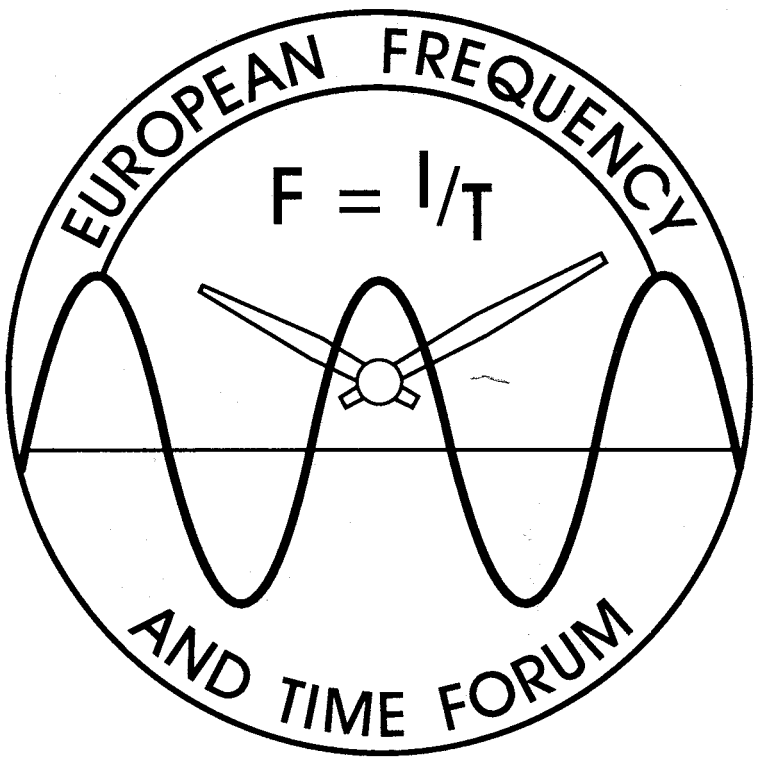
esa SP-340

Proceedings of the 6th European

FREQUENCY AND TIME FORUM



Noordwijk, The Netherlands
17–19 March 1992

Organised by:

The European Space Agency (ESA)

esa
european space agency

esa SP-340
June 1992

Proceedings of the 6th European

FREQUENCY AND TIME FORUM

Noordwijk, The Netherlands
17–19 March 1992

Organised by:

The European Space Agency (ESA)

european space agency / agence spatiale européene
8-10 rue Mario-Nikis, 75738 PARIS CEDEX 15, France

Scientific Committee *(Chairman:* S. Feltham, ESA/ESTEC)

C Audoin
A. Ballato
A. Bauch
J. Beaussier
R. Besson
R. Boloix
M. Brunet
G. Busca
L. Cutler
A. De Marchi
J. Dowsett
H. Duchaussoy
M. Ecabert
J.-J. Gagnepain
E. Graf
M. Granveaud
B. Guinot
H. Handrek
P. Hartl
J. Henaff
C.K. Hruska
K. Kalliomäki
P. Kartaschoff
D. Kirchner
D. Knight
P.W. Krempl
S. Leschiutta
C. Maerfeld
K. Nakagiri
G. Nard
B.W. Neubig
P.W. Nieuwjaar
K. Nottarp
B. Oehman
L. Prost
A. Rémond
J. Rutman
B. Schlueter
S. Starker
J. McA. Steele
P. Thomann
R. Unverdross
J. Vanier
F.K. Von Willisen
F.L. Walls
G. Winkler
N. Yannoni
W. Zingg

1992 Local Organisation Committee

S.J. Feltham *(Chairman)* ESA/ESTEC
J.J. Hunt *(Editor)* ESA/ESTEC
G. Elfering (ESTEC Conference Bureau)
R. Guerrand ESA/ESTEC

ESA SP-340

6TH EUROPEAN FREQUENCY AND TIME FORUM
ESTEC, Noordwijk, The Netherlands, 17 – 19 March 1992

Proceedings published by: ESA Publications Division, ESTEC, Noordwijk, The Netherlands

Edited by: J J Hunt

ISSN 0379-6566 *ISBN 92-9092-202-8* Price code: E4

Copyright © 1992 European Space Agency

Printed in The Netherlands

Contents

Foreword 1
S. Feltham, ESA/ESTEC

OPENING SESSION:
Chairman: S. Feltham

Welcoming Address 5
M. Le Fèvre, Director, ESTEC

Keynote address: The Advances in F&T Technology– Their Benefits of and for Space Science and Technology 9
P. Hartl, University of Stuttgart, D

SESSION 2: SPACE APPLICATIONS
Chairman: C. Audoin

Hydrogen Masers for Space 19
R. Vessot, Smithsonian Astrophysical Observatory, USA, &al

A Simple and Compact Source of Cold Atoms for Cesium Fountains and Microgravity Clocks 27
A. Clairon, LPTF, F, &al

Microwave Frequency Discriminator with a Cryogenic Sapphire Resonator for Ultra-low Phase Noise 35
G.J. Dick & D.G. Santiago, JPL, USA

The Application of the PRARE Principal for Precise Two-Way Time Transfer to Space-based Atomic Clocks 41
W. Schäfer, University of Stuttgart, D, &al

SESSION 3A: TIME AND FREQUENCY: FUNDAMENTALS
Chairman: R. Besson

Atomic Time Scales for Dynamical Astronomy 51
B. Guinot, Bureau Int. des Poids et Mesures, F

How Can Millisecond Pulsars Improve the Long-term Stability of Atomic Time Scales? 57
G. Petit, BIPM, F, &al

Relativistic Effects for Low-Earth Orbit Satellites using GPS 61
A. Spallicci, ESA/ESTEC, NL,&al

The Columbus Metrology Science Team: Report on Activities 67
A. Spallicci, ESA/ESTEC, NL

Prospects of Two-Way International Time Scale Comparisons 69
G. de Jong, Van Swinden Laboratorium, NL

SESSION 3B:
ATOMIC FREQUENCY STANDARDS: OPTICALLY PUMPED CESIUM
Chairman: L. Cutler

Preliminary Results on an Optically-pumped CS Beam Frequency Standard 77
P. Aynié, LPTF/SERT, F &al

Miniature Optically-Pumped Cesium Beam Resonator 83
P. Petit, CNRS, &al, F

Actual Limits of Fluorescence s/n Ratio in an Optically-Pumped Cesium Beam Resonator 87
N. Dimarcq &al, Lab. de l'Horloge Atomique, F

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

SESSION 4A:
QUARTZ: RESONATORS
Chairman: W. Zingg

Long-term Drift and Gamma Irradiation Effect 93
J.J. Boy, LCEP/ENSMM, F, &al

Weak Magnetic Field Sensitivity of Quartz Crystal Oscillators 99
R. Brendel, CNRS, & E. Robert, CNES, F

Synchrotron Radiation X-ray Topography Study of Lateral Field Resonators 105
B. Capelle, LMCP, &al, F

Improved PI-Network Measurement of all Quartz Crystal Parameters by appropriate Modelling of the Test Fixture 111
B. Neubig, Tele Quartz Group,& L. Schulzke, KVC, D

Measurement of the Resonant Frequency of Quartz Crystal Units: An Ongoing 'Saga' 119
E. Kentley, Roditi International

SESSION 4B:
TIME AND FREQUENCY: SPACE APPLICATIONS
Chairman: P. Hartl

DORIS Precise Orbit Determination and Location System Performances of Ultra Stable Oscillators 125
M. Brunet, CNES, F

Precise Doppler Tracking from the Medicina VLBI Station 137
R. Ambrosini, CNR, &al, I

A New Modem for Microwave Time Synchronisation via Geostationary Telecommunication Satellites 143
M. Dienert, TimeTech, D

Synchronisation, Acquisition and Tracking for Telemetry and Data Reception 149
A. Van Doninck, Alcatel Bell

SESSION 5A:
QUARTZ: OSCILLATORS
Chairman: E. Graf

A Microcomputer-compensated Crystal Oscillator MCXO 159
P. Collet &al, Composants Quartz et Electronique, F

Oscillator-Compensated with a Capacitance Array (OSCCAR) 165
W. Daniau & G. Marianneau, CNRS, F

Design of a Low-phase Noise UHF Oscillator using a Partly Metallised Dielectric Resonator 169
R.E. Elcheikh & M. Oliver, CNRS, F

Miniaturisation and Stability: a New Generation Ovenised Oscillators 173
J.P. Aubry &al, CEPE, F

SESSION 5B:
TIME AND FREQUENCY: DISSEMINATION
Chairman: D. Kirchner

A System Design for Laser Time Synchronisation via Geodetic Satellite 'AJISAI' 181
H. Kunimori, Comm. Res. Laboratory, Japan, &al

The Accuracy of SYLT and RNS-based Time Scales at MTI and SNT during 1990-91 187
T. Mansten, MTI, Finland, &al

Control of Environmental and Clock Parameters in Time and Frequency Standards Laboratory 191
K. Kalliomaki, Univ. of Oulu, & T. Mansten, MTI, Finland

Analytical Techniques for Improved Time Transfer using Common View of Direct Broadcasting Satellites 195
J.A. Davis & P.R. Pearce, NPL, UK

SESSION 6A:
QUARTZ: FILTERS AND SENSORS
Chairman: B. Neubig

Elastic Plate Mode Sensitivities to Mass Loading Applications to Gas Sensors 203
M. Hoummady & D. Hauden, CNRS, F

A State-of-the-art Crystal Filter Design Package using Windows 3.0 Techniques 209
M. Frey & B.W. Neubig, Tele Quartz, D

Monolithic Quartz Vibrating Beam Accelerometer 217
F. Deyzac &al, ONERA, F

Surface Transverse Wave (STW) Quartz Resonators in the GHz Range 219
E. Bigler, CNRS, F, &al

Comparison of Energy Trapping Resonators and Filters using the AT Cut of Quartz, Berlinite and Gallium Phosphate 223
J. Detaint, CNET, &al, F

SESSION 6B:
TIME AND FREQUENCY: GPS AND GLONASS ACTIVITIES
Chairman: R. Lucas

Rubidium Atomic Frequency Standards for GPS Block IIR 231
W.J. Riley, EG&C, USA

Mixed Constellation Synchronisation of GPS Time 239
A.L. Satin al, Aerospace Corp, USA

GPS Standardisation for the Needs of Time Transfer 243
W. Lewandowski &al, BIPM, F

Comparison of GLONASS and GPS Time Transfers 249
P. Daly, Univ. Leeds, UK, &al

An Experimental Health Monitoring Unit for GPS and GLONASS 253
B. Eissfeller, Kayser-Threde, &al, D

Collaboration in the UK on Time Transfer using the GLONASS System 259
P. Daly, Univ. Leeds, &al, UK

SESSION 7:
PLENARY POSTER SESSION

Dedicated Oscillator for Immersed Crystal Sensor 265
G. Bardèche &al, ENSMM, F

Radiation Effects in Germanium Doped Quartz 269
H. Bahadur, NPL, India

Trapped Energy Quartz Crystal Resonator on Wedge-shaped Plate 273
R. Bourquin, ENSMM, F

Time and Frequency Reference for a Trajectory Measurement System 279
G. Boime, Tekelec, F

Basic Design of a Spaceborne Hydrogen Maser 283
L.G. Bernier, Obs. de Neuchâtel, CH

Crystal Growth of Berlinite: Methods for Growing Longer Crystals and Improving the Quality 287
X. Buisson, SICN, F &al

TF-4020A TFL's New Compact Military Rubidium Standard 293
E. Eltsufin & A. Stern, TFL, Israel

Ultra Stable Oscillators for Satellite Applications 297
J. Ho, Frequency Electronics, USA

Preliminary Results of a Newly-designed 'Mini OCXO' Especially Dedicated to Space Applications 301
E. Gerard & G. Wagner, CEPE, F

Use of the TV-Signal in Stabilising a Local Rubidium Oscillator 309
R. Kivelä, Univ. of Oulu, Finland

Experiments on a Diode-Laser-Pumped 87Rb Maser 315
D. Jinquan &al, Academia Sinica, PR China

Optical Detection of Laser-Cooled Ions in a Penning Trap 319
H. Imajo &alComm. Res. Laboratory, Japan

Quartz Crystal Twinning under Mechanical Stress: Finite Element Analysis 323
J.J. Boy &al, LCEP—ENSMM, F

First Experience with the ASN-16 GLONASS Receiver 329
B. Eissfeller &al, Kayser-Threde, D

Stability of Nonlinear Phase Motion from the Study of Oscillator Mappings 333
M. Planat, CNRS, F

Experimental Results of the Performance of the New Phase and Frequency Controller 337
J. Mannermaa & K. Kalliomäki, Univ. of Oulu, Finland

Time Stability: an Improvement of the Multi-Variance Method for the Oscillator Noise Analysis 343
F. Vernotte, Obs. de Besançon, & E. Lantz, CNRS, F

Anisotropic Etching of Silicon in KOH Solutions 349
A. Brahim-Bounab & C.R. Tellier, LCEP/ENSMM, F

Microwave Synthesizers in Space 355
A. Vulcan & M. Bloch, Frequency Electronics, USA

Very Wide Bandwidth HF Crystal Filters 361
P.K. Wall, Peter Wall ElectroAcoustics, UK

Measurement Results of the Crystallographic Orientation of the SC Cut Round Plates 365
K. Weiss, Tele- and Radio Research Institute, Poland

A Frequency-generating Source and Frequency Synthesis 369
N. Yahyabey, CDTA, & S. Hassani, CDTN, Algeria

Phase Differences of the Fields in a Ramsey Resonator for Submillimetric Frequency Standards 371
E. Bava, Politecnico di Milano, & P. Tavella, Inst. Elettrotecnico Nazionale, I

SESSION 8A:
QUARTZ: MATERIALS
Chairman: H. Duchaussoy

Dissolution controlee d'un materiau piezoelectrique, la berlinite (A1P04) 377
O. Cambon, CEPE, &al, F

Crystal Growth and Physical Characterisation of GaPo4 383
E. Philippot,CNRS-URA, F, &al

Mesures de courants thermoioniques dans le quartz 389
G. Jeandel, Lab. IR, F, &al

The Third-order Permittivity of Quartz Obtained by Different Methods 393
C.K. Hruska, York University, Canada

Thermoelasticity in Quartz 397
J.A. Kosinski &al, US Army Elec. Techn. & Devices Lab, USA

SESSION 8B:
ATOMIC FREQUENCY STANDARDS:
MASERS AND TRAPPED ION FREQUENCY SOURCES
Chairman: C. Salomon

The Development Progresses of Hydrogen Masers at Shanghai Observatory 405
Z.C. Zhai &al, Shanghai Observatory, Shanghai

Investigation of the Frequency Instability of CH1-75 Hydrogen Masers 409
N.A. Demidov, Inst. of Elec. Measurements KVARZ, Russia, &al

Ion Trap Frequency Standard: Research at PTB 415
D. Schnier &al, Physikalisch-Technische Bundesanstalt, D

Controlling an Optical Frequency by a Single Ion 421
I. Siemers, Universität Hamburg, &al

SESSION 9A:
QUARTZ: SPACE OSCILLATORS
Chairman: G. Busca

Quartz Oscillators, Frequency Synthesisers, and Timing Systems for Geodesy and Radioscience Applications 427
J.J. Suter &al, Johns Hopkins University, USA

Space Radiation Susceptibility of Ultra-Stable Crystal Oscillators 431
M. Brunet, CNES, &al, F

Frequency Generation and Synthesis in the DRSS 441
P. Van Loock & J. De Vlieghere, Alcatel Bell Telephone

Oscillators for Mobile GPS Receivers 445
J.P. Aubry &al, CEPE, F

SESSION 9B:
ATOMIC FREQUENCY STANDARDS: ELECTRONICS
Chairman: R. Vessot

Microwave Interrogation Frequency Noise and Clock Frequency Stability: Experimental Results 453
R. Barillet &al, Laboratoire de l'Horloge Atomique, F

High Resolution Digital Synthesiser for a Rb Frequency Standard 457
E. Detoma, FIAT CIEI, I, &al

Solid State Frequency Synthesis up to 600 GHz: an Application to the Mg Frequency Standard 471
A. Godone & C. Novero, Ist. Elletrotecnico Nazionale, I

Designing for Frequency and Time Metrology at the 10^{-18} Level 477
F.L. Walls &al, Nat. Inst. of Standards Techn., USA

SESSION 10A:
TIME AND FREQUENCY: NETWORK SYNCHRONISATION
Chairman: P.Kartaschoff

Time Generation and Distribution System for a Military National Synchronisation 485
V. Mourier, Tekelec Systems, F

Timing and Time Signal Distribution in Digital Communications Networks 489
M. Kihara & A. Imaoka, NTT Transmission Systems Laboratories, Japan

Timing a National Telecommunications Network Using GPS 493
E.W. Butterline &al, AT&T, USA

Strategies for Synchronisation in an Evolving Telecommunications Network 497
R. Avery, Telecom Solutions Europe Ltd, UK

Synopsis of Timing Measurement Techniques used in Telecommunications 503
G. Zampetti, Telecom Solutions

SESSION 10B:
ATOMIC FREQUENCY STANDARDS: MISCELLANEOUS
Chairman: F. Walls

A Compact and High-performance Reference Cell for Laser Frequency Stabilisation 511
G. Dudle & P. Thomann, Obs. Cantonal de Neuchâtel, CH

Line Inhomogeneity Effects and Power Shift in Miniaturised Rubidium Frequency Standards 515
G. Mileti &al, Obs. Cantonal de Neuchâtel, CH

Noise Measurements in Single-mode GaAlAs Diode Lasers 521
N. Sagna, Obs. Cantonal de Neuchâtel, CH, &al

Fundamental Characteristics of a Laser-diode-pumped Rubidium Maser with an External Feedback Loop 527
N. Oura &al, Tokyo Inst. of Technology, Japan

Progress Towards Laboratory Frequency Standards for the Optical Communication Bands 531
D.J.E. Knight, National Physics Lab, UK, &al,

LIST OF EXHIBITORS 539

LIST OF PARTICIPANTS 543

Foreword

1 - Introduction

The European Frequency and Time Forum (EFTF) is an international conference and exhibition, providing information on recent advances and trends of scientific research and industrial development in the fields of Frequency and Time.

The EFTF was inaugurated in 1987 as a meeting point and focus of discussion for the European Time and Frequency community. In five years the reputation of the EFTF's high standards has spread around the world and the Forum now attracts large international participation.

2 - Purpose

The fact that time can be measured so precisely, far better than any other physical parameter, represents a technological asset of great importance. Nearly all space programmes and many terrestrial applications, require some form of stable timing for successful operation. Typical examples are digital network synchronisation, spread spectrum communication, range and range-rate determination, navigation and position, very long baseline interferometry, scientific research (gravity and relativity experiments), time transfer by satellites, Earth observation, geodesy, etc.

3 - Participants

The 6th EFTF was held at the European Space Research and Technology Centre (ESTEC) in Noordwijk (Holland).
The 234 external participants from 24 countries world wide, were drawn from a wide variety of backgrounds including national observatories, research institutes, universities and industry. This diversity was reflected in the spectrum of topics addressed which ranged from fundamental physics and advanced research to materials and applications.

Feed back from participants and the scientific committee indicate that the Forum was a major success technically, organisationally and socially. For the majority of attendees, this was their first direct contact with ESA; I am sure they left ESTEC with a positive impression.

4 - Proceedings

Over 80% of the papers were handed in during the Forum. For the first time the Forum proceedings will be registered in a formal way (via the ESA system). This will increase the circulation of proceedings and possibly improve the general recognition of the Forum world wide. The executive committee will investigate the possibility of registering future Forum Proceedings (ESA, IEEE etc).

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

5 - Statistics

Event:	6th European Frequency and Time Forum (EFTF)
Date:	17-19 March 1992
Location:	ESTEC/ESCAPE
Participants:	266
External:	234
ESA:	32
Countries:	24
Invited papers:	7
Contributed papers:	62
Poster papers:	25
Sessions:	17 of which 14 (7 X 2) were in parallel
Exhibitors:	13

The next EFTF will be held in Neuchatel Switzerland between 16-18 March 1993.

S. Feltham

Session 1

Opening

Chairman: S. Feltham

Welcoming Address

Mr. Le Fèvre
Director of ESTEC
Noordwijk, The Netherlands

Ladies and Gentlemen,

It gives me great pleasure to welcome you, the European Frequency and Time community and scientists from around the world, to the European Space Research and Technology Centre (ESTEC) and to this, the 6th European Frequency and Time Forum (EFTF). The EFTF has been held annually since 1987 alternatively in Besançon, France and Neuchâtel, Switzerland. I am pleased that the Executive Committee has entrusted ESTEC to organize the first Forum to break this tradition. I am sure they and you, will not be disappointed.

I would like to take this opportunity to highlight the importance of Time and Frequency to ESA activities and outline the objectives of the Forum.

The challenge of space is a powerful driver for technological progress. No matter whether the application is:

- Communications
- Ranging, positioning or navigation
- Planetary exploration
- Fundamental scientific research
- Earth observation and geodesy

Space applications usually require hardware to have state-of-the-art performance and to operate unattended under severe environmental conditions.

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

Equipment must survive the launch and deployment phase and function for the duration of the mission, possibly up to 15 years. It must be realised with design guidelines that specify minimum mass, volume and power consumption. Add to the recipe a little spice in the form of quality control requirements, budget limitations, plus a schedule driven by a launch deadline, and you begin to have an insight into the complexities of developing equipment for space missions.

Today the spotlight on space time and frequency activities have never been more intense. This is particulary highlighted by the deployment of the American and Russian Global positioning satellite systems (GPS and GLONASS) (a total of up to 24 spacecraft each). At ESTEC we are aware of the growing importance of time and frequency for ESA programmes. Amongst the most recent measures taken to underline this are:

- Coordination of time and frequency activities
- An industrial policy to develop and maintain expertise in industry
- Briefing meeting between ESA and industry
- Sponsorship of the EFTF at ESTEC

Approved Agency programmes and those presently under study and evaluation offer many opportunities to the Time and Frequency Community. Without going into detail I would like to mention a few examples:

- Columbus and EURECA for microgravity, fundamental physics and ultra stable clock technology.
- ARISTOTELES, Polar Platform and Hermes for GPS receiver development for positioning, navigation and guidance.
- VULCAN, MARSNET, CASSINI ARTEMIS etc... will continue to generate technological challenges for equipment and system designers.

Fundamental to the ESA charter is the need for cooperation amongst member states and their national space agencies. Cooperation at international level is also important and should be encouraged. This is in line with the wishes expressed at the November 1991 ministerial conference held in Munich. Space missions are complex, costly undertakings. It is the responsibility of all involved to ensure that existing expertise and experience is used to the full benefit of the programme and to maximise the return of information and data. I was interested to learn of a joint German/Swiss/American Proposal for a joint ESA/NASA experiment to demonstrate in-orbit performance of hydrogen masers and a time transfer system. Not wanting to prejudge the issue, I welcome such initiatives and wish the experimenters every success during the evaluation phase of their proposals.

An essential part of ESA activities involves the ground segment. ESA operates 9 Tracking Stations world wide with the central control at the European Space Operations Centre (ESOC) in Darmstadt - Germany. Future ESOC activities related to Time and Frequency may include:

- Transition to GPS based network timing
- Interferometry for Geostationary satellites
- High precision time and frequency synchronisation
- Reference frequency transfer to specialised missions for space VLBI, relativistic and gravity wave experiments.

Turning to the European Frequency and Time Forum: an important element in the advance of science and technology is the dialogue between researchers working in common fields of interest. The EFTF is part of such a dialogue. It serves as a meeting point for:

- Cross fertilisation of ideas
- Exchange of views and experiences
- Identification of applications and experiments

- Exhibition of latest products and technology
- Side meetings for specialists groups and committees

This is achieved via the formal presentations, detailed discussions with authors during the plenary poster session and informally at the conference social events, coffee breaks and side meetings.

An indication of the importance of the EFTF is the number of participants, more than 250 scientists from over 25 countries world wide. Clearly the EFTF is an international event providing the framework for discussing recent advances, trends in scientific research and international developments in the field of time and frequency.

In concluding, I would like to thank those companies supporting the Forum by exhibiting their latest products.

I would like this Forum to be the occasion for everyone to appreciate the importance of cooperation and dialogue to our work. I wish you all a successful Forum, a good visit to ESTEC and an enjoyable stay in Holland.

THE ADVANCES IN F&T TECHNOLOGY - THEIR BENEFITS OF AND FOR SPACE SCIENCE AND TECHNOLOGY -

Ph. Hartl
University Stuttgart
Institute of Navigation

ABSTRACT

Time is an indispensable parameter for numerous types of physical measurements. It is also the most accurately measurable physical quantity. Space Science and Technology apply time and frequency measurements for time tagging of scientific and technical data, for time multiplexing in the fields of telecommunication, for radiometric tracking and orbit determination etc. Frequency and time (F&T) measurements constitute means for detection and quantification of various phenomena of interest for fundamental science, geoscience, and navigation. Space technology offers special opportunities for time distribution and for new markets of F&T applications, one of which being the satellite navigation techniques.

Keywords: Sidereal time, atomic frequency standards, radiometric systems, VLBI, orbit determination, relativistic effects, time distribution.

1. INTRODUCTION

1.1 Distinguished Role of F&T

Time is defined as the dimension of the physical universe which orders the sequence of events at a given place. As everything occurs in a continuously changing world, time is an indispensable parameter of any physical observation. In addition, frequency and time (F&T) are by far the most accurately measurable physical quantities. For this reason the general tendency can be observed in metrology to relate quantities of physical phenomena to measures of time, time intervals and / or frequencies, thus trying to achieve by this indirect way the utmost accuracy and reliability (Ref. 17). This distinguished role of F&T techniques has evolved in the past especially in accordance with the progress in microelectronics and will continue this way.

1.2 Outstanding Performance

Various reasons exist for the F&T measurements to achieve the outstanding performances:

- The operation of the signal sources of F&T standards is based on nearly ideal physical conditions. In the rf- and microwave region both, the atomic frequency standards as well as the crystal oscillators have arrived at a state of development, where almost all effects can be avoided, which could deteriorate the stability of radiation. The achieved signal qualities are very close to the theoretical limits of the physical principles, they are based upon. And yet, efforts continue successfully to find out and to make use of new physical effects and to invent new technological methods which allow to shift the limits of stability further towards new thresholds of quality. Since several decades almost every 7 years the F&T stability was improved by an order of magnitude, as is well known (C. Audoin, 1990). Frequency stabilities in the order of 10^{-15} and better can be achieved (Ref. 4). There is no reason to believe that the final quality limits has been reached now.

- For the optical region the progress of laser based frequency standards is important, as many practical applications evolve in the areas of optical terrestrial and space- to- space telecommunication systems as well as in the laser based optical remote sensing technique.

- The theoretical background of stable frequency generation and the modelling of signal characteristics w.r.t. frequency and phase stabilities has an excellent state. On the basis of well defined criteria the special features of the different signal sources can be assessed. They allow to identify the optimal source for the individual application of interest (Ref. 19).

- Also the methods of synthesizing signals with frequencies different from the original atomic frequency standards, or with characteristics combining the advantages of different types of frequency standards have been developed to an excellent status. They are about to become further improved. For example, with a double phase-locked loop receiver Wang and Dick designed a new frequency standard as a combination of a hydrogen maser and superconductiong cavity-maser oscillator (SCMO). It offers a frequency stability of 2×10^{-14} for a 1s, and 1×10^{-15} for 1.000s measuring time. This standard delivers output frequencies of 100 kHz, 5 MHz, 10 MHz, and 100 MHz. It combines the short time stability advantage of the SCMO with the long term stability features of the

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

hydrogen maser (Ref. 23).

- The techniques of signal processing and data handling have been improved as well.

- In the hardware developments considerable progress has been made as regards advantages in weight, volume, power, mobility, reliability, lifetime, and cost of the instrumentation, and in the provision of proper interfaces to digital systems etc.

- High quality time transfer techniques via satellites have become available, which allow time comparison in the ns and sub-ns accuracy level (Ref. 4, 10, 12, 21, 22). The global availability of highly synchronized time will lead to further applications in the areas of navigation, telecommunication, and science.

- An expanding market drives the efforts to improve the quality of mass products and cost reductions. The increased use of the GPS- based satellite navigation system becomes important for the near future: Millions of cars will be equipped with GPS- receivers. Within a few years the cost of GPS receivers has been reduced by 2 orders of magnitudes and, as of now, the cheapest low cost models run around 1.000 $. Further cost reductions are expected. One can expect the same tendencies as in the wrist watch techniques two decades ago.

2. APPLICATION IN SPACE

2.1 The Different Missions

Space Science and Technology are main drivers for further F&T improvements in all respects because of the numerous different and ambitious applications. The space missions comprise very different types of orbits ranging from Low Earth Orbit (LEO) for geoscientific satellites to the Medium Earth Orbits (MEO) for navigational satellites, to Geosynchronous Earth Orbits (GEO) for telecommunication satellites, to highly elliptical orbits for space- VLBI satellites and for comet missions. There exists a great variety of deep space trajectories, for example for planetary, interplanetary, and solar missions, and for missions leaving the solar system, the extreme case being the "TAU" proposal, which was studied recently by JPL. TAU stands for Thousand Astronomical Units, indicating that a probe was investigated which should observe the solar system from outside, namely from a distance of 1.5×10^{11} km (Ref. 18).

Apart from the common task of orbit determination and data transmission some missions require special services like real time navigation for rendezvous and docking, for planetary fly-by and planet landing and for long time autonomous mission execution

2.2 The Different Accuracy Demands

All space missions require time tagging of events. This poses, fortunately, only modest accuracy demands, because the spacecrafts move with speeds very low in comparison to the velocity of light. A tolerance of 1 μs to 1 ms can be tolerated in most cases as a spacecraft moving with 10 km/s, for example, would move during this time only 10 mm to 10 m.

More demanding can be tasks for proper operation of a telecommunication satellite system, especially if time multiplexing is applied (Ref. 20). The signals of the different ground stations must arrive at the transponder of the geosynchronous satellite in a well defined sequence. They must not overlap. To assure this some tolerance ΔT in the time of arrival of the signals must be accepted. Any range uncertainty $\Delta R = c\Delta T$ corresponds with a reduced channel capacity due to a "loss" of $\Delta b = \Delta T r$ bit/s, if r is the bitrate and c is the speed of light.

Considerably higher accuracy demands exist for the radiometric range and velocity measurements themselves. All distances in space must be measured in light- seconds. A tolerance of 1 ns corresponds with 30 cm uncertainty, for example. In addition, positioning accuracy is generally a product of that range uncertainty and some "dilution factor" relating ranging and Doppler errors to positioning errors. It accounts for the unfavourable geometry of the measurement. The dilution factor is in general in the order of 2 to 6, but can also be as large as 10 or even worse. A ranging error of 30 cm corresponds in case of a dilution factor of 10 with a position error of 3m, for example. This degradation of accuracy is well known to users of the space- based navigational system GPS for position determinatio.

This Global Positioning System GPS of the U.S. plays an extremely important role. But in the following we shall use the abbreviation "GNSS", which stands for "Global Navigation Satellite System". It comprises GPS as well as to the Russian GLONASS and/or navigation future systems of such types.

2.3 Satellite Geodesy

For satellite geodesy the scientists have established a network of high power laser ranging stations. The stations transmit very short, but powerful light pulses towards the satellite, where they are returned by an arrangement of optical corner reflectors. The delay time between transmission and reception is measured with an accuracy of the order of 1 cm. The operation of that system is expensive and weather dependent. It is only in use for special projects demanding very precise orbit determination.

For satellite to satellite tracking the geodesist expressed demands a special tracking tracking technique based on low power cw- laser measure velocity variations of the order of 10 micrometer per sec. They want to determine high order Earth gravity field components from space.

Apart from those special cases satellite tracking is generally based on microwave Doppler and Ranging techniques. Precise velocity and acceleration measurements in ranges reflect in demands for precise carrier phase measurements. Often fully coherent systems are needed because of the high accuracy and link efficiency demands. High quality oscillators, frequency synthesizers, very stable filters, excellent mixers, programmable VCO´s, and higher order phase locked loops are needed.

Very precise absolute and relative positioning demands exist in Geodesy and surveying techniques also for microwave systems. The ranging accuracies should be as good as 1dm to 1mm, and the relative range rate measurement should allow velocity measurements of 1mm/s to 0.01 mm/s accuracy and even better (Ref. 3, 13, 14).

Various relativity effects, which have been neglected so far, must be taken into account in the future: The more precise the radiometric measurements will become, the more effects must be modelled. With the aim of achieving precisions in the order of mm in the future a group of competent geodesist met recently at USNO and agreed to use in space geodesy in the future a "consensus model" of relativity with picosecond accuracy as a new geodetic standard (Ref. 5).

2.4 Oceanography

For the application of satellite based radar altimetry in the Oceanography it is necessary that both, the radiometric tracking of the orbiting satellite as well as the ranging from the spacecraft to the ocean surface is performed with accuracies in the order of some cm or better. This allows to determine the ocean surface topography and thus to measure the ocean circulation: The amount of water transport of the Gulf stream and its location is indicated, for example, by a tilt of the ocean surface across the stream amounting to about 1 m over a distance of 150 km (Ref. 7).

2.5 Fundamental Science

Gravitional and non- gravitational forces and fields will generally affect the orbit determination´s accuracies. They can be misinterpreted as technical effects like aging, temperature influences, drift etc. and vice versa.

"Tracking error"- effects are therefore to be studied carefully in most space missions. They are not only subject of investigations for the purpose to improve the accuracy of the tracking results, but are also of scientific interest:

- Deviations of the emitted nominal frequency can be caused by relativistic effects.

- The medium between the spacecraft and the ground station of the radiometric system can affect the line of propagation (ray bending) and the velocity of wave propagation due to plasma (ionosphere, corona of the sun, interplanetary plasma etc.) and neutral particles (troposphere). (Ref. 16)

- Forces and fields can effect the Doppler signature and / or the signal delay. Sub-ns delays and Doppler frequency shifts with special signatures in the order of 10^{-16} to 10^{-18} should become detectable in the near future and are subject of relativistic experiments, especially towards gravitational waves (Ref. 1, 16).

- If the ground station is not at the nominal location, i.e. if its coordinates are incorrect in the Geo Reference and/or Inertial Reference Frame, then additional errors in orbit determination can occur.

3. TIME, VLBI AND TRACKING

Time is an instant and is measured by relating that instant or "epoch" to a sequence of "recurrent events" or "periods". A period represents the time interval between successive events. These events can be related to the macrocosmos or the microcosmos, i.e. to the sidereal time, or the atomic time. Both, the sidereal time, and the atomic time are important for space.

The atomic time constitutes the basis for the determination of the distance in light-seconds, for the Doppler shift in Hertz, and for the relationship between wavelength and phase with distance. The atomic time is related to the transition radiation of the Cesium-133 atom between two of its hyperfine levels.

The sidereal time refers to the earth´s rotation, as is well known. The mean sidereal day is the time interval between two successive meridian transits of the vernal equinox: 23:56:4.09054 s.

3.1 Reference System

In order to determine the siderial time an Earth Centered Earth Fixed Coordinate Frame (ECEF) must be defined: The Z-axis represents the spin axis of the Earth and is defined by CIO. The X-axis lies in the mean equatorial plane and passes through the mean Greenwich astronomical meridian which is defined by BIH. The Y- axis is orthogonal to X and Z. The origin of the

coordinate system is the centre of mass of the Earth.

Observations made from some ground station are only well defined in space and time if that ground station's coordinates are precisely known for that instant of the observation. As long as crustal movements can be ignored it would be sufficient to monitor the systematic and random variations of the Earth's motion (polar motion and variation of the Earth's rotation) by means of the VLBI stations (Very Long Baseline Interferometry) and to relate the ground stations' coordinates to those of the VLBI stations just once: The VLBI stations would represent the fiducial points and the transformation of the radiometric results of the ground station into the inertial space reference frame would be well defined.

3.2 VLBI

The VLBI ground stations observe radio sources, as is well known. Opposed to the classical astronomy, which observes the visible stars VLBI systems track the remote quasars, or more generally, the extragalactic radio sources. The distances to them are so large that the line of sight (l.o.s.) from the Earth's centre to these radio sources remains quasi unchanged over hundred thousands of years: Even if these radio sources would move across l.o.s. with velocities close to the speed of light no existing instrument would be able to detect such a change. Extragalactic radio sources are the ideal reference points in a world, where everything is moving.

The basic concept of VLBI is well known: At two VLBI stations the signal of the quasar is observed and recorded on memory as a function of time. The weak signal of the radio source is identical at both stations. However, different additive noise is superimposed at each station and the time of arrival of the signal is also different. By correlation of the recordings during post processing in a VLBI data centre the noise is extracted and the time difference ΔT is determined. It corresponds to the range difference ΔR of the quasar signal to the two stations, expressed in light- seconds. It is a function of the length of the baseline B between the two VLBI stations and of the siderial time. The distance B is usually many thousands of km long.

The sensitivity of this technique is extremely high. From observations of many quasars and by involvement of various VLBI- stations it is possible to determine the baseline vectors with accuracies around 1 cm and milli-arcsec resp., provided the VLBI stations are equipped with very precise clocks, preferably hydrogen clocks.

Polar motion and changes of the Earth's rotation, i.e. UT1, can be determined by VLBI with highest accuracy. This technique will become even more precise in the future, when Space-VLBI will be used (Ref. 6). A Russian (RADIOASTRON) and a Japanese Space-VLBI- mission (VSOP) are under development.

3.3 Satellite Tracking

Also geodetic laser ranging to satellites in high orbits, for example LAGEOS, allows the determination of Polar motions to 0.01 to 0.02 arcsec (or 30 to 60 cm) rms accuracy. By means of Doppler satellite tracking monthly means of polar motion can be achieved with accuracies of 10 to 20 cm. GNNS positioning can provide even better results (Ref. 3) and will be used globally for operational purposes.

The ambitious request is now to monitor Earth rotation and Polar motion on an hourly basis. High precision radiometric tracking, combined with space based data collection, is requested. PRARE (Ref. 9) could be used for this purpose, if installed in an high orbiting satellites and if some PRARE- ground stations would be colocated with VLBI- stations.

3.4 Geocenter

Apart from polar motion also the centre of the Earth must be defined. The Geocenter is the centre of gravity (c.g.) of the Earth. It cannot be determined by VLBI, but only by orbit tracking of Earth rotating bodies. The laser ranging techniques of the geodesists are used for that. They track the orbits of satellites and/or the motions of the corner reflectors on the moon's surface which were established there during the Apollo missions. The position of the geocenter is presently known to an accuracy between 1m and 1dm (Ref. 13). GNSS and tracking techniques like PRARE and DORIS can improve it further.

3.5 Crustal Motion

Above it was assumed that the crustal movements can be ignored. However, this is not the case, as we are talking about the cm and mm level. These changes must be measured as well.

As the ground stations, which are used for such measurements, are changing their positions against the c.g. and against each other the tracking results and many non- gravitational and gravitational forces affect the orbit. The ground stations of the tracking network must, therefore, be arranged on Earth in a manner that some of them are located on relatively stable positions. They allow the orbit determination. Others must be located on places with considerable drifts and rotations of the plates occur. They can be used to determine the crustal movements. Sensitivity improvements towards 1mm/year are requested for the future.

3.6 Ground Station Coordinates

The uncertainties Δ in the ground stations coordinates (geometry and time) lead to orbital errors. In deep space tracking this Δ-value is to be amplified by a factor of R/R_S, to get the positioning error of the probe. R is the distance of the probe from the centre of the Earth and R_S is the radius of the Earth. This can effect position errors of many km.

If the position of the deep space probe is known, then the tracking error can be determined and the result taken as an information source to investigate Δ. Random polar motion and Earth rotation can thus be investigated. Geoscientists derive from the result -in combination with models and in situ measurements- answers about the interaction between the Earth´s surface and the atmosphere, but also about the processes in the Earth´s interior and in the crust. This is essential for earthquake research, for example.

4. DOPPLER AND RANGING

Radiometric tracking can be basedon Doppler and range measurements.

4.1 DOPPLER

Frequency is measured in Hertz = 1/s, where the second s is defined by the atomic time standard. The frequency of a received signal remains constant as long as the transmitting signal source and the receiver keep their operational conditions unchanged. A frequency change is thus an indication that either the signal source and/or the observer and/or the medium in between are subject of changes.

Changes of the frequency can be measured directly as Doppler effects or as phase measurements, or via the time delay. Phase shift is the displacement of a wave in comparison to its origin. It is measured either in multiples and fractions of that wavelength or in radians modulo 2π.

One-way and two-way Doppler systems are in practical use. The one-way Doppler system requires moderate long-term frequency stability for the transmitter and the receiver, and medium term stability during the tracking time of a station, which is for low orbiting satellites of the order of 10 to 20 minutes. Good examples of the one-way Doppler tracking techniques are the TRANSIT satellite system and the DORIS system.

Two-way Doppler tracking requires only short term stability for the comparison of the transmitted to the received frequency during the time T=2R/c, if R is the range between S/C and ground station. PRARE applies both, the two-Doppler- and the two-way Ranging technique, but allows one-way Doppler tracking as well.

The Doppler effect is a function of the orbital height. For low orbiting satellites it offers very excellent information. For geosynchronous satellites the Doppler effect is zero. For very far distances the Doppler frequency contains two effects: The radial velocity of the probe with respect to the Geocenter, and the rotational motion of the ground station due to the Earth´s rotation. The observation time for deep space missions is very long and, for this reason, one-way systems can be used there only, if extremely stable oscillators are applied on board as well as and on ground.

4.2 Ranging

The distance can be measured in either multiples and fractions of λ or in light- time $T = \lambda/c$. In technical terms: One can either use the time interval counter (TIC) or the phase detector (PD), or a combination of both. For very high precision it is the PD, which is to be preferred. It compares the transmitted signal with the received one. This is actually the same technique which is applied in VLBI.

If the phase detector is used in form of a correlator then the precision is in the order of 1o/o to 0.01 o/oo of a wavelength or even better. Improvement of the precision by a factor 10 should be possible by use of a 10 times higher carrier frequency. If the corresponding wavelength is, for example 10 cm, then precisions of fractions of mm can be achieved. For the purpose of very high resolution the PD-method is preferable, for the measurement of very long distances it is easier to use the TIC-method, because the TIC- method avoids the ambiguity problem. Ambiguity means that the distance is measured as a fraction of a wavelength plus integer multiples of the wavelength. The integer is not known a priori and must be determined additional a priori knwoledge and calculations.

The values of precision mentioned above don´t correspond, of course, with the accuracy. Ambiguity problems, drifts, biases, aging etc. impose, problems and can degrade the result. The long term stability is a main issue of nowadays technological tasks in the F&T area and concerns the whole system, especially the frequency generator, synthesizer, transmitter, receiver, the phase centre of the antenna, and the media between the transmitter and receiver.

One-way ranging is only possible with very stable clocks at the transmitter and receiver. If satellite and ground timing systems were to be maintained to an equivalent accuracy of one meter over a five day time period an oscillator stability of 8 parts in 10^{15} would be needed.This hardly achieved within reasonable costs.

Alternatively it is also possible to use a set of 4 one-way links, as is well-known from the GNSS-navigation system: The 4 transmitters (or receivers) must be equipped with very stable clocks, the demands for the receiver clock (transmitter clock) can be very moderate: Under this assumption it is possible to solve for the time uncertainty as well as for the 3 geodetic coordinates in the receiver equipment which applies the crystal clock of moderate accuracy.

Two-way ranging requires only short term stability of the clocks during the time between transmission and reception of the signal. An additional advantage is obvious: It allows direct calibration. The disadvantage is the additional effort in hardware.

Ranging is less precise than integrated Doppler, because the ranging signal is modulated onto the carrier signal. For the modulation spectrum of a carrier signal one can only allow a few percent of the carrier frequency.

A final remark to the TIC- method as compared to the PD- method is appropriate: The TIC counts the time interval between transmission and reception of each individual pulse, the PD determines the average time delay of a random or a periodic signal of some length, not just over one period. This length corresponds with the integration time of the correlator. So, if one compares the quality of the TIC and PD results it is, therefore, appropriate to average also the TIC counts over the same time slot the correlation has been performed.

4.3 Deep Space Tracking

For deep space tracking a special network was established by NASA, the Deep Space Tracking Network DSN. Each DSN- station is equipped with at least two atomic frequency standards: A hydrogen maser and cesium beam frequency. The hydrogen maser has better spectral purity and frequency stability for periods of up to 10^5 s. The spectral purity is essential, because the frequency standard must be multiplied by some factor n for the purposes of synthesizing standard DSN transmit/receive frequencies. Frequency division or multiplication by n degrades the spectral purity (S/N) by n^2. If n=10, then the S/N degraded by 20 dB. Good quality of S/N is important for DSN: Measurement times up to 10^4s round-trip light-time must be taken into account.

Two-way Doppler tracking is the preferred method in deep space missions;only occasionally two-way tracking is applied in addition, in order to avoid ambiguity problems. Two-way Doppler tracking operates very well und normal conditions with positioning accuracies in the order of a few meters. But in cases, where the deep space probe is close to zero declination. the geometric dilution factor becomes very large and approaches infinity at declination=zero. Fortunately for a probe moving in the space of the inner planets this critical period is only a few days long. For an outer planet mission it can, however be as long as many months.

A different method must be applied in this critical case. It is based in two-way ranging: Pairs of ground stations located well above and below the equator resp. are to be used. The ranging difference defines the declination value with good accuracy.

A new method of deep space tracking represents the ΔDOR- method (Ref. 8): The Delta Differential One-Way Ranging. It applies the VLBI method to both, to the quasar's and to the deep space probe's signal. The angular distance of the two signal sources (quasar and probe) should be as small as possible, at most 10° off from each other. The deep space probe transmits a wideband signal by means of a radio beacon. For some time the ground station observes the quasar signal, then for a shorter time the deep space signal and then again the quasar signal a.s.o. The angular distance of the probe w.r.t. the quasar determines the position of the probe on the sky. The one-way Doppler determines the radial velocity of the probe.

The advantage of that method is quite obvious: There is not need for two-way ranging to the probe: No transmitter is needed on ground. The main radiometric errors are the same for the quasar and the probe and cancel, therefore. The measurement time can be longer, because transmission at the probe is not only started when the uplink signal arrived. However, the oscillator stability on board the probe must be of excellent quality.

5. TIME TRANSFER

As an example showing the benefit of the space technique for the practical application on Earth the time transfer techniques will be shortly reviewed.

It is very important that now the time and frequency comparison is possible on a global basis with reasonable costs. One-way time transfer by means of GPS is possible in case of common view with an accuracy in the order of one to a few ns.

For sub-ns accuracy the two-way time transfer via telecommunication satellites is possible by means the concept of Mitrex (Ref. 10, 12). This method of time synchronization becomes especially important for worldwide comparison of frequency standards with stabilities of 10^{-14} to 10^{-16}.

It would be interesting for the near future to

test the MITREX- concept also in itself: One could transmit its signals over a series of telecommunication systems around the globe, and receive it from East, when transmitted towards West and vice versa, after 4 to 6 hops. Each ground station could then receive its own reference signal after having passed all satellite links. This would directly show the absolute accuracy. By the two-direction measurments one could also study the Sagnac effect and other effects of interest.

Another interesting test is proposed and under preparation (Ref. 4, 22): A pair of H-masers should be flown on EURECA 3 and should be tested with respect to its performance concerning time and frequency stability. This test would use the PRARE- system for time and frequency transmission and comparison.

6. CONCLUSION

Ambitious goals can be reached in Space due to the high performance of the F&T technology. Further improvements are requested. Space techniques are drivers for the F&T promotion, but offer also excellent opportunities for it.

7. REFERENCES

1. ALLEY C.O.: "Proper Time Experiments in Gravitational Fields with Atomic Clocks, Aircraft, and Laser Light Pulses", Plenum Publishing Corporation, 1983

2. AUDOIN C. : "Classical Frequency Standards", 4th EFT, Neuchâtel 1990 FS-01, 1

3. BERTIGER W.I., LICHTEN S.M., KATSIGRIS E.C. : "Determination of GPS Orbits to Submeter Accuracy", TDA Progress Report 42-95, 1988

4. BUSCA G., BERNIER L.G., STARKER S., FELTHAM S. : "In Orbit Demonstration of a H-Maser Clock System", Proceedings 23rd PTTI Applications and Planning Meeting, Dec.1991

5. EUBANKS T.M. : "Relativistic Models for Use in Space Geodesy", Proceedings of the U.S. Naval Observatory Workshop, June 1991

6. FEJES I., ALMAR I., ADAM J., MIHALY Sz. : "Space-VLBI : Potential Applications in Geodynamics", Adv. Space Res. Vol. 6, No 9 pp.205-209, 1986

7. GOWER J.F.R. ed. : "Oceanography from Space", Plenum Press, Marine Science Vol. 13

8. GRAY D.L. : "DeltaVLBI Data Performance in the Galileo Spacecraft Earth Flyby of December 1990", TDA Progress Report 42-106, Aug. 1991

9. HARTL Ph., SCHÄFER W., REIGBER Ch., WILMES H. : "Precise Range and Range Rate Equipment, PRARE, on board ERS-1", XIX I.U.G.G. General Assembly, IAG, Aug. 1987, Vancouver,

10. HARTL Ph., GIESCHEN N., MöSSENER K.M., SCHÄFER W., WENDE C.M. :"High Accuracy Global Time Transfer via geosynchronous Telecommunication Satellites with MITREX", J. of Flight Sciences and Space Research, 7 (5), pp. 335-342, 1983

11. JORDAN J.F. : "Deep Space Navigation Systems and Operations", Proc. Int. Symp. Spacecraft Flight Dynamics, Darmstadt,18-22 May 1981; ESA SP-160, Aug. 1981

12. KIRCHNER D., et al.: "Comparison of two-way Satellite Time Transfer and GPS Common-View Time Transfer between OCA and TUG", PTTI 1991

13. MALLA R.P., WU S.-C. : "Deriving a Geocentric Reference Frame for Satellite Positioning and Navigation", TDA Progress Report 42-95, July-Sept. 1988

14. MUELLER I.I., ZERBINI S. : "The Interdisciplinary Role of Space Geodesy", Earth Sciences, Vol. 22, Springer Verlag 1989

15. MURROW D.W. : "S-Band DeltaDOR Data for Earth One Flyby", JPL Interoffice Memorandum GLL-NAV-89-60, June 1989

16. NELSON S.J., ARMSTRONG J.W. : "Gravitational Wave Searches Using the DSN", TDA Progress Report 42-94 April-June 1988

17. PETLEY B.W. : "Time and Fequency in Meteorology", 4th EFTF, Neuchatel 1990

18. POSNER E.C., HORTTOR R.L., GRANT T.L. : "Communicating from the Edge of the Solar System", IEEE Spectrum June 1986

19. RUTMAN J., GROSLAMBERT J. "Characterization and Measurement of Frequency Stability : Twenty Five Years of Progress", 4th EFTF, Neuchatel 1990

20. SPILKER J.J. : "Digital Communications by Satellite", Prentice Hall 1977

21. STARKER S., ARKO D., HAMMESFAHR J., NAU N., SAPPL E., SCHILD H., SCHIMMEL R., STELZEL F.: "Navigation Experiment NAVEX on Spacelab Mission D1", Final Report DLR

22. VESSOT R.F.C., LEVINE M.W. : "Gravitational Redshift Space Probe Experiment", GP-A Project Final Report, Contract NAS8-27969, Dec. 1979

23. WANG R.T., DICK G.J. : "Combined Stability of Hydrogen Maser and Superconducting Cavity Maser Oscillator", Proceedings of 45th Frequency Control Symposium 1991

Session 2

Space Applications

Chairman: C. Audoin

HYDROGEN MASERS FOR SPACE

R.F.C.Vessot[1], E M.Mattison[1], G.U.Nystrom[1], L.M.Coyle[1], R.Decher[2], S.J.Feltham[3], G.Busca[4], S.Starker[5], and S.Leschiutta[6]

[1]Smithsonian Astrophysical Observatory (SAO), Cambridge, MA USA; [2]National Aeronautics and Space Administration George C. Marshall Space Flight Center, Huntsville, Alabama USA; [3]European Space Agency (ESA), Noordwijk, The Netherlands; [4]Observatoire de Neuchatel (ON), Neuchatel, Switzerland; [5]Deutsche Forschungsanstalt für Luft-und Raumfahrt (DLR), Oberpfaffenhofen, Germany; [6]Politecnico Elettronica Torino (PET), Torino, Italy

ABSTRACT

We describe a program for the design and testing of atomic hydrogen (H) masers for general purpose use in space and methods for high precision space-to-earth time and frequency comparisons. Because of the similarity of the space masers' basic components to those of existing terrestrial masers, their frequency stability is expected to reach well below 1 part in 10^{15} for intervals between 10^3 and 10^5 sec. A NASA-funded development program at the Smithsonian Astrophysical Observatory (SAO), currently in a conceptual design phase, is specifically directed toward producing an H-maser for space, and testing it for 6 months on the European Space Agency (ESA) EURECA spacecraft. The maser design will have the capability to cope with stringent dynamic load test levels and will have sufficient lifetime for long-term continuous operation in future deep space missions such as a close-approach solar probe. The time comparison system for the EURECA test mission will use high precision (< 50 ps) pulsed laser time transfer between the spacecraft and H-maser equipped laser ranging earth stations. European studies of the proposed inclusion of a second H-maser for direct on-board frequency comparisons and of a modified PRARE microwave system (in addition to the laser system) for world-wide time synchronization will also be discussed.

1. INTRODUCTION

Future space applications, including tests of relativistic gravitation,[1] operation of radioastronomy very long baseline interferometers,[2] high precision space tracking, and global time synchronization,[3] will require high stability space qualified oscillators, such as hydrogen masers, having frequency stability better than 1×10^{-15} for time intervals between 10^3 and 10^5 seconds. The first design of a space borne maser was for the successful 1976 SAO/NASA Gravitational Redshift Experiment (GP-A), a two hour, near-vertical, test of the Einstein Equivalence Principle[4]. The next generation of space masers designed for long term continuous operation in space is now under development at SAO in a conceptual design program supported by the NASA Office of Aeronautics and Space Technology (OAST) "In-Step" program. An ESA program for space maser development has begun at the Observatoire de Neuchatel (ON).

The goal of the ongoing NASA/SAO technology experiment is to demonstrate the performance of the new maser design in space on the ESA EURECA-3 spacecraft, which is to be launched in 1996-1997. The EURECA spacecraft will be deployed from the NASA space shuttle and boosted to an altitude of 525 km in a 28.5° inclination orbit. The spacecraft will remain in orbit for six to nine months, after which it will be returned to a lower orbit and retrieved by the shuttle.

The frequency of the space maser will be compared with terrestrial clocks by means of pulsed laser time transfer using the existing NASA laser ranging station at Mt. Haleakala, Hawaii, which will be equipped with H-maser clocks. The ranging station clocks will be monitored by the National Institute of Standards and Technology using the GPS "common view" technique[5], and will thus be related to international time scales. As an extension of this program, a proposal for a collaborative program with ESA is being studied. This program extension would consist of (i) the inclusion on the spacecraft of a second H-maser to perform direct on-board frequency comparisons between an ESA space maser and the SAO maser, (ii) the participation of additional laser stations, and (iii) the inclusion of a modified PRARE microwave system to make time and frequency comparisons between the space clocks and a large number of earth station clocks. This joint NASA/ESA project would greatly broaden the technological goals of the NASA experiment and demonstrate the feasibility of international time transfer at precision in the tens of picoseconds.

