int_0^{2\pi} d\varphi \int_0^{\infty} \rho d\rho e^{-(\rho/\rho_0)^2}} = \tfrac{1}{2}\ln\frac{(n+1)^2+\zeta^2}{1+\zeta^2} - \frac{1}{\zeta}\operatorname{arctg}\frac{n\zeta}{1+n+\zeta^2} \tag{11}$$

$$\Pi_{DL}(\Omega_\mu) = \frac{\int_0^{2\pi} d\varphi \int_0^{\infty} \rho d\rho \pi_{DL}(\Omega_\mu)}{\int_0^{2\pi} d\varphi \int_0^{\infty} \rho d\rho e^{-(\rho/\rho_0)^2}} = \tfrac{1}{2}\ln\frac{(n+1)^2+\zeta^2}{1+\zeta^2} \tag{12}$$

where we have defined:

$$n = \frac{\omega_{R0}^2}{\gamma_2\Gamma^*} \quad \text{and} \quad \zeta = \Omega_\mu/\gamma_2 \tag{13}$$

We are interested to the limits of the expressions (11) and (12) in the case of low and high pumping rate ($n \to 0$, or $n \to \infty$), they can be obtained as:

$$\lim_{n\to 0}\Pi_{DL}(\Omega_\mu) = \frac{n}{1+\zeta^2} = \frac{\gamma_2\omega_{R0}^2}{\Gamma^*}\frac{1}{\gamma_2^2+\Omega_\mu^2} \tag{14a}$$

$$\lim_{n\to 0}\Pi_M(\Omega_\mu) = \tfrac{1}{2}\frac{n^2}{1+\zeta^2} = \left(\omega_{R0}^2/\Gamma^*\right)^2\frac{1}{\gamma_2^2+\Omega_\mu^2} \tag{14b}$$

$$\lim_{n\to\infty}\Pi_{DL/M}(\Omega_\mu) = \tfrac{1}{2}\ln\frac{n^2}{1+\zeta^2} = \tfrac{1}{2}\ln\frac{\frac{\omega_{R0}^2}{\gamma_2\Gamma^*}}{1+\left(\Omega_\mu/\gamma_2\right)^2} \tag{14c}$$

The low power approximation gives again a Lorentzian shape in both cases, and the linewidth can be always expressed by (5).
The line shape described by (14 c) on the other hand is quite different from a Lorentzian shape and the linewidth turns out smaller than that described by equation (5) for an equivalent value of pumping rate:

$$\Delta\omega_{1/2} = 2\gamma_2\sqrt{\frac{\omega_R^2}{\gamma_2\Gamma^*}} \tag{15}$$

The description of the experimental data with these profiles is quite satisfactory as can be seen by the fitting curves of figure 3. In particular, in figure 3a we report the experimental emission profile of the maser at low saturating power fitted with a Lorentzian curve, in 3b the profile of the dark line observed at low saturation power fitted with a Lorentzian curve, and in figure 3c the emission profile of the maser excited with a high intensity laser beam, fitted with equation (14 c).

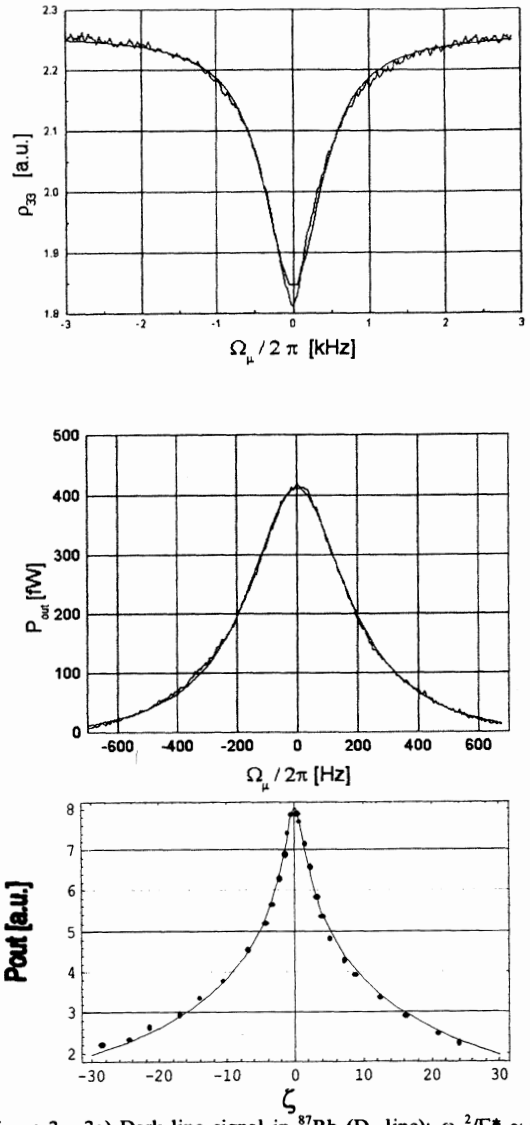


Figure 3 - 3a) Dark line signal in ^{87}Rb (D_2 line): $\omega_R^2/\Gamma^* \approx \gamma_2$ (low pumping rate). 3b) Maser emission profile in ^{87}Rb (D_1 line): $\omega_R^2/\Gamma^* \approx \gamma_2$ (low pumping rate). 3c) Maser emission profile in ^{133}Cs (D_2 line): $\omega_R^2/\Gamma^* \approx 10\ \gamma_2$ (high pumping rate).

6 LASER ASYMMETRIES IN THE EXCITATION PROCESS

Up to now we have assumed to excite the Λ transition with a perfectly symmetric spectrum utilizing atoms with symmetric dipole coefficients and with a laser well tuned with respect to the atomic levels.

In the present paragraph we will investigate what happens to the line-shape when these conditions are not respected anymore.

It is worth to notice that since in the equations appear only the optical Rabi frequency from the mathematical point of view it is the same to have an asymmetry in the spectrum or in the Clebsch-Gordan coefficients. In fact the Rabi frequency is defined as the product of the applied field time the transition dipole moment ($\omega_{R1} = E_{01}\, d_{13}/\hbar$, $\omega_{R2} = E_{02}\, d_{23}/\hbar$); the higher symmetry degree is one of the reasons why it is important to use the D_1 line; another reason that leads to the choice of the D_1 line is the absence of cycling transitions that could lead to strong differential absorption of the two radiations.

When $\omega_{R1} \neq \omega_{R2}$ and $\Delta_0 \neq 0$ simultaneously it is still possible to derive an analytical solution of the system (1) [9],even if quite complex mathematical expressions are obtained not transparent to a physical interpretation; it is possible in any case to analyze them numerically. Such analysis gives us two different behaviors for the Dark Line and the CPT maser.

When $\omega_{R1} \neq \omega_{R2}$ and $\Delta_0 \neq 0$, in the fluorescence spectrum appears a Raman type transition that can distort the line-shape, without moving the minimum of the profile, while the maser emission is shifted in frequency but its line-shape is not affected by the Raman transition.

This difference is due to the fact that while the CPT maser emission is a phenomenon related only to the coherence in the ground state, the Dark Line is observed trough the Electromagnetic Induced Transparency and then is sensitive to all those effects that affect the excited state population.

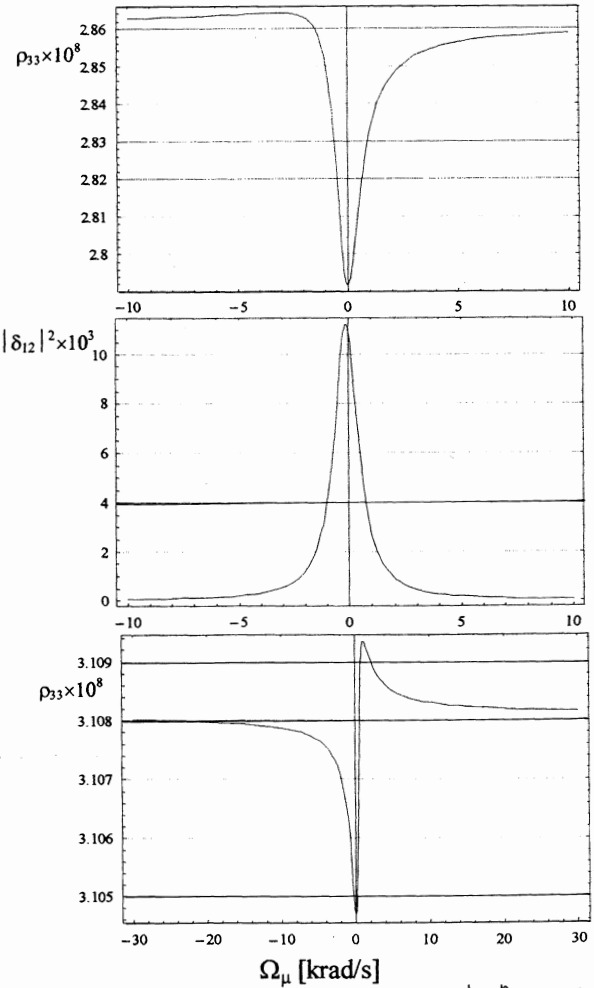


Figure 4 – 4a) and 4b) Calculated value of ρ_{33} and $|\delta_{12}|^2$ respectively: laser asymmetry 85%, carrier detuning 320MHz, γ_2 500 s^{-1}, $\omega_R^2/\Gamma^* =$ 560 s^{-1}. 4c) Calculated value of ρ_{33}: $\omega_R^2/\Gamma^* =$ 16000 s^{-1}, other parameters as before specified.

In figure 4 are reported the calculated profiles when a carrier detuning of 300 MHz is present ($\Delta_0 \approx \Gamma^*/2$) together with a laser asymmetry of 85%.
As it can be seen the maser emission profile is a Lorentzian shifted by 200 rad/s while the Dark Line is distorted, but the minimum of ρ_{33} remains always at $\Omega_\mu=0$.
With a very strong asymmetry it is possible to enhance the Raman transition, but the minimum always remains at zero detuning (figure 4 c).
Experimental observations confirm these results with respect to the line-shapes: in order to excite the Λ transition with a strongly asymmetric radiation it is possible to use the carrier and the second harmonic of the modulation frequency instead of the two first sidebands. In figure 5 some experimental profiles are reported for the Dark Line and for the CPT maser.

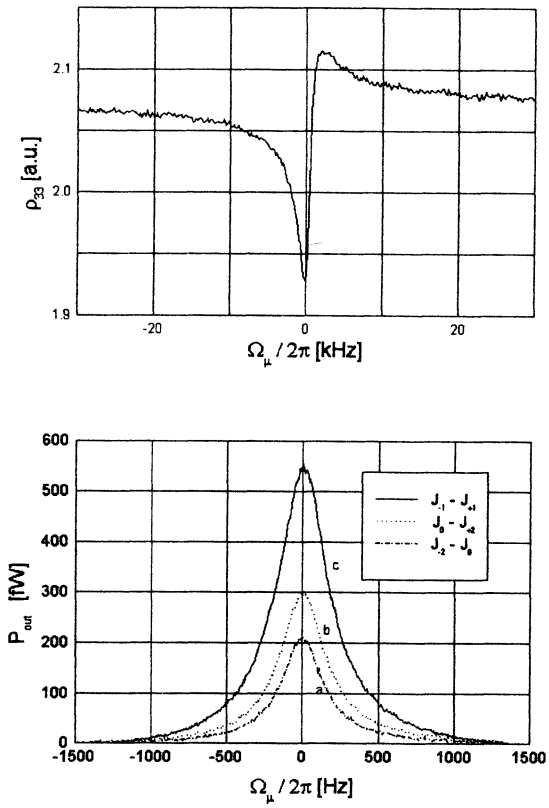


Figure 5 - 5a) Dark line signal in ^{87}Rb (D_2 line): laser power $(\omega_R^2/\gamma_2\Gamma^*) \approx 5$, carrier detuning 350 MHz. 5b) Maser emission profile in ^{87}Rb (D_1 line): $(\omega_R^2/\gamma_2\Gamma^*) \approx .5$ in the excitation process the two first sidebands are used (curve c) or the carrier and the second harmonic (curve a and b) with an asymmetry above 50%, laser detuning 100 MHz.

It is confirmed that the maser emission profile is always Lorentzian no matter which is the laser detuning and which is the laser asymmetry, while the shape of the Dark Line is sensitive to the above mentioned parameters.
The measure of the shift of the CPT-Maser profile is a more complex problem, in fact as it was discussed in [12] the total light shift (accounted for in the static Hamiltonian operator, such as other shifting effect like for example buffer gas and static Zeeman shift) is a complex function of the modulation index and of the carrier detuning. It is then extremely difficult to distinguish between the static light shift and the dynamic one described above with the set-up used for the experiments described in this paper.

7 CONCLUSIONS

We have investigated the line-shape of the Dark Line and of the CPT maser emission profile observed in rubidium and in cesium vapors with buffer gas. In particular a non-Lorentzian line shape is determined by the joint effect of a Gaussian laser field profile and a high pumping rate. Moreover the unbalance of the optical Rabi frequencies in the Λ scheme together with a detuning of the laser carrier frequency with respect to the "center of gravity" of the optical transitions produces a significant asymmetry in the line-shape of the dark line and a shift of the maser emission profile. All the above mentioned effects have been theoretically explained in the framework of a three level system and found in good agreement with the experimental results taken on the optical transitions of ^{133}Cs (D_2 line) and ^{87}Rb (D_1 and D_2 lines).

8 REFERENCES

[1] A. Godone, F. Levi and J Vanier, IEEE Trans. on Instr. & Meas.**48** p 504 (1999).
[2] N. Cyr, M. Têtu and M. Breton, IEEE Trans. Instr.& Meas. **42**, p. 640 (1993).
[3] N. Vukicevic *et. al.*, in Proceedings of 1999 Joint EFTF and IEEE- IFCS meeting, Besancon (1999), p 133.
[4] M. O. Scully and M. Fleschhauer, Phys. Rev. Lett., **69**, p 1360, (1992).
[5] G. Alzetta *et. al.*, Il Nuovo Cimento B **36**, 5, (1976).
[6] A. Godone, F. Levi and J. Vanier Phys. Rev. A 59 1 p R 12 (1999).
[7] E. Arimondo Prog. in Optics XXXV, p 257, (1996).
[8] J. Vanier, A. Godone and F. Levi, Phys. Rev. A **58**, p 2345, (1998).
[9] G. Orriols, Il Nuovo Cimento B **53**, 1, p 1, (1979).
[10] In this paper we assume that there is no RF coupling in the ground state. For the high feedback theory see the paper "Self consistent theory of the CPT maser" in these proceedings".
[11] P. Minguzzi, *et. al.* J. Opt. Soc. Am. B **3**, p. 1075 (1986).
[12] F. Levi, A. Godone, J. Vanier, IEEE on UFFC (to be published).

PROGRESS TOWARD THE REALIZATION OF A CESIUM FOUNTAIN FREQUENCY STANDARD AT IEN

Filippo Levi[+], Aldo Godone[+], Luca Lorini[++] and Davide Calonico[++]
[+]*Istituto Elettrotecnico Nazionale "G. Ferraris", Torino, Italy*
[++]*Politecnico di Torino, Torino, Italy*

1 ABSTRACT

At Istituto Elettrotecnico Nazionale "Galileo Ferraris" works are in progress in order to realize a cesium fountain to be eventually used as the Italian Primary Frequency Standard. A characterization of the atomic source is given.

2 PHYSICAL PACKAGE DESCRIPTION

2.1 Laser system

Four laser diodes are used in the optical system as shown in figure 1. A DBR laser diode (master laser) mounted in a low feedback configuration in order to further reduce its line-width, is frequency locked with a Pound-Drever-Hall technique, upon need between 150 and 170 MHz red of the F=4→F'=5 cesium D_2 transition.

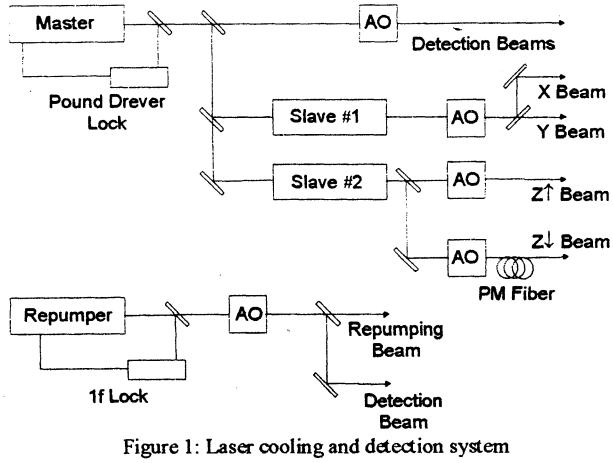


Figure 1: Laser cooling and detection system

The light of the master laser (approximate output 50 mW) is then used to injection lock two other laser diodes (slave lasers) and to realize the detection beams; this last beam is shifted back with its own acousto-optic modulator (AO) running at the same frequency of the master one.
The two slave lasers (150 mW each) are used to amplify the master laser light and generate the X and Y cooling beams and the Z↑ and the Z↓ beams respectively. Three AO (one for the X and Y beams and the other two for the Z beams) driven by independent synthesizers, locked to a common reference, are used to blue shift the laser light near to resonance.
The Z↓ beam is delivered to the vacuum system through a polarization maintaining fiber, while the other trapping beams are spatially filtered by means of pinholes. The power density of the trapping beams is around 3 mW/cm^2.
The repumping light is obtained from a second DBR laser diode frequency locked to the F=3→F'=3 transition with a first order saturated absorption technique; this laser is then blue shifted to F=3→F'=4 transition by mean of an AO; this procedure is necessary because we use the repumper light also in the detection region and we need a fast optical shutter.

2.2 Physical package

The physical package (designed in collaboration with NIST) is composed by four main regions: trapping region, lower and upper detection region, microwave cavity and drift tube figure 2.

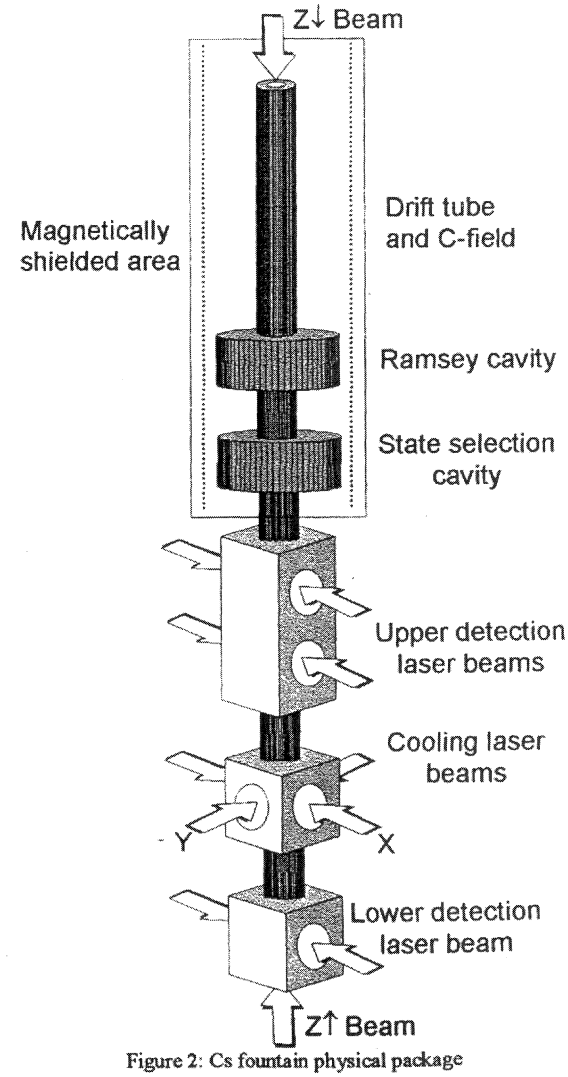


Figure 2: Cs fountain physical package

Only this latter region is magnetically shielded: four layers of molybdenum-Permalloy cylindrical shields surround the cavity, the drift tube and the C-field bobbin. Two microwave cavities separated by 10 cm of below cutoff cylindrical waveguide are used, one for doing state selection and the second for the clock Ramsey interaction. The microwave cavities are part of the vacuum structure and are described in [1].
No magnetic shields are used in the trapping and detection regions; instead compensation coils are used in the trapping region in order to improve cooling and tossing efficiency. A static quadrupole magnetic field is also present in the trapping region when the system is operated as magneto-optical trap $(\partial B / \partial z \approx 10$ Gauss/cm$)$.
In the upper detection region two distinct detectors are used to make a differential measure of the population in the state F=3 and F=4 after the Ramsey interaction [2]; a non-imaging optical system is used and a fluorescence detection efficiency of 17% is achieved.
At the time of the conference the cavities are being installed and we will report later on the first Ramsey signals.

2.3 Control system

A PC operating in Windows system superintends to all the control and acquisition functions. Three National Instruments boards are used: pattern generation, A/D converter, and GPIB control.
The pattern generation is done with a board providing 32 I/O TTL channels; the accurate timing of the sequences is guarantee by an autonomous clock on the board, the minimum time interval that can be programmed is of 1 μs and continuous operation is guaranteed by double buffering.
The A/D converter has a resolution of 16 bit and a sampling speed of 100 kS/s; it can be programmed to digitize up to 8 distinct channels.
The PC will measure the area of the time of flight signals, through the GPIB board will modulate the interrogating microwave signal and eventually lock the synthesizer on the central Ramsey fringe.

3 TRAP CHARACTERIZATION

3.1 Loading of the trap

We have done a series of experiments to characterize the behavior of the cold atom source.
The loading of the MOT/molasses is done out of a low collimated beam, which however create also a background gas in dynamical equilibrium with the graphite getters present in the vacuum chamber.
If we analyze the number of atoms loaded as a function of the Cs source temperature we can see a smooth transition from the beam loading to the gas loading in the temperature range between 10 and 30 °C (figure 3).

The present beams configuration (XYZ with light reflection on the X and Y) does not allow us to further increase the oven temperature and to complete the transition to a complete gas loading of the trap, in fact light absorption on the X and Y axis produces a strong imbalance in the standing-waves, creating a competing heating mechanism. At 40°C absorption of 15% of the X and Y beams is measured.

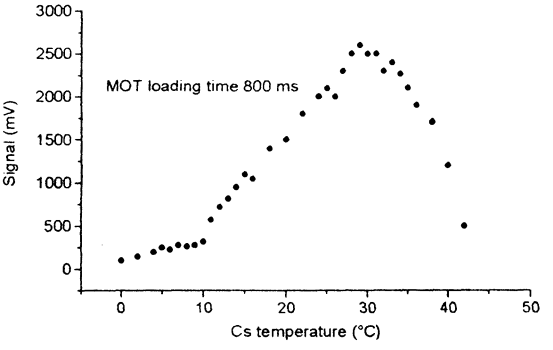


Figure 3: Signal intensity as a function of the source temperature

Not only the number of atoms but also the loading time is a function of the oven temperature as can be seen in figure 5.

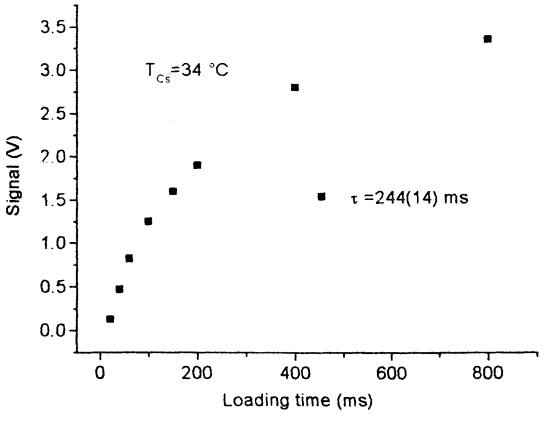


Figure 4: Signal intensity versus MOT loading time

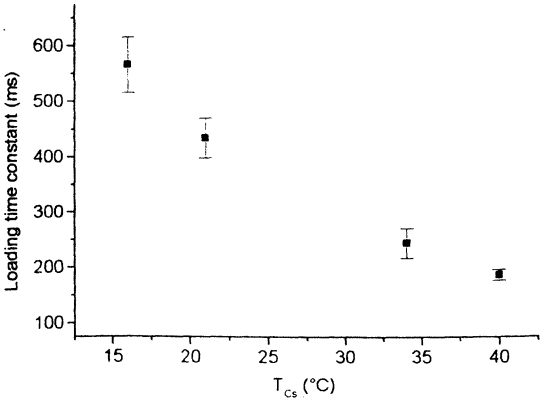


Figure 5: MOT loading time constant as a function of the Cs source temperature

The estimated number of atoms gathered at a temperature of 33°C is around 10^7.

3.2 Post cooling sequences

As it is well known sub-Doppler molasses are required to push Cs temperature down to the μK region. Sub-Doppler cooling mechanism is achieved by optical pumping between Zeeman sub-levels of hyperfine structure; polarization gradients and magnetic field control are required in the molasses area. Two different polarization configuration are effective in sub-Doppler cooling: counter-rotating polarization ($\sigma^+-\sigma^-$) and linear perpendicular polarization (lin⊥lin) [3].
Relations between molasses temperature and laser parameters are similar in both configurations: if laser intensity is well below saturation ($I << I_S$) and the detuning is larger than the line width ($|\delta| >> \Gamma$) we have a simple relation between temperature and laser parameters.

$$kT \propto \hbar\omega_R^2/|\delta|.$$

Low temperature in sub-Doppler molasses requires low intensities and large detunings: these conditions are not suitable to capture large numbers of atoms so we first load a Doppler molasses or a MOT, we then adiabatically change laser parameters, optimizing sub-Doppler conditions (Post-Cooling technique).
During the post-cooling sequence the laser detuning is changed from Γ to 10Γ and simultaneously the laser power is weakened by 40 dB, well below saturation. Parameters change is done, by the pattern generator through programmed sequences.
The efficiency of the post-cooling sequence was tested measuring molasses temperature with a standard Time Of Flight (TOF) analysis. Temperature measurements are then related to laser power and detuning; measurements confirm theory predictions, showing linear relations between power, detuning and temperature (figure 6).

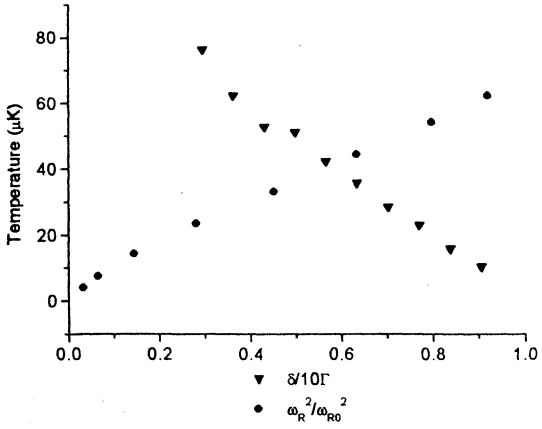


Figure 6: Relations between molasses temperature, laser attenuation (ω_R^2/ω_{R0}^2) and detuning (δ/10Γ)

Obviously the better performances are obtained when intensity and power ramps are active at the same time: In figure 7 is plotted the atomic temperature versus the light shift parameter $\left(\frac{\omega_R^2}{|\delta|\Gamma}\right)$.

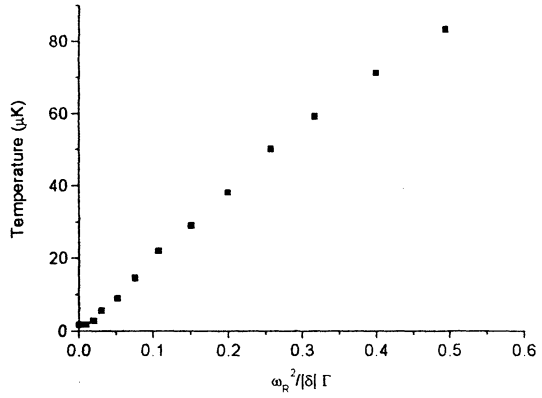


Figure 7: Molasses temperature versus light shift parameter

Figure 8 shows a TOF signal obtained with an optimized $\sigma^+-\sigma^-$ molasses the measured temperature is T=1.5(.5) μK.

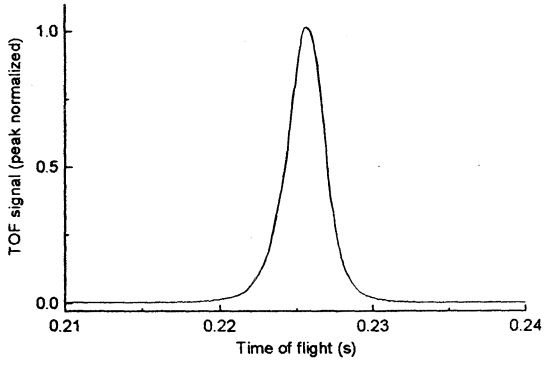


Figure 8: TOF signal corresponding to a 1.5 μK molasses temperature

3.3 Friction coefficient measurement

Another characterizing parameter of a molasses is the sub-Doppler friction coefficient α
The cooling and post-cooling processes can be represented through a friction force of the type:

$F = -\alpha v$

where α is related to damping time constant τ of the thermal velocity by $\alpha = \frac{m_{Cs}}{\tau}$.

Our measurements of α were carried out by changing the laser power from a high value, consistent with Doppler cooling, to a low value, consistent with sub-Doppler cooling. Post-cooling duration was varied to study the transient behavior of the atomic temperature

upon a variation of the laser power. The friction coefficient can then be obtained (figure 9) by fitting the measured velocities achieved at different time of transient response with the theoretical function

$$v(t) = v_R + v_0 e^{-t/\tau}$$

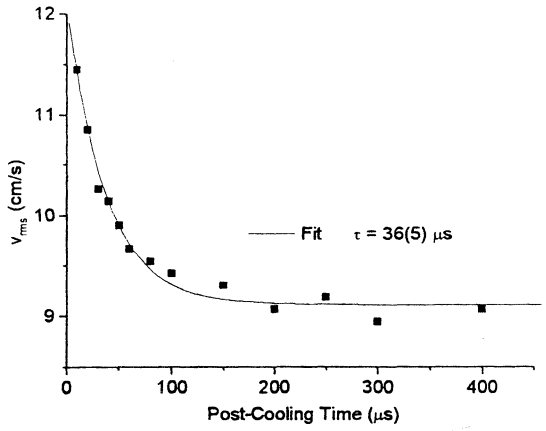


Figure 9: Post-cooling transient

The result obtained (α=6.0(8)·10^{-21} kg/s) is in agreement with the values reported in [4].

3.4 Atom launch and detection

The launch of the atoms is done with the usual moving molasses technique. The trapped atoms are temporary released switching off all the laser beams, the two synthesizer driving the Z AO are frequency shifted (in opposite) directions by $\pm\Delta\nu$, and then the atom are cooled again in the moving frame, the speed of the new frame is

$$v_0 = 2\lambda\Delta\nu$$

After the acceleration atoms are cooled with the post cooling sequence below 2 μK, and then all the beams are switched off leaving the atoms in a free ballistic flight inside the gravitational field.
Because of structure limitations up to now we were limited to a maximum launch speed of 2.4 m/s corresponding approximately to a toss height of 30 cm, with a maximum TOF of 600 ms.
As can be seen in figure 10 the atom loss is higher of what it can be estimated by the thermal cloud expansion. This is mainly due to two reasons:

1) inadequate vacuum level
2) vertical alignment of the vacuum structure respect to gravity force.

Both problems are of easy solution and will be hopefully solved in the near future.

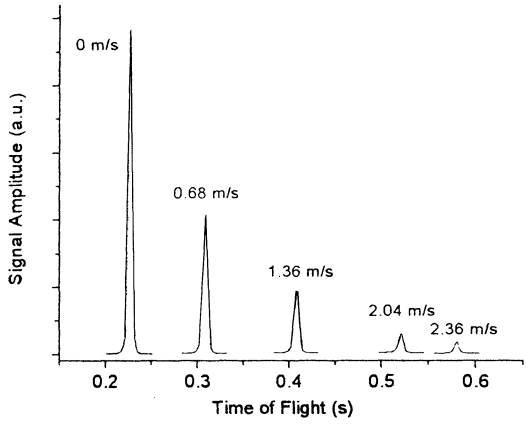


Figure 10: TOF signals corresponding to increasing toss speeds

4 CONCLUSIONS

In this work we described the new developments of IEN Cs fountain.
Our optical system was improved and fully prepared and is now ready for routine operation. A PC based control system was installed, giving a complete automation of the fountain sequences (capture, post-cooling, launch and detection).
We characterized sub-Doppler cooling by measuring molasses temperature as a function of laser power and detuning. Post-cooling transient was also investigated and a measurement of sub-Doppler friction coefficient was obtained. Atoms were cooled to 1.5 μK and launched to obtain a 600 ms TOF.
This work is performed in the frame of a scientific agreement between IEN, NIST and the Politecnico di Torino.

ACKNOWLEDGEMENTS

We would like to thanks S. R. Jefferts for his invaluable help in the realization of our fountain.
Filippo Levi would like to thanks the time and frequency division of NIST for the opportunities given during his permanence at NIST

REFERENCES

[1] S. Jefferts, M. Drewsen, A. De Marchi, Proc. of Freq. Cont. Symp. (1998)
[2] A. Clairon *et al.* IEEE Trans. On Instr. and Meas. 44, 2 (1995)
[3] *Laser cooling and trapping of atoms*, Special issue of J.O.S.A. B, 6, 11 (1989)
[4] M. Drewsen *et al.* Appl. Phys B 59, 283-298 (1994).

MULTI LAMBDA CYLINDRICAL RABI CAVITIES FOR THERMAL CESIUM BEAM FREQUENCY STANDARDS ; OPTIMIZATION OF THE PERFORMANCES

Virgile HERMANN[1], Frédéric HAMOUDA, Geneviève THEOBALD, Luc CHASSAGNE, Claude AUDOIN and Pierre CEREZ*

Franck MERSCHBERGER, Fabrice BOUST and François DEYZAC**

*Laboratoire de l'Horloge Atomique
Bâtiment 221 – Université Paris-Sud, 91405 Orsay Cedex, France

**ONERA BP 72 – F92322 Châtillon Cedex, France

[1]PhD student at the Laboratoire de l'Horloge Atomique with a CIFRE convention supported by TEMEX TELECOM

ABSTRACT

The Rabi transition probability of cesium atoms interacting with microwave fields produced in $TE_{01N(N=1,2,3)}$ cylindrical cavities is first calculated. Then, we consider a small optically pumped frequency standard operating with these cavities. We determine the operational parameters which lead to both the better short term frequency stability and the cancellation of the cavity pulling frequency shift.

Finally, we estimate the first order Doppler frequency shift which, in the Rabi cylindrical cavities, is the main shift.

INTRODUCTION

It is well known that multi wavelength cylindrical cavities present attractive features such as the ease of construction for small and inexpensive Cs beam frequency standards. In this paper we consider the possibility of using TE_{011}, TE_{012}, TE_{013} cylindrical cavities for the excitation of microwave transitions in a Cs beam. The three considered cavities have the same interaction length L=90 mm. The Rabi transition probability is first calculated for a monokinetic atom through Bloch equations. Then this probability is weighted on the atomic beam velocity distribution. The result depends on both the frequency detuning and the Rabi angular frequency b. In operating conditions, the microwave field frequency is square wave modulated with a given modulation depth ω_m. The operational parameters b and ω_m are set as to provide the maximum slope of the frequency discriminator as well as to cancel the cavity pulling frequency shift. Such operational parameters are then used to evaluate the frequency performances of small optically pumped frequency standards operating with these cylindrical Rabi cavities. The first evaluation concerns the short term frequency stability when the predominant source of noise is the noise of the optical detection system.

The second evaluation is devoted to the calculation of the shift due to the first order Doppler effect which in the Rabi cylindrical cavities is the main shift. This calculation needs the knowledge of the microwave magnetic field variation along the cavity axis.

1. CAVITY FIELD DESCRIPTION

The 3 considered cavities resonate on the TE_{011}, TE_{012}, TE_{013} modes respectively, at the cesium frequency. When these cavities are weakly coupled on one hand and have negligible losses on the other hand, the microwave magnetic field of pulsation ω is longitudinal close to the cavity axis z. Its profile along the z axis is well represented by the sine function $\sin\left(\frac{N\pi z}{L}\right)$ where N = 1, 2, 3 is the number of modes in the cavity (fig. 1).

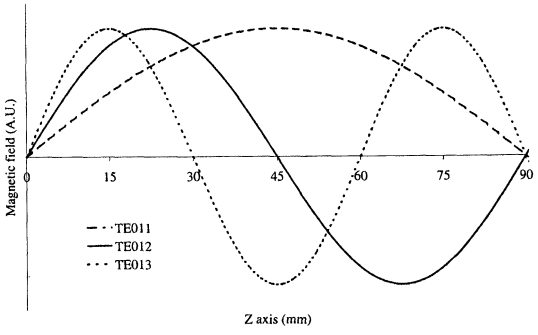


Fig. 1 Magnetic field amplitude versus longitudinal cavity z axis.

It is judicious to introduce the variable $\theta = \frac{z - z_0}{v}$ in which z_0 is the abscissa at the entrance of the cavity, v is the atomic velocity. θ represents the interaction time with the microwave field. Let us define $T = \frac{L}{v}$ and $\tau = \frac{T}{N}$. T appears to be the whole interaction time in the cavity while τ is the interaction time in one mode. Therefore, we may introduce a dimensionless interaction time variable $\xi = \frac{\theta}{\tau}$. ξ varies from 0 to N in the cavity TE_{01N}, with $N = 1, 2$ or 3.

2. RABI TRANSITION PROBABILITY IN A TWO LEVEL SYSTEM

The time evolution of the atomic quantum state is described by a set of three differential equations [1].

$$\begin{cases} \frac{\partial a_1(\xi)}{\partial \xi} + \frac{\Omega_0 T}{N} a_2(\xi) = 0 & (1.a) \\ -\frac{\Omega_0 T}{N} a_1(\xi) + \frac{\partial a_2(\xi)}{\partial \xi} - b\tau . sin(\pi\xi) . a_3(\xi) = 0 & (1.b) \\ b\tau . sin(\pi\xi) . a_2(\xi) + \frac{\partial a_3(\xi)}{\partial \xi} = 0 & (1.c) \end{cases}$$

a_1 and a_2 are the hyperfine coherence terms while a_3 is the population difference between the 2 states with an energy separation $\hbar\omega_0$. b is the maximum value of $b(\xi)$.

The set of equations (1) is numerically solved using a Runge Kutta method. We are interested in $a_3(\xi)$ which is related to the transition probability by:

$$p(\xi) = \frac{1}{2} . \left[1 - \frac{a_3(\xi)}{a_3(\xi = 0)} \right] \quad (2)$$

At the output of the cavity the transition probability $p(\xi=N)$ depends on the value of 3 parameters which are τ, b and the frequency detuning $\Omega_0=\omega-\omega_0$.

We have to take into account the distribution $f(\tau)$ of the interaction time. When the resonator is optically pumped with a single laser, $f(\tau)$ has the following expression [2] :

$$f(\tau) = \frac{2}{\tau_1} . \left(\frac{\tau 1}{\tau} \right)^5 . e^{-\left(\frac{\tau_1}{\tau}\right)^2} \quad (3)$$

with $\tau_1 = \frac{L}{\alpha N}$. α is the most probable velocity in the cesium beam.

Finally we get :

$$P(b, \Omega_0) = \int_0^\infty f(\tau) . p(\tau, b, \Omega_0) . d\tau \quad (4)$$

3. OPERATING CONDITIONS

The microwave field frequency is square wave modulated with a modulation depth ω_m. The operational parameters $\omega_m T_1$ and $b\tau_1$ are set in order to optimize both the short and long term frequency stability. T_1 is a peculiar value of the transit time T in the cavity for atoms with velocity α.
For purpose of optimizing the short term frequency stability, it needs to obtain the largest slope for the error signal, that is :

$$\left. \frac{\partial P}{\partial \Omega_0 T_1} \right|_{b\tau_1, \Omega_0 = \omega_m} \text{max} \quad (5)$$

Moreover, the long term frequency stability can be improved if the frequency shift due to cavity pulling is cancelled. This is achieved when :

$$\left. \frac{\partial P}{\partial b\tau_1} \right|_{b\tau_1, \omega_m T_1} = 0 \quad (6)$$

The values of the operating parameters $b\tau_1$ and $\omega_m T_1$ which give the largest value of the slope with the condition (6) satisfied are gathered in table 1.

Cavity	$b\tau_1$	$\omega_m T_1$	$\left. \frac{\partial P}{\partial \Omega_0 T_1} \right\|_{b\tau_1, \omega_m T_1}$	$P(b\tau_1, \omega_m T_1)$
TE_{011}	6.55	3.26	0.115	0.523
TE_{012}	2.88	2.29	0.172	0.260
TE_{013}	6.99	2.77	0.087	0.603

Table 1 Values of $b\tau_1$ and $\omega_m T_1$ which give the maximum slope of the error signal when the condition (6) is satisfied for the 3 considered cylindrical cavities.

4. SHORT TERM FREQUENCY STABILITY

In most cases, the predominant source of noise in an optically pumped cesium beam resonator is the noise of the optical detection system. It then results [3] that the Allan variance of the relative frequency

fluctuation is given by :

$$\sigma_y(\tau)=\frac{1}{\omega_0 T_1}\left(\frac{F}{2I_0}\right)^{1/2}.H^{1/2}.\tau^{-1/2} \qquad (7)$$

for an averaging time $\tau \geq 1$s. F is the noise factor of the atomic detection, I_0 is the atomic flux. $H^{1/2}$ caracterizes the short term frequency stability; it is inversely proportional to the slope of the error signal.

It appears from table 1 that the TE_{012} cavity gives the best short term frequency stability.

This result agrees with the conclusions obtained by comparing resonators designed with Ramsey cavities which deliver a signal with either a maximum or a minimum at resonance [2]. Finally the TE_{013} cavity gives the worst result.

5. FREQUENCY SHIFT DUE TO THE FIRST ORDER DOPPLER EFFECT

Up to now, we deal with ideal cavities in which the Rabi frequency $b(\xi)$ is a sinusoidal function of ξ.

$$b(\xi)=b.sin(\pi\xi) \qquad (8)$$

Actual cavities exhibit wall losses on one hand and imperfections due to coupling to the main line on the other hand. In this case, the Rabi angular frequency can be approximated by the following expression :

$$b(\xi)=b.sin(\pi\xi)e^{-i\phi(\xi)} \qquad (9)$$

$\phi(\xi)$ is the residual phase of the microwave field.

It follows that :

$$b(\xi)=b_1(\xi)+ib_2(\xi) \qquad (10)$$

with :

$$b_1=b.sin(\pi\xi).cos(\phi(\xi)) \qquad (10.a)$$

$$b_2=-b.sin(\pi\xi).sin(\phi(\xi)) \qquad (10.b)$$

The set of equations (1) becomes the set of equations (11).

The numerical treatment of equations (11) requires the knowledge of the residual phase ϕ as a function of the position on the beam axis in the cavity. The magnetic field of a real cavity coupled by holes has been calculated by our colleagues of IRCOM using a finite element method [4].

A first result is given in figure 2 for two modes cavity TE_{012} coupled by means of 2 slots. Large phase variations are observed. They are probably due to an inadequate coupling.

The frequency offset is extracted from the transition probability curve by simulating the operation of the servo system in modulation regime.

The program searches for the value of Ω_0 which satisfies $P(\Omega_0+\omega_m)=P(\Omega_0-\omega_m)$.

Figure 3 shows the result of this calculation as a function of the modulation depth. In standard conditions of operation defined by $\omega_m T_1 = 2.29$ and $b\tau_1 = 2.88$, the frequency shift is $\Delta F = -1.397$ Hz.

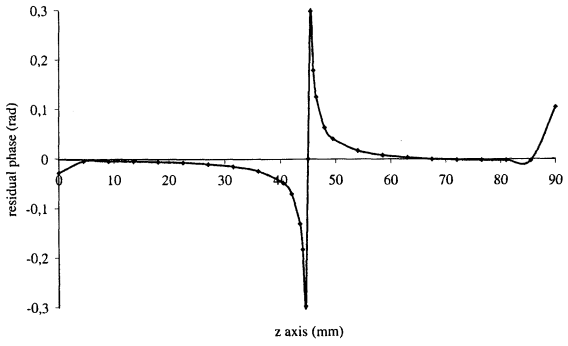


fig. 2 Residual phase versus cavity z axis.

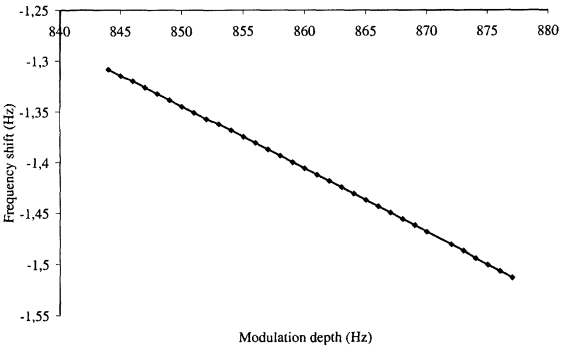


fig. 3 First order Doppler frequency shift versus modulation depth.

$$\begin{cases} \dfrac{\partial a_1(\xi)}{\partial\xi}+\dfrac{\Omega_0 T}{N}a_2(\xi)-b\tau.sin(\pi\xi).sin(\phi(\xi))a_3(\xi)=0 & (11.a)\\ -\dfrac{\Omega_0 T}{N}a_1(\xi)+\dfrac{\partial a_2(\xi)}{\partial\xi}-b\tau.sin(\pi\xi).cos(\phi(\xi))a_3(\xi)=0 & (11.b)\\ b\tau.sin(\pi\xi).sin(\phi(\xi))a_1(\xi)+b\tau.sin(\pi\xi).cos(\phi(\xi))a_2(\xi)+\dfrac{\partial a_3(\xi)}{\partial\xi}=0 & (11.c) \end{cases}$$

As a matter of fact, this large offset is not surprising if we observe that a large phase difference $\Delta\phi \approx 10$mrad occurs between the magnetic fields at $z=L/4$ and $z=3L/4$ on z axis.

A second result concerns the same two modes TE_{012} cavity coupled by a sophisticated and well balanced system feeding the two above mentioned slots. The calculated frequency shift amounts now to -1.41 mHz in standard operating conditions. It is very small in this case and this fact may be related to the very small value of the phase difference previously defined $\Delta\phi \approx 5$ µrad .

In a small cesium beam resonator, the relative frequency offset due to Doppler effect could be reasonably of the order of a few 10^{-12}. The previous calculations allow us to set an order of magnitude of $\Delta\phi \approx 100$µrad.

6. CONCLUSION

In this paper, we have studied the performances of an optically pumped Cs beam clock in which a multi lambda cylindrical Rabi cavity has been implemented. We focus our attention on the short and medium terms frequency stability as well as first order Doppler shift.

These two features were estimated through numerical evaluation of the Bloch equations governing the atom field interaction. For this purpose, finite elements methods applied to actual cylindrical cavities provide us with the necessary amplitude and phase of the microwave field. Such estimates are quite encouraging regarding the possible use of the TE_{012} cylindrical cavity in a small Cs bean resonator. However, the practical realization of the cavity coupling must ensure a value of $\Delta\phi \leq 100$ µrad, which is quite feasible.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support of the DGA (Délégation Générale de l'Armement – DSP/STTC) and the CNES (Centre National d'Études Spatiales – Département Temps-Fréquence).

REFERENCES

[1] Claude Audoin, "Comparaison de plusieurs types de cavités résonantes", Private communication.

[2] C. Audoin, V. Giordano, N. Dimarcq, P. Cérez, P. Petit and G. Théobald, "Properties of an optically pumped Cesium beam frequency standard with $\phi = \pi$ between the two oscillatory fields", IEEE Trans IM, vol. 43, n°4, pp. 515-520, 1994.

[3] V. Giordano, V. Candelier, A. Hamel, C. Audoin, G. Théobald and P. Cérez, "Noise in the optical detection of atoms in a beam", Optics communications, vol. 67, n°4, pp. 287-292, 1988.

[4] J. Bras, M. Aubourg, IRCOM, Limoges, "Étude de la cavité micro-onde de PHACS par la méthode des éléments finis", Private communication, Sept. 99.

ANALYSIS OF THE SIGNAL OBTAINED WITH THE CS-BRAZILIAN ATOMIC CLOCK

F. TELES, M. S. SANTOS, D. V. MAGALHÃES, A. BEBEACHIBULI AND V. S. BAGNATO
Instituto de Física de São Carlos, USP/SC
Caixa Postal 369, São Carlos /SP
13560-970 - Brazil - Tel: 55 16 2712012 Fax: 55 16 2713616
fteles@if.sc.usp.br monica@if.sc.usp.br dvarela@if.sc.usp.br

Abstract *– As a continuation of our effort to implement scientific time and frequency metrology in Brazil, we report several improvements to our clock which allow us to increase the short-term stability by two orders of magnitude. An analysis of the Ramsey pattern allows the determination of the Rabi frequency and main shifts (Doppler and cavity shifts). The obtained results are within the expected values.*

I) Introduction

For the last three years we have made considerable efforts towards the establishment of a laboratory for the development of scientific time and frequency metrology. As an initial step we have constructed the first atomic clock in Latin America. It is an optically pumped cesium clock, which has been already presented in details [1]. During our first evaluation we identified some frequency shifts and obtained a short term stability of $(1.2 \pm 0.1) \times 10^{-9}$. The relatively high value obtained had been attributed to several causes: thermal isolation in the rf-chain, mechanical and thermal noise in the room and a limited signal to noise ratio. With several improvements and the establishment of our own evaluation system (see details in section V) we were able to improve the short term stability by two orders of magnitude and started implementing techniques to analyze the clock signal.

In this work we present the new evaluation of our clock together with an analysis of the Ramsey fringes, which allow us to obtain the Rabi frequency, the second order Doppler shift and the cavity shift for our present clock conditions.

II) Brief Description of the Clock

The Brazilian atomic clock has been fully described in a recent publication [1]. In brief, it is an optically pumped Cs-beam clock. Before getting into the rf-cavity, atoms are prepared in the $6S_{1/2}(F=3)$ ground state level. After passing through the rf-cavity, they are detected using a laser resonant with the $6S_{1/2}(F=4) \rightarrow 6P_{3/2}(F'=5)$ transition. Collecting optics and detectors allow the fluorescence observation in each region (preparation and detection). The microwave cavity has a U-shape with 19 cm between the two zones. The 9.192 GHz synthesizer was built by Dr. F. Walls (NIST – Boulder – USA), using an external oscillator (Stanford DS345) to frequency modulation. The control of the clock is done by a LabView® program implemented by our group [1]. An auxiliary system, composed by a GPS and a commercial clock allow constant evaluation of the clock performance.

III) Magnetic Field Inhomogeneity

When observing the Ramsey fringes it is known that the Rabi pedestal depends on the sum of transition probabilities in each of the zones [2]. Since those probabilities depend on the Zeeman shift of each transition, it is expected that the Rabi pedestal be shifted from the central fringe when inhomogeneities are present. This shift is well noticed in the obtained spectrum, shown in fig. 1, where all the seven transitions, from $6S_{1/2}(F=3) \rightarrow 6P_{1/2}(F=4)$ sublevels, are seen.

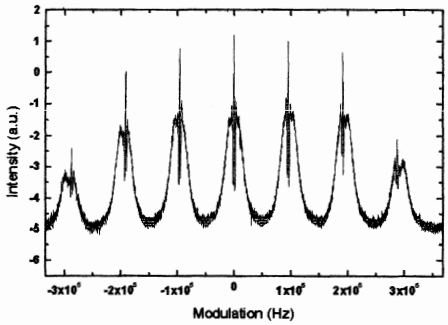


Figure 1 – Observation of the seven transitions before system demagnetization.

The inhomogeneity causing the shift observed in fig. 1 can be analyzed following Shirley et al [3]. In their procedure, the shift in the Rabi pedestal for each of the seven transitions is plotted as a function of m (the Zeeman sublevel quantum number). The result presented in fig. 2 shows that the pedestal shift is linear with m, showing that indeed it is associated with magnetic field inhomogeneity and not cavity pulling or even Rabi pulling. For comparison the last two shifts are also represented in the same diagram of fig. 2.

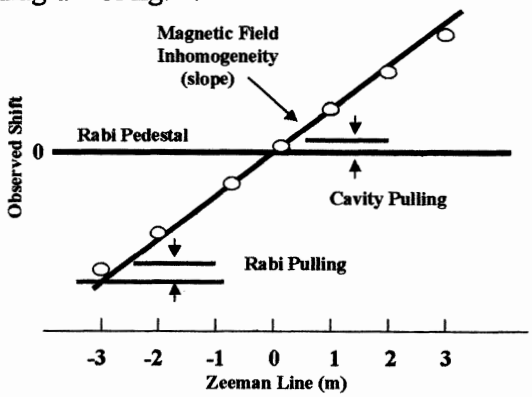


Figure 2 - Pedestal shifts for the seven Zeeman lines as a function of m

After a full procedure of demagnetization the pedestal shift observed in the transitions is fully removed, as observed in fig. 3. The control of the magnetic field inhomogeneity is one of the most important aspects for the improvement in the short term stability recently obtained for our clock (see section V).