Figure 1 shows the expected frequency stability of the spaceborne maser, represented by the Allan deviation, $\sigma(\tau)$, as a function of averaging interval τ. Maser frequency stability over short time intervals – less than roughly 10^4 seconds – is governed by the inherent thermal noise within the maser's oscillation linewidth and by the signal power to the receiver system[6]. These characteristics are determined by the design of the maser's physics unit, and by the receiver's noise figure and bandwidth. These statistical properties are not likely to be affected by spaceborne operating conditions. However, *systematic* frequency variations, which are important for times beyond about 10^4 seconds and which cause the up-turn in the $\sigma(\tau)$ curve shown in Figure 1, can well be different in space than on earth, and are the focus of interest in our investigation.

The laser system will make time interval measurements between the earth and space clocks that will yield the clocks' frequency variations. The heavy slanted lines in Figure 1 show the limits on frequency stability determination that result from uncertainties of 20 ps and 50 ps, respectively, in the laser time transfer system. With a 20 ps system, for example, we will be sensitive to maser frequency variations on the order of 1×10^{-15} for intervals of about six hours.

2. THE EXPERIMENT CONCEPT

The concept for making frequency comparisons between a spaceborne clock and an earth station originated in the early 1980s with a NASA sponsored study of a Satellite Time and Frequency Transfer (STIFT)[7] mission, for which a combination of laser time transfer and microwave time and frequency comparison was proposed. Here the idea was to use laser timing, which is highly precise but is limited to clear

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

sky conditions, to calibrate a pseudo-random-noise (PRN) modulated microwave time transfer system, which would be useable under nearly all weather conditions. The proposed use of a modified PRARE system for time transfer and frequency comparison, in addition to the laser pulse system, follows the general concepts of the STIFT project, and would allow access by large number of users to a world-wide time sub-nanosecond time transfer.

The EURECA spacecraft is shown in Figure 2 with its solar panels extended. Magnetic torquers and cold nitrogen gas jets maintain its orientation in space to keep its solar panels facing the sun. Because the surface of EURECA that faces the ground will change as the spacecraft circles the earth, laser reflector and photodetector arrays will be mounted above and below the spacecraft to allow laser time transfer under nearly all orientations relative to the earth station.

Each day the spacecraft will be visible from the Hawaiian laser ranging site during several consecutive passes, separated by 94 minutes, EURECA's revolution period about the earth. Thus the space maser's frequency will be measured over averaging intervals of approximately 94 minutes, as well as for intervals of a day and longer. With a laser timing precision of 20 ps, we expect to be able to make short-term frequency measurements with a precision of $\sigma_y(94\ \text{min}) \sim 4 \times 10^{-15}$. For longer intervals, the precision improves with the time interval between measurements, as indicated by the straight lines in Figure 1. A goal of the planned test is to measure any systematic effects on the operation of the maser. Such effects will be observable at averaging intervals beyond one day. Of particular interest are the effects of magnetic field and temperature variations. Continuous monitoring will be done of all relevant temperatures, the maser's internal vacuum, hydrogen source pressure and dissociator efficiency, and the maser's output signal level. These environmental data will be telemetered to earth and correlated with the observed systematic frequency variations.

Figure 3 shows a block diagram of the experimental system, including the major EURECA electronic systems and the maser's control electronics, receiver system, clock and event timer. The arrival times of the incoming laser pulses are registered by the event timer with a resolution of about 20 ps and are stored in memory for subsequent telemetry to the EURECA ground control station, along with readings of the monitored system parameters.

During the mission, telecommands from earth will occasionally be sent to the maser's microprocessor to adjust the maser's operating parameters, such as its source H_2 pressure, cavity resonance frequency, and internal magnetic fields. The microprocessor performs several functions, including controlling maser parameters; monitoring maser operation; and carrying out programmed sequential operations, such as determining the maser's internal magnetic field by varying the Zeeman oscillator's frequency and measuring the corresponding maser output power. Control, monitoring and operation of the ESA space maser would also follow this general approach.

3. THE SAO SPACE MASER

Much of the internal structure of the new space SAO maser is based on the maser used in the 1976 GP-A Gravitational Redshift mission, in that the cavity resonator, storage bulb and beam optics are similar. The thermal control system of the new maser has been substantially improved by taking advantage of the longer mission duration and the vacuum of space, which permits us to use multi-layer insulation. Figure 4 shows a cross section view of the new space maser identifying its seven zones of thermal control. The new thermal design evolved from the development of four small passive masers under sponsorship by the U.S. Naval Research Laboratory.[8] This approach permits operation under atmospheric conditions, for testing, as well as in the vacuum of space. The technique uses a set of segmented aluminum isothermal oven enclosures that control heat loss conducted primarily through the multi-layer insulation under atmospheric conditions, as well as heat flow mainly through structural members when under vacuum.

The maser's cavity resonator is isolated from dimensional variations in the belljar that result from barometric pressure changes encountered when going from earth to space. This isolation is achieved by mounting the glass-ceramic CER-VIT[9] cavity resonator and storage bulb, as a subassembly, on a baseplate attached to one of the belljar necks near the center of the belljar endcap. The effects of axial thermal expansion of the cavity mounting structure are minimized by clamping the cavity onto its baseplate with a zero-rate Belleville spring, while radial expansion is reduced by supporting the cavity on the baseplate by a quasi-kinematic roller mount. A three-section printed circuit magnetic field solenoid fits closely within the innermost magnetic shield that surrounds the titanium alloy vacuum belljar. Four hydrogen-sorbing cartridges surround the hexapole state-selector magnet. A small ion pump scavenges non-hydrogen gases.

As in the GP-A maser, hydrogen for the maser is obtained from $LiAlH_4$ contained in a thermally-controlled vessel and maintained at about 3.8×10^5 Pa (3.7 atm). Hydrogen flow is regulated by servo control of the temperature of a palladium-silver diaphragm to maintain a constant pressure in the hydrogen dissociator. Molecular hydrogen is dissociated into atoms by a plasma discharge powered by an external r.f. supply.

The SAO maser's size, weight and expected power consumption are summarized in Table 1. Because EURECA's magnetic torquers produce magnetic field variations as high as 0.1 Gauss at the proposed location of the maser, we have added a fifth layer of magnetic shielding enclosing the entire maser, in addition to the four layers that surround the cavity and belljar, to improve the overall shielding factor and to prevent magnetic field inhomogeneity frequency shifts[10] that can result from these field variations.

A photograph of the engineering demonstration model H-maser that was built to test the thermal design is shown in Figure 5. Thermal tests in vacuum to verify the design will be made using this maser.

Table 1. Space Maser Characteristics

Dimensions (mm/inch)		Mass (kg)	Power Requirements (W)	
diam:	442/17	67	Physics unit:	17
length	863/34		Receiver:	10
			Total:	27

4. LASER TIME TRANSFER FOR FREQUENCY COMPARISONS

The space maser will be compared with a ground maser clock by laser time transfer, which at present is the most precise method of comparing widely separated clocks.

The laser ground station transmits short laser pulses (~200 ps duration) at a rate of about 5 per second. Timing of the laser pulses is controlled by a hydrogen maser at the ground station that is compared with primary time scales by means of GPS time transfer using the common-view technique. When the pulses reach the spacecraft they are both reflected back to the ground station by corner reflectors mounted on the spacecraft, and detected by a photodiode located near the reflectors. The time transfer concept is illustrated in the light time diagram shown in Figure 6. The transmission time, t_{e1}, and the arrival time, t_{e2}, of the reflected laser pulse are measured in terms of the time scale, t_e, of the ground clock. One half the pulse round trip interval, $t_{e2} - t_{e1}$, provides the propagation delay to the spacecraft and allows us to calculate the time of arrival, $t_{ae} = (t_{e1} + t_{e2})/2 = (t_{e1} + [t_{e2} - t_{e1}]/2)$, of the laser pulse at the spacecraft as measured in the ground clock's time scale. The spacecraft photodetector senses the arrival of the laser pulse and provides a signal to the event timer controlled by the onboard maser clock. The event timer determines the arrival time, t_{as}, of the laser pulse in the time scale t_s of the onboard clock. The difference between t_{as} and t_{ae} (onboard and ground pulse times) yields the difference between the ground and space clock time scales. We expect that a precision of 20 to 50 ps in the comparison of the space and ground clocks should be achievable during each period of visibility from the ground station.

Accurate position and velocity data of the spacecraft are needed to determine relativistic corrections. For precision of frequency comparison at a level of 1 part in 10^{16}, we require position tracking accuracy of approximately 1 meter, and velocity accuracy of approximately 1mm/sec. Several options are available to obtain the necessary accuracy of orbit determination, including tracking the spacecraft using the high precision laser ranging data, and operating an onboard GPS receiver.

6. A PROPOSED JOINT ESA-NASA H-MASER TEST PROGRAM

A joint NASA/ESA H-maser space experiment is being considered.[11] This proposed extension of the current SAO/NASA experiment consists of an addition of a second space H-maser built by the Observatoire de Neuchatel and of a microwave time and frequency transfer system built by the Deutsche Forschungsanstalt für Luft und Raumfahrt, based on a modification of the Precision Range and Range Rate Equipment (PRARE) ranging system[12]. The combined experiment, shown in Figure 7, would yield direct on-board frequency comparison of two H-maser clock systems and would permit high precision time transfer to a large number of ground stations.

Microwave links operating with wide-band PRN codes for time transfer[13] and with continuous-wave Doppler cancellation systems[4] for frequency transfer have been shown to be highly effective. For high precision operation, these techniques include provision for measuring or eliminating ionospheric frequency dispersion and for simultaneous two-way transmissions to remove propagation path errors.

A method for realizing these microwave measurements, based on the existing PRARE[14,15] system developed for the European Remote Sensing Satellites ERS 1 and ERS 2, is being studied. This system makes precise orbit determinations by simultaneous range and range rate measurements to four ground stations. Thirty PRARE ground stations have been deployed for worldwide geodetic position determinations. While the microwave frequency comparisons during each pass would last for only about 7 minutes, and direct *frequency stability* comparisons would be correspondingly limited to about 1 part in 10^{14}, the resulting *time synchronization* (from the correlation of PRN phase modulation of the microwave links between earth and ground) is expected to be well into the subnanosecond domain.[16,11] By recovering the *phase* of the microwave carrier signal from orbit to orbit so as to maintain the phase coherence of the successive frequency comparisons.[17] A much higher precision in frequency comparisons can, in principle, be achieved. This process will require very high precision tracking for relativistic and gravitational corrections and precise measurements of the magnitude of the atmospheric propagation delay to permit "reconnecting" the phase of frequency measurements across the dead times between observations.

7. POSSIBILITIES FOR DEMONSTRATING VERY HIGH PRECISION INTERNATIONAL TIME TRANSFER

Laser ranging stations at Shanghai, China and at Matera, in southern Italy, are associated with nearby VLBI stations equipped with H masers, and could perform time transfer operations at the 100 ps level of precision. EURECA is visible from these stations at low elevation angles as indicated in the spacecraft's ground track shown in Figure 8. To perform a test of a future microwave system for global coverage, there are about 10 existing PRARE ground stations within range of EURECA; in addition, there are other H-maser equipped VLBI stations within EURECA visibility, including the Noto station in Italy, that could be equipped with PRARE ground equipment and laser ranging. Furthermore, additional mobile laser stations could be made available in the USA and in Europe and set up in locations providing visibility of the spacecraft.

8. CONCLUSIONS

Development of space-qualified H-maser clocks and methods for highly precise time and frequency transfer provide enabling technologies for a great many applications, both terrestrial and in space. These applications include (1) time synchronization of primary time keeping centers at the 10 ps level, (2) global synchronization and syntonization of communication and navigation systems, (3) spaceborne and lunar terminals for astronomical and astrophysical research using VLBI radioastronomy techniques over immensely long baselines, (4) arrays of space terminals for gravitational wave detection and (5) deep space probes to the sun for tests of relativistic gravitation. For these technologies to be available for space programs in the next decade, designs for these applications require extensive testing in space and long term life testing to establish their frequency performance in the space environment and to assure reliable uninterrupted operation. The present work is a significant beginning in this important area of technological progress.

9. ACKNOWLEDGEMENTS

This study is supported by the Smithsonian Institution Research Opportunities Fund; maser work at SAO is supported by the NASA under contract number NAS8 39194. The DLR is supporting preparatory studies for proposals to ESA in cooperation with the Institute of Navigation in Stuttgart. The ASI (Italian Space Agency) will support preparatory studies and participation with two laser ranging and timing stations. The initial design study phase of the ESA space maser was supported under the Advanced System

Technology Program (ASTP-3). The European element of the maser clock experiment is under evaluation in the Agency's PRODEX program. The initial design of the Neuchatel Observatory maser was supported by ESA under the Advanced system technology Program (ASTP-3). The European element of the maser clock experiment is under evaluation for inclusion in the Agency's PRODEX program.

REFERENCES

1 R.F.C.Vessot 1991, Applications of Highly Stable Oscillators to Scientific Measurements, *Proc. IEEE* **29**, 7, 1040-1053.

2 G.S. Levy et al. 1986, Results of a very long baseline interferometry demonstration using the NASA tracking and data relay satellite, *Acta. Astron.* **15**, 6/7, p. 481.

3 R. Decher et al. 1980, A space system for high precision global time and frequency comparison of clocks, *Proc. 12nd Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting*, p. 99

4 R.F.C.Vessot et al. 1980, Test of relativistic gravitation with a space-borne hydrogen maser, *Phys. Rev. Lett.*, **45**, 26, 2081.

5 D. W. Allan et al. 1987, Accuracy of international time and frequency comparisons viaa global positioning system satellites in common-view, *IEEE Trans. on Instrumentation and Measurement*, **IM-34**, 572.

6 R. F. C. Vessot, M. W. Levine and E. M. Mattison 1977, Comparison of theoretical and observed hydrogen maser stability limitations due to thermal noise and the prospect for improvement by low temperature operation *Proc. 9th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting*, 549-566.

7 R. Decher, D.W. Allan, C.O. Alley, R.F.C. Vessot,and G.M.R. Winkler 1981, Shuttle experiment to demonstrate high-accuracy global time and frequency transfer, *Proc 1981 International Geoscience and Remote Sensing Symposium*, IGARSS-81, p 370.

8 E.M.Mattison, G.U.Nystrom and R.F.C.Vessot 1979, Design, construction and testing of a small passive hydrogen maser, *Proc. 33rd Ann. Symp. on Freq. Control*, 549-553.

9 CER-VIT is an Owens- Illinois trademark for their glass-ceramic material having high mechanical stability and near-zero coefficient of thermal expansion

10 S. B. Crampton and H. T. M. Wang 1975, Duration of hydrogen-atom spin-exchange collisions, *Phys. Rev.* **A12**,1305.

11 G. Busca, L. G. Bernier, S. Starker, and S. Feltham 1991, In orbit demonstration of a H-maser system, *Proc. 23rd Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting* (in press).

12 P. Hartle, W. Shafer, C. Reigber, H. Wilmes 1987, Precise range and range rate equipment, PRARE, *XIX I.U.G.G. General Assembly, International Association of Geodesy*.

13 S Starker et al. 1990, Navigation experiment NAVEX on Spacelab mission D1, Deutsch Forschungsanstalt für Luft-und Raumfahrt, Oberpfaffenhofen, D-8031 Wessling, Germany.

14 H. Wilmes, C. Reigber, W. Schaefer and P. Hartl 1987, Precise Range and Range Rate Equipment, PRARE, on board ERS 1 orbitology in support of of radar altimtery and a tool for precise geodetic positioning, XIX I.U.G.G.General Assembly, International Association of Geodesy.

15 W. Schaefer and H. Wilmes 1986, Precise Range and Range Rate Equipment, PRARE- system design and status of development, *Workshop on Advances in Satellite Radio Tracking*, Austin, Texas USA.

16 H. Penfield, E. Imbier and R.F.C. Vessot 1982, Design of the Satellite Time and Frequency Transfer (STIFT) microwave ground terminal, *Proc. 14th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting*, p. 223.

17 D.W. Allan, C.O. Alley, N. Ashby, R. Decher, R.F.C. Vessot and G.M.R.Winkler 1981, Ultra-accurate international time and frequency comparison via an orbiting hydrogen maser clock, *Journal de Physique*, **42**, 12, 395.

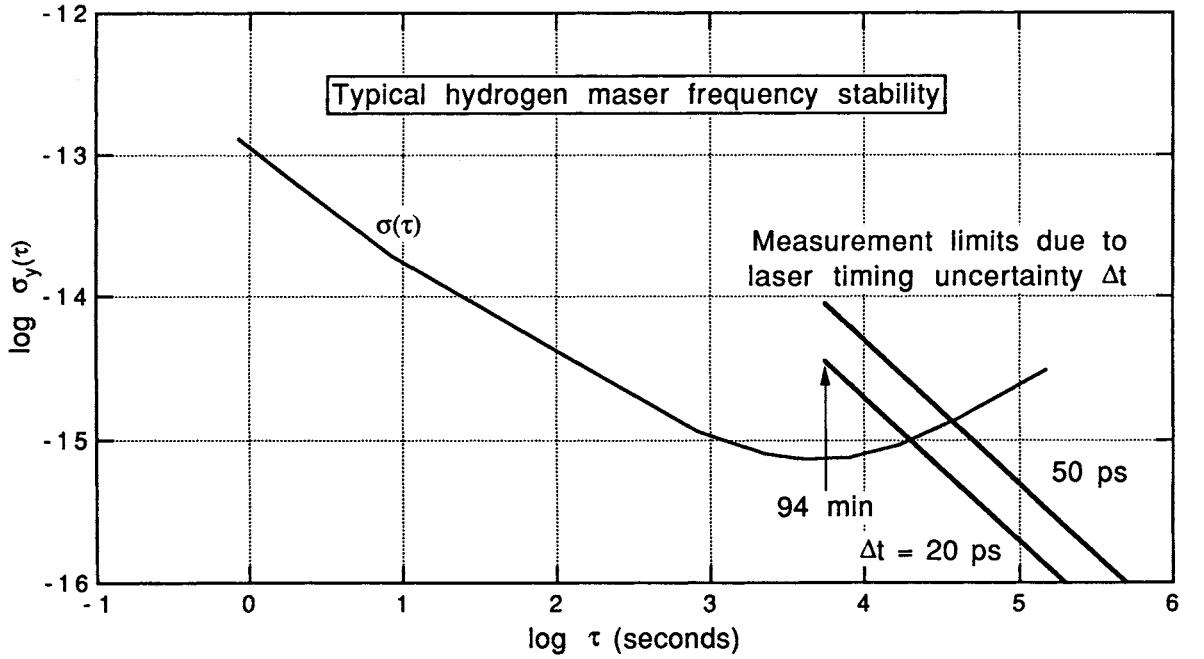


Fig. 1. Space maser expected frequency stability

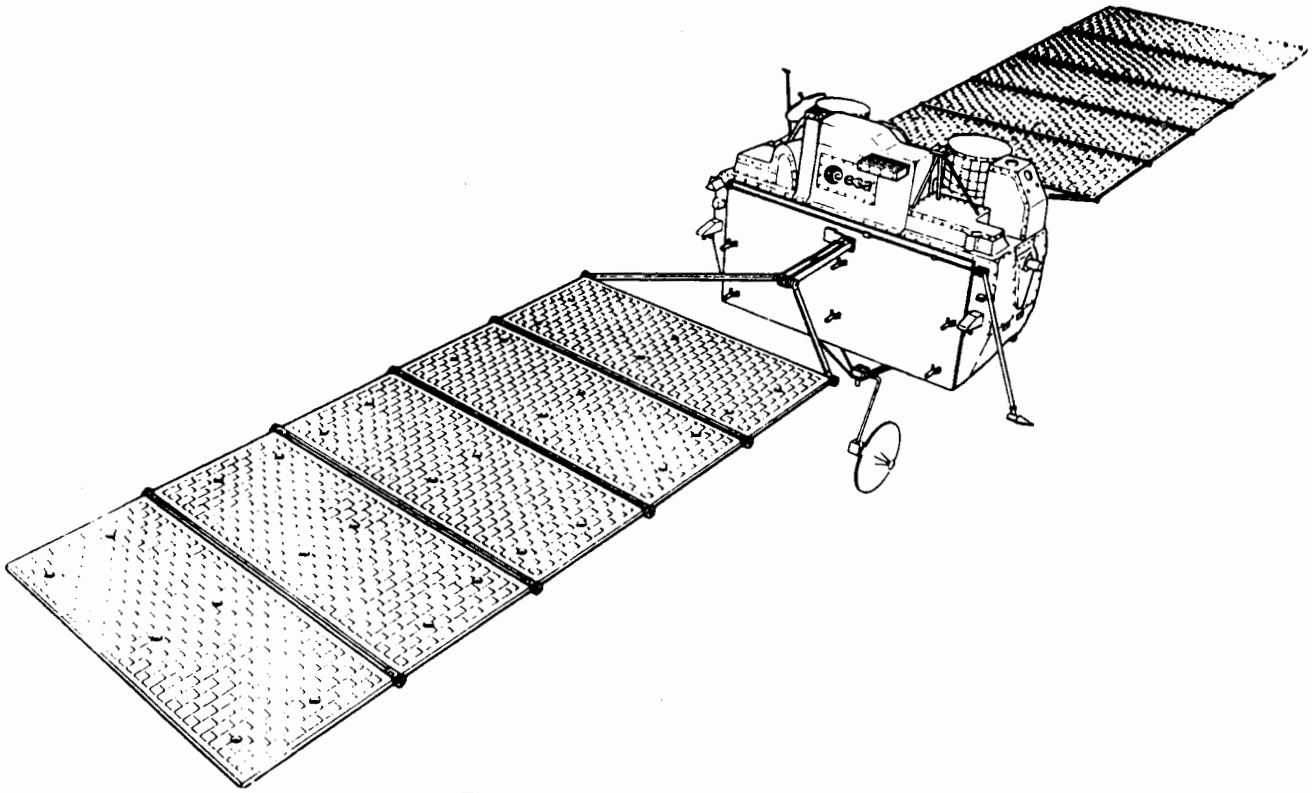


Fig. 2. EURECA spacecraft

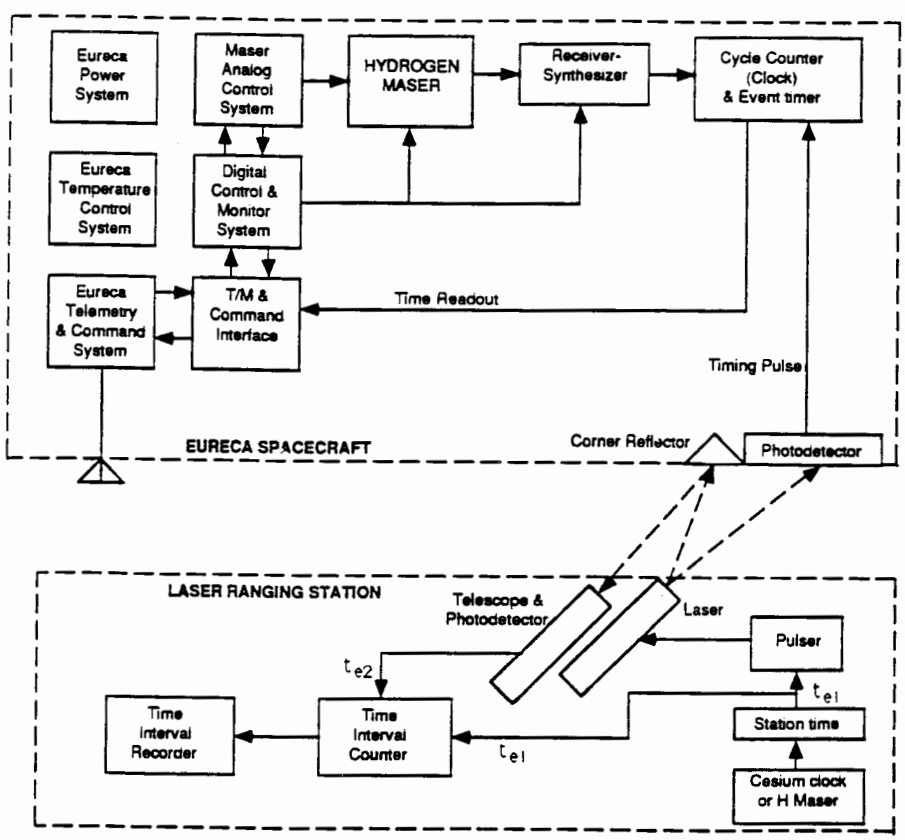


Fig. 3. Electronic systems block diagram

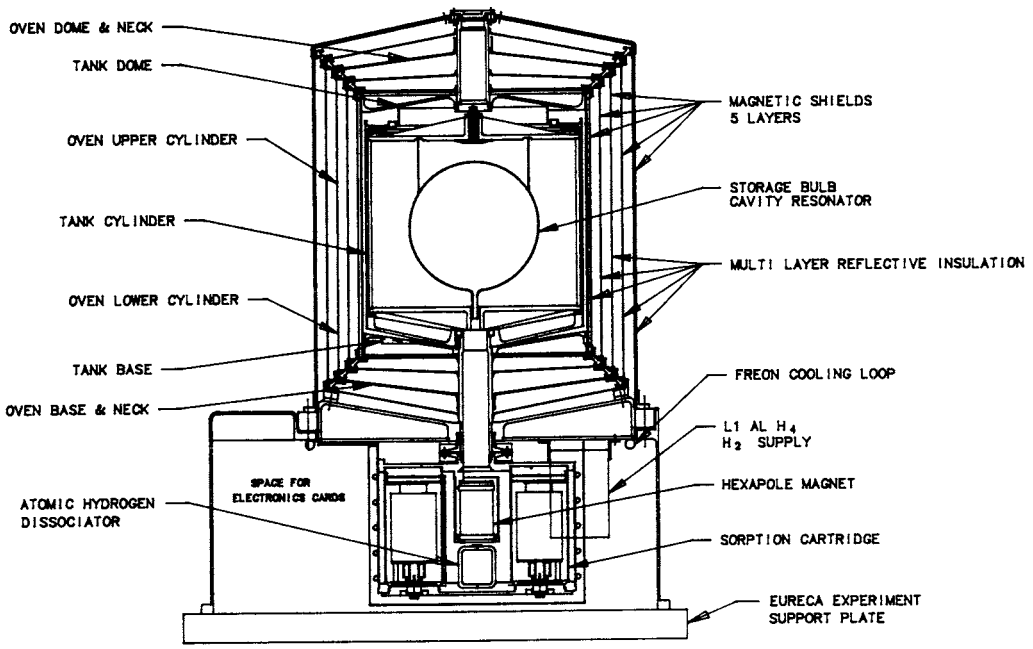


Fig. 4. Cross section diagram of space maser

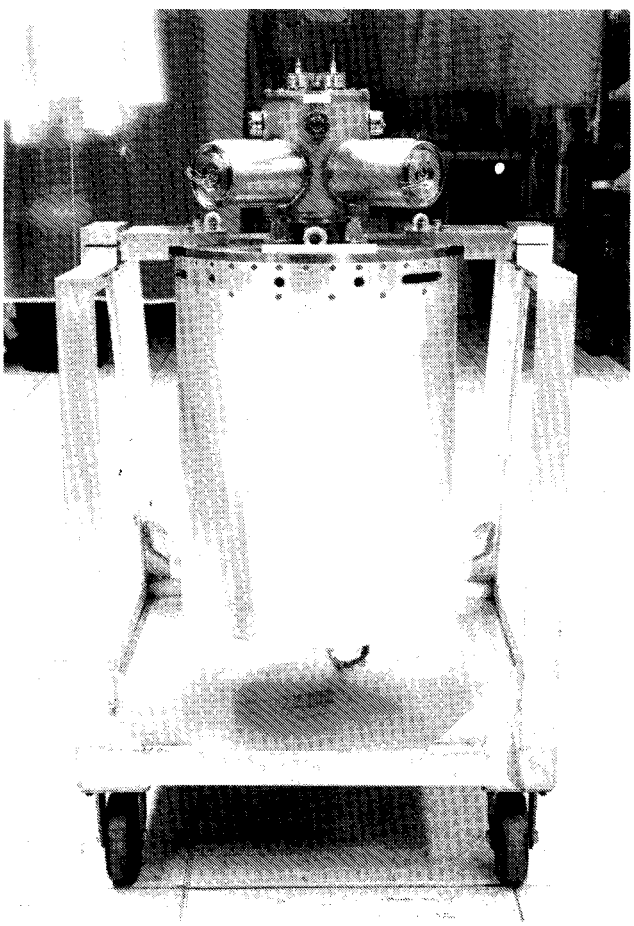
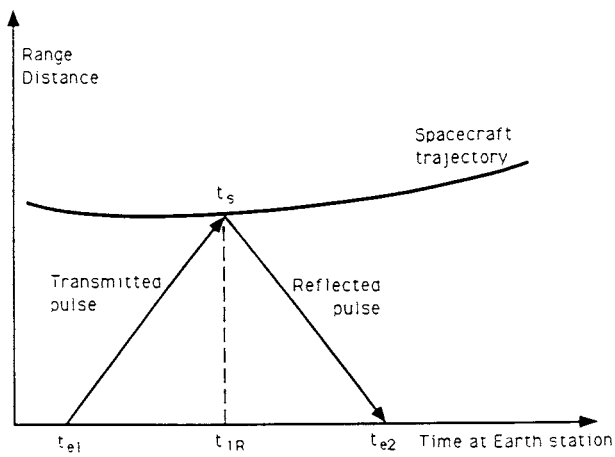


Fig. 5. Photograph of demonstration model of space maser



t_{e1} = time of laser pulse transmission as measured by earth station clock

t_s = time of detection at spacecraft as measured by spacecraft clock

t_{1R} = time of pulse arrival at spacecraft in terms of earth clock $=(t_{e2}-t_{e1})/2$

$\Delta t_{se}=t_s-t_{1R}$ is the difference between the pulse arrival times measured by the space and earth clocks, corrected for propagation delay:

$$\Delta t_{se}=(t_s-t_{1R})=t_s-\left(\frac{t_{e1}-t_{e2}}{2}\right)-t_{e1}=t_s-\left(\frac{t_{e1}+t_{e2}}{2}\right)$$

Fig. 6. Light time diagram of laser time transfer system

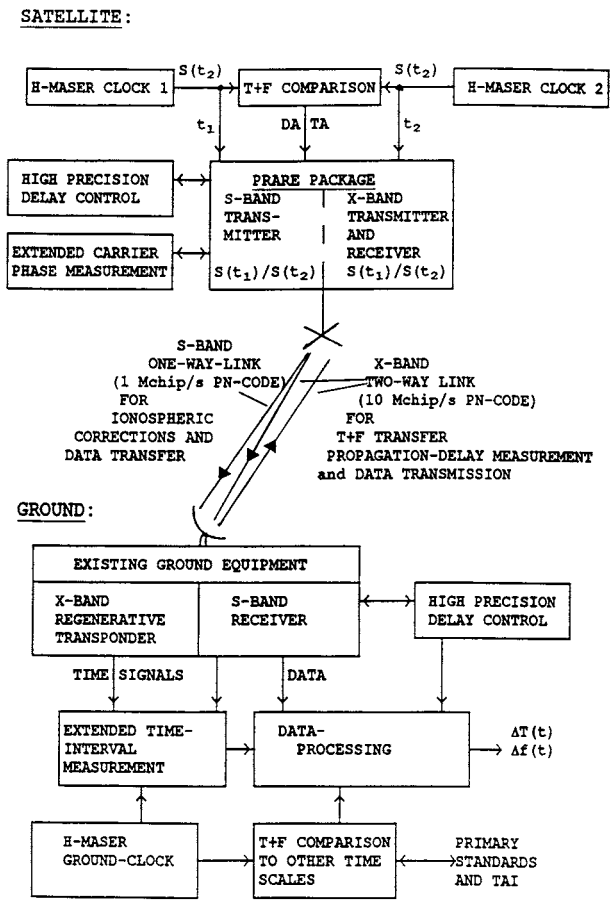


Fig. 7. Blockdiagram of PRARE microwave ranging system

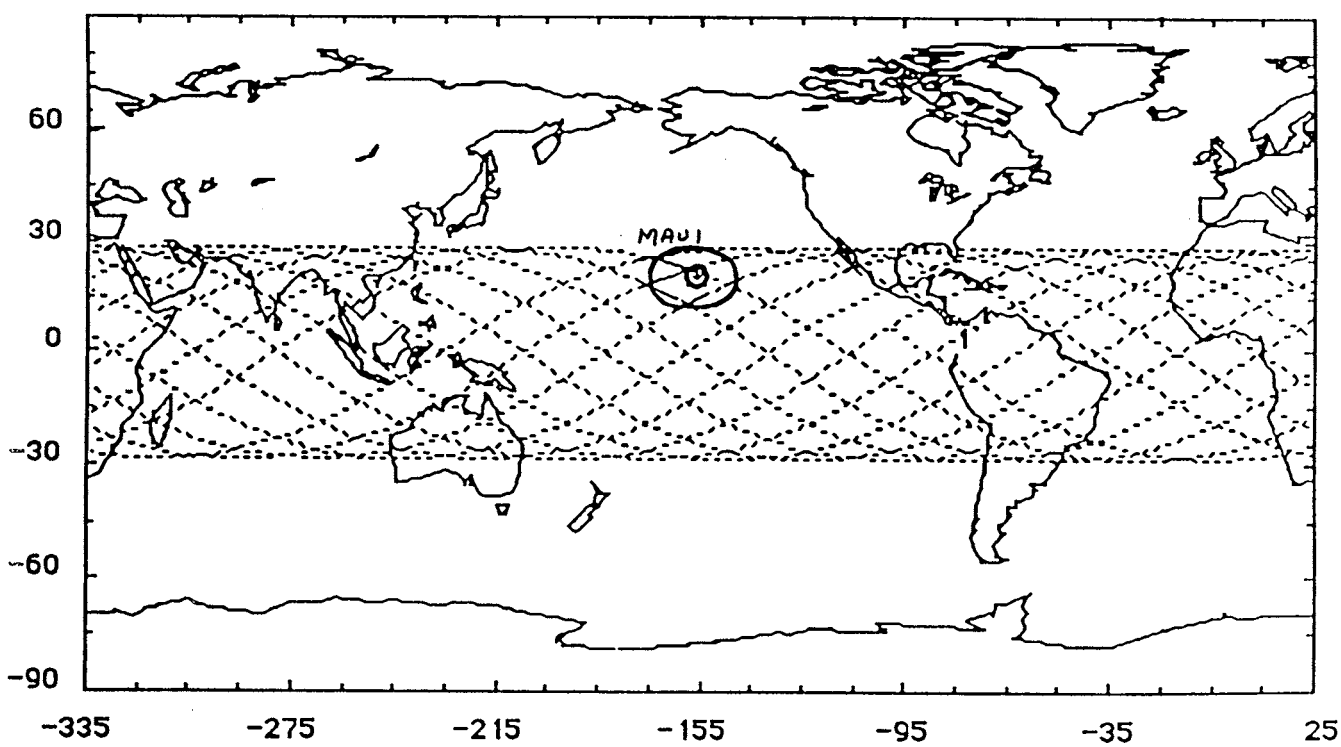


Fig. 8. Simulated EURECA ground track

A SIMPLE AND COMPACT SOURCE OF COLD ATOMS FOR CESIUM FOUNTAINS AND MICROGRAVITY CLOCKS

A. Clairon, Ph. Laurent, A. Nadir
Laboratoire Primaire du Temps et des Fréquences
61, avenue de l'Observatoire, 75014 Paris, France

M. Drewsen*, D. Grison, B. Lounis and C. Salomon
Laboratoire de Spectroscopie Hertzienne de l'E.N.S.
24, rue Lhomond, 75005 Paris, France
*Permanent address : Institute of Physics, Aarhus University, NY Munkegade, 8000 Aarhus C., Denmark

ABSTRACT

Laser cooling and trapping of atoms [1,2] has opened up new exciting possibilities in the field of ultra-stable atomic frequency standards. We recently observed a narrow (2 Hz) Ramsey resonance in a Cesium fountain and we discussed the expected performances of an optimized fountain clock [3]. A crucial element in such a fountain as well as in future microgravity clocks is the source of cold atoms. Here we present a simple and compact source using a magneto-optical trap [4,5], or simply, optical molasses in a low vapor pressure cell and we discuss their performances.

Introduction.

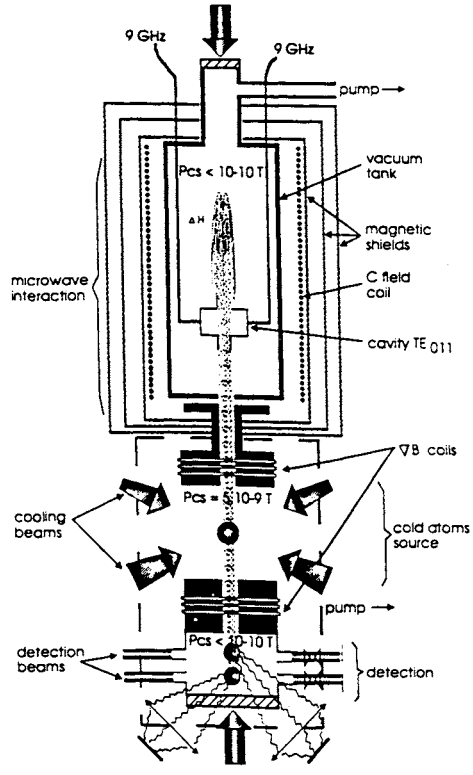
When designing a cold atoms source for a tall (~ 1 m) pulsed fountain many parameters must be considered simultaneously : the number of cold atoms N_i, the initial density, the minimum achievable temperature with the polarization configuration used and the possibility to launch efficiently the atoms. For accuracy reasons (cold collisions [6] and distributed phase shift) the atom density must be low and the atomic position distribution must be nearly uniform over the microwave interaction region. For a tall fountain, the number of detected atoms varies as N_i/T_t (where T_t is the transverse temperature). The final size of the atomic cloud only depends on the temperature; the detected signal is then proportional to the final atomic density. In order to optimize simultaneously the short term stability (which only depends on the number of detected atoms) and the accuracy, a large (Ø ≈ 1 cm), cool and low density initial atomic cloud seems better. Optical molasses work also in a low vapor cell and could be a better atomic source for a clock. However more atoms can be cooled with a $\sigma^+\sigma^-$ magneto-optical trap than with pure optical molasses. Nevertheless, optimum conditions (laser intensity per beam I = 5-6 mW/cm^2, detuning $\delta = -3\Gamma$ where Γ is the natural line width of the 4-5 Cs line, magnetic field gradient ∇B = 8G/cm) for our $\sigma^+\sigma^-$ trap, correspond to an atomic cloud radius of about 1 mm in diameter for 3. 10^8 atoms cooled. Obviously, if the magnetic field gradient is reduced the size of the atomic cloud increases; but simultaneously the number of cooled atoms decreases to reach the atom number in $\sigma^+\sigma^-$ molasses. On the other hand, the number of cold atoms in pure optical molasses increases with the Cs pressure in the cell and approaches the one obtained in a magneto-optical trap at high pressure. The number of cold atoms in molasses also strongly depends on the laser polarization.

These points will be discussed in the next sections.

The Cs fountain clock prototype.

A General design

The prototype being built at LPTF is shown in Fig. 1.



Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

The maximum launching height above the cooling region is 0,8 m, and the maximum height above the microwave cavity is 0,6 m. This geometry allows an interaction time of about 0,7 s with the microwave field. The corresponding width of the Ramsey resonance will then be 0.7 Hz, a factor 100 improvement over the existing conventional Cs clocks. The Cs pressure in the laser cooling region can be adjusted [10^{-9} to 10^{-7} Torr] by acting on the temperature of a cold finger. The laser cooling region is separated from the microwave and detection regions by graphite tubes in order to reduce the background Cs pressure in these regions.

The three magnetic shields, the C field solenoid and the temperature control system are outside the copper vacuum tank. Two 1cm diameter holes are drilled in the cylindrical TE011 cavity. The atoms, after the microwave interactions, are detected by fluorescence. In order to reduce the noise due to fluctuations of the number of atoms two detection beams are used to normalize the signal to the total number of cold atoms [12].

B The cold atom source.

The experimental set-up allows both magneto-optical trapping (∇B on, $\sigma^+\sigma^-$ polarized beams) and pure optical molasses experiments. For preliminary studies of the cold atom source, independently controlled laser beam pairs are used. This arrangement allows measurements of the temperature and damping time of the atomic kinetic energy in 1 Dimension or 3 Dimensions.

The three cooling beams pairs are provided by a 150 mW (5422) Spectra Physics diode laser injection locked. The master diode laser (HLP 1400) is "spectrally narrowed" by optical feedback from a confocal Fabry-Perot and frequency stabilized 2 MHz below the center of the F = 4 -> F' = 5 saturation peak in a vapor cell. 50% of the power of the master laser is also used to realize the probe beam for the temperature measurements.

In order to control accurately and independently the trapping laser frequency and intensity we use acousto-optic modulators in a double path configuration. The first one, A.O.1, is placed on the master laser before injecting the slave laser. By tuning the A.O.1, one is able to control quickly (10-20 μs) the trap laser frequency. The slave beam is split in 3 equal intensity cooling beams by using a combination of half wave plates and polarizing beam splitters. Then, each beam is passed through A.O.2 to partly compensate the 2x75 MHz shift of A.O.1 and to allow a flexible control of the cooling intensity.

The three cooling laser beams are spatially filtered and expanded to 1,4 cm in diameter ($1/e^2$ points). The three standing waves are simply obtained by retroreflecting the beams. The intensity imbalance in each arm is less than 5% (without Cs absorption); the maximum laser intensity is about 6 mW/cm^2 per beam.

We noted that the trap was working very well with polarizations far from being circular as long as the ellipticity of the polarization along each arm keeps the correct sign. The horizontal and vertical cooling beams are orthogonal within ~ 1 mrd, and the counter propagating beams are collinear to within 0,3 mrd.

With this arrangement, we are able to load the trap or molasses with reproductible conditions and constant atom number (I, δ and ∇B fixed) and then switch with an adjustable time constant to the desired I and δ final values. An adjustable thermalization period [10 μs to 550 ms] is allowed before the cooling beams are turned off in a few μs and the magnetic field in less than 2 ms. A repumping laser resonant with the F = 3 -> F' = 4 transition is superimposed with the horizontal cooling beams.

The experimental set-up used to study the cold atom source is shown in Fig.2. The temperature is measured by a time of flight method [TOF]. The number of cold atoms can be measured from the fluorescence of the cold atoms cloud or deduced from the TOF signal. The size of the atomic cloud is measured by a CCD camera. The friction coefficient along the vertical beam is simply deduced from the damping time τ_α of the temperature $M_{Cs}/2\alpha = \tau_\alpha$ where M_{Cs} is the Cs atomic mass and α is the friction coefficient. The spatial diffusion coefficients of atoms in molasses can be determined from the measurement of the radius of the atomic cloud as a function of time after the trapping magnetic field ∇B is turned off.

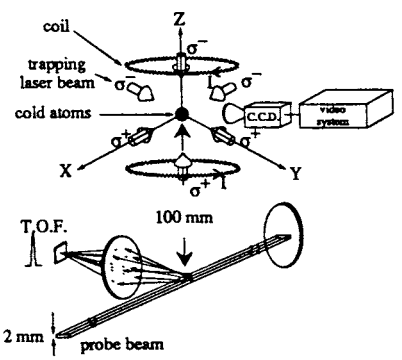


Fig. 2 : experimental set-up.

Temperature measurement of atoms in $\sigma^+ - \sigma^-$ molasses and traps

A Molasses.

Temperature of molasses using linear polarizations in 3D have been previously measured for two different configurations [7] :
- each beam pair is made of counter propagating beams with orthogonal polarizations (lin⊥lin)
- each beam pair is made of counter propagating beams with parallel polarizations (lin//lin).

The $\sigma^+\sigma^-$ configuration presents a particular interest since it corresponds to the configuration used for the magneto-optical trap [MOT]. For temperature measurements in $\sigma^+\sigma^-$ molasses the atoms are first loaded in the MOT and then released in a $\sigma^+\sigma^-$ molasses situation by switching off the ∇B field in 2 ms.

The first remarkable point is that the time needed to reach a steady state molasses temperature, starting from trapped atoms, depends on the initial density. This thermalization time is 10 to 50 ms or 100 to 1000 times the expected typical cooling time $\tau_\alpha = M_{Cs}/2\alpha$ of Sub-Doppler cooling mecanisms [see Fig. 3].

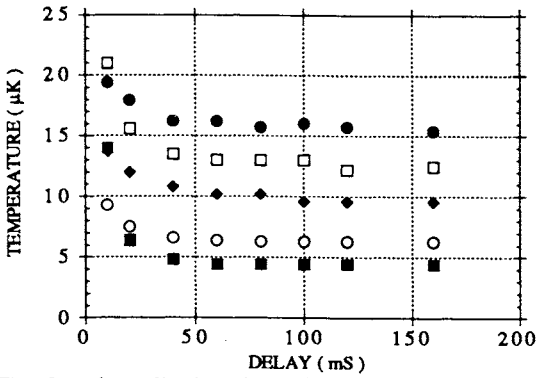


Fig. 3 : thermalization time of molasses temperatures.
$\delta = 8\Gamma$. Ω^2/Γ^2= ● 2 □ 1,2 ◆ 0,8 ○ 0,4 ■ 0,2
where Ω is the Rabi frequency.

As a consequence, it is natural to attribute this anomalously long time to interatomic interactions. In order, to properly measure the temperature of the atoms in the molasses without this effect, the atoms were left diffusing in the molasses for 100 ms before switching off the laser beams for the TOF measurement.

For a given laser detuning, there exists also a threshold intensity below which the temperature rises again indicating that the Sub-Doppler $\sigma+\sigma-$ cooling mecanism ceases to function properly; this is accompanied with a dramatic decrease of the number of cold atoms in the molasses volume. When the temperature is plotted as a function of $\Omega^2/\delta\Gamma$ (Fig. 4), where Ω is the Rabi frequency of each of the travelling waves, we do not find a single curve as in the lin⊥lin case [7]. There is an excess temperature for small detunings. We recover a single curve only for detunings greater than 3Γ. This observation is also in qualitative agreement with the 3-D numerical calculations of K. Möllmer [9]. A least square fit of the data for $|\delta| \geq 5\Gamma$ gives the following result :

$$k_B T = C_{\sigma^+\sigma^-} \hbar \frac{\Omega^2}{|\delta|} + C_0$$

where k_B is the Boltzmann constant, Ω and δ are expressed in units of Γ, $C\sigma+\sigma- = 0.24 \pm 0.05$ and $C_0/k_B = 2.6 \pm 0.4\ \mu K$.

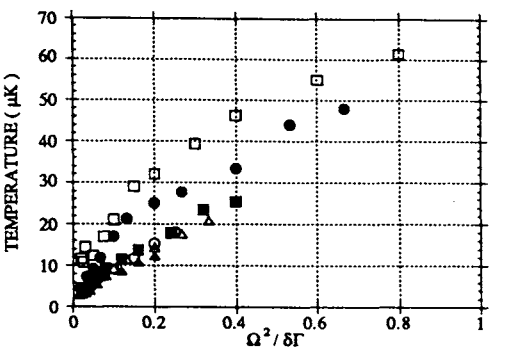


Fig. 4 : molasses temperatures.
δ= □ 2Γ ● 3Γ ■ 5Γ △ 6Γ ○ 8Γ ▲ 10Γ

In Fig. 5 is plotted the minimum achievable $\sigma+\sigma-$ molasses temperature as a function of detuning and compared with the minimum temperature for linear polarizations measured in [7]. For large detunings, the minimum temperature is quite similar (2.5 μK), whereas at small detunings the temperature is systematically higher for the $\sigma+\sigma-$ case.

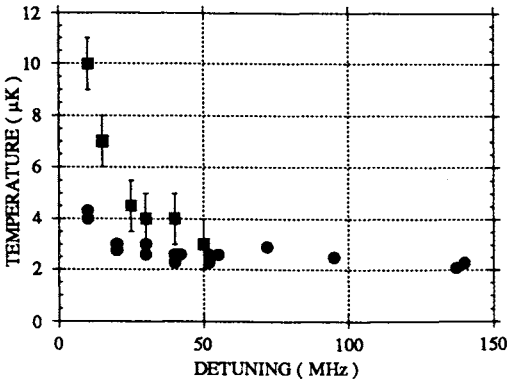


Fig. 5 : minimum molasses temperature.
●linear polarization ■ $\sigma^+ - \sigma^-$ polarization

B Magneto-optical trap.

The number of atoms that can be trapped in a MOT depends on laser detuning δ, laser intensity I, ∇B field and on the area of the intersection volume of the laser beams.

The steady state number of atoms at equilibrium near the center of the MOT can be written as [4,8]:

$$N_{st} = \frac{\Pi}{2}\frac{S}{\sigma}\left(\frac{v_c}{\bar{v}}\right)^4\left(1-\frac{\tau_p}{\tau_c}\right) \quad (1)$$

where S is the area of the capture volume, σ is the cross section for collisions that eject a cold atom from the trap, v_c is the maximum velocity captured by the trap ($kv_c \equiv \delta$ for optimized conditions. k is the laser wave number), $\bar{v}$ is the atomic thermal velocity in the cell, τ_p is the damping time constant of the atom positions in the trap and $\tau_c = \frac{1}{\Gamma_c} = \frac{2\Pi^{3/2}}{\rho v \sigma}$ is the mean lifetime of the atoms in the trap due to a collision between a cold atom and a thermal atom in the cell. ρ is the atom density in the cell. The time constant for loading the trap is simply $\tau_{load} = \tau_c$. The loading time is typically 1 s at 10^{-8} torr partial Cs pressure. Equation 1 is only valid if collisions with background gases other than Cs, cold collisions and atom-atom interactions due to multiple photons scattering are negligible. In our trap, I = 6 mW/cm^2 and $\emptyset_{laser} \sim$ 1 cm, the optimum loading conditions are $\delta = -3\Gamma$ and $\nabla B \approx$ 8 G/cm. In these conditions 3.10^8 atoms can be trapped.

When sub-Doppler mechanisms are taken into account in the trap operation [5], it is anticipated that, even in the presence of the magnetic field gradient, the trapped atoms should have a temperature identical to that of $\sigma^+\sigma^-$ molasses, i.e. in the 2 to 100 μK range. Our temperature measurements confirm very well this prediction : more than 10^8 atoms can be trapped and still have a sub-Doppler temperature. However the molasses temperature is only reached when the density in the trap is low enough, typically 10^{10} atoms/cm^3 at small detunings and 10^{11} atoms/cm^3 at large detunings (6 to 10 Γ). For a large number N of atoms in the trap, we observe as in [12] that the size of the atomic cloud increases as $N^{1/3}$ because of atom-atom interactions. This effect is accompanied by an increase of the atoms temperature as shown in Fig. 6. We find a temperature excess proportional to $N^{1/3}$ and nearly proportional to $1/|\delta|$. Clearly this temperature excess is not due to the magnetic field gradient since it persists several tens of ms after switching off the magnetic field (fig.3); it could be due to atom-atom interactions such as multiple photons scattering or cold atoms collisions. More work is necessary to understand more completely this regime in the trap.

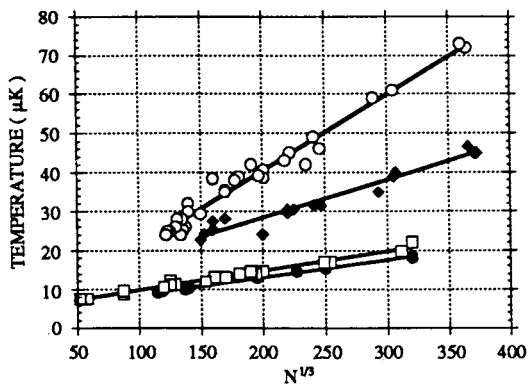


Fig. 6 : trap temperatures as a function of the cubic root of the atom number in the trap.
δ = —●—6Γ —□—5Γ —◆—3Γ —○—2Γ

Friction coefficient measurements.

The temperature measurements solely give information on the momentum diffusion coefficient over the friction coefficient ratio. The two 1D gradient polarization theories [1] predict very distinct behavior of the friction coefficients : $\alpha_{lin\perp lin}$ (resp. $\alpha_{\sigma+\sigma-}$) should increase proportionnaly (resp. inversely) to the laser detuning. In order to test these predictions in the 3D geometry, we have performed direct measurements of the friction coefficients in optical molasses. These measurements are also useful to anticipate the maximum acceleration γ_{max} that can be given to the atoms by using "moving molasses" for the launching [2,3]. The maximum acceleration is indeed approximately given by :

$$\gamma_{max} \leq \frac{\alpha}{M_{Cs}} V_c$$

where Vc is the critical velocity for which the sub-doppler friction force is maximum. V_c is not accurately known in a 3D experiment but it can be estimated from the launching efficiency given in [3] (about 10 cm/s for I= 4 mW/cm^2 and $\delta = 2\Gamma$). V_c decreases when increasing the laser detuning and increases with the laser intensity.

The friction coefficient of 3D molasses can be determined by measuring the damping time constant τ_α of the temperature after a sudden reduction of the laser intensity. The friction coefficient is measured along the z axis and is simply :

$$\alpha = \frac{M_{Cs}}{2\tau_\alpha}$$

With the apparatus described previously we can also measure τ_α in 1 dimensional molasses. 1D measurements are performed by measuring the time constant to reach 1D equilibrium by

starting from a 3D molasses temperature. This is simply done by switching off the horizontal beams and measuring the temperature as a function of the delay between 1D and 3D molasses.

The results are together presented in Fig. 7 with some theoretical predictions for various polarization configurations. The first remarkable point is that the friction coefficient in 3D is much larger than the Doppler friction coefficient but is nearly a factor 10 smaller than in 1D molasses; note that our measurements in 1D agree very well with the 1D $\sigma+\sigma-$ theoretical prediction [11]. The second point is that the friction coefficient is nearly independent of the polarization configuration. This result contradicts the 3D predictions of K. Möllmer [9].

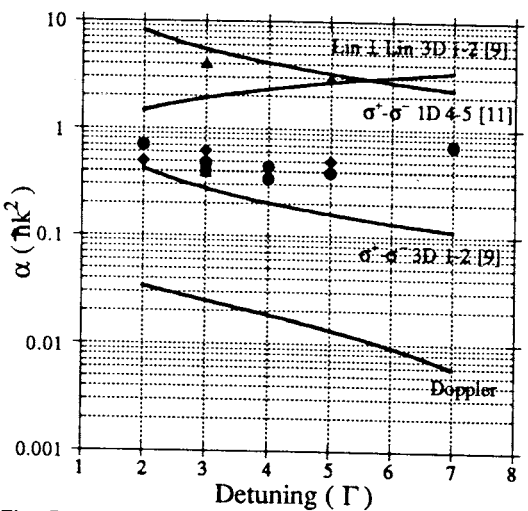


Fig. 7 : α values for different polarization configuration as a function of the laser detuning. Plain line : theoretical predictions. Experimental results : ● $\sigma^+-\sigma^-$ 3D ◆ lin ⊥ lin 4-5 ▲ $\sigma^+-\sigma^-$ 1D ■ lin // lin 4-5

We have also tried to deduce indirectly α from the diffusive motion of the atoms in molasses. For a pure Brownian motion we expect a mean square atomic cloud radius $\langle x^2(t)\rangle$ given by:

$$\langle x^2(t)\rangle - \langle x^2(t_0)\rangle = 2\,D_x(t-t_0) \quad (2)$$

where Dx is the spatial diffusion coefficient. If the friction coefficient α is independant of the atomic velocity and isotropic, the spatial diffusion coefficient can be written as:

$$D_x = \frac{k_B\,T}{\alpha}$$

We have simultaneously measured the spatial diffusion coefficient and the temperature of $\sigma+\sigma-$ molasses at various laser detunings. The deduced friction coefficients are systematically lower, by a factor from 4 to 10, than those directly measured.