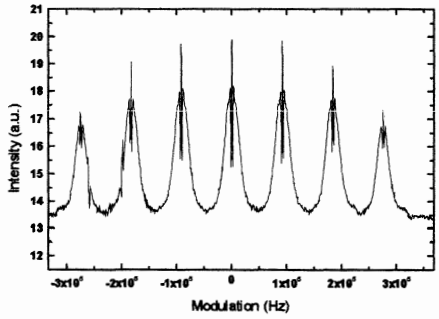


Figure 3 - Seven transitions after demagnetization

IV) Analysis of the Ramsey pattern

After having optimized the clock, we used the analysis procedure proposed by Makdissi et al[4] to determine the values for the Rabi frequency, the second order Doppler shift and the phase shift in our Cs-beam clock.

A. Determination of the Rabi frequency

The output signal in a cesium beam standard is the Ramsey probability averaged over the time of flight (TOF) for all atoms.

If the frequency detuning is small, compared with the Rabi frequency, the observed signal can be written as

$$s(\omega,b)=\frac{1}{2}\int_0^\infty \sin^2(b\tau)\cos(a\omega\tau)f(\tau)d\tau + C_1(b), \quad (1)$$

where ω is the detunning from the atomic resonance, b is the Rabi angular frequency, τ is the TOF for one atom from the velocity distribution across one arm of the Ramsey cavity, $f(\tau)$ is the TOF distribution, a is the ratio of the drift region length L to the excitation region length l and $C_1(b)$ is an ω-independent function of b [2].

Using the cosine transform operator for the function $f(\tau)$, defined as

$$F(\omega)=\int_0^\infty f(\tau)\cos(\omega\tau)d\tau \quad (2)$$

we can rewrite eq. (1) as

$$s(\omega,b)=\frac{1}{2}F(a\omega)*\left[\delta(a\omega)-\frac{1}{2}\delta\left(\frac{2}{a}b-\omega\right)-\frac{1}{2}\delta\left(\frac{2}{a}b+\omega\right)\right], \quad (3)$$
$$+C_1(b)$$

where $\delta(\omega)$ is the Dirac delta function and (*) represents the convolution operation.

Equation (3) shows that the Ramsey pattern can be seen as the convolution of

$F(a\omega)$ with three Dirac pulses located at $\omega_0=0$, $\omega_{-1}=-2b/a$, $\omega_1=2b/a$. Estimating the relative positions of these Dirac pulses we are allowed to determine the Rabi angular frequency in the microwave cavity as

$$b = a\frac{(\omega_1 - \omega_{-1})}{4}. \quad (4)$$

To experimentally determine b we use a derivative method. The second derivative of $s(\omega)$ with respect to ω is

$$s_2(\omega) = \frac{d^2 s(\omega)}{d\omega^2} = -\frac{a}{2}\int_0^{\infty} \tau^2 f(\tau)\sin^2(b\tau)\cos(a\omega\tau)d\tau, \quad (5)$$

which can be written as

$$s_2(\omega) = -\frac{1}{2}F_2(a\omega)^* \left[\delta(\omega) - \frac{1}{2}\delta\left(\frac{2b}{a} - \omega\right) - \frac{1}{2}\delta\left(\frac{2b}{a} + \omega\right)\right]. \quad (6)$$

Here $F_2(a\omega)$ is the cosine transform of $\tau^2.f(\tau/a)$. The multiplication of $f(\tau)$ by τ^2 causes a broadening in the τ domain, therefore $F_2(a\omega)$ is narrower compared to $F(a\omega)$, allowing better determination of the peak position in the frequency domain. Fig.4 shows the Ramsey pattern and Fig. 5 its second derivative. The interval between the two peaks in $S_2(\omega)$ corresponds to 2b. In our case we obtain b = 12957rad/s, a value that is consistent with the used operational parameters.

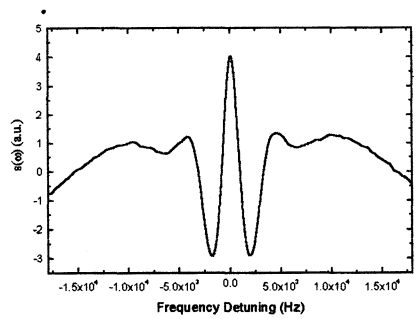


Figure 4 - Ramsey pattern

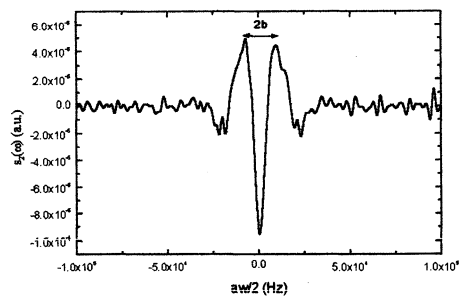


Figure 5 - Second derivative of the Ramsey pattern of Fig. 1

B. Second-order Doppler shift and Cavity phase shift

Using a slow modulation frequency ω_m, the second-order Doppler shift can be obtained by:

$$\frac{\omega_D - \omega_0}{\omega_0} = \frac{\frac{L^2}{2c^2}\int_{-\infty}^{w_m}[s(\omega) - C_1(b)]d\omega}{\left.\frac{ds(\omega)}{d\omega}\right|_{\omega=\omega_m}}, \quad (7)$$

where c is the speed of the light, ω_D and ω_0 are the shifted and rest transition frequencies, respectively [4].

Fig. 6 shows the relative frequency shift versus the modulation amplitude ω_m for two different temperatures of the Cs-oven. The obtained values for the Doppler shift are consistent with the velocity distribution of the atoms previously measured in our system.

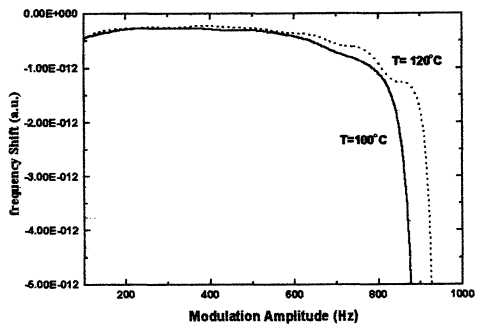


Figure 6 - Second-order Doppler shift

The relative frequency shift due to the cavity phase difference ϕ is given by the cavity phase shift and its value can be obtained from

$$\frac{\omega_\phi - \omega_0}{\omega_0} = \frac{\phi\,\hat{s}(\omega)|_{\omega=\omega_m}}{\omega_0\left.\frac{ds(\omega)}{d\omega}\right|_{\omega=\omega_m}} \quad (8)$$

where $\hat{s}(\omega)$ is the Hilbert transform of $s(\omega)$ [5].

Fig.7 shows the relative frequency shift versus the modulation amplitude ω_m for several employed power levels. Here P_1 is our typical operation power (30mW).

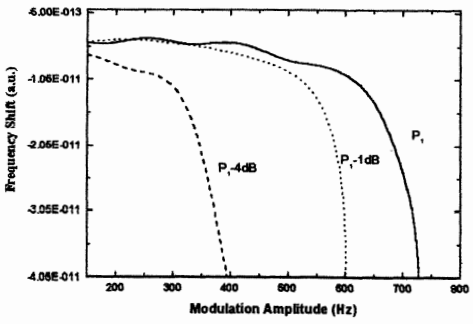


Figure 7 - Cavity phase shift for different power levels.

V) New Evaluation of the Clock Performance

As reported in a previous work [1], we were working on the improvement of an evaluation system.

Our present evaluation system comprises a GPS receiver (Model 9390-6000 – Datum), a commercial atomic clock (HP5061B - HP), a counter (SR620 – Stanford) and a computer, to store the data reading. Using one of the reference sources (GPS receiver or commercial clock) and the signal coming from our cesium beam atomic clock, we compare these signals in the SR620 counter and store the data in a computer connected by a GPIB interface.

With the HP5061B as a reference, we measure the stability of our clock. Using the Allan standard deviation $\sigma(\tau)$ we obtain the results presented in fig. 8.

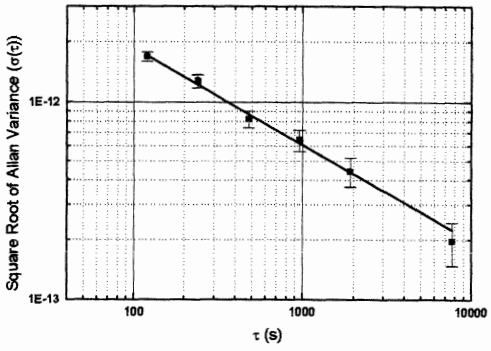


Figure 8 - Allan standard deviation $\sigma(\tau)$ for the Brazilian cesium beam atomic clock. The clock signal is compared with a HP5061B as a reference source.

The data fitting of Fig. 8 results in $\sigma(\tau) = (1.78 \pm 0.2) \times 10^{-11} \tau^{-(0.49 \pm 0.02)}$. The value for the short term stability is two orders of magnitude better than our previous evaluation [1], showing that the changes in our system really had improved it.

Several facts can be associated with the obtained improvement. The better control of magnetic field inhomogeneity (section III), the 50 times improvement of the signal to noise ratio as well as a complete digital amplification scheme.

VI) Conclusion

In this work we report a sequence of improvements which allow us to obtain two orders of magnitude improvement for our Cs-beam clock. Using an analysis of the Ramsey pattern we obtain the values of Rabi frequency, the Doppler shift and the cavity shift for the power levels in the rf-field.

Our present clock is undergoing a series of modifications and about one more order of magnitude improvement is expected. With this, we will then have reached the limits of our designed clock.

We are planning to use the presented evaluation procedure techniques for our atomic fountain presently under construction.

Acknowledges: This work has been supported by Pronex, PADCT and Fapesp.

VII) References

[1] F. Teles, D.V. Magalhães, M. S. Santos, G. D. Rovera and V.S. Bagnato, *"Evaluation of the first Brazilian atomic clock"* in Proceedings of the 14th European frequency and Time Forum, pp. 66-69, 1999.

[2] J. Vanier and C. Audoin, *The Quantum Physics of Atomic Frequency Standards*. Bristol, U. K, Adam Hilger, 1989

[3] J. H. Shierley, W. D. Lee, G. D. Rovera and R. E Drullinger, *"Rabi pedestal shifts as a diagnostic tool in primary frequency standards"* IEEE Trans. Instrum. Meas.,vol 44, no 2 pp.136-139, April 1995.

[4] Ala'a Makdissi and E. Clercq, *"A signal approach analysis of the Ramsey pattern in cesium beam frequency standards"*, IEEE Trans. Instrum. Meas.,vol 46, no 2 pp.112-116, April 1997.

[5] V. Oppenheim and R.W. Schafer, *Discrete-Time Signal Processing*. London, U.K.: Prentice-Hall, 1989.

DESIGN OPTIMIZATION OF A SAPPHIRE LOADED CAVITY FOR A SPACEBORNE HYDROGEN MASER

Takao Morikawa* ,
Ken'ichiro Takahei, Masaro Uehara**, Kenjiro Mori**, Masahiro Tsuda****
*Communications Research Laboratory, 4-2-1 Nukuikitamachi Koganei Tokyo 184 Japan,
Phone: 81-423-27-7554, FAX: 81-423-27-6687, E-mail: tak@crl.go.jp
** Anritsu, 1800 Onna Atsugi Kanagawa 243 Japan

ABSTRACT

TE_{011} mode of a small sapphire loaded dielectric cavity for a spaceborne hydrogen maser is analyzed and its design principle is discussed. The analysis shows that the frequency stability is optimized when the ratio of the outer diameter of the dielectric tube to the inner diameter of the cavity cylinder is between 0.7 and 0.5. The weight of the physics package is minimized when the diameter and the height of the cavity cylinder are equal to each other. The diameter or the volume of the cavity cylinder is determined by the frequency stability required in the application program.

1. INTRODUCTION

For a spaceborne H-maser (SHM), reduction in size and weight is crucial issue, and a sapphire cavity is very effective for this problem without degrading the excellent stability of maser. A few analyses of a sapphire cavity for SHM have been reported [1, 2]. In the analyses, however, the used dielectric loss is inaccurate or even neglected, and there is discrepancy between the cavity quality factors obtained by the analyses and by the theoretical formula when the thickness of the sapphire is zero.

We have solved the Maxwell's equations analytically for TE_{011} mode of a sapphire cavity, and calculated the parameters such as the cavity quality factor Q_o and the filling factor η ' for various shape and volume V_c of the cavity as well as the frequency stability of maser $\sigma_y(\tau)$. From these parameters, the design principle of a sapphire cavity for SHM is discussed in this paper.

Based on the design principle, a sapphire cavity has been made and its measured Q_o agreed very well with the calculation using the measured dielectric loss of sapphire, which is 1×10^{-6} at 1.42GHz and smaller by one order magnitude than the value reported.

2. ANALYSIS

For design of a sapphire cavity for SHM that is shown in Fig.1, four design parameters are necessary. They are a, b, c, and h, or ρ_1, ρ_2, α, a. In this section, it is discussed how to optimize these design parameters from viewpoint of stability and weight of the maser.

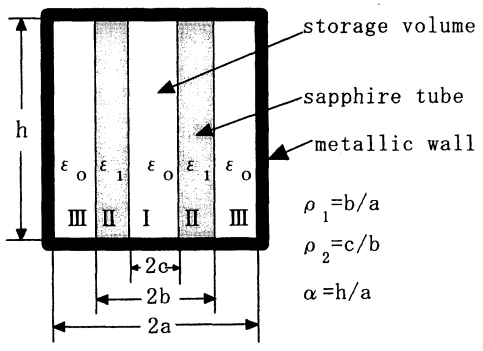


Fig.1 Structure of a sapphire cavity

By solving the Maxwell equation under appropriate boundary conditions for TE_{011} mode of the sapphire cavity, the following characteristic equation is obtained.

$$\frac{S_0J_1(y)+R_0N_1(y)}{S_0J_0(y)+R_0N_0(y)}=\frac{y}{\rho_1 x}\frac{J_1(\rho_1 x)N_1(x)-J_1(x)N_1(\rho_1 x)}{J_0(\rho_1 x)N_1(x)-J_1(x)N_0(\rho_1 x)} \quad (1a)$$

$$S_0=\frac{y}{\rho_1 x}J_1(\rho_1\rho_2 x)N_0(\rho_2 y)-J_0(\rho_1\rho_2 x)N_1(\rho_2 y) \quad (1b)$$

$$R_0=J_0(\rho_1\rho_2 x)J_1(\rho_2 y)-\frac{y}{\rho_1 x}J_1(\rho_1\rho_2 x)J_0(\rho_2 y) \quad (1c)$$

where $\rho_1=b/a$, $\rho_2=c/b$, $x=a\xi_0$, $y=b\xi_1$

$k_i^2=\omega^2\mu_0\varepsilon_i$, $\xi_i=k_i^2-(\pi/h)^2$ (i=0,1)

By solving Eq. (1a), the field in the cavity is obtained, and the cavity dimensions are determined. From Fig.2, which shows the volume of a sapphire cavity V_c for various shape of sapphire tube, one sees that inserting a sapphire tube is very effective to reduce the cavity volume. ρ_2 represents the thickness of the sapphire tube, and $\rho_2=1$ means that the thickness of sapphire tube is zero. As ρ_2 decreases, or the sapphire tube becomes thicker, the cavity volume sharply decreases. One should also notice that larger ρ_1 gives smaller V_c.

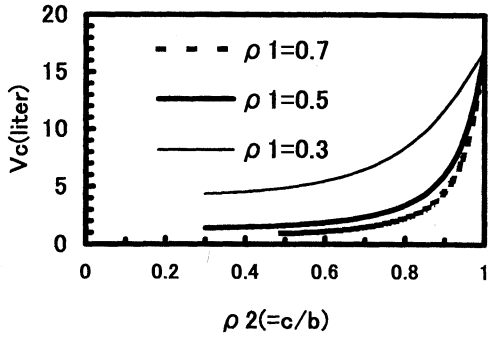


Fig.2 Volume of a sapphire cavity($\alpha=2$)

In order to calculate Q_o, the dielectric loss of sapphire tan δ and the skin depth of the cavity wall δ_s should be known. We have measured tan δ of a sapphire at 13.5GHz and deduced tan δ at 1.4GHz, using the relationship of tan δ /f=constant. The deduced tan δ at 1.4GHz is 1.0×10^{-6}, which is smaller by one order magnitude than the value reported in the Ref.[1]. Fig.3

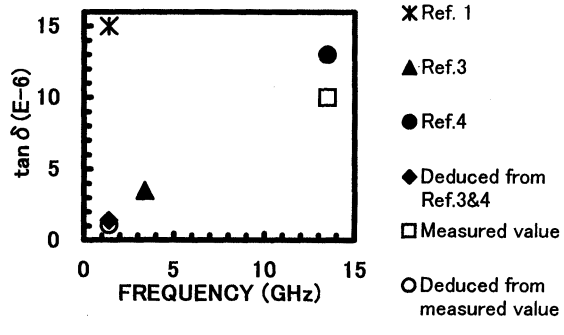


Fig.3 Frequency dependency of tan δ of sapphire

shows the frequency dependency of tan δ of sapphire, and shows that the above tan δ agrees very well with tan δ at 1.4GHz deduced from the data in Ref.3 and 4 . δ_s is estimated to be $3.5\times10^7/\Omega$m from the measured Q of cavity without sapphire.

Though Q_o calculated using the above tan δ and δ_s greatly depends on ρ_1, as shown Fig.4, it converges to the theoretical Q_o of the cavity without sapphire at V_c=16.9*l*. The detailed investigation revealed that the dielectric loss in the sapphire is so small that Q_o mainly depends on the wall loss. Though smaller ρ_1 gives higher Q_o, it gives larger V_c. Thus, some compromise is required.

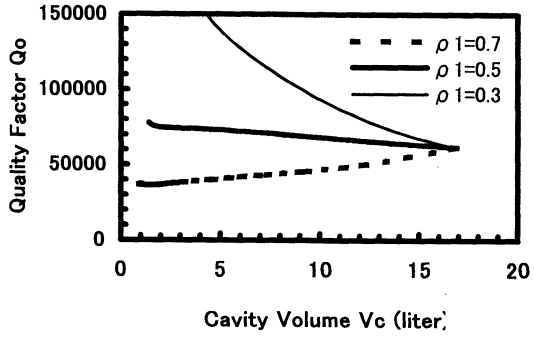


Fig.4 Calculated Q_o of a sapphire cavity ($\alpha=2$)

As shown in Fig.5, η' increases as V_c decreases, and reaches to a maximum at a certain value of V_c. η' decreases very sharply if V_c becomes smaller than this value. η' also depends on ρ_1 and the condition of ρ_1 =0.5 gives the largest maximum value of η', which is 0.5 and slightly larger than the filling factor of

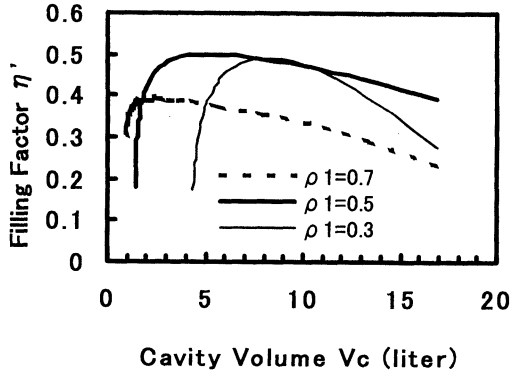


Fig.5 Filling factor of a sapphire cavity ($\alpha=2$)

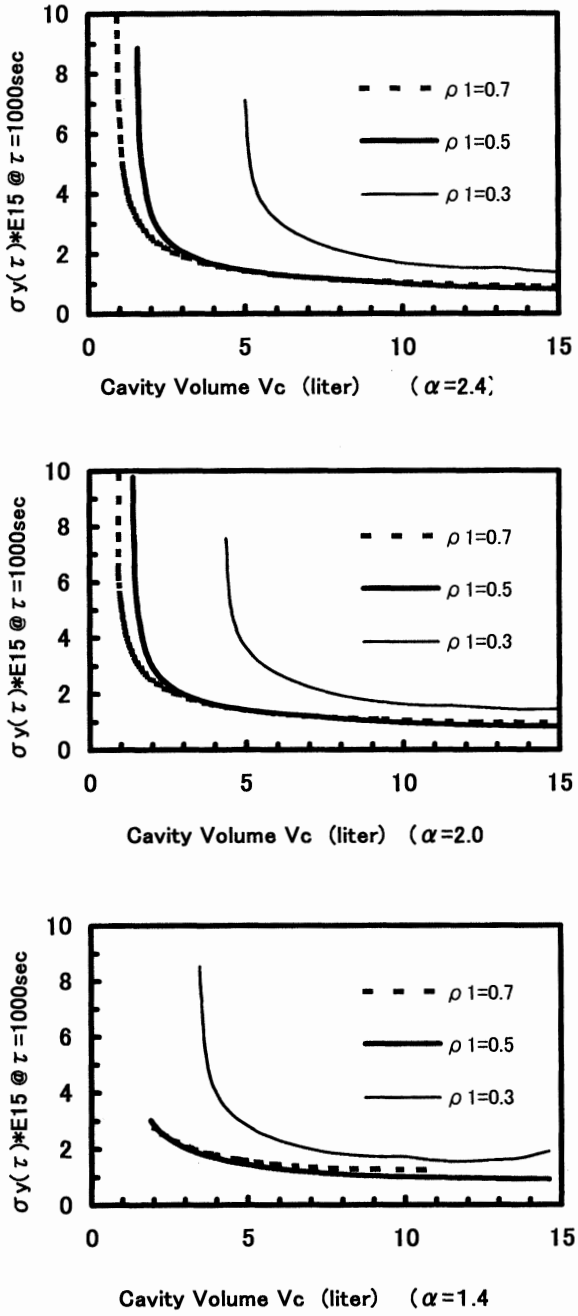


Fig.6 Stability of SHM with a sapphire cavity

conventional full size cavity.

Fig.6 shows the calculated frequency stability of an SHM with the sapphire cavity at averaging time of 1,000s for various α. ρ_1 between 0.7 and 0.5 gives the best frequency stability, which is almost same for all α. Thus, the first design parameter ρ_1 can be determined. Smaller V_c gives worse frequency stability, and the stability becomes worse very sharply if V_c becomes smaller than about 2 liter. Thus, there should be a certain compromise between the frequency stability and the cavity volume, and the minimum V_c can be determined by the stability required in the application program of SHM. As shown in Fig.6, the cavity can be as small as 1 liter, if the stability requirement is only 5×10^{-15} for τ =1,000s. Once V_c is fixed, the second design parameter ρ_2 can be determined from Fig.2.

The third design parameter, α, can be determined from the viewpoint of minimum weight of SHM. Since the cylindrical parts outside the sapphire cavity such as magnetic shields and vacuum envelope account for large proportion of total weight of SHM, it is very important to minimize their weight. The weight of a thin cylindrical shell with endplates can be minimized when the diameter equals with the height if the volume of the shell is constant. If the radial and vertical spacings between the cavity and the magnetic shield outside the cavity are same to each other, the weight of magnetic shield is minimized when α=2.

The fourth design parameter, a, can be easily obtained from V_c and α.

3. EXPERIMENT

Based on this analysis, a sapphire cavity which is shown in Fig.7 has been fabricated. Its dimensions are shown in Table1, and the volume is 3.3 liter.

Table2 shows the calculated and measured parameters of the cavity. The temperature coefficient of the resonance frequency df_c/dT is calculated using the

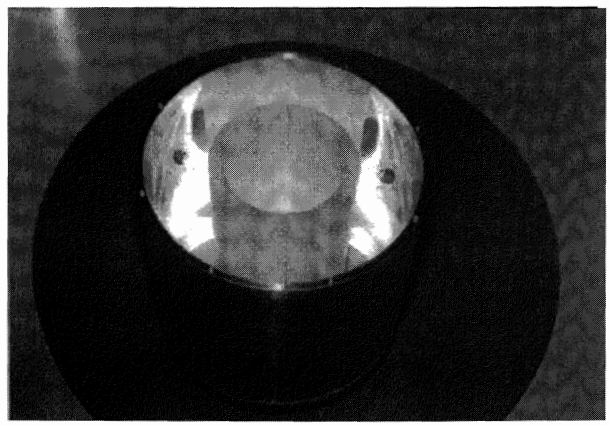


Fig.7 The fabricated sapphire cavity

measured temperature coefficient of ε of sapphire, which is 1.06×10^{-4}/K. As shown in Table 2, the measured loaded Q agrees very well with the calculated one when the above $\tan\delta$ and δ_s is used in the calculation.

Table 1 Dimensions of the sapphire cavity

2a	161.9mm	2c	71.82mm
2b	87.42mm	h	161.9mm

Table 2 Calculated and measured parameters of the sapphire cavity

	calculated	measured
loaded Q	53380	55900
df_c/dT	-80kHz/K	-70.9kHz/K
dfc/da	-6.7MHz/mm	-7.2MHz/mm
dfc/db	-57.2MHz/mm	N.A.
dfc/dc	60.3MHz/mm	N.A.
dfc/dh	-1.13MHz/mm	N.A.

As already mentioned in the analysis, $\tan\delta$ of sapphire is small enough that Q is only slightly affected even if $\tan\delta$ is doubled. Q is more affected by δ_s. Two types of metallic walls were tested. The cylindrical surfaces of both types and the endplates of the first type were chemically silver-plated without brightening agent, while the endplates of the second type is with the agent. The first type gives the loaded Q of 55,900, while the second type gives only 36,600.

4. CONCLUSION

TE_{011} mode of a small sapphire loaded dielectric cavity for an SHM has been analyzed and it has been shown that the design parameters of the sapphire cavity, ρ_1, ρ_2, α, and a can be optimized from the viewpoint of frequency stability and weight of SHM. Based on the analysis, a sapphire cavity was fabricated and its Q_0, df_c/dT, and dfc/da were measured. They well agreed with the analysis using $\tan\delta$ of 1.0×10^{-6}, which is smaller by one order magnitude than the value reported.

5. References

[1] Laurent-Guy Bernier, "Preliminary Design and Breadbording of a Compact Space Qualified Hydrogen Maser Based on a Sapphire Loaded Microwave Cavity", in Proceedings of EFTF 1994, pp.965-980

[2] F. S. Rusin, B. A. Gaygerov, and S. F. Ponomarenko, "Calculation of a Cylindrical Cavity Partially Filled with a Dielectric", Measurement Techniques, vol.28, No. 5, 1985, pp.396-398

[3] A. Nakayama, "A Measurement Method of Complex Permittivity at Pseudo Microwave Frequencies Using a Cavity Resonator Filled with Dielectric Material", IEICE Trans. Electron, vol. E77-C, No.8 June 1994, pp.894-899

[4] Y. Kobayashi and H. Tamura, "Round Robin Test on a Dielectric Resonator Method for Measuring Complex Permittivity at Microwave Frequency", IEICE Trans. Electron., vol. E77-C, No.6 June 1994, pp.882-887

A MODEL AND A MEASUREMENT OF THE CAVITY PULLING EFFECT IN COLD ATOM FOUNTAINS

Y. SORTAIS[1], S.BIZE[1], C. NICOLAS[1], C. MANDACHE[2], G. SANTARELLI[1], C. SALOMON[3] AND A. CLAIRON[1]

[1] BNM-LPTF, 61 Av. de l'Observatoire, 75014 Paris, France,

[2] Institutul National de Fizica Laserilor, Plasmei si Radiatiei, P.O.Box-MG36, Bucuresti, Magurele, Romania,

[3] Laboratoire Kastler Brossel, Ecole Normale Supérieure, 24 rue Lhomond, F-75231 Paris Cedex 05, France.

Thanks to a low collisional frequency shift, ^{87}Rb is a very promising atom in order to further improve cold atom frequency standards accuracy [2, 3, 4]. When using 5×10^6 atoms at launch, in an optical molasses, the collisional frequency shift is near 10^{-16} and the quantum limited short term stability is $5 \times 10^{-14}\tau^{-1/2}$ [1]. On the other hand, the cavity pulling can be as high as 10^{-15} in the same conditions, when using a microwave resonator with a quality factor of 2×10^4. It could be one of the limiting factors of the accuracy of cold Rb fountains in the 10^{-16} range.

The cavity pulling effect is, as in a H-maser, due to the interference between the electro-magnetic field radiated by the magnetic dipole of the atomic sample and the field sustained by the microwave resonator. This interference causes a time-dependent phase of the microwave field when the atoms cross the cavity. It generates a phase difference between the two Ramsey microwave fields, leading to a frequency shift of the clock. This shift exhibits a linear dependence with the number of atoms crossing the microwave cavity and a dispersive shape as a function of the detuning between the resonator frequency and the atomic transition frequency. It increases when the quality factor of the resonator is increased.

We will describe a measurement of the cavity pulling effect in a ^{87}Rb fountain. We will compare this measurement with theoretical predictions. We will discuss the trade-off between the cavity pulling effect and the clock frequency shift due to phase gradients inside the microwave resonator. Finally, we will discuss possible schemes to reduce and evaluate this effect with an accuracy in the 10^{-16} range.

References

[1] Santarelli G., *et. al*, Phys. Rev. Lett. **82** (23), 4619 (1999).

[2] Kokkelmans S. J., *et. al*, Phys. Rev. A **56**, R4389 (1997).

[3] Bize S., *et. al*, Europhys.Lett. **45** (5), 558 (1999).

[4] Bize S., *et. al*, "Collisional frequency shift in ^{87}Rb fountains", submitted to EFTF 2000.

Corresponding author: Y.Sortais[1], Fax: 33 01 43 25 55 42, e-mail:yvan.sortais@obspm.fr

STUDIES ON AUTOMATIC CAVITY PHASE TUNING FOR A CESIUM BEAM FREQUENCY STANDARD

Koji Nakagiri*
Kinki University
Yasusada Ohta, and Noboru Kotake
Communications Research Laboratory, M. P. T. of Japan, and
Mitsuru Kawakita,
Kinki University
*Iwade-Uchita, Wakayama, 649-6493, Japan
Tel +81-736-77-0345, Fax +81-736-77-4754
E-mail: nakagiri@info.waka.kindai.ac.jp

ABSTRACT

In order to realize an automatic Ramsey cavity phase tuning in a cesium beam frequency standard, the cesium beam velocity and the cavity phase control characteristic have been investigated. Our calculation using beam orbit simulation showed that velocity peaks 120 m/s, 170 m/s, and 210 m/s were obtained by the 4-pole magnet of maximum fields at the pole piece 0.4 T, 0.6 T and 0.8 T, respectively. The phase difference between the both ends of Ramsey cavity conducted by a small loop antenna using varactor diode and a screw of 5 mm diameter at the antinode of magnetic field on the E-plane of TE_{10} mode waveguide was investigated. It became evident that the loop antenna which was located at 1.25 λ_g away from the cavity T-feed center. increased the phase of both ends to about 0.7 degree and made the phase difference of 15 % between the both ends.

1. INTRODUCTION

A vertical cesium beam experimental standard has been under construction[1,2] at Kinki University to develop new basic technology to obtain the accuracy of 7×10^{-15}.

In this paper the beam velocity calculation and the preliminary experiments for an automatic Ramsey cavity phase tuning are reported.

2. CESIUM BEAM VELOCITY

The cavity phase shift is proportional to the product of cesium beam velocity by phase difference between 2 microwave interaction regions in a Ramsey cavity. We can set the beam velocities by changing the coil current of 4-pole magnet and measure the frequency dependence to detect and control the phase difference of Ramsey cavity.

Figure 1 shows the structure of cesium beam system and the Ramsey microwave cavity. The beam velocity distribution is calculated by using beam orbit simulation[3], as shown in Fig.2. The parameters 0.4 T, 0.6 T, 0.8 T are the maximum magnetic field intensity at the pole piece, respectively. From the simulation data of Fig.2 the frequency stability is estimated as shown In Table 1. The stability is not so much dependent on the beam velocity. The stability using 2 Frequency optical pumping and magnetic energy selection jointly is 1 order better than that by using magnetic energy selection.

Table 1 Beam characteristic and frequency stability.

	Magnetic field of 4-pole magnet		
	0.4 T	0.6 T	0.8 T
Detected atom number ($\times10^7$)	5.3	10.8	14.1
Mean velocity (m/s)	150	190	220
Resonance width (Hz)	140	190	210
Frequency stability ($\times10^{-12}/\tau^{1/2}$)	7.4	8.4	8.0
Frequency stability by 2 Frequency optical pumping ($\times10^{-13}/\tau^{1/2}$)	6.2	6.7	6.6

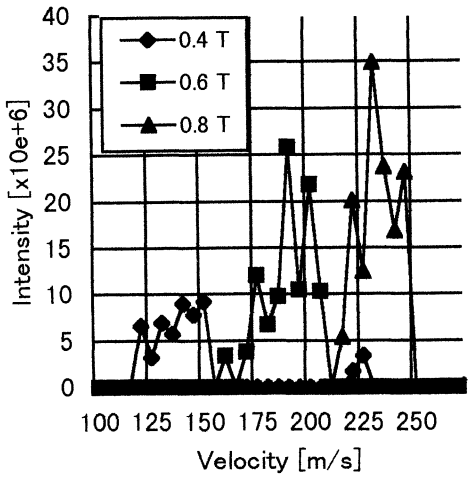


Fig. 2 Velocity distribution for different magnetic fields of 4-pole electromagnet.

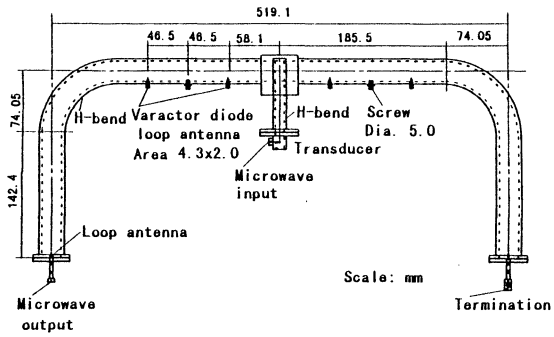


Fig.3 Microwave phase measurement of Ramsey cavity using varactor diode loop antennas and screws.

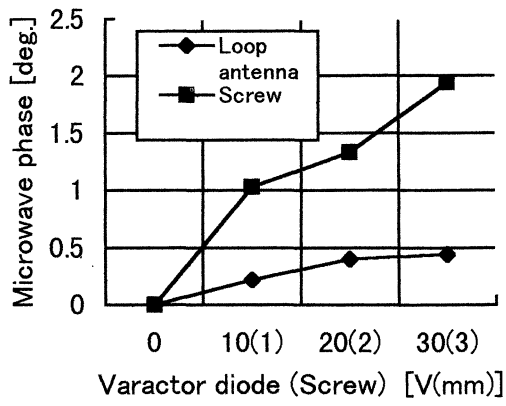


Fig. 4 Microwave phase change in a straight waveguide by a varactor loop antenna and a screw.

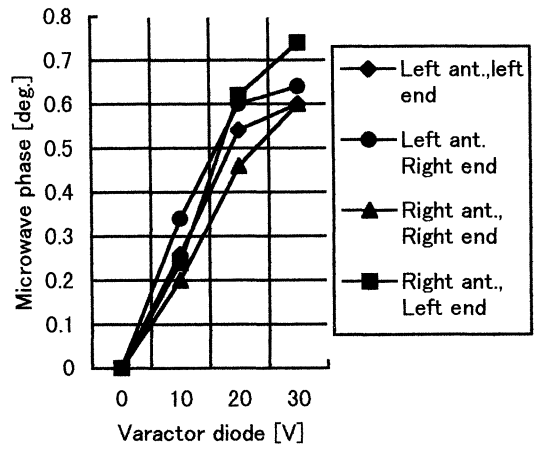


Fig. 5 Microwave phase change of the left and right ends by varactor loop antennas at the position of 1.25 λ_g away from the T-feed center of Ramsey cavity.

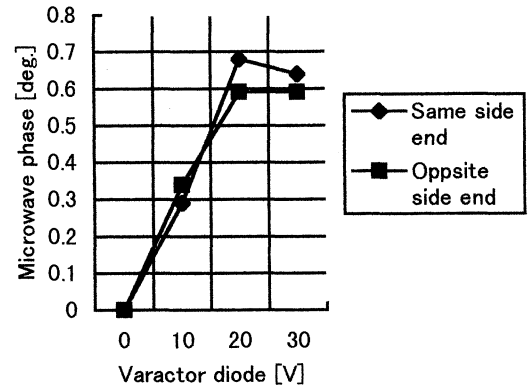


Fig. 6 Microwave phase change of the both ends by left varactor loop antenna at the position of 3.25 λg away from the T-feed center of Ramsey cavity.

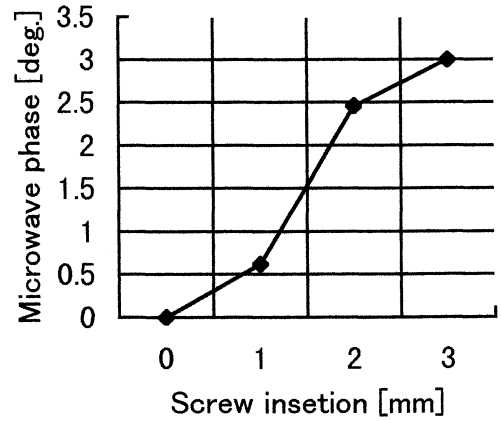


Fig. 7 Microwave phase change by right varactor loop antenna.

3. MICROWAVE PHASE IN RAMSEY CAVITY

The tuning of Ramsey cavity S parameter S_{21} phase by means of a variable capacitance diode placed in a loop antenna and a screw is tested. The measurement setup in an even mode cavity is shown in Fig.3. The loop antenna is connected through a 141 semi-rigid coaxial cable with the external dc bias stripline using a capacitance and a short termination. The varactor diode is tip size commercially for UHF and SHF tuner. The capacitance is 3.85 pF and 0.74 pF by dc bias 2 V and 25 V, respectively.

At fast the phase change through a straight waveguide by loop antenna and screw is measured as shown in Fig.4. The loop antenna is the same of the left one in Fig. 3.

Figure 5 shows the loop antenna characteristic at the position of 1.25 λ_g away from the T-feed center for different output ports. The influence on the opposite end is larger than that on the same side end. Figure 6 shows the left antenna effect at the position of 3.25 λ_g away from the T-feed center. In this case the effect on the same side end is large, but a small quantity. The right loop antenna effect on the same and opposite end is almost the same. The different effect may be due to the small difference between dc bias stripline circuits of left and right antenna.

Figure 7 shows the phase change due to the screw insertion. This amount is about 1.5 times larger than that In Fig.4. In the case of loop antenna is also about 1.5 times larger than that in Fig.4. The effect of screw insertion on the same side and opposite side end is same within the measurement errors.

4. CONCLUSION

We examined the possibility of an automatic cavity tuning for the cesium beam frequency standard. It becomes clear that the beam velocity is changeable from 120 m/s to 210 m/s and the cavity phase control of about 0.1 deg. may be possible by means of a variable capacitance diode placed at the position of 1.25 λ_g away from the T-feed center on the E-plane inside Ramsey cavity.

5. ACKNOWLEDGMENT

The authors would like to thank Mr. Takao Morikawa and Dr. Masatoshi Kajita for their supports and comments. The authors would also like to thank the National Research Laboratory of Metrology and Science and Technology Agency of Japan for their financial help.

6. REFERENCES

[1] K.Nakagiri, J.Umezu, Y.Ohta, M.Kajita, N.Kotake, T.Morikawa, Design and Preliminary Results of a Vertical Cesium Beam Frequency Standard: Proceedings of the 5th Symposium on Frequency Standards and Metrology., Woods Hall, Oct. 1995, pp. 417-418.

[2] K.Nakagiri, N.Kotake, Y.Ohta, M.Kajita, Beam Directivity and Efficiency of Recirculating Oven Using Cylindrical Collimator for Cesium Beam Frequency Standard, Mem. School. B.O.S.T. Kinki University, No.4, 6-53(1998)

[3] K.Nakagiri, H.Okazawa, M.Shibuki,S.Urabe, Beam Trajectry in Cesium-Beam Primary Frequency Standard and Its influence on Frequency Stability and Shift: Japanese Journal of Applied Phisics, Vol. 27, No.12, December, 1988, pp. 2383-2391.

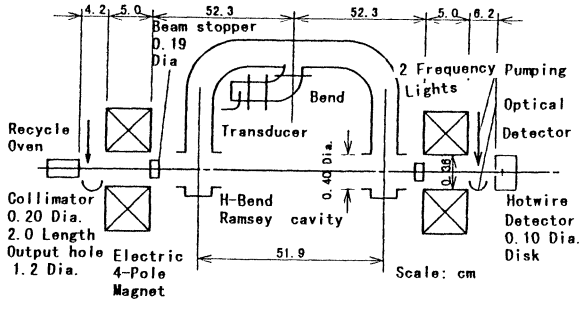


Fig. 1 Structure of cesium beam system and Ramsey cavity.

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

THEORETICAL STUDY OF THE DICK EFFECT IN A CONTINUOUSLY OPERATED RAMSEY RESONATOR

Alain JOYET*, Gaetano MILETI, and Pierre THOMANN
Observatoire cantonal, 58 Rue de l'Observatoire, CH-2000 Neuchâtel, Switzerland
*Tel. (++41) 32 889 6870 Fax (++41) 32 889 6281 e-mail: alain.joyet@ne.ch

Gregor DUDLE
Federal office of metrology, Lindenweg 50, CH-3003 Bern-Wabern, Switzerland

ABSTRACT

It is well establish, that passive frequency standards operated in pulsed mode suffer a degradation of their frequency stability due to the frequency (FM) noise of the Local Oscillator (LO).

In continuously operated frequency standards, it has been shown that a similar degradation of the frequency stability may arise, depending on the used modulation-demodulation scheme. In this paper, we report a theoretical analysis on the possible degradations of the frequency stability of a continuous fountain due to the LO FM noise. A simple model is developed to evaluate whether or not aliasing persists. This model is based on a continuous frequency control loop of a frequency standard using a Ramsey resonator.

From this model we derive a general formula, valid for all usual modulation-demodulation schemes, for the frequency fluctuations of the LO due to aliasing in closed loop operation. We demonstrate that in an ideal situation and for all usual modulation waveforms, no aliasing occurs if the half period of modulation equals the transit time of the Ramsey cavity. We also deduce that in the same condition, square-wave phase modulation is expected to cancel stronger the instability of the LO in closed loop operation. Finally, we show that the "Dick formula" for the specific case of the pulsed fountain can be simply recovered from the model by a sampling operation.

1. INTRODUCTION

In passive frequency standards operated in pulsed mode, like ion traps or cold atom fountains, it has been experimentally demonstrate [1] that this particular mode of operation leads to a degradation of the frequency stability. In this kind of devices the frequency (FM) noise of the Local Oscillator (LO) around harmonics of the pulse rate is downconverted by aliasing into the bandpass of the frequency control loop. This mechanism, called "Dick effect" in the literature because it was first predicted and described by Dick at JPL [2], limits the achievable stability even with state-of-the-art quartz local oscillators.

In continuously operated frequency standards, a similar degradation of the frequency stability may arise. It depends on the scheme of modulation-demodulation used to generate the error signal wich controls the LO and on the value of the modulation frequency. This mechanism, the so called "Intermodulation effect" was first pointed out by Kramer [3], then described in detail by Audoin [4] in the case of a passive cell standard in the quasi-static approximation (interrogation frequency much smaller than atomic resonance linewidth).

We are interested in possible degradations of the frequency stability of a continuous Ramsey resonator due to the LO FM noise. In this type of standards, atoms are continuously interrogated and any phase step in the LO will be detected and will produce an error signal. It is expected that this absence of dead times will constitute a significant advantage over the pulsed fountains, where aliasing is unavoidable.

A generalisation of the Dick formalism for pulsed Ramsey resonators to the continuous case was proposed in [5,6], and briefly addressed in [7]. Our approach is complementary to [6,7] in that we start out with the error signal generated in *the continuous case* and derive the "Dick formula" as a particular case where the atomic detection is gated.

To evaluate whether aliasing nevertheless persists, we consider a simple model, described in section 2, of a continuously operated frequency standard using a Ramsey resonator. In section 3, the error signal of the frequency control loop is derived in a general case involving any modulation-demodulation waveform. Then, we calculate in section 4 the frequency spectrum of the error signal in closed loop operation from the knowledge of the power spectral density $S_y^{LO}(f)$ of fractional frequency fluctuations of the LO. Finally, we deduce from it the power spectral density $S_y^{LLO}(f)$ of fractional frequency fluctuations of the Locked Local Oscillator (LLO) due to aliasing. The result constitutes a generalisation to continuous or modulated fountain Ramsey resonators of the Dick formula, wich was

derived specifically for pulsed resonators. To end, we show in the last section that it is possible to recover from the model the "Dick formula" for the specific case of the pulsed fountain by a sampling operation.

2. MODEL OF FREQUENCY CONTROL LOOP

In the present derivation of the frequency stability limitation of the LLO, we focus our attention on passive frequency standards using a continuous Ramsey interrogation scheme. A sketch of the considered continuous fountain frequency standard is represented in Fig.1.

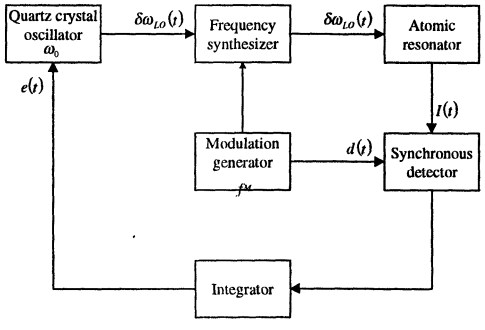


Fig. 1. Block-diagram of the continuous fountain frequency standard

We assume a conventional quartz crystal oscillator characterised by its power spectral density of fractional FM noise $S_y^{LO}(f)$. The atomic resonator is supposed to be an ideal Ramsey resonator with two infinitely short Rabi interactions. The signal at the output of the resonator is synchronously detected at modulation frequency f_M. Then, after lowpass filtering, it forms the continous error signal used to control the LO. We still assume that the frequency synthesizer, the synchronous detector and the modulation generator add no noise to the error signal.

3. ERROR SIGNAL OF THE FREQUENCY CONTROL LOOP

The signal at the output of the assumed resonator is given, in a good approximation, by [8] :

$$I(t) = I_b + \frac{1}{2} I_0 \{1+\cos[\phi(t,T)]\} \qquad (1)$$

for monokinetic atoms and optimum power. $\phi(t,T)$ is the apparent phase difference of the interrogation signal between the two pulses and T is the transit time between these pulses. If we suppose that no permanent phase difference exists between the two cavity arms and that there is no detuning between the carrier frequency of the interrogation signal and ω_0 the resonance angular frequency of cesium atoms, the remaining apparent phase difference is composed of two terms :

$$\phi(t,T) = \Delta\phi_{\text{mod}}(t,T) + \delta\phi_{LO}(t,T) \qquad (2)$$

The first one is due to the generated frequency (or phase) modulation :

$$\begin{aligned}\Delta\phi_{\text{mod}}(t,T) &= \phi_{\text{mod}}(t) - \phi_{\text{mod}}(t-T) \\ &= \phi_M\, c(t,T)\end{aligned} \qquad (3)$$

where $\phi_{\text{mod}}(t)$ is assumed to be an odd and periodic function of period T_M. ϕ_M is the phase modulation depth and $c(t,T)$ is a periodic function of period T_M containing odd harmonics only, such that the maximum of $|c(t,T)|$ is unity. The second term in eq. (2) is a random component due to the frequency (or phase) fluctuations of the LO :

$$\begin{aligned}\delta\phi_{LO}(t,T) &= \delta\phi_{LO}(t) - \delta\phi_{LO}(t-T) \\ &= \int_{t-T}^{t} \delta\omega_{LO}(\tau)\, d\tau\end{aligned} \qquad (4)$$

where $\delta\omega_{LO}(t)$ is a random function representing the angular frequency fluctuations of the LO. With these assumptions, the signal at the output of the resonator reads :

$$\frac{I(t)}{I_0} = \frac{I_b}{I_0} + \frac{1}{2}\{1+\cos[\phi_M c(t,T) + \delta\phi_{LO}(t,T)]\} \qquad (5)$$

With the actual LOs, the standard deviation of $\delta\phi_{LO}(t,T)$ is much smaller than the modulation depth ϕ_M wich is usually of the order of $\pi/4$, that is :

$$\delta\phi_{LO}(t,T) << \phi_M \qquad (6)$$

Then we can expand the above expression limited to the first order to get :

$$\begin{aligned}\frac{I(t)}{I_0} = \frac{I_b}{I_0} &+ \frac{1}{2}\{1+\cos[\phi_M c(t,T)]\} \\ &- \frac{1}{2}\delta\phi_{LO}(t,T)\sin[\phi_M c(t,T)]\end{aligned} \qquad (7)$$

This signal is then synchronously detected at f_M and lowpass filtered with cut-off frequency $f_c << f_M$. Since the spectrum of the periodic function

$c(t,T)$ contains only odd harmonics, it follows that the Fourier series expansion of $\cos[\phi_M c(t,T)]$ and of $\sin[\phi_M c(t,T)]$ contains only even and odd harmonics respectively. Therefore, only the third term of eq. (7) will provide the error signal.

If $d(t)$ denotes the demodulation waveform of the synchronous detector, the error signal wich controls the LO is given by :

$$e(t) = -\frac{1}{2} K \delta\phi_{LO}(t,T) \sin[\phi_M c(t,T)] d(t) \qquad (8)$$

where K is a constant wich depends on the synchronous detector gain. Eq. (8) is the basic relationship of our simple model and will be analysed below in order to find the power spectral density of the frequency fluctuations of the LLO.

4. FREQUENCY SPECTRUM OF THE LLO

We are now interested to calculate the spectrum of the error signal defined by eq. (8). It is composed of two parts :

- a random part $\delta\phi_{LO}(t,T)$ depending on the angular frequency fluctuations $\delta\omega_{LO}(t)$ of the LO. We assume in the following that the process $\delta\omega_{LO}(t)$ is stationary and of zero mean value.
- a deterministic part $\sin[\phi_M c(t,T)] d(t)$ wich depends on the used modulation-demodulation scheme.

Let us start by calculating the two-sided power spectral density $S'_{\delta\phi_{LO}}(f)$ of the random part $\delta\phi_{LO}(t,T)$ given by :

$$\delta\phi_{LO}(t,T) = \int_{t-T}^{t} \delta\omega_{LO}(\tau)\, d\tau = \delta\omega_{LO}(t) * h(t) \qquad (9)$$

where the symbol $*$ stands for convolution and $h(t)$ is a rectangular impulse equal to 1 for $0 \le t \le T$ and 0 elsewhere. The autocorrelation function $R_{\delta\phi_{LO}}(\tau)$ of $\delta\phi_{LO}(t,T)$ is given by :

$$R_{\delta\phi_{LO}}(\tau) = R_{\delta\omega_{LO}}(\tau) * h(-\tau) * h(\tau) \qquad (10)$$

where $R_{\delta\omega_{LO}}(\tau)$ is the autocorrelation function of $\delta\omega_{LO}(t)$.

The two-sided power spectral density is obtained by taking the Fourier transform of $R_{\delta\phi_{LO}}(\tau)$, we get :

$$S'_{\delta\phi_{LO}}(f) = |H(f)|^2 S'_{\delta\omega_{LO}}(f) \qquad (11)$$

$H(f)$ is the transfer function of $h(t)$ and $S'_{\delta\omega_{LO}}(f)$ is the two-sided power spectral density of the angular frequency fluctuations of the LO.

Let $a(t)$ be the deterministic part of the error signal :

$$a(t) = \sin[\phi_M c(t,T)] d(t) \qquad (12)$$

As the demodulation waveform $d(t)$ is assumed to contain only odd harmonics of the frequency of modulation f_M, $a(t)$ will contain even harmonics only. Its autocorrelation function $R_a(\tau)$ is then the following :

$$R_a(\tau) = a_0^2 + \sum_{k=1}^{+\infty} (a_k^2 + b_k^2)\, (e^{i2k\omega_M\tau} + e^{-i2k\omega_M\tau}) \qquad (13)$$

where a_k, b_k are the Fourier components of the Fourier series of $a(t)$. These coefficients are in principle calculable from the knowledge of the modulation and demodulation waveforms. The two-sided power spectral density follows immediately :

$$S'_a(f) = a_0^2\, \delta(f) + \sum_{k=1}^{+\infty} (a_k^2 + b_k^2)\, [\delta(f - 2kf_M) + \delta(f + 2kf_M)] \qquad (14)$$

with $\delta(f)$ the Dirac impulse function.

The autocorrelation function of the error signal is given by :

$$R_e(\tau) = \frac{K^2}{4} R_{\delta\phi_{LO}}(\tau)\, R_a(\tau) \qquad (15)$$

since $\delta\phi_{LO}(t,T)$ and $a(t)$ are two independent variables [9]. The two-sided power spectral density $S'_e(f)$ of the error signal is then readily obtained :

$$S_e'(f) = \frac{K^2}{4} S_{\delta\phi_{LO}}'(f) * S_a'(f) \qquad (16)$$

$$= \frac{K^2}{4} a_0^2 |H(f)|^2 S_{\delta\omega_{LO}}'(f) + \frac{K^2}{4} \sum_{k=1}^{+\infty} (a_k^2 + b_k^2)$$

$$\times \left[|H(f - 2kf_M)|^2 S_{\delta\omega_{LO}}'(f - 2kf_M) \right.$$

$$\left. + |H(f + 2kf_M)|^2 S_{\delta\omega_{LO}}'(f + 2kf_M) \right]$$

This spectrum is composed of two parts. The first one wich contains $S_{\delta\omega_{LO}}'(f)$ corresponds to the error signal that actually controls the frequency fluctuations of the LO. The second one , wich contains all the even multiples of f_M , corresponds to a spurious error signal generated by down-conversion of frequency fluctuations of the LO at higher harmonics of the frequency of modulation, this is the aliased part of the spectrum.

In closed loop operation, the LLO frequency spectrum is readily obtained from that of the error signal in open loop. It is the power spectral density of the LO divided by the open loop gain plus the aliased part of the spectrum of the error signal, into the bandwith f_c of the frequency control loop. If we assume that $f_c << f_M$ wich is usually the case in synchronous detection and introducing one-sided power spectral densities, the LLO frequency spectrum is then given by :

$$S_{\delta\omega_{LLO}}(f) \cong \frac{4 S_{\delta\omega_{LO}}(f)}{K^2 a_0^2 |H(f)|^2} \qquad f \leq f \qquad (17)$$

$$+ 2 \sum_{k=1}^{+\infty} \frac{a_k^2 + b_k^2}{a_0^2} \frac{|H(2kf_M)|^2}{|H(f)|^2} S_{\delta\omega_{LO}}(2kf_M)$$

We are interested here only in the aliased part of the spectrum. Introducing the expression of the transfer function $H(f)$ and evaluating the aliased part of the spectrum at $f = 0$, we obtain finally the following expression for the power spectral density of fractional frequency fluctuations of the LLO due to aliasing :

$$S_y^{LLO}(0) = 2 \sum_{k=1}^{+\infty} \frac{a_k^2 + b_k^2}{a_0^2} \sin_c^2\left(2k\pi \frac{T}{T_M}\right) S_y^{LO}(2kf_M) \qquad (18)$$

where $\sin_c(x)$ is the cardinal sine function and $S_y^{LO}(f)$ the power spectral density of fractional frequency fluctuations of the LO. This formula is the basic relationship to analyse the effect of any type of modulation-demodulation scheme on the stability of the LLO frequency.

Without considering any particular scheme of modulation-demodulation, we can already point out some general features of the result (18). We can see immediately that if the half period of modulation $T_M / 2$ is equal to the transit time T, the power spectral density $S_y^{LLO}(f)$ of the LLO vanishes, because the term involving the ratio T / T_M is equal to zero. This is true for *all usual modulation-demodulation schemes*, since the latter intervenes in the formula only through the Fourier coefficients a_k and b_k. This result has already been obtained by Makdissi [7] in the context of an extension of the sensitivity function [2] from the pulsed to the continuously operated resonator.

Moreover, we can see easily that for the same condition, square wave phase modulation is expected to cancel stronger the instability of the LLO due to aliasing, since in this case the Fourier coefficients a_k and b_k of $a(t)$ are all equal to zero ($k \neq 0$), the function $a(t)$ being a constant. This fact will probably be a good feature in actual cases with atomic velocity distribution and finite Rabi interactions.

We can summarize these results in few words in saying that, in these cases, the sensitivity of the LLO frequency to frequency fluctuations of the LO is a constant over the time.

5. REDUCTION TO THE PULSED CASE

It is interesting now, before analysing further the model, to compare it to the well-known case of the Dick effect appearing in the pulsed mode operation of a cesium fountain frequency standard. In this kind of device, the error signal is not always available to control the LO but only at periodic instants, inherent to the pulsed operation of the fountain itself.

In order to reduce to this case, we assume that the function $a(t)$ is periodic of period T_c, where T_c is the period of one cycle.

The periodic availability of the error signal can be taken into account in our model by sampling the assumed constant atomic beam intensity in such a way that it would be different from zero only at periodic instants, T_c apart in time. The error signal takes then the following form :

$$e(t) = -\frac{1}{2} K \, \delta\phi_{LO}(t,T) \, a(t) \, u(t) \qquad (19)$$

where $u(t)$ is a Dirac comb given by :

$$u(t) = \sum_{k=-\infty}^{+\infty} \delta(t - kT_c) \qquad (20)$$

Within these assumptions and doing again the spectral analysis of the so defined error signal, we find for the power spectral density of fractional frequency fluctuations of the LLO :

$$S_y^{LLO}(f) = 2 \sum_{k=1}^{+\infty} \sin_c^2\left(k\pi \frac{T}{T_c}\right) S_y^{LO}(kf_c) \qquad (21)$$

This formula corresponds exactly to the well-known "Dick formula" in the case of a pulsed fountain. Indeed, the cardinal sinus term is equal to the ratio g_k^2 / g_0^2 of the Fourier coefficients of the Fourier series development of the sensitivity function $g(t)$, as defined in [1].

6. CONCLUSION

A simple model of an ideal assumed continuously operated fountain frequency standard has been developed to investigate possible degradations of the frequency stability of the LLO. This model has shown that if the half period of modulation $T_M / 2$ is equal to the transit time T, no frequency instability due to an aliasing effect appears. This is true for all usual modulation-demodulation schemes. The case of a pulsed frequency standard can be recovered by sampling the continuous atomic beam at the cycle frequency, leading to the well-known "Dick formula".

The extension of the model, to take into account the effects of a velocity distribution of atoms and the finite Rabi interactions, is under investigation. We expect that these effects will introduce only a small degradation of the frequency stability, not preventing from reaching the "shot noise" stability limit with a conventional quartz crystal oscillator and state-of-the-art frequency chain.

This model also sets a solid basis for future experimental investigations.

7. REFERENCES

[1] G. Santarelli, C. Audoin, A. Makdissi, P. Laurent, G.J. Dick, and A. Clairon, "Frequency Stability Degradation of an Oscillator Slaved to a Periodically Interrogated Atomic Resonator", IEEE Trans. Ultrason., Ferroelect., Freq. Control., vol. 45, n° 4, pp. 887-894, July 1998.

[2] G.J. Dick, "Local Oscillator Induced Instabilities in Trapped Ion Frequency Standards", in Proc. of the 19th Precise Time and Time Interval (PTTI), Redendo Beach, 1987, pp. 133-147.

[3] G. Kramer, "Noise in passive frequency standard", in Proc. of CPEM74, 1974, pp. 157-159.

[4] C. Audoin, V. Candelier, and N. Dimarcq, "A Limit to the Frequency Stability of Passive Frequency Standards due to an Intermodulation Effect", IEEE Trans. On Instr. and Measur., vol. 40, n°2, pp. 121-125, April 1991.

[5] A. Makdissi, E. de Clercq, C. Audoin, and A. Clairon "Generalisation of the Dick Effect to Continuous Beam Frequency Standards", in Proc. of 1997 IEEE Int. Freq. Contr. Symp., 1998, pp. 263-269.

[6] A. Makdissi, "Traitement de Signal Appliqué aux Etalons Primaires de Fréquence: Amélioration de leur Exactitude et de leur Stabilité", Thèse, Université de Paris-sud, Orsay, 1999, ch. 6.

[7] L. Lo Presti, G.D. Rovera, and A. De Marchi, "A simple analysis of the Dick effect in terms of phase noise spectral densities", IEEE Trans. Ultrason., Ferroelec., Freq. Contr., vol. 45, n° 4, pp. 899-905, July 1998.

[8] J. Vanier and C. Audoin, The Quantum Physics of Atomic Frequency Standards, Bristol, UK: Adam Hilger, 1989, ch. 5, p. 634.

[9] W. B. Davenport Jr. and W. L. Root, An Introduction to the Theory of Random Signals and Noise, New York : IEEE Press, 1987, ch. 4.

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

HIGHER-ORDER STARK EFFECT AND BLACKBODY RADIATION SHIFT IN Cs

Yu. S. Domnin and V. G. Pal'chikov, IMVP GP VNIIFTRI

Institute of Metrology for Time and Space at National Research Institute for Physical-Technical and Radiotechnical Measurements, Mendeleevo, Moscow Region, 141570 Russia. Phone: 007095 5350849, Fax: 007095 5359334, E-mail: ydomnin@imvp.aspnet.ru, vitpal@mail.ru

ABSTRACT

The purpose of the present work is to develop a theory of higher-order Stark effect and blackbody radiation shift in Cs atoms. We have calculated the scalar and tensor dipole polarizabilities (α) and hyperpolarizabilities (γ) of the ground and excited states of cesium. Our theory includes both fine- and hyperfine structure effects. Semiempirical and accurate electron-correlated wave functions have been used to determine the static values of (α) and (γ). Numerical calculations are carried out using sums of oscillator strengths and, alternatively, with the Green function for the excited valence electron including integral over continuum for the radial matrix elements. The results are compared with high-precision measurements and theoretical data.

1. INTRODUCTION

Cs atom is used in a number of important applications including atomic clocks , laser cooling, parity violation experiments, the search for a permanent electron dipole moment, atom interferometry, etc. Precise measurements of Stark effect provide information about susceptibilities of atomic states that are important for describing of atomic properties including van der Waals constants and dielectric constants , anticrossings of atomic levels occuring in the Stark splitting of multiplet at an electric field.

The interaction between Cs atoms and the electric fields of blackbody radiation (BR) produces the Stark effect on the hfs of sublevels for the ground state [1], which leads to a shift in the Cs frequency standard [2,3]. It is important to note, that BR shift is one of the dominant uncertainty term, which could limit the accuracy of this kind of the frequency standards.

The main goal of this paper is to analyze the higher-order Stark effect on hfs –components of Cs atoms both in the ground and excited states as well as the temperature-dependent shift of $6S_{1/2}$ hfs splitting due to the BR. The contents of the present paper can roughly be divided into two main parts.

The first part contains the general theory of higher-order Stark effect including hyperfine structure effects. We modify and extend the theory in several ways. We consider the problem of the hfs in the presence of a uniform electric field using rather general angular momentum arguments to show how the different hfs-interactions contribute (magnetic dipole contact interaction, magnetic dipole orbital interaction, magnetic spin-dipolar interaction and electric quadrupole interaction). It is shown, in particular, that the BR shift appears in the third- and fifth orders of perturbation theory, where the electric-field interaction is taken 2 times (polarizabilities) and 4 times (hyperpolarizabilities) and the hyperfine interaction is taken only once.

We employ two alternative calculating approaches based on the sums of oscillator strengths and on the Green' function method in the framework of the Fues' model potential approximation. The results are compared with experiments [2,3] and with calculations by other authors [1].

In the second part, we considered the dc hyperfine Stark shifts on $6P_{3/2}$ excited state in Cs both theoretically and experimentally using specially designed cesium cell with a resolution on the level 10^{-11}.

2. FORMULATION OF METHOD

The dipole moment induced in an atom by a uniform electric field F is for most purposes expressed as a linear function of F, but higher terms in F may become important when the field strength becomes sufficiently large. The induced dipole moment d_e is given by

$$d_e = \alpha F + \frac{1}{6}\gamma F^3 + \ldots,$$

in which α is the electric dipole polarizability, and the γ represents the deviation from a linear polarization law. The paramter γ is often termed the electric dipole hyperpolarizability. Such a deviation from linearity corresponds to fourth-order terms in F in the phenomenogical expression

$$E - E_0 = \Delta E = -\alpha F^2/2! - \gamma F^4/4! - \ldots,$$

where E_0 is the field-free energy.

The polarizability α_{nLJM} and hyperpolarizability γ_{nLJM} are conveniently expressed in terms of the second-order through forth-order matrix elements of the atomic dipole moment

$$\alpha_{nLJM} = 2 < nLJM|DG_{nLJ}D|nLJM >,$$
$$\gamma_{nLJM} = 24[< nLJM|DG_{nLJ}DG_{nLJ}D$$
$$\times G_{nLJ}D|nLJM > -\frac{1}{2}\alpha_{nLJM}S_{nLJM}(-3)].$$

Here

$$S_{nLJM}(-3) =< nLJM|D(G_{nLJ})^2 D|nLJM >,$$

and

$$G_{nLJ} = \sum_{|m>\neq|nLJ>} \frac{|m><m|}{E_m - E_{nLJ}}$$

is the reduced Green function of the atom with the energy E_{nLJ}, the state $|nLJ>$ being excluded from the sum over the complete set of $|m>$. D is the z-component of the electric dipole moment of the atom.

The structure of α_{nLJM}, γ_{nLJM} and methods for the numerical calculations were discussed in detail in our previous paper [4]. After the angular integration, with the help of standard methods of the theory of angular momentum the static hyperpolarizability tensor may be resolved into irreducible parts,

$$\gamma_{nLJM} = \sum_j C^{JM}_{JMj0}\gamma_j(nLJ)$$

Here the standard designations for Clebsch-Gordan coefficients are used. It should be noted that the general expression for hyperpolarizability is similar to that of the polarizability, which can be written in terms of the scalar (s) and tensor (t) parts

$$\alpha_{nLJM} = \alpha^s_{nLJM} + \alpha^t_{nLJM}\frac{3M^2 - J(J+1)}{J(2J-1)}$$

The irreducible parts of hyperpolarizability are non-zero by only three invariant atomic parameters. First of them, $\gamma_0(nLJ)$,corresponds to the scalar part of hyperpolarizability, $\gamma_2(nLJ)$ corresponds to the rank 2 tensor part of hyperpolarizability, $\gamma_4(nLJ)$ determines the tensor part of the rank 4.

Finally, the hyperpolarizability γ_{nLJM} can be represented in a general form [4]

$$\gamma_{nLJM} = \gamma_0(nLJ) + \gamma_2(nLJ)\frac{3M^2 - J(J+1)}{J(2J-1)}$$
$$+\gamma_4(nLJ)\frac{3(5M^2 - J^2 - 2J)(5M^2 + 1 - J^2)}{J(2J-1)(2J-2)(2J-3)}$$
$$-\gamma_4(nLJ)\frac{10M^2(4M^2-1)}{J(2J-1)(2J-2)(2J-3)}.$$

Since the zero-field separation between fine-structure components is essentially less than the energy separations between states with different principal and/or orbital quantum numbers, degenerate perturbation theory should be used for determining the energy shift.

Thus, the energy shift $\Delta E_{nLJM} = E - E_{nLJM}$ is the solution of the secular equation

$$\det||\Delta E_{nLJM}\delta_{JJ'} - V_{JJ'}|| = 0.$$

When determining the energy corrections up to the fourth order in F, we have to calculate the matrix elements in this equation to the same order. The second- and fourth-order corrections to the diagonal matrix elements are determined by the polarizability and hyperpolarizability tensors. For the off-diagonal matrix element $V_{JJ'}$ which is of second order in F, it is sufficient to take into account only the first nonvanishing term of the corresponding power series because it appears in the secular equation always in a product with another nondiagonal matrix element.

Accurate calculations of the reduced matrix elements for the components of the polarizability and hyperpolarizability require reliable wavefunctions which should give correct values for both of the quantities. The calculations of the atomic susceptibilities were carried out using a sum of oscillator strengths. We have used precise relativistic single-double calculations of reduced dipole matrix elements [5]. To control the results we have also carried out similar calculations with the Fues' model potential method. A simple analytical presentation for the model-potential Green function enables us to evaluate the contribution of the high-n terms (including continuum) which are neglected in the finite-sum approach.

3. FUES' MODEL POTENTIAL METHOD

The Fues' model potential (FMP) for describing the states of a valence electron in a many-electron atom is

a rather useful tool for simple and straightforward calculations of the optical transition amplitudes in any order of interaction between an atom and an external electromagnetic field. It was proposed first by Simons [6] as analytically solvable example of atomic pseudopotential which may be used in generating valence and Rydberg orbitals for calculating oscillator strengths and ionization amplitudes specifically for atoms with a single valence electron.

However, it was noted, that the method, giving quite accurate results for excited states, failed to do so for the transitions involving the ground-state wave functions. In this communication we have used a new approach for determining the parameters of the FMP providing accurate numerical results in calculating static optical characteristics without resorting to any other kind of a wave function for the states of atom [4].

The FMP method is based on the Rydberg formula for the energy of an atomic level $|nl>$ with the effective principal quantum number ν_{nl}:

$$E_{nl} = -\frac{Z^2}{2\nu_{nl}^2},$$

where Z is the charge of the residual ion. Corresponding radial wave function of the state has the hydrogen-like form,

$$R_{nl} = (-1)^k \frac{2}{\nu_{nl}^2} \sqrt{\frac{Z^3 n_r!}{\Gamma(\nu_{nl}+\lambda_l+1)}} \left(\frac{2Zr}{\nu_{nl}}\right)^{\lambda_l}$$

$$\times e^{-Zr/\nu_{nl}} L_{n_r}^{2\lambda_l+1}\left(\frac{2Zr}{\nu_{nl}}\right),$$

and depends on the choice of the radial quantum number n_r, which in its turn determines the effective angular momentum quantum number, $\lambda_l = \nu_{nl} - n_r - 1$ (a smooth dependence of λ_l on energy of the level in most cases may be neglected). In order to make complete the set of valence electron states with the given angular momentum l, observed in experiments, the value of n_r =0 was usually taken for the lowest state with given l , consecutively increasing by unity for higher states. This definition caused substantial inaccuracy (about 50% in absolute value, with an opposite sign in some cases) of the radial matrix elements $<nl|r^L|n'l'>$ with L=1. That is why the nodeless radial wave functions were replaced by the one-node wave functions. Such a replacement is quite equivalent to another definition of the radial quantum number, i.e. n_r = 1 for the ground state and the effective angular momentum $\lambda = \nu_g - 2$. This definition implies the existence of an additional "imaginary" state with the effective principal quantum number $\lambda_{im} = \nu_g - 1$, but the "imaginary" state cannot provide significant contributions to the optical transition amplitudes. The value of $k = n - n_r - l - 1$ should be taken in the phase factor $(-1)^k$, in accord with the phase of corresponding "hydrogenlike" wave function.

In the present paper, we used the analytical representation of the Green function for the Fues' model potential for summing over the intermediate states by resolving into the infinite series over the spherical-harmonic basis [4]:

$$G_E(\vec{r}_1,\vec{r}_2) = \sum_{lm} g_l(E;r_1,r_2) Y_{lm}(\vec{n}_1) Y_{lm}^*(\vec{n}_2),$$

where $\vec{n} = \vec{r}/r$ is the unit vector, and the coefficients of these series are the radial Green functions, which in their turn may be resolved into Sturm-function series, as follows

$$g_l(E;r,r') = -\frac{4Z}{\nu}(xx')^l \exp\left(-\frac{x+x'}{2}\right)$$

$$\times \sum_{k=0}^{\infty} \frac{k! L_k^{2\lambda_l+1}(x) L_k^{2\lambda_l+1}(x')}{\Gamma(k+2\lambda_l+2)(k+\lambda_l+1-\nu)}$$

and $x = 2Zr/\nu$, $\nu = Z/(-2E)^{1/2}$, $L_m^l(x)$ is the Laguerre polynomial, $\Gamma(x)$ is the gamma-function.

To illustrate our method of calculation, we give in Table I the numerical values of reduced dipole matrix elements relevant to contributions to the polarizabilities and hyperpolarizabilities for the cases considered in this paper. The data in the third row were obtained using relativistic single-double approximation [5] in which single and double excitations of Dirac-Hartree-Fock wave function are included to all-orders in perturbation theory. As seen from the Table I, our results given in second row are in good agreement with these data and measured values (fourth row).

Our results for polarizabilities and hyperpolarizabilities are collected in Table II

TABLE I. Comparison of reduced dipole matrix elements (a.u.) for Cs with other theoretical values and with experiment.

TRANSITION	PRESENT WORK	REF. [5]	EXPT. [7]
$6P_{1/2}$ - 6S	4.490	4.482	4.4890(65)
$6P_{3/2}$ - 6S	6.315	6.304	6.3238(73)
$7P_{1/2}$ - 6S	0.295	0.297	0.284(2)
$7P_{3/2}$ - 6S	0.590	0.596	0.583(10)
$8P_{1/2}$ - 6S	0.088	0.091	-
$8P_{3/2}$ - 6S	0.229	0.232	-
$6P_{1/2}$ - 7S	4.200	4.196	4.233(22)
$6P_{3/2}$ - 7S	6.449	6.425	6.479(31)
$7P_{1/2}$ - 7S	10.287	10.254	10.308(15)
$7P_{3/2}$ - 7S	14.225	14.238	14.320(20)

Table II. Stark polarizabilities and hyperpolarizabilities in Cs, a.u.

COMPONENT	$6S_{1/2}$	$6P_{3/2}$
α^S	397	1625
α^t	-	-259
γ_0	8.14×10^8	-9.75×10^8
γ_2	-	9.56×10^7
γ_4	-	2.21×10^5

An agreement with the experimental results of Tanner and Wiemam ([8], α^t = - 262.4 ± 1.5) as well as with theoretical data of Zhou and Norcross ([9], α^s = 1638.3, α^t = -261.7) and of Safronova et al ([5], α^s(6S)= 399.3) is satisfactory. Thus, for Cs atoms a Fues' model potential method is apparently quite a good approximation for the purposes of calculating high-order Stark effect in this atom. We intend to use our method in future study of hyperfine structure effects on excited Cs atom.

4. HYPERFINE STRUCTURE EFFECTS: APPLICATION TO BLACKBODY RADIATION SHIFT

The Stark effect in hyperfine structure (hfs) occurs as a result of perturbation of energy levels under the interaction between bound electron and nuclear magnetic dipole moment μ as well as the interaction with external electric field F. The Stark effect leads to a shift of sublevels with different values of $f = I + J$ (I is the nuclear spin, J is the total electron momentum), and to a splitting ΔE_{fm_f} along the projection m_f of the total momentum f in the direction of the field F. Here we give the results of calculation of ΔE_{fm_f} taking into account first time the parameter γ_{nLJM} for the Cs atoms.

The non-relativistic Hamiltonian of hfs interaction including the dipole contact interaction, the magnetic dipole orbital interaction, the magnetic spin-dipolar interaction and electric quadrupole interaction, has the form

$$h = \frac{\alpha^2 \mu}{2I}\left[\frac{8\pi}{3}(\vec{I}\cdot\vec{S})\delta(r) + \frac{(\vec{I}\cdot\vec{L})}{r^3} - \frac{\sqrt{10}}{r^3}\left(\vec{I}\cdot\left\{C^2\cdot\vec{S}\right\}\right)\right] - \left(C^2\cdot Q_2\right)/r^3.$$

Taking into account the hfs interaction to the first order, and the interaction with the electric field F to the second order of perturbation theory, the leading energy correction $\Delta^{(2)} E_{fm_f}$ can be expressed by

$$\Delta^{(2)} E_{fm_f} = F^2[< n|hGDG|n > + 2 < n|DGhGD|n > - < n|h|n > S_{-3}].$$

The corresponding equation for $\Delta^{(4)} E_{fm_f}$ [10] has the similar, but more complicated structure and has been omitted here for brevity. $|n>$ denotes the coupled wave function of the electron wave function Ψ_{JLM} and the nuclear wave function Φ_{IM_I}.

The most accurate atomic frequency standards are based on hyperfine transition frequencies in $S_{1/2}$ ground stated, such as in ^{133}Cs and ^{87}Rb employing laser- cooling techniques. As it was pointed out in Introduction, the BR Stark shift of the hyperfine splitting could limit the accuracy of this kind of the frequency standards. In this paper, we estimate the temperature-dependent shift induced by an BR electric field $\langle E^2\rangle = (831.9V/m)^2\left[T(K)/300\right]^4$ on the $\nu = 6S_{1/2}(f = 3, M_f = 0) - (f = 4, M_f = 0)$ clock transition.

The theoretical approach have been described above so we shall only outline results of calculations here. We have used the third-order perturbation expression for the scalar part of hfs Stark shift as well as fifth order contributions (hyperpolarizabilities). We have calculate the radial matrix elements using exact analytical methods including continuum spectra. Including an effect of separation in frequency between BR spectrum and the transition frequency of the D1

and the D2 line in ^{133}Cs [1] the frequency shift $\delta\nu$ is given by

$$\frac{\delta\nu}{\nu_0} = -17.2\times10^{-15}(T/300K)^4$$
$$\times[1+0.014(T/300K)^2 - 1.8\times10^{-4}(T/300K)^2].$$

The last term connected with contribution to the cesium static hyperpolarizability from the contact and the spin-dipolar interaction. This result is in good agreement with precise measurements based on the use the Cs atomic fountain frequency standards [2,3]

5. EXPERIMENTAL SETUP

To measure the Stark effect we develop an experimental set-up according to the following drawing:

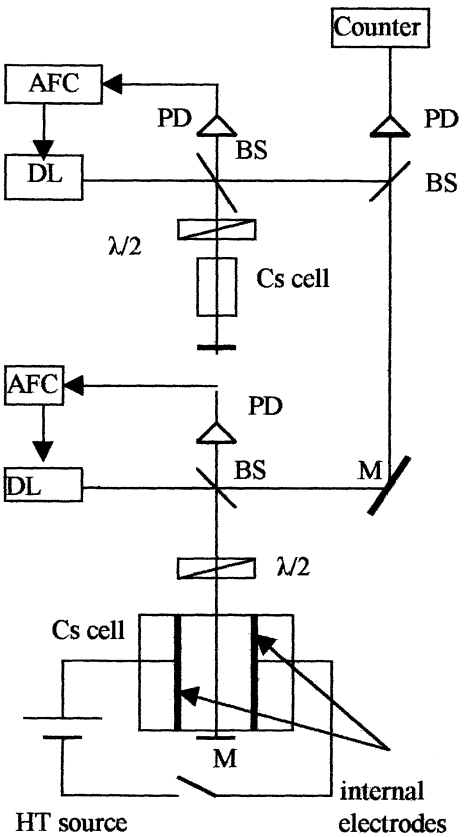


Figure 1: Experimental setup: M- mirror, PD-photodiode, BS-beamsplitter

The idea of the experiment is the following: Two diode lasers DL are locked via automatic frequency control loop AFC to the neighboring saturation absorption peacks in Cs vapors. The frequency difference is measured via counter. After high tension source HT is applied to the cell with internal electrodes we may count the frequency difference again and calculate the value of the Stark effect for a given transition. It is proposed that the sensitivity to the fractional frequency shift is of the order of 10^{-11} and the accuracy is evaluated by the value 10% and restricted mainly due to the uncertainty in determining the value of electrical field in the cell.

The preliminary results of measurements demand a separate examination and will be discussed in our forthcoming work

5. REFERENCES

[1] W. M. Itano, L. L. Lewis and D. J. Wineland, Phys. Rev. Vol. 25A, pp. 1233-1236, 1982.

[2] A.Bauch and R.Schroder, Phys. Rev. Lett. Vol. 78, pp. 622-625, 1997.

[3] E. Simon, P.Laurent and A.Clairon, Phys. Rev. Vol. 57A, pp. 436-439, 1998.

[4] A.Derevianko, W. R. Johnson, V. D. Ovsiannikov, V. G. Pal'chikov, D. R. Plante, G. Von Oppen, Phys. Rev. Vol. 60A, pp. 858-864, 1999.

[5] M. S. Safronova, W.R.Johnson and A.Derevianko, Phys. Rev. Vol. 60A, pp. 5675-5689, 1999.

[6] G. Simons, J.Chem. Phys, Vol. 55, pp.756-761, 1971

[7] S. C. Bennet, J. L. Robert and C. E. Wieman, Vol. 59A, pp. R16-R19, 1999.

[8] C. E. Tanner and C. Wieman, Phys. Rev. Vol. 38A, pp. 163-165, 1988.

[9] H. L. Zhou and D. W. Norcross, Phys. Rev. Vol. 40A, pp. 5048-55051, 1989.

[10] V.G. Pal'chikov, Hyperfine Interaction, 2000, in press

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

The Microwave Cavity in the PTB Caesium Fountain Clock

R. Schröder, U. Hübner

Physikalisch-Technische Bundesanstalt
Bundesallee 100, D-38116 Braunschweig, Germany

ABSTRACT

The microwave cavity of the PTB caesium fountain clock is designed to achieve a low transversal phase variation across the cavity opening and, at the same time, a low temperature sensitivity of its resonance frequency. This design is described and theoretically analysed. The theoretical studies confirm the validity of the design concept and will be discussed in some detail.

1. INTRODUCTION

Cylindrical microwave cavities with the field oscillating in the TE011 mode have been used in fountain clocks [1-3] hitherto. This TE011 mode has especially low losses (high intrinsic Q) which means that there is little field phase dependence on the position of the atom passage. Further reduction of the phase dependence could be achieved by symmetrically coupling at two diametrically situated positions. Then, in case of perfect coupling symmetry, the radial energy flow is zero in the center of the cavity which means that the phase dependence on the position is purely quadratic.

In the fountain clock of the LPTF (France) [1], high coupling symmetry has been achieved exciting the two equally sized small coupling areas in the cavity by two mutually decoupled cables. Additional measurements serve to equalize the amplitudes and phases of the waves running in the direction of the coupling areas. The requirements on the equality are the higher the stronger the coupling of the feeds to the cylindrical cavity is. Reasonable requirements can be met for weak couplings (e.g. loaded Q = 15000). In this case, the narrow resonance requires the stabilization of the temperature of the cavity, if the center frequency of its resonance shall be held in coincidence with the caesium frequency to get stable power in the cavity.

In PTB's fountain clock [2,4] we have chosen a different arrangement to feed the coupling slits of the cavity. Fig. 1 shows the principle. A coaxial line is tightly coupled with a waveguide which splits into two branches in an E-plane T. Each branch guides the wave via a circularly bent waveguide to its coupling slit resulting in symmetrical excitation on both sides. In contrast to the LPTF construction the excitation at both coupling sites is done not with decoupled but with coupled feeds. This guarantees that in the lossless case a pure standing wave field occurs in the cavity and in the bent waveguides, so that the dependence of the phase on the location vanishes even if the symmetry of the cavity arrangement is disturbed. In case of losses one expects that the position dependence of the phase inside the cylindrical cavity will be small for symmetry inaccuracies. Moreover, this design allows a stronger coupling between the cylindrical cavity and the feeding waveguides.

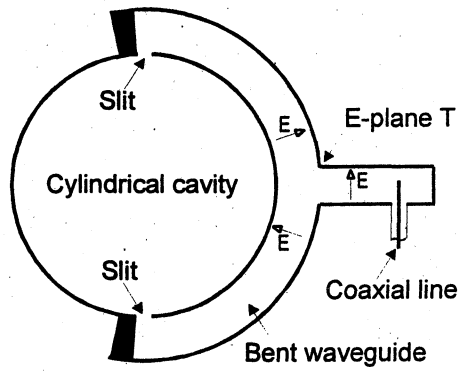


Fig. 1: Schematic view of the cavity arrangement. Excitation is achieved by two mutually coupled bent waveguides which are fed symmetrically by an E-plane T.

In our case we prepared the coupling so that the loaded Q≈2000. Therefore, the large width of the resonance guarantees that small temperature changes will not change the RF power in the cavity. Thus, any temperature stabilization can be avoided. Moreover, an adjustment of the installed cavity is not necessary. The cavity arrangement is adjusted before installation and is never changed. Additional advantages are that only one feed cable from outside has to be installed, and that separated adjustments of magnitude and phase of the coupling fields across the slits are not necessary. The disadvantage is that experiments, where one wants to

change the phase dependence on the position of the atomic transit through the cavity, are not possible.

In the following chapter we will theoretically analyze the phase dependence on the position for the arrangement described in Fig. 1 with a simplified model for the case when the symmetry (coupling and lengths of the feeding waveguides) is disturbed. The results of this chapter are used to design our final cavity arrangement.

2. A SIMPLIFIED MODEL

In our simplified model the cylindrical cavity is replaced by a rectangular waveguide cavity (Fig. 2) which is excited in the TE201 mode.

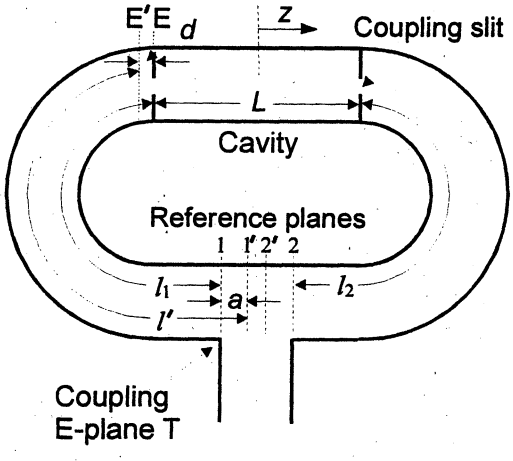


Fig. 2: Simplified model for the arrangement shown in Fig. 1. The cylindrical cavity is replaced by a rectangular waveguide cavity oscillating in TE201 mode.

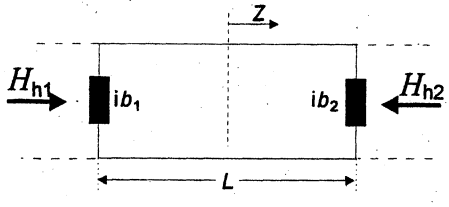


Fig. 3: Equivalent circuit for the cavity of Fig. 2. H_{h1} and H_{h2} are the waves running towards the cavity. b_1 and b_2 are the relative susceptances for the coupling.

The equivalent circuit is shown in Fig. 3. The waveguide cavity has the complex propagation coefficient $\gamma_C = \alpha_C + i\beta$, where α_C means the damping factor, $\beta = 2\pi/\lambda_g$, and λ_g is the waveguide wavelength. The damping factor is chosen to be $\alpha_C = 0.5 \cdot 10^{-5}$ mm^{-1} so that the unloaded Q of the cavity is $Q_C = 25000$ and is equal to that of the cylindrical cavity. The magnitude of the relative susceptances b_1 and b_2 is a measure for the coupling between the waveguides and the cavity.

The length L is defined so that the cavity is in resonance at the caesium frequency (9192631770 Hz) for the symmetrical case ($b_1 = b_2 = b$). Some calculations [5] give $b = (Q_{load}/\pi)^{1/2} \approx 25$ (for $Q_{load} = 2000$) and $\beta_0 L = 2\pi + \arctan(2/b)$ which determines L. β_0 is the value of β in resonance.

H_{h1} and H_{h2} are the transversal magnetic field waves running towards the cavity defined at its coupling slits. We are interested in the phase $\varphi(z)$ of the transversal magnetic field inside the cavity near z = 0. The transversal magnetic field described by $H(z)$ can be expressed as superposition of two counterpropagating waves

$$H(z) = H_h e^{-\gamma_C z} + H_r e^{+\gamma_C z}.$$

H_h and H_r are the complex amplitudes of the waves at $z = 0$. Thus, one gets the relation

$$H(z)e^{-i\delta} = (|H_r|e^{\alpha_C z} + |H_h|e^{-\alpha_C z})\cos(\beta z + \delta) + i\cdot(|H_r|e^{\alpha_C z} - |H_h|e^{-\alpha_C z})\sin(\beta z + \delta)$$

where δ is half the phase difference between the wave H_r and the wave H_h at $z = 0$. For the cases relevant to ours the calculation shows that this phase difference is $<<0.01$ so that it can be set to zero. Using the approximations $\exp(\alpha_C z) \approx 1 + \alpha_C z$ and $\exp(-\alpha_C z) \approx 1 - \alpha_C z$ and neglecting $\alpha_C z$ compared to 1 in the first term of the above equation one gets

$$\varphi(z) \approx (\frac{1}{V} + \alpha_C z)\tan(\beta z) \qquad (1)$$

where

$$\frac{1}{V} = \frac{|H_r| - |H_h|}{|H_r| + |H_h|}. \qquad (2)$$

Equation (1) shows that the influence of the asymmetry ($1/V$) and of the losses (α_C) in the cavity can be linearly superposed. V can be interpreted as a standing wave ratio. In case of asymmetry one has $|H_r| \neq |H_h|$ and therefore $1/V \neq 0$.

$1/V$ is a function of the variables b_1, b_2, $\gamma_C L$, γl_1, γl_2, and ϑ. The line lengths l_1 and l_2 are explained in Fig. 2. γ is the complex propagation coefficient $\gamma = \alpha + i\beta$ of the lines with $\alpha = 1.5 \cdot 10^{-5}$ mm^{-1} (for copper) and β the same value as mentioned when discussing the cavity. The angle ϑ is calculated from elements of the scattering matrix of the E-plane T and can be interpreted as follows. At some distance $a = \vartheta/\beta$ from the reference planes 1 and 2 (Fig. 2) towards the

symmetry axis of the E-plane T one finds the reference planes 1' and 2', where independently of the termination we have $H_{1'} = H_{2'}$.

It is our task to calculate the factor $1/V$ as a function of certain asymmetries. Fig. 4 shows $1/V$ for different line lengths ($l_1 \neq l_2$) with $\Delta l = (l_1-l_2)/2 = 2$ mm but equal couplings ($b_1 = b_2 = 25$) as a function of the average line length $l = (l_1+l_2)/2$. The shape of the curve shows remarkable features. First, $1/V$ shows resonant increases (discontinuous regions) at $l = 31.190$ mm and 54.456 mm. Second, $1/V$ vanishes at $l = 42.823$ mm. Both features can be interpreted physically. Assuming symmetry ($H_{h1} = H_{h2}$, $b_1 = b_2$) and $\alpha_C = 0$, one finds two reference planes E' (Fig. 2) at small distance ($d \approx 0.3$ mm) from the coupling slits where the reflection coefficient is equal to +1 in case of resonance and −1 far away from resonance. Now, the length l' between the planes E' and the earlier introduced reference planes 1' and 2' of the E-plane T (Fig. 2) has to be chosen to $n\cdot\lambda_g/2$ (in our case n = 2) to get the length $l =$ 42.823 mm where $1/V$ vanishes (second feature),

$$l = n\cdot\lambda_g/2 + d - a. \qquad (3)$$

If one increases or decreases the length l, calculated with Eq. (3), by $\lambda_g/4$, one arrives at values of l at which $1/V$ resonantly increases so that the phase shows the unfavorable discontinuous shape (first feature).

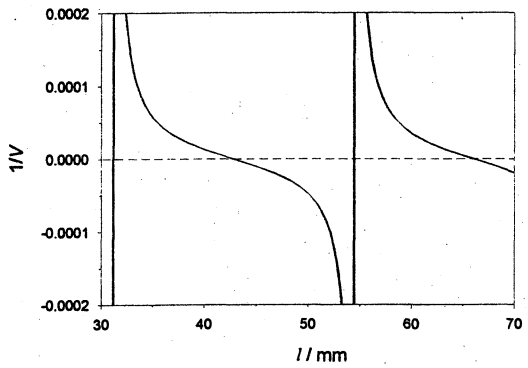


Fig. 4: $1/V$ as a function of the average waveguide length l for the E-plane T. Remarkable features are the zero crossing at $l = 42.823$ mm and the resonant increase at $l = 31.190$ mm and 54.456 mm. The parameters are $b_1 = b_2 = 25$, and $\Delta l = (l_1-l_2)/2 = 2$ mm.

Fig. 5 shows $1/V$ as a function of the average line length l for the case $b_1 \neq b_2$ but $l_1 = l_2$. We have taken $b_1 = b + \Delta b = 25 + 2.5$ and $b_2 = b - \Delta b = 25 - 2.5$. The shape of the curve in Fig. 5 shows a mostly constant course which changes resonantly at the now well-known lengths $l = 31.190$ mm and 54.456 mm.

More extended calculations [5] of $1/V$ lead to the following conclusions which hold in the complete variation region of l, excluded the above mentioned crucial regions:

- The two cases considered, $\Delta l \neq 0$ and $\Delta b \neq 0$, can be superposed linearly,
- in both waveguides the damping (α) practically does not play any role, and
- the shape of the curve for $1/V$ is not changed significantly by changing the frequency up to a few multiples of the resonance bandwidth.

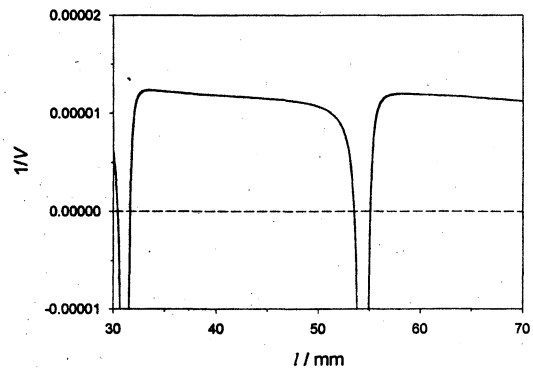


Fig. 5: $1/V$ as a function of the average waveguide length l for the parameters $l_1 = l_2 = 42.515$ mm and $\Delta b = (b_1-b_2)/2 = 2.5$. Significant deviation from zero can be seen at the length $l = 31.190$ mm and 54.456 mm.

3. THE REALIZED CAVITY ARRANGEMENT

We were forced by the actual dimensions of our fountain clock to use the more compact solution of Fig. 6. The coaxial line is coupled directly into the arms. The location of the coaxial line probe (length 7 mm, diameter 1 mm) is shifted by $\lambda_g/4$ out of the symmetry center so that the difference of the arm lengths is $l_1 - l_2 = \lambda_g/2$. This guarantees the desired inphase excitation ($H_{h1} = H_{h2}$) of the cavity coupling slits. The termination walls are placed not directly behind the coupling slits, instead shifted by approximately $\lambda_g/2$ behind the slits. This does not make a fundamental difference.

For this arrangement we have calculated $1/V$ and δ as functions of l where we have modelled the cylindrical cavity by the TE201 waveguide cavity as before. We have chosen the same asymmetries Δl and Δb as used in the E-plane T arrangement. Because of the difference of $\lambda_g/2$ between the lengths of the arms we use $l_1 = l + \Delta l$ and $l_2 = l - \Delta l - \lambda_g/2$.

We found that the shapes of the curves for $1/V$ are similar to those of the E-plane T (Figs. 4,5) not only qualitatively but also quantitatively. That means, the difference of the arms by $\lambda_g/2$ has practically no

influence on the results. Therefore, all the conclusions of chapter II can be used here. Especially, two planes 1' and 2' can be defined where $H_{1'} = -H_{2'}$. These planes are positioned symmetrically to the probe at a distance of $a = \lambda_g/4$ in the bent waveguide. The distances l' (arm 1) and l'-$\lambda_g/2$ (arm 2) are the distances between the reference planes E' near the cavity and the planes 1' and 2', respectively. Choosing $l' = n\lambda_g/2$ (n = 1,2,...) corresponds to the length l where $1/V$ crosses zero. Contrary to this, $l' = n\lambda_g/2 + \lambda_g/4$ corresponds to the length for the resonant increase of $1/V$.

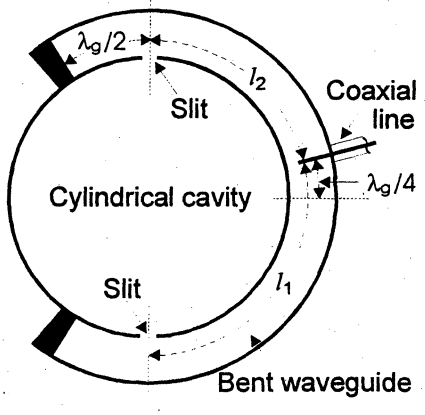


Fig. 6: Schematic view of the realized cavity arrangement. One coaxial cable feeds the bent waveguide. The feed is shifted by $\lambda_g/4$ so that both waveguide arms differ by $\lambda_g/2$. The shorts at the end of the waveguides are shifted by $\lambda_g/2$ to ease the mechanical construction.

These results were used to determine the dimensions of the waveguide. Measurements showed that the planes E' lie d = 1.5 mm before the center of the coupling slits. Therefore, the whole length between both slits is $l' + (l' - \lambda_g/2) + 2a + 2d = 2*\lambda_g/2 + \lambda_g/2 + 2*\lambda_g/4 + 2*1.5$ mm $= 2\lambda_g + 3$ mm. The value of λ_g is determined by the width (here 22.8 mm) and to a certain small amount by the curvature of the bent waveguides. One calculates λ_g = 46.72 mm [6]. This value refers to the middle of the waveguide height (10 mm). Starting with the dimensions of the bent waveguide arms and taking the wall thickness (1.5 mm) for the cylindrical cavity into account, one gets the inner diameter (48.4 mm) of the cavity. From this one calculates a height of the cavity of 28.62 mm if the cavity is unloaded. In the loaded case (Q_{load} = 2000, slit length 10,8 mm, width 2 mm) the length has to be smaller (28.12 mm). The experimental determination of the length of the cavity (the tuning to the Caesium frequency) has to be done with the finally used coaxial line including the isolator installed outside the vacuum enclosure. The unloaded Q of the cavity is 28450 and, therefore, smaller as compared to the maximum possible Q of 30924 which could be achieved for smaller diameters of the cavity. One should mention that choke slits (width 0.6 mm, depth 5 mm) and the holes for the atomic passage (diameter 10 mm) in the caps shift the TM111 resonance by 390 MHz to lower frequencies [2].

4. DISCUSSION

On the basis of our simplified model, we will discuss the requirements on the symmetry of the arrangement. One requirement can be that the phase minimum should not be shifted by more than 1 mm relative to z = 0. From this a limitation for $1/V$ follows. For small z we have $\tan(\beta z) \approx \beta z$ so that $\varphi(z) \approx (1/V + \alpha_C z)\,\beta z$. This equation represents a quadratic parabola shifted by $z_0 = 1/(2\alpha_C V)$ relative to $z = 0$. With $\alpha_C = 0.5 \cdot 10^{-5}$ mm^{-1} and the requirement $z_0 < 1$ mm we get $1/V < 10^{-5}$. Fig. 5 shows that this is almost fulfilled even for large coupling asymmetries as it is for $(b_1\text{-}b_2)/b = 0.2$. Length asymmetries can be neglected if the average line length l is chosen corresponding to $l' = n\lambda_g/2$.

Let us dicuss now what follows if the same demand $z_0 < 1$ mm (resp. $1/V < 10^{-5}$) is posed on the arrangement with decoupled feeds. The discussion is restricted to the case that the coupling into the cavity is symmetric ($b_1 = b_2$) and that the waves H_{h1} and H_{h2} (Fig. 3) at the coupling slit are in-phase (as in the ideal case) but deviate in the amplitudes. A calculation [5] gives

$$\frac{1}{V} \approx \frac{|H_{h2}| - |H_{h1}|}{|H_{h2}|} \cdot \frac{\pi}{2Q_{load}} .$$

For Q_{load}=2000 and $|1/V|<10^{-5}$ this means $\left|1 - |H_{h1}|/|H_{h2}|\right| < 0.013$ which is not easy to fulfill. However, with Q_{load} = 15750 the amplitudes have to agree only to about 10% for the same demand $z_0 < 1$ mm, but the narrow resonance would require temperature stabilization which in our case can be avoided.

ACKNOWLEDGEMENT

The authors wish to thank D. Griebsch for his engagement in the mechanical design and preparation of the cavity.

REFERENCES

[1] G. Santarelli, "Contribution à la Réalisation d'une Fontaine Atomique", PhD Thesis in French, Université Pierre et Marie Curie, Paris VI, pp. 29-31, 1996.

[2] U. Hübner, S. Weyers, J. Castellanos, D. Griebsch, R. Schröder, C. Tamm, A. Bauch, "Progress of the PTB caesium fountain frequency standard", in: Proceedings of the 12th European Frequency and Time Forum (EFTF), Warsaw, Poland, March 10-12, pp. 544-547, 1998.

[3] S. R. Jefferts, R. E. Drullinger, A. DeMarchi, "NIST Cesium Fountain Microwave Cavities", IEEE Frequency Control Symposium, pp. 6-8, 1998.

[4] S. Weyers, U. Hübner, B. Fischer, R.Schröder, Chr. Tamm, A. Bauch, "Investigation of Contributions to the Stability and Uncertainty of the PTB's Atomic Caesium Fountain", to be published in the Proceedings of the European Frequency and Time Forum (EFTF), 2000.

[5] R. Schröder, U. Hübner, to be published.

[6] N. Marcuvitz, "Waveguide Handbook", McGraw Hill, New York, pp. 333-334, 1951.

RECENT RESEARCH ACTIVITIES OF HYDROGEN MASERS AT SHANGHAI ASTRONOMICAL OBSERVATORY

Qinghua Wang, Zaocheng Zhai, Chuanfu Lin

Shanghai Astronomical Observatory, Chinese Academy of Sciences
80 Nan Dan Road, Shanghai, 200030, P. R. China
Tel: (86 21) 64388398 Fax: (86 21) 64384618
Email: qhwang.shao@mailcity.com

ABSTRACT

Since first hydrogen masers for laboratory use in China were constructed at Shanghai Astronomical Observatory in the 70's, rugged and easily transportable engineering hydrogen masers have been developed to be used in an increasing number of demanding applications. In recent years, we delivered about thirty hydrogen masers. Hydrogen masers with the cavity auto tuning system, as well as the intelligent monitor and control system were developed. Typical performance parameters of the engineering hydrogen maser are shown in this paper.

For an attempt to design small, lightweight and inexpensive hydrogen masers, we have studied the subcompact electrode-loaded cavity and the TE111 septum cavity. The maser equipped with the electrode-loaded cavity produced sustained oscillation, with the Q electronically enhanced from 3000 to 170000. Its preliminary test results are described. The small-sized TE111 septum cavity without a separated storage bulb is predicted to be able to meet the maser oscillation condition. We observed inductive radiation with the TE111 mode maser, but could not achieve self-oscillation. The system supports oscillation when the cavity is placed in an external feedback loop.

1. DEVELOPMENT OF ENGIGEERING HYDROGEN MASER

The high-stability hydrogen maser based upon quantum transitions in hydrogen atoms plays roles in science, technology and commerce. It is the most stable commercial frequency standard currently available for averaging intervals from 1 second to 10^4 seconds.

Shanghai Astronomical Observatory began to carry out the research on hydrogen masers in 1970. The first traditional hydrogen maser for laboratory use in China was built two years later.

In 1987, a new engineering active hydrogen maser, illustrated in Figure 1, was developed successfully. The entire maser is self-contained in a single standard-sized rack cabinet. It is a relatively lightweight, rugged, and easily transportable standards.

Our hydrogen masers are used in an increasing number of very demanding applications, including very long base line interferometers (VLBI) for radio astronomy, laboratory reference standards for metrology, frequency standards for satellite position system, navigation, radio communication and time keeping. In twelve years, we manufactured and delivered about thirty hydrogen masers to domestic military and civil customers, and one of them was exported to Australia. Meantime, improvements of hydrogen masers have continued for special applications:

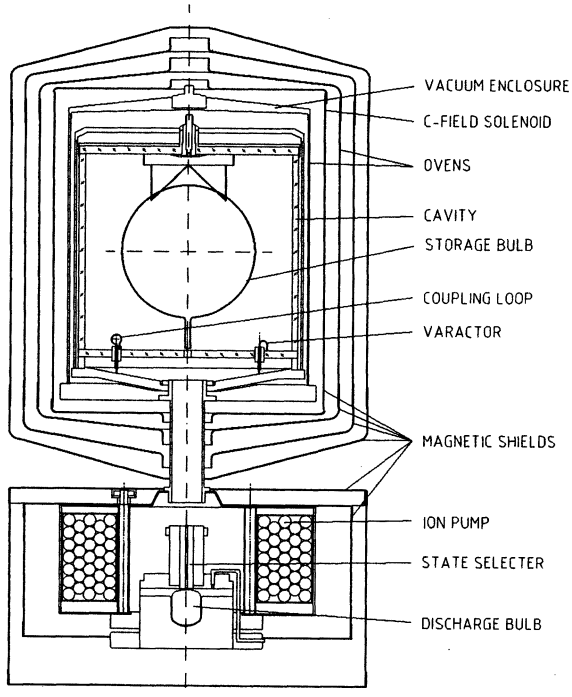


Fig. 1 Schematic diagram of engineering hydrogen maser

For time keeping, the cavity auto tuning system is applied in the hydrogen maser . It greatly improves the long-term stability of the maser, with little perturbation of their short-term stability, due to the introduction of the injected signal of the cavity auto-tuner.