There is also another interesting difference between 1D and 3D $\sigma+\sigma-$ molasses. When an homogeneous magnetic field (< 500 mG) is applied along a laser beam, the atoms, in 1D molasses, are cooled around a non-zero velocity $V=\frac{\omega_R}{2k}$ (ω_R : Zeeman frequency)[10] independent of the laser detuning; in 3D molasses, the drift velocity decreases with the laser detuning and is accompanied by an heating of the atoms. In 3D, the drift velocity remains systematically below its 1D value.

Number of cold atoms in molasses and traps

In writing equation (1), it is assumed that the losses in the trap are only due to collisions with background Cs atoms. When the Cs pressure is increased, a cooled atom cannot reach its equilibrium position before a collision ejects it from the trap. However the atoms are still cooled by Doppler and SubDoppler molasses when $\tau_C >> \tau_D$ (caracteristic damping time of the velocity for Doppler cooling). When $\tau_p >> \tau_C >> \tau_D$, the size of the atomic cloud increases up to the laser beam diameter. This case is very similar to a pure molasses configuration in a high Cs pressure cell. The presence of the magnetic gradient field in the trap modifies the Doppler capture velocity v_C in two ways: first, it modifies the atomic resonance frequency and consequently the friction force; second, it bends the atomic trajectories towards the trap center and then increases the effective slowing distance. In our optimum experimental loading conditions, the laser detuning is much larger than the Zeeman detuning and we believe that the first contribution is negligible. The second effect is much more difficult to estimate but it will not strongly modify the capture velocity when $\tau_C \le \tau_p$.

The steady state number of atoms in molasses can be obtained by a simple modification of (1):

$$N_{st} = \frac{\Pi\,S}{2\,\sigma}\left(\frac{v_C}{\bar{v}}\right)^4 \frac{\Gamma_c}{\Gamma+\Gamma_c} \quad (3)$$

where Γ is the loss rate due to other causes than collisions : spatial diffusion, atomic drift due to laser beams imbalance or parasitic magnetic fields. These losses are suppressed in a trap but not in molasses. If $\Gamma c \approx \Gamma$ then the number of atoms in the molasses becomes similar to the one in the trap.

Experimentally we find that the ratio of the number of atoms in molasses to the number in the trap, N_{mol}/N_{trap}, increases with the Cs pressure and reaches 0.2 for lin⊥lin molasses, 0.07 for $\sigma+\sigma-$ molasses and 0.04 for lin//lin molasses at a Cs pressure of 10^{-8} Torr. At higher Cs pressure, we observe a decrease of these ratios probably due to a strong absorption by the uncooled atoms in the cell leading to a laser intensity reduction and to a significant intensity imbalance. Note that, in our conditions, more than 10^7 atoms can be cooled in a lin⊥lin molasses. We also observe that the loading time constant in a lin⊥lin molasses is always larger than the one measured in the other configurations.

Toward a microgravity clock

Clearly, on earth,the gravity acceleration sets a limit to the achievable resolution: a 1m high fountain is required to produce a 1s interaction time between the two Ramsey interactions. Since the resolution increases only as the square root of the fountain height, it is tempting to investigate the limiting case where the gravity is turned off as for instance in a satellite. This may turn out to be the only way to take full advantage of the microKelvin temperature of the laser cooled atoms and to gain another factor 10 in resolution with a relatively simple and compact set-up. Such a satellite clock would have both an excellent stability and, most importantly, a superior potential accuracy. As we discussed in [EFTF1991], for long (1-10 s) Ramsey interrogation times, it is the transverse temperature of the atoms, which limits the achievable stability. More precisely, in this regime, we showed that the Allan variance will degrade only as the square root of the interrogation time. The advantage being mostly on the accuracy aspects, the small volume and simplicity of use.

A proposed microgravity clock is depicted in Fig.8 and is in essence much simpler than an Earth fountain because it does not require sophisticated launching techniques. As in Fig.1 it consists of two separate chambers: a cooling chamber where the atoms are captured and cooled in optical molasses or in a magneto-optical trap from a low vapor pressure Cs cell and a low pressure chamber ($p<10^{-10}$ Torr) in which the micowave interaction takes place. Atoms are transferred from the cooling chamber into the interrogation chamber using a weak longitudinal magnetic field in $\sigma+\sigma-$ 1D optical molasses that produces a longitudinal drift velocity of 5 to 10 cm/s. This acceleration phase is followed by a short phase of transverse cooling down to 2.5 μK without degrading significantly the longitudinal velocity spread. The atoms then drift slowly through a 30 to 50 cm long microwave TE_{01n} cylindrical cavity (or a more traditional U shaped two zone Ramsey cavity) and are finally detected optically. This scheme would yield an interrogation time in the 5 to 10 seconds range, a factor 10 improvement over the Earth fountains. The total length of the apparatus may be less than 0.5 m and the system requires only two frequency controlled diode lasers. It may easily be tested in the milligravity offered by the parabolic fligths, although the 10^{-2} g residual gravity fluctuations will impose a lower limit to the choice of the drift velocity and therefore will limit the resolution. A research program aiming at testing the basic elements of such a microgravity clock has started with the support of the french center for spatial studies CNES.

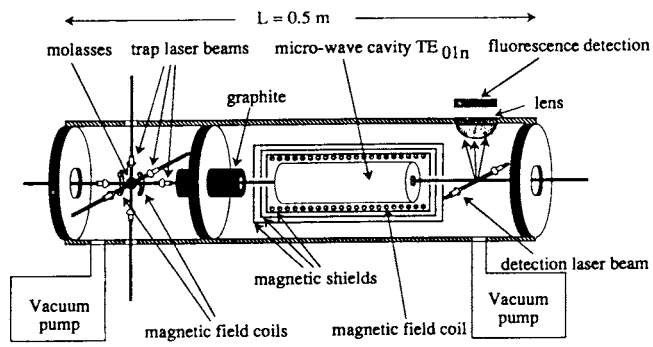


Fig.8 : a microgravity Cesium clock.
Cesium atoms captured in a low vapor pressure cell are slowly ($V_l \cong$ 5-10 cm/s) sent into a 30 cm long microwave cavity and optically detected. The interrogation time of 5-10 s should lead to a 0.05-0.1 Hz resonance width.

Conclusion

We have shown that optical molasses or magneto-optical trap in a cell can be used as a source of very cold atoms for a pulsed fountain clock. With proper operating conditions, an almost equal number of atoms can be cooled in both situations. An obvious advantage of the molasses lies in the absence of any pulsed magnetic field near the C field region. However the possibility to use an MOT in the same apparatus will be very useful to study the effect of cold collisions on the clock frequency and to estimate the distributed phase shifts.

We have measured the friction coefficient of molasses in 3D; the friction values obtained are lower than expected. For high fountains, the launching using 3D moving molasses could become a problem with low power lasers. The large friction, measured in 1D molasses, would allow a much larger acceleration of the atoms.

In this case, the acceleration phase must be followed by a 3D cooling phase to avoid a large transverse spread of atoms induced by the spontaneously emitted photons that would lead to a severe signal reduction.

To summarize, cold atoms offer a promising way to improve by two orders of magnitude the performances of the present day primary Time standard. In the future, such ultra-stable clocks will then be useful tools in various branches of physics and space sciences: from testing the fundamental laws of physics such as general relativity or CPT invariance to astrophysics for timing millisecond pulsars or VLBI and to geophysics for precise ranging on the Earth (GPS) or in space.

References

[1] See for instance Special issue of JOSA. B **6** (1989).

[2] M. Kasevich et al., Phys. Rev. Lett. **63**, 612 (1989).

[3] A. Clairon, C. Salomon, S. Guellati and W. D. Phillips, Europhys. Lett. **16**, 165 (1991).

[4] C. Monroe, W. Swann, H. Robinson and C. Wieman, Phys. Rev. Lett. **65**, 1571 (1990).

[5] A. M. Steane and C. Foot, Europhys. Lett. **14**, 231 (1991).

[6] E. Tiesinga, B. Verhaar, H.T.C. Stoof and D. Van Bragt, Phys. Rev. A**45** (1992).

[7] C. Salomon, J. Dalibard, W.D. Phillips, A. Clairon and S. Guellati, Europhys. Lett **12**, 683 (1990).

[8] D. Grison, PhD Thesis Paris VI, (1992).

[9] K. Molmer, Phys. Rev. A **44**, 5830 (1991).

[10] S.Q. Shang, B. Sheehy, P. Van Straten and H. Metcalf, LIKE conference, Elba, 79- (1990) and Phys. Rev. Lett. **65**, 317 (1990).

[11] Y. Castin, PhD Thesis Paris VI (1992).

[12] D.W. Sesko, T.G. Walker and C.E. Wieman, JOSA B **8**, 946 (1991).

MICROWAVE FREQUENCY DISCRIMINATOR WITH A CRYOGENIC SAPPHIRE RESONATOR FOR ULTRA-LOW PHASE NOISE *

G. John Dick and David G. Santiago
California Institute of Technology, Jet Propulsion Laboratory
4800 Oak Grove Drive, Pasadena, California 91009, U.S.A.

ABSTRACT

First results are presented for an X-band frequency discriminator using a cooled sapphire microwave resonator. These results show a lower close-in (1Hz - 1kHz offset) phase noise measurement floor than any oscillator presently available. This performance is made possible by a sapphire "whispering gallery" mode resonator which shows the highest quality factor (with Q's up to 30 million) for any rf, microwave, or acoustic resonator at temperatures down to 77K. Performance is further increased by the use of novel phase detection circuitry. The sapphire discriminator was used to characterize the phase noise of a single crystal quartz oscillator of the highest quality, without the use of a second similar oscillator as reference.

Keywords: Phase Noise, Cryogenics, Microwave Oscillator, Whispering Gallery Resonator

INTRODUCTION

The fundamental instrument for spacecraft Radio Science experiments is the ultra-stable oscillator (USO), providing signals with spectral purity and stability to obtain data on fluctuations in phase, frequency, and intensity of radio signals as they propagate through intervening media before Earth reception[1]. The data yields information on planetary atmospheres and ionospheres, as well as information on planetary interiors by observing the perturbation of signals due to gravity fields. The spectral purity of the signal derived from the USO determines the resolution and accuracy of measured parameters in radio science investigations.

We propose use of a sapphire phase stabilizer (SPS) operating at X-band or higher to compensate phase fluctuations in the multiplied frequency of the USO[2]. As compared to the best conventional sources at offset frequencies of 1 Hz and higher, there is a potential for 30 dB or more improvement in phase noise for a sapphire resonator operating at 80 Kelvin with a critically coupled Q of 15 million. For comparison, the Q of the best quartz crystal oscillators is about 2 million.

The SPS is based on a novel technology consisting of a cooled sapphire "whispering gallery" mode X-band resonator operating in the temperature range from 70 Kelvin to 300 Kelvin - values achievable by means of radiative and thermoelectric cooling. In support of radio science experiments, the SPS would be used to sense and compensate phase fluctuations in X-band signals derived from conventional or ultra-stable oscillators on board spacecraft and in communications support systems on the ground.

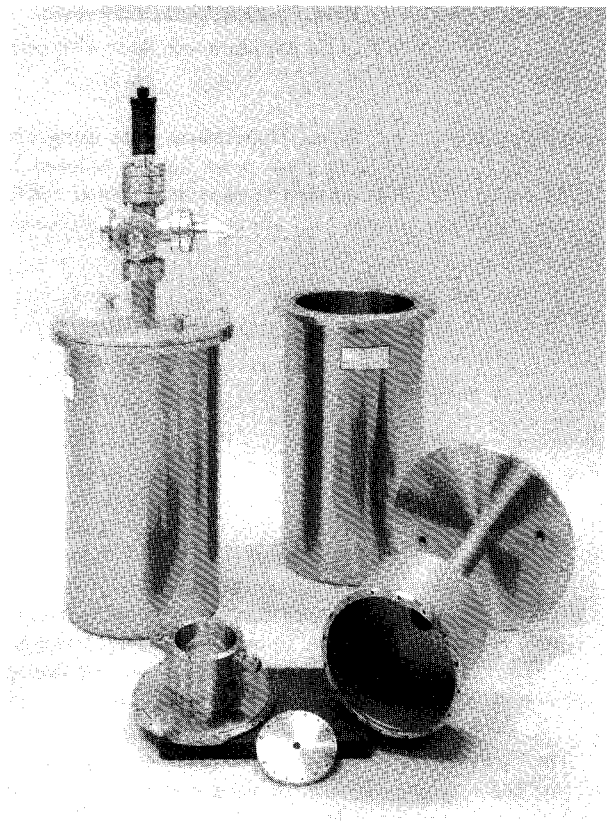


Figure 1: Cryogenic components for the two sapphire resonator subsystems.

We are now able for the first time to accurately sense and measure the phase noise of the quietest available quartz oscillator with a noise floor reduced by 10 to 45 dB for offset frequencies from 1 Hz to 1000 Hz. For operation at 80 Kelvin, this noise floor is almost exactly as predicted for the measured Q of 6 million; it would be reduced by another 8 dB at the design Q of 15 million.

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

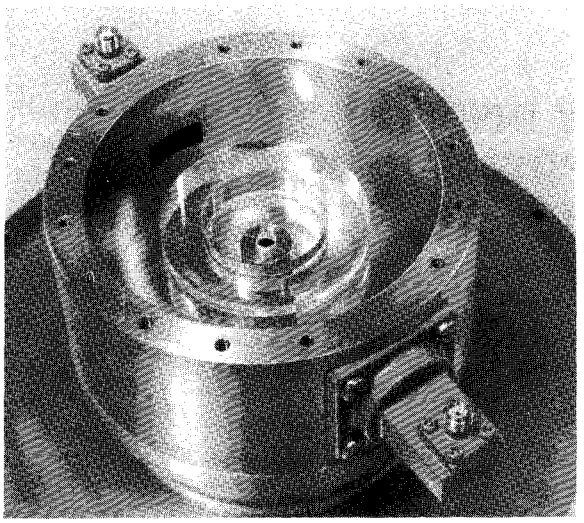


Figure 2: Close-up view of the sapphire resonator with cover removed.

Also significant are room-temperature tests using new rf circuitry[3]. These tests show noise further reduced by approximately 10 dB compared to the conventional configuration. This circuitry is also applicable to the 80 Kelvin SPS.

BACKGROUND

Resonator Performance

Dielectric losses in sapphire are rapidly reduced as the temperature is cooled below ambient, showing an approximate T^{-5} dependence down to temperatures as low as 60 Kelvin[4–7]. While metallic losses limit the achievable Q values in most microwave resonators, recently developed "whispering gallery" mode sapphire resonators isolate the resonant microwave energy to the sapphire itself, and so allow the inherent Q value of the sapphire to be realized[4–9]. Unloaded Q values at 8–10 GHz (X–band) for these resonators are approximately $2 \cdot 10^5$ at room temperature, increasing to $3 \cdot 10^7$ at 80 K.

The thermal coefficient of frequency of the sapphire resonators has also been studied, showing a value of $5 \cdot 10^{-5}$ at room temperature and reducing to $3 \cdot 10^{-6}$ at 80 K[7]. Even this reduced coefficient at 80 K limits the resonator's applicability as a frequency stabilizing element. However, application as a phase stabilizer is not adversely effected for the relatively higher fluctuation frequencies above 1 Hz, where the thermal mass of the sapphire resonator effectively stabilizes its operating frequency.

Phase Noise Considerations

Phase noise performance for rf and microwave cavity and DRO oscillators at small offsets ($f < 1000$ Hz) is determined by resonator Q and by the 1/f noise level in the active device[10–13]. For slow phase fluctuations ($f \ll \nu/Q$), the characteristic phase slope of the resonator $\partial\phi/\partial\nu = 2Q/\nu$ implies a corresponding slow fluctuation in the frequency of the oscillator. Here f represents the fluctuation frequency, ν the microwave frequency, Q the quality factor of the resonator and ϕ the phase of the microwave signal. In this way, a power spectral density of phase fluctuations for the active

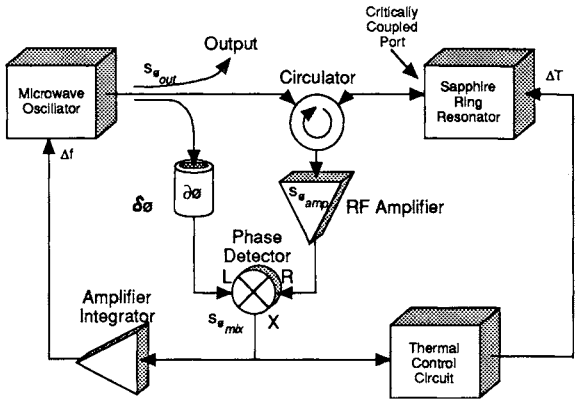


Figure 3: SPS configuration for reduced phase noise. Critical coupling to resonator and operation very near ν_o allow insensitivity to phase noise of amplifier. Amplifier gain allows reduction of phase detector noise.

device $S_\phi(f)|_a$ results in an oscillator output frequency noise spectral density of

$$S_y(f)|_{out} = \frac{S_\phi(f)|_a}{(2Q)^2} \tag{1}$$

or the mathematically equivalent output phase noise

$$S_\phi(f)|_{out} = \left(\frac{\nu}{f}\right)^2 \frac{S_\phi(f)|_a}{(2Q)^2}, \tag{2}$$

where $y \equiv \delta\nu/\nu$ is the fractional frequency deviation.

The use of a stabilized local oscillator (STALO) configuration allows a passive mixer to take the place of the rf amplifier or transistor used in simpler oscillator configurations. While somewhat more complicated due to the use of feedback at baseband (frequencies nominally zero) rather than at rf, the low noise of the mixer ($S_\phi(1\text{ Hz})|_a \approx -135dB$) at X–band equals that of active devices at lower frequencies and so essentially eliminates any performance disadvantage for high frequency operation[10–13].

Further improvement on the phase noise performance of sapphire oscillators, both at room and cryogenic temperatures, is made possible by use of novel circuitry. We have developed oscillator configurations which reduce the effect of semiconductor 1/f phase noise sources for stabilized local oscillator circuits. By appropriate use of carrier suppression, a small signal is generated which suffers no loss of loop phase information or signal–to–noise ratio. This small signal can be amplified without degradation by multiplicative amplifier noise, and can be detected without saturation of the detector. This circuitry promises a reduction in phase noise of 20 dB or more.

SPS Design using Suppressed Carrier

Figure 3 shows a version of the STALO – type design which can result in improved performance[3]. Here a small (nominally zero) signal returned from the resonator is amplified before it enters the mixer. Two effects of this addition are easy to understand. The loop gain will be increased by the added gain, an effect which may be compensated for in the design of the base–band amplifier. Secondly, the gain of the amplifier will increase the sensitivity of the mixer output to phase error in the resonator without significantly affecting mixer noise. Thus the effective mixer phase noise is reduced

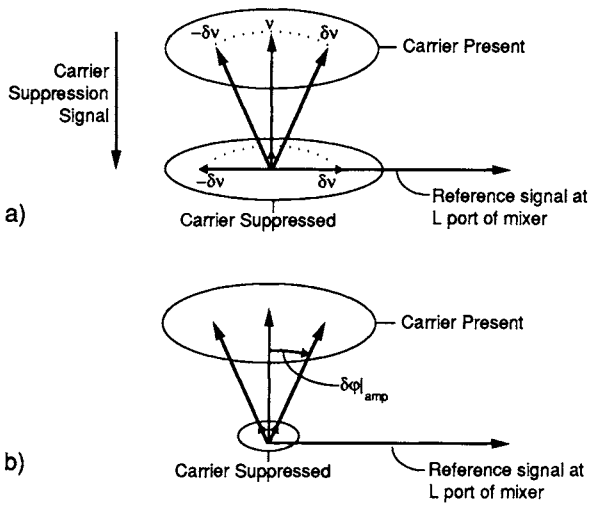


Figure 4: (a) Phasor diagram showing effect of LO frequency error on rf resonator signal with and without carrier suppression. Also shown are (constant) carrier suppression signal and mixer reference signal. Component of resonator signal in phase with reference is unchanged by 20 dB carrier suppression. (b) Phasor diagram showing effect of amplifier phase noise on resonator signal. Reduction of carrier by 20 dB results in similar reduction in signal error in phase with the mixer reference. (Horizontal scale for a) and b) both greatly enhanced.)

by the amount of amplifier gain. This can be a very substantial improvement.

The third effect of this modification is a little more complicated. Since flicker noise in X-band amplifiers is somewhat greater than for mixers (-120 dB vs -135 dB at 1 Hz offset), a crucial point is the proper analysis of the contribution of amplifier noise. The kind of noise under discussion is not additive noise, which would be independent of any large signal also present, but instead is multiplicative noise, which transforms a large signal by slightly modifying its amplitude and phase. (Additive noise is typically insignificant except at offset frequencies $f > \approx$ 10kHz.)

Figure 4(a) shows the cavity signals in phasor form, with and without carrier suppression. The added signal due to small frequency variations in the local oscillator (LO) is also shown. These added signals detected by the mixer allow feedback circuitry to cancel the frequency variations. The effect of multiplicative phase noise in the amplifier for the two cases is shown in Figure 4(b). It is clear from the figure that multiplicative noise generates signals that are indistinguishable from those caused by actual frequency variations, but which are generated only in the presence of a large carrier. Thus a reduction in amplifier noise is effected which is proportional to the degree of carrier suppression of the microwave signal.

Thus, oscillator phase noise is determined by a combination of mixer noise (reduced by amplifier gain) and amplifier noise (reduced by the degree of carrier suppression). For example, if mixer noise of $-135\text{dB}/f(/\text{Hz})$ were reduced by 25 dB of amplifier gain to $-160\text{dB}/f(/\text{Hz})$, and amplifier noise of $-120\text{dB}/f(/\text{Hz})$ were reduced by 40 dB of carrier suppression to the same value, the combined noise of $S_\phi(f)|_a = -157\text{dB}/f(/\text{Hz})$ would determine oscillator performance as given by Equation 2.

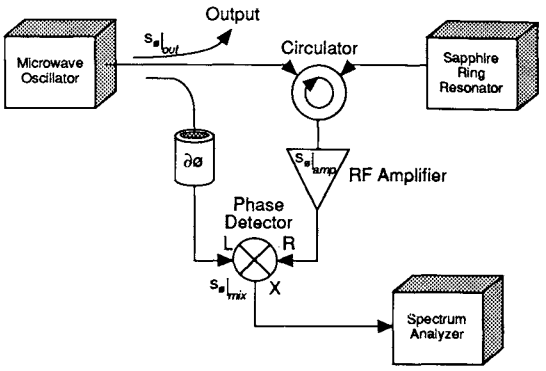


Figure 5: Measurement setup for open–loop test of sapphire phase stabilizer, configured as an ultra–low noise frequency discriminator.

EXPERIMENTAL ASPECTS

Design

The design of the experimental setup has been more completely reported elsewhere[2]. Significant elements are a) use of a small commercial Dewar for reliability and low cost, b) a stainless steel vacuum insert with copper bottom for resonator cooling, c) 80 K operating point held above the 77 K liquid nitrogen boiling point by a combination of rf and dc heating with detected rf feedback, d) reduction of thermal fluctuations by 400 second thermal time constant, and e) waveguide coupling ports which are below cutoff (non–propagating) at the 8.1 GHz operating frequency with coupling adjustment by means of easily machineable teflon (ϵ = 2) inserts that reduce the cutoff frequency.

We have verified all aspects of component performance including vacuum integrity, thermal time constants, mixer noise levels, and resonator frequency, Q, coupling constants and mode purity. The sapphire resonator element was purchased a little larger than the design value to allow tuning to frequency by grinding. The initial frequency is approximately 7.94 GHz at 80 Kelvin. All tests to date have been done at this frequency. After tuning to within 0.1% by grinding, a thin (0.1 cm) sapphire tuning disk will be installed above the resonator to bring the frequency within the required 0.001% for operation. Final tuning to the 8.1 GHz design frequency can then be accomplished by variation of the operating temperature by $\approx$ 1 Kelvin.

The waveguide coupling ports had approximately 2% of critical coupling before installation of their teflon inserts, and under this circumstance an unloaded resonator Q of 30 million was demonstrated. A first try at insert design created a somewhat overcoupled condition, with a loaded Q of 6 million for two ports. The phase–noise measurements described below were made under these conditions. A single–port configuration is probably more advantageous, allowing a critically coupled Q of 15 million. Subsequent tests will thus be done with only one teflon insert, the second port being only weakly coupled.

Phase Noise Measurements

Open–loop tests of the SPS as a frequency discriminator were made at both 80 Kelvin and 300 Kelvin using the circuit shown in Fig. 5 with the circulator replaced by a 3 dB hybrid coupler. The use of the hybrid gave somewhat lower noise at

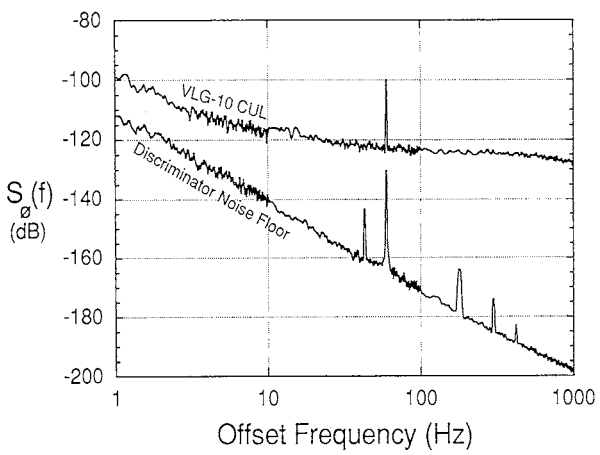


Figure 6: Phase noise characterization at 100MHz of an ultra–low noise quartz crystal oscillator using the frequency discriminator shown in Fig. 5 operating at 80 Kelvin with a loaded Q of 6 million. This represents the first time that this phase noise could be characterized without subtracting a similar noise contribution due to the reference oscillator.

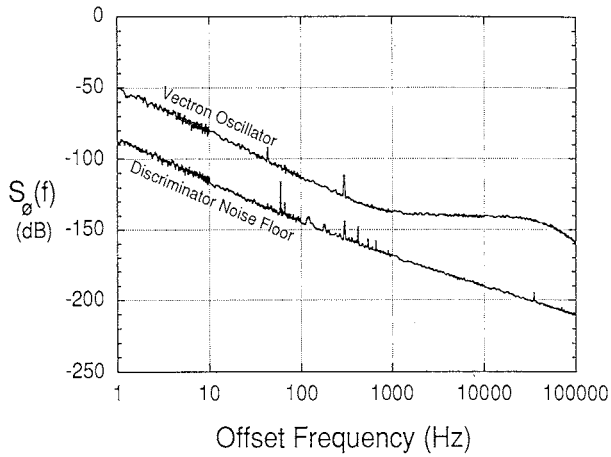


Figure 7: Phase noise characterization of a 100 MHz Vectron oscillator using a frequency discriminator shown in Fig. 5 operating at 300 Kelvin with a loaded Q of 150,000.

the expense of signal strength. These tests show (at 80 K) lower phase noise than the quietest available quartz crystal oscillator, in agreement with an expected mixer noise of -135 dB/f(/Hz) and the measured Q of 6 million.

Figure 6 shows a measurement of phase noise using the cryogenic discriminator. The oscillator under test is incorporated in a clean–up loop for the NASA VLG–10 hydrogen masers, which incorporates an ultra–low noise quartz crystal oscillator[15]. Previous tests of the VLG–10 CUL had required comparing two units, and dividing the resulting measured noise by 2 (-3dB) to compensate for the noise of the second oscillator. The SPS discriminator allowed the noise of a single unit to be characterized for the first time, and gave results in excellent agreement with previous tests of the oscillator.

These measurements were made in spite of the relatively low Q, poor mode purity, overcoupled ports, and unadjusted frequency as described above. They already show a noise floor 10 dB to 70 dB below the VLG–10 CUL. We expect an improvement of approximately 10 dB with the proper coupling, and much less ancillary equipment, not otherwise described here, when the frequency is properly adjusted.

While cryogenic results have yet to show an effective reduction of mixer noise due to carrier suppression techniques, room temperature tests gave a substantial improvement. Greater freedom of access allowed the circuitry to be more easily optimized, resulting in a 10 dB reduction in mixer noise. Figure 7 shows the results of these tests using a critically coupled resonator showing a Q of 150,000. We were able to tune the resonator to 7.800 GHz in this case by the addition of two sapphire tuning disks, and so eliminate a synthesizer which may have degraded the low temperature measurements. Coupling was precisely adjustable to the critically coupled condition, and so we were able to optimize performance better than in the cryogenic tests. The Vectron 100 MHz oscillator, with its low noise at high offset frequencies, gave a relatively small noise signal through the rf amplifier shown in Figure 5. In contrast, measurements with the cryogenic resonator were done under conditions where the noise through this amplifier severely taxed its dynamic range. This effect of uncompensated high–frequency noise may have also degraded the cryogenic measurements.

The room temperature noise floor corresponds to an active device noise of $S_\phi(1\text{ Hz})|_a \approx -145dB$ in Eq. (2) – a value 10 dB lower than expected for mixer noise, and so shows the efficacy of the carrier suppression circuitry.

Extrapolating the noise floor of the room–temperature measurements to the loaded Q of 15 million expected at 80 Kelvin indicates an improvement of almost 20 dB below the floor shown in Figure 6; this noise floor would be 30 dB and more below that of the ultra–quiet quartz oscillator incorporated in the VLG–10 clean up loop.

Mode Purity

In our measurements at 80 K we find that incidental material variations in the sapphire resonator give a mode splitting of approximately 10^{-6} in frequency for the two nominally degenerate $WGH_{10,0,0}$ whispering gallery modes. These modes show a misalignment in their angular orientation with respect to the coupling port that changes from run to run, giving rise to a variable coupling coefficient. Under optimal conditions, only the strongly coupled mode will show a reduced, loaded Q, while the other mode should show the intrinsic Q if it can be observed at all. In contrast, our observations showed the "uncoupled" mode had coupling typically reduced by only ≈50%.

To remove this dependence on details of the installation, a mode splitter, designed for 10^{-5} splitting and with variable angle control, has been installed in the SPS to align one of the modes with the coupling port; the other mode, rotated by 90 degrees, will have nominally zero coupling.

Adjusting the angular position of the mode splitter in cryogenic tests controlled mode alignment with the coupling port, while providing the desired frequency splitting. The adjustment gave critical coupling to the chosen mode and, importantly, did not degrade resonator Q. This should allow us to critically couple to the resonator for the design Q of 15 million at 80 K.

ANALYSIS

These results clearly establish the credibility of liquid nitrogen temperature sapphire technology. With open loop oper-

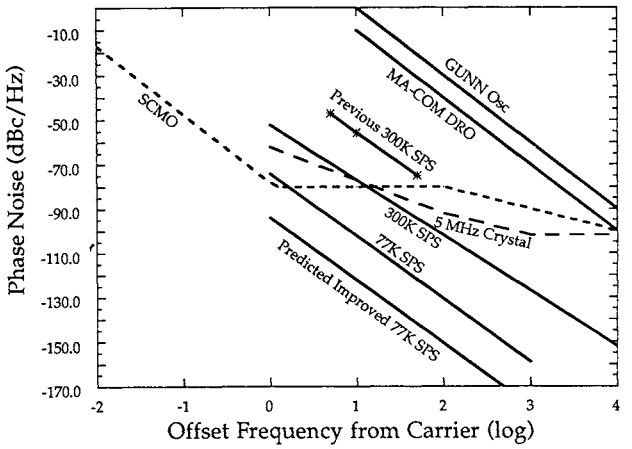


Figure 8: Phase noise, referred to 10 GHz, for various frequency sources including the results from Figs. 6 and 7 as well as our prediction for performance of the 77K SPS.

ation showing excellent noise floor performance, systematics involved in closed loop operation must now be addressed. Because these phase noise levels are below those obtainable any other way, one unit must act as a reference for a second, and phase comparisons will probably need to be done at X-band rather than at 100 MHz as was done so far.

Some problem areas are bypassed by these open loop measurements. For example, various sources of phase fluctuation which are below the VLG-10 noise level may still be above that of the noise floor. These include: a) Effects due to uncorrected noise in the VLG-10 at high offset frequencies; phase compensation by means of an X-band phase shifter would allow the effect of these high frequency fluctuations to be eliminated; b) AM noise in the signal applied to the SPS; c) AM noise in harmonics of the 8.1 GHz carrier; d) Circulator non-linearities and noise; and e) PM-AM conversion in the X-band phase shifter.

Figure 8 compares the data shown in Figs. 6 and 7 with phase noise data at X-band for various low noise sources, and with predicted performance for the 80 K SPS. Our previously reported results are also shown[2]. The cryogenic sources show significant advantages over the entire range of offset frequencies (0.01 Hz $\leq f \leq$ 10,000 Hz), with SPS performance taking up where superconducting cavity maser oscillator (SCMO) performance flattens out at $f \approx 1$ Hz[15].

It seems likely that a compromise between the expense and complication of the liquid helium-cooled SCMO and the relatively simple SPS may also be an attractive option. Operating at several milliwatts rf power, compared to ≈ 1 nanowatt for the SCMO, the SPS demonstrates much lower high-frequency noise. The SCMO, on the other hand, with a nominally zero temperature coefficient of frequency shows much greater stability at the lowest offset frequencies. Because thermal effects "freeze out" with an approximate T^3 dependence, an SPS cooled by a closed-cycle refrigerator to 10 Kelvin might be an attractive compromise, showing both low noise and higher stability.

ACKNOWLEDGEMENTS

The authors would like to express their gratitude to R. T. Wang for cryogenic and rf expertise, A. Kirk for assistance with phase noise tests, and T. K. Tucker for assistance with circuit design and component selection.

REFERENCES

[1] L. Maleki, "Frequency and Timing Requirements for Space Exploration," *Proc. 45th Ann. Symposium on Frequency Control,* 707-711 (1991).

[2] J. Dick, J. Saunders and T. Tucker, "Ultra-Low Noise Microwave Phase Stabilizer using Sapphire Ring Resonator," *Proc. 44th Ann. Symposium on Frequency Control,* 577-584 (1990).

[3] G. J. Dick and J. Saunders, "Method and Apparatus for Reducing Microwave Oscillator Output Noise", U.S. Patent #5,036,299, July 30, 1991.

[4] G. J. Dick and D. M. Strayer, "Measurements and Analysis of Cryogenic Sapphire Dielectric Resonators and DRO's," *Proc. 41st Ann. Symposium on Frequency Control,* 487-491, (1987).

[5] V. B. Braginsky, V.P. Mitrofanov and V.I. Panov, *Systems with Small Dissipation* (Univ. of Chicago Press, Chicago, 1985), 85-89.

[6] V. I. Panov and P. R. Stankov, "Frequency Stabilization of oscillators with high-Q leucosapphire dielectric resonators", *Radiotekhnika i Electronika* **31**, 213, (1986), (In Russian).

[7] D. G. Blair and I. N. Evans, "High-Q Microwave Properties of a Sapphire Ring Resonator," *J. Phys. D: Appl. Phys.,* **15**, 1651-1656, (1982).

[8] A. Giles, S. Jones and D. Blair, "A High Stability Microwave Oscillator Based on a Sapphire Loaded Superconducting Cavity", *Proc. 43rd Ann. Symposium on Frequency Control,* 89-93, (1989).

[9] X. H. Jiao, P. Guillon, and L. A. Bermudez, "Resonant frequencies of whispering-gallery dielectric resonator modes", *IEE Proceedings* **134**, Pt. H, 497 (1987).

[10] F. L. Walls, A. J. D. Clements, C. M. Felton, M. A. Lombardi, and M. D. Vanek, "Extending the Range and Accuracy of Phase Noise Measurements," *Proc. 42nd Ann. Symposium on Frequency Control,* 432-441 (1988).

[11] F. L. Walls and C. M. Felton, "High Spectral Purity X-band Source," *Proc. 44th Ann. Symposium on Frequency Control,* 542-548 (1990).

[12] T. E. Parker, "Characteristics and Sources of Phase Noise in Stable Oscillators", *Proc. 41st Ann. Symposium on Frequency Control,* 99-110, (1987).

[13] G. K. Montress, T. E. Parker, M. J. Loboda, and J. A. Greer, "Extremely Low Phase Noise SAW Resonators and Oscillator: Design and Performance," IEEE *Trans. Ultrasonics, Ferroelectrics, and Frequency Control* **UFFC-35**, #6, 657-667 (1988).

[14] Quartz oscillator model 8600.03 by Oscilloquartz SA, CH-2002 Neuchâtel 2, Switzerland.

[15] R. T. Wang and G. J. Dick, "Improved Performance of the Superconducting Cavity Maser at Short Measuring Time," in *Proc. 44th Ann. Symposium on Frequency Control,* 89-93 (1990).

THE APPLICATION OF THE PRARE PRINCIPLE FOR PRECISE TWO-WAY TIME TRANSFER TO SPACE-BASED ATOMIC CLOCKS

Wolfgang Schäfer

University of Stuttgart, Institute of Navigation, Senefelder Straße 26, 7000 Stuttgart 1, Germany

ABSTRACT

The PRARE (**P**recise **RA**nge and **R**ange-Rate **E**quipment) principle was originally intended as a universal tool for precise orbit determination of earth-orbiting spacecraft. Its primary purpose was to provide exact measurements of physical quantities required for this task, and to avoid as far as possible the use of orbit- and signal propagation models. Systematic errors have to a large extent been reduced by choosing a near-optimum topology.

This paper demonstrates that the very demanding requirements of time-transfer between an orbiting H-Maser Clock and ground could be met by the instrument in its basic configuration, provided that the carrier phase measurement resolution is increased and the on-board calibration provisions are specifically tailored for this task. The paper concludes with an estimation of the expected residual instrumentation error.

Keywords: Time-Transfer, Ranging, Range-Rate, H-Maser, Orbit Determination.

1. SYSTEM REQUIREMENTS

H-Masers in space promise improved performance compared to their earth-based counterparts. The use of these systems depends on the availability of suitable time-transfer systems linking space to ground. (**Fig. 1-1**).

Compared to existing two-way time-transfer systems, which are deemed amongst the most accurate for terrestrial use and where nearly arbitrary long integration intervals may be used, the H-Maser´s relative velocity to ground imposes more restrictions, due to the lack of a quasi static configuration. This paper does not cover relativistic effects. It is assumed, that these may be taken into account by the application of relativistic theory.

A suitable time-transfer system must provide sufficient accuracy to fully exploit the H-Maser´s stability with a noise floor of few times in 10^{-15}. The maser would accumulate a time error of 1 picosecond during one visibility period and about 10 picoseconds during one orbit.

In the case of low-orbit satellites, visibility from one ground station is restricted to several minutes and the signal travel time from space to ground varies several tens of milliseconds in this period. Furthermore, all signals are affected by high doppler shifts. The system must cope with these very high dynamic effects during initial signal acquisition and during the measurement period itself. Upon successful signal acquisition, it must provide useful results with resolutions down to the picosecond level within one visibility period. No ambiguities are allowed from one contact period to the next.

Dispersive effects of the propagation medium need have no real effect on the overall system performance. Therefore, ionospheric and tropospheric effects have to be taken into account.

2. SYSTEM LAYOUT

The satellite based instrument performs the two basic functions of two-way range and two-way carrier phase measurements on an X-band microwave channel and uses an additional S-band frequency for ionospheric correction.

Initially, the instrument was optimised for operation on board low earth-orbiting satellites. It is able to compensate for the satellite's high relative velocity versus ground. This instrument has the potential of about 60 psec for group delay measurements and 1.8 mm per sec for velocity measurements, each with an integration time of one second.

The PN-coded microwave signals are transmitted to coherent, regenerative ground transponders. Up to four return signals may be processed in parallel by the on-board receivers. Round-trip time delay and carrier phase is reconstructed from the received signals. The space-equipment is fully coherent both in the transmitter and in the receiver chains. Data pre-processing and data collection is done on board as well. (**Figs. 2-1 and 2-2**)

The ground station itself calculates the difference of time-of arrival measurements between the two microwave channels as a direct means to determine Total Electron Content in the signal path. This data is uplinked together with meteorological data gathered at

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

each side. This is stored on-board along with range- and range-rate data until interrogation by the master ground station. This scheme provides a world-wide data-collection function with minimum delay until availability of science data. Apart from on-board storage, all data is re-transmitted in real time back to the ground stations in sight.

System data and ephemeris data are transmitted regularly to the ground stations, which are designed to operate unattended over prolonged periods of time.

The basic two-way measurement is performed in X-band, whereas the S-band down link is used for ionospheric correction only.

Two-way range data is acquired at the space-craft at a rate of 100 measurements per second with an individual resolution of 100 psec. Digital post-processing is used to increase this resolution to 10 psec for an integration time of one second. Range jitter is limited by thermal noise to about 2-3 cm for this integration interval. Range data is recorded once per second.

Two-way carrier phase measurement is performed once per second with a resolution of 0.1 wavelength (i.e. 3.8 mm). A continuous counter, which is not stopped during read-out, is used to count phase cycles. Thus arbitrary long integration above one second intervals may be used for data analysis. It should be stressed that phase resolution is limited by the instruments present design and not by thermal noise, leaving room for an improvement by a factor of up to 1000.

The instrument conducts all measurements automatically. Ground signal acquisition including doppler steering on the received carrier frequency and the received PN-modulation is based on a real-time orbit calculation performed by the instrument's micro processors. The locations of 30 ground stations are stored within the instrument. An automated self calibration routine follows each ground station contact to eliminate instrument internal drifts.

The ground stations are of compact design and easily transportable. The antenna diameter is 60 cm and the weight of the electronics box is about 60 kg. They may be easily installed at sites of scientific interest. About 25 sites have been surveyed and prepared to accommodate the stations.

3. PERFORMANCE

The instrument's present performance can be summarised as follows:

Noise values:

± 1.7 cm	for X-band ranging (1 sec interval)
± 30 cm	for S-band ranging (1 sec interval)
± 0.1 mm/s	for X-band doppler (30 sec interval)

Bias values:

< 1 cm	for X-band
< 3 cm	for S-band

Considerable post-processing will be applied to all data and the residual contributions are as listed below:

Tropospheric error (vertical propagation path):

$$\delta R_T = 2 - 7 \text{ cm}$$

Ionospheric error (vertical propagation path):

$$\delta R_I < 1 \text{ cm}$$

Thermal noise and calibration errors:

$$\delta R_C = 2 - 3 \text{ cm}$$

Antenna phase centre uncertainty:

$$\delta R_A = 1 \text{ cm}$$

Root mean square:

$$3 \text{ cm} < \delta R < 8 \text{ cm}$$

In-orbit performance is monitored regularly by closing the internal delay calibration loop. This loop covers all parts of the transmitter and receiver channels including all signal filters with the exception of the instrument's antenna and antenna cabling. This part contributes about 4 nsec and is assumed to be stable throughout the mission. The calibration loop re-injects a portion of the transmitted signal into the receiver at a signal level and at a signal-to-noise density ratio equivalent to a ground station at a distance of 1500 km.

A careful calibration of thermal effects is performed prior to launch and there are about 30 temperature monitors, whose readings are contained in the house-keeping data.

The uncorrected stability of the instrument has been measured to ensure that instabilities remain within tolerable levels. Total uncorrected warm-up drift after one hour of operation was one nanosecond and the uncalibrated internal delay drift at temperatures from - 10 °C to + 50 °C was about two nanoseconds. The temperature drift of the internal calibration loop is better than 20 psec / K.

4. PLANNED IMPROVEMENTS

The present ranging capabilities are not suitable in the present configuration, even after post-processing and data correction.

However, a careful study of limiting effects shows, that there are several areas of improvement. The basic theory is to determine as far as possible, the instrument's internal delays, the phase of the transmitted carrier and the error contribution of the receivers separately by suitable on-board calibration. All signal paths shall be covered to the utmost extent.

Potential error sources are analysed in the following paragraphs and an on-board calibration set-up is presented for each case and suitable means shall be integrated into future designs.

4.1 Transmitted PN Code-Phase

Range measurement in the present design is performed by determining the phase of the received PN-code versus the transmitted PN-code (**Fig. 4.1-1**). For internal calibration, a portion of the transmitter signal is translated to the receiver input frequency in a calibration transponder located at the antenna connector. In this way the sum of signal delays within the transmitter and the receiver chains is covered (**Fig. 4.2-1**).

4.2 Signal Delay in Receiver Channel

The individual contributions of receiver and transmitter can be obtained by simple modification of the above design. To measure the receiver contribution, the PN modulation is removed from the transmitter and a clean carrier appears at the antenna terminal. This signal is translated to the receiver frequency and subsequently phase modulated directly at the receiver input port. Hence, the receiver delay can be determined with this set-up (**Fig. 4.2-1**).

4.3 Transmitted Carrier Phase

The most critical value is the carrier phase of the transmitted microwave signal, which shall be locked to the on-board clock. In the present design, the frequency generator alone is controlled by phase locked loop. An improved scheme covers all amplifier and filtering stages as well. The second harmonic of the transmitted PSK signal which is generated by a squaring device connected to the output of the transmitter filter - a clean carrier - is fed back to a sampling mixer which is directly driven by the on-board clock. This scheme is outlined in **Fig. 4.3-1**. Phase noise and drift even in the power amplifier and output filter are cancelled.

4.4 Carrier Phase in Receiver Channel

Once the transmitted carrier phase is known, the receiver phase shift is measured with the normal internal calibration transponder. The reconstituted phase is directly compared to the on-board clock. The set-up is shown in **Fig. 4.4-1**. All the above measures would require minor hardware changes only.

A simple example shall demonstrate, that the present PRARE design actually performs two-way carrier phase measurements:

Given that the space-based clock travels from the left position to the position furthest right (**Fig. 4.4-2**). As long as the result is independent of the actual path (S_i) on which the clock travels (relativistic effects are excluded here) and if it can be demonstrated, that the result is independent from the actual velocity (V_i) along each path (P_i), we are dealing with a true two-way carrier phase measurement. For example, it can be easily shown, that PRARE-results remain valid if the sign of the velocity changes along the path from the starting point to the final location.

A minor hardware re-design effort would be necessary to increase phase resolution by a factor of eight. Furthermore it is planned change the RF carrier frequencies from X-band to Ku-band. Together, this would increase the range-rate performance by a factor of about 15 and the accumulated uncertainty during a nominal visibility time of 500 sec can be determined as follows:

$$\text{Error} = \frac{\text{Phase Resolution}}{\text{Visibility Time}} = \frac{0.007 \text{ cycle}}{500 \text{ sec}} = 1 \cdot 10^{-15}$$

The instrument would profit largely from a good temperature control of the environment, which is rather easy to achieve in space.

5. CONCLUSIONS

The implementation of group delay monitoring and phase stabilisation separately for the transmitter and the receiver channels is feasible and requires minor hardware changes only. The stability of the instrument could reach $1 \cdot 10^{-15}$ in near future, which would suffice to monitor a space-based H-Maser within one visibility period from a single ground station.

But system calibration at these levels of accuracy tends to be extremely difficult and external errors like multi-path or cross-correlation effects must be kept to the absolute minimum by careful system layout. System dynamics due to relative motion of a space-based clock are causing large signal doppler shifts, amplitude variations, high doppler rates on the carrier signal and on the PN-code modulation. These effects cannot be simulated on ground in all combinations and further research will be necessary to fully understand their error contribution.

6. ACKNOWLEDGEMENTS

This research work has been funded by the German Ministry for Research and Technology under grant 01 QS 154 being jointly conducted by the Institute of Navigation, University of Stuttgart and the "Deutsches Geodätisches Forschungsinstitut", München.

REFERENCES

/1/ Ph. Hartl, W. Schäfer; "Application of Spread Spectrum Techniques in Combined Use for Navigation and Telecommunication in Satellite Systems"; FREQUENZ No. 41 & 42, 1987/88.

/2/ W. Lechner, H. Wilmes, "The Precise Range and Range-Rate Equipment PRARE on a GPS-Satellite - Simulations about the Capability for Orbit Determination and Positioning"; 5th International Geodetic Symposium on Satellite Positioning; March 13-17, 1989, Las Cruches, New Mexico, U.S.A.

/3/ W. Schäfer, H. Wilmes, "Precise Range and RangeRate Equipment (PRARE) - System Design and Status of Development -"; Workshop on Advances in Satellite Radio Tracking, May 5-6, 1986, Austin, Texas, U.S.A.

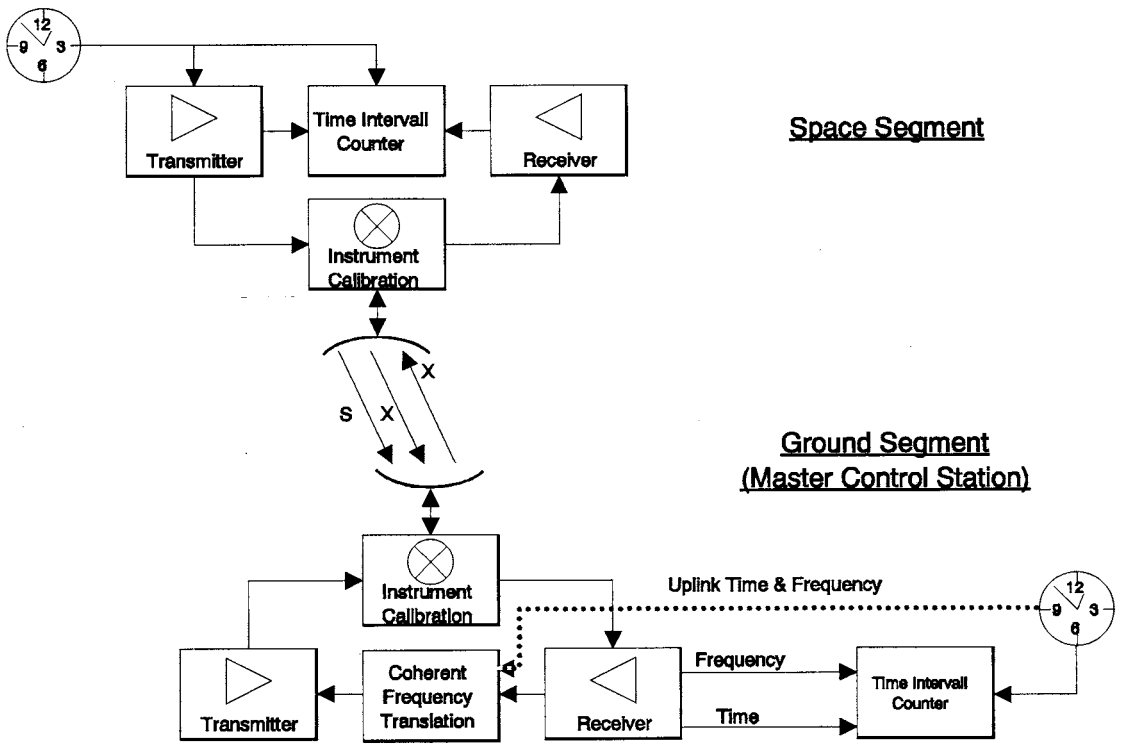


Figure 1-1: Two-way time transfer between ground and a space-based clock. Calibration loops are incorporated on both sides. The round trip time delay and time of arrival at ground is measured simultaneously using a coherent, regenerative transponder on ground.

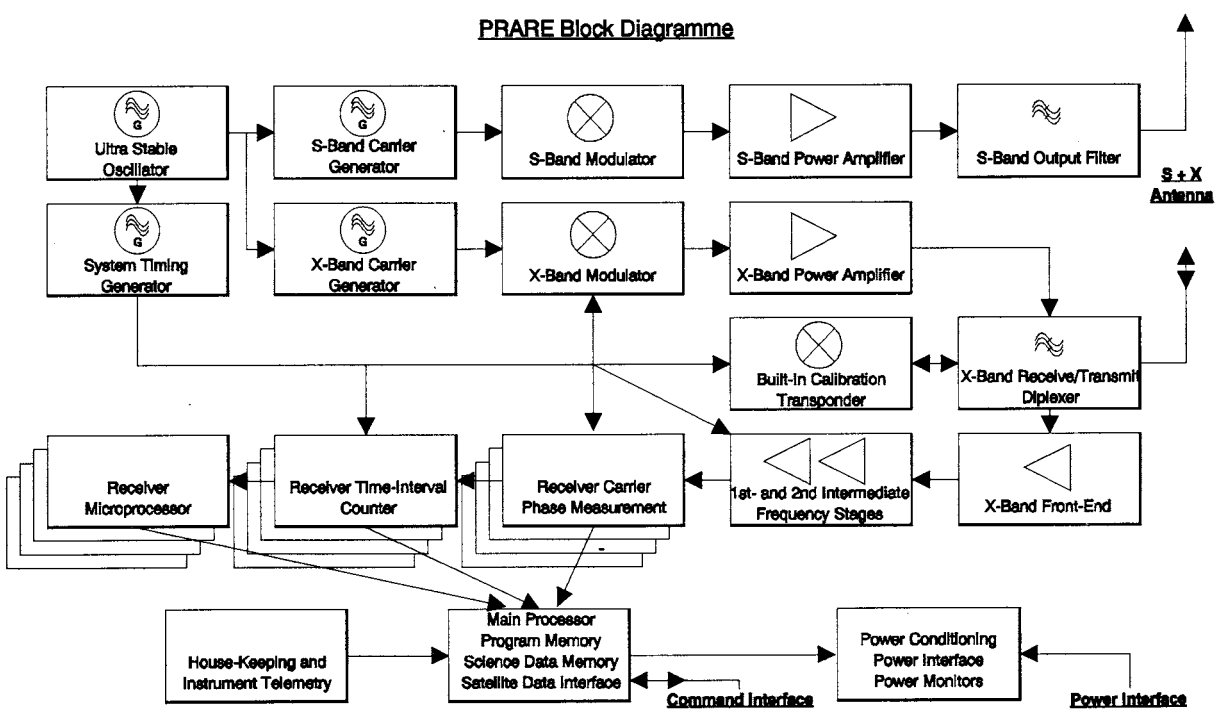


Figure 2-1: Present instrument design showing signal-paths as well as on-board processing, data collection and data storage.