To meet the need of convenient operation and remote control, hydrogen masers with intelligent monitor and control system were developed in 1998. The system uses single-chip-microcomputer as the core, with the RAM, I/O expanded circuit and outward circuit including the A/D module, interface module, serial

Table 1 Typical performance parameters of the engineering hydrogen maser

Electrical Specifications	Stability	10ms	100ms	1s	10s	100s	1000s	10000s
		3E-11	5E-12	3E-13	3~4E-14	7~9E-15	4~6E-15	1E-15
	Accuracy	5E-13						
	Drift	1~2E-15/Day						
	Retrace	2E-14						
	Synthesized Frequency Resolution	7E-14						
	Available Frequency Outputs	5MHz, Sine, 50 ohms > 1V rms						
		1 PPS, positive polarity TTL, Pulse Width: 2ms, Rise Time: <20 ns						
	5 MHz SSB Phase Noise	Offset from Carrier (Hz)		1	10	100	1000	10000
		Phase Noise (dBc)		-100	-120	-140	-150	-150
	Power	Operating voltage		Frequency		Power Consumption		
		220 VAC		50 Hz		150W		
Environmental Specifications	Temperature Sensitivity (22~31℃)				Slow coefficient		Fast coefficient	
				ΔT up	2.9E-14/℃		1.2E-14/℃	
				ΔT down	3.0E-14/℃		1.3E-14/℃	
	Magnetic Field Sensitivity (0.4Gauss)	1.1E-13 / Gauss						
	Barometric pressure Sensitivity (25mmhg)	1E-14						
Physical Specification	Size	54×74×111cm						
	Weight	140 kg						

communication module, digital clock module. It serves the functions of automatic start-up and shutdown, multiplexed external monitoring, operating status diagnosis, alarming, selection of output frequencies, standard time generation, as well as communication with the remote terminals.

Typical performance parameters of the engineering hydrogen maser [1] are shown in the table 1.

2. EXPERIMENTAL STUDIES FOR SMALL HYDROGEN MASERS

However, in some degrees, conventional hydrogen masers are limited in their usage where there are severe size and weight constraints (especially as the spaceborne frequency standards). Designing small, lightweight and inexpensive hydrogen masers is in demand.

The cavity dimensions determine the minimum size and weight of a hydrogen maser. A typical cylindrical TE011-mode cavity utilized in a conventional maser is about 28 cm both in diameter and length. In order to reduce the size of the cavity, different types of the compact cavity were developed.

Since the end of the 80's, we have studied two kinds of compact hydrogen masers with the electrode-loaded cavity and the TE111 septum cavity respectively.

2.1 Electrode-loaded cavity [2]

This design employs metal electrodes placed inside a cylindrical cavity as loading structures. The dimensions of the inductive electrodes and the capacitive gaps between them are selected to give the desired cavity resonant frequency of 1420.405MHz.

The schematic diagram of the cavity is shown in Figure 2. Four copper electrodes, with the length of 12cm and the gap of 7.8cm, are attached to the outer surface of the quartz storage bulb by low-loss epoxy. The bulb, 10cm diameter by 10cm long, is coated internally with Teflon, and is supported by a 1cm high quartz glass tube attached to the cavity cylinder bottom. The cavity cylinder of 15cm diameter × 0.5cm wall × 14cm long is made of aluminum, the loaded Q of which is measured about 3000.

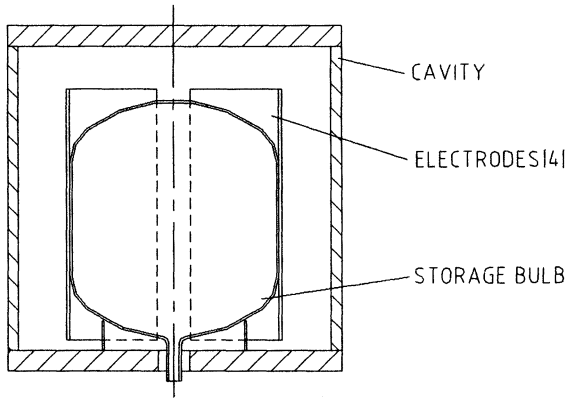


Fig. 2 Schematic diagram of electrode-loaded cavity

In order to make the system break into oscillation , we fed back a small portion of the amplified maser output into the cavity to enhance the

cavity Q.

Table 2 Physical characteristics of two metal cavities

Material	Weight (kg)	Conductivity σ_1(s/m)	Skin depth δ(μm)	Theoretical unloaded Q	Cavity frequency temperature sensitivity (kHz/°C)
Cu	17	5.5×10^7	1.8	3.1×10^4	-23.6
Al	5	3.2×10^7	2.3	2.3×10^4	-32.7

The phase shifter and the attenuator are adjusted for propagation delays and the amount of the feedback, respectively. When the feedback signal coheres with the electromagnetic field in the cavity in phase, the oscillation condition is met (The Q is enhanced to 170000).

Nevertheless, the feedback loop is actually a part of the resonant system, which is susceptible to phase shift in the feedback loop as well as environmental fluctuation. It is therefore necessary to have a cavity auto-tuning system. Additionally, the feedback loop and the front of the maser electronics are placed into an isothermal box, which is contributed to reduce the thermal fluctuation and to improve the long-term performance.

The preliminary test results are shown in Figure 3. The apparent degradation in the short-term stability could be attributed to enhanced thermal noise in the cavity.

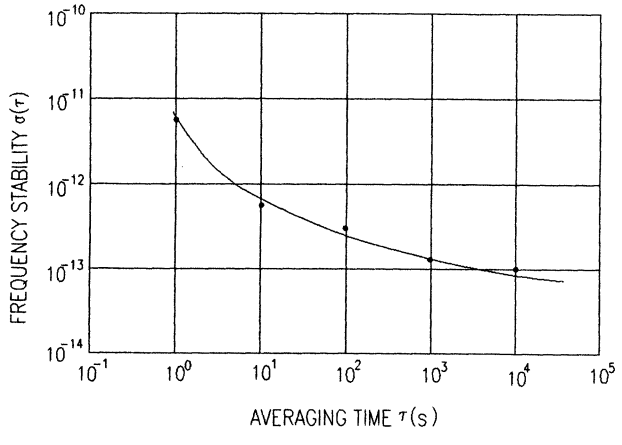


Fig. 3 Preliminary frequency stability data

2.2 TE111 septum cavity [3]

For an attempt to develop the active hydrogen maser with a small-sized cavity without the separate quartz storage bulb, E. Mattison, M. Levine, R. Vessot developed a cylindrical TE111-mode septum cavity [4].

Figure 4 shows the distribution of rf magnetic field in the TE111 septum cavity. A Teflon sheet in the longitudinal axis of symmetry divided the cavity into two areas of oppositely directed oscillating magnetic field. Hydrogen beam in the higher hyperfine energy level is separated into two parts when it enters the cavity, making radiative transitions in each region.

The resonant wavelength corresponding to the TE111-mode cavity is as follow [5]:

$$\lambda_{TE111}=\frac{1}{\sqrt{\frac{1}{4l^2}+\left(\frac{1}{3.41R}\right)^2}} \tag{1}$$

Then,

$$(Df)^2=\left(\frac{2c}{3.41}\right)^2+\left(\frac{c}{2}\right)^2\left(\frac{D}{l}\right)^2 \tag{2}$$

To avoid existing different resonant modes in one cavity, D/l should be far less than 0.95 (when TE111 is the prime mode in the cylindrical cavity), But it should not be too little to ensure the high cavity Q. We chose $\frac{D}{l}=0.8$. For $f=1420.406MHz$, $D=15cm$, $l=18.75cm$. The dimensions of the cavity approximate to the half the size of the TE011-mode cavity.

The maser oscillation condition for a septum cavity can be represented by the inequality:

$$\frac{S_0}{1+\beta}\geq 5.9\times 10^3 \tag{3}$$

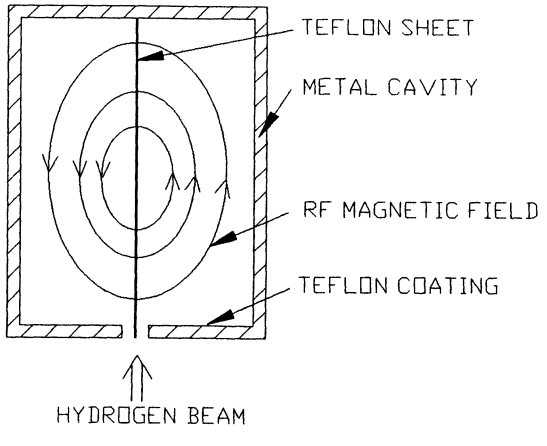


Fig.4 Distribution of magnetic field for TE111 cavity

Where the parameter S_0 is the product of Q_0, the unloaded cavity Q, and η', the storage-bulb filling factor. β is the external coupling factor of the cavity.

According to the S_0 curve for TE111-mode cavity, S_0 for the cavity with such size is above 13×10^3, When

$\beta = 0.15$, $\frac{S_0}{1+\beta} = 11 \times 10^3$, twice the threshold valve. So this type of the resonant cavity should oscillate by the theoretical evaluation.

From equation (1), we derive the coefficients of frequency sensitive to the length, radius and temperature of the cavity. When the length or the radius adds to 1mm, the cavity frequency lowers 2.4MHz and 12.9 MHz respectively.

The whole cavity consists of four parts. The Teflon sheet is clamped tightly with screws of the left and right halves to form the two atoms-confining regions. While the two parts perpendicular to the axis are pressed with six bolts. The small quartz domes cover coupling loops in the cavity, and the outside surface of the domes as well as the other interior surface of the cavity is coated with Teflon. Thus, no materials but Teflon is impacted with hydrogen atoms, which is helpful to decrease the loses of hydrogen atoms. The physical characteristics of two metal cavities in our experiments are shown in Table 2.

The resonance frequency is roughly set by machining the cavity to the proper size, then regulated during operation by adjusting the maser's thermal control circuits to vary the cavity temperature according to the calculated sensitivities of cavity frequency.

In the first-step experiment, the cavity is made of Oxygen-Free High Conductive copper to ensure relatively high Q. Its cavity Q_l is measured about 25000.

When the pulse signal of 1420.406MHz is injected into the cavity tuned to the resonant frequency, atoms with F=1, m_F=0 are at the state of stimulated radiation. Adjust various of parameters such as discharge source, flow of atoms, and C field until the signal is the strongest. But after the pulse is removed, the inductive radiation signal is attenuated and the system could not remain oscillating.

We think, when curing the Teflon coating of the cavity under the high temperature of 320℃, the OFHC copper could not avoid being oxidized, although we had taken measures such as separating the cavity from the oxygen and filling the nitrogen into the oven. The oxidized copper decreases the high conductivity of the metal, and the Teflon-coated surfaces seem not be satisfactory. In addition, interfaces of dividing planes probably lead to RF leakage. These factors cause that the cavity Q_l drops to about 20000. The cavity loses are such that the maser oscillation condition is difficult to meet in our experiment. Therefore, further work has to be done on construction, material and coating method of the cavity.

To apply the Q-enhanced technique (the Q itself needs not be very high), we meantime fabricate a lightweight aluminum cavity, which is easy to coat the Teflon. The Q is risen from 11800 to 142000. The maser achieves self-oscillation after the pulse is removed.

2.3 Other activity

We have also made an approach to the passive small hydrogen maser, where the atomic resonance is used as a narrow bandpass amplifier.

Now we are going a step further to develop the small maser to a commercial frequency standard. It would be used widely in the future of China.

3. REFERENCES

[1] Technical Manual of Engineering Hydrogen Maser , Shanghai Astronomical Observatory, May 1998.

[2] Z. C. Zhai, C. F. Lin, et al, The Development and Experimental Results on the Minature Hydrogen Maser for Time-keeping, Acta Metrologica Sinica, Vol.15, No.4, pp276-280, 1994.

[3] Q. H. Wang, Z. C. Zhai, W. Q. Zhang and C. F. Lin, An Experimental Study for the Compact Hydrogen Maser with a TE111 Septum Cavity, IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, Vol.47, No.1, pp197-201, Jan. 2000.

[4] E. M. Mattison, M. W. Levine, R. F. C. Vessot, New TE111-mode hydrogen maser. in Proceedings of the 8th PTTI, pp355-367, 1976.

[5] R. L. Gu, B. H. Li, et al, Microwave Techniques and Antenna, National Defense Industry Press, China, pp158-161, 1980.

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

ISOTROPIC COOLING OF CESIUM ATOMS FOR A COMPACT ATOMIC CLOCK: FIRST RESULTS AND EXPECTED PERFORMANCES

C. Valentin, Ch. Guillemot, E. Guillot, P.-E. Pottie, E. Fretel, P. Petit, and N. Dimarcq
Laboratoire de l'Horloge Atomique, Unité propre du CNRS,
F-91 405 Orsay,
E-mail: constance.valentin@ief.u-psud.fr.
E-mail: pottie@ief.u-psud.fr

***Abstract*— We aim to build a compact atomic clock by means of cold atoms. We briefly remind the set-up used for this purpose, and its two main features: the isotropic cooling and the temporal scheme of the clock. The largest number of cold atoms and the lowest temperature we obtained from our cold atoms source are reported. These first measurements are used to predict the frequency accuracy and the frequency stability, taking into account the temporal scheme of the clock. The means to improve the frequency performances are also pointed out. Two effects have been evaluated particularly, the first-order Doppler shift and the so-called Dick effect. The first-order Doppler shift have been evaluated to a few 10^{-15} and the relative frequency stability to a few $10^{-13}\tau^{-1/2}$.**

I. Introduction

HORACE (HOrloge A Refroidissement d'Atomes en CEllule) is an innovative project where we use cold atoms in order to make a very compact clock. Instead of using the low velocity of cold atoms to perform long interrogation times, we take advantage of it to reduce the size of the microwave cavity [1], [2]. Moreover, we group all the functions of a regular atomic resonator within the microwave cavity itself: cooling, atomic preparation, microwave interrogation and detection of the clock signal take place in the microwave cavity. The size of the resonator is as small as 40 mm. The atoms do not travel the whole cavity anymore, and every function applied to them follows a temporal scheme. Hence, it is easy to change the interrogation scheme. We can thus play with the interrogation time, the dead time between the two Ramsey-pulses, the duty cycle and the cycle-time that is impossible to do in a regular atomic clock. With this original context of a temporal scheme, we are presenting here the numerical calculation of the first-order Doppler shift versus the interrogation time and the estimation of the Dick effect versus both the interrogation time and the duty cycle.

On the other side, we are showing here the features of our cold atoms source. In previous papers, we have already demonstrated that we must use isotropic cooling for HORACE [1], [2]. In the first section of this paper, we will report on the temperatures and the numbers of cold atoms we succeeded to measure with this technique in our configuration. The following section are dedicated to the expected frequency performances of HORACE.

II. Isotropic cooling: Experimental results

Isotropic cooling has already been demonstrated to be efficient with diffusive and reflective cells [1], [3]. We succeeded to cool down cesium atoms below 50 μK in these two cases. In order to determine the best configuration for HORACE, we have begun a complete experimental study of isotropic cooling, varying the cell geometry, the number, the position and the type of the fibers. We restrict our discuss to the present set up in Horace: the polished up copper cylinder has an axis of revolution symmetry perpendicular to the axis of gravity. We use 4 multimode fibers, set up in the median plane of the cylinder. We let the atoms fall in the gravity field, and we detect them nearly 250 millimeters below with a probe laser beam by a single Time Of Flight (TOF) technique. We cannot then distinguish between the widening of the TOF distribution due to the spatial extension of the cold atoms cloud, and the widening due to the temperature of the atoms. So every temperature we indicate here is given assuming that the atomic cloud is a point, and is then strongly overestimate. A probe beam perpendicular to the median plane of the cylinder makes it possible to measure the density of thermal atoms in the middle of the cell and thus the loading time (see figure 1).

We observed single peak TOF. The maximum number of cold atoms measured in the TOF signal is 3.7×10^7 when the detuning is $\delta = -1.6\Gamma$ (δ is the detuning of the laser frequency compared to the atomic transition frequency, and Γ is the natural width of the atomic transition). For this detuning, the temperature we measure is less than 15 μK. The temperature decreases with increasing detuning. The lowest temperature we measure is below 7 μK and is reached when the detuning is $\delta = -2.0\Gamma$.

We measured also the loading time of the cold atoms (see figure 2). This curve is an exponential function with a time constant of roughly 60 ms. After 200 ms of cooling time, we do not cool any more atoms. With 20 ms cooling time, 30 percent of the maximum cold atoms are cooled down. These rather short durations are promising because they could allow us to reduce the cooling time after the first sequence and hence improve the duty cy-

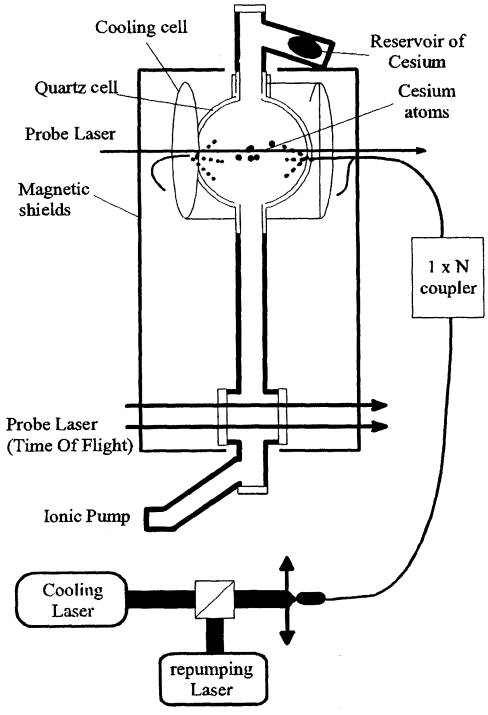


Fig. 1. *Sketch of the experiment.*

cle of the sequence. Moreover, one can assume that the atoms are not significantly heated during short laps of time in the order of 10 ms. Thus, for short interrogation time, the cooling time may be as short as a few 10 ms, and the duty cycle near 0.5.

III. Expected frequency performances

In order to estimate the frequency accuracy and the frequency stability of such a frequency standard as HORACE, we have performed numerical simulations based on the calculation of the transition probability of an atom in interaction with a microwave magnetic field [4]. We have evaluated two effects. The first one is the first-order Doppler shift, that will be compared to the collisional shift. The second one is the so-called Dick effect [6] which can dramatically limit the short-term frequency stability of a pulsed clock.

A. *First-order Doppler shift*

Every perturbation of the microwave field gives residual progressive waves and so a phase gradient in the three directions of space. The atoms travelling in this phase gradient see their frequency shifted by Doppler effect. We have numerically calculated the first-order behavior of this effect. As we are using cold atoms, the second order Doppler shift is negligible.

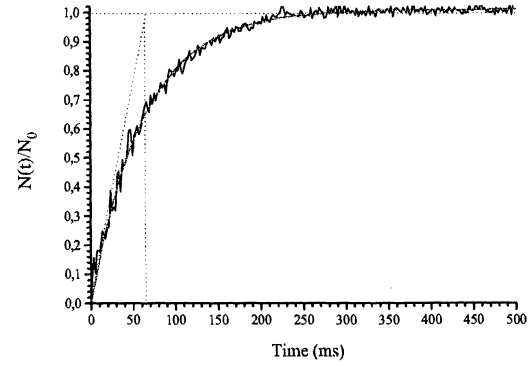


Fig. 2. *Number of cold atoms versus the cooling time in the cylindrical reflective cell. The maximum number of cold atoms is reached within* 200 *ms. The load time is roughly* 50 *ms.* 30 *percent of the maximum number of cold atoms are cooled with a* 20 *ms cooling time.*

The amplitude and the phase of the magnetic field have been calculated at IRCOM (Limoges, France) by means of a finite elements method. Along the axis of revolution symmetry of the cylinder, which is parallel to the gravity axis, the phase gradient is evaluated about $-1.12 \times 10^{-2}\ rad.m^{-1}$ from the center to the end of the microwave cavity, i.e. a total length of 21.5 millimeters. The numerical calculation was a Monte-Carlo simulations run with 10,000 cold atoms within a distribution temperature of 50 μK, and which are falling down along the gravity axis. We have assumed here that the atomic cloud was a 5 mm sphere with a gaussian distribution.

HORACE is operated with a temporal sequence so that the interrogation scheme may be either a Rabi-like interrogation or a Ramsey-like interrogation, since we can vary the interaction durations τ_1 and τ_2 and the dead time T. We calculate the first-order Doppler shift in these two schemes (Figure 3).

These curves exhibit a relative frequency shift of less than 9.0×10^{-15} when the total interrogation time is 65 ms long. We did not run the calculation for longer durations, because an atom which would start from the center of the microwave cavity in HORACE with a null velocity would begin to leave the cavity and the signal to noise ratio would dramatically decrease. On the contrary we can take advantage of the temporal scheme of HORACE operating with short cycle times and short interrogation times but with a high frequency rate. We could thus reach first-order Doppler shifts smaller than 1.0×10^{-15}.

The first-order Doppler shift may be compared with the frequency shift due to collisions between cold atoms. We currently measure about 3.7×10^7 cold atoms in

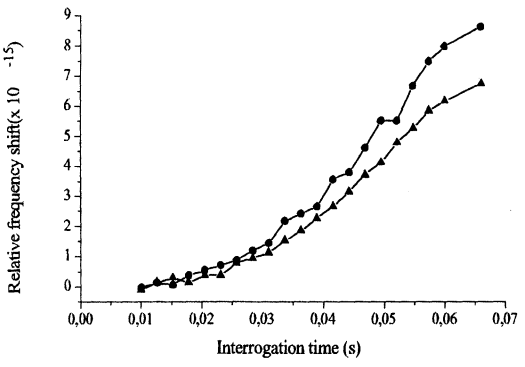


Fig. 3. *Relative frequency shift due to the first-order Doppler effect. The triangles are for the Rabi-like interrogation and the circles are for the Ramsey-like interrogation. For the Ramsey-like interrogation, the ratio between the sum of the two interrogation durations τ_1 and τ_2 and the dead-time T is constant and is roughly* 0.2.

the Time Of Flight signal. Varying the number of cold atoms in the cell, we have calculated a collisional shift between -5×10^{-14} and -5×10^{-15} [7], [8]. Thus, collisions between cold atoms and the first-order Doppler shift should limit the inaccuracy of HORACE in the few 10^{-15} range [7].

B. *Dick effect : Frequency stability limit*

As the atomic signal is sampled at a rate which is the inverse of the cycle time T_c, it gives spectrum aliasing in the Fourier-space. The noise at Fourier components different from zero and multiple of the sampling frequency $f_c = 1/T_c$ contributes to the frequency stability degradation of the atomic frequency standard [5]. The Dick effect decreases first when the the interrogation time increases, and second when the duty cycle increases and tends towards 1. As for the first-order Doppler shift, and thanks to our large number of cold atoms, we could take advantage of a shorter cooling time in order to increase the cycle frequency. However, we might not operate with too high cycle frequency because the frequency noise spectral density starts to increase at a few Hertz [9]. There is thus an optimal configuration to minimize the product of the frequency noise spectral density of the oscillator and the Fourier components of the sensibility function.

We have estimated the limit of the short-term frequency stability due to the Dick effect, using the link between the power spectral density of the relative frequency fluctuations of the locked local oscillator and both the Fourier components of the sensibility function g(t) and the frequency noise spectral density of the free-running local oscillator at Fourier frequencies kf_c [9]:

$$S_y^{LLO}(f) = 2 \sum_{k=1}^{\infty} |\frac{g_k}{g_0}|^2 \, S_y^{LO}(kf_c) \qquad (1)$$

We have calculated the sensibility function g(t) for infinitesimally small phase steps of the oscillator signal at the time t. We assume that the atoms undergo a Ramsey interrogation scheme with two interaction times τ_1 and τ_2, and a dead-time T. It is important to notice that the atoms starting from the center of the cavity are falling down into the lower part of the microwave cavity, and that the power of the second pulse is weaker than the first one for the same input microwave power. So the second interaction duration must be longer than the first one in order to provide two $\pi/2$ pulses. The sensibility function exhibits then the regular shape [9]. Moreover, we can smooth the Fourier components of the sensibility function using other shapes of pulses (e.g. gaussian shapes). When the shape is smoother, the sensibility function series vanishes more quickly and thus the Dick effect is reduced.

The Dick effect obviously depends on the spectral purity of the U.S.O. (Ultra Stable Oscillator) in use. We compute the stability for two kinds of U.S.O.: a regular local oscillator on one hand, and an ultra low-noise oscillator from Oscilloquartz on the other hand. Their spectral densities of relative frequency fluctuations are:

$$S_y^{LO}(f) = 1.0\,10^{-23}/f + 1.0\,10^{-25} f + 1.0\,10^{-31} f^2 \quad (2)$$

for the first one, and:

$$S_y^{LO}(f) = 7.2\,10^{-27}/f + 8.0\,10^{-27} f + 8.0\,10^{-29} f^2 \quad (3)$$

for the last one. For this last one, the limiting noise is the flicker phase noise of the U.S.O. [9].

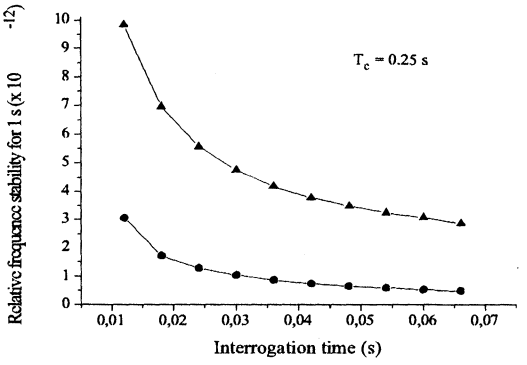


Fig. 4. *Frequency stability for $\tau = 1s$ versus the interrogation time computed for two oscillators. The circle are for the ultra low-noise oscillator from Oscilloquartz and the triangles are for the regular oscillator we currently use. The Dick effect decreases when the interrogation time increases.*

In order to overestimate the relative frequency stability, we run the calculation with a cycle time of 250 ms, according to the whole loading time of the molasses that is about 200 ms long [7]. Then the relative frequency stability limit for 1 s is about 5×10^{-12} with the first U.S.O., and is about 4.8×10^{-13} with the best U.S.O., for $\tau_1 = 3.5$ ms, $\tau_2 = 7.5$ ms and T = 55 ms.

However, the duty cycle may be enhanced (cf figure 5). If we decrease the cooling time to 100 ms (corresponding to a load of 80 percent of the maximum cold atoms) [7], and taking into account the probe time and the dead-times, we increase the duty cycle to 0.35, and the relative frequency stability for one second may be 4×10^{-13}. With a cooling duration of 50 ms (corresponding to a load of 50 percent of the maximum cold atoms), the duty cycle is then 0.5, and the relative frequency stability for one second may be 3×10^{-13}, with a 65 ms interrogation time.

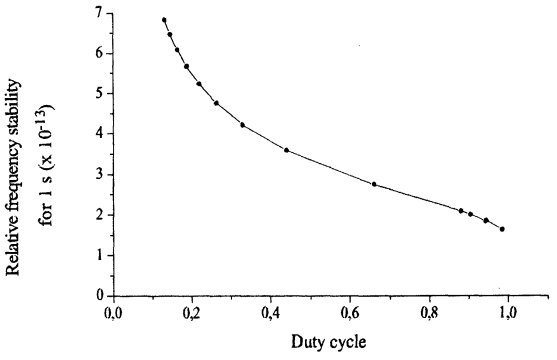


Fig. 5. *Relative frequency stability for one second versus the duty cycle computed in the case of the best oscillator.*

The short-term stability limit due to the Dick effect may be compared with the frequency stability given by the Allan standard deviation for a pulsed clock where the only limiting effect would be the S/N ratio (equation 4):

$$\sigma_y(\tau) = \frac{2}{\pi}\frac{\Delta\nu}{\nu_0}\frac{1}{S/N} \times \sqrt{\frac{T_c}{\tau}} \qquad (4)$$

where $\Delta\nu$ is the FWHM (Full Width at Half Maximum) of the clock signal, ν_0 is the microwave frequency, S/N is the signal to noise ratio and τ is the integration time.

We state that $S/N = 1/\sqrt{N_{at}}$, where N_{at} is the number of atoms which contributes to the clock signal. We have calculated the frequency stability $\sigma_y(\tau)$ at one second by varying the number of atoms contributing to the clock signal when the cycling time is 250 ms long. The frequency stability will be limited by the Dick effect at 4.8×10^{-13} level when $N_{at} > 3.0 \times 10^4$. As we cool down a large number of atoms, the Dick effect may be the main limit to the stability of HORACE.

IV. CONCLUSION

We demonstrate that we have obtained cold atoms with isotropic cooling in a quartz cell surrounded by a copper cavity when the gravity axis is perpendicular to the fiber plane. In this case, temperatures lower than 10 μK have been measured. The HORACE clock presents a potentially good short-term frequency stability, mainly limited by the Dick effect. The temporal scheme of HORACE will allow us to improve experimentally the short-term frequency stability playing with the duty-cycle, the cycle time, and the shape of the interrogation pulses. The inaccuracy may be as well improved. Since the best parameters for the short-term stability are not the same that the ones for the accuracy, there is a trade-off between accuracy and short-term frequency stability for HORACE. But HORACE may also favor accuracy against stability (or the contrary) depending on the frequency performances one wants for its own application.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge CNES (Centre National d'Etudes Spatiales) and DGA (Délégation Générale pour l'Armement) which support this work. We are very grateful to Daniel Guitard and Jacques Dupont for their technical support. We are specially thankful to Roland Barillet for his careful read of this paper and his judicious corrections.

REFERENCES

[1] E. Guillot, P. Petit, C. Valentin, and N. Dimarcq, *HORACE: a new advance for compact clocks with cold atoms*, in Proc. of 12^{th} European Frequency and Time Forum, Warsaw, Poland, 1998.

[2] E. Guillot, P.-E. Pottie, C. Valentin, P. Petit, and N. Dimarcq, *HORACE : Atomic clock with cooled atoms in a cell*, in Proc. of 13^{th} Joint Meeting EFTF - IEEE IFCS, Besançon, France, pp 81-84, 1999.

[3] C. Guillemot, Gh. Vareille, C. Valentin, and N. Dimarcq, *3D cooling of cesium atoms with isotropic light*, in Proc. of 11^{th} European Frequency and Time Forum, Neuchâtel, Switzerland, pp 156-160, 1997.

[4] C. Valentin, N. Dimarcq, E. Fretel, G. Dudle, A. Joyet, and P. Thomann *Numerical simulations of atomic behaviour for several configurations of cold cesium atomic clocks*, in Proc. of 13^{th} Joint Meeting

EFTF - IEEE IFCS, Besançon, France, pp 100-103, 1999.

[5] G.J. Dick *Local oscillator induced instabilities in trap ion frequency standards*, in Proc. of Precise Time and Time Interval (PTTI), Redendo Beach, pp 133-147, 1987.

[6] G.J. Dick, J.D. Prestage, C.A. Greenhall, and L. Maleki, *Laser oscillator induced degradation of medium term stability in passive atomic frequency standards*, in Proc. of 2nd Annual Precise Time and Time International (PTTI), Applications and Planning Meeting, pp 487-508, 1990.

[7] Emmanuel Guillot, *Nouveau concept d'horloge atomique : HORACE, HOrloge à Refroidissement d'Atomes en CEllule*, Thesis from the Université Paris-Sud, Orsay, 2000.

[8] S. Ghezali, Ph Laurent, S.N. Lea, and A. Clairon, *An experimental study of the spin-exchange frequency shift in a laser-cooled cesium fountain frequency standard*, Europhys. Lett., vol. 36, pp 25-30, 1996.

[9] Giorgio Santarelli, Claude Audoin, Ala'a Makdissi, Philippe Laurent, G. John Dick, and André Clairon, *Frequency stability degradation of an oscillator slaved to a periodically interrogated atomic resonator*, IEEE Trans. Ultra., Ferro. and Freq. Control, vol. 45, pp 887-894, 1998.

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

AN OLD METHOD FOR MAGNETIC FIELD MAPPING APPLIED IN A NEW WAY IN A CS FOUNTAIN FREQUENCY STANDARD

G.A. Costanzo [#], D. M. Meekhof [ψ], S.R. Jefferts [ψ], A. De Marchi [#]

Dipartimento di Elettronica, Politecnico di Torino
C.so Duca degli Abruzzi, 24
10129 - TORINO, Italy
e-mail: costanzo@polito.it

ψ Time and Frequency Division
National Institute of Standard and Technology
325 Broadway
80303 - Boulder, USA

ABSTRACT

Cesium fountains are now being implemented as primary frequency standards in many important metrological laboratories in the world. Results demonstrating 10^{-15} relative frequency accuracy have been obtained. The nearly monokinetic velocity distribution of the atoms and the high Q of the atomic resonator facilitate the accuracy evaluation of the standard.

Furthermore, in the Cs fountain scheme, C-field evaluation is considerably easier compared to the classical Cs beam standards because the static magnetic field uniformity can be measured by using the atomic cloud as a spatial probe of the magnetic field.

An old method using low frequency $\Delta F=0$, $\Delta m_F=1$ transitions is proposed as a different way for C-field mapping. Results obtained with the NIST-F1 Cs fountain apparatus are here reported.

1. INTRODUCTION

A few years ago the technique of laser-cooling was presented to the scientific community as a possible solution to overcome the limits of Cs frequency standards based on continuos thermal beams.

Presently Cs atoms can be cooled down to a few microkelvin thanks to sub-Doppler cooling mechanisms which allow velocities of the order of cm/s and long interaction times with the microwave interrogating signal. As a matter of fact neutral atom manipulation opened a new way to reach accuracies in the low 10^{-15} range.

In the early 90's LPTF preliminary evaluations [1,2] gave excellent results encouraging several laboratories in the world to start projects in atomic frequency standard field.

Time and Frequency Division of NIST has recently completed the realization of a Cs fountain. The prototype has now operated as a frequency standard and preliminary results were presented during the 1999 Joint Meeting EFTF-IEEE FCS Conference [3].

This project is inserted in a collaboration program between NIST (USA), Istituto Elettrotecnico Nazionale "G. Ferraris" (Italy) [4] and Politecnico di Torino (Italy).

2. FOUNTAIN SETUP OVERVIEW

A full analysis of the NIST fountain can be found in several papers proposed in the past [5,6] with a schematic diagram of the main subcomponents. Only a short description will be presented in this section.

NIST F-1 fountain uses a pure optical molasses to capture and cool atoms to a few μK in a low pressure vapor cell connected to a simple Cesium oven. This relatively high Cs pressure region is separated from the drift zone by a graphite getter which allows very low Cs pressure in the upper vacuum tank where atoms perform the ballistic flight for Ramsey interrogation process.

Atoms are accelerated upwards by detuning the vertical beams symmetrically by $\Delta\nu$: a moving frame is created in order to launch the atoms at velocity $v=\Delta\nu \cdot \lambda$ in the laboratory frame.

Thereupon the intensity and frequency of the lasers are controlled by acousto-optical modulator performing a post-cooling process. After the loading-launching sequence atoms enter a shielded region where a first microwave cavity is used as state selection.

A π-pulse at 9.192GHz moves the atoms in the

G. A. Costanzo and A. De Marchi are affiliated with Instituto Nazionale di Fisica della Materia (INFM), Torino, Italy

$|F=4,m_F=0>$ to the $|3,0>$ state. All the remaining atoms in $|F=4, m_F\neq0>$ are then pushed away applying an optical pulse.

Only atoms in $|3,0>$ state move to the second cylindrical cavity where the Ramsey interrogation process is accomplished thanks to the parabolic flight of the atomic sample.

The state selection cavity, Ramsey cavity and drift zone are magnetically shielded by four moly-Permalloy cylindrical tubes. An innermost cylinder is used to sustain the C-field coil necessary to provide the quantization field along the parabolic flight in the Ramsey region. A highly stable current supply is used to obtain the appropriate magnetic field.

Atoms detection and normalization is finally achieved after the detection of the $|F=4>$ and $|F=3>$ populations in a scheme similar to that used in LPTF fountain. The population of both the hyperfine levels is measured and Ramsey pattern is given after normalization which reduces signal variations due to shot-to-shot atom number fluctuation.

The error signal is obtained as difference of normalized signals measured when the microwaves are switched from one side of the central fringes to the other side.

3. C-FIELD RELATED SHIFT

In a Cs beam frequency standard the main source of frequency shift is due to the static magnetic field. The latter is necessary to remove the degeneracy of the magnetic substates in the Cs atom.

The Zeeman effect causes a frequency shift of the order of 10^{-10} when thermal beams are used and the obtainable field uniformity along the microwave interaction zone is often a limit to the accuracy of the standard.

Of course lowering the field helps to reduce this shift but the effects due to the existence of neighboring transitions can be huge because the Rabi pedestal is very wide. These effects are Rabi and Ramsey pulling as well as Majorana transitions which are difficult to be evaluated in classical frequency standard. Finally the chosen value of the field in the classical thermal beam frequency standard is a trade-off between opposing requirements.

However in fountain the shift is much smaller due to the small (10^{-7}T) C-field strength used and the Rabi pedestal is of the orders of 100Hz, two order of magnitude smaller than in the optically pumped Cs standards.

From Breit-Rabi formula the clock transition shift has a second order coefficient of $427\cdot10^8$[Hz/T^2]. At a field value of about 0.1μT as normally used in NIST F-1 standard this shift is of the order of $4\cdot10^{-14}$ in terms of relative units. The uncertainty due to this shift must be less than 1% when accuracy in the low 10^{-15} level is required.

The uncertainty ascribed to the quadratic terms of Zeeman effect comes from three distinct contributions:

1.random errors in the measurements of the static magnetic field,

2.magnetic field inhomogeneity,

3.magnetic field instabilities.

First an uncertainty in the measured magnetic field causes an uncertainty in the correction which can be calculated as

$$\delta(\Delta\nu_B)= 2\times427\cdot10^8\, B\cdot\delta B$$

when the uncertainty δB on the measured magnetic field is taken into account. In order to evaluate this shift the temporal average of the magnetic field as seen by the atoms is measured by a comparison of the central Ramsey fringe for the clock transition with the $m_F=1$ linearly dependent transition.

The second type of uncertainty in C-field evaluation arises from the magnetic field inhomogeneity. The resonance signal from $\Delta F=1, m_F=1$ transition is a good test in order to evaluate the average magnetic field along the trajectory of the atoms. Moreover the fringes contrast is a measure of the field homogeneity as sampled by the atoms.

The third uncertainty arise from temporal instabilities of the C-field. The central fringe of the field dependent transition can be used to lock the fountain and to infer the magnetic field fluctuations inside the standard through the frequency stability measurements.

4. C-FIELD MEASUREMENTS USING LOW FREQUENCY $\Delta F=0, \Delta m_F=1$ TRANSITIONS

In the past low frequency transitions were used as a method to monitor the homogeneity of the C-field along the beam path [7]. This technique is an useful tool which permits C-field uniformity evaluation and also a simple check for the beam tube because a low frequency resonant signal can be obtained. The main limit to the measurements were the rather broad resonance signal and the bias induced by Millman effect.

We use a different approach during an accuracy evaluation for mapping the magnetic field. A field map is generated by measuring the frequency from the $|4,0>$ to the $|4,1>$ magnetically sensitive transition.

Low frequency excitation is induced by using coils parallel to the flight path. These coils generate a magnetic field normal to the C-field direction. The low frequency field causes $\Delta m=1$ transitions when the frequency and field strength are properly adjusted. By applying the low frequency pulse symmetrically about apogee the Millman effect can be suppressed. Figure 1 shows a typical resonance signal (dotted curve) of the

rotary field method when LF power is set to the optimum level. At higher power level a resonant signal broadening is observed (triangle).

Thanks to the relatively high interaction time of the atoms probe with the LF excitation a 10Hz (FWHM) linewidth is obtained. These data, after numerical interpolation, are used to evaluate the magnetic field shape at different heights above the Ramsey cavity.

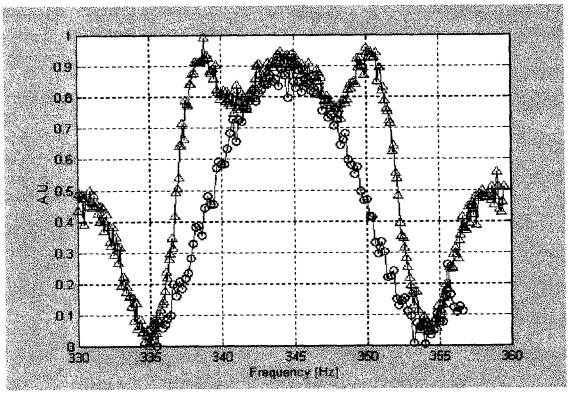


Figure 1: low frequency transition when LF power is set to the optimum (circle) or higher (triangle) values.

The C-field is mapped during each evaluation. The results of one such measurement are shown in Fig.2. The effects of the field leakage from the holes in the end caps can be clearly seen (above 80cm), as can the effect of the highly permeable vacuum feedthroughs on the microwave cavities (state selection cavity and Ramsey cavity).

5. RESULTS

NIST F-1 fountain uses the low frequency transition as a check to the field homogeneity which, in general, can be measured in two different manners.

The first one is the so called antenna method which uses a 10ms long microwave pulse to excite the atomic transition when the atoms are at the apogee of their parabolic flight above the Ramsey cavity.

The frequency difference between the clock transition and the field sensitive transition is measured. The C-field value is inferred from the frequency difference and a map is generated by varying the launch height.

The second method uses a 100ms long low frequency pulse in order to excite the atoms at their apogee. The frequency of this transition has a sensitivity with the magnetic field of about 350kHz/gauss.

The magnetic field inhomogeneity causes the central fringe on the magnetically sensitive transition to be displaced from the center of the underlying Rabi pedestal. In order to evaluate the second-order Zeeman correction and to unambiguously recognize the central fringe position, a plot of the time average magnetic field is necessary.

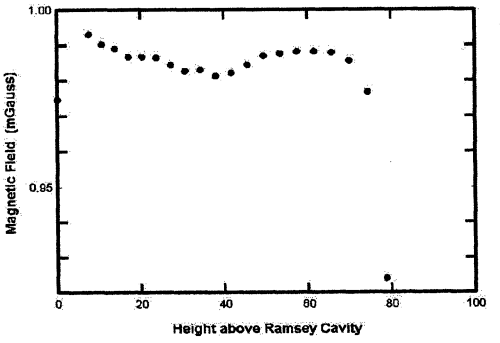


Figure 2: magnetic field map inside the C-field region.

Figure 3 shows the fringes position measured by different tossing height above the Ramsey cavity. The dots are the position of various Ramsey fringes from $m_F=1$ hyperfine transition at different distance above Ramsey cavity. The continuous line is the predicted position of the central fringe from the time integral of the magnetic field measured through the low frequency transition excitation as shown in Fig.2.

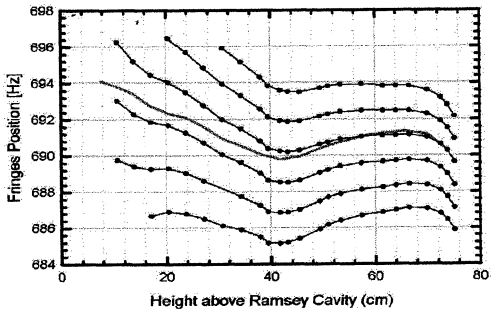


Figure 3: fringes position of the field dependent transition (dotted line).

The agreement is quite good for apogee more than 40cm above the cavity. We believe that the disagreement for distances between 10cm and 40cm is a result of a lack of knowledge of the magnetic field between the Ramsey cavity and the top of the cutoff waveguide.

6. SUMMARY

We described a new method for C-field homogeneity measurements which are necessary to evaluate one of the most important shift in atomic fountain frequency standard. Low frequency resonant signals are induced by using coils parallel to the flight path. Measurements of C-field homogeneity obtained with this method were reported.

7. REFERENCES

[1] A. Clairon, P. Laurent, G. Santarelli, S. Ghezali, S.N. Lea, M. Bahoura, "A Cesium fountain frequency standard: preliminary results", IEEE Trans. on I&M, vol.44, no.2, April 1995

[2] A. Clairon, S. Ghezali, G. Santarelli, P. Laurent, S.N. Lea, M. Bahoura, E. Simon, S. Weyers, K. Szymaniec, "Preliminary accuracy evaluation of a Cesium fountain frequency standard", in Proc. 5th Symp. on Freq. Standards and Metrology, 1997

[3] S.R. Jefferts, D.M. Meekhof, J.H. Shirley, T.E. Parker, F. Levi, "Preliminary accuracy evaluation of a Cesium fountain primary frequency standard at NIST", in Proc. of Joint Meeting EFTF-IEEE FCS, 1999

[4] F. Levi, A. Godone, D. Calonico, L. Lorini, "Progress toward the realization of a Cesium fountain frequency standard at IEN", these Proceedings

[5] S.R. Jefferts, D.M. Meekhof, L.W. Hollberg, D. Lee, R.E. Drullinger, F.L. Walls, C. Nelson, F. Levi, T.E. Parker, "NIST Cesium fountain frequency standard: preliminary results", in Proc. of IEEE-FCS 98

[6] S.R. Jefferts, R.E. Drullinger, A. De Marchi, "Design and realization of the NIST fountain cavity", in Proc. of IEEE-FCS, 1998

[7] J.Vanier, C. Audoin, "The quantum physics of atomic frequency standards", ed. Hilger, 1989

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

A DYNAMIC MODEL OF THE LO NOISE TRANSFER IN VAPOR CELL PASSIVE FREQUENCY STANDARDS FOR GENERAL MODULATION SCHEMES

Massimo Ortolano†, Nicolò Beverini‡, Andrea De Marchi†
† Politecnico di Torino, Torino, Italy
‡ Università di Pisa, Pisa, Italy
The authors are affiliated to Istituto Nazionale di Fisica della Materia

Andrea De Marchi
Dipartimento di Elettronica — Politecnico di Torino
Corso Duca degli Abruzzi, 24
10129 TORINO, ITALY.
Phone: +39-011-5644049 — Fax: +39-011-5644099
E-mail: demarchi@polito.it

ABSTRACT

The method described in [1] as a mean to study the local oscillator noise transfer in Rb-Cell frequency standards, with square wave frequency modulation of the interrogation signal, is here generalized for analyzing different modulation schemes. The dynamic analysis shows a strong dependence of the noise transfer on the operational parameters of the cell, while the quasi-static analysis of [2] does not.

1. INTRODUCTION

Different approaches have been used, in the last few years, to describe the transfer of the local oscillator (LO) frequency noise, at even harmonics of the modulation frequency, to the output of a passive frequency standard [2, 3, 4]. Both theoretical analysis [3, 4] and experimental data [5] suggest that the aliasing effect could be reduced by using a square-wave modulated interrogation signal. However, at fast modulation frequencies, large discrepancies between theoretical models, developed using the quasi-static approximation, and experimental data have been found [3, 5]. This means that, for an effective optimization of a passive frequency standard, the dynamic behavior of the atomic system, and its dependence on the operational parameters of the cell, should be considered.

In this work, we briefly describe a dynamic model of a cell frequency standard, based on the density matrix formalism and on signal analysis methods, which is suitable to analyze different modulation schemes. Although several simplifications have been introduced in the model, the basic behavior of the cell should not be significantly altered. The model is here developed for a Rb-cell frequency standard, but it can be easily modified to describe a Cs-cell standard. Unfortunately, currently available experimental data are not sufficient for a complete validation of the model.

2. THE CELL MODEL

The evolution of an atomic system can be calculated by means of the density matrix formalism. The rate equations for the density matrix of an optically pumped Rb-cell are derived in [6]. In [1] the rate equations were modified to take into account, as a perturbative term, the frequency noise of the local oscillator. That model was then applied for analyzing the aliasing effect in a square-wave frequency modulated Rb-cell standard. Here the model is generalized in order to describe the response of the cell to an arbitrary waveform (phase and frequency) modulation. The basic assumptions used in the development of the model are: 1) The operating conditions of the cell are supposed to be uniform over its entire volume; 2) The cell contains a buffer gas; 3) Optical pumping is assumed to take place on the D_2 line between levels $5S_{1/2}$, $F = 2$ and $5P_{3/2}$; 4) The contribution to the optical signal coming from Zeeman sublevels with $m_F \neq 0$ is negligible; 5) Zero light shift.

The ensemble of atoms is described by means of the density matrix $\boldsymbol{\rho} = [\rho_{ij}]$, which is an 8×8 hermitian matrix. By means of condition 1), the matrix elements are functions of time but not of space. Condition 4) means that the optical signal is proportional to $\rho_{77} - \rho_{33}$. The microwave interrogation signal is supposed to be phase modulated at frequency f_m with phase $\phi(t)$. The angular frequency of the carrier is ω, and the LO phase noise is represented by a random process $\delta\phi(t)$. The description of parameters

used is given below.

Parameter	*Description*
ω_0	Rb clock angular frequency
Γ	optical pumping rate
γ_1	longitudinal relaxation rate
γ_2	transversal relaxation rate
$\Gamma_1 = \gamma_1 + \Gamma/2$	total longitudinal relaxation rate
$\Gamma_2 = \gamma_2 + \Gamma/2$	total transversal relaxation rate
$\Delta\nu = \Delta\omega/2\pi$	FWHM linewidth (includes power broadening)
f_m	modulation frequency
ω	LO angular frequency
$\phi(t)$	LO signal phase
$\delta\phi(t)$	LO phase noise
$\delta\omega(t) = \delta\dot{\phi}(t)$	LO angular frequency noise
$\Delta = \omega - \omega_0$	angular detuning
b	Rabi angular frequency
$S \propto b^2$	Saturation factor

It can be shown that with the following change of variables

$$\begin{aligned} u &= \mathrm{Re}\left\{\rho_{37}\, \mathrm{e}^{-\mathrm{j}(\omega t+\delta\phi(t))}\right\} \\ v &= \mathrm{Im}\left\{\rho_{37}\, \mathrm{e}^{-\mathrm{j}(\omega t+\delta\phi(t))}\right\} \\ z &= \frac{1}{2}(\rho_{77} - \rho_{33}) \\ p_+ &= (\rho_{11} + \rho_{22} + \rho_{44} + \rho_{55})/4 \\ p_- &= (\rho_{66} + \rho_{88})/2 \end{aligned} \tag{1}$$

the rate equations can be put in the compact matrix form

$$\dot{\mathbf{x}} = [\mathbf{a}(t) + \delta\omega(t)\,\mathbf{s}]\,\mathbf{x} + \mathbf{c}, \tag{2}$$

where

$$\mathbf{x} = (u, v, z, p_+, p_-)^T$$

$$\mathbf{c} = (0, 0, \Gamma/4, \Gamma_1/8, \Gamma_1/8)^T$$

$$\mathbf{a}(t) = \begin{pmatrix} -\Gamma_2 & -\Delta & -b\sin\phi(t) & 0 & 0 \\ \Delta & -\Gamma_2 & -b\cos\phi(t) & 0 & 0 \\ b\sin\phi(t) & b\cos\phi(t) & -\Gamma_1 & -\Gamma & -\Gamma/2 \\ 0 & 0 & -\Gamma/8 & -\Gamma_1^+ & -\Gamma/8 \\ 0 & 0 & -\Gamma/8 & \Gamma/4 & -\Gamma_1^- \end{pmatrix} \tag{3}$$

$$\mathbf{s} = \begin{pmatrix} 0 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

and

$$\begin{aligned} \Gamma_1^+ &= \Gamma_1 + \frac{\Gamma}{4} \\ \Gamma_1^- &= \Gamma_1 - \frac{3}{8}\Gamma \end{aligned} \tag{4}$$

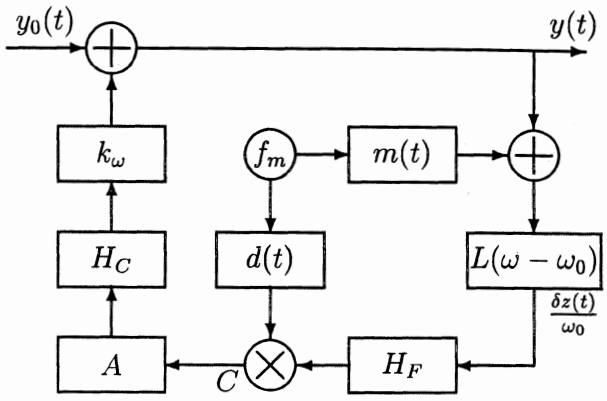


Figure 1: Block diagram of a vapor-cell frequency control loop.

The matrix $\mathbf{a}(t)$ is periodic with frequency f_m. Since the LO noise $\delta\omega$ is assumed to be small, equation (2) can be solved by means of perturbation theory, by expanding $\mathbf{x}(t)$ in series of $\delta\omega$, as in $\mathbf{x}(t) = \mathbf{x}_0(t) + \delta\omega\,\mathbf{x}_1(t)$. By substituing the latter in equation (2), and equating separately terms up to the first order in $\delta\omega$, the following two differential equation are obtained for $\mathbf{x}_0$ and $\delta\mathbf{x} = \delta\omega\,\mathbf{x}_1$

$$\dot{\mathbf{x}}_0 = \mathbf{a}(t)\,\mathbf{x}_0 + \mathbf{c} \tag{5}$$

$$\dot{\delta\mathbf{x}} = \mathbf{a}(t)\,\delta\mathbf{x} + \delta\omega\,\mathbf{s}\mathbf{x}_0 \tag{6}$$

The components of $\mathbf{x}_0(t)$, which are periodic at frequency f_m, and of $\delta\mathbf{x}(t)$, which are random processes, are

$$\mathbf{x}_0(t) = (u_0(t), v_0(t), z_0(t), p_{0+}(t), p_{0-}(t))^T \tag{7}$$

$$\delta\mathbf{x}(t) = (\delta u(t), \delta v(t), \delta z(t), \delta p_+(t), \delta p_-(t))^T. \tag{8}$$

Equation (5) can be solved for arbitrary modulation waveform by means of FFT techniques, and an approximate solution to equation (6) is found in [1].

By means of assumption 3), the optical signal at the output of the cell is proportional to $z(t) = z_0(t) + \delta z(t)$. In a frequency standard, this signal is synchronously demodulated at frequency f_m. The demodulated signal is then used to voltage control the LO frequency, as shown in the block diagram of figure 1, which represents a typical frequency control loop of a cell frequency standard.

The cell is represented by the block marked $L(\omega - \omega_0)$, and $\delta z(t)$ is its response to the LO noise. The modulation and demodulation waveforms, which are periodic with frequency f_m, are $m(t)$ and $d(t)$,respectively. The LO fractional frequency instability is $y_0(t) = \delta\omega(t)/\omega$, while $y(t)$ is the output fractional frequency instability. Let μ_d and μ'_m be the Fourier coefficients of $d(t)$ and $u_0(t)$,

respectively. It is shown in [1] that the power spectral density of the output frequency noise $S_y(f)$ can be written as

$$S_y(f) = \frac{S_{y_0}(f)}{K^2|H_C(f)|^2 D^2} + W \tag{9}$$

where

$$\begin{aligned} W &= \sum_{r=1}^{\infty} a_r^2 S_{y_0}(r f_m) \\ a_r &= \frac{\sqrt{2}}{D}\left|\sum_{d=1}^{\infty} \mu_{-r,d} H_G^*(d f_m) + \mu_{r,d}^* H_G(d f_m)\right| \\ D &= 2\sum_{d=1}^{\infty} \mathrm{Re}\{\mu_{0,d}^* H_G(d f_m)\} \\ \mu_{r,d} &= \mu_d \mu'_{r-d} \end{aligned} \tag{10}$$

and $S_{y_0}(f)$ is the power spectral density of $y_0(t)$.

Usually, only odd harmonics of the modulation frequency appear in $d(t)$ and $u_0(t)$. This is sufficient to have $a_r = 0$ if r is odd. In this case the contribution of the white noise term W to the short-term frequency instability of the clock was written in [5, 3] as

$$\sigma_y^2(\tau) = \sum_{r=1}^{\infty} C_{2r}^2 S_\phi(2r f_m)\, \tau^{-1} \tag{11}$$

where the C_{2r} coefficients are simply related to the a_{2r} coefficients of equation (10) by the following formula

$$C_{2r} = \frac{\sqrt{2}\pi f_m}{\omega_0} 2r\, a_{2r} \tag{12}$$

3. RESULTS

The model described in section 2 is here used to analyze the cases of square-wave and sine-wave frequency modulations, both with sine-wave demodulation. Preliminary investigations [1] have shown a strong dependence of the aliasing effect on the working conditions of the cell. The latter must be known in order to compare any experimental result with this theoretical analysis. Unfortunately, this was not the case for the experiment described in [5], since the operating conditions have not been reported.

Figure 2 shows the first aliasing coefficients as a function of modulation frequency for square-wave frequency modulation and sine-wave demodulation. Figure 3 shows the same for sine-wave frequency modulation and demodulation. This analysis has been carried out with the following parameters: $\gamma_1 = 300\,\mathrm{s}^{-1}$, $\gamma_2 = 290\,\mathrm{s}^{-1}$, $\Gamma = 850\,\mathrm{s}^{-1}$, and $b = 880\,\mathrm{rad/s}$ which gives a saturation factor $S \simeq 2$ and a linewidth $\Delta\nu \simeq 400\,\mathrm{Hz}$. These are the same of a typical cell described in [6], although large changes should be expected from case to case.

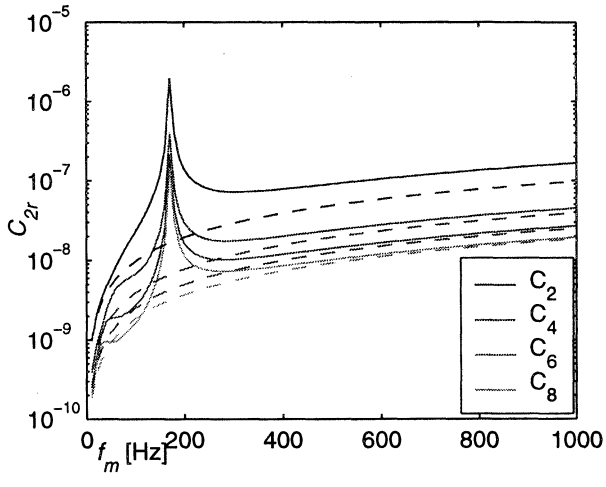


Figure 2: First aliasing coefficients as a function of modulation frequency for square-wave frequency modulation and sine demodulation. Coefficients calculated in the quasi-static approximation are shown as dashed lines, for comparison.

Figure 2 shows that, for low modulation frequencies, the aliasing coefficients converge toward the values given by the quasi-static analysis of [2, 3], but at about the Rabi frequency a strong peak appears. Far and above the Rabi frequency the aliasing coefficients are close again to the quasi-static values. Moreover, it appears from figure 2 that dynamically calculated aliasing coefficients are greater than those calculated with the quasi-static approximation. Although this could be expected, it should not be taken as a general behavior, as it was shown in [1] with different parameters. The same Rabi resonance is shown in figure 3, where the coefficients have been calculated by using sine-wave modulation. It can be seen that the coefficient C_2 is slightly greater with sine-wave modulation than with square-wave modulation, but higher order coefficients become negligible very quickly with sine-wave demodulation.

4. CONCLUSIONS

The dynamic analysis of the cell behavior shows a strong dependence of the aliasing effect on the operational parameters of the cell, while quasi-static analysis does not. In particular, aliasing coefficients greatly increase when the modulation frequency is close to the Rabi frequency of the cell, which is unfortunately what it's usually chosen to be. Furhermore, coefficients a_r don't vary if both the modulation frequency and the Rabi frequency are changed, keeping their ratio constant. This means that further investigations, both theoretical and experimental, should be done before deciding which is the best modula-

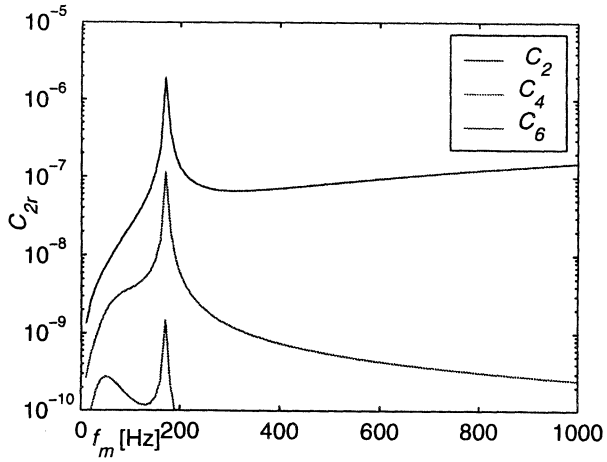


Figure 3: First aliasing coefficients as a function of modulation frequency for sine-wave frequency modulation and demodulation.

tion scheme for high stability vapor-cell frequency standard.

5. REFERENCES

[1] Massimo Ortolano, Nicolò Beverini, and Andrea De Marchi, "A dynamic analysis of the LO noise transfer mechanism in Rb-Cell frequency standards", IEEE Transasctions on Ultrasonics, Ferroelectrics and Frequency Control, vol. 47, pp. 471–474, March 2000.

[2] Claude Audoin, Vincent Candelier, and Noël Dimarcq, "A limit to the frequency stability of passive frequency standards due to an intermodulation effect", IEEE Transactions on Instrumentation and Measurement, vol. 40, pp. 121–125, April 1991.

[3] A. De Marchi, L. Lo Presti, and G. D. Rovera, "Square wave frequency modulation as a mean to reduce the effect of local oscillator instabilities in alkali vapor passive frequency standards", In Proceedings of 1998 International Frequency Control Symposium, 1998.

[4] G. J. Dick and C. A. Greenhall, "L.O. limited frequency stability for passive atomic frequency standards using "square wave" frequency modulation", In Proceedings of 1998 International Frequency Control Symposium, 1998.

[5] J. Q. Deng, A. De Marchi, F. L. Walls, and R. E. Drullinger, "Reducing the effects of local oscillator PM noise on the frequency stability of cell based passive frequency standards", In Proceedings of 1998 International Frequency Control Symposium, 1998.

[6] J. Vanier and C. Audoin, The quantum physics of atomic frequency standards, Adam Hilger, Bristol (UK), 1989.

THE USNO CESIUM FOUNTAIN

Christopher R. Ekstrom, Eric A. Burt, and Thomas B. Swanson

U. S. Naval Observatory
3450 Massachusetts Ave. NW
Washington, D. C. 20392 U.S.A.

1. ABSTRACT

We are pursuing a program that will integrate atomic fountain-based clocks into the USNO Master Clock. We will present data from our atomic fountain, including initial measurements of our device relative to an active hydrogen maser. As this maser is part of the USNO timing ensemble, we can easily relate the frequency stability of the fountain to any of our clocks or internal timescales. We will also present several possible continuous operation strategies for the incorporation of this new class of clock into the local clock ensemble.

2. INTRODUCTION

The last decade has seen advances in laser cooling and trapping that have allowed the practical construction of atomic fountain frequency standards. These devices are now producing results at several national standards institutions [1].

We have undertaken a program to integrate atomic fountain clocks into the timing ensemble at the U. S. Naval Observatory (USNO). The mission of the Observatory does not require that any of our standards be accurate realizations of the second, only that they be stable and run continuously. Our goal is to have several atomic fountain standards running with a short term stability of $1-2x10^{-13}\tau^{-1/2}$ and a statistical floor of $1-3x10^{-16}$.

3. EXPERIMENTAL LAYOUT

The physical layout of our fountain is shown in Figure 1. In addition to the small optical table that houses the vacuum chamber, there is a larger table that houses all of the lasers, with optical fiber coupling of all light onto the table with the vacuum chamber. All background pressures in the vacuum chamber are below $4x10^{-8}$ Pa.

We collect atoms in either a MOT or molasses and launch them in a (1,1,1) geometry. The laser light for the upward and downward directed laser beams are generated by two injection seeded tapered amplifiers. The light is then transported to the vacuum chamber with optical fibers. Both of these beams (and the detection light optical fiber) have power servos closed around the fiber path to reduce amplitude noise at the atoms. We can operate with up to 30 mW of laser light per collection and launching beam.

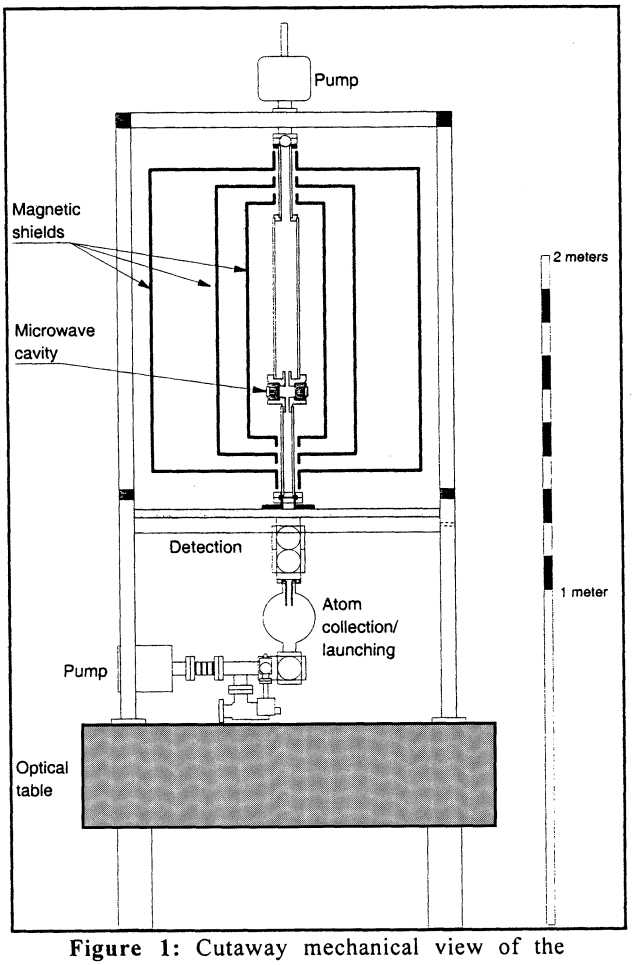


Figure 1: Cutaway mechanical view of the fountain.

The atoms are launched in two phases. The first phase applies a violent acceleration for 1.4 milliseconds with an average detuning from resonance of 6 MHz. A second phase follows immediately with an average detuning of 50 MHz for 0.8 milliseconds, half the laser intensity, and a linear ramp of the intensity to zero at the end of the launch. We measure launch temperatures of 1.6 (0.2) μK by monitoring the vertical width of the launched cloud when it passes the detection region both on the way up and on the way down, which allows us to remove effects of the initial cloud size.

Immediately after launch, the atoms are pumped into the F=4 hyperfine levels with a vertical repumping beam tuned to the F=3 to F'=4 transition. The atoms are state selected at the detection zone. They are exposed to a 3 millisecond long shaped pulse of 9.2 GHz microwaves from an axial loop antenna inside the vacuum chamber. This pulse transfers the F=4, m_F=0 atoms to the F=3, m_F=0 state. The remaining F=4 atoms are removed from the atomic sample with a transverse traveling wave laser beam in the lower detection zone.

The microwave cavity and drift region are temperature stabilized to 0.1 °C and enclosed in a set of three magnetic shields. The shields are made from 1.6 mm thick moly permalloy (MIL-N-14411B, "comp 1") and have an axial shielding factor of 30,000. Details of the construction of the shields have been provided previously [2]. An axial solenoid provides a 230 nT magnetic field for the cavity and free precession regions.

After making two transits of the microwave cavity the atoms return to the detection region. The upper detection zone monitors the F=4 population, and the non-retroreflected laser beam removes these atoms from the sample. The second detection zone monitors the F=3 population by collecting fluorescence with both F=4 to F'=5 resonant light and F=3 to F'=4 repumping light present. The signals are collected, background levels are subtracted from each signal, and the F=4 signal is normalized by the sum of the F=3 and F=4 signals.

We run the fountain with a total cycle time between 1.1 and 1.9 seconds.

3. RESULTS

Figure 2 shows a microwave Ramsey resonance pattern from our fountain. The local oscillator is an active hydrogen maser from our clock ensemble. The microwave frequency chain is locked to a 100 MHz output of the maser.

The signal to noise at the fringe half height points is roughly 40, while the signal to noise at the fringe peaks is 140, indicating that we are limited by technical and not shot noise. One possible cause of this noise is the long run (approximately 100 meters) of cable between the maser and our frequency chain. In the near future, we plan to move a maser to the same room as the atomic fountain.

We have also collected data while sitting on one side or hopping between sides of the central microwave fringe. The transition probability can be easily converted to a fractional frequency measurement of the local oscillator. The local oscillator for this measurement was an active hydrogen maser from our timing ensemble. Figure 3 shows the Allen deviation from a typical data set. The maser is more stable than the fountain on all time scales for this data set.

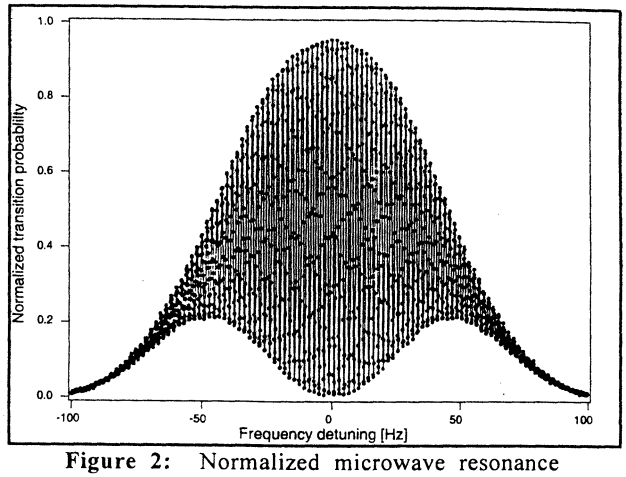


Figure 2: Normalized microwave resonance pattern with no averaging. Cycle time is 1.4 seconds and the FWHM fringe width is 0.95 Hz. Lines connect data points only.

Because the maser is also monitored by our measurement systems, we can reference our frequency measurement of the maser to any other clock or average within our ensemble. This will also allow us to easily reference the long term stability of our fountain to any other time scale, such as TAI or local realizations of UTC, that we monitor with our time transfer efforts.

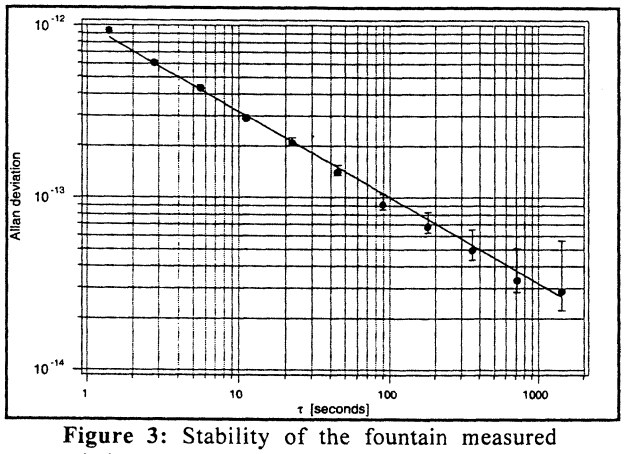


Figure 3: Stability of the fountain measured relative to an active hydrogen maser. The solid line is $10^{-12}\tau^{-1/2}$ and is intended for reference only.

5. OPERATIONAL POSSIBILITIES

We are considering several possible ways of running our fountain as a continuous clock. The first is to directly steer a high quality quartz crystal. The short term stability of our best crystal will require us to close our steering control loop in 2 to 4 seconds. The steering interval also sets the time scale over which the fountain can be allowed to not steer the local oscillator.

We are planning on a using a different method of producing a steered, continuous output that employs an active hydrogen maser as a local oscillator. This would allow a steering time of roughly one hour for the

class of masers at our disposal. The steered output would be monitored by our local measurement systems and would be introduced into our timescales in an identical manner to all other classes of clocks at the USNO. We anticipate that with our expected fountain performance and the measured drift rates of the masers in our ensemble, that the fountain should be steering out the maser drift starting at 1 to 20 days. The longer steering interval greatly reduces the stress on the operating duty cycle of the underlying atomic fountain at the addition of considerable cost for the maser and steering generator. In addition, the medium term stability of the unsteered maser is much better than that of an unsteered crystal.

It is our intention to build one more research device with rubidium atoms and to then produce up to five atomic fountains for operational use at the USNO and our alternate master clock facility in Colorado. We will use rubidium-based devices due to the dramatically smaller cold collision frequency shift [3].

6. CONCLUSIONS

In conclusion, we have observed reasonably high signal to noise microwave fringes in our atomic fountain. We have presented preliminary stability measurements of the device relative to an active hydrogen maser from our clock ensemble and outlined potential strategies for moving this class of device into continuous operation at the USNO.

7. REFERENCES

[1] A. Clairon, et. al., "Preliminary Accuracy Evaluation of a Cesium Fountain Frequency Standard" in Proceedings of the Fifth Symposium on Frequency Standards and Metrology, 1995, pp. 49-59; S. Bize, et. al. "Interrogation Oscillator Noise Rejection in the Comparison of Atomic Fountains", pp. 9-11; S. R. Jefferts, et. al., "Preliminary Accuracy Evaluation of a Cesium fountain Primary Frequency Standard at NIST", pp. 12-15; S. Weyers, et. al., "First Results of PTB's Atomic Caesium Fountain", pp. 16-19; P. B. Whibberly, "Development of a Caesium Fountain Primary Frequency Standard at the NPL", pp. 24-26, all in Proceedings of the conference EFTF and IEEE FCS, 1999

[2] E. Burt, et. al., "Cesium Fountain Development at USNO", in Proceedings of the conference EFTF and IEEE FCS, 1999, pp. 20-23.

[3] Y. Sortais, et. al., "An Evaluation of the Collisional Frequency Shift in a ^{87}Rb Cold Atom Fountain", pp. 34-38; C. Fertig, et. al., "Laser-Cooled Rb Fountain Clocks", pp. 39-42, both in Proceedings of the conference EFTF and IEEE FCS, 1999

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

ATOMIC CLOCK ENSEMBLE IN SPACE (ACES): PROCESSING STRATEGIES FOR THE FREQUENCY COMPARISONS BETWEEN ACES CLOCKS AND EARTH CLOCKS.

P. Uhrich[1], P. Laurent[1], P. Lemonde[1], F. Vernotte[2],
[1]BNM-LPTF, Observatoire de Paris, 61 avenue de l'Observatoire, F-75014 Paris, France.
[2]Observatoire de Besançon, 41bis avenue de l'Observatoire, B.P. 1615, F – 25010 Besançon, France.

Corresponding author: Pierre Uhrich, BNM-LPTF,
Observatoire de Paris, 61 avenue de l'Observatoire, F – 75014 Paris, France.
Phone: +33 - (0)1 - 4051 - 2216, Fax: +33 - (0)1 - 4325 - 5542, e-mail : pierre.uhrich@obspm.fr

ABSTRACT

The Atomic Clock Ensemble in Space (ACES) is a scientific project of the European Space Agency (ESA) on-board the future International Space Station (ISS). It consists mainly in two atomic frequency standards, a Space Hydrogen Maser (SHM) and a cold atom cesium clock called "PHARAO" (for "Projet d'horloge atomique à refroidissement d'atomes en orbite"). Optical and Microwave links to Earth clocks will also be included into ACES.

The low ISS orbit at a nominal altitude of 450 km produces a nominal average rotation period around the Earth of 5610 s. As a baseline, the duration of each pass over an Earth station is assumed to be 350 s. After a few consecutive passes, the ISS will remain invisible from that Earth station before the next pass, which will start after a complete rotation of the ensemble [ISS + Earth] in an Earth reference system. Each session will be separated from the next one by a changing dead time much larger than the session itself. This situation calls for a specific processing strategy of the future data.

The first step is a gathering of the measurements based on their respective durations or analysis periods: Short Term Data (one continuous pass), Middle Term Data (a few consecutive passes separated only by one ISS rotation), or Long Term Data (one day or more). As a second step, assuming different phase noise modulations issued from the comparison mean, a spectral density analysis or an Allan variance analysis is proposed on the Short Term Data, taking into account the dead times or other effects when necessary.

1. INTRODUCTION

The Atomic Clock Ensemble in Space (ACES) [1] is a scientific project of the European Space Agency (ESA) on-board the future International Space Station (ISS). It consists mainly in two atomic standards, a Space Hydrogen Maser (SHM) [2] and a cold atom cesium clock called "PHARAO" (for "Projet d'horloge atomique à refroidissement d'atomes en orbite") [3]. The means to compare these clocks to any clock on the Earth will also be included: a Time Transfer by Laser Link (T2L2) [4], and a Microwave Link (MWL) [5].

The low ISS orbit at a nominal altitude of 450 km will be one of the major issues for the frequency comparisons between the ACES clocks and any Earth clock. The ISS nominal average rotation period around the Earth is 5610 s, about 93.5 min, and, as a baseline, the duration of each pass over an Earth station is assumed to be 350 s. Moreover, it is expected that after a few consecutive passes the ISS will remain invisible from that Earth station for about one day before the next pass, which will start after a complete rotation of the ensemble [ISS + Earth] in an Earth reference system. Each session where a frequency comparison between an ACES clock and an Earth clock could be achieved following the outstanding MWL requirements [5] will be separated from the next one by a changing dead time much larger than the session itself. This situation calls for a specific processing strategy, which is described in this first approach report.

The following lines are based on the input that the measurements will rely on phase difference readings. It is assumed that the Noise Modulation brought by the MWL in the clock comparisons will be limited to either a White Phase Noise, either a Flicker Phase Noise, or a Random Walk Phase Noise. There could be some periodic terms appearing in the phase measurements. It is assumed that no ambiguity compensations are needed.

The first step is a data gathering based on their respective analysis periods: over each continuous pass, over a few consecutive passes separated only by one ISS rotation period, or over one day or more. As a second step, assuming the phase noise modulations issued from the comparison mean, a spectral density analysis or an Allan variance analysis is proposed on the Short Term Data, taking into account the dead times or other effects when necessary.

2. DATA GATHERING

Following the ISS orbit considerations, the first step of any processing will be a gathering of the measurements based on their respective durations or analysis periods. Therefore, the data are separated into three different domains, or three different groups.

The first group of data is limited to each single session duration without any dead times due to the ISS

orbit. These data are phase differences between the system [MWL + clock] on-board and the system [MWL + clock] on the Earth, measured during each ISS pass over the same Earth station on a continuous basis. The sampling period is called τ_0. The total duration of each session is called d_{ISS}. This first group is referred to as "Short Term Data".

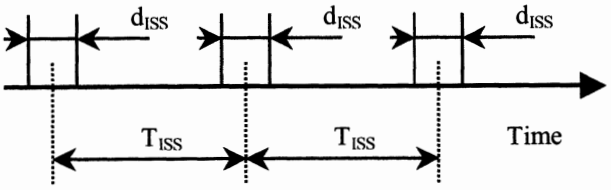


Figure 1. Time diagram of the Middle Term Data.

The second group of data is made by gathering a few consecutive sets of data of the first group. Here the word "consecutive" means separated by only one rotation period of the ISS. Each consecutive session is assumed to be of equal duration d_{ISS}. The starting date of two consecutive sessions is supposed to be separated by exactly one ISS orbit duration T_{ISS}, assumed to remain constant over the analysis period. Therefore, the dead time between two consecutive sessions is simply equal to $T_{ISS} - d_{ISS}$. The figure 1 shows such a set of "Middle Term Data". It could be assumed $T_{ISS} \approx 16 \times d_{ISS}$.

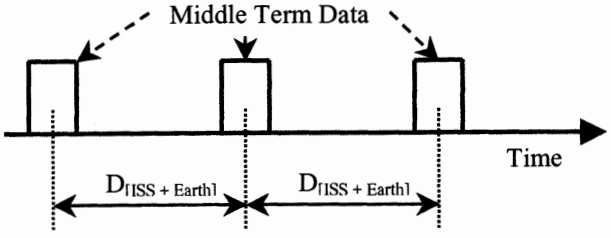


Figure 2. Time diagram of the Long Term Data.

The third group of data is obtained from the Middle Term Data as it is shown on the figure 2. It could appear as a new sampling of the phase or frequency deviations by a long term analysis period $D_{[ISS + Earth]}$. One "average" phase or frequency difference value is set at a "middle" date of each sequence of consecutive ISS passes, separated by one complete rotation of the ensemble [ISS + Earth] in the Earth reference frame. The words "average" and "middle date" are depending on the noise modulation affecting the "Long Term Data" over the considered analysis period.

3. SHORT TERM DATA PROCESSING

The n phase differences are given as an ensemble $x_i = x(t_i)$, for i = 1, ..., n, affected each by a phase noise modulation $\sigma_{xi}(\tau_0)$. The analysis is carried on a period $\tau = (n - 1) \times \tau_0$, where $\tau \leq d_{ISS}$. The variance analysis is based on the noise modulation assumptions. Assuming a noise degradation of 20 % due to the MWL in the clock readings, the Short Term Data requirements are [5]:

$$\sigma_x(\tau = 350 \text{ s}) \leq 0.3 \text{ ps} \quad (1)$$

$$\sigma_y(\tau = 350 \text{ s}) \leq 2.1 \times 10^{-15} \quad (2)$$

3.1. White Phase Noise Modulation.

Assuming a White Phase Noise Modulation over the entire Short Term Data set, the phase noise modulation associated with each sample x_i has the same value: $\sigma_{xi}(\tau_0) = \sigma_x(\tau_0)$. In this case, it can be shown that the best average frequency estimate is given by the weighted mean square linear regression over the phase difference ensemble [6]. The average frequency deviation $y(\tau)$ is then given for $\tau = (n - 1) \times \tau_0$ by:

$$y(\tau) = \frac{K_n}{\tau_0} \times \left[\sum_{i=1}^{n} (i-1) \times x(t_i) - \frac{n-1}{2} \times \sum_{i=1}^{n} x(t_i) \right] \quad (3)$$

where K_n is a function of n given by:

$$K_n = \frac{12}{(n+1) \times n \times (n-1)} \quad (4)$$

The variance $\sigma_y^2(\tau = (n - 1) \times \tau_0)$ associated to this average frequency deviation $y(\tau)$ is obtained as:

$$\sigma_y{}^2(\tau) = K_n \times \frac{\sigma_x{}^2(\tau_0)}{\tau_0{}^2} \quad (5)$$

The White Phase Noise Modulation could be considered as being averaged out when this variance fulfils the requirements (2) for $\tau = 350$ s.

It is very easy to show that the average frequency estimate would be less precise by simply integrating between the first and the last phase differences of the ensemble. In this case the variance of the average would be given by:

$$\sigma_y{}^2(\tau) = \frac{2 \times \sigma_x{}^2(\tau_0)}{(n-1)^2 \times \tau_0{}^2} \quad (6)$$

These variances in (5) and (6) are equal for a number of samples n which are solving the equation:

$$(n-2) \times (n-3) = 0 \quad (7)$$

For n = 2, there is one single straight line between two points. For n = 3, the average slope between the samples is the same, only the origins of the straight lines are different. As soon as n > 3, the weighted mean square average is a more precise estimation.

3.2. Flicker Phase Noise Modulation.

In the presence of a Flicker Phase Noise Modulation, the Time variance $\sigma_x^2(\tau)$ is characterized on a bi-logarithmic diagram by a slope 0 which does not depend on the analysis period. It means that averaging over the set will not improve the precision, given by:

$$\sigma_x(\tau = (n-1) \times \tau_0) = \sigma_x(\tau_0) \quad (8)$$

3.3. White and Flicker Phase Noise Modulations.

It could happen that both phase noises would modulate the phase differences. In the usual case of the

power law model, the White Phase Noise would modulate up to an analysis period $\tau \leq \tau_1$, with $\tau_1 = p \times \tau_0$, immediately followed by the Flicker Phase Noise Modulation for the longer analysis periods $\tau \geq \tau_1$, as it is shown on the figure 3.

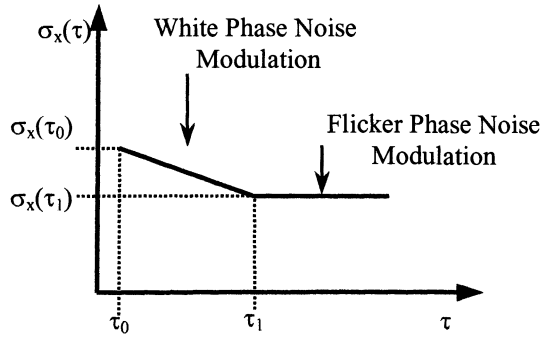


Figure 3. Time Standard Deviation in the case of a White Phase and a Flicker Phase Noise Modulation.

In this case, what is recommended is to compute the average frequency over the ensemble of phase differences by a weighted mean square linear regression. However, the computation of the variance following the equation (3) has to be limited to the analysis period over which the White Phase Noise Modulation is preponderant, which means for $\tau = \tau_1$, or in other words for a function K_p, instead of K_n in (4), given by:

$$Kp = \frac{12}{(p+1)\times p\times(p-1)} \tag{9}$$

3.4. Random Walk Phase Noise Modulation.

It happens from the definition of the instantaneous phase deviation x(t) and frequency deviation y(t) that a Random Walk Phase Modulation affecting the phase differences is equivalent to a White Frequency Noise Modulation affecting the frequency. However, a White Frequency Noise Modulation is very close to what is expected from the SHM, or even from PHARAO over an analysis period corresponding to one single pass [5]. It means that a Random Walk Phase Noise Modulation could not be averaged out if the level of noise is not lower than the anticipated clock noise. The best estimate of the average frequency is given by the simple frequency mean, affected by a noise given by the usual Two-sample or Allan variance (AVAR) computed with overlapping estimates for $\tau = m \times \tau_0$ [7]:

$$\sigma_y^2(\tau) = \frac{K_m}{\tau^2}\sum_{i=1}^{n-2m}(x(t_{i+2m}) - 2\times x(t_{i+m}) + x(t_i))^2 \tag{10}$$

where K_m is equal to:

$$K_m = \frac{1}{2\times(n-2m)} \tag{11}$$

3.5. A periodic signal in the phase deviation x(t).

If a periodic signal of period $T_0 = 1/f_0$ appears in the measurements, its signature will decrease the precision. Such a signal can be described in the instantaneous fractional frequency deviation by:

$$y(t) = a\times\sin(2\times\pi\times f_0\times t) \tag{12}$$

For a signal at a reference frequency ν, this periodic term appears in the instantaneous phase deviation as:

$$x(t) = \frac{a}{2\times\pi\times\nu}\times\sin(2\times\pi\times f_0\times t) \tag{13}$$

The Power Spectral Density of such a signal is then [8]:

$$S_x(f) = \frac{a^2}{2\times(2\times\pi\times\nu)^2}\times\delta(f - f_0) \tag{14}$$

From this, the Time variance is computed by [9, 10]:

$$\sigma_x^2(\tau) = \frac{8\times\tau_0^2}{3\times\tau^2}\times\int_0^\infty S_x(f)\times\frac{\sin^6(\pi\times\tau\times f)}{\sin^2(\pi\times\tau_0\times f)}\times df \tag{15}$$

One obtains for the Time variance: (16)

$$\sigma_x^2(\tau) = \frac{a^2}{6\times(\pi\times\nu)^2}\times\frac{\sin^6(\pi\times f_0\times\tau)}{(\pi\times f_0\times\tau)^2}\times\frac{(\pi\times f_0\times\tau_0)^2}{\sin^2(\pi\times f_0\times\tau_0)}$$

This function of the variable τ reaches its first maximum for a value u_M defined by the equation:

$$\tan(u_M) = 3\times u_M \tag{17}$$

from where comes the analysis period τ_M corresponding to this first maximum:

$$u_M = \pi\times f_0\times\tau_M \tag{18}$$

The amplitude of this first maximum is given by: (19)

$$\sigma_x^2(\tau_M) = \frac{a^2}{6\times(\pi\times\nu)^2}\times\frac{\sin^6(u_M)}{(u_M)^2}\times\frac{(\pi\times f_0\times\tau_0)^2}{\sin^2(\pi\times f_0\times\tau_0)}$$

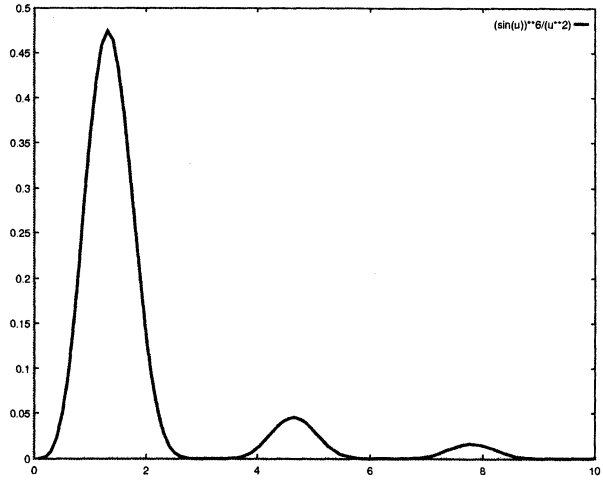


Figure 4. The plot of the function g(u) given by (20).

The variance (19) computed with the ACES parameters will show the noise limit allowed to any periodic term when compared to the requirements (1). Each following maximum appears as a bump on the Time Standard Deviation plot, for each analysis period given by $(2k+1)T_0/2$, but decreasing fast in amplitude as the analysis period increases. In the figure 4 is proposed

for u varying from 0 to 10 the plot of the function g(u) given by:

$$g(u) = \frac{\sin^6(u)}{u^2} \qquad (20)$$

One can see that the second maximum is already ten times lower than the first one. A graphical resolution gives as first maximum: $u_M = 1.3242$.

4. HOW TO DETERMINE THE NOISE MODULATING THE SHORT TERM DATA

All the developments made in the chapter 3 are based on the preliminary knowledge of the noise modulation. But when computing an Allan variance, the last possible term is for an analysis period roughly equal to half the interval duration (see (10)). For determining the preponderant noise modulation over the whole Short Term Data set, it is possible to join a few consecutive sessions together, and to take into account the effect of the dead times through Bias Functions. The use of the Total Time Variance (Total TVAR) to improve the confidence intervals on the computed values could help to discriminate between the noise modulations.

4.1. The Bias Functions.

Allan and Barnes have proposed to use three Bias Functions [11]. The two-sample variance $\sigma^2(2, p, T, \tau)$ of p consecutive grouped sessions, separated by dead times $T - \tau$, can be corrected with the dead times collected at the end of the time interval, by:

$$\sigma_y{}^2(\tau) = \frac{\sigma^2(2,p,T,\tau)}{B_2(r,\mu)\times B_3(2,p,r,\mu)} \qquad (21)$$

where $r = T / \tau$, and μ is the exponent of the power law model τ^μ of the variance. To determine μ, an other Bias Function based on the ratio between the N-sample variance and the two-sample variance, is given by:

$$B_1(N,r,\mu) = \frac{\sigma^2(N,T,\tau)}{\sigma^2(2,T,\tau)} \qquad (22)$$

The Bias Functions B_1, B_2 and B_3 are tabulated in [11] for the usual noise modulations of the power-law model.

Vanier and Audouin, quoting Lesage, have also proposed one Bias Function [12] which allows to relate the Allan Deviation of p consecutive sessions $\sigma_{y,p}(p \times \tau)$, separated by dead times $T - \tau$, to the Allan Deviation of the joined consecutive sessions with all the dead times collected to the end of the time interval $\sigma_y(p \times \tau)$:

$$B(p,T,\tau) = \frac{\sigma_y(p\times\tau)}{\sigma_{y,p}(p\times\tau)} \qquad (23)$$

Note that this definition is based on standard deviations, and not on variances. B(p, T, τ) is tabulated in [12] for the usual noise modulations of the power-law model.

4.2. The Total Time Variance.

The Total variance approach has been developed for increasing the confidence of the estimation of the AVAR, particularly for large analysis periods [13]. This method is based on an extension of the original data sequence called the mirror-reflection, which increases the equivalent degrees of freedom (edf) of each AVAR estimate. However, Total AVAR is not suitable for the Phase Noise Modulations analysis.

This approach was generalized to the Time variance, the Total TVAR [14] and to the Modified Allan variance, the Total MVAR [15]. These variances are specially designed for estimating the level of the Phase Noise Modulations.

By using Total TVAR, the edf of the estimates near the longest analysis periods, limited to half the total duration of the data sequence (see (10)), increases by a factor between 2 and 4 with respect to the classical TVAR. Since the edf are defined as :

$$edf = \frac{2\times Mean[estimateTVAR]}{Variance[estimateTVAR]} \qquad (24)$$

the confidence interval over each Total variance estimate is reduced from 70% to 50% relatively to the classical variance estimate. Thus, the noise levels are more precisely determined and the use of Total TVAR over a sequence of duration T is equivalent to the use of classical TVAR over a sequence of duration 2T.

5. MIDDLE TERM DATA PROCESSING

The processing of the Middle Term Data will differ depending if the noise of the link would be averaged out or not at the end of the Short Term Data processing.

5.1. The noise of the link is already averaged out.

In this case, the computation of the Short Term Data leads to the average clock frequency difference, which is anticipated in ACES to exhibit a White Frequency Noise Modulation. Therefore, the influence of the dead times or of the sampling can be completely neglected in the computation of the related Allan Variance. The Middle Term Data processing is reduced to a simple mean of the average frequency differences.

5.2. The noise of the link cannot be averaged out.

In the case the MWL would not satisfy the short term requirements (1) or (2), there will still remain some Phase Noise Modulation in the average short term frequency differences. The time distribution of the measurements could lead to a T_{ISS}-sampling aliasing effect on the Power Spectral Density, written as:

$$S_{y[MT]}(f) = \frac{1}{T_{ISS}} \sum_{n=-\infty}^{+\infty} S_y\left(f + \frac{n}{T_{ISS}}\right) \qquad (25)$$

where $S_y(f)$ is the Power Spectral Density of the measurements, and $S_{y[MT]}(f)$ is the Power Spectral Density of the Middle Term Data. The high cut-off frequency for this signal is $1 / d_{ISS}$.

As it is well known, such a T-sampling effect may, for analysis periods long enough, which means for frequencies low enough, lead to a constant value depending on d_{ISS} and T_{ISS}. In other words, the sampling may lead to a White Frequency Noise Modulation in the average frequency differences. What remains to be computed is the level of this Noise Modulation. This should lead to a comprehensive study of the Long Term frequency comparisons between the ACES clock and any Earth clock.

6. CONCLUSION

The outstanding requirements of ACES, in direct relationship with the ISS low orbit, are calling for a specific fine processing of the future measurements. Some of the numerous options are presented here.

What is already clear is that the Short Term Data processing will be based on the knowledge of the preponderant Phase Noise Modulation. Assuming that, the average frequency difference will be most of the time given by a weighted mean square linear regression. The associated Standard deviation will however look different depending on the noise modulation. The detailed computation includes the limits produced by a periodic term in the phase measurements.

The potential complexity of the Middle Term Data processing will rely on the results of the Short Term Data processing. Either the noise of the MWL can be averaged out, and the processing is reduced to a simple mean of the frequency differences affected by a White Frequency Noise Modulation, or the noise of the MWL cannot be averaged out. In this case, what could be anticipated is that the dead times, much larger than the measurement session durations, would produce aliasing in the Power Spectral Density. This still remains to be investigated more deeply.

The Long Term Data processing will merely be considered as a new sampling of frequency differences. This too requires some more investigations before the start of the ACES flight scheduled today for 2005.

REFERENCES

[1] C. Salomon, C. Veillet, Proc. of the 1st ESA symposium on Space Station Utilisation, SP385, 295 (1996).

[2] "The SHM-ACES Hydrogen Maser Space Clock : A high performance time and frequency standard for the ACES experiment on the ISS", Observatoire de Neuchâtel, ON/SHMA-001/P-1, Issue 1, 11 April 97.

[3] P. Laurent, P. Lemonde, E. Simon, G. Santarelli, A. Clairon, N. Dimarcq, P. Petit, C. Audoin, and C. Salomon, "A cold atom clock in absence of gravity", Eur. Phys. J. D 3, pp. 201-204 (1998).

[4] E. Samain, P. Fridelance, "Time Transfer by Laser Link (T2L2) experiment on Mir", Metrologia, 1998, 35, pp. 151-159.

[5] P. Uhrich, P. Guillemot, P. Aubry, F. Gonzalez, C. Salomon, "ACES Microwave Link Requirements", Proc. of the IEEE, Special Issue on the 1999 Annual Symp. on Freq. Control, accepted (in press).

[6] D.W. Allan, "Tutorial: Clocks and Clock Systems Performance Measures", Proc. of the 27th Annual PTTI Meeting, 1995, pp. 209-233.

[7] D.B. Sullivan, D.W. Allan, D.A. Howe, F.L. Walls (Edited by), Characterization of Clocks and Oscillators, NIST Technical Note 1337, March 1990, p. TN-164.

[8] A. Papoulis, Probability, Random Variables, and Stochastic Processes, McGraw-Hill, Second Edition, 1984, pp. 265 - 271.

[9] D.B. Sullivan, D.W. Allan, D.A. Howe, F.L. Walls (Edited by), Characterization of Clocks and Oscillators, NIST Technical Note 1337, March 1990, p. TN-9.

[10] D.W. Allan, M.A. Weiss, J.L. Jespersen, "A Frequency - Domain View of Time - Domain Characterization of Clocks and Time and Frequency Distribution System", Proc. of the 45th Annual Symp. on Frequ. Contr., May 1991, pp 667-678.

[11] J.A. Barnes, D.W. Allan, "Variances based on Data with Dead Time between the Measurements", NIST Technical Note 1318, 1990 (reproduced inside NIST Technical Note 1337, March 1990, pp. 296 – 335).

[12] J. Vanier, C. Audouin, The Quantum Physics of Atomic Frequency Standards, Adam Hilger, 1989, Vol. 1, Appendix 2F, pp. 233 – 245.

[13] D.A. Howe, C.A. Greenhall, "Total Variance: a Progress Report on a New Frequency Stability Characterization", Proc. of the 29th Annual PTTI Meeting, 1997, pp. 39-48.

[14] M.A. Weiss, D.A. Howe, "Total TDEV", Proc. of IEEE International Frequency Control Symposium, 1998, pp. 192-198.

[15] D.A. Howe, F. Vernotte, "Generalization of the Total variance approach to the modified Allan variance", Proc. of the 31st Annual PTTI Meeting, 1999 (in press).

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

ACCURATE DIVISION OF OPTICAL FREQUENCIES BY CW OPTICAL PARAMETRIC OSCILLATORS : APPLICATION TO THE DESIGN OF NEW OPTICAL SYNTHESIZERS

A. DOUILLET, J.-J. ZONDY, G. SANTARELLI, A. MAKDISSI, A. CLAIRON
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)
albane.douillet @obspm.fr

ABSTRACT

We report on an accurate phase-locked 3:1 division of a 843.06 *nm* optical frequency. The divider is achieved by using a continuous-wave, diode laser pumped, doubly resonant optical parametric oscillators (DRO) ($3\omega \rightarrow 2\omega+\omega$), based on periodically poled lithium niobate (PPLN). The idler output (λ_i=2.53 μm) from the DRO is externally frequency doubled and the resulting beat note with the signal wave (λ_p=1.26 μm), used to phase-lock the divider. A fractional frequency stability of 2×10^{-17} of the division process has been achieved for 100 *s* integration time. The technique developed in this work can be generalized to the accurate phase and frequency control of any cw optical parametric oscillators.

INTRODUCTION

In the past, optical frequency measurement were performed with harmonic frequency chains, which create successive harmonics from a primary frequency standard. A new generation of optical frequency measurement chains, much more flexible is under development. It combines the measurement of frequency intervals and the use of broad frequency combs, provided, for example by a ϕ stabilized high repetition rate femtosecond laser [1]. Recently, a difference frequency in the near infra-red, with a frequency gap as large as 100 *THz* has been bridged up with a wide-span femtosecond Ti:Sa laser comb generator [2]. Provided we are able to measure, in the optical spectrum, an interval between ω_{st} and $\omega_{st}+\omega_{st}/n$ (obtained by sum frequency generation in a non linear crystal) by an optical comb generator, we also measure ω_{st}. Thus, the use of an optical division by n can considerably reduce the span required by the comb generator, and at the same time, reduces the ϕ noise requirement of the comb generator by n^2.

Continuous-wave optical parametric oscillators, are promising optical frequency dividers. Divide-by-two cw OPO's have been demonstrated [3,4]. For by-three division based on OPOs there are two methods. The first one makes use of an all-optical self phase-locking of the subharmonics to the exact 3:1 frequency ratio, by inducing their mutual injection locking via cascaded OPO/second harmonic generation interactions [5]. In the second method, the idler wave must be externally frequency doubled. Then the RF beat between the signal and the doubled-idler waves is used to phase lock the device to zero or to an RF reference frequency [6,7]. The advantage of these OPO-based dividers is that they require only a single pump laser source to produce the subharmonics. Furthermore, the coherent nature of the division process leads to subharmonic linewidths even narrower than the pump one. They are ultimately limited by the phase diffusion noise stemming from the spontaneous parametric fluorescence. We report here on an accurate phase-locked optical frequency divider of a diode laser operating in the range 840-850 *nm*. We show that even with a weak beat signal level (-50dBm typically) and without a fast electro-optic cavity length actuator, it is possible to achieve a very high fractional stability and long-term operation of the divider.

1. EXPERIMENTAL SET-UP

The doubly resonant optical parametric oscillator (DRO) is pumped by a master-oscillator power amplifier (MOPA) AlGaAs diode laser which is optically injected by a home-made AlGaAs extended-cavity diode laser in the Littrow configuration. The pump wavelength is λ_p=843.06 *nm*. Its frequency, monitored by a 500 *MHz* resolution wavemeter, can be tuned over ~1.5 *GHz*. The available power at the DRO input is 0.4 W. The nearly spherical DRO cavity consists of two highly reflective (at the signal and idler wavelengths) ZnSe mirrors. Their radius-of-curvature is R=50 mm and the cavity length L=106 mm. The cavity length can be tuned using a 20-mm long piezoelectric transducer (PZT), separated into two sections (3/4 and 1/4 of the length) that can be independently driven. A multi-grating periodically poled lithium niobate (OPO-PPLN) is the non linear crystal used to performed the $3\omega \rightarrow (2\omega\pm\delta)+(\omega\mp\delta)$ interaction. The δ quantity ($\delta << \omega$) represents the small radio-frequency mismatch from perfect division ratio. The 19-*mm* long (aperture 0.5×11 mm^2) PPLN crystal has eight different poling periods (Λ) ranging from Λ=22.4 μm to 23.1 μm. The experimental set-up is sketched Fig. 1. The oscillations threshold is 15 *mW*,

which is comparable of the one reported by Lindsay *et al.* [8] for similar wavelengths. To increase the output coupling efficiency, limited to 200 μW of idler output power, we have inserted inside the DRO cavity a CaF_2 Brewster plate. This yields an oscillation threshold increase, up to 65 *mW*. The available idler power to generate the second harmonic (SHG) of this wave is 3 *mW*, and the signal power 6 *mW*. We perform the SHG, outside the OPO cavity, in a second periodically poled lithium niobate sample (SHG-PPLN), having two different grating periods (34 μm and 35 μm). For both OPO and SHG interactions, the three waves are extraordinary polarized. The signal ($2\omega\pm\delta$) and the doubled-idler ($2\omega\mp2\delta$) waves are mixed on a 7- *GHz* bandwidth avalanche photodiode. The resulting 3δ beat frequency signal is used to control the division ratio of the DRO.

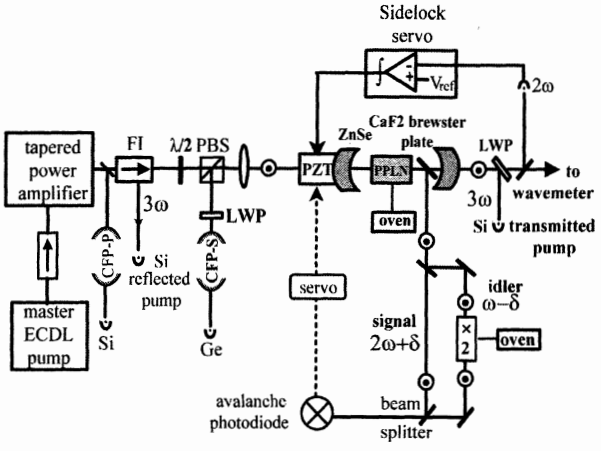


Fig. 1. Type I [$(3\omega\rightarrow(2\omega\pm\delta)+(\omega\mp\delta),\delta\ll\omega$)] phase-locked DRO experimental set-up. FI : -40dB Faraday isolator, PBS : polarizing beamsplitter, LWP : long wave pass filter, CFP : confocal Fabry-Pérot. The ◉ symbol refers to the beam polarization. The three interacting waves are extraordinary polarized (in the vertital plane of the figure). The fast photodiode detects the 3δ frequency beatnote between the signal ($2\omega\pm\delta$) and the doubled idler ($2\omega\mp2\delta$). The CaF_2 plate is almost at Brewster angle (0.1% reflection). The operating temperature is 105°*C* with the Λ=22.9 μm grating for the OPO-PPLN, and 70°*C* with the Λ=35 μm grating for the SHG-PPLN.

2. EXPERIMENTAL DIFFICULTIES

2.1 Mode pair selection and beatnote detection

Since the thickness of the OPO-PPLN is only 500 μm, to overcome diffraction losses, the OPO cavity is a nearly spherical resonator. Furthermore, the dispersion of the extraordinary index of refraction of lithium niobate from 2.53 μm to 1.26 μm is weak. This yields a signal and idler free spectral ranges of about 1.15 *GHz*, differing by only 1 or 2 *MHz*. Thus, this oscillator behaves as a type-I degenerate DRO. Frequent axial mode hops (a few *GHz* away) and cluster hops (a few tens of *GHz*) occur. The selection of the appropriate mode-pair (ie those giving a 3δ frequency beat within the photodiode bandwidth) is rather difficult to achieve since more than 700 signal-idler mode pairs can experience similar gain within a cluster, for this cavity configuration.

A standard side-of-fringe (*sidelock*) servo is first implemented. The error signal of this servo loop is built by comparing the signal wave power detected by an InAs photodiode to a reference voltage. It is then integrated and feeds back to the largest section (3/4 of the length) of the PZT transducer. The sidelock servo acts as a hybrid frequency/intensity servo : under negligible pump frequency fluctuations, it maintains a constant cavity frequency detuning of the signal wave, by correcting OPO cavity length perturbations, while under pump intensity fluctuations, it updates the signal detuning so as to keep the intracavity power constant.

Thanks to this sidelock servo, the DRO fortunately oscillates on a single mode pair during 5-15 minutes, which leaves enough time to proceed to the phase-locking. The mode pair selection is the first experimental difficulty we have to overcome. Fine tuning of the OPO-PPLN temperature and of the sidelock reference voltage have to be repetitively processed until a beat signal in the frequency range 200 *MHz*-5 *GHz* is detected. This beat note is mixed with a microwave syntesizer and down converted to 100 *MHz*.