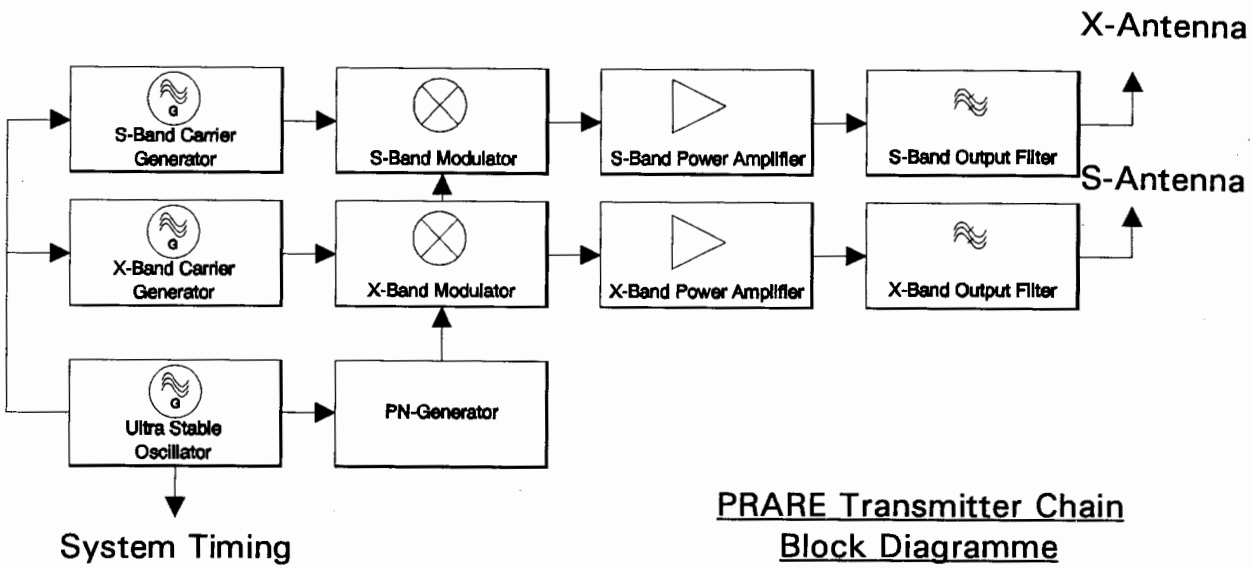


Figure 2-2: Block diagram of the two transmitter sections, showing full coherency between signals and carriers with respect to on-board clock.

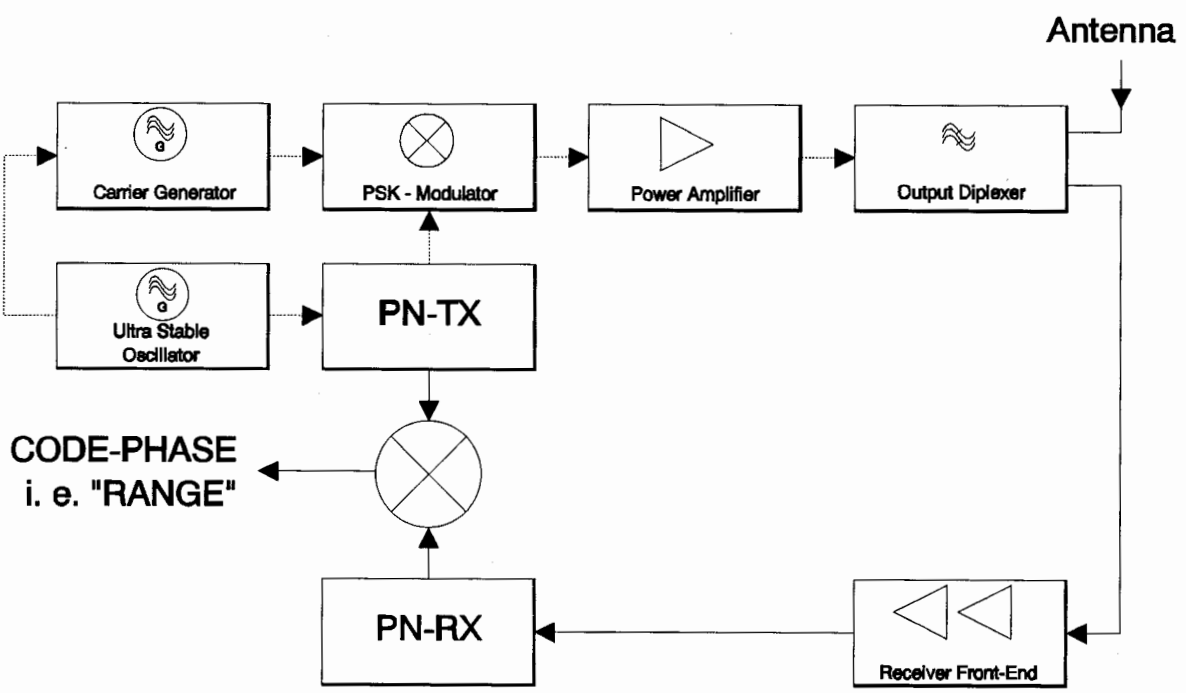


Figure 4.1-1: Basic scheme of range measurement which is actually performed as a phase measurement between transmitted code and reconstituted code in the receiver channel.

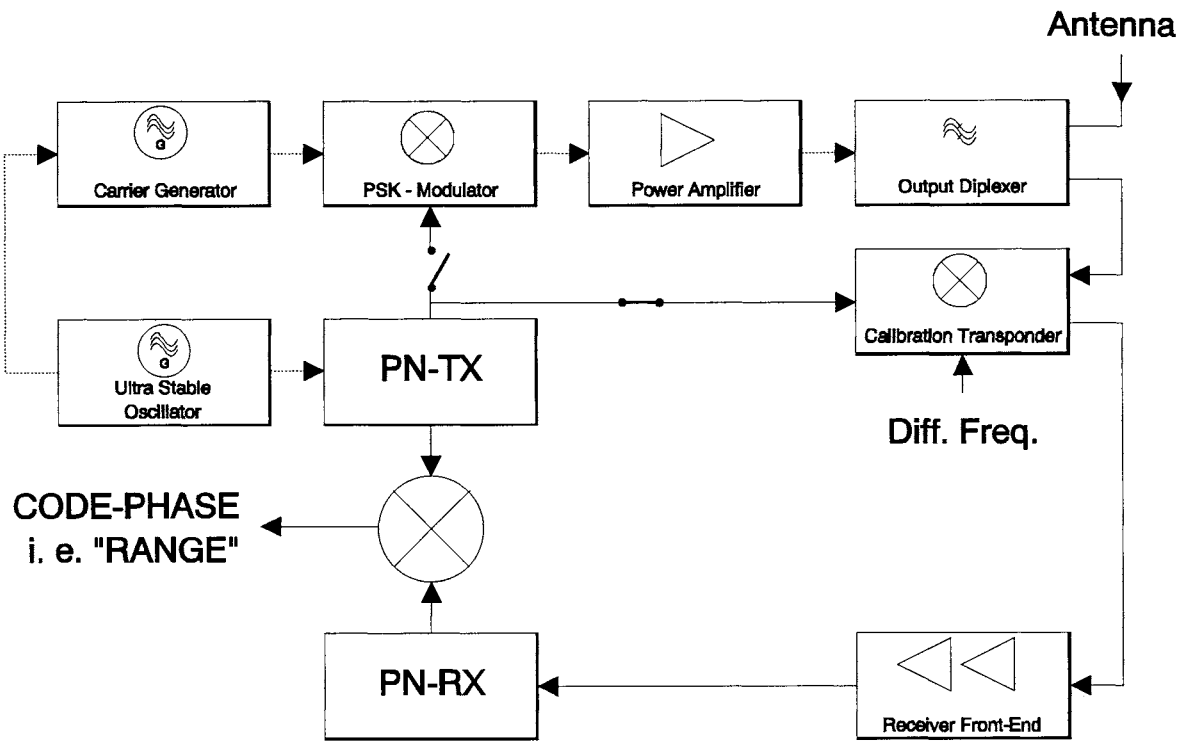


Figure 4.2-1: Block diagram of delay calibration. Total instrument delay is determined with an built-in calibration transponder. For calibration of the receiver section only, a PN modulated signal referenced to the on-board clock is generated at the front-end. PN modulation is removed from the transmitter during receiver calibration.

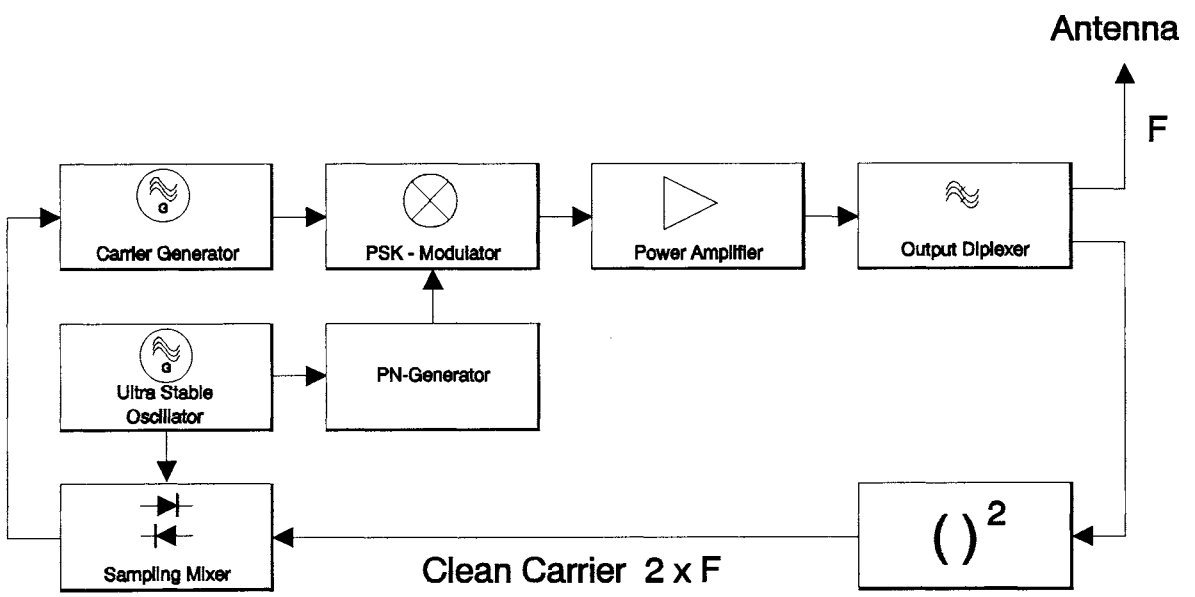


Figure 4.3-1: Proposed transmitter phase control loop referencing carrier phase at the instrument output directly to on-board clock.

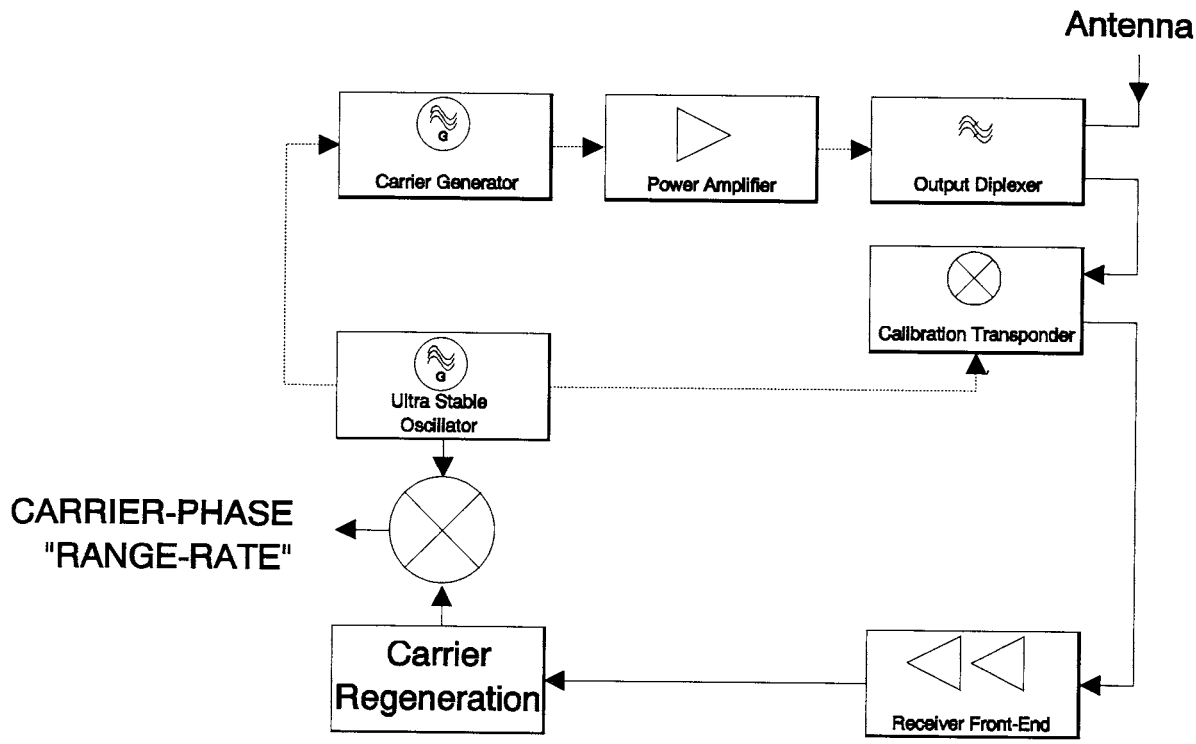


Figure 4.4-1: Block diagram for carrier phase measurement. Phase of a signal received from a ground station is compared to the phase of a signal coherently generated by the calibration transponder.

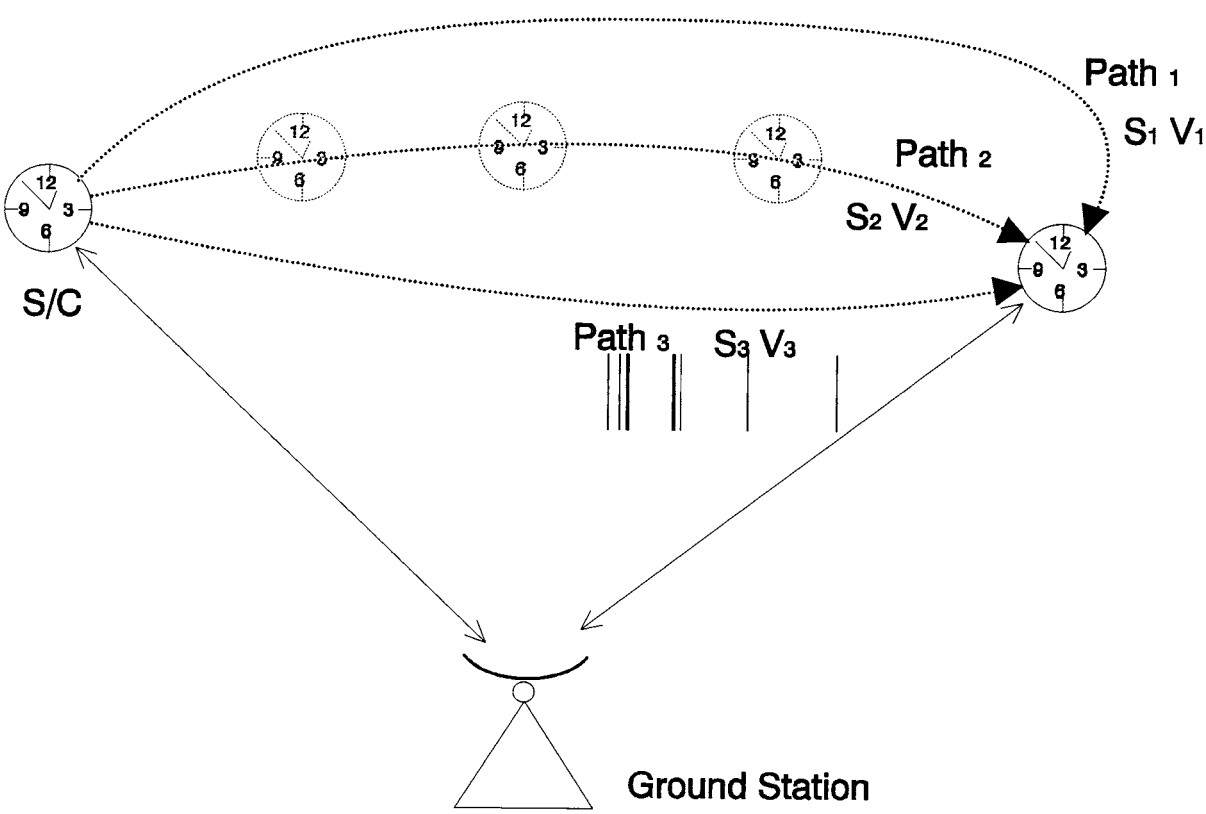


Figure 4.4-2: Scheme to illustrate necessary conditions for two-way carrier phase measurement.

Session 3A

Time and Frequency: Fundamentals

Chairman: R. Besson

ATOMIC TIME SCALES FOR DYNAMICAL ASTRONOMY

B. Guinot

Bureau International des Poids et Mesures, Sèvres

ABSTRACT

In 1991, the International Astronomical Union (IAU) issued a series of recommendations on space-time references, in the framework of the General Theory of Relativity. The paper presents these new recommendations, their background, the difficulties that were met, the connection with previous recommendations, with an emphasis on time. For simplicity of theoretical developments in celestial mechanics, several coordinate systems must be defined. This implies the definition of new coordinate times, these are presented together with their relation between themselves and to other times. International Atomic Time (TAI) is not affected by these recommendations, but appears now as a realization of the ideal Terrestrial Time having a clear scientific definition.

Keywords: Time scales, reference systems, relativity

1. INTRODUCTION

In solar system dynamics we deal with a small number of quasi-point masses. Dissipation by rotation and tidal deformation is very weak and can be modelled well or even can be neglected over periods of several centuries. On the other hand, angular measurements spread over decades or centuries, and more recently 'distance' measurements, by time of flight of electromagnetic signals, bring excellent relative accuracy. For example, the product GM of the gravitation constant, G, and the mass of the Sun, M, is now known in SI units with a relative uncertainty of 10^{-10}.

We thus have a physical system to which dynamical theories can be applied with great precision, both in applications, especially to space navigation, and in the testing of theories. However, to take full advantage of the observational accuracy, the time scales to which the observations are referred should have matching qualities. The definition and realization of such time scales is a matter of continuing concern to astronomers and remains a controversial topic. There are two main approaches possible:

(a) The time scale can be 'external' to the orbital motions and based on postulates, for example on the periodicities of rotations or on the constancy of atomic frequencies (Ref. 1).

(b) The time scale can be derived from a selected orbital motion, which implies the expectation that all orbital motions can be expressed as a function of a single chosen motion.

Method (a) was applied to the rotation of the Earth, providing Universal Time UT1 (present designation). Then, faced with inconsistencies of observation and theory generated by the irregularity of the Earth's rotation, the astronomers shifted to method (b), the reference orbital motion being that of the Earth so providing Ephemeris Time TE. The gestation of TE took the whole first half of the 20th century. This long period was mainly devoted to scientific investigations but, in the background, appeared a logical problem perfectly summarized by Danjon in 1929 (Ref. 2). I quote (English translation):

> "Having to choose between the constancy of the Earth's rotation and the validity of Newton's law, the astronomer cannot hesitate, he sacrifices the first one. (...) Newton's law is saved, but is subjected to a singular adventure: it is henceforth called to provide a measure of time, it becomes partially non-verifiable and ceases to be, properly speaking, a law. (...) Let us only hope that, some day, we shall discover a good terrestrial time standard and let us then abandon these difficulties of pure logics..."

The "good terrestrial time standard" has existed since 1955, in the form of atomic time which marks a return to method (a)... However, we find that its use in astronomy is not so simple as might have been expected. Nevertheless, a resolution adopted by the International Astronomical Union (UAI) in 1991 seems to be an important step forward. Before summarizing this resolution it is helpful to consider its background.

2. FUNDAMENTAL PROBLEMS

2.1. Dynamical time and atomic time

The theory presently applied to solar system dynamics is the General Theory of Relativity (GTR). But the conceptual definition of dynamical time does not depend on the choice of the theory: it was expressed in almost the same terms by Poincaré in 1905 (Ref. 3), for Newtonian mechanics, and by

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

Misner et al. in 1973 (Ref. 4), in GTR: "Le temps doit être défini de telle façon que les équations de la mécanique soient aussi simples que possible" (Poincaré), "Time is defined so that motion looks simple" (Misner et al., p. 23). In GRT, the assumed measurable quantity is proper time, but we must stress that GTR is a dynamical theory and that we cannot be certain *a priori* that an atomic clock, based on quantum physics, measures proper time, although, so far, no discrepancy has been found.

2.2. Uncertainties of realization of atomic time

The experimental data of an astrometric observation that is to be modelled by a theory is usually a position at some date read in a realized time scale such as International Atomic Time, TAI. The total uncertainty which imposes a limit to the development of applications and to tests of the theory has two components.

o One is the uncertainty ϵ in some variable quantity at the dated instant, e. g. in an angle or a distance. If v denotes the rate of variation of this quantity as a function of time, we can define a corresponding time uncertainty by

$$\delta t_{obs} = \epsilon / v. \qquad (1)$$

o The other one is the deviation δt between the time scale and the ideal time that it represents.

To be satisfactory, the time scale should be such that the contribution of δt is negligible:

$$\delta t << \delta t_{obs}. \qquad (2)$$

But the properties of these two uncertainties are quite different. Uncertainty δt_{obs} has an upper limit which progressively decreases with technical advance. In contrast δt increases indefinitely because it takes the form of an integral, since the time origin, of an elementary time interval error. Thus an instant necessarily occurs when (2) is no longer satisfied. When is this? The best response takes the form of examples based on TAI.

(a) Orbital motions

Among the most important contributions to orbit modelling are the distance measurements to Mars in the Viking experiment. The measurement uncertainty ϵ was of order 10 m, but a realistic estimate of the uncertainty in the distance between centres of mass of Mars and the Earth is, according to Bretagnon (Ref. 5), of the order a few hundreds of metres. Let us take ϵ = 200 m. Using as v the maximum velocity of Mars with respect to the Earth, this corresponds to δt_{obs} = 12 ms. This is not very demanding. At the present performance level of TAI, δt would not reach this value in less than several millenia. But if we could have ϵ = 0,1 m, which is not unconcievable in manned exploration of Mars, then δt_{obs} would reduce to 6 μs, which may be attained by TAI in two decades.

(b) Earth's rotation

With Very Long Baseline Interferometry (VLBI), the rotation angle of the Earth is measured with uncertainty ϵ = 0,0007". As v = 15"/s, δt_{obs} = 50 μs. The uncertainty in TAI is significant after less than a century. However the theory of the geophysical perturbations of the Earth's rotation in the long term cannot match the measurement accuracy: TAI is, therefore, practically an ideal reference for this purpose.

(c) Millisecond pulsars

The TAI date of reception of a given pulse of pulsar PRS1937+21 can be read with an uncertainty of about 0,4 μs after a daily averaging which forms normal points (Ref. 6 and 7).

For some given site of observation, let us define t_R as a non-uniform time scale constructed as follows. The reading of t_R at reception of pulse number n (from an arbitrary origin, pulse number 0) is nP_e, where P_e is a fixed value of the period in seconds, chosen conventionally, but close to the true period of the pulsar. The slowly variable quantity

$$\Theta = TAI - t_R \qquad (3)$$

is submitted to analysis. The aim is to model all the effects contributing to Θ so that, ideally, the post-fit residuals of normal points is of order 0,4 μs. Unfortunately, the contribution of TAI exceeds 0,4 μs in less than one year.

2.3. Is atomic time a quasi-ideal time for astronomy ?

Atomic time is certainly not a quasi-ideal time for astronomy, as shown by the example of millisecond pulsars. But for the study of orbital motions in the solar system, it seems to be a quasi-ideal reference over many centuries. Nevertheless, we must pay attention to the accuracy of astrometric measurements and to a possible fundamental difference between quantum time and dynamical time.

Thus the geocentric ephemerides could have been expressed with TAI as time argument, with the practical interest that TAI is an available time scale. This has been proposed, by Guinot (Ref. 8) and repeatedly by Winkler (Ref. 9). However this direct use of TAI has been rejected for at least three reasons.

- One is practical: there exists an offset of $32,2^s$ with respect to Ephemeris Time, the previous time argument of astronomical ephemerides.

- One is theoretical: it is necessary to define another time argument for ephemerides with space origin at the barycenter of the solar system.

- One is conceptual: it is desirable to express ephemerides in terms of the theoretical time argument in which they are developed and to separate this from the problem of the realization of this time argument using atomic clocks.

These reasons are quite good. But they lead to a complex situation with many 'Times' and so to the question: Why so many time scales? (Ref. 10).

3. A FIRST ATTEMPT TO DEFINE NEW DYNAMICAL TIMES IN CONJONCTION WITH ATOMIC TIME (1976, 1979)

In the 70's, although the General Theory of Relativity (GTR) had been established for about 50 years, it was not yet accepted officially as a working theory for astronomical ephemerides. However, the consequences of a treatment of time scales for dynamics in GTR were sufficiently important, although small, to justify the definition by the International Astronomical Union (UAI), in 1976, of two 'coordinate times' for geocentric and barycentric reference frames. The expression 'coordinate time' was carefully avoided in the designation of these 'times', as being too specific to a particular theory, the GTR. In 1979

these times were given the names Terrestrial Dynamical Time (TDT) and Barycentric Dynamical Time (TDB). Even worse, the scale units of TDT and TDB were biased with respect to the theory by the use of 'scale factors', in order to avoid a secular divergence from the readings of an ideal caesium clock located at a fixed point on the geoid.

Soon afterwards, the definition of TDT and TDB raised misunderstandings and controversies (Ref. 11). Clearly a re-examination of the problem of time scales was desirable. In 1988 the IAU took the opportunity of the creation of a Working Group on Reference Systems (WGRS) to include 'Time' as one of the topics to be considered by this group.

4.WORKING GROUP ON REFERENCE SYSTEMS (1988-1991)

As is the case for other scientists, astronomers have to agree on quantities, on units and on conventional values for constants. They have, in addition, specific problems: the realization of directions of reference and of time scales which can be modelled in dynamical theories with high accuracy and in the very long term.

There are two conflicting requirements: to update reference frames and constants in order to keep pace with the progress of observational techniques; to avoid the inconvenience of frequent changes. During the last decade, the development of VLBI and the emergence of space techniques (HIPPARCOS, HUBBLE Telescope) reduced the uncertainties of angular measurements by a factor of order 1/100. At the same time, radar and laser telemetry progressed. The previous IAU resolution on Astronomical Constants, Time Scales, and the Fundamental Reference Frame was issued in 1976. It is no longer adequate for accurate observations and in several programmes different constants and standards have been used (for example, *IERS Standards*, Ref. 12). A revision was becoming urgent: it was the task of the WGRS.

The WGRS was chaired by the late and much regretted J. Hughes. It was divided in four Sub-group, on Reference Frames (leader, J. Kovalevsky), Time (B. Guinot), Astronomical Constants (T. Fukushima) and Nutation (D.D. McCarthy). After intensive work by correspondence, the WGRS established draft recommendations to the IAU at a meeting held in Virginia Beach (USA) in October 1990 (Ref. 13). These recommendations, after minor changes, were officially adopted as an IAU resolution, consisting of nine recommendations, at the IAU General Assembly, Buenos Aires, July 1991. This resolution, with its extensive explanatory notes, forms an important document and is published in both the IAU Information Bulletin (Ref. 14) and the IAU Proceedings.

The following section deals with the background to this resolution and presents its recommendations concerning 'time'. In section 6. some brief comments are given on the other recommendations on space references and astronomical constants.

5. IAU RESOLUTION ON REFERENCE SYSTEMS, RECOMMENDATIONS I TO V (1991)

At an early stage in the discussion, the leaders of the Sub-groups on Reference Frames and Time rapidly reached the conclusion that they should establish a firm theoretical basis, in the framework of GTR, for a global and homogeneous treatment of space and time reference through their respective Sub-groups. They sought the advice of relativity experts: among these should be mentioned V. Brumberg (Russsia), T. Damour (France), T. Fukushima (Japan) and H. Soffel (Germany), to all of whom much thanks are due for their fruitful cooperation.

5.1.General recommendations on space time coordinate systems

Recommendation I, initially called 'Conceptual recommendation', explicitly introduces the GTR as the theoretical background for the definition of celestial coordinate systems.

In GTR, coordinates are not physically measurable quantities and can be selected in an arbitrary manner, but a practical requirement is simplicity of the equations. This simplicity cannot be achieved with a single coordinate system covering all space time. It is evident, for example, that the theory of the motion of an articial satellite of the Earth is more complicated in a heliocentric frame than in a geocentric frame, because, in the latter, the action of all the bodies, with the exception of the Earth, is quite small (tidal potential).

The success of Newtonian mechanics suggests the adoption relativistic coordinates which are not too different from absolute time and from the Cartesian coordinates of Euclidian space. This led to metric forms, linking coordinates to measurable quantities, known as Newtonian approximations. These forms appear as power series in $1/c$, c being the velocity of light. All reasonable metrics agree on the first terms, but may differ at higher order depending on the coordinate conditions adopted. The IAU has chosen a metric form in which the low order terms are common to those of more sophisticated models. This is sufficient for present purposes at the current level of accuracy of the observations (with the exception of the modelling of geodesic precession). Later, when desirable and possible, other terms may be added without changing other recommendations.

Thus, the IAU recommends that the four space-time coordinates ($x^0 = ct$, t being coordinate time, x^1, x^2, x^3) be selected in such a way that in each coordinate system centred at the barycentre of any ensemble of masses the squared interval ds^2 be expressed with the minimum degree of approximation in the form:

$$ds^2 = -c\,d\tau^2 = -(1 - 2U/c^2)\,(dx^0)^2 + (1 + 2U/c^2)\,[(dx^1)^2 + (dx^2)^2 + (dx^3)^2], \quad (4)$$

where τ is proper time (measurable with a local clock), and U is the sum of the gravitational potentials (positive sign convention) of the above mentioned ensemble of masses, and of a tidal potential generated by bodies external to the ensemble , the latter vanishing at the barycentre.

The metric form (4) defines theoretically the coordinate systems, makes possible the transformations between observables and coordinates and allows coordinate transformations between different coordinate systems (Ref. 15).

Then follows a 'Constraint recommendation' (no II) specifying how to realize the conceptual coordinate systems of Recommendation I. For space, the realization will be based on direction of extragalactic sources, for time, on atomic clocks on the Earth. That means we may assume that atomic clocks give a good realization of proper time. It is stated that the units for proper time and proper length are the SI units.

5.2. Recommendations on coordinate times

It still remains to fix the origin of the coordinate times. This is done by Recommendation III, which states that the readings of all the coordinate times be 1977 January 1, $0^h\ 0^m\ 32{,}184^s$ on 1977 January 1, $0^h\ 0^m\ 0^s$ TAI, at the geocentre (the specification of a location is necessary, because the relation between coordinate times is fully four-dimensional). This recommendation also gives the names Geocentric Coordinate Time (TCG) and Barycentric Coordinate Time (TCB) to the most important coordinate times centred at the centres of mass of the Earth and of the solar system.

To a sufficient approximation, at present, the link between τ and the coordinate time t of some coordinate system is obtained by integration of

$$dt = (1 + U/c^2 + v^2/2c^2)\ d\tau, \qquad (5)$$

derived from (4) neglecting terms in c^{-4}, where v is the velocity of the clock in the considered system. Equation (5) shows that the scale unit of t is always shorter than one SI second. At fixed point on the geoid the differences are 7×10^{-10} s for TCG and $1{,}6 \times 10^{-8}$ s, on average, for TCB, leading to divergences of, respectively, 22 ms/year and 0,5 s/year.

These divergences were a major source of dissent in the work of the WGRS. The main objection to the introduction of TCG and TCB was that important ephemerides, established by the Jet Propulsion Laboratory (USA) and widely distributed, have been developed and issued with the time argument TDB defined in 1976/79. The definitions of TDB and TCB are indeed the same, except for a scaling factor, and the use of one or the other of these two times is theoretically equivalent. However three reasons favored TCB and TCG and led a large majority to vote their adoption.

- The removal of the secular part of TDB - TDT depends on the adopted procedure, in particular on the averaging time.
- The use of scaling factors in TDT and TDB has the consequence that some astronomical constants and quantities take different values, depending on the coordinate system which is adopted.
- It is always preferable to stick to the theory, without introducing arbitrary conventions (experience has shown that such conventions are often regretted later).

On account of the important work already performed using TDB, Recommendation V recognizes that TDB may be used where discontinuity with previous work is deemed to be undesirable.

5.3. Recommendation on Terrestrial Time

International Atomic Time, TAI, is not a realization of TCG because the scale unit of TAI should have, by definition, a duration of one SI second on the rotating geoid. It would be a purely academic exercise to evaluate the pros and cons of setting the scale unit of TAI on that of TCG: this was not done and now it is too late to change. On the contrary, the IAU defined, in Recommendation IV, another form of the ideal TCG, called Terrestrial Time (TT), which differs only by the scale unit, so that the scale unit of TT is the SI second on the rotating geoid. Thus TT appears as the ideal form of TAI. Recommendation IV provides a scientific definition of TAI, in a wider framework than the definition by the Comité International des Poids et Mesures (1970) completed by a declaration of the Comité Consultatif pour la Définition de la Seconde (1980) (Ref. 16, p. 110, and 17).

Heated discussions took place on the subject of the origin of TT.

As noted previously, there is a time offset between Ephemeris Time and TAI: approximately,

$$TE - TAI = 32{,}2^s. \qquad (6)$$

This also is a circumstance one might deplore. While the atomic definition of the SI second was established so that its duration would coincide with the best estimate of the ephemeris second, the origins of atomic times (including the early form of TAI) established by several laboratories were conventionally set, with respect to Universal Time UT1, by the condition that an event on 1958 January 1, 0^h UT1 receive this same date in atomic time. This inconsistency is explained by the poor accuracy with which Ephemeris Time may be read, especially in near real time, the uncertainy being of order of $\pm\ 1^s$.

In 1973, a resolution of the IAU proposed to move the origin of TAI by 32^s in order to ensure a sufficient agreement with TE. Perceiving some hesitation on the part of the astronomers, the Comité International des Poids et Mesures preferred to postpone a modification on the origin of TAI. Then the IAU withdraw its resolution and the occasion was lost to have a unique time for astronomy and the rest of human activities...

The WGRS had an opportunity to leave TAI unchanged, but to chose the origin of TT so as to cancel the time offset. This was first proposed in preliminary discussions and rejected quasi-unanimously. However, at the Virginia Beach Colloquium, discussion of this subject was re-opened and it appeared that the idea of aligning TT on TAI had gained acceptance, but not enough to carry a vote. Thus we are left with a time offset: on 1977 January 1, $0^h\ 0^m\ 0^s$, TAI, the reading of TT is, by definition, 1977 January 1, $0^h\ 0^m\ 32{,}184^s$.

The Terrestrial Time, thus defined, is strictly identical to TDT defined in 1976/79. However we have now a clear, unambiguous definition. The adjective 'dynamical' was removed in order to avoid the confusion with a realized time scale based on dynamics, such as TE. Nevertheless, TT, TCG, TCB appear, in these recommendations, as being both ideal coordinates of a dynamical model (GTR) and ideal forms of atomic time. This is a potential source of difficulties (Ref. 18).

5.4. General organization of time scales for astronomy

Figure 1 summarizes the relations between the time scales defined in 1976/79 and in 1991.

5.5. Notations

Time scales TT, TCG, TCB are ideal 'Times'. The IAU recommends notations such as TT(xxx) for realizations of these times, where xxx is an identifier. For example, a convenient realization of TT is

$$TT(TAI) = TAI + 32{,}184^s. \qquad (7)$$

Another example is the publication of TT(BIPM90) which is a time scale established in retrospect (Ref. 19).

The use of notations ^{h}, ^{m}, ^{s} (superscript) should be noticed. This is an IAU notation to express the date (Ref. 20), which has been adopted in the text of the Resolution on Reference Systems. This notation has the advantage of distinguishing the

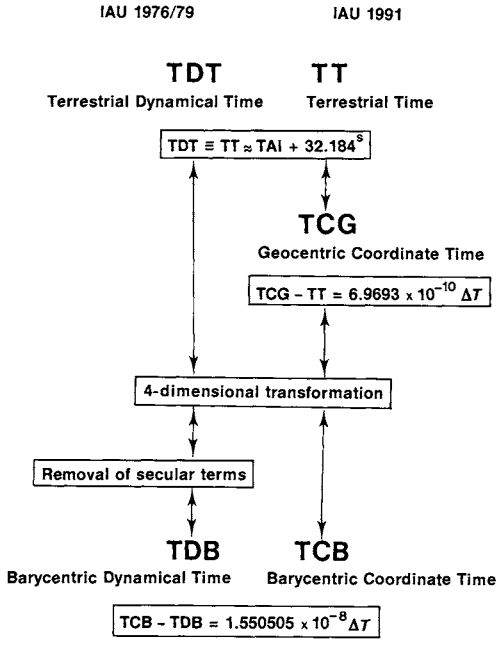


Fig. 1. Coordinate times defined for solar system dynamics in 1976/79 and 1991.

units of duration in proper time (in particular the SI second) and the designation of time markers in a specified time scale.

6. IAU RESOLUTION ON REFERENCE SYSTEMS RECOMMENDATIONS VI TO IX (1991)

It has not yet been possible to adopt a specific reference frame in conformity with Recommendation II (i. e. a catalogue of positions of extragalactic sources). A Working Group is required first to select useful sources and these sources need to be included in observing programmes. The resolution then indicates how the new reference frame should be initially set in continuity with the present frame (the FK5 star catalogue). A revolutionary conceptual decision was taken, to abandon the connection of the new reference system with the equatorial system (*i. e.* with the pecularities of the Earth's kinematics).

Concerning the astronomical units, constants and quantities, a permanent working group was created, under the leadership of T. Fukushima, to prepare reports before each IAU General Assembly (every three years) and also to prepare procedures for computation. Attention is also given to the need for electronic access to this information.

For nutation, the existing model is kept, but, when its accuracy is not sufficient, users are encouraged to take observed and predicted values published by the International Earth Rotation Service, as long as a better theory of nutation does not exist.

7. CONCLUSIONS

The history of the measurement of time shows many examples of unhappy decisions. We have seen some of them here: the time offset between International Atomic Time TAI and Ephemeris Time ET, the scale factors of time scales introduced in 1976. This can be explained by the complexity of time measurements and also by the pressure of users who wish to take immediate advantage of better time scales. Such circumstances probably occur in other scientific domains, but the consequences are worse in the treatment of time, because the heavy weight of continuity makes it difficult to revert to better decisions.

The International Astronomical Union, IAU, has tried to avoid future difficulties by remaining as closely as possible to the theory. Unfortunately this has led to contradiction with previous IAU decisions. The success of these new recommendations depends on their implementation in ephemerides, in the expression of constants and quantities, and in their use in procedures and software.

The implementation of these recommendations does not concern astronomers only. Geodesists, for example, had to decide whether they will use the coordinate system associated with Terrestrial Time TT (TAI, in practice), or the one associated with Geocentric Coordinate Time, TCG. The difference in distance scale is 7×10^{-10}, still small, but not negligible. They have chosen the TCG system (Resolution 2 of the International Union of Geodesy and Geophysics, 1991; see for example Ref. 21). This is in the spirit of the UAI resolution.

The IAU resolution provides us with a better definition of TAI and clarifies its role in coordinate systems, but it does not change it in any manner.

In the future, when definitions applicable to more accurate time standards become necessary, it will certainly be wise to retain the present situation in which TAI is a realization of TT (with time offset). But more precise definitions will require a better definition of the geoid, and decisions concerning the use of the tidal term (a few parts in 10^{-17} in normalized frequency).

8. AKNOWLEDGMENT

The preparation of the resolution which has been presented required many oral and written contributions, which, although important, where not published. At least for acknowledgment, some of their authors are mentioned below in 'References'.

9. REFERENCES

1. MacNamara M 1982, Basic postulates and time standard stability or the pursuit of fa, *S. Afr. J. Philos.*, 1, 83-97.

2. Danjon A 1929, Le temps, sa définition pratique, sa mesure, *Bull. Soc. astron. de France*, 43e année, 13-23.

3. Poincaré H 1905, *La valeur de la science*, Flammarion, Paris, 278 p.

4. Misner C W Thorne K S Wheeler J A. 1973, *Gravitation*, Freeman and Co, San Francisco, 1279 p.

5. Bretagnon p 1992, personal communication.

6. Stinebring D R & al 1990, Cosmic gravitational-wave background: limits for milli-second pulsar timing, *Phys. Rev Let*, 65, 3, 285-288.

7. Petit G 1992, Personal communication.

8. Guinot B 1971, Contribution to IAU Working Group on Units and Constants.

9. Winkler G M R 1990, Contribution to WGRS (letter, 9 March 1990).

10. Seidelmann P K & Fukushima T 1992, Why new time scales?, *Astron. Astrophys.* (in preparation).

11. Guinot B & Seidelmann P K 1988, Time scales - their history, definition and interpretation, *Astron Astrophys*, 194, 304-308.

12. McCarthy D D 1989, *IERS Standards (1989)*, Intern. Earth Rotation Service, Central Bureau, Paris.

13. Hughes J A Smith C A Kaplan G H (eds) 1991, *Proc. 127th IAU Col. "Reference System"*, US Naval Observatory, Washington, 419 p.

14. International Astronomical Union 1992, *Information Bulletin*, 67, 7-16.

15. Brumberg V A 1991, *Essential Relativistic Celestial Mechanics*, A. Hilger, Bristol, Philadelphia and New-York, 263 p.

16. Bureau International des Poids et Mesures 1991, *Le système international d'unités (SI)*, 6e édition, BIPM, Sèvres, 118 p.

17. Thomas C 1992, Personal communication.

18. Seidelmann P K 1990, Contribution to WGRS.

19. Guinot B 1988, Atomic time scales for pulsar studies and other demanding applications, *Astron. Astrophys.* 192, 370-373.

20. Astronomer's Handbook 1966, *Trans IAU vol XIIC*, Academic Press, London & New-York, 207 p.

21. Willis P 1991, IAG Newsletter, *Bull. Géodésique*, 65, 3, 189-208.

HOW CAN MILLISECOND PULSARS IMPROVE THE LONG TERM STABILITY OF ATOMIC TIME SCALES?

G. Petit [1], P. Tavella [2], C. Thomas [1]

[1] BIPM, Pavillon de Breteuil, 92312 Sèvres CEDEX, France
[2] IEN, Strada delle Cacce 91, 10135 Torino, Italy

ABSTRACT

Millisecond pulsars are galactic objects that exhibit a very stable period of rotation. Since the first discovery in 1982, radioastronomers have carried out timing observations, which measure the time of arrival of the radio pulses relative to an atomic time scale. Several tens of millisecond pulsars have now been discovered, located in different parts of the sky. If they are observed regularly for many years, they can be considered as an ensemble of clocks which, although they are affected by frequency drifts, allows an average time scale to be deduced through a stability algorithm. We present the basis for such an algorithm and apply it to simulated pulsar data. We discuss the properties of the resulting time scale and how it could monitor the stability of the atomic time scale.

Keywords: Pulsars, Atomic time, Time scales, Stability algorithms.

1. MILLISECOND PULSARS

The first millisecond pulsar was discovered in 1982 (Ref. 1) and it was quickly recognized that its period of rotation is very stable. Since that time, astronomers have carried out timing observations, which measure the time of arrival (TOA) of the radio pulses relative to an atomic time scale. They have also discovered many more millisecond pulsars which are now regularly observed at a number of observatories.

The technique allows to reach a resolution in time of order 1 μs over the whole period of observation, which is extremely good for an astronomical measurement. This turns to an angular resolution of order 10^{-9} rad, which is presently the ultimate astrometric precision. Such resolutions open new paths in astronomy and astrophysics: detection of gravitational waves, probe of relativity in strong fields and for high velocities, detection of orbiting planets, etc.

As the measurements are related to atomic time, pulsar analysis demands an ideal standard of atomic time. Conversely it must be possible to transfer the stability of the rotation of pulsars to atomic time if this appears better.

2. WHAT MILLISECOND PULSARS DEMAND OF ATOMIC TIME

2.1 Description of the pulsar data analysis

The analysis can be described simply in the ideal case of a pulsar with a constant period, without motion relative to the solar system, and fictitiously observed at the geocentre. In this case the TOA observations should be regularly spaced except for an annual variation due to the orbital motion of the Earth. This signature would determine the astrometric position of the pulsar, and the period would be determined as the average spacing between TOAs.

The true situation is more complicated, but the general description is similar: the linearly varying period (which produces a parabolic signature in the TOAs) and the position and proper motion of the pulsar (which produce annual signatures) are parameters determined by a fit to the observations. The residuals to the fit include information on whatever physical phenomenon has not been accounted for in the model. These may include interstellar propagation, instabilities due to the pulsar, and instabilities in the atomic time scale (AT), which are the subject of our study.

2.2 Requirements on the atomic time scale

It follows from the above description that, if astronomers and astrophysicists want to extract some significant information from the residuals, the atomic time scale must not introduce errors larger than the measurement uncertainty. If σ is this uncertainty, and T the duration of the period of observation, this analysis sets two requirements on the atomic time scale: that the order of magnitude of its annual variation be smaller than σ, and that its stability for an averaging time τ be better than σ/τ (for values of τ up to T).

In the present state of pulsar observations, σ is as low as 0.3 μs in the best case (Ref. 2), but more generally is at the level of 1 μs, while T has reached 9 years for the first discovered pulsar PSR1937+21. The first requirement is thus probably nearly satisfied by the best atomic time scales, which should have annual variations not larger that 0.3 μs (Ref. 3). The second requirement is not fulfilled for τ larger than about one year, the stability of the present best realization of an uniform atomic time being estimated at about 2.10^{-14} (Ref. 4). Statistical analysis of the timing residuals of PSR1937+21 over several years of observation (Ref. 2,5) indicates that the intrinsic stability of this pulsar is indeed at least as good as that of the atomic time scale for averaging times larger than one year.

3. WHAT MILLISECOND PULSARS CAN BRING TO ATOMIC TIME

3.1 Annual variations of atomic time

As the position of the pulsar is directly related to an annual signature in the TOAs, it is possible to get information on an annual variation of AT if the position can be taken from another source. This in theory is possible because Very Long Baseline radio-Interferometry (VLBI) can determine the position of the pulsars in an independent way. In practice however, it is unlikely that the position derived from VLBI be accurate enough to detect the expected variation of AT: it would require an accuracy of 5.10^{-10} rad in the position to detect a 0.5 μs amplitude in AT, while the uncertainty in the relationship between the

VLBI and pulsar celestial reference frames is at least one order of magnitude larger.

Another possibility would be to have at least two pulsars close to the ecliptic: in this case the accuracy requirement on the ecliptic latitudes can be relaxed so that the latitudes can be adopted from other sources. It would then be possible to solve for the longitudes of the pulsars and the amplitude and phase of a yearly sine function representing the error in AT.

3.2 Long term stability of atomic time

It is not possible to define a pulsar time scale independently from atomic time. This is due to the variation of the rotation period of the pulsar, which cannot be described by a physical model. Its variation is assumed to be linear, and its rate can be determined only with reference to an atomic time scale (Ref. 3).

Any instability in the atomic time should show up in the residuals of the TOA analysis unless it is either below the level of the measurement noise σ, or filtered out by the fit of the parameters. The first case includes all noise sources that dominate at high Fourier frequencies (period lower than about one year), which should not reach the level σ. The second case includes the annual variations which are filtered out when solving for the pulsar position and proper motion, and noise sources that dominate at low Fourier frequencies (periods of order T), which are filtered out when solving for the rotation period and derivative. Thus the stability of AT could be estimated for averaging times above one year and up to about $T/2$, assuming the pulsar rotation is perfectly stable and not degraded by external phenomena.

We cannot separate the intrinsic instability of the pulsar rotation from other sources of noise, including atomic time, but, having collected observations from many pulsars, we can try to smooth the random noise of each pulsar data by making a weighted average of them. We then define an "ensemble pulsar time", PT, whose random noise will be reduced below that of any individual pulsar data, and below that of atomic time. To do this we have to devise an appropriate stability algorithm, a procedure similar to that used for atomic clocks.

Since pulsar data can bring a significant contribution to atomic time only after many years of observations, we are interested in building up a pulsar time for use as a long term time reference and optimized to provide long term stability. Because of the long observation time needed, it is more convenient to calculate PT a posteriori. It would then be possible to determine the instabilities of the atomic time scales, for averaging times of some years, and to correct for identifiable systematic trends, if any.

4. Simulation of a Pulsar Time Scale

4.1 Description of the simulated data

We have simulated data describing 8 pulsars observed continuously over a very long period. For this purpose, we generated 8 series of 1200 numbers following statistical criteria defined below. They represent the residuals of the TOA measurements relative to a model whose basic parameters are the period of rotation at a given time, the variation rate of the period, the position and the proper motion of the pulsar. This is equivalent to defining an individual time scale PT_i for pulsar i as follows: if A is the event "arrival of pulse n of pulsar i at a given observatory", $PT_i(A)$ is the TOA computed from the model for this pulse. With a radiotelescope, any event can be compared to the arrival of a pulse, at least in theory, just as any event can be compared to the 1pps of an atomic clock.

The residual errors of the TOA model come mainly from the following sources: measurement noise, errors in the ephemerides of the solar system, interstellar propagation noise and intrinsic instabilities of the pulsar rotation. They are simulated as follows: the measurement noise is represented as white phase noise at a level σ_i, varying with each pulsar. Ephemerides errors are represented as a one year sine function of amplitude smaller than σ_i (terms with longer periods could exist, but they will be mixed with the stochastic terms below). The remaining errors are simulated as random walk frequency noise at a level consistent with the quality of present day atomic time scales over a few years. This hypothesis, which corresponds to a phase spectral density proportional to f^{-4}, may not exactly represent the true errors: it has been suggested (Ref. 6) that spectral densities with slopes -3, -4, -5 and beyond would be present coming from interstellar density fluctuations, intrinsic noise, and gravitational radiation, respectively. The most important source is probably the first, but it is expected that dual-frequency observations will allow a substantial decrease in its influence on the measurements. Finally a small residual frequency drift is added. The period and phase of rotation at a given time, which are the remaining parameters of the TOA model, are not significant for the stability analysis.

Each of the series represents the simulated residuals $(AT\text{-}PT_i)$ over 33 years, with one observation every ten days, of pulsar i relative to an ideal atomic time scale AT. Their Allan deviations are plotted in figure 1.

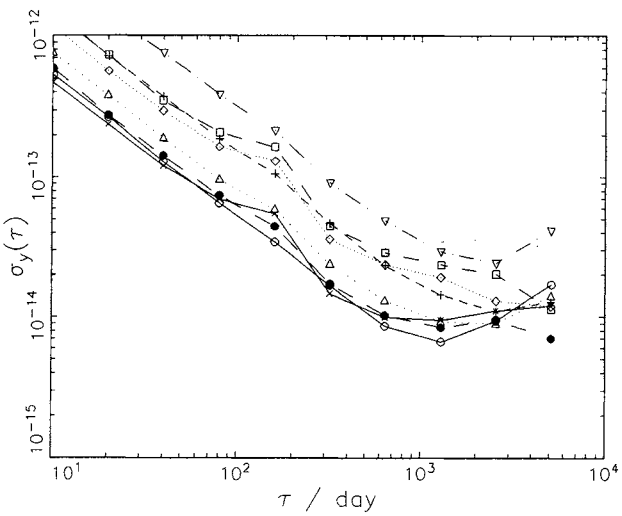


Fig. 1: Square root of the Allan variance of 8 simulated series $(AT\text{-}PT_i)$. Individual series from 1 to 8 are designated by circle, square, triangle (apex up), plus, diamond, triangle (apex down), star, solid circle.

4.2 Description of the algorithm

In analogy to what it done for atomic time (Ref. 7), the definition of the "ensemble pulsar time" PT is a simple weighted average of the basic measurements $(AT\text{-}PT_i)$. It is defined in the form of the difference AT-PT, as:

$$AT - PT = \sum_{i-1}^{8} w_i \, (AT - PT_i)$$

where w_i is the relative weight assigned to pulsar i and is characteristic of its long term stability.

We suppose that we have a fixed number of pulsars at our disposal for the whole period. We also assume that the intrinsic stability of each pulsar does not significantly change with time so that we can compute the average with a fixed set of weights. In this particular case, a frequency prediction to ensure continuity is not necessary [Ref. 8].

Experiments have been carried out using two different weighting choices. In a first step, PT has simply been computed with a fixed set of equal weights ($w_i = 1/8$). In a second step we have evaluated the instability of each pulsar time by computing its Allan variance and have set its weight inversely proportional to such variance. Since we are interested in building an ensemble pulsar time to be used as reference time scale for averaging times of some years, we have chosen to optimize the stability of the ensemble over averaging intervals of 1000 days and, consequently, we have defined the weights as:

$$w_i = \frac{1}{\sigma_i^2(1000d)} \left[\sum_k \frac{1}{\sigma_k^2(1000d)} \right]^{-1}$$

that is as the inverse of the Allan variance for 1000 days, normalized to unity.

In a simulated experiment, we have a perfect knowledge of the variances of the individual pulsar data. In the real case, it is necessary to estimate them. This can be done with the n-cornered hat technique (Ref. 9) in which, assuming that the pulsars are independent, it is possible to estimate individual variances from the pair differences PT_i-PT_j. Such differences can easily be obtained by differencing (AT-PT_i) and (AT-PT_j), eliminating the contribution of AT. We used this technique to estimate the Allan variance of each pulsar time.

4.3 Optimum estimation of the PT scale unit

It is known that each PT_i has a residual error in the period and its derivative, thus the average PT has also such errors and its scale unit does not conform with the definition of the atomic second. Some processing has to be done to derive from PT a new time scale, PT', in which the scale unit conforms with AT in an optimal way. We do not apply this to our simulated data because we are interested in stability, and the stability of PT' is similar to that of PT. However it will be necessary to address this problem when applying the method to real data.

4.4 Results

We have carried out the computation of PT for the two choices of weights mentioned above: the equal weight average and the Allan variance weighted average. In the second case, the 1000-day Allan variances lead to the relative weights listed below, where the numbering of pulsars refers to figure 1.

Pulsar	1	2	3	4	5	6	7	8
Weight	0.35	0.03	0.16	0.04	0.04	0.01	0.19	0.18

The instability of the ensemble pulsar time PT can be estimated from the difference (AT-PT). As AT is ideal, the stability of (AT-PT) is an estimate of the quality of PT and, hence, of the adequacy of the algorithm used. Estimates of the Allan deviation for the two cases considered are reported in figure 2 where it can be seen that the equal weight and the Allan variance weighted averages both have better stability than any single pulsar for the averaging times of several years that we have used.

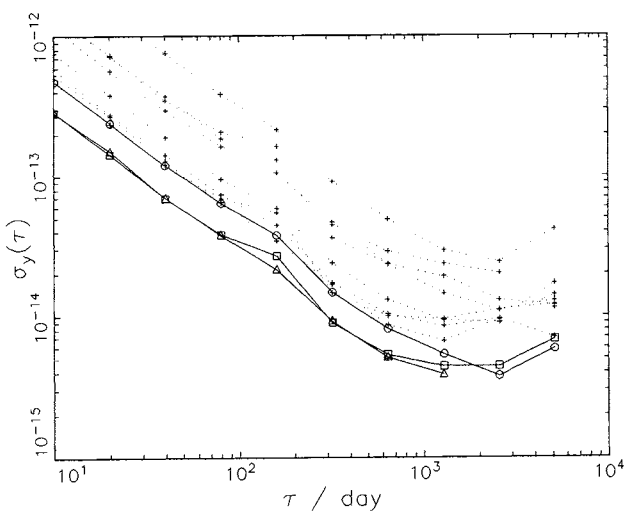


Fig. 2: Square root of the Allan variance of the average pulsar time scale: equal weight (circles), weights estimated from the 1000-d Allan variance with the whole data set (squares), and with the first 5000 days (triangles). Individual variances from figure 1 are shown as light dotted lines.