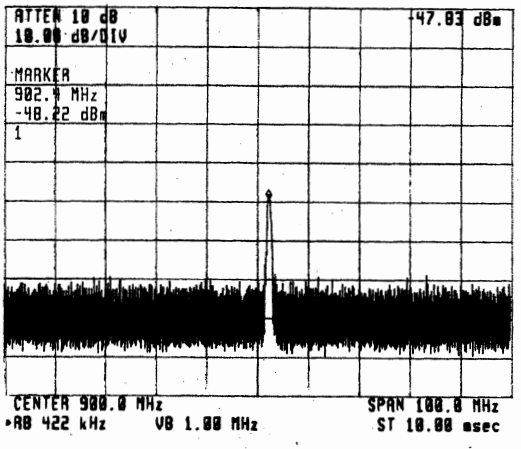


Fig. 2. The 3δ free running beatnote between the doubled-idler ($2\omega-2\delta$) and the signal ($2\omega+\delta$) at 902.4MHz with a 100MHz span and a 422kHz resolution bandwith. The center is at 900MHz. The vertical scale is 10dB/div.

A typical detected beat note is shown on Fig. 2. Its signal-to-noise ratio is 30 *dB* in a 100 *kHz* resolution bandwidth of the spectrum analyzer. The beat signal frequency drift is less than 1 *MHz* per minute and its jitter, mainly caused by acoustic perturbations, less than 400 *kHz*.

2.2 Beat note regeneration by use of a tracking VCO

The second experimental difficulty is a low beat note signal-to-noise ratio which prevents a direct accurate frequency counting. To solve it, a tracking voltage-controlled oscillator (VCO) is phase locked to the beat note. This servo loop is denoted the *tracking loop*. Since the beat note amplitude noise is rejected and its phase noise filtered by this loop bandwidth, the signal-to-noise ratio of the beat signal at the VCO output is much higher than that of the input beat signal. Thus, division and counting without any false triggering of this beat note is practicable. A 1.15 *MHz* bandwidth is achieved for this loop.

2.3 Phase lock bandwidth limitation

The third experimental difficulty, and the less easy to solve, is to replace the sidelock servo by a phase lock, using the single PZT transducer, given the quite different nature of the information carried by their respective error signals. Furthermore, the mechanical resonance of the PZT will ultimately limit the loop bandwidth to a few *kHz*. Despite all these intensive difficulties, we shall demonstrate in the next section that is possible to control the relative phase and frequency of the DRO output waves to a high stability level, by use of an intricated electronic phase/frequency loop.

3. THE PHASE-COHERENT 3:1 FREQUENCY DIVIDER

The phase-coherent 3:1 frequency divider is performed in two steps. First, we control the beat note frequency, and then its phase. Fig. 3 represents the whole developed phase-locking set-up, made of 5 overlapping loops

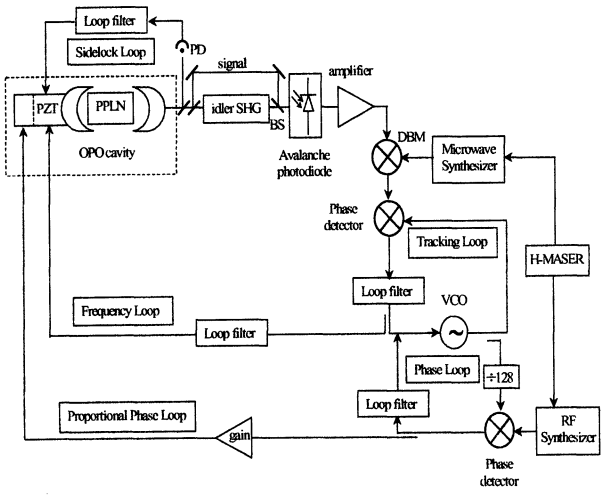


Fig. 3. The whole phase-lock loop. PPLN, periodically poled lithium niobate; PZT, piezoelectric transducer; SHG, second harmonic generation; BS, beam splitter; DBM, double-balanced mixer; VCO, 100 *MHz* voltage-controlled oscillator; PD, InAs photodiode.

3.1 Frequency locking the divider

In first step toward the phase-locked divider, we pre-stabilize the beat note, by reducing its frequency fluctuations. This is achieved using a *frequency lock*, which will take over the *sidelock*. The varicap voltage of the VCO, which carries the beat frequency fluctuations, is summed to the *sidelock* error signal. The gain of the frequency loop is progressively increased, and when this loop is stable enough, the *sidelock* loop is disabled, thus, avoiding a long term conflict between these two servo loops acting on the same physical quantity (cavity length). The DRO tendency to mode hop is then reduced.

In Fig. 4 we have plotted the Allan standard deviation deviation $\sigma_y(\tau)$ of the DRO, under *sidelock* (upper curve) and frequency lock (lower curve). The quantity y is the ratio between the mean value of δ, and the idler frequency ν_i=118.6 *THz*. The integration time is τ.

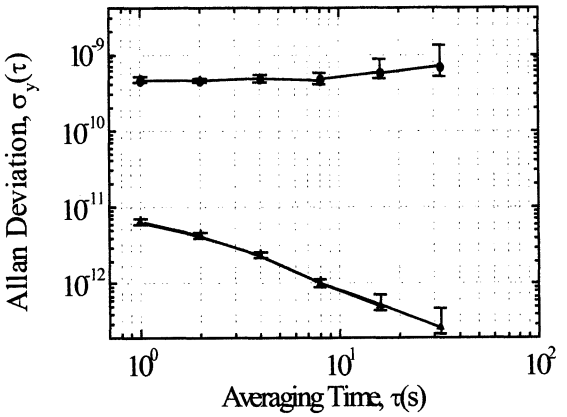


Fig. 4. Allan standard deviation of the 3:1 divider frequency noise, relative to the idler frequency under *sidelock* (circles) and *frequency lock* (triangles).

Under frequency lock, the beat fluctuations are reduced to the *kHz* level, in comparison to a few hundreds of *kHz* under sidelock.

3.2 Phase-locking the divider

In a second step, the phase control of the divider is performed. The 100 *MHz* signal at the VCO output is divided by 128, using a digital division circuit. A triangular response phase detector, compares the phase of this divided beat note with a 781 *kHz* reference synthesizer. In this set-up, each synthesizer and the frequency counter are synchronized to an H-maser, whose frequency stability is 10^{-15}, for one day integration time. By-128 division of the beat note, yields a reduced phase lock loop bandwidth, since the phase error signal has an 128π peak-to-peak dynamic range.

We have first unsuccessfully tried to phase lock the OPO divider by directly acting on the PZT transducer. Actually, the same transducer is used to correct the derivative of the phase (information obtained via the frequency loop) and the phase

difference (information from the phase loop). Signs of the two corrections are not necessarily the same, and so, both information conflict.

To implement a stable phase-locking, the correction signal must be applied to the VCO varicap voltage. This is the phase loop sketched in Fig. 3. The phase correction signal is built by filtering the phase error signal in a first order loop filter. As the frequency loop and the phase loop are overlapping via the VCO, gains of these both locks are optimized by minimizing, at low frequencies, the ϕ noise level, measured using a FFT spectrum analyzer.

In a final step, a proportional phase loop is added. A small part of the phase error signal is proportionally amplified and directly feeds back the small part of PZT (1/4 of the length). The goal of this last loop is to increase the bandwidth of the whole phase locking set-up. However, as the gain of this loop needs to remain small to avoid conflicts with the frequency loop (both servos act on the same physical quantity), it reduces only a little bit the phase noise.

This phase-frequency lock loop is extremely robust and can tolerate very large frequency jumps. When the phase goes out of lock, the frequency lock shortens 100 times the recapture delay of the phase loop. The device can remain phase-locked for up to one hour, and the 3δ beat frequency is stabilized to the *mHz* level for 100s integration time. Fig. 5 represents the excess phase noise of the OPO divider, compared to an ideal divider. A perfect by-n divider reduces the input signal phase noise spectral density by a factor n^2. This spectrum shows the numerous mechanical frequency resonances of the device.

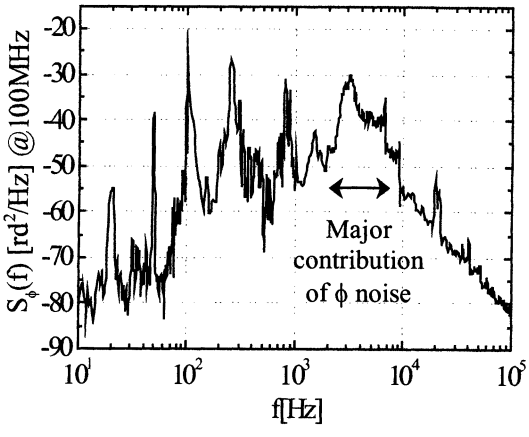


Fig. 5. Phase noise spectrum of the phase-locked divider, measured using a FFT spectrum analyzer for the δ frequency beat note.

To observe a coherent beat signal, the total phase noise has to be less than 1 rd^2. Because of the limited bandwidth due to the PZT mechanical resonance, we can not increase enough the phase lock loop gain to observe a coherent beat signal We deduced from Fig. 5, from a numerical integration, a 1.2 rd^2 total phase noise at the idler frequency (30% of the energy in a narrow coherent peak). Without the by-128 division, the residual phase noise exceeds the phase detector dynamic range. Phase-locking the device is then impossible. The achieved phase lock bandwidth is about 10 *kHz*. The obtained 3δ beat note linewidth is ~3 *kHz*.

The corresponding measured Allan standard deviation of the 3:1 optical divider noise, is plotted in Fig. 6. It is $\sigma_y(\tau)=2\times10^{-15}/\tau$. This is the expected white phase noise behavior which demonstrates that no cycle slip occurs during the measurement.

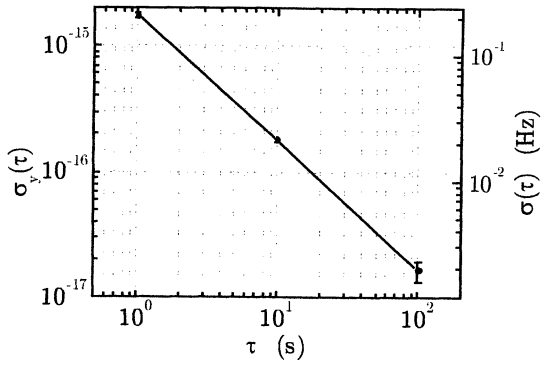


Fig. 6. Left axe : Allan standard deviation of the 3:1 divider frequency noise, measured at the idler frequency, as a function of the integration time. Right axe : absolute Allan standard deviation.

3.4 Model of the phase-lock loop

This experimental phase locking loop has been modelized. The transfer function of the whole closed loop is expressed in equation 1.

$$\Phi_{DIVISEUR}(s) = H_{REF}(s)\,\Phi_{REF} + H_{OPO}(s)\,\Phi_{freeOPO}(s) + H_{VCO}(s)\,\Phi_{freeVCO}(s) \qquad (1)$$

The phase noise of the locked divider ($\Phi_{DIVISEUR}(s)$) depends of the phase of the free beatnote ($\Phi_{freeOPO}(s)$), of the free VCO ($\Phi_{freeVCO}(s)$) and of the reference synthesizer ($\Phi_{REF}(s)$). The quantity s is the defined by $s = j2\pi f$, where f is the Fourier frequency.

The fifth order transfer functions $H_{OPO}(s)$ and $H_{VCO}(s)$, respectively describe the divider behavior, with noise source stemming from the OPO and the VCO. With such a phase-frequency servo, all types of noise are eliminated, even high amplitude noises, the frequency lock avoiding large phase jumps.

The Bode diagram of the fifth order transfer function $H_{REF}(s)$ is plotted in Fig.7. It represents the

relation between the reference fluctuation (input) and the locked divider (output). For high frequencies ($s \rightarrow \infty$), $H_{REF}(s) \rightarrow 0$. Fast reference fluctuations are not corrected by this loop. Because of the division by 128, the transfer function value for low frequencies is 42dB (20log(128)) instead of 0dB. The locked divider phase follows, for long time operation, the reference phase.

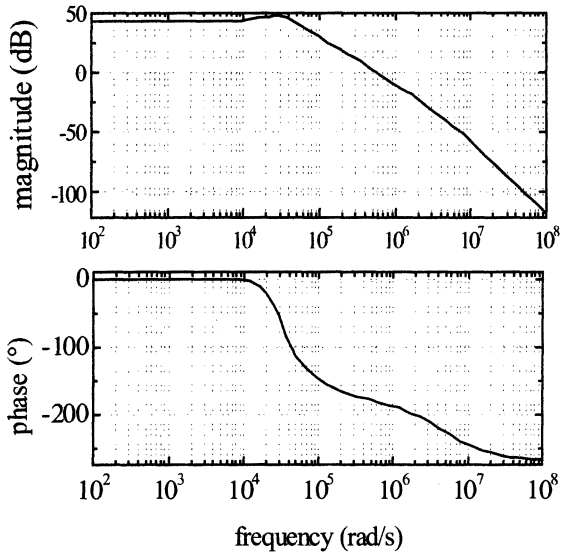


Fig. 7. Bode diagram of the transfer function between the phase of the reference synthesizer (input) and the phase of the phase-locked beatnote (output).

5. CONCLUSION

We have realized an optical frequency divide-by-3 of 843 *nm* light, using a doubly resonant optical parametric oscillator. A relative optical divider frequency stability of 2×10^{-17} for 100 *s* integration time, at the idler wave is achieved, which highlights the potential resolution of parametric dividers. A shorter OPO cavity, allowed by a thicker PPLN chip, would solve the inconvenience linked to the delicate mode selectivity and make easier the beatnote selection. To further increase the stability of this 3:1 divider, it is possible to insert an electro-optic modulator in the DRO cavity as a second transducer, which would overcome bandwidths limitations associated to the PZT mechanical resonances. This PPLN-DRO can also generates a comb of frequencies ($\omega_p/3$, $2\omega_p/3$, ω_p, $4/3\omega_p$, $5\omega_p/3$, $2\omega_p$) from various non-phase matched or higher-order quasi-phasematched up-and-down conversion mixings that occur within the chip. Measuring any interval between these sidebands leads to a simple detection of all frequencies of the comb. Compared with conventional methods, it could reduce by one third the required span of high repetition rate femtosecond laser.

REFERENCES

[1] Th. Udem, J. Reichert, R. Holzwarth, and T.W. Hänsch, "Accurate measurement of large optical frequency differences with a mode-locked laser," Opt. Lett. **24**, 881-883 (1999)

[2] S. A. Diddams, D. J. Jones, L-S Ma,S. T. Cundiff, and J. L. Hall, "Optical frequency measurement across a 104 THz gap using a femtosecond laser frequency comb," Opt. Lett. **25**, 186-188 (2000)

[3] D. Lee and N. C. Wong, " Tunable optical frequency division using a phase-locked optical parametric oscillator", Opt. Lett., 17, 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., 127, 69-72 (1996).

[5] D.-H. Lee, M. E. Klein, J.-P. Meyn, P. Groß, S.Marzenell,R. Wallenstein, and K.-J. Boler, "Self-injection-locking of a CW-OPO by intracavity frequency-doubling the idler wave", Opt. Express, 5, 114-119 (1999).

[6] S. Slyusarev, T. Ikegami, S. Ohshima, "Phase-coherent optical frequency division by three of 532-nm laser light with a continuous-wave optical parametric oscillator", Opt. Lett., 24, 1856-1858 (1999).

[7] A. Douillet, J.- J. Zondy, A. Yelisseyev,S. Lobanov and L. Isaenko, "Toward a 3:1 frequency divider based on parametric oscillations using $AgGaS_2$ and PPLN crystals", to be published in the Special Issue on Frequency Control and Precision Timing of the IEEE trans. UFFC

[8] 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., 23, 1889-1891 (1998).

[9] 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", J. Opt. Am. B, 16, 1481-1498, (1999)

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

DEVELOPMENT OF A FREQUENCY SYNTHESIS CHAIN FROM 29.1 THz TO 563.2 THz : APPLICATION TO THE ABSOLUTE FREQUENCY MEASUREMENT OF $^{127}I_2$-STABILIZED Nd:YAG LASER AT 563.2 THz/532 nm

F. Ducos, Y. Hadjar, G.D. Rovera and O. Acef

BNM-LPTF/CNRS-UMR 8630/Observatoire de Paris,
61 Avenue de l'Observatoire, F-75014 Paris (France)
e-mail : Franck.Ducos@obspm.fr

ABSTRACT

We propose a new method to measure the absolute frequency of $^{127}I_2$-stabilized Nd:YAG laser at 563.2 THz/532 nm based on the frequency combination of two well known optical frequency standards CO_2/OsO_4 at 29.1 THz/10.3 μm and LD/Rb at 385.3 THz/778 nm, previously measured against primary cesium clock at the BNM-LPTF [1, 2]. We use two intermediate laser diodes operating in the near-infrared at 281 THz/1.067 μm (LD1) and 192.6 THz/1.556 μm (LD2) linked to the previous OFS. In this paper, we report our preliminary results.

1. INTRODUCTION

Several atomic and molecular radiations recommended by the Consultative Committee of Length (CCL) since the 17th Conference Générale des Poids et Mesures, have been measured with respect to a well known OsO_4 frequency reference used to stabilize the CO_2 laser in the 10 μm spectral range: CH_4 at 3.39 μm [3], Rb at 778 nm [2], I_2 at 633 nm [4] and H at 389 nm [5]. We propose to measure the absolute frequency of the R(56), 32-0, a_{10} iodine transition at 532 nm, which is commonly used for the stabilization of frequency doubled Nd:YAG lasers. This radiation is also recommended by the Comité International des Poids et Mesures (CIPM) for the *mise en pratique* of the definition of the meter. This particular line is very promising as an optical frequency standard owing to its intrinsic frequency stability and reproducibility [6]. Nevertheless, its absolute frequency is known with an uncertainty of 7×10^{-11} as reported in [7].

Our measurement involves two OFS. The first operates at 29.1 THz/10.3 μm and is based on a sealed CO_2 laser, frequency stabilized on an OsO_4 molecular transition known with 1.3×10^{-13} uncertainty [1]. The second at 385.3 THz/778 nm uses a laser diode frequency locked on a two photon transition in rubidium which is known with 2.5×10^{-12} uncertainty [8]. The later will determine the ultimate accuracy of the measurement of the OFS at 563.2 THz/532 nm. Thus, we expect to improve this accuracy by more than one order of magnitude with respect to the previous measurements.

2. EXPERIMENTAL SET-UP

The frequency synthesis chain set-up is shown in Fig. 1. We employ two compact laser diode sources. The first one is an α-DFB (angled-grating distributed feedback) laser diode (tunable from 1.063 to 1.067 μm) operating in this case at 1.067 μm. It delivers up to 500 mW cw output power. The second is a laser diode operating at 1.556 μm with an output power of 20 mW. This later is frequency doubled using a Periodically Poled Lithium Niobate (PPLN) crystal in order to phase lock its frequency on the LD/Rb OFS, which uses the $5S_{1/2}, F_g=3 - 5D_{5/2}, F_e=5$ two-photon transition in ^{85}Rb at 778 nm as frequency reference.

On the other hand, the LD1 laser is phase locked on the $^{127}I_2$-stabilized Nd:YAG laser standard (1.064 μm). The 611 GHz frequency gap between LD1 and the Nd:YAG reference, is bridged using an harmonic mixing in a Schottky diode with a far infrared laser itself phase-locked against H-maser. The final step consisting of a difference frequency generation (DFG) in a $AgGaS_2$ crystal will be achieved in a very near future. The theoretical generated power is reported further in this paper. The issued radiation from this nonlinear mixing at 3.39 μm will be compared to the light delivered by a 2 m long He-Ne transfer laser. This laser is linked to the third harmonic of OsO_4 standard using a MIM diode [9]. Finally, the frequency in the green will be determined as

$$\nu_{I2} = \nu_{Rb} + 6 \times \nu_{OsO4} + 2 \times \nu_{FIR} + \delta$$

where δ is the algebraic sum of all frequency beat notes involved in our chain.

3. PRELIMINARY RESULTS

3.1 Bridging the 611 GHz frequency gap

The two laser beams issued from LD1 and Nd:YAG laser (FWHM ~ 5 kHz) are carefully superimposed with the same polarization before focusing into the point contact Schottky diode [9]. The total power reaching the Schottky diode was 65 mW. An optical fiber was used in order to ensure a good mode-matching between the two laser beams illuminating the Schottky diode. The resulting beat note between the two laser sources exhibits a linewidth of less than 50 kHz

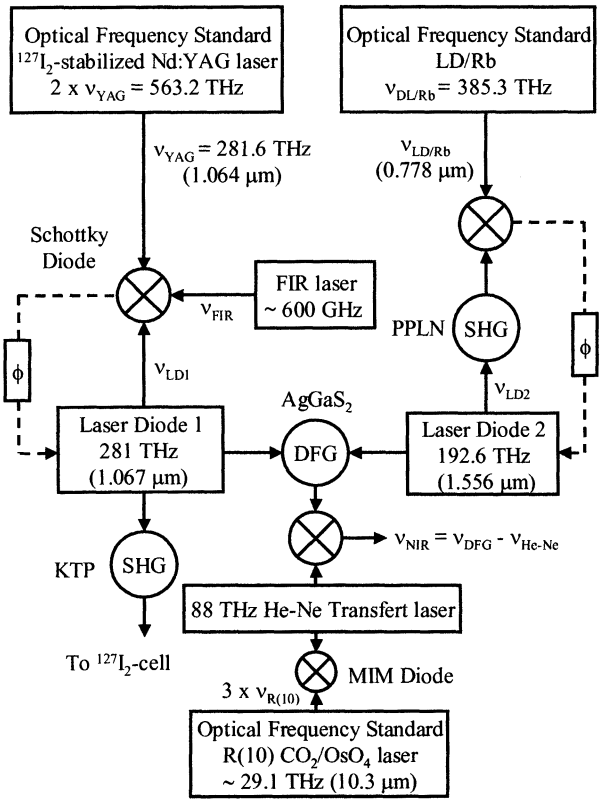


Fig. 1 Frequency synthesis chain set-up

In order to test the harmonic mixing behavior of the Schottky diode in the 1.06 μm range, the LD1 frequency was varied. As a preliminary result, we have been able to observe beat notes corresponding to optical frequency differences with harmonics of a 10 GHz radiation. We report in Fig. 2 the behavior of the Schottky diode as a function of the lasers frequency difference for three harmonics of the microwave radiation (k=1, 2, 3). The point k=0 corresponds to SNR of the beat note between the two lasers with the RF synthesizer off.

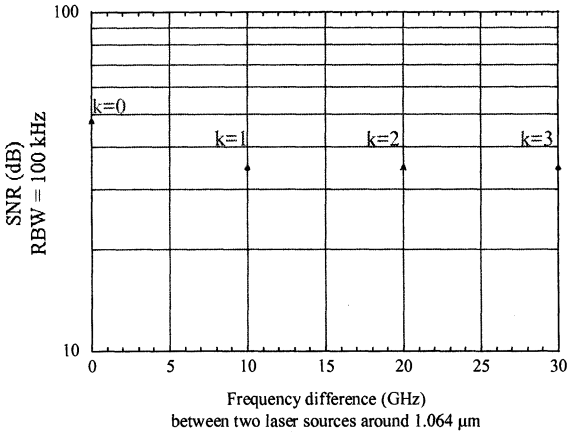


Fig. 2 SNR of beat note between the reference laser and DL1 for different harmonics of 10 GHz/15 dBm radiation versus frequency difference between lasers.

The response of the Schottky diode at 1.064 μm (1 μA/mW) was found nearly the same to the one observed at 850 nm (3 μA/mW) [9]. In this range, a large frequency difference (up to 1 THz) was already observed. These results show the feasibility of the measurement at 611 GHz. However, this mixing needs a operation temperature of LD1 (1.067 μm) around 45 °C. In this situation, the lifetime of the laser diode is strongly reduced. So that, we will use the Schottky diode only when all other parts of the chain will be ready. On the other hand, in order to overcome this high temperature operation, we plan to develop an optical frequency comb generator [10, 11] around 1.55 μm, in order to bridge the 611 GHz frequency gap.

3.2 SHG at 1.064 μm in KTP at room temperature

An interesting aspect of the laser diode (LD1) is its large wavelength tunability from 1.063 to 1.067 μm associated to its narrow linewidth and to the high output laser power (up to 500 mW cw). It can be usefully used to observe the iodine transitions over 2 THz in the green region with a high level of resolution. This frequency span is inaccessible with conventional frequency doubled Nd:YAG lasers. For these reasons, we have developed a setup for the frequency doubling as shown in Fig. 3.

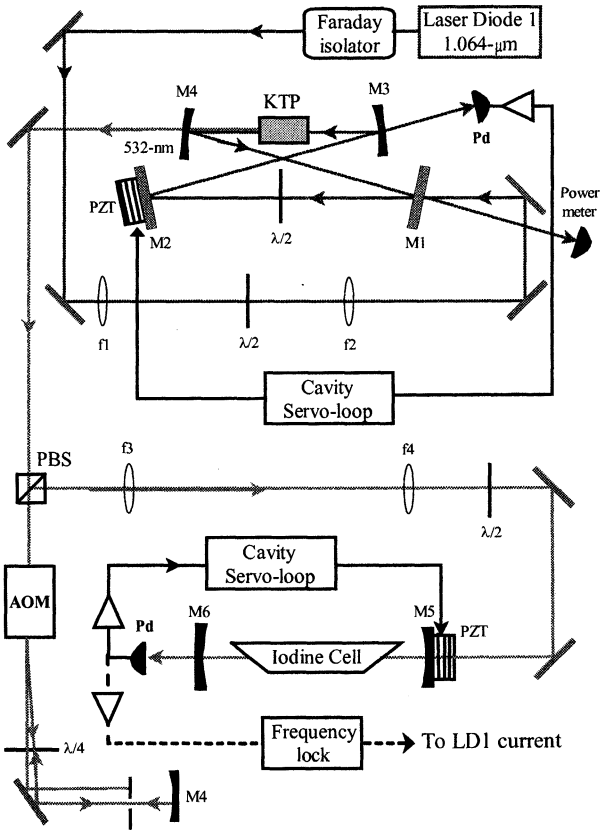


Fig. 3 Experimental set-up for the frequency doubling with a KTP crystal at room temperature in a ring cavity. The harmonic light at 532 nm is sent into an iodine cell placed in a Fabry-Perot cavity.

The LD1 laser is isolated from feedback using two successive 45 dB-Faraday isolators. The beam is then mode-matched to the fundamental TEM_{oo} mode of the doubling cavity, by means of two lenses. Our doubling cavity is a conventional ring-cavity of 90 cm total length with two curved mirrors (M_3, M_4) of 77.5 mm radius of curvature and two flat mirrors (M_1, M_2). The mirrors are tilted with an angle equal to 6°. The used non-linear crystal for the frequency doubling operation is a 3×5×10 mm *a*-cut KTP crystal (θ=90°) with two faces antireflection coated for low loss at both fundamental and harmonic wavelengths. The KTP crystal is placed between M_3 and M_4 at the position of the cavity waist w_{cav}=32 μm. We note that this value is quite close to the optimum waist for second harmonic generation (26 μm for 10 mm KTP crystal) [12]. The two curved mirrors and the flat mirror M_2 have coatings that are highly reflecting at 1.064 μm (≥ 99.9%) with high transmission at 532 nm (T_4=98%). The available input coupler has 1.4% transmission at 1.064 μm.

A servo-loop monitors the length of the cavity via a piezoelectric ceramic glued on M_2 so that we keep the cavity at resonance with the laser frequency. The cavity length is modulated at 80 kHz and the servo-loop bandwidth is 4 kHz, limited by the ceramic intrinsic resonance. The green light is filtered to eliminate the residual fundamental light at 1.064 μm before sending it into an iodine cell.

The type II non-critical phase matching in the ring cavity is achieved by means of a half-wave plate inside the cavity since it compensates the rotation of the polarization due to the birefringence of the crystal [13]. This use allows us to overcome high temperature operation, needed to achieve a non-critical phase matching in a ring cavity [14]. A second half-wave plate before the cavity is used both to set the polarization at the input to 45° and to optimize this polarization taking into account the depolarization induced by the input coupler.

When the non-critical phase matching condition is fulfilled, we measure both the second harmonic output power $P_{2\omega}$ and the fundamental incident power P_ω. In Fig. 4, we display the harmonic output $P_{2\omega}$ as a function of the incident infrared power P_ω (Fig. 4a) and the dependence of the total conversion efficiency η of the doubling cavity on P_ω (Fig. 4b),

$$\eta = P_{2\omega}/P_\omega = \left(\frac{4\rho T_1}{\sqrt{T_4}}\right)^2 \frac{\eta_{SHG} P_\omega}{\left(T_1 + p + \sqrt{\eta}\sqrt{\eta_{SHG} P_\omega}\right)^4} \quad (1)$$

where η_{SHG} is the single-pass conversion efficiency, p=1% represents the total round-trip loss and ρ = 66% is the fraction of input light effectively coupled to the fundamental mode of the cavity (mode matching). The theoretical curves in Fig. 4 are derived by solving the third degree equation (1). The single-pass conversion efficiency η_{SHG} of the crystal, which is set to the value η_{SHG}=1.3 10^{-3} W^{-1}, is the only adjustable parameter. The maximum second harmonic power generated by the doubling cavity is 119 mW for an input infrared power of 350 mW, corresponding to 34% conversion efficiency at room temperature.

To perform the iodine molecular spectroscopy , the green light is sent into a second Fabry-Perot cavity placed around a 120 mm long iodine cell. This allows to use a well defined laser mode minimizing then the line broadening due to first order Doppler effect. We have used an acousto-optic modulator (AOM) to isolate the doubling cavity from the iodine cavity and also to compensate the power drift at 532 nm due to thermal effects in the KTP crystal. Let us note that the iodine saturated absorption has not yet been observed because of a quite important jitter of LD1. This frequency noise is converted in amplitude noise by the two cavities. To reduce this jitter, we plan to frequency lock LD1 on a stable external cavity following the Pound-Drever technique.

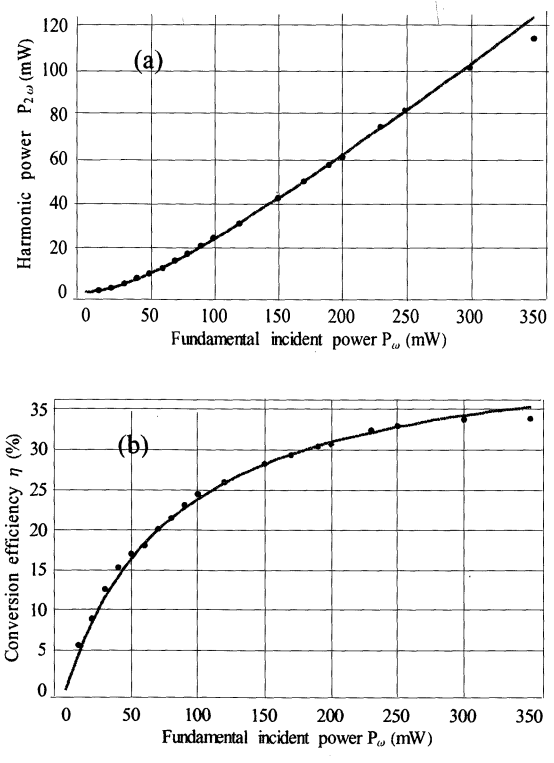


Fig. 4 (a) Second harmonic power $P_{2\omega}$ as a function of incident infrared power P_ω, (b) Conversion efficiency η as a function of infrared incident power. The solid curve is derived from equation (1).

3.3 SHG at 1.556 μm in a PPLN crystal

To frequency lock the laser diode 2 (LD2) on the two-photon Rubidium standard, we frequency double LD2 in a 19 mm long PPLN crystal (multi-grating period from 18.6 to 20.4 μm). The sample is heated at 65 °C to achieve the optimum second harmonic conversion at 1.556 μm. This temperature operation corresponds to a 19.0 μm period grating. In these conditions, we obtain 6 μW at 778 nm with 20 mW input power on the crystal. A 1.5% / W single-pass

conversion efficiency has thus been achieved. The output beam will be separated from the fundamental with a dichroïc beam-splitter and then compared to the Rubidium standard.

3.4 DFG in a type II cut $AgGaS_2$ crystal

The final step of this work is to generate the frequency difference between LD1 and LD2 using a 15 mm long type II cut $AgGaS_2$ crystal. We will generate the 88 THz/3.39 μm frequency radiation which will be linked to the third harmonic of our 29.1 THz/10.3 μm OsO_4 standard via a 10 mW intermediate He-Ne laser. Our calculation result is presented in Fig. 6. It shows that with 200 mW at 1.067 μm and 20 mW at 1.556 μm powers, we can generate up to 4 μW of 3.39 μm radiation. That suffices to realize a beat note with the 88 THz intermediate laser locked with respect to the third harmonic of the CO_2/OsO4 optical frequency standard.

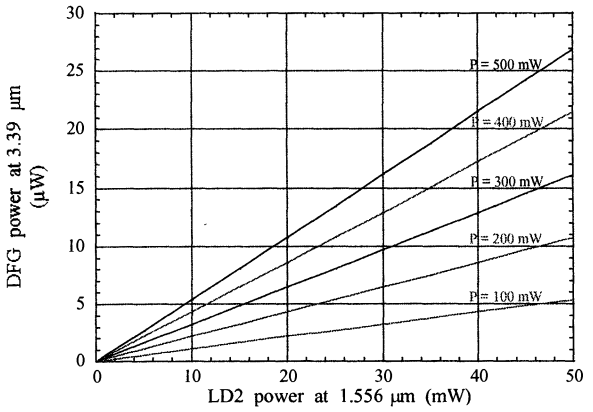


Fig. 6 Calculated power at 3.39 μm as a function of LD2 power for different values of LD1 power.

4. CONCLUSION

We have shown that a Schottky diode can be used to measure large frequency difference between two lasers at 1.06 μm. The SNR that we have obtained for an harmonic mixing up to 30 GHz is almost the same as the one observed in the 850 nm range where gaps of more than 1 THz have been spanned in previous works. The Schottky diode confers in a simple way a very large frequency tunability to the frequency synthesis chain. We have achieved SHG from 1064 to 532 nm in a KTP crystal at room temperature and from 1556 to 778 nm in a PPLN crystal with respectively output powers of 119 mW (34% conversion efficiency) and of 6 μW (1.5%W single-pass conversion efficiency). In the near future, we will generate the frequency difference in a $AgGaS_2$ crystal and compare it to the third harmonic of the OsO_4 standard. In the mean time, we plan to develop the experimental setup for iodine spectroscopy. As soon as LD1 will be frequency locked on iodine transition, we will be able to measure both the absolute frequency of any OFS around 563.2 THz/532 nm and the iodine transitions in 2 THz frequency range. This opens the way to a new investigation of iodine spectroscopy, up to now inaccessible with conventional frequency doubled Nd:YAG lasers.

5. ACKNOWLEDGEMENTS

This work is a part of the scientific developments of the Bureau National de Métrologie (BNM-France). We gratefully thank J. J. Zondy for pertinent remarks, in the same way as P. Yelisseyev and L. I. Isaenko from DTIM (Novossibirsk) for providing the KTP crystal. We are gratefully indebted to Y. Millerioux, and P. Juncar (from BNM-INM) for useful discussions and loaning optical components needed to achieve this experiment. We also acknowledge L. Robertson and J. M. Chartier (from BIPM) for their valuable help for the iodine set-up. Great thanks are due to the electronic staff and to A.H. Gerard for their help in the course of this work.

6. REFERENCES

[1] F. Ducos, G.D. Rovera, C. Daussy and O. Acef, " Performances of OsO4 stabilized CO2 lasers as optical frequency standards near 29 THz", in Proceedings of the 1999 Joint Meeting of the European Frequency and Time Forum and the IEEE International Frequency Control Symposium, Vol.2, 1999, pp.714-717.

[2] D. Touahri, O. Acef, A. Clairon, J.J. Zondy, R. Felder, L. Hilico, B. Beauvoir, F. Nez and F. Biraben, "Frequency measurement of the $5S_{1/2}$-$5D_{1/2}$ two photon transition in rubidium", Optics Comm., Vol. 133, p. 471, 1997.

[3] O. Acef, A. Clairon, G.D. Rovera, F. Ducos, L. Hilico, G. Kramer, B. Lipphardt, A. Shelkovnikov, E. Kovalchuck, E. Petrukhin, D. Tyurikov, M. Petrovsky, M. Gubin, R. Felder, P. Gill, S. Lea, "Absolute frequency measurements with a set of transportable methane optical frequency standards", in Proceedings of the 1999 Joint Meeting of the European Frequency and Time Forum and the IEEE International Frequency Control Symposium, Vol.2, 1999, pp.742-745.

[4] 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 HeNe/I2 laser", Optics Comm., Vol. 97, pp. 29-34, 1993.

[5] 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.

[6] J.L. Hall, L.S. Ma, M. Taubman, B. Tiemann, F.L. Hong, O. Pfister and J. Ye, "Stabilization and frequency measurement of the I2 stabilized Nd:YAG laser", IEEE Trans. Instrum. Meas., Vol.48, No. 2, pp. 583-586, 1999, and references therein.

[7] T. J. Quinn, "Practical realization of the definition of the meter (1997)", Metrologia, Vol. 36, pp. 211-244, 1998.

[8] L. Hilico, R. Felder, D. Touahri, O. Acef, A. Clairon and F. Biraben, "Metrological features of the rubidium two-photon standards of the BNM-LPTF and Kastler Brossel Laboratories, Eur. Phys. J. AP, Vol. 4, pp. 219-225, 1998.

[9] O. Acef, L. Hilico, M. Bahoura, F. Nez and P. de Natale, "Comparison between MIM diodes as harmonic mixers for visible lasers and microwave sources", Optics Comm., Vol. 109, pp.428-434, 1994.

[10] M. Kourogi, K. Nakagawa and M. Ohtsu, "A highly accurate frequency counting system for 1.5 μm wavelength semiconductor lasers", SPIE, Vol. 1837, pp. 205-215, 1992.

[11] S. A. Diddams, D. J. Jones; L.S. Ma, S. T. Cundiff and J. L. Hall, "Optical frequency measurement across a 10-THz gap with a femtosecond laser frequency comb", Optics. Letters, Vol. 25, No. 3, pp. 186-188, 2000.

[12] G. D. Boyd and D. A. Kleinman, J. Appl. Phys, Vol. 39, pp. 3597, 1968.

[13] J. J. Zondy, M. Abed, and S. Khodja, "Twin-crystal walk-off-compensated type-II second-harmonic generation: single-pass and cavity-enhanced experiments in KTiOPO4", J. Opt. Soc. Am. B, Vol. 11, No. 12, pp. 2368-2378, 1994.

[14] Z. Y. Ou, S. F. Pereira, E. S. Polzik, and H. J. Kimble, "85% efficiency for cw frequency doubling from 1.08 to 0.54 μm", Optics Lett., Vol. 17, pp. 640-642, May 1, 1992.

FREQUENCY STABILIZATION OF A 1.54 μm DFB-LASER DIODE TO DOPPLER-FREE LINES OF ACETYLENE

U.Sterr and T. Kurosu*

Physikalisch-Technische Bundesanstalt, Bundesallee 100, D-38116 Braunschweig, Germany
uwe.sterr@ptb.de, Tel. +49 531 591 4312, FAX +49 531 591 4305

*National Research Laboratory of Metrology, 1-1-4 Umezono, Tsukuba, Ibaraki 305, Japan

ABSTRACT

We present results for the stabilization of a DFB laser diode to Doppler-free lines of acetylene. The fast linewidth of an standard DFB-laser diode was reduced by optical feedback from a partially reflecting mirror which allows to lock its frequency to a narrow linwidth build-up resonator. The method of cavity-enhanced noise-immune molecular spectroscopy (NICE OHMS) is used to detect the the saturated absorption lines with high signal to noise ratio.

1. INTRODUCTION

Optical frequency standards in the 1.5 μm region are of major interest for metrology in optical telecommunication systems. In this wavelength region a number of molecular overtone lines are available as references. Because of their ro-vibrational structure they provide a grid of lines that can be very useful for the calibration of measurement equipment. The small molecular dipole moment leads to small pressure shift (1.5 kHz/Pa) and broadening (230 kHz/Pa) and to a narrow natural linewidth $\Delta\nu \sim 1$ kHz. With high-pressure cells Doppler broadened lines are easily detectable. However, the attainable accuracy is limited by the linewidth and by the pressure shift. For a fractional frequency accuracy of better than 10^{-9} Doppler-free spectroscopic methods have to be used.

Because of the small transition matrix elements of these molecules, high laser intensities are required to saturate these transitions. As the saturation power also increases with pressure and to limit pressure shifts, it is necessary to work at pressures of a few Pa, which leads to very small total absorption in the usual length of the absorben cells. One way to overcome both problems is to use an separate resonator for power build-up and to increase the effective absorption length.

2. EXPERIMENTAL SETUP

2.1 Diode - Laser

We use a DFM-laser diode with a free-running line width of 10 MHz. This fast line width is reduced to below 100 kHz by optical feedback from a partially reflecting mirror (R=15%) at a distance of 100 mm from the laser's output facet. The laser can be tuned over a few GHz by translation of the feedback mirror with a piezoelectric transducer (PZT). To allow for the correction of fast frequency fluctuations, an AR-coated $LiNbO_3$ crystal (25mm x 5 mm x 10 mm) acting as phase modulator (EOM) is introduced in the optical feedback path.

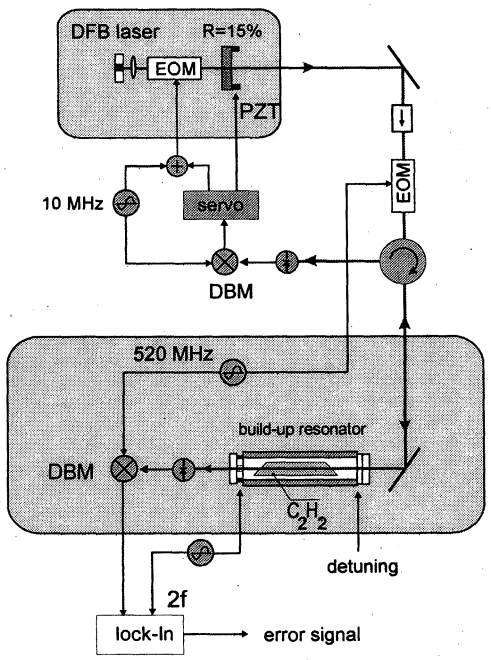


Fig. 1: Experimental setup.

2.2 Build-up resonator

The build-up resonator of length L = 290 mm contains an absorption cell filled with 1.3 Pa of acetylene $^{13}C_2H_2$. The resonator length can be tuned and modulated by two separate PZTs. An intracavity cell with Brewster windows is used to prevent contamination and to allow for operation as a sealed system while still be able to tune the resonator length. With the cell included the finesse is F = 300 and approximately 85% of the incident power (3 mW) can be coupled in the cavity's fundamental mode.

2.3 Stabilization

The laser is locked to the build-up resonator using the Pound-Drever-Hall method [1]. The laser frequency is frequency-modulated at 10 MHz using the EOM inside the feedback path with a modulation index $M \approx 0.2$. This modulation creates sidebands on the laser spectrum. When the light consisting of carrier and sideband is sent to the resonator, the carrier is coupled in while the sidebands are reflected with amplitude and phase nearly independent of the actual frequency offset between laser- and resonator frequency. However the amplitude and the phase of the carrier leaking from the cavity strongly depend on that detuning. Therefore in the reflected light any detuning of the laser frequency is transformed to an imbalance between the frequency components and subsequently to an amplitude modulation (AM). After phase sensitive detection, this AM leads to an error signal suitable for locking the laser to the resonator. After integration and filtering the signal is applied to the PZT (low frequency part) and to the EOM (high frequency part).

3. SPECTROSCOPIC METHOD

While the locking of the laser to the build-up resonator would already allow to observe molecular absorption by monitoring the transmission of the resonator while tuning its eigenfrequency across the line, the achievable signal to noise ratio (SNR) is severely degraded by residual amplitude and frequency noise of the laser. The resonator's narrow fringe effectively transforms frequency fluctuations to amplitude noise that directly affects the absorption signal. In addition, low frequency laser amplitude noise and geometry fluctuations that affect the coupling efficiency degrade the SNR. FM methods widely applied for high-resolution spectroscopy are not easily applicable because of the discrete resonance frequencies of the build-up cavity. One way to improve the SN-ratio is the use of an FM-type spectroscopy, where the modulation frequency is chosen to match the free spectral range of the build-up resonator. There, the sidebands are also resonant with the cavity and can serve as references for the detection of the nonlinear phase shift of the carrier due to the molecular absorption. In addition, frequency fluctuations between laser and resonator lead in first order to the same phase shift for carrier and sideband and therefor no conversion from FM to AM takes place. This noise immune cavity enhanced overtone molecular spectroscopy (NICE-OHMS) [3] was first introduced by Ye et al. for the detection of weakly absorbing molecular overtone lines.

In our setup we use a external EOM, driven by a stable frequency of approximately 520 MHz to generate the sidebands. The frequency was generated by a synthesizer and adjusted by minimizing the noise around the modulation frequency as observed with a rf spectrum analyser. The modulation-index was 0.6 leading to 8% of the total power in the optical sidebands. The transmission signal of the cavity was detected phase sensitively with a fast InGaAs photodetector, an amplifier and a double-balanced mixer (DBM). This signal is sensitive to the phase difference between carrier and sidebands that is introduced by the linear and by the saturated dispersion of the absorber. As the modulation frequency is of the same magnitude as the inhomogeneous Doppler broadening, the signal also contains a broad background contribution from the broad dispersion curve. To reduce this background the resonance frequency of the build-up resonator and correspondingly the laser frequency was dithered at a frequency of 400 Hz and an amplitude of 2 MHz. Detecting the output of the DBM with a lock-in amplifier at the low-frequency modulation frequency or at twice the modulation frequency leads to a symmetric or antisymmetric background-free signal respectively. We have used the 2f-detection to derive a dispersion-like signal suitable for frequency stabilization. This approach corresponds to the 3f-detection widely employed for e.g. the frequency stabilization of iodine-stabilized He-Ne lasers. The signal of the Doppler broadened line, the Lamb-dip and the dispersion-like 2f-signal are shown in fig. 2.

The error signal was then integrated and sent to the second tuning PZT of the build-up resonator to stabilize the laser frequency to the center of the molecular line.

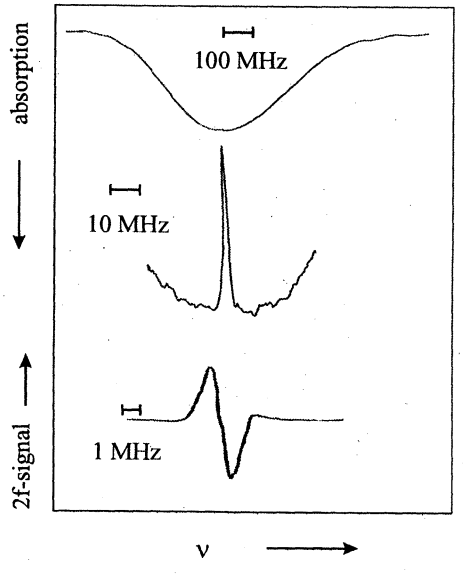


Fig. 2: Observed transmission signal around the Doppler broadened line ($\nu_1+\nu_3$) P(6) (top) and magnified central part showing the Lamb-dip (center trace). The dispersion-like 2f-signal is shown in the bottom trace.

4. RESULTS

To assess the stability of the laser frequency an interferometric wavelength comparison to an iodine-stabilized 633 nm He-Ne Laser was used. The length of an evacuated Fabry-Perot resonator was stabilized using the 1.5 μm radiation of the acetylene stabilized laser. An additional 633 nm He-Ne laser was also stabilized to the same resonator, using a standard 3f-locking method (fig.2). The frequency of this laser was in turn compared to an iodine stabilized He-Ne laser by counting the beat-frequency between both radiations.

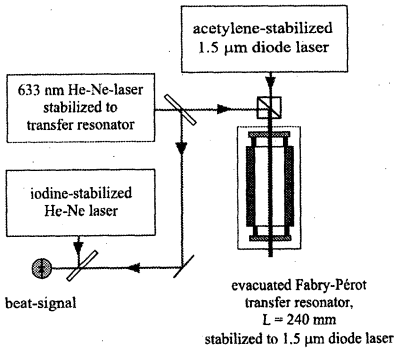


Fig. 3: Setup for interferometric wavelength comparison between the acetylene–stabilized diode laser and an iodine stabilized He-Ne laser.

The resulting Allan deviation of the frequency difference that is equal to the Allan deviation of the 1.5 μm laser with respect to the iodine-stabilized laser is shown in fig. 3. Up to now this Allan deviation is nearly one order higher than the deviation between two He-Ne lasers, so the Allan deviation is mainly due to the acetylene-stabilized diode laser.

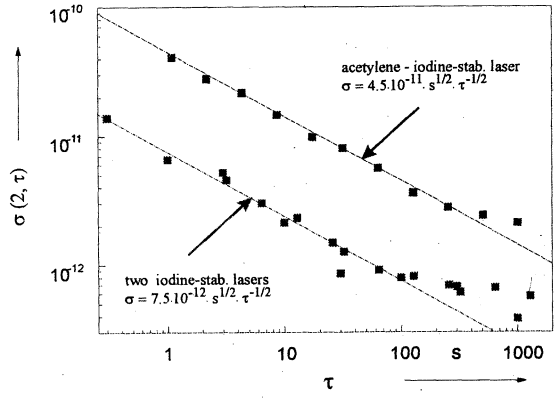


Fig. 4: Allan standard deviation between an acetylene-stabilized diode laser and a iodine stabilized He-Ne laser. For comparison the Allan deviation between two He-Ne lasers is also shown.

We assume that the stability is limited because of the mechanical stability of the setup and because of spurious feedback in the optical setup that leads to additional etalons that produces additional amplitude modulation at the employed FM modulation frequencies.

5. CONCLUSION

We have presented a DFB laser diode stabilized to saturated absorption lines of acetylene. The setup allows with only few mW of available power to achieve a good SNR and a stability of $\sigma_y = 2 \cdot 10^{-12}$ at 1000 s averaging time. With an improved optical and mechanical setup further improvements seem possible.

6. ACKNOWLEDGEMENTS

T. Kurosu acknowledges financial support from the Alexander-von-Humbold foundation.

7. REFERENCES

[1] K. Nakagawa, M. de Labachelerie, Y. Awaji, M. Kourogi, "Accurate optical frequency atlas of the 1.5-μm bands of acetylene", J. Opt. Soc. Am. B vol. 13, pp 2708 , 1996

[2] R.W.P. Drever, J.L. Hall, F.V. Kowalski, J. Hough, G.M. Ford, A.J. Munley, H. Ward, "Laser Phase and Frequency Stabilization Using an Optical Resonator", Appl. Phys. B, vol. 31, pp 97-105, 1983.

[3] Jun Ye, Long-Shen Ma, J.L. Hall, "Sub-Doppler optical frequency reference at 1.064 μm by means of ultrasensitive cavity-enhanced frequency modulation spectroscopy of a C_2HD overtone transition", Opt. Lett. 21, pp 1000, 1996.

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

FREQUENCY-NOISE SENSITIVITY AND AMPLITUDE-NOISE IMMUNITY OF DISCRIMINATORS BASED ON FRINGE-SIDE FABRY-PEROT

E. Bava, G. Galzerano, and C. Svelto

Dipartimento di Elettronica e Informazione del Politecnico di Milano, INFM and CNR-CSTS
Piazza Leonardo da Vinci 32, 20133 Milano, Italy
Phone: +39 02 2399 3609 Fax: +39 02 2399 3413 E-mail: Elio.Bava@Polimi.it

ABSTRACT

Optical frequency discriminators employing Fabry-Perot cavities are analyzed studying their sensitivity to laser frequency fluctuations with particular attention to their amplitude-noise immunity. Comparisons between different schemes, with intensity noise compensation, based on transmission, reflection or a proper combination of both these signals have been performed, pointing out similarities and differences in their operation properties.

1. INTRODUCTION

This contribution is concerned with the analysis of frequency-discriminator systems based on the use of Fabry–Perot resonators and of direct detection for the characterization of laser-source frequency fluctuations. Although frequency discriminators are provided of some compensation for amplitude fluctuations, they are not, however, completely insensitive to amplitude noise, which therefore must be independently measured and then subtracted from the output of the particular discriminator system used in order to know the actual frequency noise of the laser.

The problem of evaluating the responses of discriminator schemes to frequency fluctuations associated to the signal impinging into the discriminator itself has been faced in the frequency domain in [1], extended in [2] and [3] by including the responses due to both frequency and amplitude noise, studied for three kinds of discriminator systems.