A second problem we have studied with this simulation is the degree to which the duration of the data collection period affects the stability of the derived PT. Supposing we have access only to the first 3000 or 5000 or 8000 days of data, the stability over 1000 days has been estimated and the corresponding average PT computed in a way similar to that used with the complete data set. While it is recognized that the confidence level of the Allan variances improves with the number of data used, the estimated relative weights do not change very significantly, nor does the stability of the ensemble PT. As an example the Allan deviation of PT using 5000 days of data is shown in figure 2.

5. Discussion of results

5.1 Determination of instabilities in atomic time

Pulsar data have been simulated assuming that the long term random errors in the pulsar timing residuals result in a stability comparable to that of the present day atomic time scales, i.e. 2.10^{-14}. This assumption is based on the present knowledge of these errors and on the results of the analysis of several years of pulsar data (Ref. 5).

For averaging times of a few years, the stability of the averaged pulsar is much better than that of the present day atomic time scale, and that should allow to estimate its long term systematic trends, according to the aim of this experiment. Figure 3 shows the deviation of the weighted average PT from the ideal time scale AT, together with the deviations (AT-PT_1) and (AT-PT_4) of pulsar 1 and 4, which are among the best and the worst in 1000-day stability, respectively. It appears that the averaging procedure effectively smooths some of the noise from each pulsar time and allows the definition of an improved PT for use as a long term reference. In the simulated case, it would be possible to detect instabilities in the atomic time of order 10^{-14}, over averaging times of a few years.

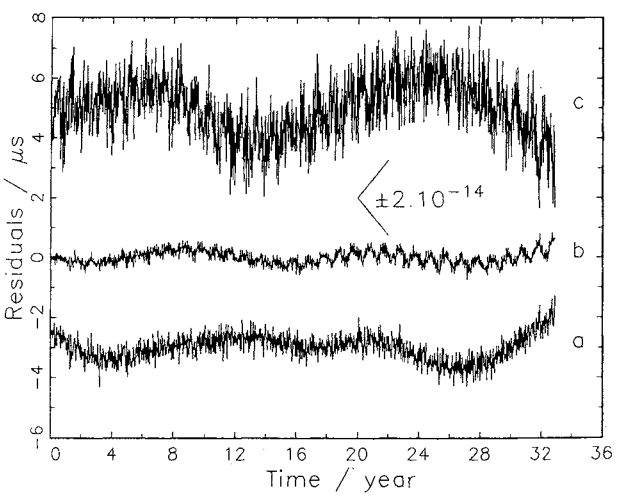


Fig. 3: Residuals (PT-AT) for (a) a pulsar with a good stability, (c) a pulsar with one of the worse stabilities, and (b) the average pulsar time. Weights are estimated from the 1000-d variances. The slope in the insert indicates the level of instabilities of atomic time. A parabolic function of time has been removed from each series.

As far as averaging times of 1000 days are concerned, ten years of observation may be sufficient for satisfactory results and this is nearly reached for pulsar 1937+21. Unfortunately most of the known millisecond pulsars have been discovered very recently, and the received signals from most of them are very weak, so the noise pattern of their residuals is still dominated by measurement noise. Apart from 1937+21, two other pulsars, 1855+09 (Ref. 5) and 1821-24 (Ref. 10) are good candidates, having been observed for several years and having reached the stability level of a few parts in 10^{14}.

5.2 Transfer of the accuracy of atomic time

It is possible to take advantage of the long term stability of PT to transfer the accuracy of atomic time from one period of time to another, provided continuity in the pulsar observations is maintained. After an initial period of data collection of length T, PT is computed using the proposed algorithm, and PT' is derived, in which the scale unit conforms to the atomic second over this initial period. As the long term frequency stability of PT is much lower than the uncertainty in the realization of the atomic second, σ_u, the inaccuracy of PT' is at the level σ_u. Due to the linear variation of the rotation rate of pulsars, a residual frequency drift is present in a pulsar time scale. For PT', this residual frequency drift is of order σ_u/T. Having collected pulsar data over a duration τ of a few years after the initial period, it is possible to compute a new time scale, starting from PT' and applying the stability algorithm on the new pulsar data. The inaccuracy of this new time scale over the extension period is dominated by the initial error σ_u. This is because the contribution from the residual frequency drift is negligeable compared to σ_u as long as τ remains much smaller than T, and because the intrinsic stability of PT over a duration τ is also much smaller than σ_u.

As an example, assume that we compute PT and PT' after 10 years of pulsar observations. The error σ_u, as provided by the present frequency standards, is of order 2.10^{-14}. Then assume that the frequency standards that provide the atomic second fail simultaneously. After a few more years (period of duration τ, smaller than T which is 10 years) without access to the atomic second, the scale unit of PT still has an uncertainty of order 2.10^{-14}, while that of an atomic time scale based on industrial clocks could be off by several parts in 10^{14} over the same period (such variations are typically observed over two years when comparing an unsteered atomic time scale with a frequency standard). Furthermore, when frequency standards come back to operation, they will allow to compute a new PT' over the whole period, that will provide full access to the atomic second even during the period when the standards were off.

Similarly, when a future frequency standard with improved stability has been in continuous operation for years (and eventually provides a new definition of the second), PT will allow to transfer it backwards so that it will be possible to evaluate the accuracy of our present atomic time scales. In that case, however, the limitation could come from the stability of PT itself, which may be worse than that of the new standard for long averaging times.

5.3 Adaptation of the stability algorithm

In the simulated experiment, we assumed a fixed number of pulsars. The real case is likely to be different because new pulsars will always be discovered and observed and therefore entered in the ensemble. No removal of a pulsar from the ensemble should happen from a failure in pulsar rotation, as it is expected that millisecond pulsars may live billions of years in an observable state (Ref. 5), but it is quite possible that an observatory may stop operation. For these reasons, the definition of the average should contain a suitable frequency prediction to ensure continuity of the ensemble time, when the number of pulsars changes.

A contribution to the instability pattern of PT, over very long averaging times, will arise from the effect of the residual period derivative being incompletely corrected by the model. Such residual drift, leading to an Allan variance proportional to τ^2, could obscure the random noise characteristics of a pulsar for the same averaging times. This problem should arise only for very long averaging times but, if necessary, the residual drift can be filtered in the instability estimation, by using a different estimator with third order difference, such as the Hadamard variance weighted by binomial coefficients (Ref. 11), which overcomes the problem.

6. Conclusions

Without making any assumption on the unmodelled phenomena which corrupt millisecond pulsar timing observations, it is possible to process such data from an ensemble of pulsars to derive an average pulsar time scale that has a stability better than atomic time and better than the time derived from individual pulsar data. This improvement holds for averaging times above one year up to about half the period of observation.

It seems adequate to use a weighted average, with weights inversely proportional to the variances, over an averaging time of a few years. The data from each pulsar being independent, the variances can be estimated from a n-cornered hat analysis.

The application of the above procedure to real data requires first the availability of many years of data on a number of pulsars, then the possibility to intercompare data taken at different observatories. It may be needed to upgrade the simple algorithm presented here to account for real world events (irregularly spaced data, apparition and disparition of pulsar data), but it should be possible to estimate the instabilities of the present realizations of Terrestrial Time by the end of the century. In return, this improved time scale will yield a better knowledge of the physical phenomena that affect individual pulsars.

References

1. Backer D C, Kulkarni S R, Heiles C, Davis M M, Goss W M 1982, A millisecond pulsar, Nature, 300, 615-618.

2. Rawley L A, Taylor J H, Davis M M, Allan D W 1987, Millisecond pulsar PSR1937+21: a highly stable clock, Science, 238, 761-765.

3. Guinot B, Petit G 1991, Atomic time and the rotation of pulsars, Astron. Astrophys., 248, 292-296.

4. Bauch A, Dorenwendt K, Fischer B, Heindorff T, Müller E K, Schröder R 1987, CS2: The PTB's new primary clock, IEEE Trans., IM-36, 613-616.

5. Taylor J H 1991, Millisecond pulsars: Nature's most stable clocks, Proc. IEEE, 79, 7, 1054-1062.

6. Blandford R, Narayan R, Romani R W 1984, Arrival-time analysis for a millisecond pulsar, J. Astrophys. Astr., 5, 369-388.

7. Tavella P, Thomas C 1991, Comparative study of time scales algorithms, Metrologia, 28, 57-63.

8. Guinot B 1987, Some properties of algorithms for atomic time scales, Metrologia, 24, 195-198.

9. Barnes J A 1982, Time scale algorithms using Kalman filters: insights from simulations, 2nd Symp. on atomic time scale algorithms, Boulder CO.

10. Lestrade J-F 1992, Personal communication.

11. Rutman J 1978, Characterization of phase and frequency instabilities in precision frequency sources: Fifteen years of progress, Proc. IEEE, 66, 9, 1048-1075.

RELATIVISTIC EFFECTS FOR LOW EARTH ORBIT SATELLITES USING GPS

A. Spallicci[1] , C. Jiménez[1,2], G. Prisco[1], N. Ashby[3]

1. European Space Research and Technology Center, ESTEC, Noordwijk, the Netherlands
2. Facultad de Ciencias, Sección de Físicas, Universidad de Zaragoza
3. Department of Physics, University of Colorado at Boulder

ABSTRACT

The relativistic corrections for a low earth orbit satellite are evaluated. The GPS satellite clock rate is slowed before launch by 4.465×10^{-10}, called the "factory offset", for time dilation and gravitational frequency shift. This offset cancels the main constant relativistic effects for terrestrial users, who in order to operate in coordinate time have only to process the GPS orbital eccentricities, a sinusoidal function whose peaks are in the order of tens of ns, and the Sagnac effect. For a space user the situation greatly differs, because a large part of the relativistic effects are still present due to the high velocity of the satellite and its location in the Earth gravitational field. Past tests and proposals for future measurements with GPS (perigee advance, Shapiro time delay, preferred frame independence, Lense Thirring effect, light bending and gravitational waves), are briefly reviewed.

Keywords: Relativity, Fundamental Metrology, GPS, Satellites.

1 INTRODUCTION

The GPS system consists of a constellation of satellites ultimately numbering 18 plus three active spares, in six orbital planes, in 12 hour orbit, at 20183 km altitude and 55 degrees inclination. The Navigation data message (Van Dierendonck et al., 1980) is contained in a frame , subdivided in five subframes each including system time. The first subframe contains the GPS satellite clock correction (including relativistic effects) and ionospheric propagation delay parameters. The second and the third subframes contain the ephemeris of the GPS under tracking. The fourth subframe contains alphanumeric characters. The fifth subframe contains the ephemeris, the clock corrections and the status of the other GPS satellites. This constitutes an almanac aiding the User to acquire other GPS satellites.

Each satellite carries a set of atomic clocks (GPS block 1 has one Cesium and three Rubidium, block 2 has two Cesium and two Rubidium, whereas for the future Maser clocks are considered), which control the value and the stability of two continous radio frequency transmissions at 1.2276 and 1.57542 GHz, called L2 and L1 respectively. Appropriate code division multiplexing techniques allow to differentiate among the several GPS satellites.

There are three relativity physical phenomena to be considered:

- Time dilation (Harkins, 1979; Hafele et al., 1972 a,b), amounts to about -8.35×10^{-11} for GPS.
- Sagnac effect (Czajko, 1990; Hafele et al., 1972 a,b; Post, 1967; Sagnac, 1913, 1914).
- Gravitational frequency shift (Harkins, 1979; Vessot, 1979), amounts to about 5.3×10^{-10} for GPS.

The nominal frequency of the reference source is 10.23 MHz as it appears from the ground, but the frequency in the satellite is 10.22999999543 MHz: when the satellite is launched into its orbit, relativity causes the clock frequency to increase by +4.568 mHz, which corresponds to about 38.5 μs per day.

The compensation performed before launch consists simply in tuning the frequency standard by an offset equal to -4.568 mHz , $\Delta f/f = -4.465 \times 10^{-10}$ (Rockwell Int. Corp., 1987; McCaskill et al., 1978). The GPS orbital eccentricity produces further relativistic effects of the order of tens of ns and the correction parameters are sent in the Navigation message. GPS system time is controlled with reference to UTC. The present uncertainty of the system is 10 ns in time, or 3 m in position, but improvements down to 1 ns accuracy might occur (Will, 1986).

2 RELATIVITY TESTS AND FUTURE MEASUREMENTS WITH GPS

The first two tests quoted in this section are concerned with past measurements, which are actually incorporated in the system design of GPS, whereas the remaining six are potential future measurements with an improved GPS system operating at subnanosecond accuracy (Ashby, 1990 a,b; Will, 1986).

2.1 Relativistic Doppler shift

During tracking of GPS-NTS 2, a relativistic Doppler fractional frequency shift of 4.473×10^{-10}, consistent with theoretical predictions was observed (Harkins, 1979).

2.2 Sagnac effect

The Sagnac effect is proportional to the equatorial projection of the area swept out by the radius vector from the center of the Earth to the electromagnetic pulse as it propagates from the point of transmission to the point of reception.

An around-the-world Sagnac experiment was performed with electromagnetic signals from GPS satellites, using timing centres to accomplish the circumnavigation: Astron. Obs. in Tokio, Nat. Bureau of Standards in Boulder and the Phys. Techn. Bundesanstalt in Braunschweig, Federal Germany (Allan et al., 1985, 1984). The time

shift was 1.6227×10^{-21} sec/meter^2 A equal to $240 \div 350$ ns ± 5 ns, in one circumnavigation.

2.3 Perigee advance

The GPS perigee advance would be given by (Will, 1986, 1981):

$$\delta\omega = \frac{6\pi GM}{pc^2}\left[\frac{1}{3}(2+2\gamma-\beta)+J_2\frac{R^2c^2}{2GMp}\right] \quad (1)$$
$$= 8 \text{ cm/orbit or } 60 \text{ m/year.}$$

where:

G =constant of gravitation,

$M \simeq$ Earth mass,

$p = (1-e^2)\, a_G$,

e = orbit eccentricity,

a_G = semi-major axis of the orbit of the GPS satellite,

γ = PPN (Will, 1981) formalism parameter measuring how much space curvature is produced per unit rest mass; in General Relativity $\gamma = 1$,

β = PPN (Will, 1981) formalism parameter measuring the nonlinearity in the superposition law for gravity; in General Relativity $\beta = 1$,

R = Mean radius of the oblate body,

$J_2 = (C-A)/MR^2$, C and A are the moments of inertia about the body rotation and the Earth equatorial axis respectively,

c = speed of light.

2.4 Shapiro time delay

It is the retardation of an electromagnetic wave passing near a massive body (Shapiro, 1964), experimented ranging Mariner and Viking spacecrafts, Mercury and Venus. The travel time delay, due to Earth mass, for GPS has been considered by Will (1986).

2.5 Preferred frame independence

The velocity of light is supposed to be independent from the frame it propagates in. GPS time transfer at 1 ns accuracy may set a limit on the fractional variation at 10^{-8} level (Will, 1986).

2.6 Dragging of inertial frames (Lense Thirring effect)

A rotating mass drags the space near it into rotation. The plane of a satellite orbit is dragged around an axis aligned with the Earth rotation axis. For GPS the amount would be given by (Will, 1986):

$$\Omega_{dragging} = 4 \times 10^{-10}\, \omega_e \left(\frac{R_e}{r}\right)^3 = \quad (2)$$
$$= 1.4 \times 10^{-8} \text{ rad/year or } 40 \text{ cm/year.}$$

where ω_e = Earth rotational angular velocity.

2.7 Light bending

The effect would be in the order of few centimeters as it results from finding the null geodesic along which the electromagnetic signals propagate (Ashby, 1988).

2.8 Gravitational waves

The use of the GPS constellation for the search of gravitational waves has been proposed (Hansen et al., 1986; Hansen, 1981). The entire 21 satellites constellation is considered a gravitational wave antenna made up of 42 space to ground links. Doppler tracking of the carriers L1 and L2 could provide continuous, 24 hours and global 4π stereoradians coverage. The latter property provides the direction of the wave origin and the polarisation components. The major difficulty is tracking through the ionosphere frequency whose noise is about 10^{-11}, while the Cesium frequency stability is 10^{-13}. A computer simulation showed the experimental feasibility, as discussed by Thorne (1987) for gravitational waves amplitudes h_g well above $10^{-13 \div 14}$, corresponding to improved Cesium clocks or to Maser clocks, via modelling of the ionosphere, reduction of the ephemeris errors and use of an accurate Earth gravity model.

The GPS system maximum sensitivity occurs between 1.4 and 3.7 Hz in the gravitational frequency spectrum, according to:

$$\frac{\Delta f}{f_0} = h_g \sin\left(\frac{2\pi L}{l_g}\right) \quad (3)$$

where:

f_0 = reference carrier frequency,

h_g =dimensionless amplitude of the gravitational wave. It is the perturbation of the spacetime metric from its Minkowsky (flat) value,

L = GPS altitude = 20280 km or global diameter of the constellation = 53094 km.

l_g =wave length of the gravitational wave ($= 4L$ for the maximum sensitivity).

The detection of gravitational waves is based upon small additional Doppler shifts in the reference rf carrier frequencies used for space to ground GPS links, between two free masses (GPS satellite and the Earth). Two pulses signature (one way link), instead of the classical three pulses for interplanetary spacecrafts, would then be produced at the monitoring station. A long integration time would allow to enlarge the frequency spectrum towards lower frequencies.

3 THE GPS SYSTEM AND GROUND BASED USERS

A group of scientists, in the last 70's and early 80's have claimed that the relativistic corrections were not taken into account in the GPS system design and that the GPS system would have failed to meet its specifications (Skalafuris, 1985). A large polemic arose until it was later shown that relativity was properly taken into account (Will, 1986; Eardley et al., 1985).

Ashby and Allan (Allan et al., 1986; Ashby et al., 1984, 1979) suggest an underlying Earth nonrotating frame (ECI= Earth centered inertial), in which a fictitious set of standard synchronised clocks results in a time grid always valid everywhere (also on the surface of the rotating Earth, where time t' is related to coordinate time t).

The expression for the elapsed coordinate time on the GPS satellite clock (Ashby, 1990 a; Jorgensen, 1986) is:

$$\Delta t' = \int dt_G \left[1 + \frac{3GM}{2a_Gc^2} + \frac{\Phi_0}{c^2} - \frac{2GM}{c^2}\left(\frac{1}{a_G} - \frac{1}{r_G}\right)\right] \quad (4)$$

where:

$t' = \left(1 + \frac{\Phi_0}{c^2}\right) t$ = time as measured by standard clocks at rest on the surface of the Earth (time on the geoid),

t =time as measured by standards clocks at rest at infinity (coordinate time),

t_G = proper time (time measured on the GPS satellite),

r_G = distance from the center of the earth to the GPS satellite,

Φ_0 = gravitational effective potential on the geoid,

$= -\frac{GM}{R_e}\left(1 + \frac{J_2}{2}\right) - \frac{1}{2}(\omega_e R_e)^2 =$

$= (-6.95360 \times 10^{-10} - 3.77 \times 10^{-13} - 1.203 \times 10^{-12})\, c^2 =$

$= -6.9694 \times 10^{-10}\, c^2$; the first term is the monopole contribution, the second and third ones are the quadrupole and centripetal contributions respectively.

The terms $\frac{3GM}{2ac^2} + \frac{\Phi_0}{c^2}$ represent a constant $= 4.465 \times 10^{-10}$, for which the correction is equal to -4.465×10^{-10}, "the factory offset". The last term in expression (4) results in a correction which shows a sinusoidal oscillation with the position of the GPS satellite in its orbit (Ashby, 1990 a):

$$\begin{aligned} \Delta t' &= \int \frac{2GM}{c^2}\left(\frac{1}{r_G} - \frac{1}{a_G}\right) dt' = \\ &= \frac{2}{c^2}\sqrt{GM a_G}\, e \sin E = \\ &= 4.4428 \times 10^{-10} \frac{\text{seconds}}{\sqrt{\text{meter}}}\, e \sqrt{a_G} \sin E \end{aligned} \quad (5)$$

where:

dt_G has been taken as $\simeq dt'$,

e = eccentricity of the GPS satellite orbit,

E = eccentric anomaly of the GPS satellite.

This correction must be made by the receiver. For $e = 0.01$ the amplitude is about 23 ns.

The data needed for these calculations are supplied in the navigation message, whose elaboration is discussed by Van Dierendonck et al. (1980).

The second correction to be made by the ground based receiver is due to the Sagnac effect: one should allow for extra propagation time due to motion of the receiver:

$$\Delta t' = \frac{2}{c^2}\, \vec{\omega_e} \cdot \vec{A} \quad (6)$$

where :

$\vec{A} = \frac{1}{2}\, \vec{r_A} \times \vec{r_B}$, is the vector area swept out by a vector from the center of the Earth to the electromagnetic signal as it travels from the point of emission A to the point of reception B,

$\vec{\omega_e}$ = Earth rotational angular velocity vector.

4 LOW EARTH ORBIT (LEO) SATELLITES

In the case of Space GPS users the corrections to be implemented are quite different from the case of Earth users and the literature available (Deines, 1990; Lear, 1989 a,b; Matchett et al., 1980), covers this topic only partially.

The GPS system was originally conceived as a positioning and synchronisation system for ground users, which are slowly moving, while a fast aircraft has typically a short duration mission and unaccounted relativity effects don't have the time to accumulate.

The requirement of higher precision in positioning and synchronisation for ground users and aircrafts and the appearance of a new class of users, i.e. space users, pose new matters that need to be solved.

For space users the relativistic correction differs because:

- LEO satellites have large velocity with respect to a GPS satellite, even twice.
- Due to the air drag orbital decay the position of the LEO satellites in the Earth gravitational field may change considerably , this affects the gravitational shift and the time dilation (the orbital velocity changes with the change in position). An Earth user is instead located in a different and approximately constant position in the gravitational field.
- Two eccentricities have to be taken into account (e_{LEO} and e_{GPS}).

For the GPS satellites the relativistic effects produced by the geoid surface inequalities, higher harmonics of the Earth gravitational field and Sun and Moon gravitational fields may be neglected except for their influence upon the GPS ephemeris. The gravitational effects of the Sun and the Moon are cancelled to high accuracy by a second-order Doppler shift due to the Earth orbital motion around the Sun (Ashby et al., 1984, 1979). The Sun gravitational field is transformed away by adopting an Earth centered, locally inertial, freely falling reference system . A small residual tidal effect of few parts in 10^{17} (Ashby, 1990 a) is due to the different location of the GPS clock and Earth center in the Sun gravitational field .

In view of the possible future utilisation of Maser clocks on board of the GPS satellites and the crosslink ranging among GPS satellites (Jorgensen, 1989; Ashby, 1988), the impact of relativity upon the overall system time maintenance and transfer accuracy shall demand even more concern.

4.1 LEO proper time as a function of the transmitted time by GPS

We wish to relate the proper time on the LEO satellite and the proper time on GPS.

$$dt' = dt_G \left[1 + \frac{1}{2c^2} v_G^2 - \frac{1}{c^2}(V_G - \Phi_0)\right] \quad (7)$$

$$dt' = dt_L \left[1 + \frac{1}{2c^2} v_L^2 - \frac{1}{c^2}(V_L - \Phi_0)\right] \quad (8)$$

where:

t' = time on geoid (= coordinate time),

t_L = proper time on LEO,

$v_{L,G}$ = velocity of LEO or of the satellite vehicle of the GPS system, respectively (ECI coordinates),

$V_{L,G}$ = potential including J_2 at LEO or GPS position, respectively , $= -\frac{GM}{r}\left[1 - J_2\left(\frac{R_e}{r}\right)^2 P_2(\cos\theta)\right]$,

R_e = equatorial radius of the Earth,

r = distance from the Earth's center,

P_2 = Legendre polynomial.

θ = polar angle of the point of observation measured down from the axis of rotation.

Neglecting the effect of the quadrupole moment on the GPS satellite, we can write:

$$dt_L = \frac{1 + \frac{3GM}{2a_G c^2} + \frac{\Phi_0}{c^2} - \frac{2GM}{c^2}\left(\frac{1}{a_G} - \frac{1}{r_G}\right)}{1 + \frac{1}{2c^2}v_L^2 - \frac{1}{c^2}(V_L - \Phi_0)} dt_G \quad (9)$$

Due to the factory offset the expression (9) is simplified into:

$$dt_L = \frac{1 - \frac{2GM}{c^2}\left(\frac{1}{a_G} - \frac{1}{r_G}\right)}{1 + \frac{1}{2c^2}v_L^2 - \frac{1}{c^2}(V_L - \Phi_0)} dt_G \quad (10)$$

4.2 Order of magnitude of the relativity offset

We give a back of the envelope result in order to have an idea of the order of magnitude of the relativity offset for a typical LEO orbit. We work out the result in ECI.

We shall not take into account eccentricities, the quadrupolar moment of the Earth and the centripetal potential contribution ($V, \Phi_0 \simeq -\frac{GM}{R}$). Then:

$$dt_L \simeq \left(1 - \frac{1}{2c^2}v_L^2 - \frac{1}{2c^2}\frac{GM}{r_L}\right) dt \quad (11)$$

$$dt' \simeq \left(1 - \frac{1}{c^2}\frac{GM}{R_e}\right) dt \quad (12)$$

To the order $1/c^2$:

$$dt_L \simeq \left(1 + \frac{1}{c^2}\frac{GM}{R_e} - \frac{1}{2c^2}v_L^2 - \frac{1}{c^2}\frac{GM}{R_L}\right) dt' \quad (13)$$

This is in the order of 2.9×10^{-10} for an altitude like that of the Space Station. It corresponds to a situation in which a clock on the LEO loses about 9 milliseconds per year while the GPS transmitter clock would gain about 14 milliseconds per year. Expression (13) represents a simplified version of expression (8) and indicates the approximated relation between a clock on the LEO and a clock on the Earth surface.

It is important to notice that the effect varies along the orbit according to altitude and velocity of the LEO. This orbit decays for air drag. It is recalled that 10 ns of offset correspond to an error of 3 m in position.

Expression (13) is relevant for LEO satellites with a clock, whose stability is better than some units$\times 10^{-10}$.

4.3 Synchronization with GPS signal

The Interface Control Document for GPS (ICD–GPS–200) does not consider the necessary relativity corrections for a moving GPS receiver. If one takes the GPS signals as transmitted and takes into account the motion of the LEO during the propagation delay of the signal, the time derived from the received GPS signal has to be corrected in the amount :

$$\Delta t = \frac{\vec{v_L}}{c^2}(\vec{r_L} - \vec{r_G}) \quad (14)$$

where :

$\vec{v_L}$ = LEO satellite velocity in ECI coordinates,

$\vec{r_L}$ = LEO satellite position in ECI coordinates,

$\vec{r_G}$ = specific GPS satellite vehicle position at the instant of transmission in ECI coordinates.

Expression (14) is relevant for LEO satellites which synchronize directly from GPS satellites signals.

4.4 Accelerated frame of reference

Transforming to an accelerated frame of reference, such as one fixed on the rotating Earth, one could make use of the Nelson transformation (Nelson, 1987) which involves introducing a coordinate time differing from GPS time. An attempt in this direction has been performed by Deines (1990), but the conclusions are erroneous because of a misinterpretation of the relationship between such a local coordinate time and GPS time (Ashby, 1990 c).

5 CONCLUSIONS

Relativistic corrections for a low earth orbit satellite have been outlined and described. The growing accuracy levels needed in positioning and time synchronisation will require a careful consideration of the relativity effects.

CONSTANTS AND QUANTITIES

c = speed of light, ... 2.99792458×10^8 m/s

J_2= Earth quadrupole moment coeff., ... 1.082×10^{-3}

GM= Earth mass $\times$ grav. const., 3.986005×10^{14} m^3/s^2

R_e= Earth radius at equator, ... 6.37814×10^6 m

ω_e=Earth angular speed, ... 7.292115×10^{-5} s^{-1}

$\frac{2\omega_e}{c^2}$=Sagnac correction area coeff., 1.6227×10^{-21} s/m^2

a_G = GPS satellite semi-major axis, ... 26.561×10^6 m

REFERENCES

Allan D.W. & Ashby N., 1986, Coordinate time in the vicinity of the Earth, *Relativity in celestial mechanics and astrometry, IAU Symposium n.114* , Leningrad, CCCP, J.Kovalesky & V.A.Bromberg Eds. , 299–313.

Allan D.W., Davis D.D., Weiss M., Granveaud M., Dorenwendt K., Hetzel P., Fischer B., Aoki S., Fujimoto M., Charron L. & Ashby N., 1984, Accuracy of international time and frequency comparison via the Global Positioning System in common view, *Conference on precision electromagnetic measurement* , Delft, Nederland, IEEE–IM–34(2), 118–125.

Allan D.W., Weiss M.A. & Ashby N., 1985, Around-the-world relativistic Sagnac experiment, *Science* , 228, 69–70.

Ashby N., 1988, *Relativistic effects in crosslink ranging in the Global Positioning System* , Final Report NBS 40RANB8B5161, 58 p.

Ashby N., 1990 a, *A tutorial on relativistic effects in the Global Positioning System* , NIST Contract No. 40RANB9B8112, 43 p.

Ashby N., 1990 b, *Effect of accelerations on clock in satellites in the Global Positioning System* , NIST Contract No. 40RANB9B8112 Part II, 15 p.

Ashby N., 1990 c, *Response to Steve Deines' paper entitled "Missing relativity terms in GPS"* , NIST Contract No. 40RANB0B7835, 8 p.

Ashby N. & Allan D.W., 1979, Practical implications of relativity for a global coordinate time scale, *Radio Science* , 14, 4, 649–669.

Ashby N. & Allan D.W., 1984, Coordinate time on and near the Earth, *Phys. Rev. Lett.* , 53, 19, 1858.

Czajko J., 1990, On the Hafele-Keating experiment, *Ann. der Phys.* , 47, 6, 517–518.

Deines S.D., 1990, Missing relativity effects in GPS, *GPS-90 Proceedings*, Institute of Navigation, Colorado Springs (Co), 138–142.

Eardley D., Dyson F., Horowitz P., Press W., Ruderman M., Shapiro & Treiman, 1985, *Relativistic effects in the Global Positioning System* , The Mitre Corporation, JSR–84–104, 14 p.

Hafele J.C. & Keating R.E., 1972, Around-the-world atomic clocks: predicted relativistic time gains, *Science*, 177, 4044, 166–167.

Hafele J.C. & Keating R.E, 1972, Around-the-world atomic clocks: observed relativistic time gains, *Science*, 177, 4044, 168–170.

Hansen P.M., 1981, Prospects for using the Navstar/GPS constellation as detector of gravitational radiation, *10th Texas Symposium on relativistic astrophysics* , Baltimore, Maryland, December 1980, 447 (only abstract).

Hansen P.M., Chiu Y.T. & Chao C.C., 1986, *Prospects for using the Navstar/GPS constellation as a detector of astrophysical gravitational radiation* , The Aerospace Corporation Report, ATR–86(8421)–2, 64 p.

Harkins M.D., 1979, The relativistic Doppler shift in satellite tracking, *Radio Science* , 14, 4, 671–675.

Jorgensen P.J., 1986, Relativity correction in GPS user equipment, *PLANS '86, Position Location and Navigation Symposium* , Las Vegas, Nevada, November 4–7 1986, New York Institute of Electrical and Electronic Engineering, 177–183.

Jorgensen P.J., 1989, Special relativity and intersatellite tracking, *Navigation: Journal of the Institute of Navigation* , 35, 4, 429–442.

Lear W.M., 1989 a, *GPS navigation for Low Earth Orbiting vehicles* , NASA Lyndon B. Space Center, Houston, Texas, 87–FM–2, Rev. 1, JSC–32031, 283 p.

Lear W.M., 1989 b, *An introduction to special relativity* , NASA Lyndon B. Johnson Space Center, Houston, Texas, 87–FM–20, Rev. 1, JSC–22813, 119 p.

Matchett G.M., Vogel M.A. & Macdonald T.J., 1980, *Space Shuttle navigation analysis, Final report. Vol.1-GPS aided Navigation* , The Analytical Sciences Corporation, NASA–CR–160765, TR–1001–1–1, Contract No. NAS 9–15151, 14 p.

Nelson R.A., 1987, Generalized Lorentz transformation for an accelerated rotating frame of reference, *Journal of Mathematical Physics* , 28, 10, 2379.

Mc Caskill T., White J., Stebbins S. & Buisson J., 1978, NTS–2 Frequency stabilisation results, *Proceedings of the 32nd. Frequency control Symposium* , Atlantic City, N.J., 560-566

Post E.J., 1967, Sagnac effect, *Rev. of Mod. Phys.* , 39, 2, 475–493.

Rockwell International Corporation, Interface Control Document. *Navstar GPS space segment navigation user interfaces (1987)* , ICD–GPS–200, Rev. B, Rockwell International Corporation Space Operations and Satellite Systems Division, 115 p.

Sagnac G., 1913, *Compt. rend.* 157, 708, 1410.

Sagnac G., 1914, *J. Phys. Radium* 5th. Ser. 4, 177.

Shapiro I.I., 1964, Fourth test of general relativity, *Phys. Rev. Lett.* , 13, 789–791.

Skalafuris A.J., 1985, Current theoretical attempts toward synchronisation of a global satellite network, *Radio Science* , 20, 6, 1529-1536.

Thorne K.S., 1987, Gravitational radiation, *300 years of gravitation* , S.W.Hawking & W.Israel Eds. Cambridge University Press, 330–446.

Van Dierendonck A.J., Russel S.S., Kopitzke E.R. & Birnbaum M., 1980, The GPS navigation message, *Global Positioning System* , Vol.1, 55–73, Institute of Navigation Publ. , USA, ISBN 0–936406–00–3.

Vessot R.F.C., 1979, Relativity experiments with clocks, *Radio Science* , 14, 4, 629–647.

Will C.M., 1981, *Theory and experiment in gravitational physics* , Cambridge University Press, Cambridge, England, 1981, 341 p.

Will C.M., 1986, *Accuracy of time transfer in satellite systems* , National Academy of Science for USAF, National Research Council, Washington, Committee on accuracy of time transfer in satellite systems, AD–A176579, Contract No. F49620–85, 85 p.

THE COLUMBUS METROLOGY SCIENCE TEAM: REPORT ON ACTIVITIES

A.D.A.M. Spallicci

European Space Research and Technology Center, ESTEC,
Noordwijk, the Netherlands

ABSTRACT

The Columbus Attached Pressurised laboratory of the Freedom Space Station, the Columbus Free Flyer Laboratory , EURECA and Spacelab might offer the potential for experiments in fundamental metrology. In order to identify the specific opportunities offered by these microgravity carriers, a Columbus Metrology Science team has been sponsored by ESA. The activities of the Columbus Metrology Science team are summarised below.

1 ESTABLISHMENT OF THE TEAM

For the Columbus program the potential role for sciences like fundamental metrology has been considered (Spallicci, 1991 — also to appear on I Fairbank Meeting Proceedings — see numerous references quoted therein). A team of scientists has been established (A. Brillet–CNRS, Orsay; G. Busca–Neuchatel; F. Fuligni–IAS/CNR; A. Nobili–Pisa U.; I. Roxburgh–QMW/London U., chairman; A. Spallicci–ESTEC, secretary). The team has concentrated mainly on gravitation science experiments, but proposals and ideas in other fundamental aspects of metrology are invited.

2 MECHANICAL EXPERIMENTS

Often gravitation science experiments require acceleration noise levels below 10^{-12} g. Since the Columbus elements do not plan to offer this environment, investigators have proposed noise attenuators. The Pico Gravity Box (PGB) by A. Nobili, (K. Knott Ed. et al., 1991 a) is a passive attenuator aimed at a noise reduction below 10^{-12} g. It is a technological experiment which includes test of accelerometers. PGB would possibly become the technical facility for scientific experiments, like :

- GRAVCON (consisting of rotating masses and accelerometers) to measure the constant of gravitation to an accuracy of 10^{-5} , by I. Roxburgh (K. Knott Ed. et al., 1991 a);
- Weak Equivalence Principle experiment by A. Nobili (to appear).

3 CLOCKS

A proposal regarding Maser clocks to fly on board EURECA has been formulated by G. Busca and it is described in these proceedings. Efforts are under way to define possible relativity measurements with this technological experiment (K. Knott Ed. et al, 1991 b).

4 CONCLUSIONS

The team has so far identified four experiments, of both scientific and technological nature.

In the following months a report shall be produced. It is stressed that European scientists are invited to put forward ideas and proposals at any time. It is also under consideration that a questionnaire be sent to the scientific community aimed at an understanding of potential Metrology Users Mission Requirements for Columbus, in view of the growing interest for the Columbus program by the Metrology and Fundamental Physics science communities.

REFERENCES

Knott K., Roxburgh I. & Spallicci A., 1991 a, *Proc. of the Columbus Metrology Science Team* , 2–3 September 1991, ESTEC .

Knott K., Roxburgh I. & Spallicci A., 1991 b, *Proc. of the Columbus Metrology Science Team* , 26 October 1991, ESTEC.

Spallicci A.D.A.M., 1991, Experimental gravitation and the Columbus program, *Proc. of IX Italian Conf. on General Relativity and Gravitation physics* , R. Cianci, R. de Ritis, M. Francaviglia, G. Marmo, C. Rubano & P. Scudellaro Eds., 25-28 September 1990, Capri (Napoli, I), World Scientific, 611-614.

Spallicci A.D.A.M., The Columbus program and gravity science, *Proc. of I Fairbank meeting on Relativistic Gravitational Experiments in Space (Advanced Series of Astrophysics and Cosmology)* , C.F.W. Everitt & R. Ruffini Eds., 10-14 September 1990, Roma (I), (accepted for publication, to appear).

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

PROSPECTS OF TWO-WAY INTERNATIONAL TIME SCALE COMPARISONS

Gerrit de Jong

Van Swinden Laboratorium
P.O.Box 654
2600 AR Delft, Netherlands

ABSTRACT

The Two-Way Time Comparison technique is introduced, and the effects of non-reciprocity are discussed. The obtained results so far are reported. The support and activities of international bodies is described and prospects to use a satellite for both regional use and transatlantic use in N. America and W. Europe are given.

Key words: international time comparison, two-way time transfer, precise satellite time.

1. INTRODUCTION

The time indicated by clocks is called a time scale. Today the best, that is the most uniform and accurate, time scales are produced by caesium atomic clocks. The unit of time, the second, is defined on the quantum mechanical properties of atoms of Cs-133. Its long term stability has not been surpassed yet and for this reason caesium clocks are used by the national timing centres for time scales. These clocks can contribute to the formation of the International Atomic Time (TAI) if a good quality comparison method to the BIPM is provided.

To find out how good time scales are, we have to compare them. Therefore the clocks are co-located and the signals of the clocks are led to the start and stop input of a time difference meter, generally called a time interval counter (TIC). State of the art TIC's can measure the time difference to an uncertainty of about 10 ps (pico seconds).

If the clocks are not co-located, the signal of the remote clock is transported to the TIC by means of coaxial cables, optical fibre cables, radio links or optical links. The delays in these links have to be known to the desired uncertainty and have to be corrected for to be able to determine the time difference of the two clocks. This delay calibration in the desired accuracy regions of 10 ns for state of the art cesium clocks is a great problem, especially for radio links of which the delays are changing with distance, ionosphere, troposphere, air humidity and air density, temperature, earth conductivity and so on.

To cancel these influences in first order the two way scheme has been introduced: at both clock sites the time signals are transmitted at the same instant and on both sides the signal from the other clock is received and measured. After the exchange of the measured data, the difference of the two clocks is calculated, fortunately the delays cancel due to the in first order complete reciprocity of the signal paths. The accuracy of the result depends on the knowledge of and correction capability for the departure from the assumption of complete reciprocity (refs. 4 & 6)

2. TWO-WAY TIME COMPARISON CALCULATION

From fig. 1 it can be seen how the difference of the clocks at station 1 and 2 can be determined.

TA(k) is the time scale at station k
TI(k) is time interval reading
TT(k) is transmitter delay
TR(k) is receiver delay
TU(k) is uplink propagation delay
TD(k) is downlink propagation delay
TS(k) is satellite delay
TCU(k) is the Sagnac correction in the uplink
TCD(k) is the Sagnac correction in the downlink

The difference of the time scale (clock) at station 2 from the time scale (clock) at station 1 expressed by TA(1) -TA(2) is determined as follows.

The TIC reading at station 1 is:
$TI(1) = TA(1) - TA(2) + TT(2) + TU(2) + TCU(2) + TS(2) + TD(1) + TCD(1) + TR(1)$
The TIC reading at station 2 is:
$TI(2) = TA(2) - TA(1) + TT(1) + TU(1) + TCU(1) + TS(1) + TD(2) + TCD(2) + TR(2)$
Subtracting both expressions gives:
$TI(1) - TI(2) = 2\cdot TA(1) - 2\cdot TA(2) + TT(2) - TT(1) + TU(2) - TU(1) + TS(2) - TS(1) + TD(1) - TD(2) + TR(1) - TR(2) + TCD(1) - TCU(1) - TCD(2) + TCU(2)$
or the timescale difference is:

$$TA(1) - TA(2) = \tfrac{1}{2}\{TI(1) - TI(2)\} \text{ (=TIC diff.)} + \tfrac{1}{2}\{TS(1) - TS(2)\} \text{ (= Satellite delay diff.)} + \tfrac{1}{2}\{TU(1) - TD(1)\} \text{ (= Up/Down difference at 1)} - \tfrac{1}{2}\{TU(2) - TD(2)\} \text{ (= Up/Down difference at 2)} + \tfrac{1}{2}\{TT(1) - TR(1)\} \text{ (=Transmit/Receive diff. at 1)} - \tfrac{1}{2}\{TT(2) - TR(2)\} \text{ (=Transmit/Receive diff. at 2)} - \tfrac{1}{2}\{TCD(1) - TCU(1) - TCD(2) + TCU(2)\} \text{ (=Sagnac correction TC(12))} \quad [1]$$

The last 6 expressions are the corrections for non-reciprocity and are further discussed in section 3.

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

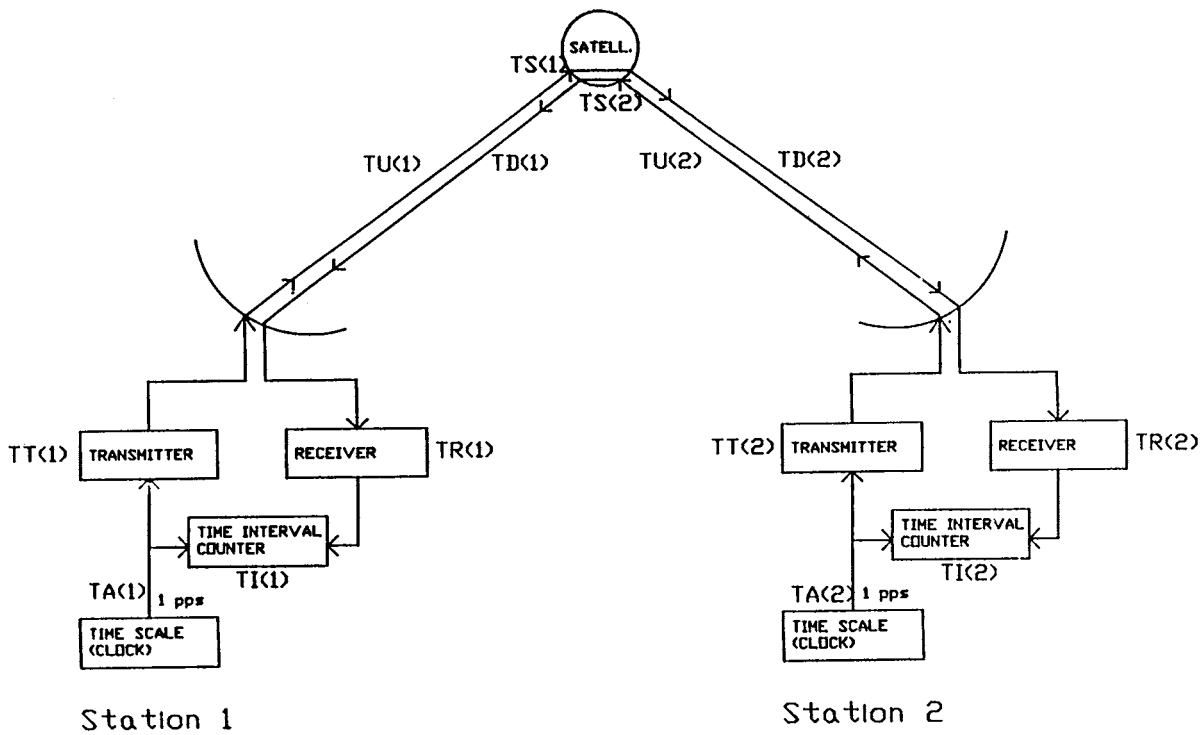


Fig. 1 Two-way time scale comparison principle.

3. DELAY DIFFERENCES FROM NON-RECIPROCITY

3.1 Satellite delay difference

If in the satellite the same frequency, the same receive antenna, transponder channel and transmit antenna is used, then the delays TS(1) = TS(2). This is not the case when different frequencies, transponders or different spot beams are used for the reception and/or transmissions from each station, i.e. the transatlantic Intelsat satellites. In this case TS(1) and TS(2) or at least the difference TS(1) - TS(2) should be measured before the launch of the satellite. Or otherwise estimated from detailed drawings of the satellite transponders and its wiring. Another problem is that the antennas for the two spot beams are not co-located, there can be a separation of several meters, and its time difference effect may vary depending on the satellite orientation with respect to station 1 and 2. Every meter difference in separation gives up to 3 ns non-reciprocity error.

If the signals from both stations cross each other at the satellite at the same moment, or at least the satellite is not moving in the time between the signals pass the satellite, then the up and down path lengths are equal. However, if the satellite is moving from the earth centre at a (Doppler) range rate of 3 m/s, then the signals experience the double value of 6 m/s, so if i.e. the two signals pass the satellite within 0.01 s then this non-reciprocity correction amounts up to $\frac{1}{2}(6 / (c \cdot 0.01)) = 100$ ps, where c = speed of light.

3.2 Propagation link delay difference

3.2.1 Ionosphere

The up and down link signals at each station differ in carrier frequency and they experience a different ionospheric delay equal to:

$$40.3 \cdot TEC \cdot (1/c) \cdot (1/fd^2 - 1/fu^2) \qquad [2]$$

where TEC = Total Electron Content; c = speed of light; fd and fu are the down and up link frequencies (ref 14).
For TEC = $1 \cdot 10^{18}$ electrons/m^2 and for fu = 14.5 and fd = 12.5 GHz this ionospheric delay = 0.859 ns - 0.639 ns = 0.220 ns. So the correction for $\frac{1}{2}\{TU(k)-TD(k)\}$ is typically smaller than -0.11 ns.

3.2.2 Troposphere

The troposphere gives a delay depending on the elevation angle, the water content of the air, air density and temperature, but this delay is not frequency dependent and so its influence on the up and down propagation delays is equal and no correction is needed (ref. 14).

3.2.3 Relativistic (Sagnac) effects

Due to the movement (around the rotation axis of the earth) of both the earth stations and the satellite during the propagation of a time signal to and from the satellite, a correction has to be applied to the propagation time of the signal. This Sagnac correction (ref. 15) amounts for one way between satellite s and station k:

$$TCD(k) = (\Omega/c^2) \cdot (Y(k) \cdot X(s) - X(k) \cdot Y(s)) \qquad [3]$$

where:
Ω= earth rotation rate = $7.2921 \cdot 10^{-5}$ rad/s
c= speed of light = 299 792 452 m/s
r= earth radius = 6367 000 m
R= satellite orbit radius = 42 150 000 m
X(k)= geocentric x-coordinate of station
X(s)= geocentric x-coordinate of satellite
Y(k)= geocentric y-coordinate of station
Y(s)= geocentric y-coordinate of satellite
X(k) = r cos(LA(k)) · cos(LO(k)) m
X(s) = R cos(LA(s)) · cos(LO(s)) m
Y(k) = r cos(LA(k)) · sin(LO(k)) m
Y(s) = R cos(LA(s)) · sin(LO(s)) m
LA()= latitude
LO()= longitude.
For geostationary satellites LA(s)= 0,

so TCD(k)=
$(\Omega/c^2) \cdot R \cdot r \cdot \cos(LA(k)) \cdot \sin(LO(k) - LO(s))$ [4]

Total Sagnac correction TC(12) for clock at station 1 as reference to measure clock at station 2 is:
TC(12)= ½(TCU(1) + TCD(2) - (TCU(2) + TCD(1)))
Further is TCU(k)= -TCD(k), so
TC(12)= -TCD(1) + TCD(2) [5]

Examples for satellite at -53° East:
LA(VSL)= 52° N, LO(VSL)= 4° E, LO(sat)= -53° E, difference in LO = 57°, TCD(VSL)= +112.42 ns
LA(USNO)=39° N, LO(USNO)= -77° E, LO(sat)= -53° E, difference in LO = -24°, TCD(USNO) = -68.83 ns
TC(VSL->USNO): -TCD(VSL) + TCD(USNO)= -181.25 ns
TC(USNO->VSL): -TCD(USNO) + TCD(VSL)= +181.25 ns

3.3 Station delay difference

The difference of the transmit and receive section including the up and down convertors, modulator and demodulator (modem), feeds, wiring, etc. has to be determined at each station. Methods to obtain this have been:
- co-location of both stations, or
- subsequent co-location of a third (transportable) earth station at both stations.
- using a satellite-simulator + calibrated cable.

The last method (ref. 2) is the least expensive and

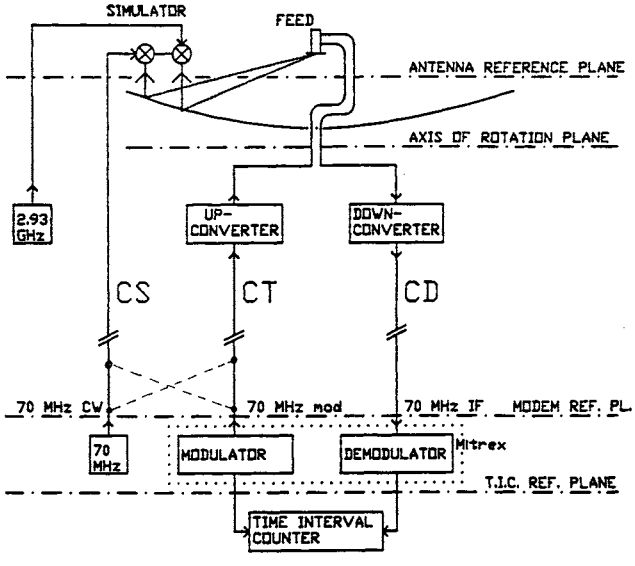


Fig. 2. Earth station calibration method.

can be used frequently. This method (see fig. 2) consists of the calibration of a auxiliary cable CS, measurement of the sum of the transmit and receive delay, measurement of the sum of the auxiliary cable delay and the receive delay and calculation of the receive and transmit delay from the measurements.
The internal transmit and receive delay of the modems have to be determined too. This can be done by co-locating the modems and measuring the sum of the transmit delay of one modem and the receive delay of the other. Another method is indicated in Ref. 2, this needs opening the modem and making a temporary modification.

4. BI-PHASE MODULATION AND DEMODULATION WITH PSEUDO NOISE CODES

An example of a two-way modem is the MITREX modem (ref.1). This has a modulator section (fig. 3) in

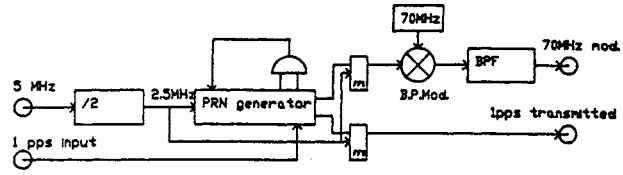


Fig. 3. MITREX modulator section

which the 1 pps signal is bi-phase modulated on a carrier. The PRN generator is clocked at a 2.5 MHz rate and every 4 ms a 10000 bit sequence is generated. As soon as the 1 pps signal appears, then one sequence of 4 ms is inverted. The sequences are bi-phase modulated on a 70 MHz carrier frequency. The signal is band pass filtered and is fed as the IF signal to the up-converter and then transmitted to the satellite.

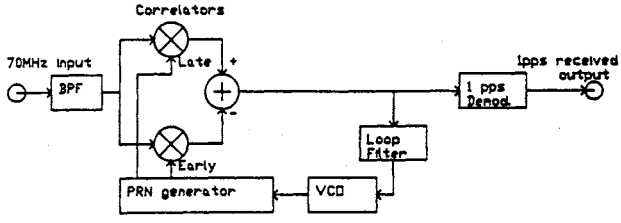


Fig. 4. MITREX demodulator section.

The demodulator section (fig.4) receives the 70 MHz down-converted signal, correlates it to a locally generated replica of the selected PRN code and provides a 1 pps received pulse synchronous to the detected 4 ms inverted PRN sequence.

5. RESULTS

The two-way method using microwave carrier frequencies and satellites has been used already in 1962 on the Telstar satellite between England and the USA (ref. 9). The use of bi-phase modulation with PRN code has been reported in 1974 and the use of a dedicated modem (Mitrex) for high precision transatlantic time transfer has been reported in 1983 giving 0.5 ns precision (ref. 1).
In the German NAVEX navigation and timing experiment flown on board of the Space Shuttle in 1985 a space-born caesium clock was compared with earth-based clocks using a two-way scheme with bi-phase PRN coded signals; a total systematic error of 25 ns and a precision of 1 to 7 ns was reported (ref. 7).
Since Aug. 1987 the time scales of USNO, NIST and NRC are routinely compared using the two-way Mitrex method 3 times per week on a US domestic satellite. Results show a precision of ≈1 ns at 1 s averaging time.
The most recent result (ref. 3,5) has been reported, using a EUTELSAT ECS satellite for two-way time transfer between France and Austria during LASSO experiments and also common view GPS measurements were taken in 1990/91. This excellent work, including careful calibration, has shown a difference between two-way and GPS of 3 ns with a precision of 3 ns.

6. INTERNATIONAL SUPPORT AND PROSPECTS

The remarkable results with the Mitrex modem since 1983 have been recognized by international bodies that deal with accurate time scales and its dissemination.
The Comité Consultatif pour la Définition de la Seconde (CCDS) made a recommendation in 1985 which was confirmed in 1987 by the 18th Conférence Générale des Poids et de Mesures (CGPM) to implement two-way satellite links using PRN-coded signals for the transfer of International Atomic Time (ref. 16). In 1989 the Bureau International de Poids et Mesures (BIPM) was asked by the CCDS to establish ad-hoc working groups to coordinate activities on such time links (ref. 10). The task of these ad-hoc working groups includes to recommend about:
- what satellites and frequency bands
- earth station specifications
- calibration of stations
- measuring procedures and schedules
- data processing.

The idea was to establish regional time links with two way on regional satellites, i.e. in the US and Canada, in Western Europe and Asia and to establish a few interregional time links to connect the regions. This method would give an alternative for the common view GPS method (which is affected by measures taken by the military GPS-owners) as well as an improvement of the precision and accuracy of the time links. In Europe such an ad-hoc working-group has first met in June 1989 at VSL in Delft, Netherlands. In this group a recommendation was made to use in Europe the 12.50 to 12.75 GHz down link and the 14.0 - 14.5 GHz up link on i.e. ECS or Intelsat satellites.

In the International Telecommunication Union (ITU), the Comité Consultatif International de Radio (CCIR) Study Group 7 has issued an urgent question to study two-way time transfer (ref. 11) and for this purpose a Task Group 7/2 has been established, which the author has the privilege to chair. Its task will be to study two-way time transfer and to draft Recommendations about: long-term stability, time accuracy, frequency comparison capability, performance compared to other methods, causes and cures of systematic delay variations, and standard data exchange format for measurement results.