On the other hand, to the same purpose, a time-domain analysis can be carried out. A few situations of responses to frequency-step variations have been discussed in [4], whereas the evolutions of the overall system response to amplitude steps has been studied in [5]. The method used in [4] uses a signal with an initial frequency until $t = 0$, then a new signal with a step frequency variation at $t=0$, introducing at $t=0$ two transients inside the resonator at different frequencies. Moreover many approximations are necessary, even if reasonable, and when the responses are evaluated it is not clear whether the results of the time-domain analysis could coincide with those of the frequency-domain analysis. This coincidence is not mandatory in general because the transients start with signals already present in the resonator and relaxing in the system.

In this work the response to frequency noise in the time domain has been studied through the evaluation of the transient determined by a phase step variation of the signal. The model works without introducing approximations, unless for the linearizations consisting in omitting noise contributions of higher orders. Therefore non necessary approximations introduced in [5] are here avoided. The use of the phase step and of the amplitude step shows directly that the two kinds of analysis, in the time and in the frequency domain, must give the same results, because there is just one transient even if the resonator is full of steady radiation at the time $t = 0$. Moreover insights in the system behaviors appear more direct with the time domain analysis. Numerical results confirm the uniqueness of the solutions for both the frequency discrimination and the amplitude noise immunity.

2. FREQUENCY DISCRIMINATION

The setups for measuring laser frequency fluctuations by using the discriminating effect of a Fabry-Perot are shown in Fig. 1 for the three configurations discussed in the following. The two schemes (a) and (b) are quite simple, they use the fringe side slope in transmission or in reflection, however the third (c), by using both reflection and transmission, has a higher sensitivity to frequency variations and a better rejection of amplitude noise as it will be shown later. In all the three cases 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 therefore converted into amplitude fluctuations by the opposite slopes of the resonator transmission or reflection coefficients.

A straightforward method to study a frequency discriminator turned out to be the application of a phase step to the input signal of the system, corresponding to an impulsive excitation of the input frequency. From this transient analysis it is possible therefore to evaluate the frequency response of the overall system, provided it is linear. A former non-linearity is the assumption that the photodetector responds linearly to the incident intensity, however linearization in the neighborhood of the operation point faces this problem without difficulty

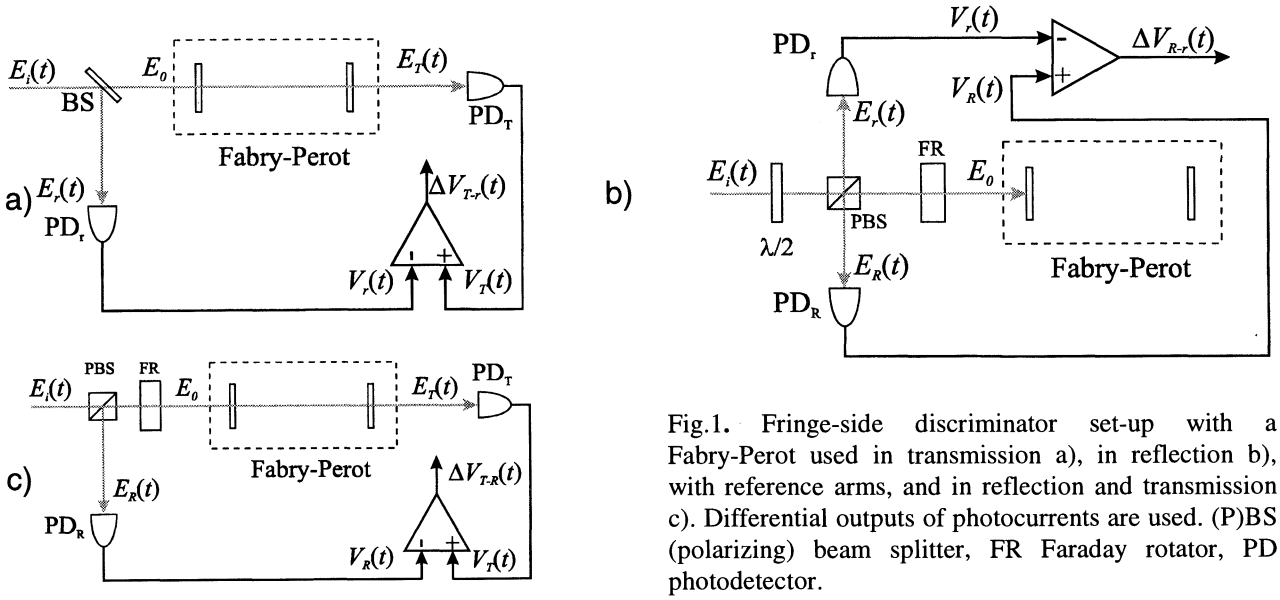


Fig.1. Fringe-side discriminator set-up with a Fabry-Perot used in transmission a), in reflection b), with reference arms, and in reflection and transmission c). Differential outputs of photocurrents are used. (P)BS (polarizing) beam splitter, FR Faraday rotator, PD photodetector.

thinking that fluctuations are responsible for output variations. A latter non-linearity is determined in principle by the relation between phase variation and input signal, as it will be discussed in the following. The requirement to avoid this effect is that the phase step must be << 1 rad. This last condition, or the equivalent one about the total amount of phase noise << 1 rad^2 , appears necessary to face this problem also with the method of the frequency-domain analysis.

Let us consider the incident field into the Fabry-Perot resonator be represented by a sinusoidal electric field, $E_i(t)$, evolving at the angular frequency ω_0 and suffering a sudden phase jump $\Delta\varphi$ at $t = 0$, that is $E_i(t) = E_0 \exp\{j(\omega_0 t + \varphi_0)\}$ for $t < 0$ and $E_i(t) = E_0 \exp\{j(\omega_0 t + \varphi_0 + \Delta\varphi)\}$ for $t \geq 0$; therefore the field $E_i(t)$ is supposed to be made of a sinusoidal component, unaltered for t varying from $-\infty$ to $+\infty$, and of an additional step-function variation at $t = 0$, namely

$$E_i(t) = E_0 \exp\{j(\omega_0 t + \varphi_0)\}\{1 + U(t)[\exp(j\Delta\varphi) - 1]\} \cdot \quad (1)$$

The main result to this point is that there is no relaxation at $t = 0$ of the energy trapped in the resonator, because there is just the transient determined by the step function. Moreover, if a linear relationship is sought between $\Delta\varphi$ and the output variation, $\Delta\varphi$ << 1 must hold. Following the procedure outlined in [4] and [5], by using the mirrors amplitude reflection and transmission coefficients r_k and t_k ($k = 1, 2$ for the first and second mirror, respectively), it is possible to write the expressions for the transmitted (E_T) and reflected (E_R) fields, after the phase step $\Delta\varphi$, as a function of the current number, n, of round trip time τ

$$E_T(n\tau) = \frac{t_1 t_2 \exp\left\{j\left[\omega_0\left(t - \frac{\tau}{2}\right) + \varphi_0\right]\right\}}{1 - r_1 r_2 \exp(-j\omega_0\tau)} E_0 \times \left\{1 + U\left(t - \frac{\tau}{2}\right)[\exp(j\Delta\varphi) - 1] \times \left[1 - (r_1 r_2 \exp(-j\omega_0\tau))^{n+1}\right]\right\} \quad (2)$$

$$E_R(n\tau) = \frac{\exp\{j(\omega_0 t + \varphi_0)\}}{1 - r_1 r_2 \exp(-j\omega_0\tau)} E_0 \times \left\{[r_1 - r_2 \exp(-j\omega_0\tau)] + U(t)[\exp(j\Delta\varphi) - 1] \times \left[r_1 - r_2 \exp(-j\omega_0\tau) - t_1^2 r_2 (r_1 r_2)^n \exp[-j(n+1)\omega_0\tau]\right]\right\} \quad (3)$$

and these expressions hold for $n = 0, 1, 2, 3, \ldots$etc.

To carry out the transient evaluation of the system, after the phase step perturbation, the detected photocurrents are written according to the general relationship

$$i(t) = \eta S\left(\frac{c\varepsilon_0 e}{2h\nu_0}\right) E\, E^* = \frac{1}{2} k_i E\, E^* \quad (4)$$

where η is the quantum efficiency, S the equivalent detector area, $\nu_0 = \omega_0/2\pi$, c the velocity of light, ε_0 the vacuum permittivity, e the electron charge and h the Planck constant. The factor k_i, different in principle for each detector, transforms the incident light intensity into a detected photo-current. These current values when multiplied by the transimpedence gain of the detector electronics k_V yield the detected voltages V applied by the discriminator channels to the two inputs of the differential stages. Assuming $\Delta\varphi$ << 1, the following expressions are obtained for $V_T(n\tau)$ and $V_R(n\tau)$:

$$V_T(n\tau) = \frac{1}{2} k_{i,T} k_{V,T} \frac{|t_1 t_2|^2}{k_\rho^2} E_0^2 \times$$

$$\left\{1 - 2\Delta\varphi U\left(t - \frac{\tau}{2}\right)(r_1 r_2)^{n+1} \operatorname{sen}(n+1)\omega_0\tau\right\} \quad (5)$$

$$V_R(n\tau) = \frac{1}{2} k_{i,R} k_{V,R} \frac{\rho^2}{k_\rho^2} E_0^2 \times$$

$$\left\{1 - 2\Delta\varphi U(t)\frac{t_1^2}{\rho^2} \left[(r_1 r_2)^{n+1} \operatorname{sen}(n+1)\omega_0\tau - r_2^2 (r_1 r_2)^n \operatorname{sen} n\omega_0\tau\right]\right\} \quad (6)$$

where both $k_\rho^2 = 1 + (r_1 r_2)^2 - 2 r_1 r_2 \cos\omega_0\tau$ and $\rho^2 = r_1^2 + r_2^2 - 2 r_1 r_2 \cos\omega_0\tau$ depend on the operating conditions of the Fabry-Perot through the detuning $\Omega_0 = \omega_0 - \omega_c$ of the laser angular frequency ω_0 from the cavity resonance ω_c.

In the reference arms the electric field incident onto the detector are proportional to E_0 through the coefficients k_a or k_b, according to the case (a) or (b), respectively. The steady field produces in these detectors the photocurrents $i_{T,r} = \frac{1}{2}\, k_{i,r} |k_a|^2 E_0^2$ and $i_{R,r} = \frac{1}{2} k_{i,r} |k_b|^2 E_0^2$ and as a consequence the following detected voltages $V_{T,r} = \frac{1}{2} k_{i,r} k_{V,r} |k_a|^2 E_0^2$ and $V_{R,r} = \frac{1}{2} k_{i,r} k_{V,r} |k_b|^2 E_0^2$, respectively. Taking into account the balance conditions before the step $\Delta\varphi$, following the method widely described in [5], the differential voltages are obtained for the circuits of Fig.1.

Considering that the phase step $U(t)\Delta\varphi$ corresponds to a frequency variation $\delta(t)\ \Delta\varphi/2\pi$, the expressions of the differential voltages are the transient responses determined by a frequency pulse of amplitude $\Delta\varphi / 2\pi$. To obtain the discriminator frequency response we perform the Laplace transform of $2\pi\ \Delta V/\Delta\varphi$, then $j\omega = s$ is substituted. The time variable is present in the integer n and, for each n value, a constant output value is obtained in the time interval $n\tau \le \theta \le (n+1)\tau$, where θ is a continuous time variable shifted of amounts which account for delays in different arms. The evolution of the trigonometric functions is determined by Ω_0; in fact for $\omega_0 = \omega_c$, $\cos(n\omega_0\tau) = 1$ is a constant independent of n. The Laplace transform of these series of constant values leads to geometric series which can be easily evaluated. The final expressions for these frequency responses are

$$F_a(j\omega) = -2\pi k_{i,T} k_{V,T} \frac{|t_1 t_2|^2}{k_\rho^2} \tau E_0^2 \exp(-j\omega\tau/2) \frac{\sin(\omega\tau/2)}{\omega\tau/2} \times$$

$$\frac{\exp(\alpha\tau)\operatorname{sen}\Omega_0\tau}{[1 - \exp\{(p_1 - j\omega)\tau\}]\cdot[1 - \exp\{(p_2 - j\omega)\tau\}]}$$

$$= -2\pi k_{i,T} k_{V,T} E_0^2 f_a(j\omega) \quad (7)$$

$$F_b(j\omega) = -2\pi k_{i,R} k_{V,R} \frac{t_1^2}{k_\rho^2} \tau E_0^2 \exp(-j\omega\tau/2) \frac{\sin(\omega\tau/2)}{\omega\tau/2} \times$$

$$\frac{\exp(\alpha\tau)\operatorname{sen}\Omega_0\tau\left[1 - r_2^2 \exp(-j\omega\tau)\right]}{[1 - \exp\{(p_1 - j\omega)\tau\}]\cdot[1 - \exp\{(p_2 - j\omega)\tau\}]}$$

$$= -2\pi k_{i,R} k_{V,R} E_0^2 f_b(j\omega) \quad (8)$$

$$F_c(j\omega) = -2\pi k_{i,T} k_{V,T} \frac{|t_1 t_2|^2}{k_\rho^2} \tau E_0^2 \exp(-j\omega\tau/2) \frac{\sin(\omega\tau/2)}{\omega\tau/2} \times$$

$$\frac{\exp(\alpha\tau)\operatorname{sen}\Omega_0\tau\left[\left(1 - \frac{t_1^2}{\rho^2}\right) + \frac{t_1^2}{\rho^2} r_2^2 \exp(-j\omega\tau)\right]}{[1 - \exp\{(p_1 - j\omega)\tau\}]\cdot[1 - \exp\{(p_2 - j\omega)\tau\}]}$$

$$= -2\pi k_{i,T} k_{V,T} E_0^2 f_c(j\omega) \quad (9)$$

where $p_1 = \alpha + j\Omega_0$, $p_2 = \alpha - j\Omega_0$, $\alpha = \frac{1}{\tau}\ln r_1 r_2 = \frac{1}{\tau_s} = -\pi\cdot\Delta\nu_c$, $\Delta\nu_c$ is the Fabry-Perot linewidth. The parameters k_ρ^2, ρ^2 and $\Omega_0\tau$ can be expressed in terms of cavity and operation parameters, (see [5]). In usual operations on a fringe side $\cos\Omega_0\tau \approx 1$, $k_\rho^2 << 1$, and $\rho^2 << 1$.

The diagrams of the normalised responses $|f_a(j\omega)|$, $|f_b(j\omega)|$, $|f_c(j\omega)|$, are shown in Fig. 2, for equal and lossless mirrors with the particular values $r_1 = r_2 = 0.99$, $|t_1| = |t_2| = 0.14$ and $\tau = 1.667$ ns; according to this normalisation the three systems are compared at equal Fabry-Perot input fields and overall electronic gains. At

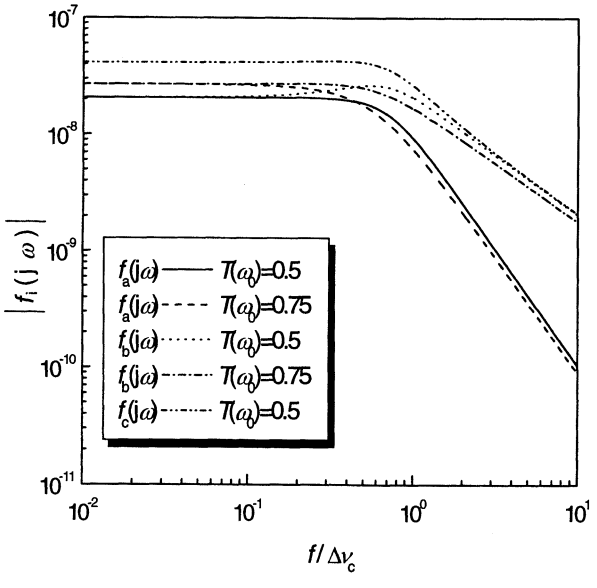


Fig. 2. Normalized frequency-noise response for differential schemes operating in reflection, transmission or both.

very low frequencies the ratio $|f_a(j\omega)| / |f_b(j\omega)| = 1$, whereas, according to expression (16) in [5],

$|f_a(j\omega)|/|f_c(j\omega)|=T(j\omega_0)$ the cavity power transmission dependent in general on the operation point adopted in system (c). In this last case T = ½ and system (c) shows its higher discriminating power as expected. In usual conditions the common factor $\frac{sin(\omega\tau/2)}{\omega\tau/2}$ depends slowly on ω as compared to the position of the complex conjugate poles, which on a fringe side can be substituted by $[\alpha+j(\Omega_0-\omega)]\ [\alpha-j(\Omega_0+\omega)]\ \tau^2$, and of the zeros of $f_b(j\omega)$ and $f_c(j\omega)$. For the complex conjugate poles, the real part is directly linked to the Fabry-Perot characteristics and the imaginary part is dependent on the working point conditions (Ω_0). For T=1/2, $|\alpha|=|\Omega_0|$, for T=0.75, $|\alpha|=2|\Omega_0|$. Their frequency cut-off is near to $\Delta\nu_c/2$. The zero of $f_b(j\omega)$ is located, for equal mirrors, close to - α, therefore the 3 dB frequency is $f_{3dB}=\Delta\nu_c/2$ and the effect of one pole is compensated as it is shown in Fig. 2. As regards the expression (15) the zero is on the real axis for $s = 2\alpha$ corresponding to $f_{3dB}=\Delta\nu_c$.

3. AMPLITUDE NOISE SENSITIVITY

The amplitude-noise immunity of each scheme has been studied by evaluating the transient response to a step variation in the amplitude of a steady excitation. To this purpose a steady signal has been considered with amplitude E_0 at the angular frequency ω_0 and, at $t = 0$, an amplitude step $\Delta E << E_0$ has been applied while keeping ω_0 fixed. The transients in the reflected and transmitted signals have been studied according the procedure described in [5], but following the same method of Section 2 for the evaluation of the Laplace transform. After multiplication by s and the substitution $j\omega = s$, the frequency responses corresponding to output voltages $\Delta V(\omega)$ excited by $\Delta E(\omega)$ are obtained. The results are

$$H_a(j\omega)=-2ik_{i,T}k_{V,T}\frac{|t_1t_2|^2}{k_\rho^2}E_0\exp(-i\omega\tau/2)sin\frac{\omega\tau}{2}\times\frac{\exp(\alpha\tau)[\cos\Omega_0\tau-\exp(\alpha\tau-i\omega\tau)]}{[1-\exp\{(p_1-j\omega)\tau\}]\cdot[1-\exp\{(p_2-j\omega)\tau\}]}=k_{i,T}k_{V,T}E_0\,h_a(j\omega) \quad (10)$$

$$H_b(j\omega)=2i\frac{k_{i,R}k_{V,R}}{k_\rho^2}E_0t_1^2r_2\exp(-i\omega\tau/2)sin\frac{\omega\tau}{2}\times\frac{r_1\cos\Omega_0\tau-r_2+(r_2\cos\Omega_0\tau-r_1)\exp(\alpha\tau-i\omega\tau)}{[1-\exp\{(p_1-j\omega)\tau\}]\cdot[1-\exp\{(p_2-j\omega)\tau\}]}=k_{i,R}k_{V,R}E_0\,h_b(j\omega) \quad (11)$$

$$H_c(j\omega)=\frac{-2j\frac{k_{i,T}k_{V,T}|t_1t_2|^2}{k_\rho^2}E_0\exp(-i\omega\tau/2)sin\frac{\omega\tau}{2}}{[1-\exp\{(p_1-j\omega)\tau\}]\cdot[1-\exp\{(p_2-j\omega)\tau\}]}\times\left\{\begin{array}{l}\exp(\alpha\tau)\left(1-\frac{t_1^2}{\rho^2}\right)[\cos\Omega_0\tau-\exp(\alpha\tau-i\omega\tau)]\\+\frac{t_1^2r_2^2}{\rho^2}[1-\cos\Omega_0\tau\exp(\alpha\tau-i\omega\tau)]\end{array}\right\}=k_{i,T}k_{V,T}E_0\,h_c(j\omega) \quad (12)$$

These expressions show the same complex conjugate poles of Section 2 and two zeros. The former zero, common in each of the expressions, is at the origin, the positions of the latter, dependent on the set-up considered, is on the real axis of the variable s. Because $\omega\tau << 1$, the values of the zeros are

$$z_a=\frac{\cos\Omega_0\tau-\exp(+\alpha\tau)}{\tau\exp(\alpha\tau)}\approx\alpha \quad (13)$$

$$z_b=\frac{1+\frac{r_1\cos\Omega_0\tau-r_2}{r_2\cos\Omega_0\tau-r_1}\exp(-\alpha\tau)}{\tau}\approx\frac{2}{\tau}\quad \text{if } r_1=r_2 \quad (14)$$

$$z_c=-\frac{\left(1-\frac{t_1^2}{\rho^2}\right)[\cos\Omega_0\tau-\exp(\alpha\tau)]+\frac{t_1^2r_2^2}{\rho^2}[\exp(-\alpha\tau)-\cos\Omega_0\tau]}{\tau\left[\exp(\alpha\tau)\left(1-\frac{t_1^2}{\rho^2}\right)+\frac{t_1^2r_2^2}{\rho^2}\cos\Omega_0\tau\right]} \quad (15)$$

for cases (a), (b), and (c), respectively. For these reasons it is possible to recognize a common behavior for the three functions as the angular frequency tends either to zero or to infinite.

The amplitudes of the normalized functions $h(j\omega)$ are reported in Figs. 3 and 4 as a function of the Fourier frequency, f, normalized to the Fabry-Perot linewidth $\Delta\nu_c$. In Fig. 3 the amplitudes of $h(j\omega)$ in the case T=0.5 are shown for all the considered schemes and for the particular values r_1=r_2=0.99, $|t_1|=|t_2|=0.14$ and τ=1.667 ns. The effect of the zero z_b can not be seen on a normalized frequency scale up to 10 as it occurs approximately for $f/\Delta\nu_c \cong F/\pi$, where F is the resonator Finesse (in this case $F\cong156$). The trend of $|h_c(j\omega)|$ appears more complex, namely as $\omega\to0$ the slope, increased by the position of z_c for $10^{-2} < f/\Delta\nu_c < 10$, again reaches 20 dB/decade for smaller value of $f/\Delta\nu_c$, with a strong reduction of sensitivity to amplitude noise. This advantage is lost for frequencies higher than the resonator linewidth, in a region, however, of reduced capabilities as a frequency

discriminator. The high reflection, here no longer compensated by a too low transmission, transfers all the amplitude noise of this channel to the detector. The adoption of a reflection-transmission differential scheme is thus confirmed as a good solution both from the point of view of frequency discrimination sensitivity and from that of amplitude noise immunity. In Fig. 4 $|h_a(j\omega)|$ and $|h_b(j\omega)|$ are compared for two values of T, namely the condition T=0.5 which exhibits a larger bandwidth as a discriminator and the condition T=0.75 which shows a higher sensitivity and a better immunity to amplitude noise.

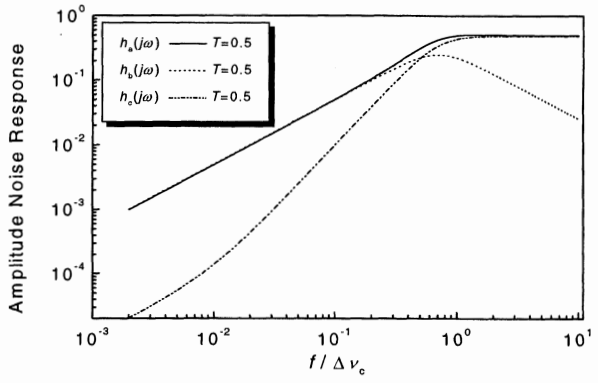


Fig. 3. Normalized amplitude-noise response for differential schemes operating in reflection, transmission or both.

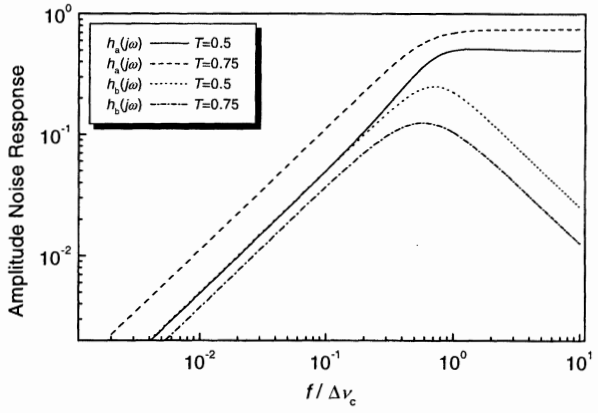


Fig. 4. Normalized amplitude-noise response for a) and b) schemes at different operating T levels.

These evaluations on frequency-noise sensitivity and amplitude-noise immunity suppose that quantum and electronic noise in the system are lower than the noise to be measured. A short discussion including these kinds of noise is reported in [2],[3] and [5].

Frequency-noise measurements carried out with these discriminating circuits should be corrected for unwanted amplitude-modulation responses according to the results expressed by (10), (11), and (12). Amplitude noise has to be measured independently, taking into account that the immunity is quite high for Fourier frequencies close to zero and rather low close to discriminator cutoff.

4. CONCLUSIONS

The frequency- and amplitude-noise responses of optical-frequency-discriminator systems based on Fabry-Perot cavities within an intensity noise compensating circuit have been analyzed, considering the effect of the resonator transient. In the adopted model mismatches between the incident field and the resonator mode have not been taken into account. From the point of view of frequency-fluctuations sensitivity and amplitude-noise immunity, systems based on transmission or reflection alone behave in a quite similar way within the discriminator bandwidth. Conversely, a proper combination of both the transmitted and reflected signals increases the frequency noise sensitivity by a factor of two and, furthermore, it allows for a higher rejection to amplitude noise.

5. REFERENCES

[1] 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.

[2] E. Bava, G. Galzerano and C. Svelto, "Measurement methods of frequency noise in optical sources based on Fabry-Perot discriminators", Proceedings of 13th EFTF, Besancon, April 1999, pp. 655-658.

[3] E. Bava, G. Galzerano, and C. Svelto, " Fabry-Perot frequency discriminators: sensitivity, amplitude noise immunity and noise limits" accepted at IMTC 2000, Baltimora(MD), May 1-4, 2000

[4] 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.

[5] E. Bava, G. Galzerano and C. Svelto, "Amplitude response of Fabry-Perot frequency discriminators", in press on *IEEE Trans. on UFFC.*

NOVEL FIBER-PUMPED QUASI-MONOLITHIC Er-Yb:GLASS LASER WITH SINGLE-FREQUENCY OUTPUT POWER OF 60 mW AT 1.534 μm

C. Svelto, G. Galzerano, F. Ferrario, and E. Bava

Dipartimento di Elettronica e Informazione – Politecnico di Milano, INFM, 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

ABSTRACT

A novel high-power fiber-pumped Er-Yb:glass quasi-monolithic microlaser operating in single-frequency at 1.5 μm has been developed. Due to the high performance in terms of available cw output power and to wavelength tunability, this laser source is attractive for a large number of applications and, in particular, for optical-fiber communications and electrooptical measurements in the 1530 nm—1560 nm wavelength region. Single-frequency output powers as high as 60 mW have been achieved with 10% slope efficiency, as well as excellent linear polarization and spatial quality of the output beam.

1. INTRODUCTION

Miniaturized, practical, and efficient all solid-state lasers operating in the near infrared have become of increasing interest due to their extremely desirable operating characteristics such as high efficiency and high power, oscillation on a single longitudinal and transverse mode with linear polarization, narrow linewidth, enhanced frequency and intensity stability by active feedback loops. In fact, the combination of these characteristics makes these lasers attractive for a variety of applications, including spectroscopy and metrology, laser radars, optical communications and photonics.

Since its first operation in the pulsed regime in 1965 [1] the erbium glass laser has attracted much interest, in particular for range-finding applications, due to its emission at the eye-safe wavelength of 1530 nm. More recently, for potential use in the third optical communication window in the 1530 nm—1560 nm range, laser physicists have in particular been actively considering cw laser oscillators based on Er^{3+} doped glasses and optical fibers. Erbium-doped glasses, when operating in this wavelength range, act as a three-level laser system and, to decrease the threshold pump rate, codoping with ytterbium is generally used. The threshold reduction arises from an effective ytterbium to erbium transfer mechanism of the excitation energy.

Diode-pumped Er-Yb:glass microlasers [2], at 1.5 μm, are very interesting sources both for optical communications and near-infrared metrology. In fact, these all-solid-state lasers oscillate on a single mode at several milliwatt output power and exhibit very narrow linewidths with extremely wide wavelength tunability compared e.g. to single-frequency laser diodes. For both high-density wavelength division multiplexing (HDWDM) [3] and frequency metrology, absolute stabilization of the emitted wavelength(s) is of great interest.

In this paper a description of the novel fiber-pumped Er-Yb:glass laser cavity is reported as well as new results achieved with this compact and reliable 1.5 μm solid-state laser.

2. THE Er-Yb ACTIVE MEDIUM

A simplified energy levels diagram of the Er-Yb:glass laser system pumped at 980 nm is shown in Fig. 1. The main processes involved include radiative absorption and emission, with absorption arising both from ground (GSA) and excited state (ESA), energy-transfer and non-radiative decay, cooperative up-conversion, ions pair formation and clusterization. Most of these processes have been investigated theoretically and/or experimentally in different glass hosts by several authors, to determine the corresponding rates. While a

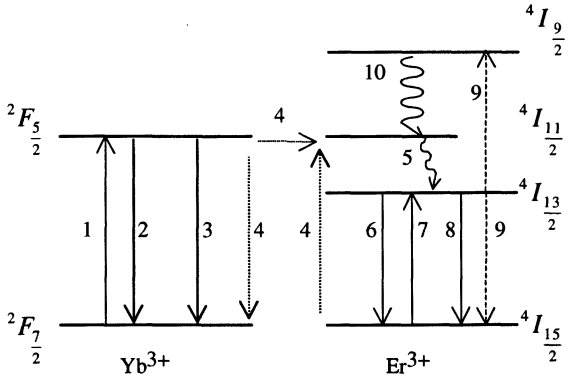


Figure 1. Energy levels diagram of the erbium-ytterbium system. The letters refer to the processes discussed in the text. The solid-line arrows refer to radiative phenomena, the dashed-dot arrows to Er-Yb energy transfer process and the dashed arrows to non-radiative transitions.

complete rate-equation system turns out to be rather involved, it can be shown that a few phenomena can be neglected on the basis of their weak influence on the behavior of Er-Yb bulk laser devices, namely clustering, ion pair formations and ESA. The processes which are thus of major concern for modeling the Er-Yb laser systems are the following (refer to the numbering in Fig. 1): (1,2) Radiative absorption and emission of a pump photon from $F_{7/2}$ and $F_{5/2}$ ytterbium levels, respectively. (3) Spontaneous decay from $F_{5/2}$ ytterbium level. (4) The Yb to Er, $Yb(F_{7/2})+Er(I_{15/2}) \rightarrow Yb(F_{5/2})+Er(I_{11/2})$, energy transfer process, that acts as indirect pumping of Er ions. (5) Erbium non-radiative decay from the $I_{11/2}$ level to the $I_{13/2}$ upper laser level. (6,7) Stimulated emission and absorption between $I_{13/2}$ upper and $I_{15/2}$ ground Er laser levels. (8) Erbium spontaneous decay from $I_{13/2}$ level. (9) Cooperative up-conversion (UC) process between two excited Er ions that promotes one ion to the $I_{9/2}$ upper level, while the second ion decays to the ground level. The first ion then relaxes, through a fast non-radiative decay, to its original $I_{13/2}$ state so that the overall effect is the loss of one, $I_{13/2}$, excited ion. To further simplify the resulting set of rate equations, the following assumptions are introduced: (i) the $I_{11/2}$ and $I_{9/2}$ levels areconsidered empty due to the fast non-radiative decays; the sum, $N_{1Er}+N_{2Er}$, of erbium populations in the lower and upper laser levels therefore reduces to the total Er concentration, N_{Er}. (ii) The main pumping process of the upper laser level is the energy-transfer rate from Yb ions and the direct absorption of pump radiation by ground state Er ions is neglected. This yields two rate equations for the upper level populations of Yb and Er ions, respectively:

$$\frac{dN_{2Yb}}{dt} = \left(\sigma_{a,Yb}N_{1Yb} - \sigma_{e,Yb}N_{2Yb}\right)F_P - KN_{2Yb}N_{1Er} - \frac{N_{2Yb}}{\tau_{Yb}} \quad (1)$$

$$\frac{dN_{2Er}}{dt} = \left(\sigma_{a,Er}N_{1Er} - \sigma_{e,Er}N_{2Er}\right)F_L + KN_{2Yb}N_{1Er} - \frac{N_{2Er}}{\tau_{Er}} - CN_{2Er}^2 \quad (2)$$

where N_{1Yb}, N_{2Yb}, N_{1Er}, N_{2Er} are the populations of the $F_{7/2}$, and $F_{5/2}$ levels of Yb and of the $I_{15/2}$, $I_{13/2}$ levels of Er, respectively; $\sigma_{a,Yb}$, $\sigma_{e,Yb}$, $\sigma_{a,Er}$, and $\sigma_{e,Er}$ are the absorption and emission cross sections for the $F_{7/2}$-$F_{5/2}$ Yb pump and $I_{15/2}$-$I_{13/2}$ Er laser transitions; τ_{Yb} and τ_{Er} are the lifetimes of the $F_{7/2}$ and $I_{13/2}$ states of Yb and Er; F_L and F_P are the total laser and pump photon flux; K and C are the energy-transfer and the UC coefficients. The K coefficient, to a first approximation, is assumed to be independent of Er concentration and ranges between 1×10^{-16} and 5×10^{-16} cm^3/s depending on Yb concentration[9]. The C coefficient in phosphate glass was determined to be of the order of 1×10^{-18} cm^3/s.

Equations (1) and (2) are spatially dependent because the pump intensity and the pump and laser mode sizes change both with longitudinal and radial coordinates inside the active material. Nevertheless a simplified solution can be found analytically near threshold. From these calculations some important conditions to minimize threshold pump power were worked out, providing useful guidelines to laser design and optimization. Following an approach similar to that used for analyzing the fiber optical amplifier[15], a numerical solution of the pump and laser field equations can be obtained by an iterative process using a Runge-Kutta method. To this aim, equations (1) and (2) are used to evaluate the local population inversion. The iterative solution can be simplified and the computational time is consistently reduced if we also assume two further approximations: (1) Negligible depletion of ground-state Yb population, so that the pump power is exponentially attenuated. (2) Negligible UC rate compared to the spontaneous decay rate (for $C \cong 1\times10^{-18}$ cm^3/s, we obtain that UC becomes negligible for Er concentration lower than 3×10^{19} ions/cm^3). In this case, in fact, equations (1) and (2) give a second order characteristic polynomial.

3. THE FIBER-PUMPED CAVITY DESIGN

The schematic cavity configuration for these fiber-pumped Er-Yb microlasers is shown in Fig. 2. The active medium is a plane-plane, phosphate glass platelet doped with erbium and ytterbium ions with the first face coated as high-reflective mirror for the laser radiation. The output coupler is a spherical mirror with -20 mm radius of curvature and 1% transmission at 1.5 μm. This mirror is mounted on an annular piezoelectric transducer (PZT) to allow for fine frequency tuning of the laser output. A 50 μm—150 μm thick BK7 etalon is inserted into the laser cavity to achieve single longitudinal mode oscillation whereas an additional Brewster plate will be adopted to steadily polarize the output beam. The pump radiation, coming from a fiber-coupled 1-W 980 nm laser diode (OptoPower, Mod. OPC H01-A001-FC/100), is focused into the active medium with a circular spot of 50 μm (radius) ensuring proper mode matching with the laser mode and hence oscillation on the fundamental transverse mode TEM_{00}. The new cavity configuration also provides for a double pass of the pump radiation within the active medium (being the second facet of the medium HR coated at 980 nm , hence assuring for more uniform pump power absorption and increased average population inversion

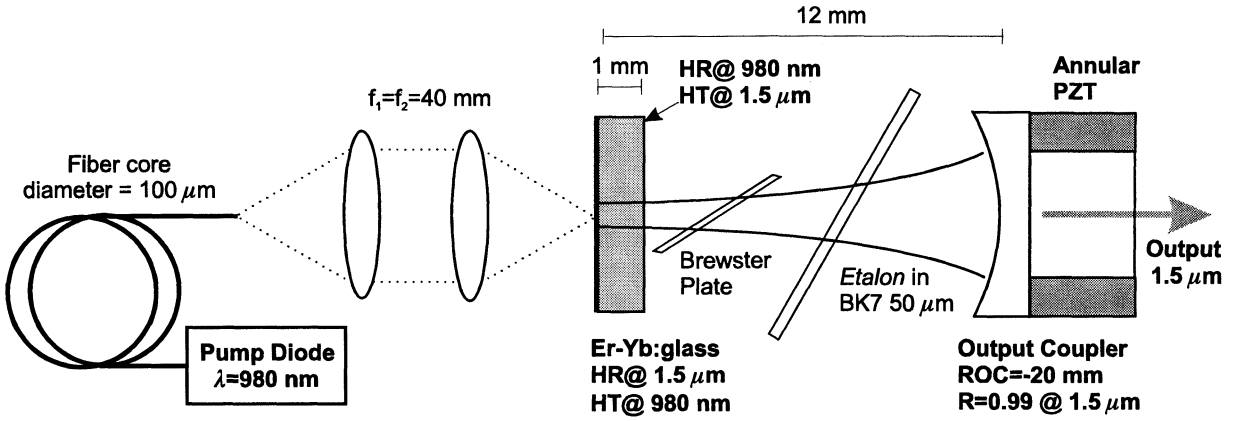


Figure 2. Block diagram of the fiber-pumped Er-Yb:glass microlaser.

for the same length of the active medium (1-mm long as shown in Fig. 2).

Precise drawings of the different mechanical parts needed to mount the optical resonator have been prepared and a self-contained aluminum block was constructed in order to hold the different optical components.

4. LASER CONSTRUCTION

A photograph of the adopted pump system is reported in Fig. 3. The 100 μm optical fiber carrying the pump radiation is shown as pointing to the following focusing optics. A 1:1 imaging system, adopting simple antireflection-coated spherical lenses, was used in order to collect the pump radiation within the 1 mm long Er-Yb doped glass platelet. In order to match the ytterbium absorption peak at 978 nm fine wavelength control of the emission spectrum is achieved by means of temperature tuning the pump diode. A single Peltier element is used to temperature stabilize the case of the OPC laser diode.

The compact one-piece laser cavity is shown in Fig. 4 where the aluminum block only takes up 4.5×5×4.5 cm^3. The entrance aperture is 3×3 mm^2 wide and the Er-Yb:glass platelet is fixed to the aluminum holder by means of a thermal conductive glue with enough flexibility to allow for the glass thermal expansion.

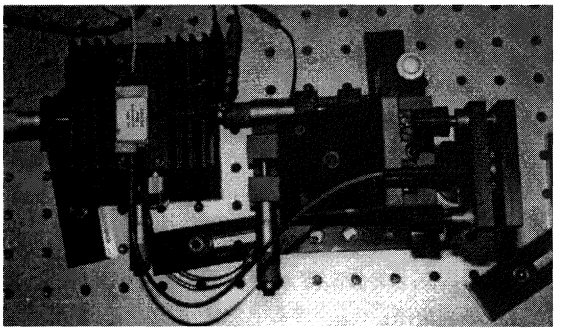


Figure 3. The fiber-coupled pump diode.

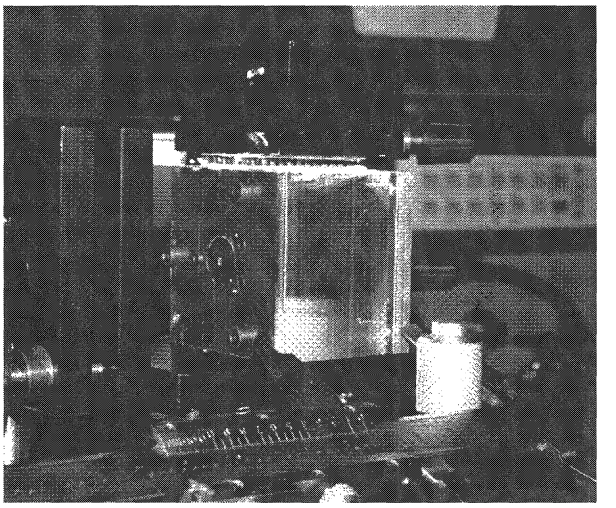


Figure 4. The novel Er-Yb laser.

With this cavity configuration, pump threshold values as low as 85 mW and 45 mW have been obtained with maxima output powers of 60 mW (slope ~12%) and 150 mW (slope ~25%) in single-mode and multi-mode operation, respectively. The spatial quality of the 1.5 μm output beam is diffraction limited with an M^2 value of 1.14. These results were achieved using only 50% of the available pump power in order to avoid thermal damage of the phosphate glass disc.

Recent results in terms of laser portability opened the possibility for interesting international experiments where the narrow-linewidth and high power of these lasers can be exploited.

5. INPUT-OUTPUT CHARACTERISTICS

With the described laser system a powerful output beam with excellent spatial beam quality has been achieved at relatively low pump power levels. Figure 5 depicts a measurement of the output beam transverse profile showing a well-circular TEM_{00} fundamental mode.

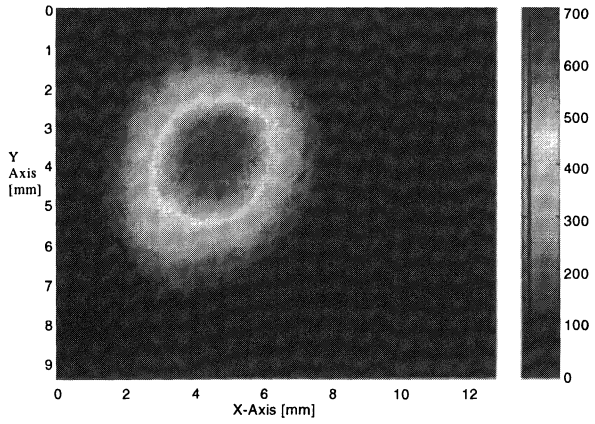


Figure 5. Transverse profile of the 1.5 μm output beam.

A beam divergence of θ_d<29 mrad was measured as well as an M^2 factor below 1.2 (i.e. diffraction limited output beam).

The input-output characteristics are shown in Figs. 6a and 6b where the multimode and single-frequency output power at 1.5 μm is plotted as a function of the incident pump power.

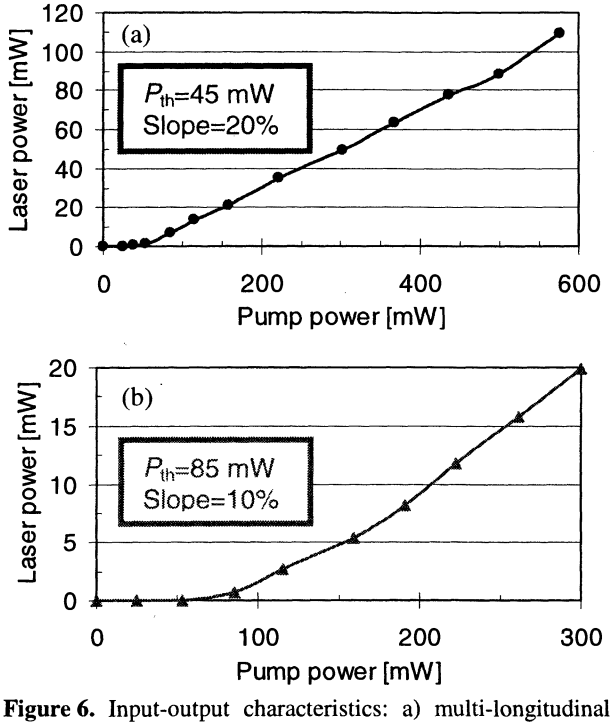


Figure 6. Input-output characteristics: a) multi-longitudinal mode regime; b) single-frequency operation.

The multimode output power reaches the remarkable 100 mW level at approximately 600 mW of pump power and with a slope efficience of 20%. In single-frequency operation a maximum output power of 60 mW was achieved with a pump power of 680 mW. This high pump power level, however, caused breaking the glass after a few hours of continuous operation. Henceforth, in the following experiments the pump power level was limited to 300 mW in order to avoid thermal problems while waiting for the availability of the new strengthened glass that should be able to whitstand up to 1 W of pump power in the adopted beams geomenry. The recorded input-output curve in single-frequency operation shows a 10% slope efficiency and 85 mW of pump power threshold.

6. CONCLUSIONS

A novel fiber-pumped Er-Yb:glass microlaser was developed. The compact cavity design allows for increased reliability and portability of this high-power and narrow-linewidth 1.5 μm laser source. Multimode output powers in excess of 100 mW have been achieved and efficient operation in the single-frequency regime was demonstrated. Power scaling of the present laser device needs the adoption of the lastly available ion-strengthened phosphate glasses. One of the two laser developed on this new design was delivered to the Tel-Aviv university in order to perform spectroscopic measurements and frequency stabilization experiments on atomic samples of potassium (^{39}K D_2 transition at 770 nm) and rubidium (two-photon transition at 778 nm).

REFERENCES

[1] E. Snitzer and R. Woodcock, *Appl. Phys. Lett. 6*, 1965, p. 45.

[2] P. Laporta, S. Longhi, S. Taccheo, O. Svelto, and C. Svelto, "Diode-pumped bulk erbium-ytterbium lasers", *Appl. Phys. B 63*, n. 5, 1996, pp. 425-436.

[3] 'Optical interfaces for multichannel systems with Optical Amplifiers', ITU-T Recommendation G.692, May 1997.

FREQUENCY NOISE AND TUNABILITY MEASUREMENTS FOR DIODE-PUMPED Er-Yb BULK LASERS AT 1.5 μm: OPTICAL FREQUENCY STANDARD BASED ON SATURATED ACETYLENE

C. Svelto, G. Galzerano, A. Onae*, F. Ferrario, and E. Bava

Dipartimento di Elettronica e Informazione – Politecnico di Milano, INFM, 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

*National Research Laboratory of Metrology, 1-1-4 Umezono, Tsukuba, Ibaraki 305, Japan
E-mail: onae@nrlm.go.jp Phone: +81 298 54.4045 Fax: +81 298 54.4135

ABSTRACT

The frequency noise and wavelength tunability of a novel diode-pumped Er-Yb:glass microlaser have been thoroughly investigated. Measurements of the spatial beam quality, linear polarization state frequency and amplitude noise have been performed.

The frequency noise spectral density was measured with a high sensitivity Pound-Drever technique. On one side, the solid-state nature of the laser source combined with a low-noise diode pumping provides for negligible frequency fluctuations above a few hundred kilohertz. On the other side, the quasi-monolithic structure of this novel optical resonator strongly reduces the technical noise effects at low frequencies. In addition, due to a low-pass filtering action taking place in the Er-Yb codoped active medium, also the amplitude noise becomes negligible at Fourier frequencies above the relaxation oscillation frequency (typically 100-200 kHz for these lasers). This results in a very stable optical carrier useful for applications requiring analog amplitude modulation above 0.5 MHz.

1. INTRODUCTION

Miniaturized, practical, and efficient all solid-state lasers operating in the near infrared have become of increasing interest due to their extremely desirable operating characteristics such as high efficiency and high power, oscillation on a single longitudinal and transverse mode with linear polarization, narrow linewidth, enhanced frequency and intensity stability by active feedback loops. In fact, the combination of these characteristics makes these lasers attractive for a variety of applications, including spectroscopy and metrology, laser radars, optical communications and photonics.

Diode-pumped Yb-Er:glass microlasers [1], oscillating at 1.5 μm, are very interesting sources both for optical communications and near-infrared metrology. In fact, these all-solid-state lasers oscillate on a single mode at several milliwatt output power and exhibit very narrow linewidth with extremely wide wavelength tunability compared e.g. to single-frequency laser diodes.

In this paper a throughout characterization of the novel fiber-pumped Er-Yb:glass laser is reported. The frequency and amplitude noise properties of this laser source are described in some detail looking for use of this solid-state laser source in a stable and accurate optical frequency standard at 1.5 μm based on saturated $^{13}C_2H_2$ absorptions

2. Er-Yb LASER CHARACTERIZATION

The novel fiber-pumped Er-Yb:glass laser is shown in Fig. 1 (for further details on the construction of this laser device, refer to the specific paper in these Proceedings). For the first time a Brewster plate has been successfully inserted in the low-gain laser cavity

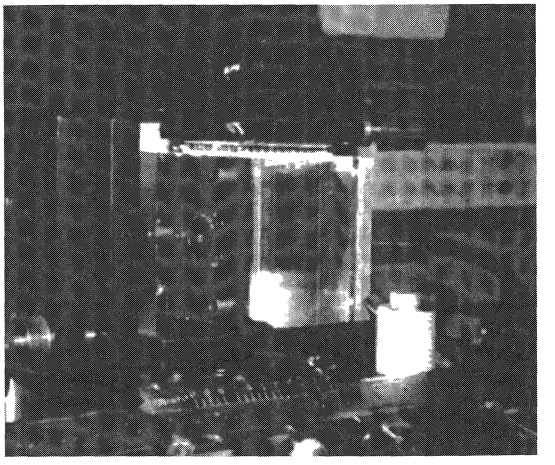


Figure 1. The erbium-ytterbium microlaser.

still allowing for several milliwatt of 1.5 μm output power. Due to the insertion of the Brewster plate within the optical cavity, excellent linear polarization has been achieved as shown in Fig. 2. A linear polarization purity of 632:1 (~28 dB) was measured using a polarizer/analyzer made of a polarizing cube beamsplitter (PBS) with a 30 dB extinction ratio. This means that the laser beam has itself a ~30 dB polarization extinction ratio, which is quite a record for these 1.5 μm optical sources.

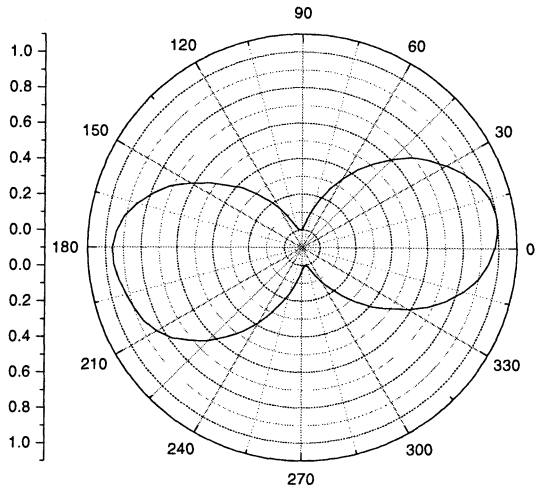


Figure 2. Polar diagram of the measurement data showing the transmitted optical power through the PBS analyzer.

A very wide continuous wavelength tunability of ~1 THz or 7 nm (from 1532 nm to 1549 nm) was measured when tilting the intracavity etalon. Figure 3 shows a typical spectrum of the laser beam showing a single line at an operating wavelength of ~1532 nm. The

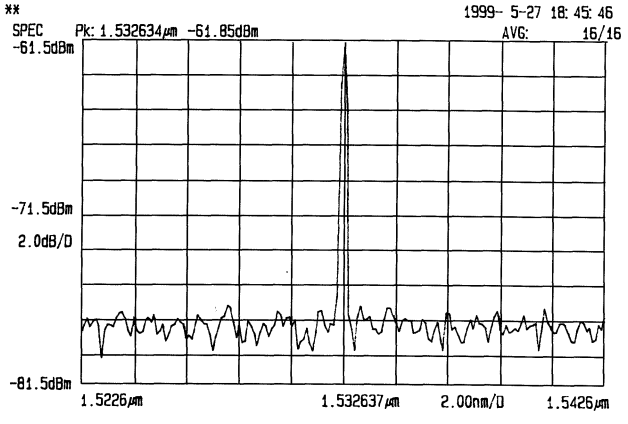


Figure 3. Output spectrum of the Er-Yb laser when operating at 1532 nm wavelength.

typical output power in this wavelength range is in excess of 20 mW with an oscillation linewidth below 50 kHz over an observation time of 1 ms. Fine frequency tuning of the oscillation frequency is achieved by means of an annular piezoelectric transducer used to vary the optical cavity length with a sub-micrometric resolution.

3. FREQUENCY AND AMPLITUDE NOISE MEASUREMENTS

In order to measure the laser frequency noise, a Pound-Drever-Hall frequency stabilization setup was adopted and the voltage signals within this optoelectronic locking loop were analyzed. Figure 4 shows the measured frequency noise spectral density as detected from the error and control signals in the Pound-Drever-Hall stabilization loop. From the control signal, curve labeled (P) in Fig. 4, a noise spectral density with a f^{-2} slope is visible for Fourier frequencies below 20 kHz whereas a higher, yet unexplained, slope is present above 20 kHz.

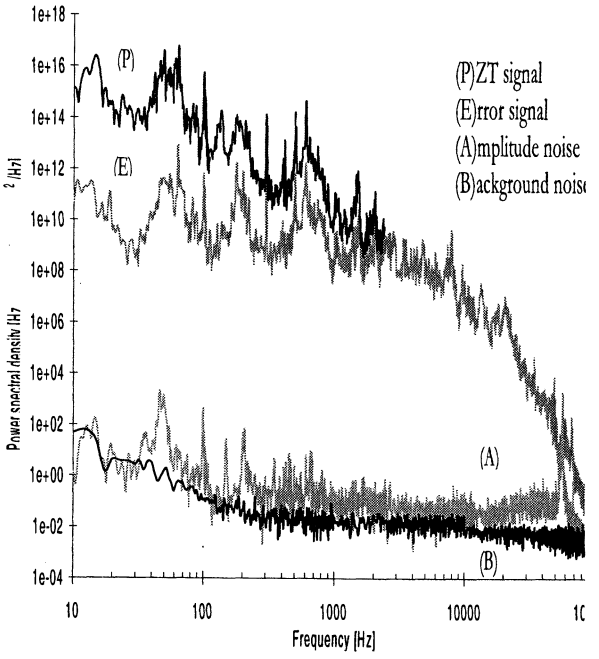


Figure 4. Frequency noise spectral density for the Er-Yb microlaser: (E) for the stabilized laser; (P) for the free-running laser.

In order to determine the laser sensitivity to pump power fluctuations, we experimentally characterized the laser transfer functions between a 980 nm amplitude noise stimulus and the corresponding 1.5 μm output response. A root mean square output signal of 100 mV is the reference output for low Fourier frequencies where the laser response is unitary (i.e. it is equal to the DC slope of the input output characteristic). Figure 5 shows the amplitude response clearly indicating the relaxation oscillation peak located at ~200 kHz. After this peak the amplitude noise decreases with a -40 dB/dec. slope thus resulting in negligible contribution (shot noise limit) for Fourier frequencies above 1 MHz.

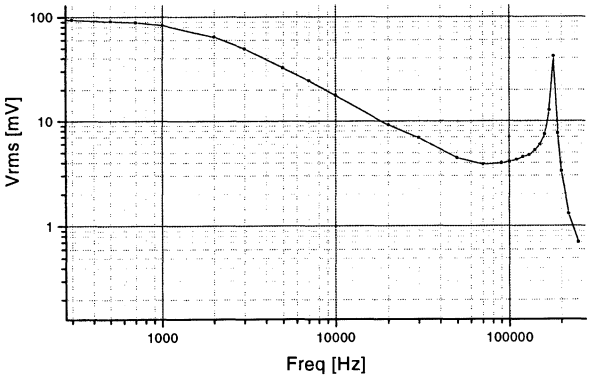


Figure 5. Amplitude response of the laser output, to pump power variations, as a function of the stimulus Fourier frequency.

The measured phase response is reported in Fig. 6 where a -180° phase shift is evident at the relaxation oscillation frequency. The additional −180° delay is due to a pole at a few kHz introduced by the Yb→Er energy transfer process and a high-frequency pole at ~100 kHz.

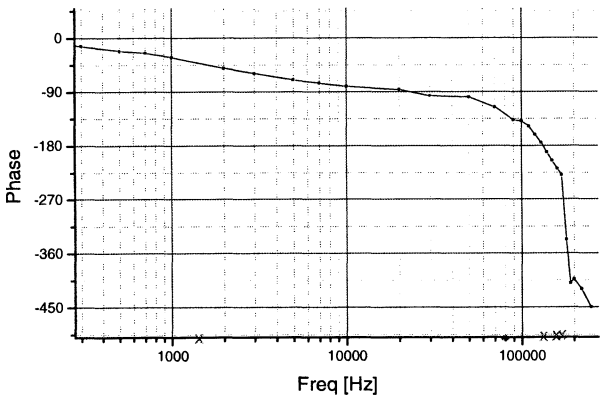


Figure 6. Phase response of the laser output, as a function of the stimulus Fourier frequency, with respect to pump power fluctuations.

4. THE ACETYLENE SATURATION SPECTROSCOPY

In order to achieve narrow reference lines at 1.5 μm, saturated acetylene lines have proved [2] to be a useful candidate. This gas, in fact, is stable and does not show a permanent electric dipole moment allowing for good reproducibility of the absorption frequencies. An accurate optical frequency atlas of the C_2H_2 and $^{13}C_2H_2$ lines at 1.5 μm was recently measured with an accuracy level better than 10^{-9} [2]. Unfortunately, the saturation intensity for the acetylene transitions in this spectral region is quite high, ~1 W/mm^2, and hence a power enhancement by means of resonant optical cavities is typically needed. Using the Fabry-Perot build up cavity depicted in Fig. 7, saturation spectroscopy of $^{13}C_2H_2$ was successfully performed leading to the measurement results of Fig. 8.

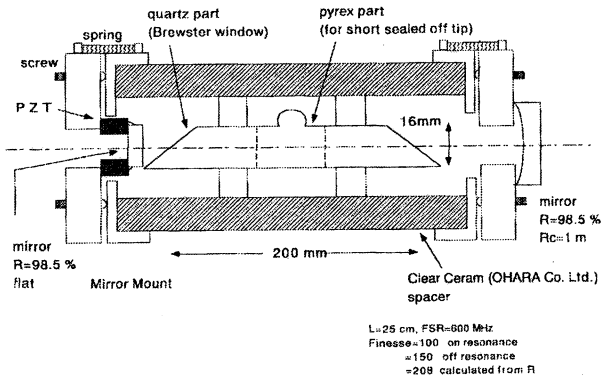


Figure 7. Plano-spherical Fabry-Perot resonant cell containing a Brewster cell with $^{13}C_2H_2$ at ~4 Pa pressure.

In the experiment, the frequency of the Er-Yb laser was locked to a resonance of the Fabry-Perot cavity and the PZT mounted on the Fabry-Perot cell was used to scan the molecular line. The transmitted power trough the gas cell was detected with an InGaAs photodiode. A saturation dip of ~1 MHz width was observed within the ~500 MHz Doppler line.

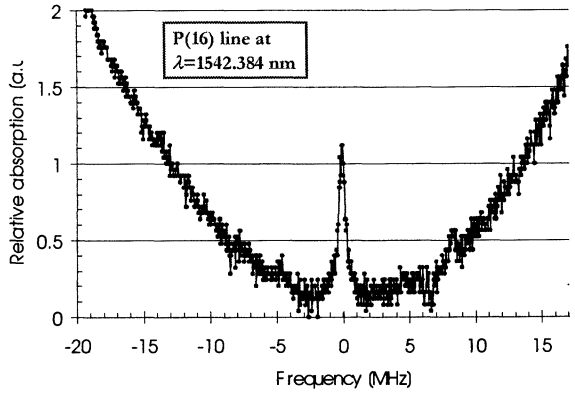


Figure 8. Saturation spectroscopy of the $^{13}C_2H_2$ P(16) line at λ=1542.384 nm. The saturation line contrast is ~3%.

5. CONCLUSIONS

The novel fiber-pumped Er-Yb:glass microlaser was characterized in terms of frequency and amplitude noise. developed. For the first time, the use of a intracavity Brewter window provided for a single linear polarization with a ~30 dB extinction ratio. Using this powerful and narrow-linewidth laser source at 1.5 μm, saturation spectroscopy of the acetylene molecule was performed achieving a saturation dip of ~1 MHz with a 3% line contrast.

Frequency locking of the erbium laser to acetylene Doppler-free absorptions is saturated is in progress and the obtained frequency stability will be experimentally

evaluated in order to establish a new optical frequency standard in this spectral region.

REFERENCES

[1] P. Laporta, S. Longhi, S. Taccheo, O. Svelto, and C. Svelto, "Diode-pumped bulk erbium-ytterbium lasers", *Appl. Phys. B 63*, n. 5, 1996, pp. 425-436.

[2] M. de Labachelerie, K. Nakagawa, Y. Awaji and M. Ohtsu, Opt. Lett. 20, 1995, pp. 572-574 .

[3] K. Nakagawa, M. de Labachelerie, K. Awaji, and M. Kurogi, "Accurate optical frequency atlas of the 1.5-mm bands of acetylene", J. Opt. Soc. Am. B, Vol. 13, No. 12, 1996, pp. 2708-2714.

NEXT DEVELOPMENTS OF TRANSPORTABLE HE-NE/CH4 OPTICAL FREQUENCY STANDARDS

M.Gubin, A.Shelkovnikov, E.Kovalchuk, D.Krylova, E.Petrukhin, M.Petrovsky, D.Tyurikov

P.N.Lebedev Physical Institute, Leninsky pr.53, Moscow 117924, Russia

Several steps for better performance of the double-mode Transportable He-Ne/CH4 Optical Frequency Standards,TOFS, stabilised over Saturated Dispersion (SD) resonances of resolved MHPS of the $F_2^{(2)}$ line (λ = 3,39μm) are in progress.

Previous absolute frequency measurements with the TOFS [1] have shown that these systems demonstrate frequency reproducibility (for different devices) of $2*10^{-12}$ (200 Hz), frequency repeatability (for one device during several months) of $2*10^{-13}$, and frequency stability: $1*10^{-14}$ (for averaging time 10^3s), $1*10^{-13}$ (for averaging time in the range 1s -100s).

The main restriction for the up-today reproducibility from our point of view is connected with *different field structure* in the telescopic part of the standards. (The telescopic part of the TOFS is united with the metnane cell). Now we introduced additional control of the spatial field distribution inside the telescope. We hope that repeatability of the field's form in the telescope will essentially decrease the frequency deviation between new types of the TOFS.

One of the general approaches to improving an accuracy of the standard is *increasing of the optical resonator finesse* in the Telescopic laser. At present the typical value of total resonator losses in this laser is ~50%. Just now it is real to decrease this value in about ten times. It should lead to cardinal changes of optimum parameters of the whole system and in particular to decreasing all types of asymmetry of SD resonances, connected with transverse effects, as all of them are proportional to the total losses.

The new scheme for narrowing a spectrum of radiation. In a previous version of the TOFS a two channel (slow/fast)) feedback frequency loop, narrowing the spectrum of the Reference laser was based on two different physical references. The "slow" channel, (0 – 50) Hz, used a 1st harmonic of modulation of the Saturated Absorption (SA) resonance of $F_2^{(2)}$ CH4 line. The "fast" channel (50 Hz–10 kHz) used a direct signal of the Saturated Dispersion (SD) resonance of the same line. At present a 1st harmonics signal of the SA resonance is substituted for a 2nd harmonics of modulation of the SD resonance. The new scheme, based only on the SD resonance, doesn't lead to deterioration of Allan variance but has some additional advantages. The recording of the SD signal in the intermode beat frequency doesn't need so much power as the recording of SA resonance. Due to that we used more short resonator with higher finesse (the transmittance of the output mirror is changed from 15% to 3%) and the LN cooled highly sensitive photodetector (used in previous TOFS in the "slow" channel) was substituted for small size room temperature photodiodes.

All togerther these improvements promise that the new generation of TOFS started now will permit to decrease dimensions/weight parameters as well as to resolve the recoil doublet of methane line and to reach the repeatability of $\sim 10^{-14}$.

1. O. Acef, A. Clairon, G.D. Rovera, G. Kramer et al, "1998 absolute frequency measurements with a set of transportable He-Ne/CH_4 optical frequency standards",

M. A. Gubin, A S. Shelkovnikov, et al, "Present performance of transportable He-Ne/CH4 frequency standards", in Proceedings of Joint Meeting of the 13th EFTF and 1999 IEEE Intern. Frequency Control Symposium, Besancon, France, April 13-16, 1999.

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

FREQUENCY STABILIZATION OF DBR-DIODE LASER AT 852 nm BY MEANS OF MODULATION TRANSFER METHOD AND SATURATED ABSORPTION OF Cs

F. Bertinetto[(1)], P. Cordiale[(1)], G. Galzerano[(2)], and E. Bava[(2)]

[(1)]Istituto di Metrologia Gustavo Colonnetti - CNR, Strada delle Cacce 73, 10135 Torino, Italy
E-mail: F.Bertinetto@imgc.to.cnr.it Phone: +39 011 3977454 Fax: +39 011 3977459
[(2)]Dipartimento di Elettronica e Informazione – Politecnico di Milano, INFM, and CNR - CSTS
Piazza Leonardo da Vinci 32, 20133 Milano, Italy
E-mail: Gianluca.Galzerano@PoliMI.IT Phone: +39 02 2399.3609 Fax: +39 02 2399.3413

ABSTRACT

Preliminary results on absolute frequency stabilization of Distributed Bragg Reflector (DBR) diode-lasers with emission wavelength at 852 nm will be reported. As a frequency reference saturated absorption D_2 lines of cesium atom have been adopted and to lock the laser frequency against the resonance the non-linear spectroscopy method of Modulation Transfer has been used. From the experimental analysis of the obtained signal to noise ratio a relative frequency stability in the order of $\approx 10^{-12} \cdot \tau^{-1/2}$ should be achieved.

1. INTRODUCTION

Due to technology progresses in the fabrication of semiconductor lasers with low amplitude and low frequency noise levels, diode lasers are very promising optical sources to realize high-accuracy frequency standards in a wide electromagnetic spectrum, ranging from the visible to the near infrared regions [1, 2]. Moreover diode lasers find interesting applications also in high-resolution spectroscopy, as high-sensitivity sensors, and in optical communications.

Experimental results on absolute frequency stabilization of Distributed Bragg Reflector (DBR) diode-lasers with emission wavelength around 852 nm are reported. As a frequency reference saturated absorption D_2 lines of cesium atom have been adopted and to lock the laser frequency against the atomic resonance the non-linear spectroscopy method of Modulation Transfer (MT) has been used [3].

2. DBR LASER AT 852 nm

Semiconductor lasers with good characteristics in term of power, amplitude noise, wavelength tunability, and emission linewidth, are recently available in the Distributed Bragg Reflector (DBR) technology.

To further reduce the emission linewidth and to obtain fine tuning and control of the output laser frequency an optical feedback to the diode laser has to be realized (extended cavity topology).

In our case, the extended cavity configuration reported in Fig. 1 has been used for the DBR laser (mod.) at 852 nm.

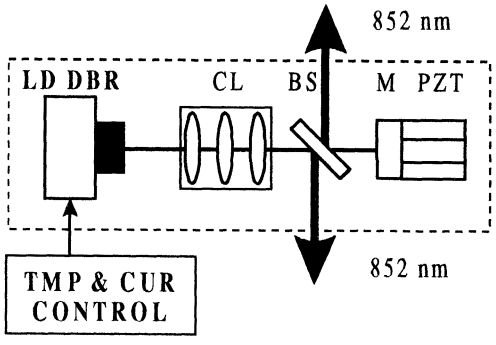


Figure 1. Extended cavity configuration. BS beam splitter, M mirror, CL reshaping optics, PZT piezoelectric actuator.

The output face of the DBR diode is antireflection coated at 852 nm and the extended cavity is closed by a plane mirror with a reflectivity of 95%. In order to reshape the astigmatic diode beam, a focusing element (CL) was placed inside the cavity. The output laser beam comes from the cavity by a 50% reflectivity beam splitter. For tuning the output laser frequency a PZT ceramic was mounted on the cavity mirror to fine control the cavity length.
A photograph of the practical realization of the extended cavity is depicted in Fig. 2.

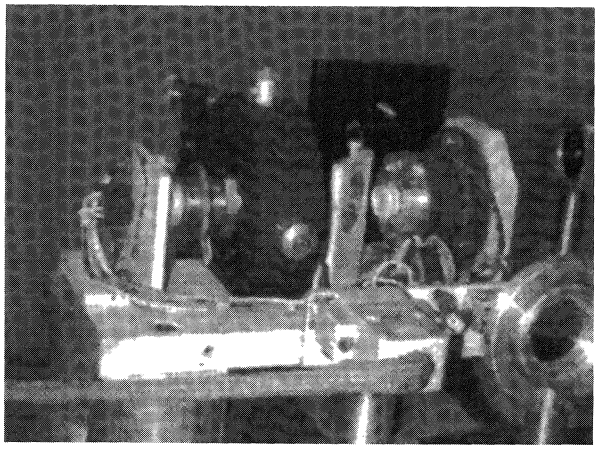


Figure 2. Photograph of the extended cavity practical realization.

At a diode temperature of 32 °C and a pump current of 170 mA an output power in excess of 10 mW has been obtained at an emission wavelength of 852 nm with a linewidth lower than 1 MHz.

3. MODULATION TRANSFER SPECTROSCOPY AND LASER FREQUENCY STABILIZATION

As a frequency reference for laser stabilization at 852 nm the saturated D_2 line of the Cs have been selected. To lock the laser frequency against the D_2 lines the Doppler-free spectroscopy method of the MT has been used.

To describe the MT spectroscopy method a nearly degenerate four-wave-mixing effect approximation can be used [3, 4, 5]. Indeed, by means of the non-linearity of the absorber (the third order susceptibility, $\chi^{(3)}$), the modulation of an intense pump beam can be transferred to a counter-propagating originally unmodulated probe beam [5]. Only when the sub-Doppler resonance condition is satisfied, the modulation transfer takes place and, in this way, the lineshape baseline stability becomes nearly independent from the residual linear absorption effect. The generated sidebands in the probe field and the probe field itself collected to a photodetector combine generating a photocurrent at the modulation frequency (heterodyne detection). Using a phase detection scheme is therefore possible to recover the dispersion and the absorption components of the sub-Doppler resonance related, respectively, to the in-phase and to the in-quadrature demodulated component [3, 4, 5]. These signals are odd functions of the frequency detuning between the laser frequency and the resonance center frequency. When the modulation frequency, δ, is lower than the sub-Doppler linewidth, γ, both the absorption and dispersion signals have a high slope in crossing the center of the resonance (frequency discriminator) and they can therefore be used to accurately lock the laser frequency against the resonance center. For phase modulation index, β, lower than 1, the detected signal MT at the modulation frequency is given by [4]

$$S(\delta) \cong J_0(\beta)J_1(\beta)\sum_{\alpha=a,b}\frac{C_\alpha}{\gamma_\alpha + i\delta}\left\{\begin{array}{l}\dfrac{1}{\gamma + i[(\omega-\omega_o+\Delta/2)+\delta]}+\dfrac{1}{\gamma - i[(\omega-\omega_o+\Delta/2)-\delta/2]}+\\ -\dfrac{1}{\gamma + i[(\omega-\omega_o+\Delta/2)+\delta/2]}-\dfrac{1}{\gamma - i[(\omega-\omega_o+\Delta/2)-\delta]}\end{array}\right\}\exp[i\delta t] \qquad (1)$$

where J_0 and J_1 are the zero and first order Bessel functions, ω-ω_0 is the angular frequency detuning, Δ is frequency shift between the pump and probe beams, γ_α is the decay time of the level α, and C_α is a coefficient related to the angular moment and to the polarization directions of the optical beams. The explicit relation between the in-phase and in-quadrature components and the absorption and dispersion profiles of the resonance is given by

$$S(\delta) \propto J_0(\beta)J_1(\beta)\left[\begin{array}{l}(L_1 - L_{1/2} + L_{-1/2} - L_{-1})\cos(\delta t)+\\ +(-D_1 + D_{1/2} + D_{-1/2} - D_{-1})\sin(\delta t)\end{array}\right] \qquad (2)$$

where L_i is the lorentzian function describing the absorption at the frequency of (ω-ω_0+Δ+δ) and D_i is the first derivative of L_i.

The adopted experimental setup for laser frequency stabilization is reported in Fig. 3. The pump beam is phase modulated by means of an electro-optic modulator at a modulation frequency of 3.1 MHz and a phase index of ~1.5. To reduce the optical feedback at the photodetector, the unmodulated probe beam is frequency shift by 80 MHz using an acoustic-optic modulator. The pump and probe beams were counter-propagating with orthogonal polarizations through the 2-cm long Cs cell with powers of approximately 0.5 mW and with beams diameter of about 2 mm. With these polarization directions the re-emitted beams (Four-Wave Mixing effect [5, 6]) combine with the probe at the photodetector to produce MT lineshape signals with enhanced efficiency only for the $\Delta F=0$ transition. With the usual modulation conditions and at a Cs cell temperature of 20 C, a slope at resonance center of the MT lineshape signal of the order of 0.25 V/MHz has been obtained. In these preliminary results the achieved signal to noise ratio is 25 dB (in a measurement bandwidth of ≈10 kHz). When the laser frequency is stabilized using this signal as an error signal in an optical feedback loop closed at the PZT transducer, the residual white noise should set the optical frequency stability at a level of $\approx 10^{-12}\cdot\tau^{-1/2}$.

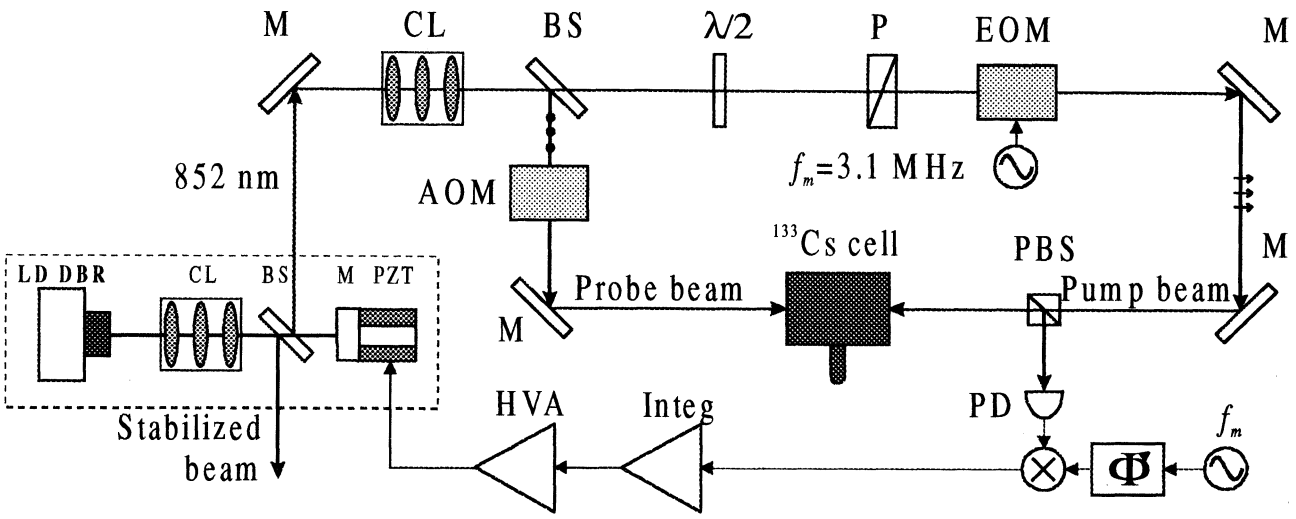


Figure 3. PBS polarizing beam splitter, L lens, P polarizer, λ/2 half wave plate, EOM electrooptic modulator, AOM acousto-optic modulator, PD photodiode, HVA high voltage amplifier, and Φ phase shifter

An example of the MT lineshape signal is reported in Fig. 4. Due to the choice in the polarization directions of the pump and probe beams, the modulation transfer effect takes place only for the principal D_2 line of the Cs, as it is shown in Fig. 4.

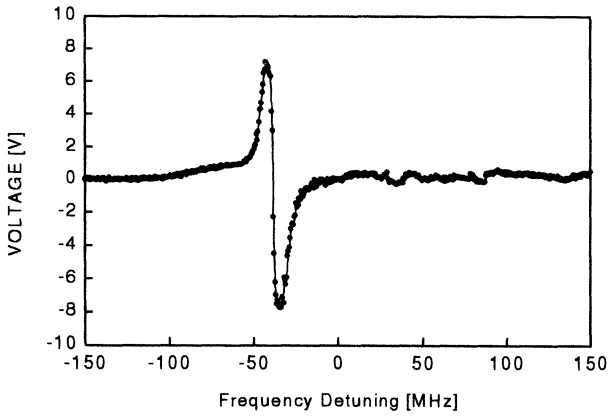


Figure 4. Modulation transfer lineshape recording.

4. CONCLUSIONS AND DEVELOPMENTS

Preliminary frequency stabilization of DBR laser against saturated absorption D_2 line of Cs at 852 nm has been reported. By means of the Modulation Transfer spectroscopy method a signal to noise ratio of 25 dB in a 10 kHz measurement bandwidth has been achieved. Using the modulation transfer lineshape as a frequency control signal, an estimated frequency stability of $10^{-12} \cdot \tau^{-1/2}$ should be obtained.

A second set-up has just been realized and in the next months we will completely characterize the whole system with particular attention to the frequency stability and reproducibility by means of beat frequency measurements between the two independent systems. Moreover, by means of more accurate lineshapes recording (employing RF offset phase-locking technique) the sensitivity and the signal to noise ratios will be characterized for different modulation frequencies and modulation indexes. Furthermore, different configurations of the pump and probe beam polarizations (linear, and circular) will be adopted in order to evaluate modulation transfer efficiency.

To our knowledge, it is the first time that modulation transfer method is used in an experiment of atom spectroscopy.

REFERENCES

[1] H. R. Simonsen, "Iodine-stabilized extended cavity diode laser at λ=633 nm", IEEE Trans. on IM, vol. 46, 141-144, 1997.

[2] Miao Zhu and R. W. Standridge, "Optical frequency standard for fiber communication based on the Rb 5s→5d two-photon transition", Optics Letters, vol. 22, n. 10, 730-732, 1997.

[3] J. H. Shirley, "Modulation transfer processes in optical heterodyne saturation spectroscopy", Optics Letters, vol. 7, n. 11, 537-539, 1982.

[4] Camy, Bordé Ch. J., and Ducloy M., "Heterodyne saturation spectroscopy through frequency modulation of the saturating beam", *Opt. Comm.*, 1982, 41, 325-330.

[5] Ducloy M. and Bloch D., "Polarization properties of phase-conjugate mirrors: angular dependence and disorienting collision effects in resonant backward four-wave mixing for Doppler-

broadened degenerate transitions", *Phys. Rev. A*, 1984, 30, 3107-3122.

[6] M. Ducloy and D. Bloch, "Polarization properties of phase-conjugate mirrors: angular dependence and disorienting collision effects in resonant backward four-wave mixing for Doppler-broadened degenerate transitions", Phys. Rev. A, 30, n. 6, 3107-3122, 1984.

[7] Long-Sheng Ma and J. L. Hall, "Optical heterodyne spectroscopy enhanced by an external optical cavity: toward improved working standard", IEEE J. of Q.E., n. 26, 2006-2012, 1990.

MODULATION-FREE LOCKING OF A DIODE LASER ON THE CS D_2 RESONANCE LINE

N. Beverini, E. Maccioni, P. Marsili, A. Ruffini, F. Sorrentino
Dipartimento di Fisica dell'Università di Pisa, and INFM

Nicolò Beverini, Dipartimento di Fisica dell'Università di Pisa, Via F. Buonarroti, 2, I-56127 PISA (Italy),
Tel +39 050 844 516, Fax +39 050 844 333, E-mail beverini@difi.unipi.it

ABSTRACT

Hyperfine optical pumping of a Cs vapor cell through a diode laser in frequency standards requires a high purity of the laser emission spectrum, and a high stability of the emission frequency. We present here a frequency locking scheme that is able to produce a discrimination signal without any modulation, by using the dichroism induced by a magnetic field on the atomic vapor. The larger line-width of the reference signal (Doppler broadened) is compensated by its larger amplitude, when compared with the saturated absorption signal. As a final result, the slope is similar, and the large line-width warrants a large locking range, making robust the lock against external perturbation. This error signal is used to lock the frequency of an External Cavity Diode Laser (ECDL), with a grating reflector in Littman configuration. A fast correction band is obtained by changing the cavity length through an intra-cavity Electro-Optic Modulator (EOM).

1. INTRODUCTION

The frequency stability of an alkali atom cell frequency standard, optically pumped through a Diode Laser (DL), requires and a high stability and a high spectral purity of the laser emission.

The typical DL jitter can be strongly reduced by optical feedback, using a diffraction grating as a frequency-selective mirror in an Extended Cavity (EC) configuration [1,2]. By this way, a line-width as narrow as some hundreds kHz can be easily obtained over a time scale of the order of few ms. The EC laser is usually locked to the peak of a saturated absorption signal in a reference cell, in order to correct for mechanical and thermal fluctuations at longer times. Frequency stabilization of the EC is performed using conventional FM side-band technique by modulating the laser frequency through an Electro-Optic Modulator (EOM). A broader correction band can be achieved by sending the error signal directly on the laser injection current. In this way the laser line-width can be reduced further at some tens kHz level.

In the following we will present a stabilization system, that uses the dichroism induced by a magnetic field on a reference vapor cell, as demonstrated by Corwin et al. [3] in the case of Rb. This technique gives a dispersion-like frequency error signal, which allows an efficient modulation-free locking of a diode laser to the D_2 resonance line of Cs at 852.1 nm.

2. DIODE LASER FREQUENCY STABILIZATION

It is well known that the frequency of a DL depends on the current and on the temperature, so these two parameters must be carefully controlled. The DL emission width can be reduced by optical feedback from a diffraction grating in Littrow or Littman configuration. The presence of the external cavity can force the DL to work at a frequency that is different from its bare central frequency. Line-widths of the order of one hundred kHz are reached, at short observation time (few ms). Over longer times, various noise sources (acoustic wave propagation, air movements, current instabilities of the DL driver, and mechanical and thermal drifts) contribute to broaden the line-width. A lot of methods have been exploited to stabilize at long term the frequency of EC systems to ultra-stable external reference cavities or to atomic resonance lines. These techniques can also allow a reduction of the fast laser frequency fluctuations, which are responsible for frequency jitter and laser line-width. This goal can be reached if the servo bandwidth is large enough, that is through fast control of the DL current or of the voltage of an intra-cavity EOM.

The typical method for stabilizing the frequency of an ECDL to the D_1 or D_2 alkali lines makes use of the saturated absorption technique. The DL frequency is locked to one of the narrow saturated absorption peaks of an hyperfine component of the transition. A discrimination (dispersion like) signal can be obtained by a frequency modulation of the DL radiation, followed by phase sensitive detection. Conventional FM side-band Pound-Drever method can be adopted [4]: a fraction of the DL radiation travels through an external EOM, which produces frequency side-bands, without introducing serious amplitude modulation in the main laser beam. The saturated absorption signal is detected by a fast photodiode and demodulated in a double balanced mixer. The error signal is fed to the DL current driver for fast frequency control and to a PZT (which changes the length of the EC by moving the diffraction grating or the cavity mirror) for long

term stabilization. The limits of the methods are related to the small capture range, of the order of 20 MHz.

3. EXPERIMENTAL APPARATUS AND DAVLL TECHNIQUE

An alternative method for laser frequency stabilization, called of Dichroic-Atomic-Vapor Laser Lock (DAVLL), was originally suggested for Rb by K. L. Corwin et al. [3]: we have adapted this method to the case of the Cs atom. The ECDL is a Littman configuration device, with a SDL 850-nm, 150 mW single mode DL as active source. The light beam is collimated, and sent at a large angle onto a diffraction grating. A PZT mounted mirror reflects the first order of diffraction back to the DL and can be employed for frequency tuning and for low frequencies corrections. The zero order of diffraction constitutes the output of the ECDL (about 30 % of the total power). An intra-cavity EOM, placed between the grating and the mirror, is used as fast actuator for the error signal. All the optical parts are positioned on a glass base, to reduce thermal drifts. The glass sheet lies on a sand layer to attenuate vibrations and the whole apparatus is enclosed in a double perspex-wooden box to shield from air turbulence and from acoustic waves.

A small fraction of the output power is sent onto the reference Cs cell, set in a uniform magnetic field

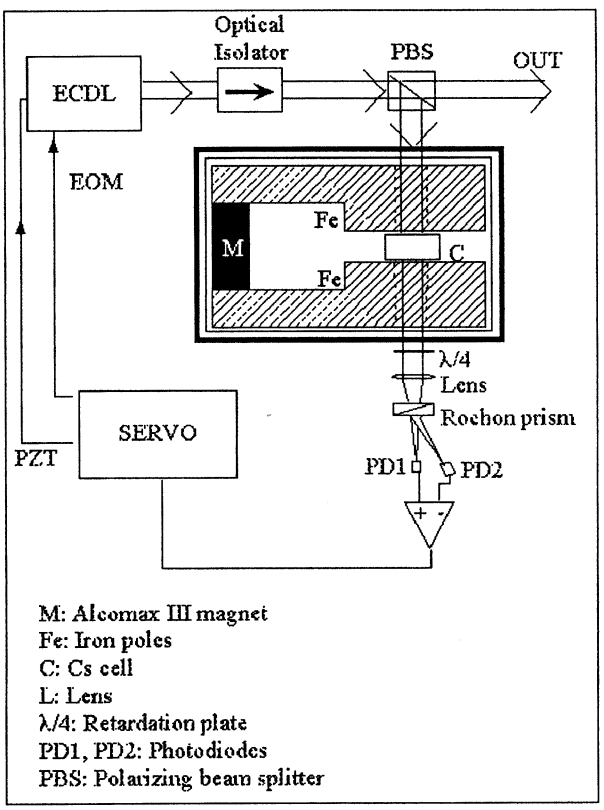


Fig. 1 – The apparatus for DAVLL technique.

collinear with the light. The linear polarization of the light can be considered as the coherent sum of σ^+ and σ^- circular polarization states of equal intensity. A Doppler-broadened profile is obtained when the laser beam is sent onto the Cs vapor cell and the frequency is scanned across an atomic resonance transition. The magnetic field splits the absorption of σ^+ and σ^- polarization components in the opposite directions. A discrimination dispersion-like signal is obtained by differentiating the two absorption profiles and can be used for locking purposes. In spite of the much larger line-width of this signal as compared with the saturated absorption signals (about 500 MHz and about 20 MHz respectively for the two cases), its slope is almost the same, due to the fact that the amplitude of the saturated absorption features is only a small fraction of that of the Doppler-broadened absorption. The large line-width of the error signal warrants a large locking range, making the lock robust against external perturbations. As we will see in the following, the stability of this kind of atomic reference depends on the temperature stability of the Cs cell, on the intensity of the magnetic field, and on the birefringence induced by the optical elements. In Fig. 1 the reference apparatus is shown: the magnetic field is produced by two permanents magnets of Alcomax III and the magnetic circuit is closed by two iron poles. In the center (position of the cell) the field is about 14 mT: this value has been chosen after an experimental study to maximize the slope of the discrimination DAVLL signal. At the same time we have observed a sensitivity of the locking frequency point of a DAVLL laser from the magnetic field (without special care in controlling the input polarization) of the order of 20 kHz/μT. The circuit and the cell are placed inside three concentric boxes of aluminum, polystyrene and wood in order to protect them from thermal variations and from acoustic perturbations. An iron box, which shield the external magnetic field perturbations, encloses the whole device. The Cs cell temperature is thermalized around 43±0.1 °C, a value chosen to have a good vapor pressure and with a precision high enough to have no observable variations on the transmitted light level. The light transmitted by the cell is finally sent onto a circular polarization analyzer, composed by a λ/4 retardation plate, followed by a polarization beam splitter, with the axis at 45° with respect the retardation plate axis, like a Wollaston or a Rochon prism. The two polarization signals are differentially detected, and the resulting dispersion-like discrimination signal is used for DL locking at the central zero crossing. An experimental DAVLL line-shape, obtained by tuning the frequency around the D_2 component leaving from F=4 ground hyperfine level is shown in Fig. 2. The zero in the frequency scale is set in correspondence of the F=4 → F'=5 hyperfine component. The Cs vapor temperature was set at 43 °C.

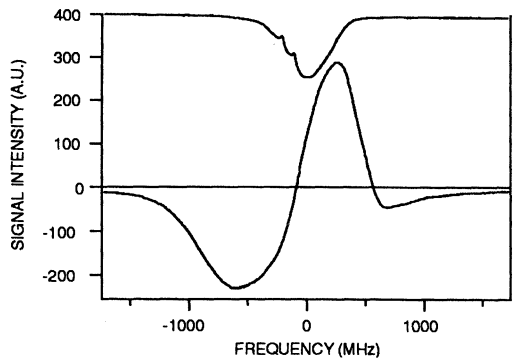


Fig. 2 – Experimental DVALL signal. The upper trace is a Doppler zero-field absorption profile with saturation Lamb-dips, recorded for calibration on a second Cs cell.

It is possible to tune the locking point inside the capture range either by adding an electronic offset or by rotating the quarter-wave plate. This second optical method acts on the relative weight of one circular polarization in comparison with the other, and is more stable than the electronic offset. In this way the lock point is virtually insensitive to laser light intensity variations. However, some disturbance (small fluctuations of the offset level and appearance of asymmetries on the DAVLL signal) may arise from the temperature-induced birefringence on the various optical elements that must be protected as better as possible from environmental temperature drifts. In particular the use of good quality zero-order retardation plate is essential, in order to avoid large locking point instabilities.

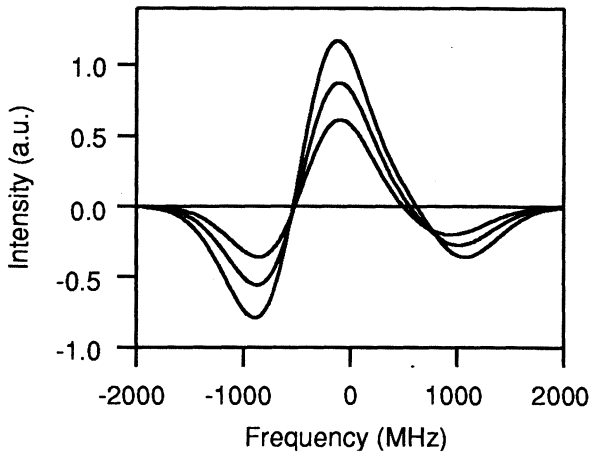


Fig. 3 – Computed DAVLL signals for three increasing values of the magnetic field: 12 mT, 13 mT, 15 mT.

In order to evaluate the effect of the different parameters to the locking point, we have numerically computed the line shape of the DAVLL signal. In a well aligned experiment, where the input radiation is perfectly linear polarized, and the detector is equally sensitive to the two circular polarizations, the expected absorption coefficient is shown in Fig. 3 as a function of the laser frequency, for different values of the magnetic field. The zero in the frequency scale is set in correspondence of the $S_{1/2} \rightarrow P_{3/2}$ transition, when Zeeman and hyperfine interactions are not considered. The field value that optimizes the error signal slope lies around 15 mT. In this hypothesis, the lock point is not sensitive to the laser intensity fluctuations, but demonstrates a weak sensitivity to the field value (of the order of 5 kHz/μT) and to the cell temperature, through the absorption Doppler width (of the order of 300 kHz/°C).

4. FREQUENCY STABILITY OF THE DAVLL SYSTEM

The servo loop is composed of two separated stages that control two different actuators: the PZT mounted cavity mirror, which corrects frequency drifts and low frequency noise (up to about 500 Hz), and the intra-cavity EOM for the highest frequency. Fig. 4a shows a measure of the spectrum of the error signal E(ν) in the two cases of open and closed loop. E(ν)is defined as the Fourier transform of the instantaneous laser frequency shift from the locking point. In Fig. 4b we show the loop spectral gain G(ν)=O(ν)/E(ν), where O(ν) is the Fourier transform of the quantity p(t)+e(t), sum of the frequency shifting of the PZT and of the EOM. The bandwidth is presently limited to 40 kHz, due to electronic limitations in the preliminary detection circuit. New servo electronics has been projected, to extend the correction rate up to the maximum value allowed by the response characteristic of the EOM (about 1 MHz).

An evaluation of the frequency drift of the system has been made observing on a spectrum analyzer the beat note between the frequencies of the DAVLL laser and of an EC laser in Littrow configuration. The DAVLL device was locked on the Cs D_2 line, while the Littrow laser was locked by means of a conventional FM technique on a saturated absorption peak of an hyperfine component of the same line (cross-over F'=4 – F'=5 of the transition F=4 – F'= 3, 4, 5). In the actual apparatus, the radiation from the Littrow system is frequency modulated, and it is thus impossible to obtain accurate evaluation of the short term stability (the laser frequency jitter can be roughly estimated to be less than 200-300 kHz). In any case, we can state a long term maximum drift lower than 1 MHz over a 20 hours period.

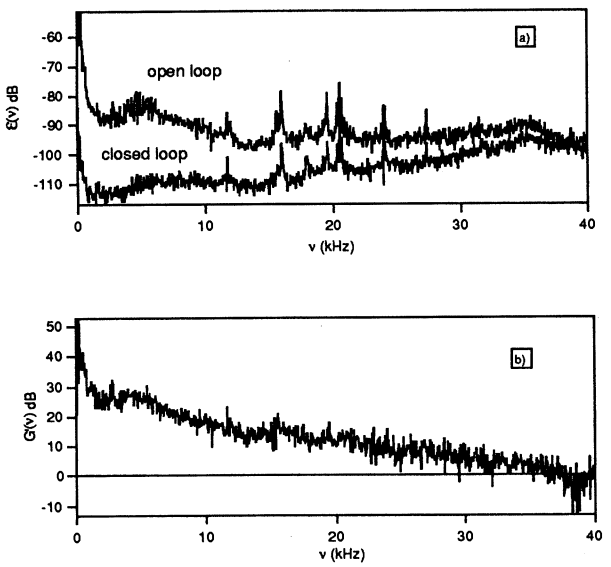


Fig. 4 – a) Open and closed loop error signal frequency spectrum; b) Spectral gain of the loop

5. CONCLUSION

We have applied the DAVLL technique to the stabilization of an EC Littman configuration diode laser, providing an atomic reference for the D_1 and D_2 Cs resonance lines. This technique is particularly robust against mechanical perturbations, because of its large capture range. Besides it offers a series of advantages in comparison with the saturated absorption methods, requiring less optical and opto-electronic components and no FM of the DL current.

This is a simple and inexpensive atomic reference scheme that is well suitable to realize a high stability diode laser source to be used for hyperfine optical pumping of a Cs vapor cell. Work is now in progress to better evaluate the frequency stability of the device over long time periods, and to increase the servo-loop bandwidth up to the MHz level. In this way it will be possible to improve the Signal to Noise ratio of the discrimination error signal and to reduce the laser source frequency jitter.

We thank M. Francesconi for the realization of the opto-electronic servo-loop.

6. REFERENCES

[1] C.E. Wieman and L. Hollberg, "Using diode lasers for atomic physics", Review of Scientific Instruments, vol. 62, pp. 1-20 , 1991.

[2] R.W. Fox, A.S. Zibrov, and L. Hollberg, "Semiconductor diode lasers", in Atomic, Molecular and Optical Physics, vol. III, (Electromagnetic Radiation), eds. F.B. Dunning and R.G. Hulet, Academic Press, 1995.

[3] K.L. Corwin, Z. Lu, C.F. Hand, R.J. Epstein, and C.E. Wieman, "Frequency-stabilized diode laser with the Zeeman shift in an atomic vapor", Applied Optics, vol. 37, No. 15, pp. 3295-3298, May 1998.

[4] R.W. Drever, J.L. Hall, and F.V. Kowalski, "Laser phase and frequency stabilization using an optical resonator", Applied Physic B, vol. 31, pp. 97-105, 1983.

14th European Frequency and Time Forum, 14-16 March 2000, Torino (Italy)

CESIUM REPUMPING-LASER FREQUENCY STABILIZATION USING CROSS-SATURATION

W.X. Ji[(1)*], B.Q. Pan[(1)], Chr. Tamm[(2)], S. Weyers[(2)]

[(1)]National Institute of Metrology (NIM), Beijing 100013, China;
[(2)]Physikalisch-Technische Bundesanstalt (PTB), Braunschweig, Germany
*Present address: Labor 4.32, PTB, D-38116 Braunschweig, Germany.
Email: jiwx@mailserv.buct.edu.cn; guest432@ptb.de.

ABSTRACT

Atomic frequency standards based on laser cooled cesium atoms typically require frequency-stabilized laser excitation of two hyperfine components of the Cs 6s $^2S_{1/2}$ - 6p $^2P_{3/2}$ resonance transition for cooling and repumping. In this contribution we describe a new advantageous technique to obtain an atomic reference signal for stabilization of the laser to the Cs repumping transition. A laser beam resonant with the F=4 - F'=5 cooling transition is passed through a room-temperature Cs cell. We observe the cross-saturation absorption signal that results if the frequency of a counter-propagating superimposed laser field is scanned through the F=3 - F'=4 repumping transition resonance. The observed cross-saturation absorption signal is significantly less perturbed by linear absorption than the conventional saturated-absorption signal. Our experimental observations indicate that similar Doppler-free cross-saturation signals can also be obtained if the saturating laser field is not tuned to the F=4 - F'=5 transition but to other hyperfine transitions or crossover resonances.

1. INTRODUCTION

The saturated absorption spectroscopy of cesium has been well studied[(1-2)]. The application of saturated absorption is a commonplace in many laboratories for high resolution spectroscopic studies and as a frequency reference. Such references are widely used in many active research fields such as quantum optics, frequency standards etc. because of the development of diode lasers and narrowband laser sources technology.

Atomic frequency standards based on laser cooled cesium atoms typically require frequency-stabilized laser excitation of two hyperfine components of the Cs 6s $^2S_{1/2}$ - 6p $^2P_{3/2}$ resonance transition. While the F=4 - F'=5 component of this transition is excited for cooling, the F=3 - F'=4 component is excited for hyperfine repumping. Conveniently the frequencies of the employed lasers are stabilized to these transitions using saturated-absorption signals from Cs cells as frequency references. However, frequency stabilization to the F=3 - F'=4 saturated-absorption signal has the disadvantage that this signal is relatively weak and appears superimposed on a strongly frequency-dependent linear-absorption background.

In this contribution we describe a new advantageous technique to obtain an atomic reference signal for laser stabilization to the Cs repumping transition. The observed cross-saturation absorption signal is significantly less perturbed by linear absorption than the conventional saturated-absorption signal. Our experimental observations indicate that similar Doppler-free cross-saturation signals can also be obtained if the saturating laser field is not tuned to the F=4 - F'=5 transition but to other hyperfine transitions or crossover resonances. We give a simple theoretical interpretation of these observations.

2. EXPERIMENTAL SETUP

The experimental setup is similar to the standard configuration for the hyperfine repopulation pumping with two diode laser as shown in Figure 1. The two diode lasers are home-made extended-cavity diode lasers. The linewidth of the lasers is below 1 MHz. The signal observed is from photodiode I, while photodiode II provides a reference signal in our experiment. The power of the two beams is at the level of 0.5 mW. The length of the cesium cells covered with magnetic shield is about 2 cm with a diameter of 2cm.

The two laser beams are carefully aligned so as to overlap with each other over the whole length inside the cesium cell. The two windows of the cell are non-perpendicular to the laser beams for avoiding possible

perturbations due to the reflection of the laser beams from the windows.

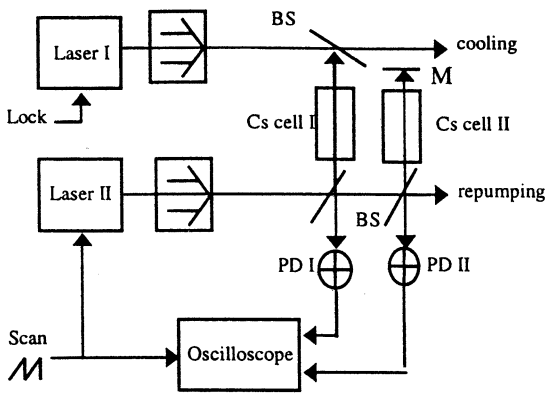


Fig.1. Experimental setup for the cross-saturation spectroscopy. While the probe beam splitted from laser I is locked at the cooling transition $^2S_{1/2}(F=4)$ to $^2P_{3/2}(F'=5)$, an absorption signal from PD I is observed by scanning the frequency of Laser II over the transitions of $^2S_{1/2}(F=3)$ to $^2P_{3/2}$ (F'=2,3,4). The signal from PD II is used as a reference to determine the position of the peaks in the signal of PD I. In the figure, PD denotes photodiodes, BS beam splitters, M mirror.

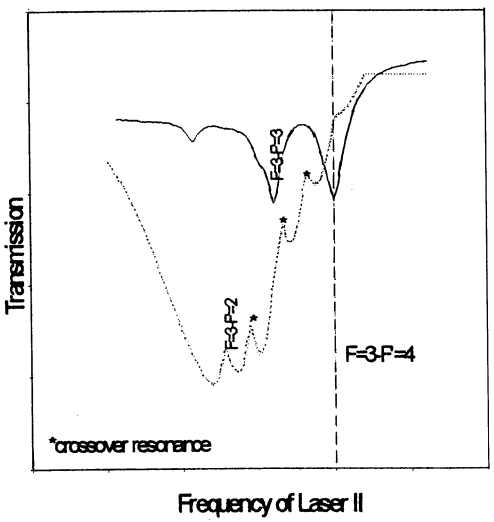


Figure 2. Cross-saturation and conventional saturation signals. The transmission signal is observed if the frequency of laser I is locked at the transition F=4 - F'=5, and the frequency of laser II is scanned over the range of F=3 - F'=2, 3, 4 components of the 852nm Cs $^2S_{1/2}$ - $^2P_{3/2}$ transition. The dotted line denotes the conventional saturation signal from PD II, the full line is the cross saturation spectrum signal observed from PD I. The vertical dashed line shows the position of the F=3 - F'=4 transition.

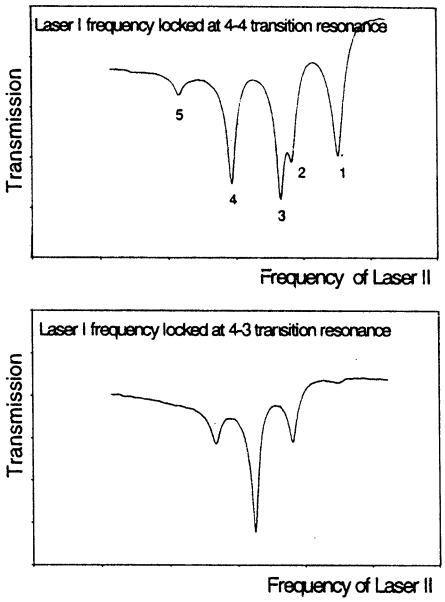


Fig. 3. Other cross saturation signals observed by locking the probe beam at different hyperfine or crossover transitions. The full line is the cross saturation spectrum from PD I, while the dotted line is the conventional F=3 - F'=2, 3, 4 saturation absorption spectrum from PD II.

3. OBSERVATIONS AND THEIR INTERPRETATION

The spectrum shown in figure 2 was recorded under the standard laboratory conditions. The negative peaks mean enhanced absorption. The reason for the enhanced absorption is the population accumulation due to the pumping effect of Laser II. More atoms at level F=4 provide more effective absorption of the probe beam of Laser I. There is significantly less background disturbance because the probe beam is locked at a fixed frequency.

Figure 3 shows the spectra observed when Laser I is locked at other different resonant transitions and crossover resonances of the Cs $^2S_{1/2}$ (F=4) - $^2P_{3/2}$ (F'=3, 4, 5) transition. The position of the peak can be interpreted with the Doppler effect of different velocity groups. Taking the spectrum obtained by locking Laser I at the frequency of the 4 - 4 transition resonance for instance, we can see that the two peaks on the right hand (1 and 2) are shifted from the two peaks nearby (3 and 4) by 251 MHz. This is because at this velocity group, the atoms see the Laser I at the transition of 4 -

5. The frequency difference between the two peaks 1 and 2 is the same as that between peaks 3 and 4, namely 201 MHz.

4. CONCLUSION

A new experimental setup was demonstrated for to obtain strong and unshifted absorption structure which can be used for locking a laser to the F=3 - F'=4 repumping transition. This transition is needed in laser cooling experiments. The scheme is very simple and makes effective use of the F=4 - F'=5 radiation which is available in the optical setup anyway. Some new spectra were observed with the new experimental setup. A simple interpretation of the observed spectra was given. The interpretation is based on the population pumping effects and the Doppler effect, which predicts rather good the position and the sign of the observed resonances.

5. ACKNOWLEDGEMENT

W. X. Ji gratefully acknowledges the financial support under the agreement between CSBTS and PTB, and the financial support for participating the conference from NIM.

6. REFERENCES

1.O. Schmidt, K. -M. Knaak, R. Wynands, D. Meschede, 'Cesium saturration spectroscopy revisited: How to reverse peaks and observe narrow resonances', Appl. Phys. B 59, pp167-178(1994) and references therein.

2.GAN Jianjua, CHEN Xuzong, LI Yimin, JI Wangxi, HUA Jingshan, YANG Donghai, WANG Yiqiu, 'The dependence of Cs saturated absorption spectra on the intensity of the pumping light', Acta Physica Sinica Vol.45, No.10, October, 1996 and references therein.

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 Jacquet-Draz 1 / Case Postale 20
CH-2007 Neuchâtel, 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 50 05 69
Fax : + 33 3 81 88 57 14

e-mail: isabelle.coutrot@ens2m.fr

AUTHOR INDEX

Conference Secretariat

Logistic Organization
Elisabetta Melli
IEN Galileo Ferraris
Strada delle Cacce, 91
10135 Torino (Italy)
Tel. +39 011 3919 524
Fax +39 011 346384
e-mail: eftf@ien.it

Scientific Organization
Antonella Alotto
Politecnico di Torino
Dipartimento di Elettronica
Corso Duca degli Abruzzi, 24
10129 Torino (Italy)
Tel. +39 011 564 4124
Fax +39 011 564 4124
e-mail: eftf@polito.it