In the European Cooperation in Standards and Metrology Organization EUROMET, in the field of Time and Frequency, a project coordinated by VSL concerning the use of two-way has been agreed (ref. 12); in several European countries steps have been taken to implement two-way at the national standards institutes for time, i.e. Austria (BEV, TUG), Germany (PTB, FTZ), United Kingdom (NPL), Netherlands (VSL), France (LPTF, OCA), Italy (IEN) (see also fig. 5).

EUROMET

Project 93

Satellite Time and Frequency intercomparison using PRN codes (Two-Way method)

Participants:

NL (VSL - coord.)	G. de Jong
AT (BEV, TUG)	D. Kirchner
DE (PTB, FTZ)	P. Hetzel, A. Söring
FR (LPTF)	M. Granveaud, P.Uhrich
GB (NPL)	D. Knight
IT (IEN)	F. Cordara

- get earth stations ready
- select satellite, get transponder time
- measurement schedule

Fig. 5. EUROMET Project 93 participants

The use of a satellite usable for transatlantic as well as regional North America and regional Western Europe has been suggested by Veenstra of INTELSAT (ref.8). The footprints are given in figs. 6 and 7, and characteristics in fig 8. This suitable satellite has 14/12 GHz full interconnectivity transponders, so the transmitted signal can be received back by each station. A complication is the fact that for the two (back and forth) transatlantic signal paths different antennas and transponders in the satellite have to be used, which gives rise to a non-reciprocity error. The elevation angle for this satellite at 53° West from Europe is relative low, but even an elevation of 6° from TUG (Austria) has been successful.

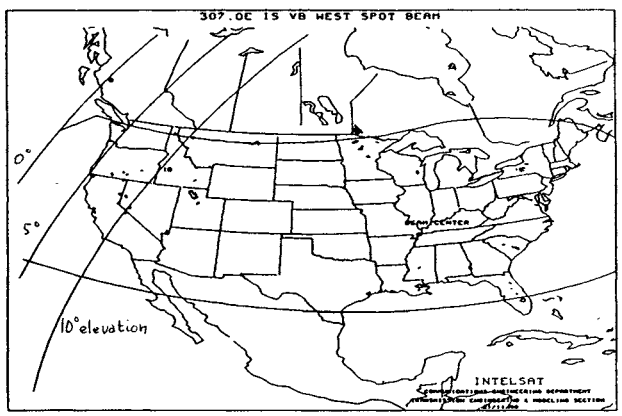


Fig. 6. Footprint West spot INTELSAT V at 53°W

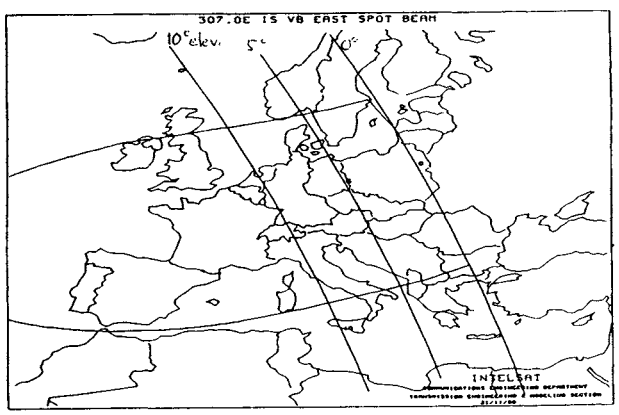


Fig. 7. Footprint East spot INTELSAT V at 53°W

Summarizing it can be said that in many major timing centres preparations are taken to implement the two-way time comparison technique.
Some problems are the cost and delivery time of earth stations, the availability of reliable, stable modems and the procedures to get the necessary permissions for the transmitting part of the earth stations. We expect that the intensive cooperation in the above mentioned organizations will facilitate the use of two-way time comparisons.

Transatlantic link
INTELSAT V-A(F13) at 307 E (53 W)

Stations: NIST, USNO, (TUG, VSL, PTB, NPL, LPTF, NRC,..?)

Occasional Use / Full Connectivity

West Spot (USA) Uplink 14.09216 GHz
West Spot Downlink 11.79716 GHz

East Spot (W-Eur) Uplink 14.09216 GHz
East Spot Downlink 12.59716 GHz

Useful for transatlantic Time Link for BIPM
International Atomic Time (TAI)
and for regional links in USA and W.-Europe

Requests to National Intelsat Signatory for approval

Fig. 8. Main characteristics of a combined transatlantic and regional two-way link.

7. CONCLUSION

Very promising results with Two-way Time Transfer have been obtained. The use of small earth stations at the premises of the timing centres is now achievable. Suitable satellites are present. In the timing community the realization of a long-time transatlantic link between NIST, USNO, NRC and European labs as well as a mutual link in Western Europe can be realized. To actually obtain the inherent performance of the two-way time transfer much calibration work has to be done. Especially efforts should be done to obtain the promising accuracy of 3 ns over long-term (>1 year).

8. REFERENCES

1. Hartl P et al 1983, Spread Spectrum time transfer experiment via INTELSAT, Proc 15th Annual Precise Time and Time Interval Applications and Planning Meeting, Washington, pp. 331-354.

2. Jong, G de 1989, Accurate delay calibration of satellite ground stations for two-way time transfer, Proc. 3rd European Frequency and Time Forum, pp 198-203

3. Kirchner D et al 1991, Comparison of two-way satellite time transfer and GPS common-view time transfer between OCA and TUG, Proceedings PTTI, Pasadena, Dec. 1991, in press.

4. Kirchner D 1991, Two-Way Time Transfer Via Communication Satellites, Proceedings of the IEEE, Vol 79, No. 7, 983-990.

5. Uhrich P et al 1991, Preliminary Comparison of Time Transfers via LASSO, GPS and Two-Way Satellite, Proc. 5th European Frequency and Time Forum, Besançon, pp. 96-104

6. Hanson D W 1989, Fundamentals of two-way time transfers by satellite, 43rd Annual Symposium on Frequency Control, pp. 174-178.

7. Starker S et al 1990, Navigation experiment NAVEX on spacelab mission D1, DLR, Oberpfaffenhofen p 139

8. Veenstra L B 1990, International two-way satellite time transfer using INTELSAT space segment and small earth stations, Proc. 22nd annual PTTI Meeting, pp. 383-400.

9. Steele J McA et al 1964, Telstar time synchronization, IEEE Trans. on Instr. and Meas., Vol. IM 13, No. 4, pp 164-170.

10. CCDS 1989, Report of the 11th Meeting of CCDS, BIPM, Pavillon de Breteuil, F-92312 Sèvres Cedex, pp S56, 57, 66.

11. CCIR 1991, Documents Period 1990-1994, Working Party 7A, Geneve, Doc. No. 7/9-E.

12. EUROMET 1988, Satellite Time and Frequency intercomparison using PRN codes, Agreed Project 93, EUROMET, Field of Time and Frequency, PTB, W-3300 Braunschweig.

13. Veenstra L et al 1981, Frequency and Time coordination via satellite, Comsat Technical Review, Vol. 11, No. 2, pp. 369-402.

14. Spilker JJ 1977, Digital Communications by Satellite, Englewood Cliffs NJ,Prentice Hall, pp. 493-608.

15. Ashby N & Allan D 1979, Practical implications of relativity for a global coordinate time scale, Radio Science, Vol.14, No. 4, pp. 649-669.

16. CGPM 1987, 18th Conférence Générale des Poids et de Mesures, Comptes Rendus, Resolution 4, Paris, BIPM.

Session 3B

Atomic Frequency Standards: Optically Pumped Cesium

Chairman: L. Cutler

PRELIMINARY RESULTS ON AN OPTICALLY PUMPED Cs BEAM FREQUENCY STANDARD

P. Aynié+, J.P. Berthet, S. Bouzid, A. Clairon, E. de Clercq, B. Dahmani, A.H. Gérard, P. Leclerc, M.Lours, G.D. Rovera and G. Santarelli

Laboratoire Primaire du Temps et des Fréquences
+ SERT
Bureau National de Métrologie et Observatoire de Paris
61, avenue de l'Observatoire
75014 Paris - France

ABSTRACT

The construction of an optically pumped primary frequency standard has been undertaken in our laboratory a few years ago. This paper describes the main features of the device under development. Preliminary results and first frequency measurements are then reported.

Keywords : Frequency standard, Cs, optical pumping.

1. INTRODUCTION

The advent of laser diodes in the near infrared opened the way to the practical realization of optically pumped cesium clocks. In these clocks the deflecting magnets of classical Cs standards are replaced by optical interactions with laser beams. The expected advantages are no strong magnetic field near to the cavity region, a not related to atomic velocities retrace of the atomic trajectories when the beam is reversed, an increased signal to noise ratio and some freedom given by the possibility of changing the used optical transitions. Another interest is the realization of the second by a different way of the classical one, which could allow a check of possible systematic offsets in its realization.

Several laboratories in the world are working now on this new kind of standard (Ref. 1). To our knowledge only one optically pumped standard has been evaluated (Ref. 2). The construction of such a primary frequency standard has been undertaken in our laboratory a few years ago. This paper describes the main features of the device under development. First results are then reported.

2. DESCRIPTION OF THE DEVICE

2.1 The Cesium Tube

The atomic beam is ribbon shaped (3x8 mm^2), coming from one of the two ovens, one on each end. The atomic beam is parallel to the East-West direction. The beam is collimated by a multichannel array made of a stack of crinkled nickel foils. The overall oven to oven length is 2,3 m. The static magnetic field or C field is transverse in the vertical direction. It is produced by four current carrying rods. These rods are 2 m long and cover the cavity length plus the interaction regions with laser beams. Between each oven and the microwave cavity there are two optical interaction regions where laser beams can cross the atomic beam. The fluorescence light is collected and focused on a Si photodiode by means of a spherical mirror in front of two fast aspheric lenses. The C field rods, the cavity and the optical devices fit inside an aluminium support cylinder. It is surrounded by three coaxial mumetal cylinders for magnetic shielding purpose. The measured homogeneity of the C field in the transverse or longitudinal direction is better than 5.10^{-4}.

2.2 The Ramsey Cavity

The Ramsey Cavity is of the usual type for a transverse magnetic field, the waveguide is bent in the E plane. The RF power is fed in by means of a coaxial cable in order to avoid holes in the magnetic shields. The cavity is in copper with three parts : a central tee with an $41\lambda g/2$ electrical length and two end-parts with a $7\ \lambda g/4$ electrical length for each one. The drift length between the two arms is L = 1,01 m. The atomic beam pass through the arms at $\lambda g/2$ from the short circuits. The dimensions of the slots are 3x8 mm^2 with a 43 mm long cut-off with the same dimensions on each side of each arm. The balance of the electrical length of the two cavity arms was adjusted with an uncertainty of a few hundreth of mm. The asymmetry of the cavity is determined by the level of excitation of odd modes.
The cavity is tuned on the Cs frequency at 22,0°C. Its quality factor Qc is about 4300. The measured temperature sensitivity is -155

kHz/°C. The relative value is -1,7 10^{-5} in very good agreement with the dilatation coefficient of electrolytic copper. Then we use the cavity temperature in order to know the cavity detuning.

2.3 The optical set-up

We use a single laser set-up. It is an extended-cavity semiconductor laser (Ref. 3). The laser diode is a 850 nm, 150 mW single mode GaAlAs diode from SDL company without additional anti-reflection coating. As usual it is temperature regulated (Ref. 4) and needs a low noise current supply. The output light is collimated and directed to a Littrow-mounted diffraction grating which ends the extended cavity. The cavity length is about 10 cm. The wavelength is first selected by rotating the grating and tuning the temperature and the d.c. current. Then for fine frequency tuning and locking purpose the cavity length is adjusted by means of a Piezoelectric Ceramic. The FWHM linewidth of a beat-note between two identical devices is 70 KHz. The device reliability is good, the laser frequency can be kept locked for a week without troubles.

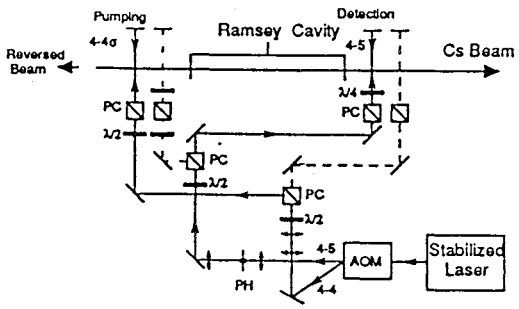


Fig. 1 - Optical Scheme

A.O.M. : acousto optic modulator, P.H. : Pinhole, P.C. polarizing cube

Fig. 1 shows the optical scheme. The 4-5 transition, which does not allow optical pumping, is used for detection of the microwave resonance. The optical pumping is performed with a σ polarized beam resonant with the 4-4 transition. In the following the laser beam tuned on the 4-5 transition is called detection beam, the 4-4 laser frequency beam is called pumping beam. We use a single laser, its frequency is locked on the 4-5 transition using the saturated absorption spectrum of a Cs cell. A part of the laser beam is frequency shifted on the 4-4 transition by means of an acousto-optic modulator. The detection beam can then be divided in two beams with a variable beam-splitter(halfwave plate and polarizer). A detection beam is then directed to the east end of the Cs beam, and the other to the west end. Before crossing the atomic beam the polarization is adjusted by polarizing components. And the same for the pumping beam. We use standing waves for pumping and detection. The perpendicularity of the laser beam and the atomic beam is adjusted to better than 1 mrad.

The east and west oven are continuously heated. Each beam can be cut by a shutter. The beam reversal is simply performed by rotating the shutters and acting on the variable beam splitters. The beam is reversed in a few minutes only. The two atomic beams can be used simultaneously.

3. FIRST RESULTS

3.1 Signal and interaction time distribution

Fig. 2 shows the whole microwave spectrum. The frequency scan is 500 KHz. The central Ramsey fringes (2 KHz scan), of the clock transition are shown on the Fig. 3. The signal-to-noise ratio in 1 Hz bandwidth is about 3000. The linewidth is 100 Hz in agreement with theoretical predictions based on the cavity length, oven temperature and a Maxwellian velocity distribution weighted by 1/v.

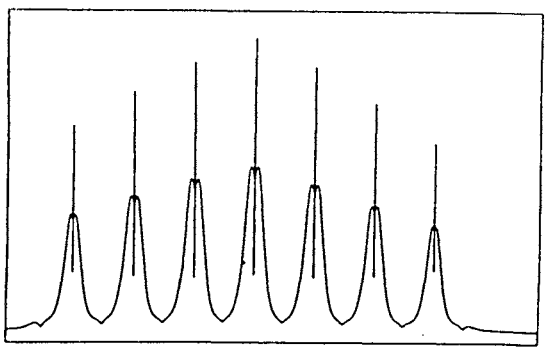


Fig. 2 - Whole microwave spectrum (500 KHz scan)

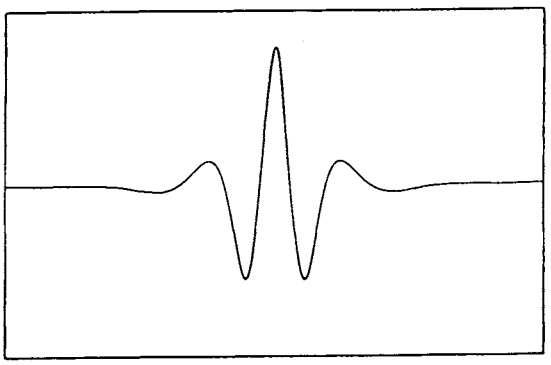


Fig. 3 - Central Ramsey Fringe
(2 KHz scan)

The experimental velocity distribution was measured by the Boulanger method (Ref. 5). We record the amplitude of the top of the Ramsey fringe versus the microwave power. The cosine Fourier transform of this curve is the time of flight distribution $f(\tau)$ with an unknown scale factor on τ, where τ is interaction time in one cavity arm (1 cm width). The scale factor is determined by fitting the calculated Ramsey fringes from the approximated interaction time with the experimental record. Fig. 4 shows the obtained interaction time distribution compared with a theoretical thermal weighted by τ distribution as expected (Ref. 6).

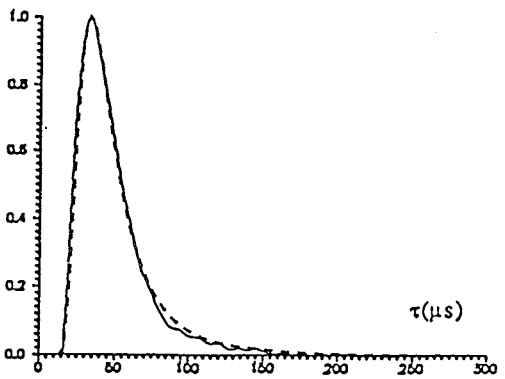


Fig. 4 - Interaction time distribution

Full line : experimental
dashed line : theoretical

3.2 Frequency Measurements

3.2.1 Experimental conditions

We use to probe the atomic resonance and to lock the quartz crystal frequency a slow square wave frequency modulation (2 Hz). The modulation width is 51 Hz very close to the halfwidth at halfmaximum of the peak to valley central Ramsey fringe at Popt. The optimum microwave power Popt is defined as the microwave power which maximizes the cw detected signal at center frequency. We used two totally different frequency synthesis and control systems, one home made and the other loaned by PTB (Ref. 7). The observed frequency difference between two electronics is less than 10^{-13}, limited by the measurement uncertainty. The frequency of our standard is measured against a H Maser from Sigma Tau Company. The measured short term stability is $\sigma y(\tau) = 2{,}2\ 10^{-12}/\sqrt{\tau}$ for $\tau > 100$s.

3.2.2 Main frequency shifts

The larger shift is caused by the C field effect, however it can be known with a very small uncertainty. The C field magnitude is determined from a measurement of the difference between the microwave frequencies of the clock transition and of the $(3,\pm 1) <-> (4, \pm 1)$ transitions. In our case, the relative shift is : $\Delta\upsilon/\upsilon = +\ 2{,}4\ 10^{-10}$ with an uncertainty below 10^{-14}.

The second order Doppler effect is given by the equation (Ref. 7) :

$$\left(\frac{\Delta\nu}{\nu}\right)_D = -\Lambda_D \frac{\langle v^2 \rangle}{2c^2} \quad (1)$$

Λ_D is a parameter depending on the velocity distribution and on the parameters of the modulation. $\Lambda_D = 0{,}93$ at Popt in our case. $\langle v2 \rangle$ is the average on the velocity distribution of the square of the atomic velocities and c in the speed of light. For a Maxwellian weighted by τ interaction time distribution at Popt :

$$\left(\frac{\Delta\nu}{\nu}\right)_D = -3{,}5\ 10^{-13}$$

The cavity pulling effect is (Ref. 8) :

$$\left(\frac{\Delta\nu}{\nu}\right)_C = \Lambda_C \left(\frac{Q_c}{Q_l}\right)^2 \frac{\Delta\nu_c}{\upsilon} \quad (2)$$

Q_c is the quality factor of the cavity, Q_l the cavity factor of the line. In the experimental conditions, the cavity temperature is 35,5°C and the cavity detuning $\Delta\nu_c = -2{,}1$ MHz $\Delta\nu_c$ is larger than the width of the cavity resonance. Then the slope of the cavity response is smaller than near resonance. Taking into

account this effect and the parameters of the modulation $\Delta\upsilon_c$ = 0,09 at Popt and $\left(\frac{\Delta\nu}{\nu}\right)_C$ = - 0,4 10^{-13} at Popt.

The end-to end phase shift can be written (Ref. 8) :

$$\left(\frac{\Delta\nu}{\nu}\right)_\phi = \wedge_\phi \frac{\Delta\phi}{\Pi Q_l} \quad (3)$$

$\Delta\phi$ is the phase difference between the two arms of the cavity. For our velocity distribution, our modulation parameters and at Popt $\Delta\phi$ = 1,17. The single method of measurement of phase shift is the beam reversal. At Popt $\left(\frac{\Delta\nu}{\nu}\right)_\phi$ = 11,1 10^{-13} . The calculated $\Delta\phi$ from eq. 3 is $\Delta\phi$ = 0,27 mrad. This value is larger than expected but not impossible. In order to check if it is well an end to end phase shift we have measured the frequency change with microwave power.

3.3 Microwave power shift

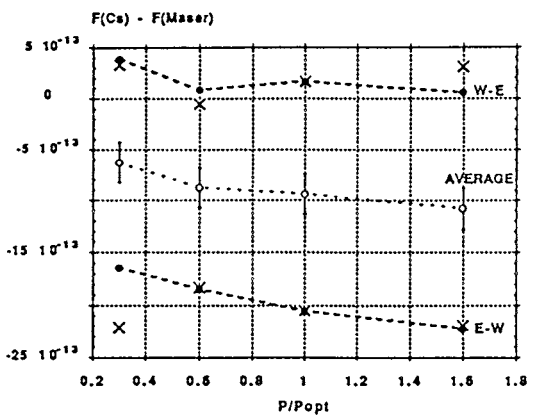


Fig. 5 - Microwave power shift

Fig. 5 shows the normalized power dependent frequency difference $\Delta\nu/\nu$. $\Delta\nu$ is the difference ν(Cs) - ν(Maser), ν(Cs) is corrected for taking into account the C field value, the second order Doppler shift and the cavity pulling shift. The last two effects are Microwave power dependent. W - E means that the beam direction is from West to East, and E - W is the opposite. Average is the averaged frequency over the two directions. Taking into account the uncertainty on the value of Popt, the uncertainty on the velocity distribution which was not measured at the time of these measurements and the frequency measurement uncertainty, the uncertainty on the average is of the order of $\pm$ 2.10^{-13}. The crosses are the calculated values for ν(Cs) = Average at Popt + $\Delta\nu(\Delta\phi,P)$ with $\Delta\phi$ calculated from the frequency difference between E - W and W - E at Popt. The agreement with experimental value is not good. Contrary to what is expected the frequency average is not power independent. This means that there is a microwave power dependent frequency shift which is not expressed by (1), (2) or (3).

4. CONCLUSION

This paper describes our laboratory optically pumped Cs beam frequency standard. The measured short term stability is $\sigma y(\tau)$ = 2,2 $10^{-12}/\tau^{-1/2}$ for τ > 100s. We have verified that the velocity distribution in the atomic beam is Maxwellian weighted by 1/v as expected. The main frequency shifts have been estimated. There is an unexpected power dependent frequency shift which needs further experiments to be explained.

Acknowledgment

We are especially grateful to Dr. A. Bauch and Dr. R. Schröder from the P.T.B. for the loan of a frequency control system. We wish also to thank Mr. G. Fréon from L.P.T.F. for his collaboration.

References

1. See for example papers in Frequency Standards and Metrology, Proceedings of the fourth Symposium, A. de Marchi editor, 1989, Springer Verlag.

2. Koga, Y., Ohshima, S., Nakadan, Y. and Ikegami, T., 1989, Performance of an optically pumped Cs frequency standard, electromagnetic Metrology, Proceedings of international symposium on electromagnetic metrology, ISEM 89, Pergamon Press.

3. De Labachelerie, M. and Cerez, P., 1985, An 850 nm semiconductor laser tunable over a 300 Å range, Opt. Comm., 55, 3, 174-178.

4. Rovera, G.D., De Marchi, A. and Santarelli, G., 1991, A simple temperature control for semiconductor lasers, Proc. 5th European Frequency and Time Forum, Besançon, France.

5. Boulanger, J.S., 1986, A new method for the determination of velocity distributions in Cs beam clocks, Metrologia, 23, 37-44.

6. Hamel, A., Giordano, V., Petit, P. and Audoin C., 1990, Velocity distribution measurement in an optically pumped Cs beam resonator, Proc. 4th european frequency and time forum, Neuchâtel, Switzerland, 169-174.

7. Schröder, R., 1991, Frequency synthesis in primary cesium clocks, Proc. 5th European Frequency and Time Forum, Besançon, France.

8. De Marchi, A., Rovera, G.D. and Premoli A., 1987, Effects of servo-loop modulation in atomic beam frequency standards employing a Ramsey cavity, IEEE Trans. on UFFC, 34, 6, 582-591.

MINIATURE OPTICALLY PUMPED CESIUM BEAM RESONATOR

P. Petit, V. Giordano, N. Dimarcq, P. Cerez, C. Audoin, G. Théobald

Laboratoire de l'Horloge Atomique Unité Propre de Recherche du CNRS associée à
l'Université Paris-Sud,Bât. 221, Université Paris Sud
91405 Orsay Cedex, FRANCE

Abstract

We have designed a miniature optically pumped cesium beam resonator. It is of the flop in type but with a phase difference of 180° between the two oscillating fields. In the present circumstances, this small machine operating at low oven temperature (85°C) exhibits a rather out standing short term frequency stability. More over, this ϕ = 180° cavity configuration makes it possible a lowering of a number of frequency shifts.

I - INTRODUCTION

In the laboratory, we have already developed and operated a high performance optically pumped cesium beam tube. This resonator called Cs3 has been primarily devoted to physical studies which have resulted in a good understanding of the light-atom interactions. High S/N ratio are always easily obtained leading to a short term frequency stability of $2.10^{-12}.\tau^{-1/2}$. Such a performance is obtained with a cavity length L = 21 cm and a distance between the oven to the detection zone equal to 55 cm. All the components of the resonator (at the exception of the pump) are contained in a cylinder of diameter 170 mm and length 800 mm. The corresponding volume is about 17 dm^3.

The filling factor of this large vacuum envelope was poor. We then decided to compact the physical package. Our challenge was to design a miniature atomic resonator exhibiting a short term frequency stability better then $10^{-11}.\tau^{-1/2}$ in a volume of a few liters.

In this paper we first discuss the choices which have been made for the design of this small optically pumped cesium beam resonator. A brief description of the machine called Cs4 is given. Finally preliminary results are presented.

II - Cs4 RESONATOR DESIGN

1) Vacuum chamber and cesium beam

The vacuum chamber is a parallelepiped (78 mm x 78 mm x 250 mm) of amagnetic stainless steel (*figure 1*). A 25 $l.s^{-1}$ ion pump external to the chamber maintains a residual pressure of the order of 10^{-5} Pa . The laser light penetrates in the chamber throught two sealed optical windows. After its interaction with the cesium beam, this laser light is efficiently absorbed by an internal light trap which is properly designed to reduce, as well as possible, both the stray light into the resonator and the optical feedback into the laser diode.

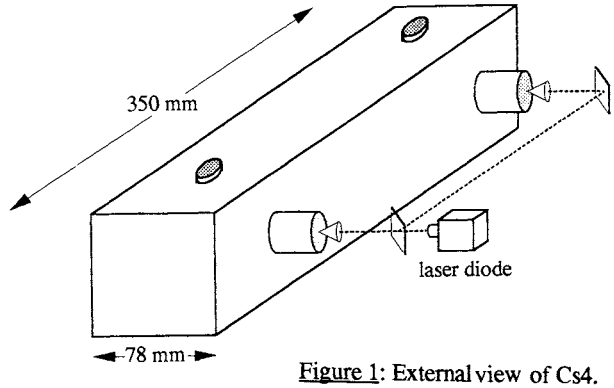


Figure 1: External view of Cs4.

The section of the emissive area of the oven is 2 mm x 4 mm with a transparency of 80 %.

At 85°C, this insures in the detection region having a cross section 3.9 mm x 7.8 mm an atomic flux of $\approx 3x10^{11}$ at/s.

2) Static magnetic field

In this machine, the magnetic field direction is perpendicular to the atomic beam axis. This geometry provides the best immunity to external magnetic field perturbation ΔB_e. Moreover, we have designed three independent magnetic field regions in order to avoid the detrimental non linear Hanle effect [1].

Two magnetic shields are placed into the vacuum envelope. A third magnetic shield can be added at the outside to insure that $\Delta B_e \simeq 2$ gauss should lead to a relative frequency shift smaller than 5×10^{-14} for a static magnetic induction B_0 of 50 mG.

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

3) Optical set up and microwave interaction

In the detection region, the light is collected and focussed on a large area photodiode which is placed outside of the vacuum chamber. The measured collection efficiency is 11 %.

The laser light comes from a commercial laser diode (Hitachi HLP 1400) tuned to the 3-3 σ optical transition: a light power of 5 mw is used for the pumping process. Less than 1 mw is enough for the optical detection.

A Ramsey type interrogation is used to induce the clock transition but in our compact resonator, the two oscillatory microwave magnetic fields are set 180° out of phase. As shown below, this configuration makes it possible some additional improvements of the frequency performances of the machine.

III - EXPERIMENTAL RESULTS

Figure 2 shows the Ramsey fringe delivered by the resonator as a function of the microwave frequency. The oven temperature was 85°C. The central fringe is 630 Hz wide. The peak to valley amplitude is 239 mV. The fringe is superimposed on a background mainly due to stray light.

Figure 3 recalls the shape of the calculated transition probability when the oscillating fields are 180° out of phase. At resonance, the transition probability is zero. However, this is not exactly the case in practice due to the overlap of the two neighbouring lines which appear on each side of the clock $\Delta m_F = 0$ line. This is due to the fact that the ground state levels are insufficiently separated when the static magnetic field is set around 50 mG. *Figure 4* shows the whole microwave spectrum recorded in such conditions. The symmetry of the spectrum is good and the Ramsey fringes are well centered on their Rabi pedestal even for large m_F value. This reflects directly the magnetic field uniformity along the atomic path in the whole microwave interaction region.

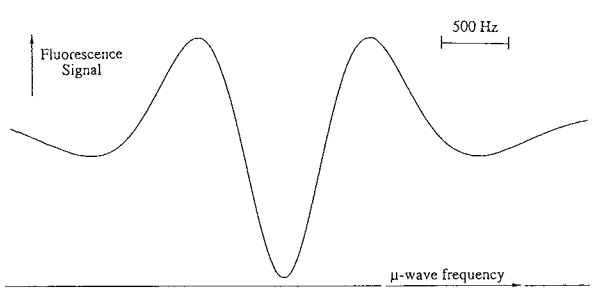


Figure 2: Experimental record of the Ramsey fringe.

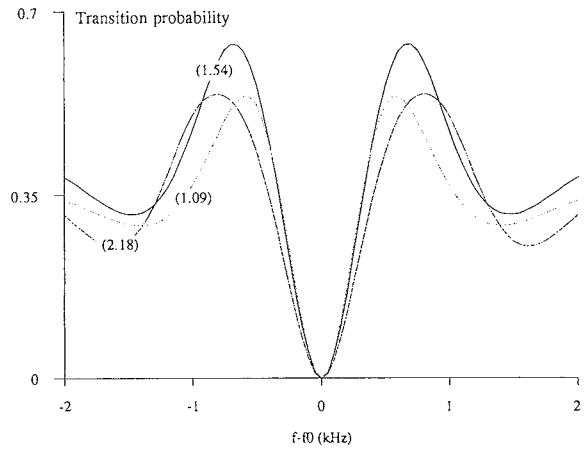


Figure 3: Calculated transition probability for Cs4 cavity. The parameter is $b\tau_0$.

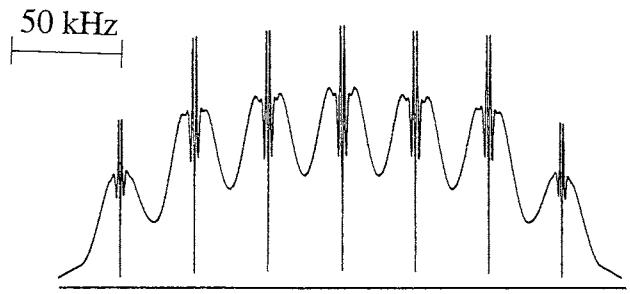


Figure 4: Experimental microwave spectrum.

IV - SHORT TERM FREQUENCY STABILITY

A good understanding of the advantages procured by the use of 180° out of phase microwave fields needs to take into account the modulation of the interrogating microwave signal [2], the noise level to be considered is determined by the mean value of the detected beam flux which depends on the frequency modulation scheme. We then have :

$$\sigma_y(\tau) = \left(\frac{FH}{2\,\omega_0^2\,T_0^2\,I_0} \right)^{1/2} \times \tau^{-1/2}$$

ω_0 is the angular frequency at resonance.

T_0 $= L/\alpha = 7.5\times10^{-4}$s; $\alpha = \left(\frac{2kT}{m}\right)^{1/2} = 212$ ms^{-1} at an oven temperature of 85°C.

I_0 is the atomic flux on the sublevel populated by the atomic preparation. In our experiment, the sublevel of interest is F = 4, $m_F = 0$ and $I_0 = 3.3\ 10^{10}$ at x s^{-1}.

F is the noise factor of the optical detection. In the present case F = 3.8.

H is a dimensionless parameter representing the effect of noise sources on the frequency stability. The best frequency stability is obtained when H is a minimum.

We consider two cases : the first one applies to conventional Ramsey cavities where $\phi = 0$. The corresponding H value is noted H^+. The second concerns cavities where $\phi = 180°$. H is referred as H^-.

In the case of slow sine wave frequency modulation we have :

$$H^{\pm} = 2\,\frac{2\,I'_b/I_0 + \int_0^{\infty} f(\tau)\,\sin^2 b\tau\,[1 \pm J_0(\omega_m T)]\,d\tau}{\left[\int_0^{\infty} (\tau/\tau_0)\,f(\tau)\,\sin^2 b\tau\,J_1(\omega_m T)\,d\tau\right]^2} \qquad (2)$$

I'_b is the fictitious atomic flux which takes into account the shot noise of the stray light and the detector noise. Our machine is characterized by $I'_b = 3.4\ 10^{10}$ at x s^{-1} .

$f(\tau)$ is the distribution function of the microwave interaction time τ. In the present time, it corresponds to a Mawwellian velocity distribution in the beam.

b is the Rabi angular frequency.

We have $T = (L/l)\tau$ and $\tau_0 = l/\alpha$.

ω_m is the modulation depth.

$J_{0,1}$ are Bessel functions of first and second order.

The numerator of H is proportional to the noise power which degrades the detection of the clock signal, whereas the denominator is proportional to the square of the slope of the frequency discriminator. Let us considered first the case $I'_b << I_0$. This means that the detection of the clock signal is mainly limited by the atomic shot noise. *Figure 5* shows a plot of $H^{\pm}$ as a function of the normalized amplitude of modulation $\omega_m T_0$.

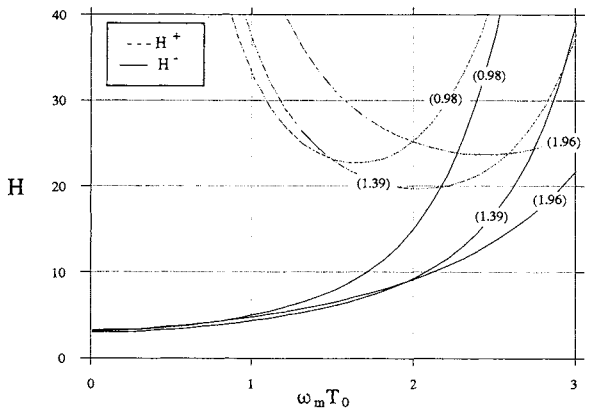


Figure 5: Variation of the H factor as a function of the normalized modulation depth in absence of background noise.

The parameter is $b\tau_0$. We may notice that H^+ exhibits a minimum value (H^+_{min}) for $\omega_m T_0 \simeq 2$ and $b\tau_0 = 1.39$. On the other hand, H^- is minimum (H^-_{min}) for $\omega_m T_0 = 0$ whatever $b\tau_0$ is. In that situation, when $\omega_m T_0$ vanishes, we have obviously a poor value of the slope of the frequency discriminator. But this detrimental effect for the clock servo system is balanced by a low noise level associated to the weak mean value of the detected atomic flux. In practice, values of $\omega_m T_0$ much smaller than unity can be used, without changing notably the value of (H^-_{min}). However that may be, H^+_{min} is always larger than H^-_{min} indicating that the short term frequency stability $\sigma_y^-(\tau)$ is better than $\sigma_y^+(\tau)$. In practice, I'_b may not be negligible compared to I_0. In that case, *figure 6* shows the variation of H^- as a function of the modulation depth. In this network, the parameter is I'_b/I_0. H^-_{min} increases rapidly with I'_b/I_0. The advantage $\sigma_y^- < \sigma_y^+$ still holds but is less pronounced than in the case $I'_b/I_0 << 1$.

In our optically pumped cesium beam resonator where $I'_b/I_0 \approx 1$ we have $H^-_{min} = 47$ (see *fig. 7*). The formula (2) gives finally $\sigma_y(\tau) = 1.2x\ 10^{-12}\ \tau^{-1/2}$.

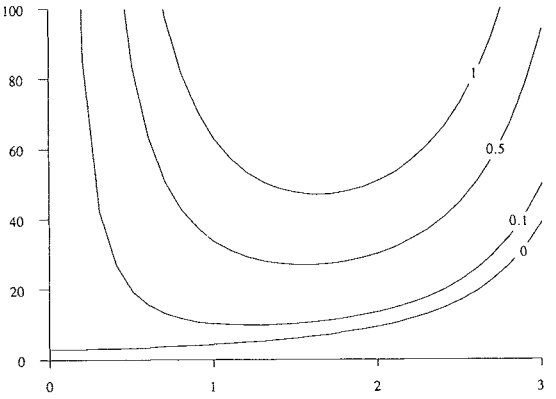


Figure 6: Variation of the H^- factor as a function of the normalized modulation depth in presence of background noise.

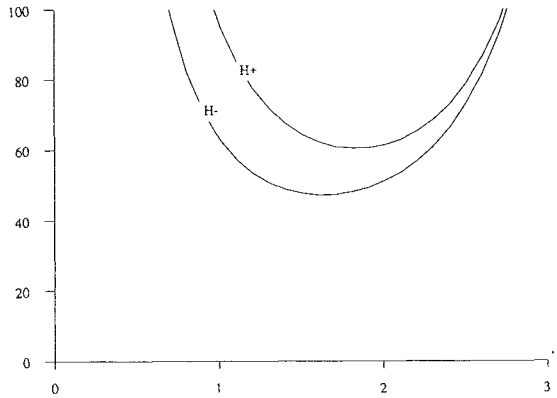


Figure 7: Comparison between H^+ and H^- when $I'_b/I_0=1$.

This prediction has been compared to experimental values. The frequency stability of one 10 MHz crystal oscillator has been measured when controlled by the cesium beam microwave resonance. The reference frequency source was the LHA hydrogen maser. We obtained :

$$\sigma_y(\tau) = 2.4 \times 10^{-12}\,\tau^{-1/2} \text{ for } 1 < \tau < 100 \text{ s.}$$

V - OTHER INTERESTING FEATURES OF THE UNCONVENTIONALMICROWAVE EXCITATION

The use of 180° out of phase microwave fields in the cavity has a number of additional advantages:

The first one concerns the cavity pulling frequency shift. When $\phi = 0$, the probability transition at resonance $P(\Omega_0)$ is dependent on the Rabi frequency b. Any detuning of the cavity distorts the variation of b around the atomic transition frequency, which induces a frequency shift.

When $\phi = 180°$, $P(\Omega_0)$ is always equal to zero whatever the value of b. Any cavity detuning distorts the line profile but without changing the resonant frequency. Actually, the microwave frequency is modulated and, for slow sine wave frequency modulation, the frequency offset due to cavity pulling is given by [2, chap. 5]

$$\delta\omega^{\pm} = (\omega_c - \omega_0)\left(\frac{T_c}{T_0}\right)^2 \Delta_c^{\pm} \qquad (3)$$

with $\Delta_c^{\pm} =$

$$\frac{\int_0^{\infty} \left[1 \pm J_0(\omega_m T) \mp J_2(\omega_m T)\right] \frac{\tau}{\tau_0} \sin b\tau \cos b\tau \, f(\tau)\, d\tau}{\omega_m T_0 b\tau_0 \int_0^{\infty} \frac{\tau}{\tau_0} J_1(\omega_m T) \sin^2 b\tau \, f(\tau)\, d\tau}$$

T_c is the time constant of the microwave field amplitude change. $T_0 = L/\alpha$ has the same value as above (see par. IV). $\Delta_c^{\pm}$ is plotted as a function of $\omega_m T_0$ in *fig. 8* for $b\tau_0 = 1.39$. This last value is that which gives the best short term frequency stability when $\phi = 0$.

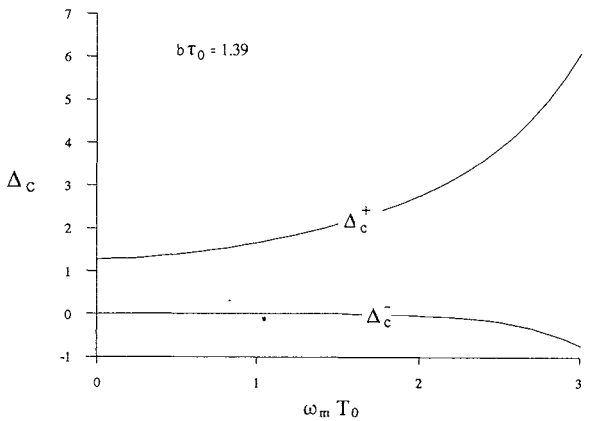


Figure 8: Variation of the cavity pulling frequency shift for the two cavity configurations as a function of the normalized modulation depth.

It is clear that the $\phi = 180°$ cavity configuration leads to a negligible cavity pulling frequency offset. However, the neighbouring lines give a small residual cavity pulling effect which will be considered in a forthcoming paper.

The other advantages of the $\phi = 180°$ configuration concern all the frequency shifts whose magnitude depends on the amplitude of the microwave frequency modulation [3,4]. In that case, the frequency shifts are usually an increasing function of $\omega_m T_0$ and it is obviously favorable to operate the resonator with a small modulation amplitude. This beneficial property tends to balance the increase of the velocity dependent frequency shifts occuring in optically pumped cesium beam resonators, due to the relatively large value of the velocity of the atoms detected. Furthermore, the sensitivity to a microwave power change is reduced.

It is worth pointing out that the advantages just summarized hold for slow square wave modulation.

VI - CONCLUSION

We have operated a small optically pumped cesium beam tube in which the microwave fields in the cavity are 180° out of phase. Compared to the conventional design this mode of operation is attractive from the short term frequency stability point of view if the condition $I'_b/I_0 \ll 1$ is fulfilled. This should be achievable with a more efficient design of the optical interaction zones, a better cleaning of the diode laser beam, and use of silicon detectors showing a lower NEP. In the present circumstances our small machine operating at low oven temperatures (85°C) exhibits a rather outstanding short term frequency stability.

In this paper, we have also pointed out the advantages of the $\phi = 180°$ cavity configuration concerning the lowering of a number of frequency shifts. This technique could be advantageously implemented in cesium beam resonator with magnetic state solution, as well.

ACKNOWLEDGMENT

We are grateful to STEI for its financial support. The authors would like also to thank BNM and PTF for their constant interest in the work.

REFERENCES

[1] V. Giordano, A. Hamel, P. Petit, G. Théobald, N. Dimarcq, P. Cérez, C. Audoin. *"New design for a high performance optically pumped cesium beam tube"*. IEEE trans. on UFFC, 38 (1991) p. 350-357.

[2] J. Vanier and C. Audoin. *"The Quantum physics of atomic frequency standards"*, vol. 2, chap. 5, Ed. A. Hilger.

[3] G.D. Rovera, A. De Marchi, E. de Clercq. *"Estimation of some frequency shift in an optically pumped clock taking into account modulation parameters"*. Proc. 4th EFTF, Neuchâtel 1990, p.161-162

[4] C. Audoin, N. Dimarcq, V. Giordano, J. Viennet. *"Physical origin of the frequency shift in Cs beam frequency standards Related environmental sensitivity."* IEEE trans. on UFFC, to be published, March 1992.

ACTUAL LIMITS OF FLUORESCENCE S/N RATIO IN AN OPTICALLY PUMPED CESIUM BEAM RESONATOR.

N. Dimarcq, V. Giordano, P. Cérez and G. Théobald

Laboratoire de l'Horloge Atomique
Unité Propre de Recherche du CNRS associée à l'Université Paris-Sud
Bât. 221, Université Paris-Sud, 91405 Orsay Cédex, FRANCE.

ABSTRACT

The frequency stability of an optically pumped cesium beam resonator crucially depends on the fluorescence S/N ratio in the detection region. The total noise results from various noise sources : atomic shot noise, photon noise, laser frequency noise and signal detection noise.

We developed a model which gives the fluorescence S/N ratio as a function of the main parameters characterizing the interaction between the atoms and the laser beam. Variations of the S/N ratio with the atomic beam intensity and laser linewidth are presented for various excitation schemes. A good agreement is found between experiments and theory.

Keywords : Optically pumped cesium beam resonator, S/N ratio, laser frequency noise.

I - INTRODUCTION

Optical methods have been implemented in cesium beam frequency standards in order to improve their performances. In the frequency domain, these performances are characterized by the short term frequency stability related to the amplitude S/N ratio measured in a one hertz bandwidth by the relation :

$$\sigma(\tau) \approx \frac{\tau^{-1/2}}{S/N}$$

where τ is the sampling time.

The S/N ratio first depends on the optical transition used for detection. Therefore, two ways for detecting atoms occur (Ref. 1) :

- detection with a pumping transition. The number of photons emitted is of the order of a few units per atom. A typical pumping transition is the F=3 → F'=3 σ polarized transition of the cesium D_2 line. Such a transition delivers 4 fluorescence photons per atom.
- detection with a cycling transition. The two cycling transitions of the cesium D_2 line are F=4 → F'=5 and F=3 → F'=2. When these transitions are used, the atoms "cycle" between the two considered levels and emit a large number of photons during their transit in the laer beam. A typical yield is 40 fluorescence photons per atom.

When the clock signal detection is performed with a pumping transition, a single laser is used to perform atomic preparation and optical detection of the microwave transition : it is the one-laser configuration. On the other hand, when a cycling transition is used for detection, two lasers are required (two-laser configuration). One of them is tuned to a pumping transition to prepare the atom state and the other one is tuned to a cycling transition for detection purpose.

Improvement of the frequency stability by a factor 5 was demonstrated at L.H.A. by using optical methods. It was essentially the result of an increased S/N ratio. In experiments performed with commercial laser diodes at 852 nm, the measured amplitude S/N ratio in a one hertz bandwidth is equal to 10 000 when a one-laser configuration is considered and for an atomic flux equal to 10^{10} at.s^{-1} (Ref. 2). Moreover, the S/N ratio increases when the atomic flux increases. However, for a two-laser configuration, the previous S/N ratio was not reached although a noticeable enhancement of the signal occurs. The S/N ratio is limited to 3000 and is independent of the atomic flux I_{at}. Inversely, it appears that the S/N ratio noticeably increases when the laser linewidth is narrowed (Ref. 3).

The interpretation of the noise features encountered in the optical signals requires a theory which can account for the fluctuations of the exciting light. In this paper, our purpose is to report how the S/N ratio can be strongly affected by laser noise and to derive the actual limits of this ratio.

II - THEORETICAL CONSIDERATIONS

II.1 Noises sources in the resonator

When considering detection of the clock signal by means of emission of fluorescent light, different kinds of noises have to be considered. These noises sources are the following :

- the atomic shot noise which comes from the fluctuations with time of the instantaneous atomic flux. It is related to a Poissonian process. This noise source determines the largest S/N ratio which may be obtained in a laser-atomic beam interaction.

- the photon noise which is associated to the random emission of the fluorescence photons.

- a noise source referred to the laser frequency noise. Indeed, if the laser is not perfectly monochromatic, non-linear coupling between atoms and radiation leads the laser frequency fluctuations to induce amplitude fluctuations in the light emitted by atoms. We consider a white frequency noise.

- a partition noise linked to the imperfect detection of the fluorescence photons.

II.2 Theoretical predictions

A theoretical approach for dealing with the noise problem in an atomic beam excited by a laser has been developed at L.H.A.. It enables us to determine the amplitude S/N ratio through the evaluation of
(i) the mean value of the fluorescence yield $<\mathcal{J}>$ delivered by the light-atom interaction region.

(ii) the power spectral density $\mathcal{S}_{\Delta\mathcal{J}}$ of the fluorescence yield fluctuations.

It then results that :

$$\frac{S}{N} = \frac{<\mathcal{J}>}{\sqrt{\mathcal{S}_{\Delta\mathcal{J}}}}$$

$<\mathcal{J}>$ is proportional to I_{at} and to the number $\mathcal{N}$ of photons emitted per atom. $\mathcal{S}_{\Delta\mathcal{J}}$ takes into account the three noise contributions as follows :

$$\mathcal{S}_{\Delta\mathcal{J}} = \mathcal{S}_{shot} + \mathcal{S}_{photon} + \mathcal{S}_{frequency}$$

It is found that the two first contributions are proportional to I_{at}. Moreover, in the case of a cycling transition, $\mathcal{S}_{photon}$ is negligible compared to the atomic shot noise contribution. On the other hand, in the case of a pumping transition, $\mathcal{S}_{shot}$ and $\mathcal{S}_{photon}$ have the same order of magnitude, depending on the branching ratios of the transitions.

Concerning the laser frequency noise contribution, a noteworthy result is that $\mathcal{S}_{frequency}$ varies as $(I_{at})^2$, increases when the laser spectral linewidth increases or when the laser intensity decreases. These results let us to predict a saturation of the amplitude S/N ratio at high atomic fluxes and a limit value of S/N independent of I_{at}. This limit depends on the kind of atomic transition - pumping or cycling transition - and on the laser features such as the linewidth $\Delta\nu_L$ and the laser intensity I_L.

Until now, the necessary condition for improving the S/N ratio was thought to be an increase of the atomic flux I_{at}. This condition fails as long as the the laser frequency noise contribution is predominent.

III - EXPERIMENTAL RESULTS

We have experimentally test the theoretical predictions concerning the influence of the laser frequency noise on the S/N ratio.

The experimental set-up is a cesium tube with two optical interaction regions (Fig.1). The first zone offers the possibility of locking the laser frequency at the top of the fluorescence line. The fluorescent light is collected by means of aspherical lenses and mirror which focus the fluorescence light onto a silicium photodiode. The collection efficiency is 10 %. In the second zone, the collecting set-up is made of two spherical mirrors and a light pipe which guides fluorescence

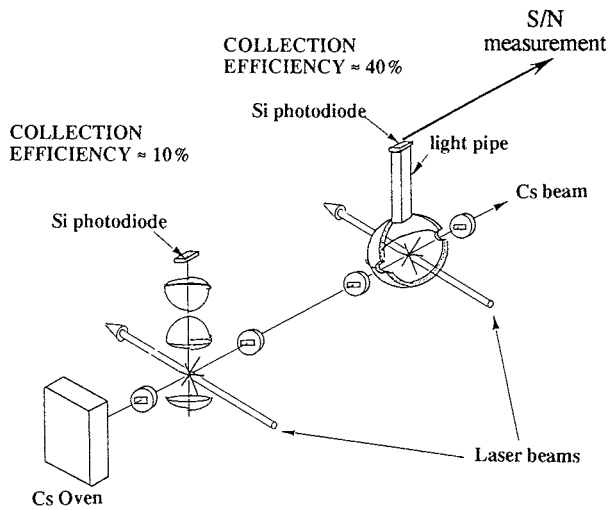


Figure 1. Schematic view of the experimental set-up.

light towards a 10 x 10 mm^2 silicium photodiode. The collection efficiency is 40 %.

The laser sources, which operates at 852 nm, have continuously varying spectral linewidth from 100 kHz to 5 MHz.

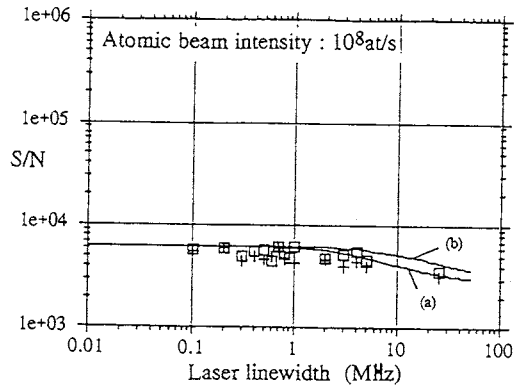
III.1 Cycling transition F=4 → F'=5

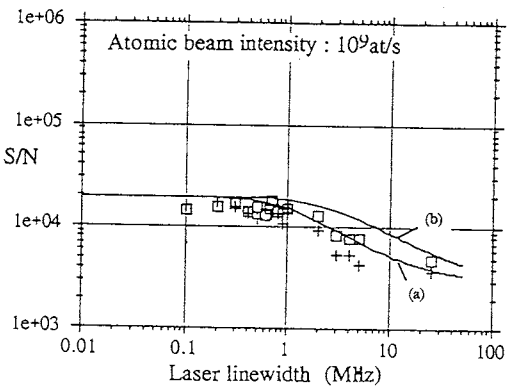
Fig. 2 reports the variations of the S/N ratio as a function of the laser spectral linewidth for atomix flux respectively equal to 10^8, 10^9, 10^{10} and 10^{11} $at.s^{-1}$ and laser intensity equal to 1 or 10 $mW.cm^{-2}$. For weak atomix flux (10^8 $at.s^{-1}$), the S/N ratio does not depend on the laser features as long as the laser linewidth is not large. The atomic shot noise is the predominant noise. For higher atomic beam intensities, the S/N ratio quickly decreases when the laser linewidth increases. The variations of S/N versus the atomic beam intensity are reported in Fig.3 for different laser linewidths (100 kHz, 1 MHz, 5 MHz and 25 MHz). The saturation of the S/N ratio at high atomic flux clearly appears on these curves. It takes place for atomic beam whose intensity is all the weaker as the laser linewidth is large.

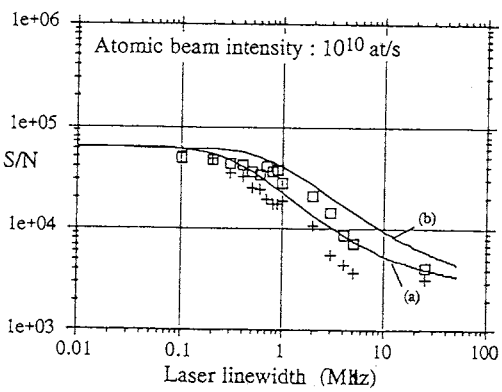
Theoretical predictions and experiments are in good agreement. As thermal atomic beam intensities are usually between 10^9 and 10^{11} $at.s^{-1}$, it appears from these results that :

(i) with a commercial laser diode (25 MHz linewidth), the frequency noise is the main contribution to the noise at detection. It is not necessary to operate the resonator with an intense atomic beam.

(ii) with narrowed laser line, the atomic shot noise is the limiting factor of the S/N ratio. S/N ratio larger than 150 000 may be obtained when the atomic beam intensity is 10^{12} at/s.







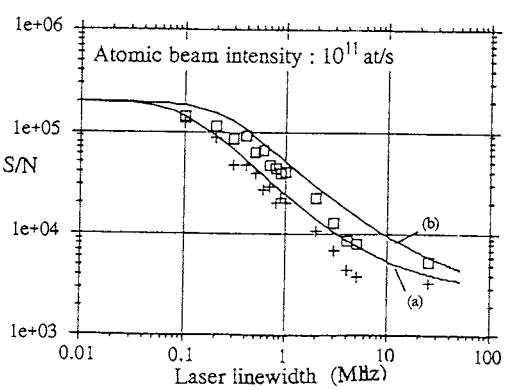
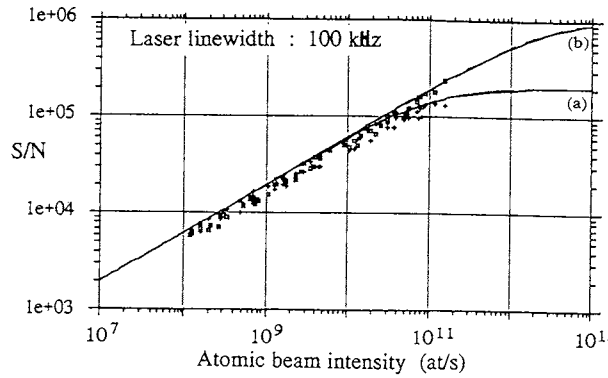
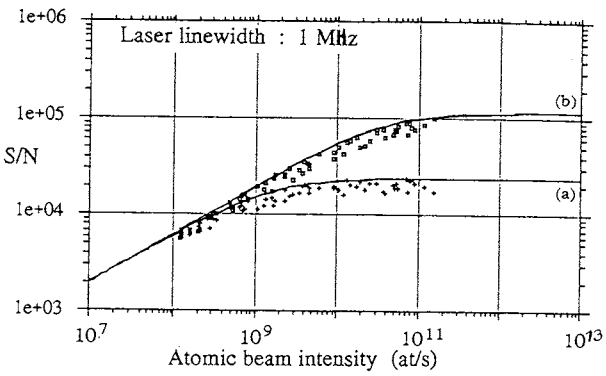
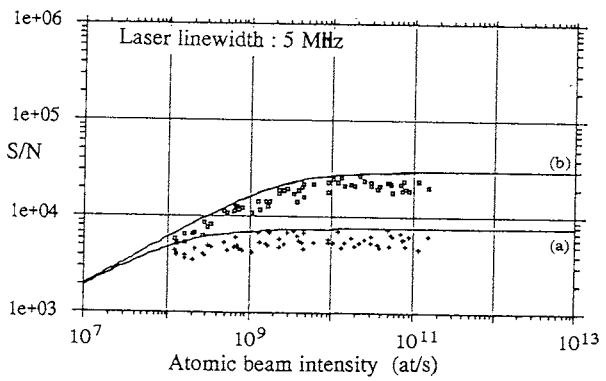


Figure 2. S/N versus laser linewidth for the F=4 → F'=5 transition. Solid lines are the theoretical results calculated with I_L = 1 mW.cm^{-2} (a) and I_L = 10 mW.cm^{-2} (b). Crosses and squares are the experimental results.







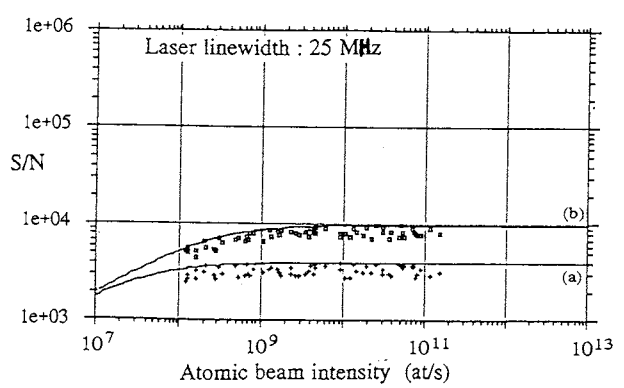


Figure 3. S/N versus atomic beam intensity. The same as Fig.2.

III.2 Pumping transition F=3 → F'=3

Fig.4 shows that the saturation of S/N ratio is also observed at high atomic flux. Nevertheless, this effect is not so significant that is for a cycling transition. With a narrowed laser line, the S/N ratio is 100 000 when the atomic beam intensity is 10^{12} at/s.

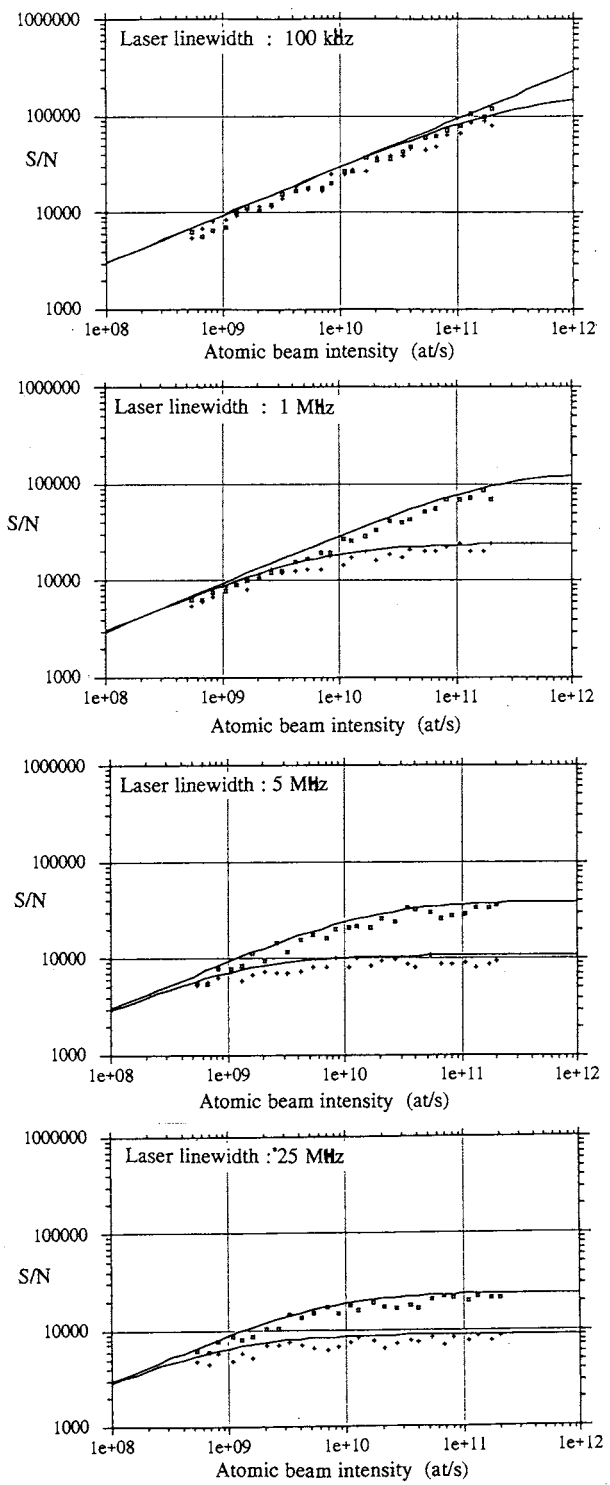


Figure 4. S/N versus atomic beam intensity for the F=3 → F'=3 transition. Solid lines are the theoretical results calculated with I_L = 1 mW.cm^{-2} (lower curve) and I_L = 10 mW.cm^{-2} (upper curve). Crosses and squares are the experimental results.

III.3 Detection noise.

We will mention some conclusions concerning the influence of the photon detection noise on the signal to noise ratio. Let us call η the detection efficiency and $\mathcal{N}$ the mean number of fluorescence photons emitted per atom. The detection does not alter S/N as long as the number of photons emitted per atom and detected (i.e. $\eta\mathcal{N}$) is large compared to 1. In the case of a cycling transition delivering large fluorescence signal, this condition is well satisfied. Consequently, a simple collection optics is sufficient. In the case of a pumping transition where $\mathcal{N}$ is weak, the photon detection modifies the S/N ratio. In order to reduce its influence, a efficient collection optics is required.

IV - CONCLUSION

A significant contribution of the laser frequency noise to the optical detection of atoms in a beam has been pointed out. It leads to a saturation of the S/N ratio versus the atomic beam intensity when the laser is not spectrally purified. When the laser line is narrowed, the S/N ratio is only limited by the atomic shot noise and S/N ratios equal to 150 000 in the case of the F=4 → F'=5 transition and 100 000 in the case of the F=4 → F'=5 transition have been obtained.

ACKNOWLEDGMENTS

The authors want to thank the Bureau National de Métrologie (Paris) for its financial support.

REFERENCES

1. G. Avila & al 1987, State selection in a cesium beam by laser-diode optical pumping, Phys. Rev. A, **36**, 3719-3728.
2. V. Giordano & al 1991, New design for an efficient optically pumped cesium beam tube, IEEE Trans. UFFC, **38**, 350-357.
3. M. de Labachelerie & al 1988, Extended-cavity semiconductor lasers for optical pumping, Proc. 2nd European Frequency and Time Forum, Neuchâtel, Switzerland, 547-554.

Session 4A

Quartz: Resonators

Chairman: W. Zingg

LONG TERM DRIFT AND γ-IRRADIATION EFFECT

J.J. Boy*, A. Berthaut*, D. Thiebaud*, K. Yamni**

* LCEP / ENSMM - Route de Gray - La Bouloie - 25030 Besançon Cedex - FRANCE
** LCCM - Faculté des Sciences - Route de Gray - 25030 Besançon Cedex - FRANCE

ABSTRACT

The frequency and motional resistance characteristics versus temperature of SC-cut resonators were measured before, during and after irradiation in order to correlate the influence of the selected chemical impurities on their electrical behavior.

The long term drifts have been discussed and compared to the Al^{3+} ion content. These measurements were performed on BVA SC-cut, which were manufactured using or synthetic - swept or unswept - or natural quartz crystals.

The results obtained show that a link could be established between the γ-irradiation sensitivity and the aluminum content and hence, the large difference, on one hand, between natural and synthetic raw material and, on the other hand, between unswept and swept quartz crystals.

Key-Words: Quartz crystal, SC-cut, BVA resonator, γ-irradiation, long term drift, frequency-temperature, Q-factor.

1. INTRODUCTION

In the "Laboratoire de Chronométrie, Electronique et Piézoélectricité", extensive measurements have been undertaken to better understand the behavior of quartz resonators. Indeed, for each "serial" production of BVA resonators, we collect results about quartz material on one hand, and about electrical parameters behavior of the resonator, on the other one. We attach us to vary the supplying of the raw material to manufacture our "daily" SC-cut BVA production for which the *design and the manufacturing parameters are settled.*

Furthermore, we perform evenly AT-cut resonators to present results up to international standards as, for example, the IEC publication 758 recommandations. So, many characterisations have been carried out on these samples including chemical and physical defect analyses such as induction coupled plasma emission spectrometry spectrochemical (ICP) analysis, X-ray topography or counts of etch pits and channels.

The frequency stability measurements under irradiation presented here are the first works performed by the lab's manufacturing team with the view to optimize stability of 10 MHz BVA SC 3rd overtone resonators under spatial telecommunication radiation doses. This paper deals only with *high dose radiation sensitivity* of quartz resonators with different aluminum content.

2. SHORT REVIEW ABOUT INTERACTIONS OF RADIATION WITH QUARTZ

It is well known that resonance frequency and Q-value (i.e. stored energy of vibration) of quartz resonators are altered by exposure to X- or γ-ray. Measurements show that these radiations induce around a selected resonance frequency long lived and transient changes in the acoustic parameters of the equivalent circuit, R_m, L, C and, to a less degree, C_0.

Moreover, these measurements show limitations and, sometimes, large frequency shifts, due to material properties. Indeed, sensitivities to ionizing radiation can vary strongly from one to other one quartz blank, and even, from one growth zone to other one of a same synthetic quartz block. So, we can deduce that the ionic impurities and, consequently, the dislocations are mainly involved in the degradation of the wait performances of a "time-base" quartz oscillator. We have reported here the first interpretations :

a. the *transient frequency change* is interpreted to be due to a *variation of the elastic modulus* of quartz brought about by "the detrapping of a positively charged compensating ion at substitutional Al^{3+} (or Fe^{3+}) defects in the quartz lattice" (1).

b. the *large permanent frequency offset* experienced by natural quartz or synthetic crystal with high level of impurities was first interpretated by King (2) and treated more generally by King and Sanders in (3). We just report that the annealing of this frequency change is simply the diffusion-controlled uncorrelated recombination of the cation H^+ (probably), with the Al^{3+} trap.

Therefore, we show that limitations in performances (ageing, long term drift, retrace,...) of a resonator can be essentially due to *migration of impurities* as Na^+, Li^+... or H^+. The frequency shift is correlated to the concentration of these impurities, its variation being related to their motion and nature (in term of migration energy).

Consequently, to use quartz resonators in severe environmental conditions as in space applications, it is necessary to *select the raw quartz material* (and, particulary, to control the parameters of the synthetic quartz growth) or to determine treatments to improve frequency stabilities. Of course, the bad performances being due to the diffusion of interstitial impurities, they could be modified using the *sweeping process* (which is an electromigration under high electric field and at high temperature) which theoretically permits the elimination of these mobile impurities.

We reported, at least, that authors have discussed specifically the behavior of AT-cut resonators (4) (5) whichfor frequency changes are induced by ionizing radiation or high DC field. They explain that *frequency offsets are due to specific crystal defects* which cause :

- stress relaxation

- increasing or decreasing of the mechanical coupling of the resonator with its stress field (coupling between stress and the acoustically active defect).

So, if radiations have altered a defect structure in a way to cause stress, the effective thickness-shear elastic modulus determining the frequency is modified.

3. AT-cut AND SC-cut "BVA" RESONATORS

Even if some special cares are not always taken to manufacture our resonators, we are absolutely sure that the crystal units were identically prepared and mounted. Moreover, as LCEP realized ordinarily some tens of SC-cut BVA resonators a year, a few of them are used for our researchs concerning as well the electronic parts than the resonator itself. For this reason and as far back as possible, we have stored all the possible data about the raw -natural or synthetic- quartz material used.

So we have had the possibility to analyse the behavior under irradiation of nine "10 MHz 3rd Ov. SC-cut BVA resonators" issued from 5 different kinds of crystals : natural, unswept and swept synthetic.

Furthermore, we have irradiated one AT-cut BVA (10 MHz 3rd Ov. too) and the four "classical" AT-cut resonators presented in (6) and issued from a HQ quartz crystal block whichfor just a half part was swept.
The manufacturing process includes :

1 - adjustement and check control to obtain blanks for which the angular dispersion is better than 10".
2 - grinding, polishing and chemical etching operations avoiding the "dead lattice" and microcracks on the surface of the resonator.
3 - electrodeless resonator according to the BVA_2 patent (the 3 parts of the "sandwich" being, of course, of the same cut).
4 - long time burn-in
5 - frequency adjustement by chemical etching and final mounting
6 - annealing at about 280°C and under high vaccum just before coldweld enclosure.(in a MC 40/U can without pinch off).
This last process is assumed to cancel the internal stresses introduced during sawing or polishing and to refine the blank surfaces. In the past, these operations have shown large frequency shifts.

Finally, we obtain SC-cut BVA resonators working at 10 MHz, which are, furthermore, stored in a oven controlled at a temperature close to the turnover point. During this process, the resonators are included in a oscillator loop.

4. EXPERIMENTAL PROCEDURE

4.1 - Quartz material quality set-up :

Data about raw material quality have been determined by the CCM laboratory team :

- the amount of impurities was measured by ICP chemical analysis on a few grams of crystal quartz extracted from each block.

- infrared measurements were performed on a parent quartz sample which is a Y-cut slice of 5 mm thick, to give an intrinsec Q-value. As for an established commercial method for unswept cultured quartz quality evaluation, we have obtained results at two constants wavelength, i.e. 3800 and 3500 cm^{-1}.

- furthermore, in case of synthetic quartz material, we give a data linking α to the content of hydrogen ions according to Lias et al. results.
- the amount of dislocations was checked on X-ray topographies and correlated with measurements of etch pits and channels obtained after chemical etching according to the IEC 758 publication.

4.2 - Resonators :

The resonators described in the previous paragraph have been tested under temperature up to 125 °C: with a desk computer, we check the phase difference between the input and the output levels on a "PI" bridge within which the quartz resonator is inserted. And then, during a slow temperature ramp from 20 to 125 °C, we store the resonance frequency and motional resistance. We note that the first temperature characteristic is always performed after annealing and aging steps which cancel the frequency variations dues to the removal of gaz from the surface of the resonator and the internal stresses relaxation. To illustrate this, the following table shows the difference between three manufacturing steps:

1. the resonator is between the two "condenser / electrode support" for test, just before frequency adjustements.
2. the frequency adjustement and final mounting is achieved : just the high-temperature annealing is not performed.
3. the BVA is finished and annealed at 280 °C under vaccum during three days.

For each characteristic, the three coefficients of the best fitted third degree polynomia as :

$$\frac{\Delta F}{F} = a.\Delta T + b.\Delta T^2 + c.\Delta T^3$$

where :
$\Delta T = T - 25°C$ and $\Delta F = F(T) - F(25°C)$. are given.

	T.O. (°C)	10MHz-F(TO) (Hz)	a $*10^{-6}$	b $*10^{-9}$	c $*10^{-12}$
1	74	-706	0.89	-13.7	62
2	74	-31	0.89	-13.7	62
3	77	-16.8	0.97	-13.9	59

Table 1

4.3. - γ-irradiation exposure :

The quartz crystal resonators are mounted under a ^{60}Co source (1.25 MeV). While a few of them have been irradiated alone (i.e. without electrical excitation), to almost, data on the oscillator frequency shift were collected in 1- or 100-s intervals. Furthermore, the electric circuit was shielded from the direct γ-ray beam, so effects of stray radiation were negligeable. For each test, we have irradiated two resonators whichfor the behavior under irradiation could be very different from one to other one. Finally, the frequencies have been checked by means of a frequency-counter driven by a 10 MHz reference oscillator. Results have been stored on the hard disk of a Hewlett-Packard desk computer as a function of time and / or accumulated radiation doses.

The γ-ray dose rate is about 3.5 gray/min. (350 rad/min.). So, during a 20 hours exposure, the quartz has received a dose slightly higher than 400 Krad.

4.4 - Long term drift measurements :

Our data acquisition system, driven by a Cesium atomic clock, can scan at least 10 quartz crystal oscillators working around its turnover point (mounted in its own test-bed oscillator, controlled in temperature). Every day, the value of the long term drift (which is the mean value of 20 data counted in a 10 s gate time) is given for each resonator. The resolution of the system is better than 10^{-12}.

5. RESULTS

5.1 - ICP and infrared analyses :

In the following table (Table 2), we present the data issued from the I.C.P. analyses. The levels of impurities are expressed in ppm relative to silicon content :

Quartz mat.	Al	Fe	Na	Li
Nat. Q. Madagascar	40	1.6	1.1	6
SICN 563 Unwept	2.80	1.5	1	0.4
SICN 309 Unwept	19			
SICN 272 Unwept	5.79	0.8	1	0.02
SICN 272 Swept	1.7	1.4	1	0

Table 2 : I.C.P. analyses

As we said in (6), the difference in Al. content between the swept and unswept samples (SICN 272) is very important and was not yet explained. In fact, several ICP analysis performed again in blocks issued from the same run give a more lower content. This surprising result concerns probably only one of the blocks located at the bottom of the autoclave...

The hydrogen content, according to Lias et al. (7), could be determined only for unswept quartz crystal. The results obtained are for:

SICN 563 : 100 ppm
SICN 309 : 200 ppm
SICN 272 : 95 ppm

5.2 - Frequency changes under irradiation :

The results concerning the nine BVA SC-cut resonators tested under irradiation are presented in table 3.

The three first one, issued from natural quartz of Madagascar present large frequency shifts: the resonance frequency decreases quasi-linearly with the time exposure.

The two following resonators (samples of SICN 563, unswept) are less sensitive: the slope before saturation is lower than 10^{-12} / rad. Only for one resonator (SC07.08, performed in synthetic quartz from SICN too, run 309) tested in a high dose environnement shown a clear ΔF/F shift which is positive and on the order of a few part in 10^6.

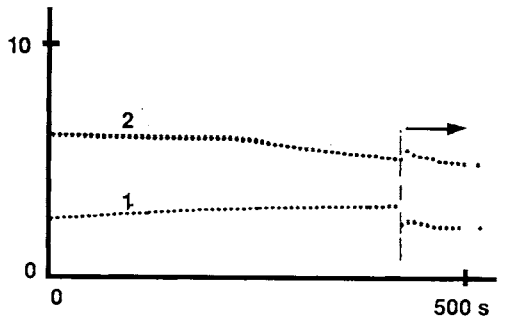
Figure 1 shows graphically the relative behaviors of 5 resonators (submitted to 3/400 Krad accumulated dose) :

1. **18.05** : natural quartz (from Madagascar)
2. **17.08** : natural quartz (- id -)
3. **07.08** : synthetic from SICN (309 unswept)
4. **15.07** : SICN 563 unswept,
5. **19.11** : SICN 272 Swept.

On these curves, we can estimate if a saturation effect exists after a very high dose of γ-ray. We are only sure that swept synthetic quartz with low Al impurity content is very rapidly insensitive to γ-ray.

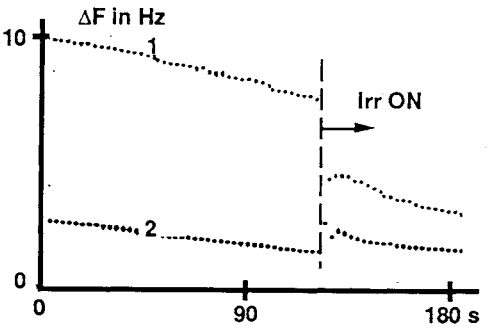
5.3 - Transient frequency changes :

The three next figures (2a, 2b and 2c) characterise the "short-term" radiation experiments, or, more exactly, the behavior of the previous five resonators tested during the first hundreds of seconds just after the beginning of the exposure to severe radiation environnements. The oscillator frequency shifts were collected in 1-s intervals.



1. SC 18.05 (Natural quartz)
2. SC 15.07 (Synt. quartz, 563 unswept)

Figure 2a : 1st irradiation



1. SC 18.05 and 2. SC 15.07

Figure 2b : Second irradiation

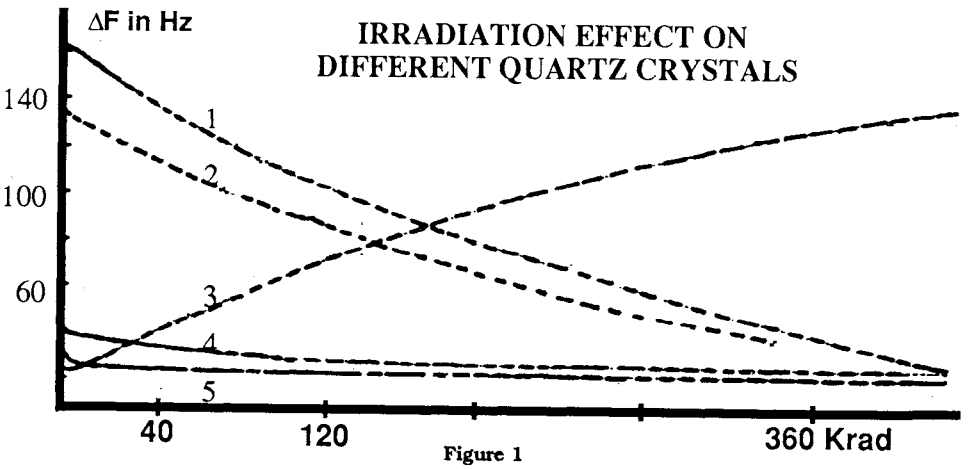
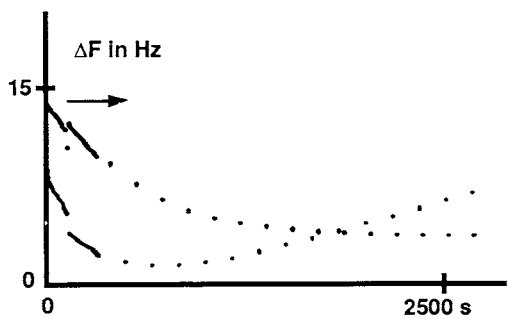


Figure 1

RESONATOR N.	17.04	17.08	18.05	15.07	16.08	07.08	19.06	19.11	20.05
IRR. DOSE (Krad)	550	300	300	300	550	300	350	300	350
In oscillator	NO	YES	YES	YES	NO	YES	YES	YES	YES
On T.O. Point: DF(After-Before) In Hertz	-197	-102	-117	-3	-0.5	+99	≈ 0	-4	-1
DF/F /rad (10^{-12})	-36	-34	-39		<-0.1	+33	<-0.01		<-0.1
Saturation (Krad)	?	NO	NO	≈ 100	?	> 400	< 100	< 100	
Before Irradiation Turn-Over (°C):	84	87	87	80	80	90	78	82	79
After Irradiation	84	87	88	76	76	77	76	79	78
Before Irradiation F(T.O.) - F(25°C):	241	250	258	201	220	?	219	235	221
After Irradiation	217		254	192	192	198	187	201	196
Before Irradiation Q-factor at 25°C:	1.32	1.3	1.28		1.38	1.05	1.36	1.38	1.35
After Irradiation	0.65	0.58	0.53		1.31	0.77			
Before Irradiation Rm (Ohm) at T.O.:	104	97	102	102	100	100	93	94	102
After Irradiation	164	145	165	130	102	260	94	95	98

Table 3



1. SC 07.08 : SICN 309, unswept
2. SC 19.11 : SICN 272, swept

Figure 2c

From these figures, we deduce that the short term behavior of quartz crystal resonators can be described by the next scheme (Figure 3) :

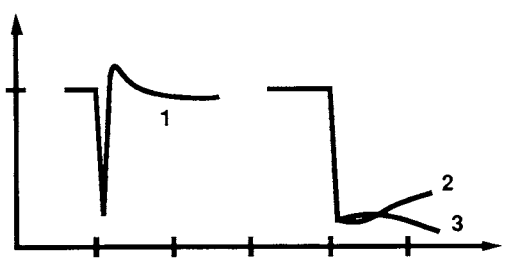


Figure 3 : Short term behavior

The curves 1 and 2 concern the behaviors of quartz resonators manufactured in synthetic quartz with respectively low and high aluminum content. As the curve 3, it can represent the frequency shift of natural quartz resonators. They do not translate exactly the time dependance of the resonance frequency, but show different responses :

-fast and large transient response,
-frequency variation due to temperature increasing (estimated to less than 0.1 °C),
-and finally long term drift due to accumulated dose.

We have, at least, tested under irradiation four "classical" AT-cut quartz resonators manufactured in synthetic quartz for which the electrical parameters have been presented in (6). We remember that the AT-cut resonators number 1 and 2 are in swept synthetic quartz (SICN 272), the others (3 and 4) in the *same quartz crystal*, but unswept).

Only comparisons between the temperature-frequency characteristics before and after irradiation have permitted to show the irradiation effect. The results are presented in the table 4 and have been associated to these concerning the AT-cut BVA 10 MHz 3rd Ov. called "AT 03.08". It was built in synthetic quartz from SICN (239, unswept) (Table 5).

RESONATOR N.	AT 3,4	AT 1,2	03.08
IRR. DOSE (Krad)	800	800	300
In oscillator	NO	NO	YES
On T.O. Point: ΔF(After-Before) In Hertz	3 : +1 4 : +0.5	1 : -0.5 2 : -0.5	5
ΔF/F /rad (10^{-12})	≈ 0.1	≈ 0.1	1
Saturation (Krad)	?	?	>400
Before Irr T.O.(°C):			78
After Irr			75
Before Irr. Rm (Ohm) at T.O.:	QUASI NO CHANGE	NO CHANGE	71
After Irr.			120
Before Irr. Q-factor at 25°C:			1.15
After Irr.			1.08

Table 4 :

	Al	Fe	Na	Li
SICN 239	8.5	0.4	6.2	1.8

Table 5 : SICN 239, unswept

Table 4 shows small frequency changes on the turnover point for the 4 "classical" AT-cut resonators, but it is interesting to note that the ΔF is positive for the 2 "unswept" resonators and negative for the other one. As the changes in motional resistances, they are negligeable to resonators n° 1 and 2 (swept mat.) and very small for the "unswept" (see fig. 5).

In the previous table (Table 5), we note a high Al. content which is translated effectively by a large frequency shift on the turnover point between before and after irradiation exposure. And as usually, the turnover point and the Q-factor decrease together.

The figure 4 describes the behavior of 2 resonators (AT03.08 and SC17.08) as we stopped the irradiation exposure. As the frequency shifts during the "relaxation" period was stored, we are able to compare the possibilities of the resonators to recover its frequency. Furthermore, this figure shows the differences between resonators manufactured in synthetic and natural quartz.

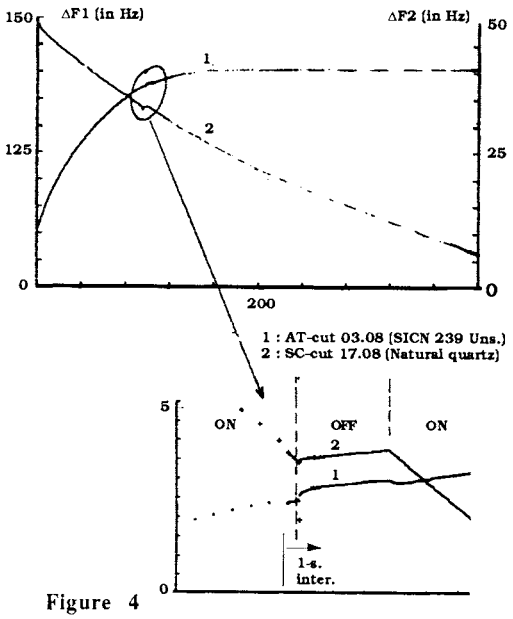


Figure 4

5.3 - Motional resistance versus temparature :

The last figures present the variations of the motional resistance vs temperature after irradiation, and at least after thermal cycles or annealing. The two first one concern the "classical" AT-cut which shows *small differences* (in temperatures lower than 90°C) in the motional resistances of the two "unswept" resonators.

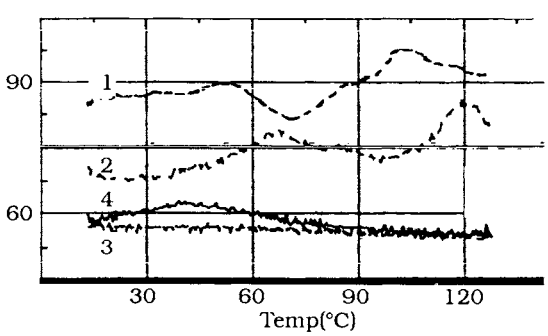


Fig. 5a : Rm before irradiation

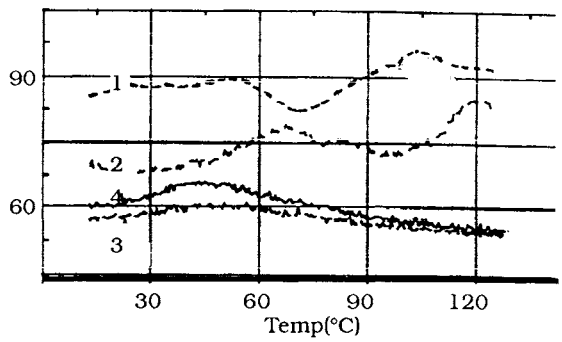


Figure 5

Fig. 5b : just after irradiation

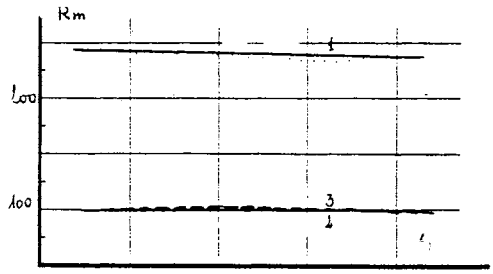
The others (Fig. 6a and 6b) present the different behaviors of synthetic and natural quartz resonators. Though the SC16.08 resonator has not been built in swept quartz, we show that the variations of resistance or other are quasi negligeable ; just the temperature of the turnover point is moving. In contrary, figure 6b shows clearly that the *effect of γ-ray is first very strong at low temperature* and second can be *desappear partially after thermal cycles* or annealings performed at high temperature.

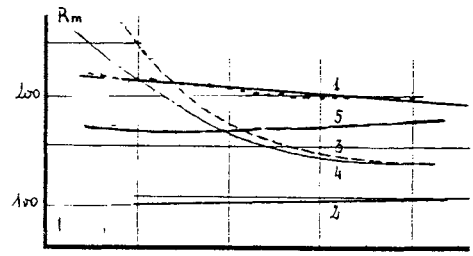
For the two figures 6, the numbers correspond to the following steps :

1. the characteristic is obtained before frequency adjustement, at ambiant pressure,
2. ... just after enclosure,
3. ... just after irradiation,

and only for the SC17.08 resonator :

4. after two thermal cycles around the turnover point,
5. just after 6 hours at about 280 °C (slightly higher than the temperature of annealing ...).





Figures 6a : SC 16.08
Figure 6b : SC 17.04

To point out a possible effect of saturation , we have submitted the SC18.05 resonator to a new 300 Krad exposure. The results are listed in the following table :

18.05 : Nat. Quartz	Before 1st Irradiation	Between 1st and 2d	After 2d Irradiation
T.O. point (°C)	87	88	89
10MHz-F(TO)	0	-117	-179
F(TO)-F(25°C)	258	254	250
Rm @ 25°C	100 Ω	310 Ω	350 Ω
Rm @ T.O.	102 Ω	165 Ω	175 Ω

table 5

So, we show that the saturation effect is more higher than 1 Mrad.

5.4 - Long term drift measurements :

Long term drifts have been tested with test-bed oscillators described above. Only few of the BVA resonators tested in temperature have been measured. Nevertheless, we have, as far back as possible, collected long term drift every 10 days as presented in the tables 6 ($\Delta F/F$ are in 10^{-10}) and figure 7 for the SC 16.08 BVA resonator :

	10 days	20 days	30 days	60days	90 days
SC14.03		3		1	0.8
SC16.08	7	4.5	4	3.5	
SC17.04		20	14	7	
SC19.06	5.5	1.0	< 1.		

Table 6a : Before irradiation

	10	20	30	60	90
SC16.08	9	5	3	1.5	
SC17.04		TOO		BAD !	
SC19.11	1.6	1.5	1.2		
SC20.05	3.5	2	1.9	0.45	

Table 6b : After irradiation

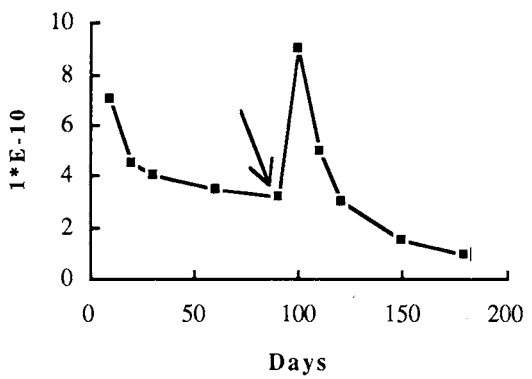


Figure 7 : long term drift on SC16.08

At least, we present in figure 8 the typical long term drift behavior of a resonator issued from the serial number 07: we can see that it exhibts a *very low ageing* (2 or $3*10^{-12}$ per day after 6 months) though the *aluminum content of the raw material is very high* (see table 2). Indeed, we may not deduce a relation between long term drift and impurities content.

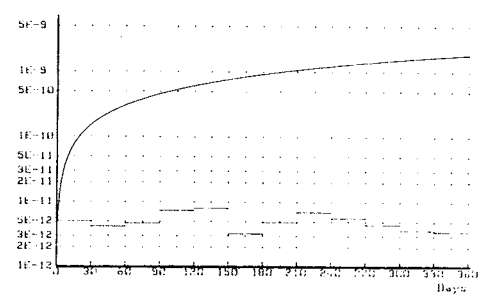


Figure 8 : Long term drift on SC 07.xx

CONCLUSIONS

The behavior under irradiation of many different SC-cut resonators have been studied. The fact of using BVA resonators permits to avoid all phenomena in the layer between electrode and vibrating material.

The results presented in this paper dealing only with high irradiation dose, we conclude that the radiation susceptibility depends on *aluminum content*, *specific crystal defects* and *manufacturing process*. About very low dose levels, we know that there is no correlation between impurity content and radiation sensitivity.

As to the long term drift, we conclude that it is essentially linked to *design* and *manufacturing process* and does not depend *in any way on the aluminum content*. It remains affected by ionizing radiation if the raw material contains large amounts of defects.

ACKNOWLEDGEMENTS

Thanks to the Professor Schraub and M. Courvoisier from the Center of Radiotherapy of the Hospital of Besançon for having facilitated our access to the Cobalt source.

REFERENCES

1. J.C. KING, H.H. SANDER : "Rapid annealing of frequency change in high frequency crystal resonators following pulsed X-irradiation at room temperature", 27th Annual Freq. Control Symp. (1973), pp. 113, 119

2. J.C. KING, Bell System Tech. J. 38, 573 (1959)

3. J.C. KING, H.H. SANDER : "Rapid annealing of frequency change in crystal resonators following pulsed X-irradiation", IEEE Trans. on Nuc. Sci.,NS-19, N° 6 (1972), pp. 23, 32

4.J.J. SUTER, A.G. BATES, J.M. CLOEREN: "The radiation susceptibility of BVA-AT resonators to simulated medium earth orbit radiation", Proc. 2d European Frequency and Time Forum. (1988), pp. 679, 686

5.R. BRENDEL, J.J. GAGNEPAIN, J.P. AUBRY : "Impurities migration study in quartz crystal resonators by using electroelastic effect", 40th Annual Freq. Control Symp. (1986), pp. 121, 126

6.J.J. BOY, P. ZECCHINI, B. VIARD : "Synthetic Quartz Material Characterisations. Connexions with the Quality of AT-cut Resonator", Proc. 5th European Frequency and Time Forum. (1991), pp.

7. J.C. BRICE, A.M. COLE : "The characterisation of synthetic quartz by using infra-red absorption, Proc. 32d Annual Frequency Control Symp. (1978), pp. 1, 10

WEAK MAGNETIC FIELD SENSITIVITY OF QUARTZ CRYSTAL OSCILLATORS

R. Brendel[1], E. Robert[2]

[1] Laboratoire de Physique et Métrologie des Oscillateurs du CNRS
associé à l'Université de Franche-Comté-Besançon,
32 av. de l'Observatoire, 25000 Besançon, France

[2] CNES - AE/TT/TF, 18 av. E. Belin, 31055 Toulouse Cedex, France

ABSTRACT

This paper presents the magnetic sensitivity curves for selected quartz crystal resonators in the range of ± 40 gauss. It is shown that in most cases the sensitivity to a small magnetic field can hardly be extrapolated from the high field measurements. Also, the frequency-versus-magnetic field curves often exhibit a strong hysteresis which makes questionable the definition of the resonator's magnetic sensitivity. Although the main cause of this sensitivity remains the ferromagnetic nature of the material used in the plate support, it is shown that the mounting technique can be of great importance.

1. INTRODUCTION

Most positioning systems use the Doppler effect. In such systems the measurement accuracy depends strongly on the frequency stability of the master oscillator. Up to now most of the frequency instabilities studied are due to temperature variation, acceleration or ionizing radiation. Many efforts have already been made to improve oscillator performance either by using selected components insensitive to the environmental variations or by shielding the oscillator against these variations. As the oscillator performance increases, new applications such as precision orbit determination need more and more accurate devices. As a consequence, some perturbation sources, ignored to date as having negligible influence compared to others, now emerge and need attention. Thus, for a few years, attention has been focused on magnetic field effects. In fact, it was observed that starting or stopping on-board magnetic stabilizing equipment may lead to slight frequency shifts in an oscillator placed nearby. Even if in most cases this effect may be ignored, being of the order of magnitude of 10^{-11}/gauss, it is unacceptable in high precision orbitographic measurements such as the DORIS mission in which the goal is to locate the satellite with a precision of 10 cm. Such a requirement needs a frequency stability of the order of 10^{-13} whatever the perturbation source may be. The actual magnetic sensitivity doesn't permit that specification to be met. In order to understand the origin of the oscillator magnetic sensitivity, CNES has initiated some studies which have identified the main mechanisms of that sensitivity. Previous papers have shown that almost all components of the oscillator circuitry may have a more or less important effect on the magnetic sensitivity. At the same time some solutions to reduce it have been proposed. Of course, among the studied components, the resonator itself has been given special attention because of its major role in the oscillator frequency. But, although the magnetic sensitivity is easily seen on the oscillator as a whole, it is rather difficult to measure the contribution of the individual components since they have to be kept in the magnetic field far from the remaining part of the oscillator. For that reason, the experiments done so far have been performed with a fairly large magnetic field (100-1 500 gauss) in order to increase the signal-to-noise ratio. In fact, the actual magnetic field strength lies in the range 0-10 gauss. This is why the experimental setup has been improved so as to make possible the magnetic sensitivity measurements of individual components with a smaller magnetic field (namely in the range 0-40 gauss).

2. EXPERIMENTAL SETUP

In the earlier experiments (Refs. 1, 2), the magnetic field was obtained by using an electromagnet but weak fields were not obtainable with this arrangement. So only the high-field sensitivity (up to ± 1 300 gauss) has been investigated so far. Fig. 1 shows the experimental setup used for this work. The magnetic field is obtained by means of Helmholtz coils the current of which is driven by a low frequency sinusoidal generator. The magnetic field can vary with an amplitude of about ± 40 gauss (± 4 mT). The coils can be oriented in two perpendicular positions so as to obtain either a vertical or a horizontal magnetic field. In the latter case, the field orientation with respect to the resonator is adjusted by rotating the crystal around its symmetry axis. Such a configuration is not exactly the one suggested by the IEEE (Ref. 3) which recommends that the device under test has to be kept in a fixed position to avoid errors due either to the earth's magnetic field or to gravitational effects when rotating the device rather than the magnetic field itself. In our case these two effects do not matter because the effective magnetic field impressed on the device is measured by a gaussmeter the Hall probe of which is placed near the resonator. With regard to gravitational effects, because of the symmetry of the device, it behaves the same way whatever its position may be. Nevertheless, a new experimental setup in which the device is kept in the same orientation is currently being built.

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

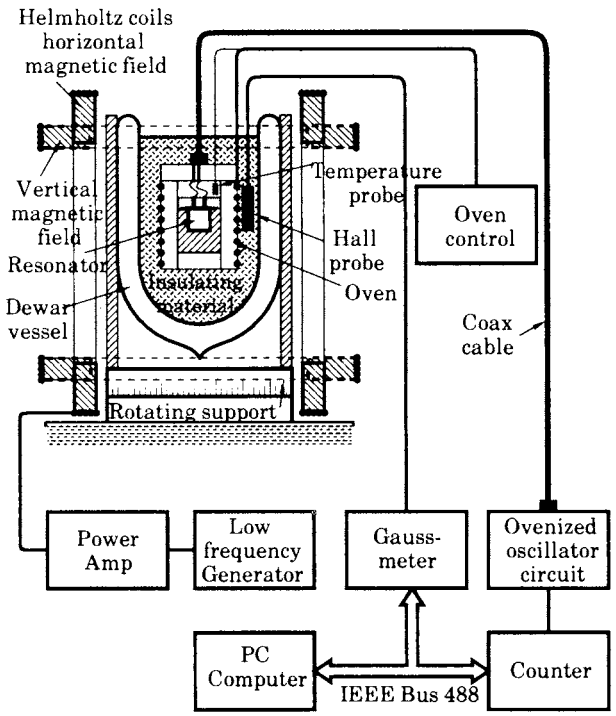


Fig. 1 : *Experimental setup*

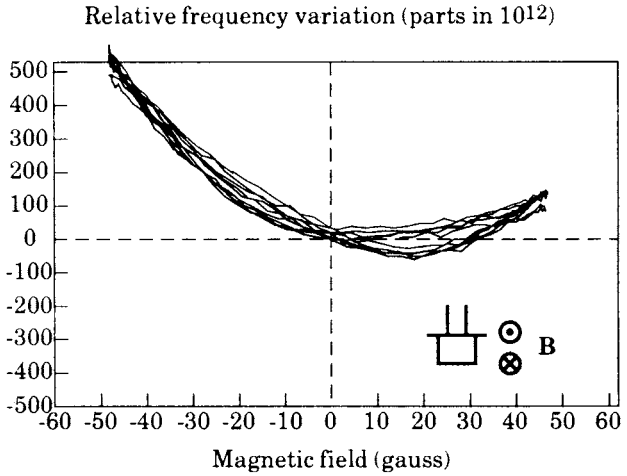


Fig. 2 : *Raw data of the magnetic sensitivity of a quartz crystal resonator. The magnetic field is perpendicular to the plane of the leads*

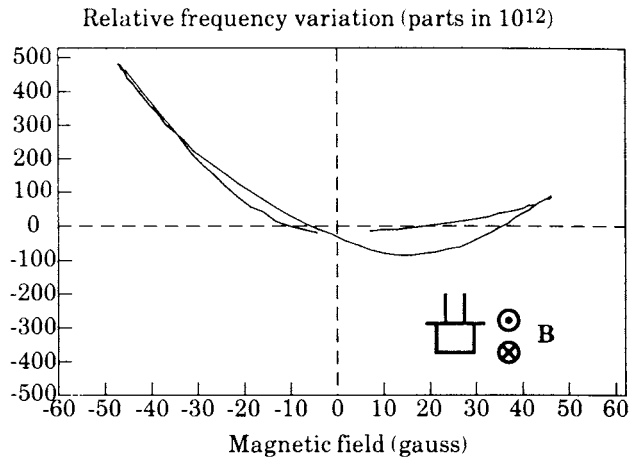


Fig. 3 : *Averaged record of the magnetic sensitivity of a quartz crystal resonator*

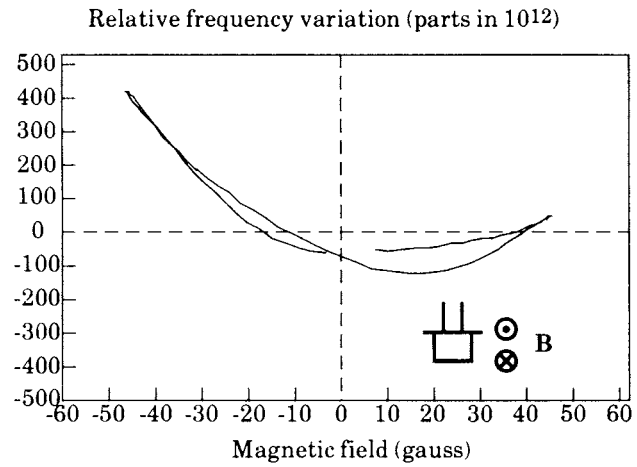


Fig. 4 : *A second record performed on the same resonator as that used for Fig. 3*

The most important source of perturbation comes from the temperature effects. To reduce these effects, the resonator under test is kept at its turnover point, in an oven placed in a Dewar flask.

The oscillator circuit is kept in a separate oven far from the magnetic field so as to avoid unwanted magnetic effects on the other parts of the oscillator. Nevertheless, temperature effects and other environmental perturbations can still affect the measurement through the rigid coax cable needed to connect the resonator to the oscillator circuit. To free the measurements from these spurious effects, the magnetic field is given a low-frequency sinusoidal modulation thus enabling the use of signal-processing methods such as the averaging process used here. Figure 2 shows raw data in which the noise and other external perturbations result in a broadening of the trace. Figure 3 shows the result obtained by averaging the record over 10 periods of the excitation signal. The period of this signal has been chosen so as to meet two opposing requirements: it has to be long enough to insure that no dynamic magnetic effects can occur and short enough to avoid long term temperature drift which could affect the accuracy of the averaging process. At first glance, the hysteresis observed in Fig. 3 could be attributed to a dynamic magnetic effect (due to a delay between the excitation and the response of the device) but we have verified that this effect does not change when the excitation frequency is divided by 10. Therefore the hysteresis is in fact a magnetization effect. After some preliminary experiments, we conclude that an excitation period of 100 s and a sampling time of 1 s constitute a satisfactory compromise. In the actual experimental setup, relative frequency variations of a few parts in 10^{12} can be seen with a reasonable reproducibility, as shown by the similarity between the two different records shown in Figs 3 and 4.

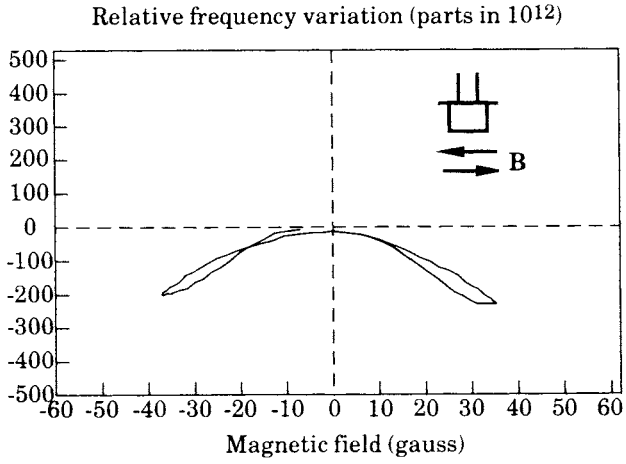


Fig. 5 : *Sensitivity when the magnetic field is in the plane of the leads (same resonator as on Fig. 3)*

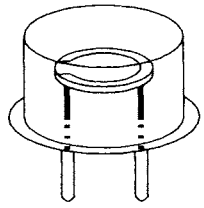


Fig. 6 : *Internal assembly of the resonator under test (10 MHz, 3rd overtone)*

3. EFFECT OF ORIENTATION

As already observed in most of the previous work (Refs. 1, 4, 5), the resonator sensitivity is strongly dependent on its orientation in the magnetic field. Fig. 5 (magnetic field is in the plane of the leads) shows a curve obtained with the same resonator used for Figs. 3 and 4.

Fig. 6 shows the internal assembly of the resonator under test. In this case, the concavity is reversed. Also, the curve shown on Fig. 3 seems to be unsymmetrical with respect to the vertical axis. This effect is probably due to device-positioning uncertainty. In fact, the device position with respect to the axis of the coils is sensitive to the juxtaposition of several connected components such as plate-support, support-case, case-oven, oven-Dewar and Dewar-coils (see Fig. 1).

Figures 7a, b and c show the effect of a slight rotation around the symmetry axis of the resonator (about ± 5°).

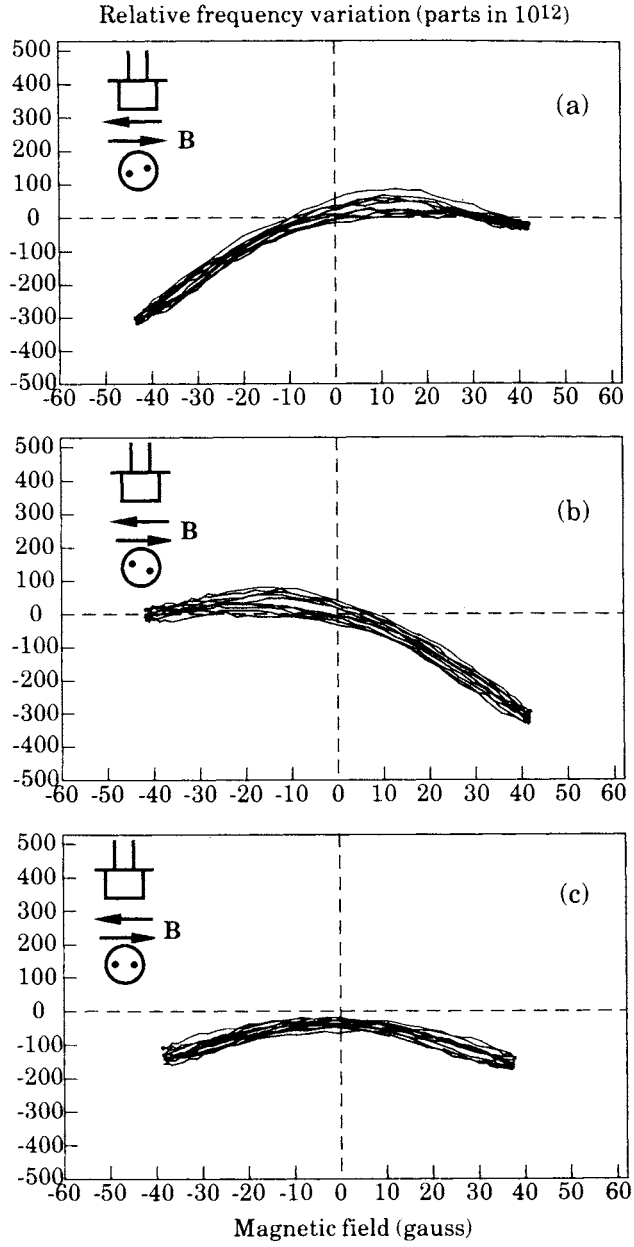


Fig. 7 : *Effect of a slight rotation around the symmetry axis of the resonator*

4. COMPARAISON WITH HIGH FIELD SENSITIVITY

Figure 8 shows the results obtained when the magnetic field is normal to the plate. In this case we observe a strong hysteresis and a rather high overall sensitivity. The results demonstrate that it is not always possible to extrapolate to the weak magnetic field sensitivity from the high field measurements. The latter could lead us to believe that the symmetry axis of the resonator is the direction of weakest sensitivity (see Fig. 1 of Ref. 2). Nevertheless, the comparison between Figs. 3 and 5 and Fig. 1 of Ref. 2 shows that the shape of curves obtained in a weak field is in accordance with what might be expected from high field experiments.

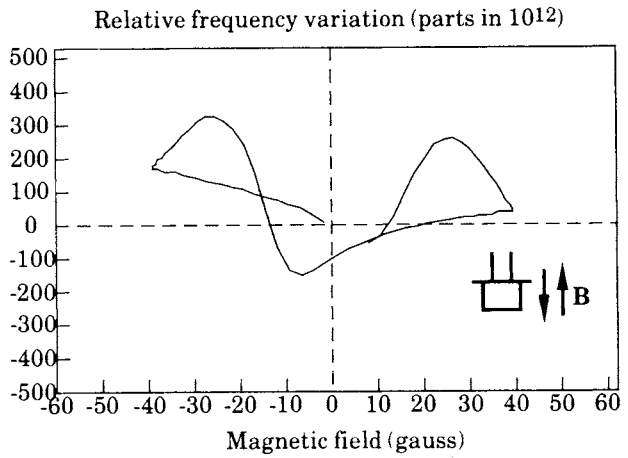


Fig. 8 : *Sensitivity when the magnetic field is along the symmetry axis (same resonator as in Fig. 3)*

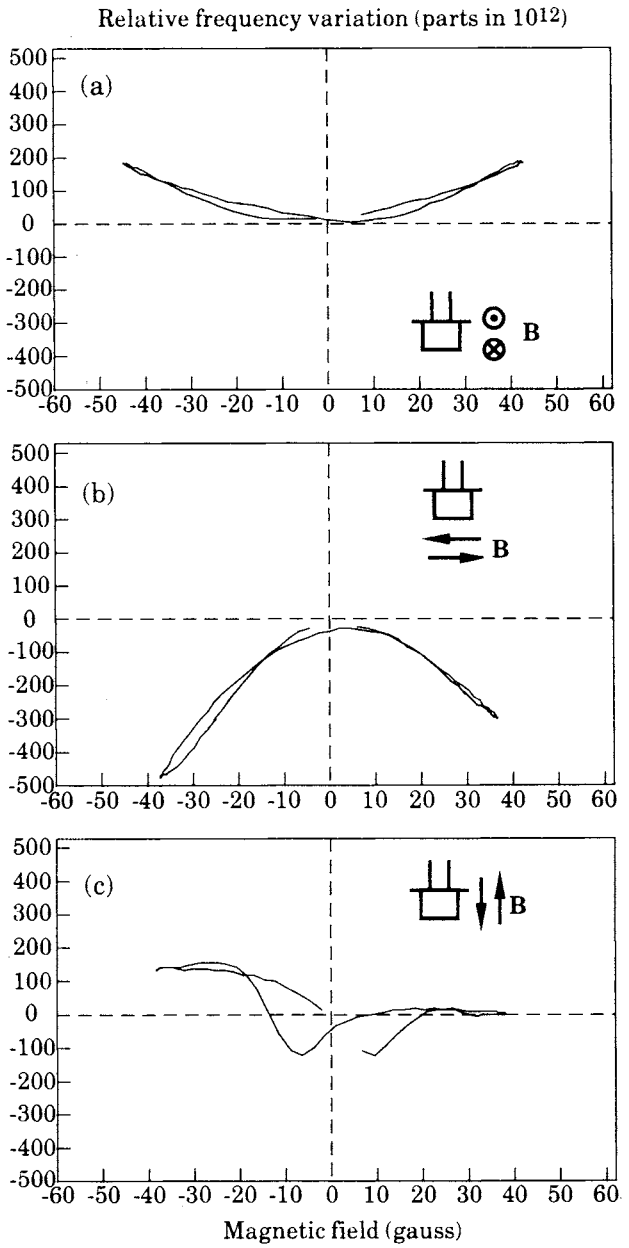


Fig. 9 : *Magnetic sensitivity of a regular resonator of same technology as in Fig. 3, 5 and 7*

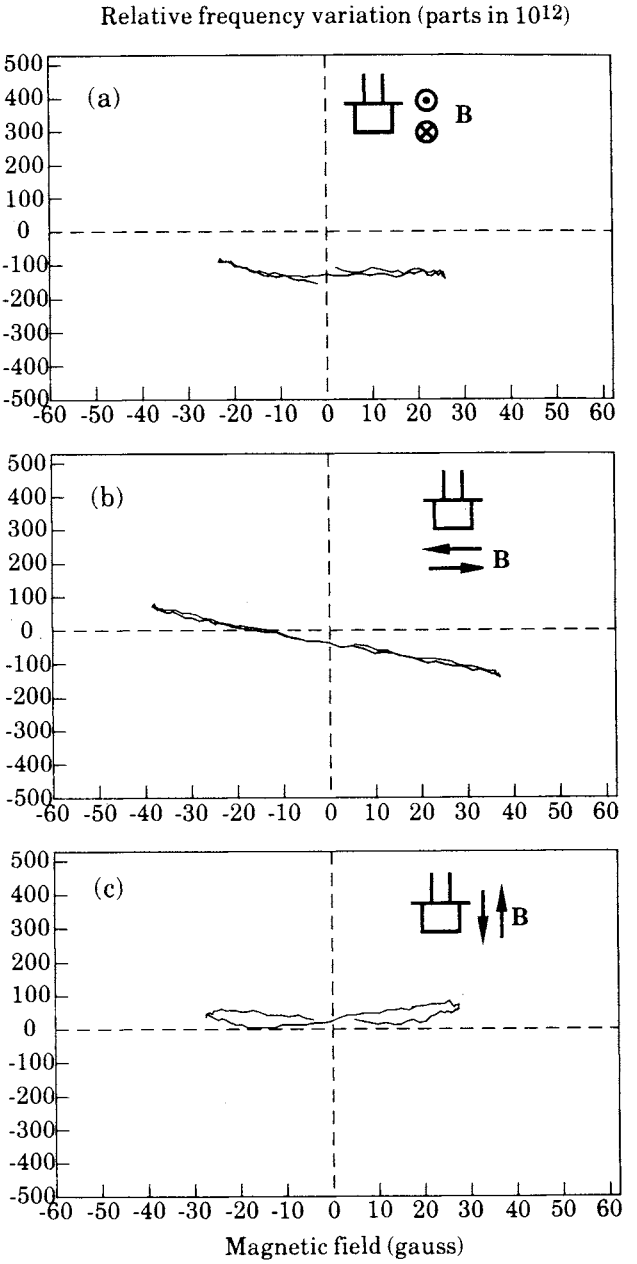


Fig. 10 : *Magnetic sensitivity of a resonator with a different technology*

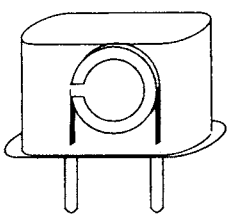


Fig. 11 : *Internal assembly of the resonator used for Fig. 10 (10 MHz, 3rd overtone)*

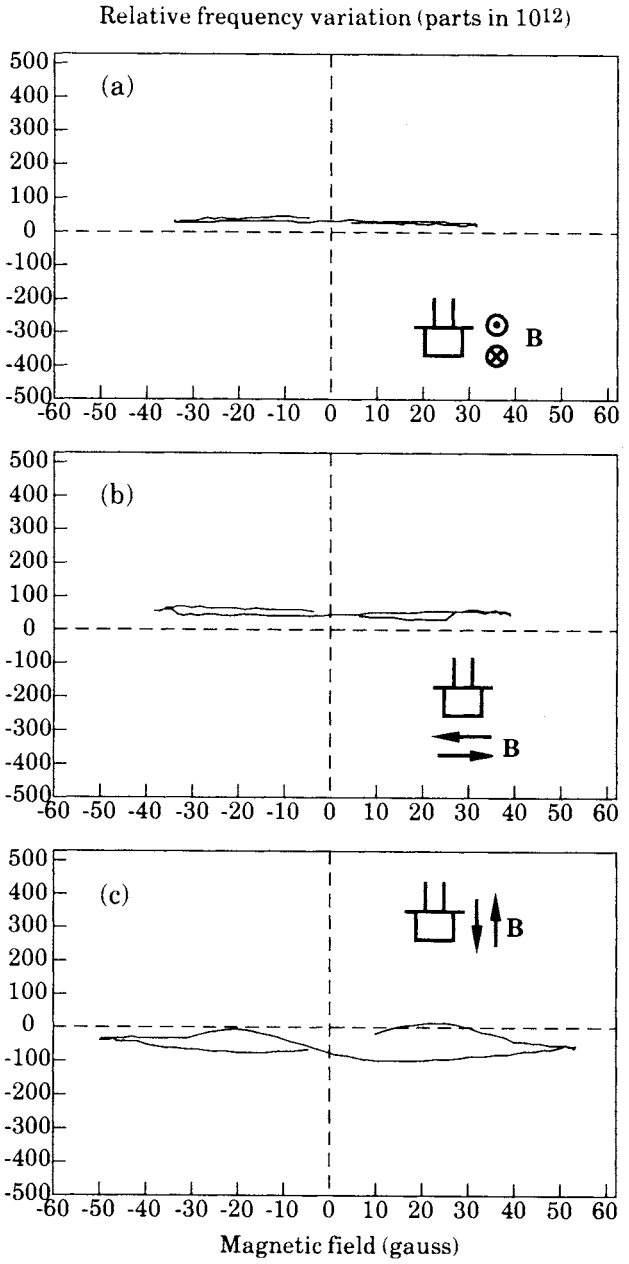


Fig. 12 : *Magnetic sensitivity of a BVA resonator (10 MHz)*

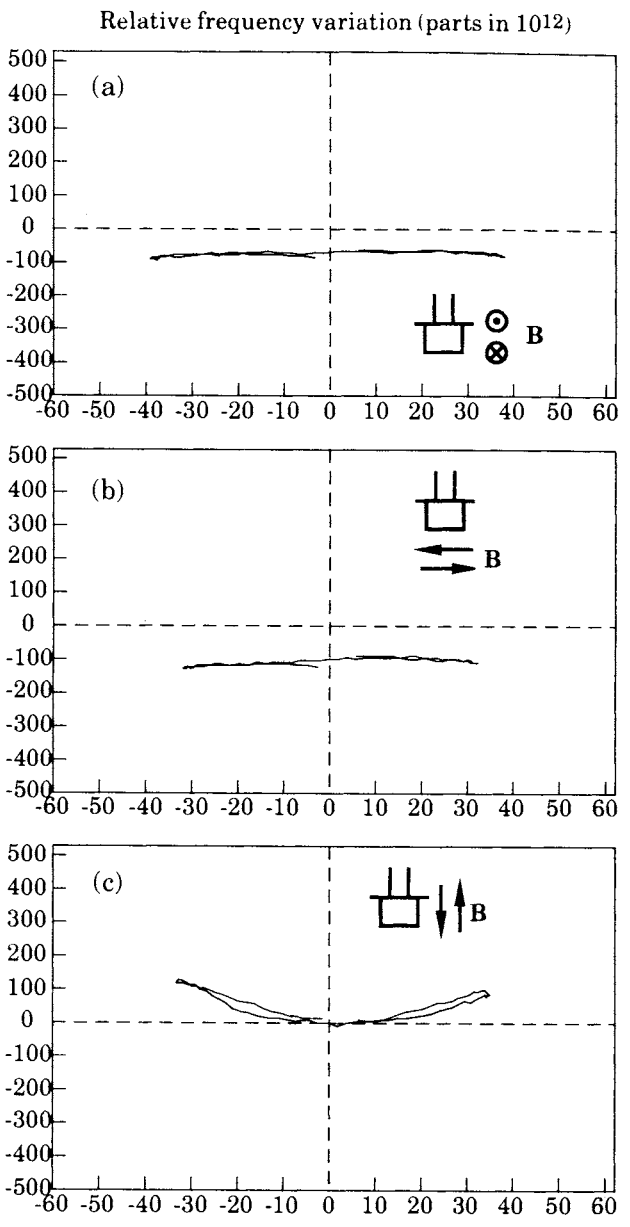


Fig. 13 : *Magnetic sensitivity of another BVA resonator*

5. COMPARAISON BETWEEN RESONATORS

Figs 9 a, b, c show data obtained using a resonator of the same technology under the same conditions as for Figs 3, 5 and 8. The overall behavior remains approximately the same. Only the amplitude of the curves is more or less pronounced. The main difference appears when the field is directed along the symmetry axis (Figs 8 and 9c) and has not yet been explained. The curves shown in Figs. 10a, b, c have been obtained with a resonator using a different technology. Its internal assembly is shown in Fig. 11. The results obtained show that the magnetic sensitivity of this resonator is substantially smaller than that of the previous resonator. The origin of the observed difference is not yet understood. It may be attributed either to the geometry, the thickness, or the magnetic properties of the crystal mounting used.

Other experiments have been carried out using resonators of the BVA type (Ref. 6). Figs. 12 a, b, c show the data thus obtained. It appears from the figures that this type of resonator has a rather low magnetic-field sensitivity due to the particular mounting system used. This fact also was evident for the high field experiments (Ref. 2). The results obtained with another sample of a BVA-type resonator are shown in Figs. 13 a, b, c and confirm the small magnetic sensitivity of this device.

6. MAGNETIC SENSITIVITY CHARACTERIZATION

The various shapes of the sensitivity curves obtained so far make specifying the magnetic sensitivity of a resonator by using a single number difficult. Such a number could be: a) The overall frequency deviation observed for a given amplitude of the magnetic field ; b) the maximum slope of the sensitivity curve; or c) the slope of the sensitivity curve near the origin so as to be able to extrapolate the small field sensitivity.

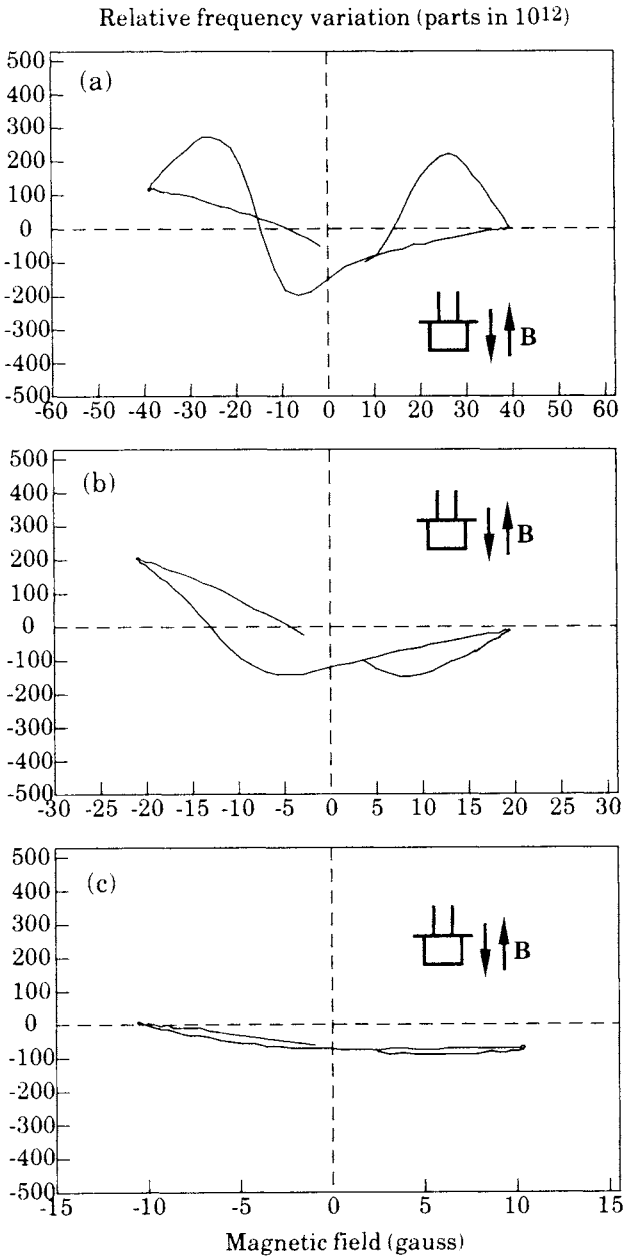


Fig. 14 : *Magnetic sensitivity curves obtained with smaller and smaller magnetic field amplitude*

Even in this latter case the experimental conditions have to be specified because the slope at the origin depends strongly on the amplitude of the magnetic field used as shown in Figs. 14 a, b, c where smaller and smaller magnetic field amplitudes have been used. In order to overcome these ambiguities, the IEEE working groups on test methods for environmental sensitivities (Ref. 3) intends to recommend that the device magnetic sensitivity be given in the form of curves such as those presented in this paper but with a magnetic field in the range of ± 1 gauss.

CONCLUSION

Although the experiments done in this work have been performed with magnetic fields much weaker than those used in the preceeding paper (Ref. 2), their magnitudes are still substantially higher than the values usually found under normal environmental conditions. The need for these somewhat higher field, was dictated by the need to improve the signal-to-noise ratio which was attenuated by the distance between the resonator under test and the remaining oscillator circuit. Nevertheless, the results obtained demonstrate that the magnetic sensitivity of the resonator depends strongly on the technology used in the crystal plate mounting. Of course, the material used for the supports remains the principal factor responsible for the magnetic-field sensitivity. A new set of experiments is being initiated in order to assess the magnetic behavior of different materials wich could be used for the quartz crystal resonator mount.

REFERENCES

1. R. Brendel, C. El Hassani, M. Brunet, E. Robert, "Influence of magnetic field on quartz crystal oscillators", Proc. 43rd Ann. Symp. on Freq. Control, pp. 268-274 (1989).

2. R. Brendel, C. El Hassani, P. Krux, E. Robert, "Magnetic sensitivity of oscillator components", Proc. 5th Europ. Freq. and Time Forum, pp. 390-395 (1991).

3. R. L. Sydnor, R. Brendel, M. Garvey, W. Wiedemann, "Guidelines for test methods for environmental sensitivities of standard frequency generators", IEEE-SCC27 working group 1 (in preparation).

4. A. Benezet, E. Robert, M. Brunet, "Compte rendu d'essais n° L5160", CNES Technical Report CT/DRT/TIT/TF 87/297 (1987).

[5] E. Robert, M. Brunet, "Susceptibilité magnétique des oscillateurs à quartz spatiaux", CNES Technical Report CT/DRT/TIT/TF 88/354, (1988).

[6] R. Besson, "A new electrodeless resonator design", Proc. of the 31st Ann. Symp. on Freq. Control, pp. 147-152 (1977).

ACKNOWLEDGEMENTS

The authors wish to thank the companies CEPE, QKP, Prof. R. Besson for providing resonators and Dr. N. Yannoni who has kindly revised the manuscript.

This work was supported by CNES and CNRS under contract 932 88 53 474 00.

SYNCHROTRON RADIATION X RAY TOPOGRAPHY STUDY OF LATERAL FIELD RESONATORS

B.Capelle*,J.Detaint**,J.Schwartzel**,Y.Zheng*,A.Zarka*

*L.M.C.P. Université P.& M.Curie,Place Jussieu 75005 Paris France
**CNET dept.PAB/BAG/MCT/CMM 192,Av.H.Ravera 92220 Bagneux France

Abstract.

Using the synchrotron radiation delivered by the DCI storage ring at the LURE (Orsay, France), plano convexe AT resonators excited by a lateral field were studied by the X ray topography technique. It was observed different overtones and anharmonics of the three types of the thickness modes.

Introduction.

In the lateral field plano convexe AT resonators, it is possible by choice of the in plane direction of the field, to excite the overtones and the anharmonics of the three types of thickness modes, whereas only one thickness mode, the fast shear mode, is piezoelectrically excited by a strictly thickness field. But the question arises to determine whether the tridimensional modes existing in the lateral field resonators and in the thickness field resonators are identical. The use of the X ray topography with the stroboscopic technique [1] permits to compare with a good precision the modes existing in these two kinds of resonators.

We have obtain, with the transmission and the reflexion Laue set-up, traverse topographs showing the three components of the displacement vector of different modes and could observe on section topographs the number of antinodal lines in the thickness of the resonator which corresponds to the overtone. Namely that we have could identify the modes by X ray topography.

The experimental set-up.

The transmission Laue set-up which uses the white X ray beam delivered by the synchrotron is shown on the figure 1. The resonator is placed in the white beam and we obtain on the film, placed behind the crystal, a Laue pattern on which the crystal is simultaneously imaged with different diffraction vectors.The contrast observed depends of the relative position of the diffraction vector and the vibration displacement vector. In particular, the contrast due to the vibration vanishes on the topograph when these two vectors are perpendicular. To visualize a particular component of the displacement vector of a mode alone, we must choose a diffraction vector parallel to this component.

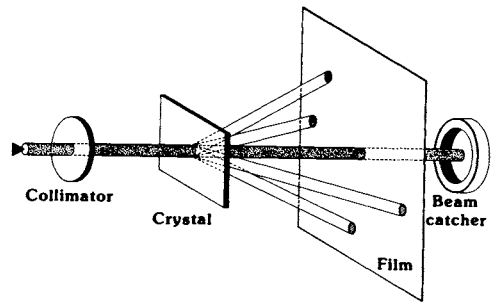


Figure-1- The transmission Laue set-up.

Two types of topographs can be obtained. The first one is a traverse topograph which is obtained when a large incident X ray beam is used, the second one is a section topograph when the incident beam is limited by a fine slit of about fifteen microns wide. With this second type of topograph we can see the section of the resonator and know how much antinodal lines we have in the thickness.

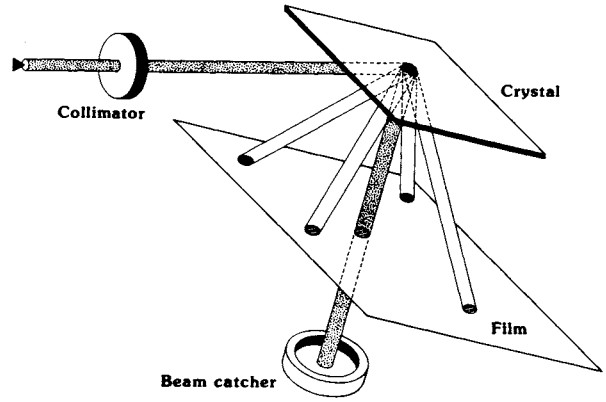


Figure-2- The reflexion Laue set-up.

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

To determine all the components of the displacement vector of a mode we must use also a diffraction vector perpendicular to the surface of the resonator. It is possible to have a diffraction vector like that, only with the reflexion Laue set-up (fig.2).

To use the stroboscopic experimental set-up the resonator must be designed to have a vibration resonant at a frequency equal to n times that of the X ray synchrotron radiation which is 3.169280 MHz.

One dimensional theory.

The figure 3 shows the axis used usually. The X, Y and Z axis are the natural ones; the X, Y' and Z' axis are the rotated axis corresponding to the AT cut, they are also named x_1, x_2 and x_3. To determine the three components u_1, u_2 and u_3 of a displacement vector **u** we must choose a

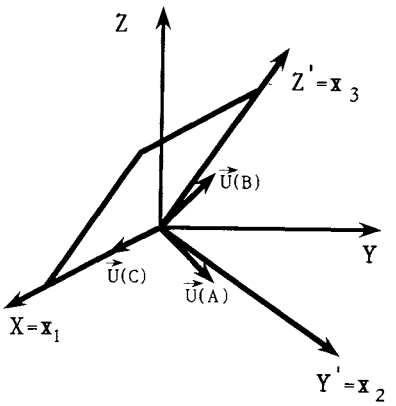
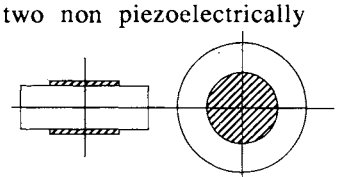


Figure-3- The different axis used usually. The X,Y and Z axis are the natural ones; the x_1, x_2 and x_3 axis are the rotated axis corresponding to the AT cut.

diffraction vector parallel to x_1, x_2 and x_3 respectively. It is possible to have a diffraction vector parallel to x_1 or to x_3 with the transmission Laue set-up and a diffraction vector parallel to x_2 with the reflexion Laue set-up.

The one dimensional theory [2,3] of the AT cut resonators using the conventional thickness field excitation predicts (Table 1) the existence of three thickness modes: one piezoelectrically excited, the slow thickness shear mode or C mode, and two non piezoelectrically excited, the fast thickness shear mode or B mode and the thickness stretch mode or A mode. Namely that the coupling coefficient is equal to zero for the A and B modes and different to zero for the C mode



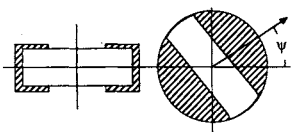
mode	longitudinal (A) (quasi)	fast shear (B) (quasi)	slow shear (C) (pure)
u (natural axis)	0.0 -0.77918 +0.62680	0.0 -0.62680 -0.77918	1.0 0.0 0.0
$V/2 = NF_a$ (kHz.mm)*	3503.6	1899.7	1660.8
k_{TE}	0	0	0.088

* $NF_r \# NF_a (1-4k_{TE}^2/\pi^2)$.

Table 1

The displacement vector **u** for the C mode is exactly parallel to x_1 direction whereas the displacement vector for A and B modes makes an angle equal to about 3°.5 with the x_2 and x_3 directions respectively.

In the case of AT cut resonators using lateral field excitation it is possible by choice of the in plane direction of the field to excite the overtones and the anharmonics of these three modes , then the correspondent coupling coefficient is different to zero. When the field is parallel to x_3 direction or Z' the C mode can be excited and when the field is parallel to x_1 direction the A and B modes can be excited, namely that with the lateral field excitation the coupling coefficients depend of the field (Table 2)[4].



	mode A quasi longitudinal	mode B quasi (fast) shear	mode C quasi (slow) shear
k_{LE} (field in the Z' direction)	0	0	0.06
k_{LE} (field in the X' direction)	0.064	0.062	0

Table 2

Some three dimensional models for the AT cut resonators with the thickness excitation have been proposed by different authors[5,6,7,8], they permit to describe with a good precision the modes and the characteristics of these resonators. The plano convexe resonators excited by a lateral field have been studied since many years [9,10,11,12] on account of the potentialities to obtain ultra stable frequencies [13], but we have not the same comprehension of the real modes that for the thickness field excitation resonators.

Lateral field excitation :

-Parallel to the x_3 direction.

The fifth overtones of the slow thickness shear mode or C mode for plano convexe resonators using lateral field excitation and thickness field excitation are imaged using the white beam transmission topography on the figures 4 and 5 respectively. These resonators are designed to have a very high Q factor for this overtone (about 6.3 MHz). In fact the frequency is equal to twice that of the synchrotron radiation and these topographs are made with the stroboscopique technique.

With a low level of the excitation it appears extra Pendellösung fringes in the section topographs which correspond to elliptical fringes on the traverse topographs [14]. These elliptical

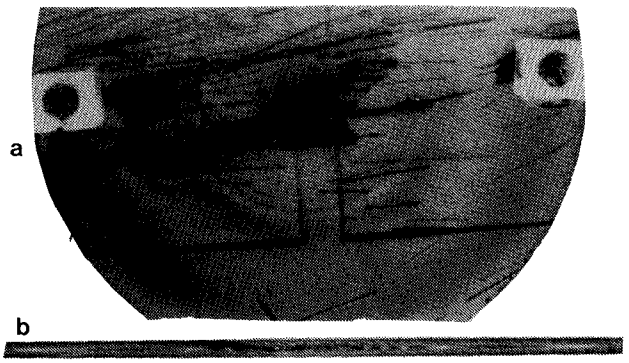


Figure-4- Traverse topograph(a) and section topograph(b) of the fifth overtones of the slow thickness shear mode for a plano convexe resonator using lateral field excitation with a low level of the excitation.

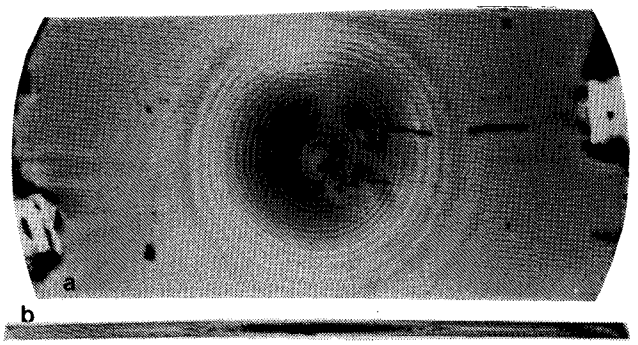


Figure-5- Traverse topograph(a) and section topograph(b) of the fifth overtones of the slow thickness shear mode for a plano convexe resonator using thickness field excitation with a low level of the excitation.

fringes result from the integration of the previous extra Pendellösung fringes over the whole resonators surface and constitute equi-level curves of the vibration amplitude. On the figure 5 the elliptical fringes are particularly well defined and present nearly the theoretical shape because this resonator have only three dislocations. Indeed these fringes are very sensitive to the presence of crystal defects.

This mode presents for these two resonators the same x_1, x_2 and x_3 dependence like it is possible to see it on the traverse topographs for the x_1 and x_3 dependence and on the section topographs for the x_2 one.

When the level of the excitation is higher the extra Pendellösung fringes on the section topographs disappear, but the antinodal planes of the vibration in the thickness of the resonator appear. The section topograph (Fig.6) of the same lateral field resonator presents four black lines corresponding to the four antinodal planes of the fifth overtone of the C mode.



Figure-6- The same section topograph that on the figure 4 but with a level of excitation higher.

-Parallel to the x_1 direction.

.B mode

It is possible to use another orientation of the lateral field to excite modes non piezoelectrically excited with a strictly thickness field. With the excitation field parallel to the x_1 direction we have excited the B mode. The topographs obtained with a diffraction vector parallel to the x_3 direction show on the figure 7 the u_3 component of different overtones.

The fundamental of the fast thickness shear mode (Fig.7a) is not trapped and presents a different lateral dependence.

The third and the fifth overtones (Fig.7b and 7d) are trapped and they have the same behaviour that the C mode. The section topographs (Fig.7c and 7e) show in the thickness of the resonator the two antinodal planes for the third overtone and the four antinodal planes for the fifth overtone. Although the polarisation of the fast thickness shear mode is not exactly in the x_3 direction, but at an angle equal to about $3^{\circ}5$, the topographs show a behaviour very close to that of a pure shear mode.

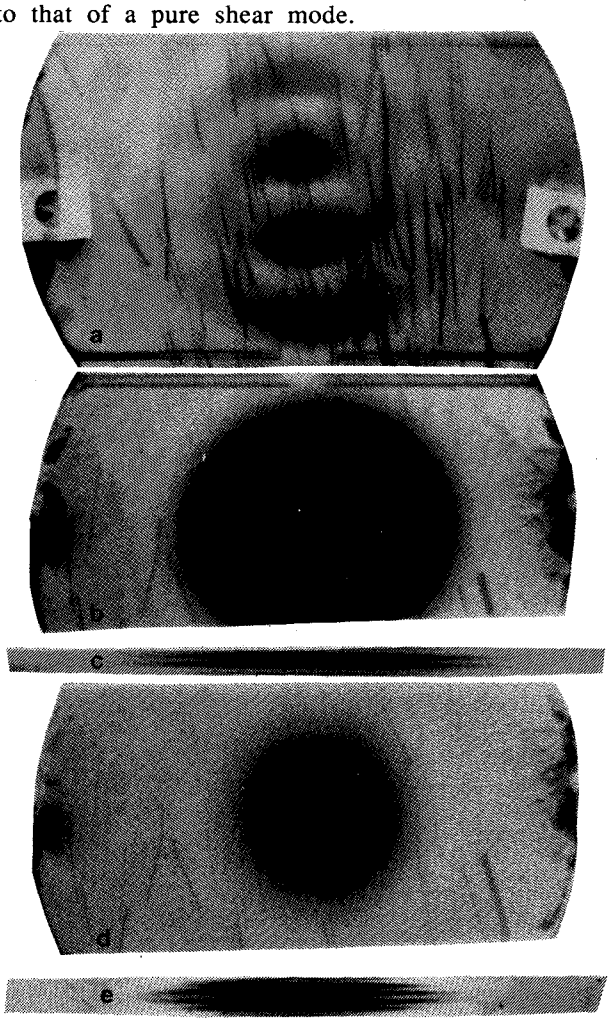
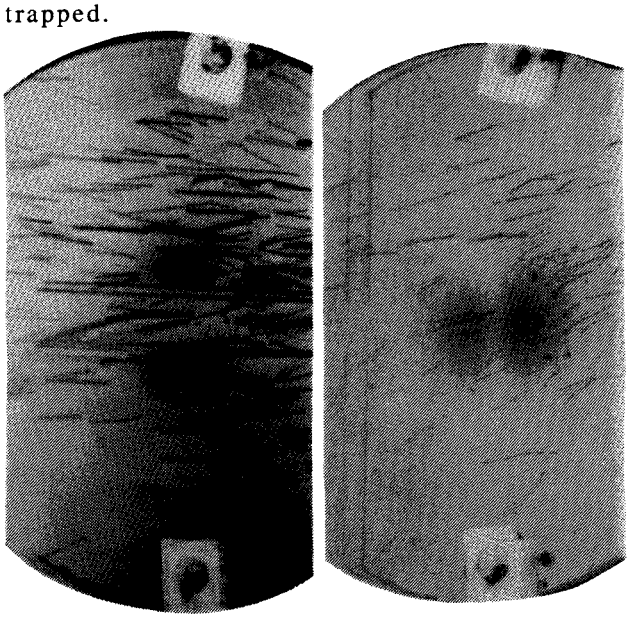


Figure-7- Topographs showing the u_3 component of different overtones of the fast shear mode.
a) The fundamental, b-c) the third overtone, d-e) the fifth overtone.

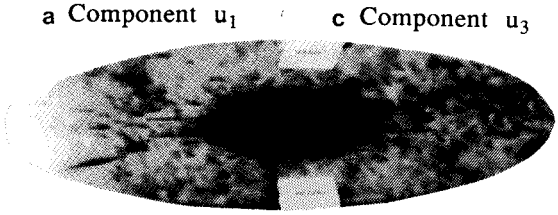
.A mode

The figure 8 shows three topographs obtained with a diffraction vector parallel to the x_1, x_2 and x_3 direction respectively, and displaying the three components of the third overtone of the A longitudinal mode which can be also excited by a lateral field parallel to the x_1 direction.

The topograph on the figure 8b obtained with the reflexion Laue set-up shows the main component u_2. This component seems to be trapped although a detailed examination indicates the existence of coupled plate modes which appear, on this type of topograph, like fine franges. The two others topographs on the figure 8 are obtained with the transmission Laue set-up. The components u_1 and u_3 which appear to be antisymetrical, are not negligible and are quite trapped.



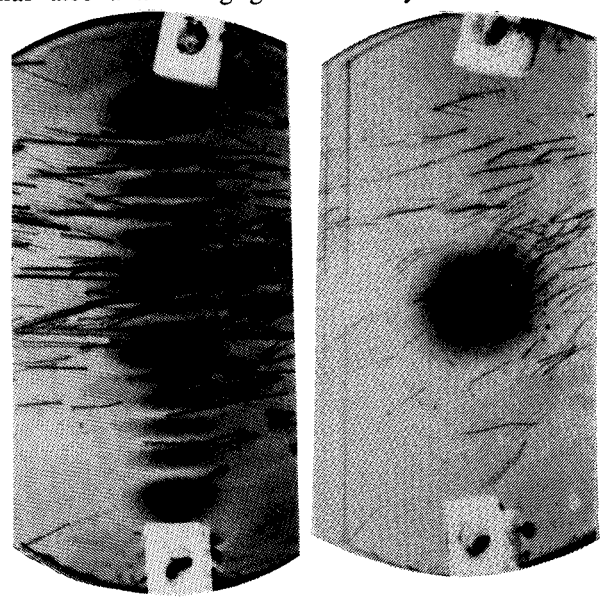
a Component u_1 c Component u_3



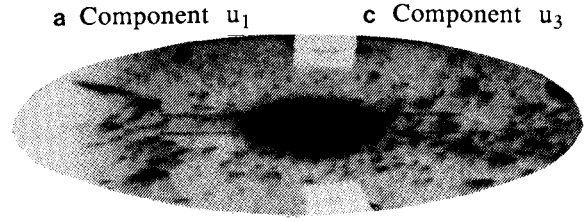
b Component u_2

Figure-8- Topographs obtained with a diffraction vector parallel to the x_1, x_2 and x_3 direction respectively, and displaying the three components of the third overtone of the A longitudinal mode.

Another example is shown on the figure 9. This is the fifth overtone of the A longitudinal mode obtained with the same conditions. Again the main component u_2 (Fig.9b) seems trapped in the center of the plate, but the energy is more confined that for the third overtone. The u_1 component (Fig.9a) extends up to the fixations in the x_1 direction and contains an important fraction of the acoustic energy. The u_3 component (Fig.9c) is confined in the center of the plate and has also a non negligible intensity.



a Component u_1 c Component u_3



b Component u_2

Figure-9- Topographs of the fifth overtone of the A longitudinal mode obtained with the same conditions that those on the figure 8.

.C mode:even overtone

The study of electrical response of a plano convexe resonator using the lateral field excitation parallel to the x_1 direction has shown the existence of anharmonics of the second overtone of the C mode. The figure 10 shows at the 2.517 MHz a strong response corresponding to the 210 C mode. This mode is imaged on the

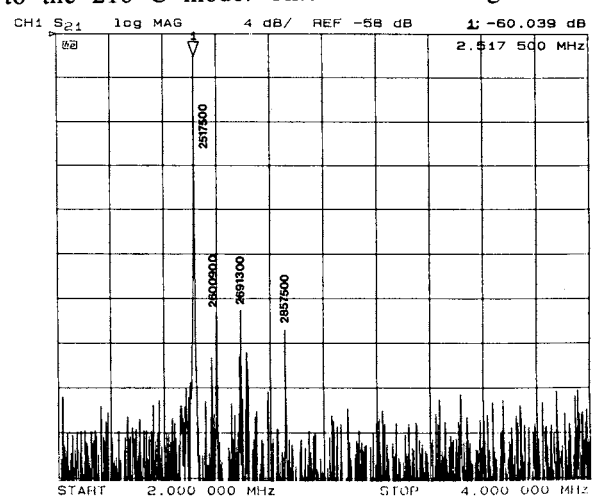


Figure-10- Electrical response corresponding to the even overtone 210 of the C mode, at 2.517 MHz, of a resonator using lateral field excitation parallel to the x_1 direction.

figure 11 by X ray topography using a diffraction vector parallel to the x_1 direction. On the traverse topograph (Fig.11a) we can observe the nodal line in the x_3 direction and on the section topograph (Fig.11b) the antinodal line identifying the second overtone.

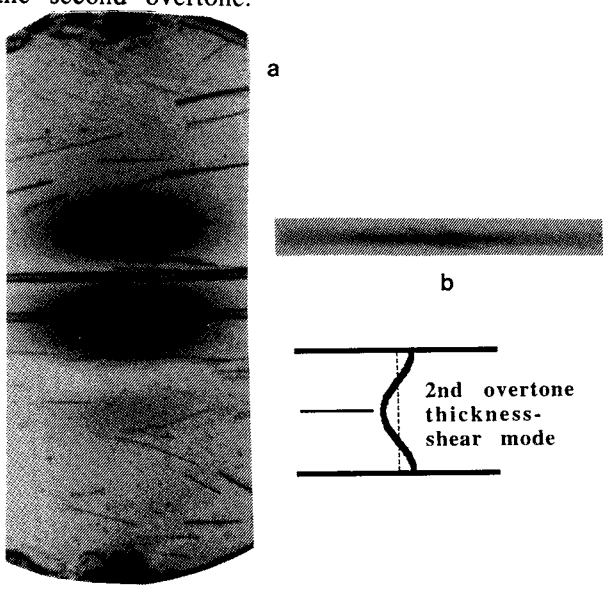


Figure-11- Traverse topograph(a) and section topograph(b) showing the even overtone 210 of the C mode of a plano convexe resonator using the lateral field excitation parallel to the x_1 direction.

Similarly we have observed in the same conditions the first antisymetrical anharmonic of the fourth overtone of the slow thickness shear mode (Fig.12). On the section topograph (Fig.12b) the three antinodal planes characteristic of the fourth overtone appear.

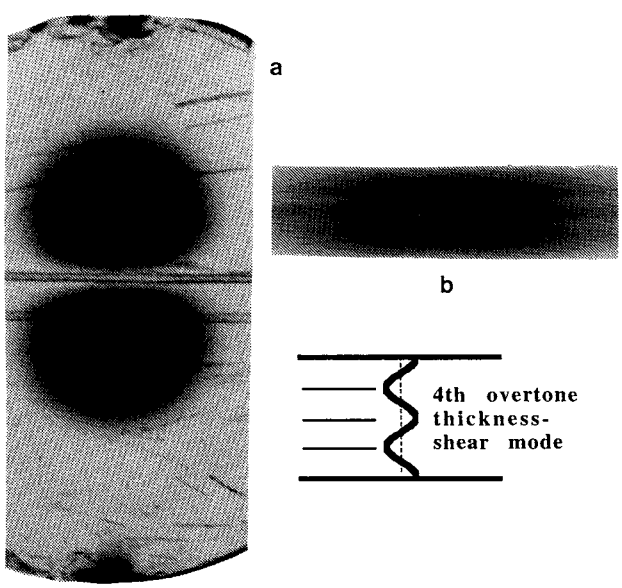


Figure-12-Traverse topograph(a) and section topograph(b) showing the even overtone 410 of the C mode of a plano convexe resonator using the lateral field excitation parallel to the x_1 direction.

Conclusion.

This has shown that it is possible to excite much more modes with the lateral field excitation that with the thickness field excitation. These modes can present intrinsic interest due to the values of the temperature coefficient and/or of the coupling coefficient. The new degree of freedom, given by the possibility of choosing the direction of the excitation field, permits interesting applications such as the suppression or excitation of one type of mode. This study can help to understand the coupled modes which appear in the thickness excitation resonators.

References.

[1] A. ZARKA, B. CAPELLE, J. DETAINT & J. SCHWARTZEL. 41st Annual Frequency Control Symposium, 27-29 May 1987, Philadelphia, U.S.A..

[2] A.GLOWINSKI. Annales des télécommunications, tome 27, n°3-4, Mars-Avril 1972, pp 147-158.

[3] T. YAMADA & N. NIIZEKI. J.Appl.Phys., Vol. 41, August 1970,pp 3604-3609.

[4] E.R. HATCH & A. BALLATO. Procedings of the Ultrasonics Symposium,Oct.31-Nov.1-2, 1983, p512.

[5] D.S. STEVENS & H.F. TIERSTEN. J.Acoust.Soc.Am. **79** (6), June 1986, pp 1811-1826.

[6] DULMET Thèse 1991 Université de Besançon.

[7] R.C. PEACH. IEEE Trans. on Ultrasonics, Ferroelectrics and Frequency Control, Vol.35, n°5, Sept. 1988, pp 593-611.

[8] J.DETAINT, H. CARRU, J. SCHWARTZEL, B. CAPELLE & A. ZARKA. Proc. 42nd AFCS, 1988, pp 19-28.

[9] W. IANOUCHEWSKY. Annales Françaises de chronométrie, 31ième année 2ième série, tome XV, 1961, pp 41-51.

[10] R. BECHMANN. Proc. 14th AFCS, May-June 1960, pp 68-88.

[11] A.W. WARNER. Proc. 17th AFCS, May 1963, pp 248-266.

[12] R.C. SMYTHE & H.F. TIERSTEN. Proc. 41st AFCS, 1987, pp 311-313.

[13] J. DETAINT, H. CARRU, R. LEFEVRE, J. SCHWARTZEL, C. JOLY, B. CAPELLE, Y. ZHENG & A. ZARKA. Proc. 44th AFCS, 1990, pp 337-348.

[14] Y. ZHENG, A. ZARKA, B. CAPELLE, J. DETAINT & J. SCHWARTZEL. Acta Cryst., 1989, A45, pp275-285.

IMPROVED PI-NETWORK MEASUREMENT OF ALL QUARTZ CRYSTAL PARAMETERS BY APPROPRIATE MODELLING OF THE TEST FIXTURE

Dipl.-Phys. B. Neubig

TELE QUARZ GROUP GmbH
Landstraße

Dipl.-Phys. L. Schulzke

KVG GmbH
Waibstadter Str.4

D-6924 Neckarbischofsheim

1. Abstract

The measurement of quartz crystal units is standardized in IEC 444. This means, that different test fixtures and additional elements must be used, e.g. for the determination of the load resonance parameters (IEC 444-4) or for physical C_0-compensation (IEC 444-3). The method to be presented in this paper introduces a general approach, which allows the precise determination of all quartz crystal parameters with one test fixture only by appropriate modelling :
- inductances in series to the inserted crystal modelling the electrical connection
- a capacitance between the PI-network ports modelling the crosstalk
- complex impedance of the PI-network.

All these elements are computed by a calibration routine with an open, a short and a calibration resistor in the complete frequency range. The quartz crystal parameters are determined by complex mathematical algorithms for the measurement data. Measurement results with different hardware will be presented in comparison to the standardized hardware based methods of IEC 444.

2. Introduction

The measurement of quartz crystal parameters is the objective of TC 49, Working Group 6 of IEC. The real problem is to find good correlation between standardized measurement methods and the requirements of the application in quartz crystal oscillators and filters, where the quartz crystal is object of circuit impedances (variable for frequency modulation) and drive levels.

Furthermore the equivalent circuit of the quartz crystal unit must be modified for some applications. In this paper the "standard circuit" is used with a series circuit of the dynamic capacitance C_1 and inductance L_1 and the series resonance resistance R1, shunted with a capacitance (static capacitance C_0 - measured at a low frequency [10 MHz ... 20 MHz] and parallel capacitance C_p - effective at the series resonance frequency f_s), given in figure 2.

The actual standard methods for the measurement of quartz crystals are given by IEC 444-1 to 444-4 with the use of a well defined PI-network, assumed to be purely resistive, and to cover the frequency range up to 200 MHz.

Due to new requirements of quartz crystal units and surface acoustic resonators (SAWR) with frequencies up to 1 GHz, new enclosure types (e.g. for SMT), and the calibration problems (especially for the reference resistors and the load capacitors), in the last years different new measurement methods were published and are in discussion for standardization (e.g. the s-parameter method or software based error corrected methods with the PI-network).

The aim of the approach given in this paper is to overcome the problems with the actual standardized methods by software solutions and to use further the existing hardware with one PI-network only.

3. Measurement Method

Definitions : f : frequency (Hz)
ω : circular frequency, $\omega = 2\pi f$

This measurement method complies with the passive transmission method according to IEC 444 with a PI-network.
The essential differences are :
- modelling of the PI-network with crosstalk capacitance C_C, series inductance L_S and complex termination T
- extended calibration routine for the determination of C_C, L_S and complex T
- definition of the parallel capacitance C_p at the series resonance frequency
- mathematical compensation of C_p in the complete frequency range instead of physical compensation with a tuned parallel circuit for frequencies above 80 MHz
- determination of the load resonance frequency f_L by an amplitude search instead of the use of physical load capacitors

3.1 Measurement circuit

The measurement circuit is given in figure 1 and complies with IEC 444

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

3.2 Model of the PI-network

The model is shown in figure 1, too, with :

T : complex termination resistance of the PI-network (IEC 444 : T = 25 Ω real)

C_C : crosstalk capacitance between the PI-network-ports

L_S : series inductance of the connection to PI-network ports

Typical value for C_C is tween 0.1 pF and for L_S between 0,2 nH and up to 20 nH (for a distance of 5 cm between the PI-network and the measurement plane). They should be determined at frequencies above 10 MHz. The influence of C_C and L_S cannot be neglected for the determination of C_O or C_P, for the load resonance frequency f_L and for all parameters of high frequency quartz crystals.
The complex resistance X, shown in figure 1, is calculated with:

$$X = \left(\frac{K}{V} - 1\right) * T$$

The resistance Z of the device under test (see figure 1) is computed with :

$$Z = \frac{1}{\frac{1}{X} - j\omega C_C} - j\omega L_S$$

3.3 Calibration of the PI-network

Definitions : $L = 2\pi f L_S$

$$C = \frac{1}{2\pi f C_C}$$

For precise measurements care should be taken for good measurement setup (low contact resistances, well defined double shielded cables, averaging of measurement data). Three measurements are taken with logaritmic frequency scaling in the interesting frequency range (e.g. 0.5 MHz to 200 MHz) :

- open ($Z = \infty$)
- short ($Z = 0$)
- with known calibration resistor (Z = R, complex, real part approx. 25 Ω)

giving the following results :

Z	V	X	calculation
∞	P	$X_P = -jC$	$= (\frac{K}{P} - 1) * T$
0	S	$X_S = +j\frac{C*L}{C-L}$	$= (\frac{K}{S} - 1) * T$
R	E	$X_E = \frac{L*C - jR*C}{R + j(L-C)}$	$= (\frac{K}{E} - 1) * T$

Column V gives the voltage ratio of the outputs of the measurement channel U_M and the reference channel U_R, defined as P,S, and E for open, short, and calibration resistor.

These three complex equations for the two complex variables K and T and the two real variables C and L can be solved principally, but not explicitely, for each frequency.

Eliminating T from these three equations results in two equations :

$$X_P - X_S = \left(\frac{1}{P} - \frac{1}{S}\right) * K * T$$

$$X_E - X_S = \left(\frac{1}{E} - \frac{1}{S}\right) * K * T$$

This can be reduced to one equation by eliminating K * T :

$$\frac{X_P - X_S}{\left(\frac{1}{P} - \frac{1}{S}\right)} - \frac{X_E - X_S}{\left(\frac{1}{E} - \frac{1}{S}\right)} = 0$$

Inserting the above resistances X_P, X_E and X_S gives for each measurement frequency the final complex equation with the unknown values C_C, L_S, the known (complex) reference resistor R and the measured voltage ratios P, S and E.

As this equation cannot be solved explicitely, the following algorithm is applied :

- calculate a first approximation for C_C with K = S , T = 25 Ohm and

$$C_C = \frac{1}{25*\omega*\mathrm{IMAG}\left(\frac{K}{P} - 1\right)}$$

- set a first approximation for L_S = 1 nH
- vary C_C and L_S in a double loop with diminuishing steps for C_C and L_S
- sum the squares of the complex equation error for each frequency above 10 MHz
- look for a minimum sum of the squares, resulting in the final values of C_C and L_S.

With the equations for the measurements with short and reference resistor the complex termination resistance T is found for each frequency :

$$T = \frac{X_S * S - X_E * E}{E - S}$$

T is then mathematically smoothed in both real and imaginary part by substituting their values with the second order parabola values calculated with the next five frequencies.

With the equation for the measurements with short the complex reference voltage ratio K is found for each frequency :

$$K = \left(\frac{X_S}{T} + 1\right) * S$$

K is then mathematically smoothed as given above for T.

3.4 Calibration of reference resistors

Excellent reference resistors with traced calibration data are commercially available e.g. from Wandel & Goltermann (W&G), Germany. Their model consists of an resistance in series with an inductance, shunted by a capacitance. Typical values are 25 Ω, 0.1 pF and 0.4 nH.

With a similar procedure as in clause 3.3 reference resistors can be calibrated together with the PI-network.

From the general equation of 3.2 the termination resistance T can be expressed by :

$$T = \frac{X}{\frac{K}{V} - 1}$$

Applying this formula to a pair of resistors X_1 and X_2 with the voltage ratios V_1 and V_2 and then eliminating K, the following equation is obtained for T :

$$T = \frac{X_1 * V_1 - X_2 * V_2}{V_2 - V_1}$$

Applying above formula to an other pair of resistances X_1 and X_3 with the voltage ratios V_1 and V_3 and eliminating then K an other equation is obtained for T :

$$T = \frac{X_1 * V_1 - X_3 * V_3}{V_3 - V_1}$$

Combining these 2 equations yields then :

$$\frac{X_1 * V_1 - X_2 * V_2}{V_2 - V_1} - \frac{X_1 * V_1 - X_3 * V_3}{V_3 - V_1} = 0$$

This equation can then be minimized with respect to the equivalent values of the calibration resistors E_1, E_2 and E_3 and the PI-network values of C_c and L_s. The calibration resistors are chosen as a short (E_1) and two resistors of about 25 Ω (E_2) and 75 Ω (E_3). Apart from the measurements with the calibration resistors an "open" measurement is performed.

The equation is solved with the following algorithm :
- define the equivalent circuit of the calibration resistors E_2 and E_3
- measure with a standard meter the static ohmic values of E_2 and E_3
- set the inductances of E_2 and E_3 to zero
- set the capacitances of E_2 and E_3 very small (< 1 fF).
- calculate a first approximation for C_c with K = S, T = 25 Ω and

$$C_c = \frac{1}{25 * \omega * \mathrm{IMAG}\left[\frac{K}{P} - 1\right]}$$

- set a first approximation for L_s = 1 nH
- vary C_c, L_s and the parameters of E_2 and E_3 in a multiple loop with diminuishing steps for these parameters
- sum the squares of the complex equation error for each frequency above 10 MHz
- look for a minimum sum of the squares, resulting in the final values of C_c, L_s and the parameters of E_2 and E_3
- calculate then T and K as in 3.3.

4. Equivalent circuit of a quartz crystal unit

The equivalent circuit is given in figure 2 with some definitions. The only difference to IEC 444 is the definition of the parallel capacitance C_p besides C_0, due to the fact, that the internal structure of the quartz crystal unit results in impedance transformations which cannot be neglected for precise measurements above about 30 MHz. Therefore C_p - effective at f_s - is no longer identical to C_0 measured at a lower frequency.

The complex resistance Z of a quartz crystal is given by :

$$Z = \frac{1}{\frac{1}{R_1 + j\left(\omega L_1 - \frac{1}{\omega C_1}\right)} + j\omega C_p}$$

5. Measurement of the equivalent parameters of a quartz crystal

The measurement is performed in five steps :
- static capacitance C_0
- parallel capacitance C_p
- coarse search of the series resonance frequency f_s
- at least two measurements close to f_s and determination of the series circuit parameters C_1, R_1, L_1, f_s, Q, Q_{eff}
- search of the load resonance frequency f_L (if applicable)

5.1 Measurement of the static capacitance C_0

At frequencies far from resonances the resistance Z of the crystal is given by:

$$Z = -j\frac{1}{\omega C_0}$$

The effective resistance X inserted in the PI-network is then :

$$X = \frac{1}{\frac{1}{\left(-j\frac{1}{\omega C_0} + j\omega L_s\right)} + j\omega C_c}$$

The measurement is performed at five equidistant frequencies between 10 MHz and 15 MHz, resulting in a matrix MC0.
The mean and spread values of C_0 are : C_{0mean} and $C_{ospread}$

The condition for validity of the C_0-measurement is :

$$C_{0spread} < 0.1 \text{ pF OR } C_{ospread} / C_{omean} < 0.03$$

If this condition is not satisfied, the non-valid values are eliminated by :

$$\mathrm{ABS}(C_{omean} - MC0) > 3 * C_{ospread}$$

and further appropriate measurements are taken.

Remark : These eliminations are necessary, when values are taken at frequenecies close to crystal resonances (other overtones or spurious modes)

5.2 Measurement of the parallel capacitance C_p

a) $f_N <= 30$ MHz :

$C_p = C_0$

b) $f_N > 30$ MHz :

Measurement at the frequency
$f_p = f_N - 1$ MHz, with $\omega_p = 2\pi f_p$:

$$Z = -j\frac{1}{\omega_p C_p}$$

$$X = \frac{1}{\frac{1}{\left(-j\frac{1}{\omega_p C_p} + j\omega_p L_s\right)} + j\omega_p C_c}$$

Correction after measurement of f_s and C_1 :

$$C_p = C_p - \frac{C_1}{1 - \left(\frac{f_p}{f_s}\right)^2}$$

5.3 Coarse determination of f_s : f_c

a) With network analyzer :
Fast sweep over the specified frequency range and search of the frequencies with maximum amplitudes :

f_{up} for sweep direction up

f_{do} for sweep direction down

$$f_c = \frac{f_{up} + f_{do}}{2}$$

b) With vector voltmeter :

f_N : nominal frequency of the quartz crystal

two measurements

$$f_N * \left(1 + \frac{1}{Q_{eff}}\right) \text{ and } f_N * \left(1 - \frac{1}{Q_{eff}}\right)$$

where Q_{eff} is an estimated value

calculation of f_c according to clause 5.4.

5.4 Measurement of f_s, R_1, C_1, L_1, Q, Q_{eff}

Measurement of the resistances Z_S of the series circuit at two frequencies f_x und f_y with the following presettings from other crystals of the batch or from standard values :

Q_{eff} = expected or estimated typical value

$$f_x = f_c * \left(1 + \frac{1}{2*Q_{eff}}\right)$$

$$f_y = f_c * \left(1 - \frac{1}{2*Q_{eff}}\right)$$

$$TIME = \frac{4*Q_{eff}}{\pi f_c}$$ rise time for oscillation of the quartz crystal

The resistance of the series circuit of the quartz crystal is :

$$Z_S = \frac{1}{\frac{1}{Z} - j\omega C_p} = R_1 + j2\pi f L_1 - j\frac{1}{2\pi f C_1}$$

Using the definition for the series resonance frequency f_s one obtains :

$$-j\frac{1}{2\pi f_s C_1} = -j2\pi f_s L_1 \quad \text{and}$$

$$Z_S = R_1 + j2\pi L_1 * \left(f - \frac{f_s^2}{f}\right)$$

For $|f - f_s| << f_s$ it can be written:

$$f - \frac{f_s^2}{f} = (f - fs)$$

and Z_S can then be expressed by :

$$Z_S = R_1 + j4\pi L_1 * (f - f_s)$$

The measured real and imaginary parts of Z_S at f_x and f_y are : R_X, I_X, R_Y, I_Y

$$R_X + jI_X = R_1 + j4\pi L_1 * (f_x - f_s)$$

$$R_Y + jI_Y = R_1 + j4\pi L_1 * (f_y - f_s)$$

From these equations calculations are done for :

$$R_1 = \frac{R_X + R_Y}{2}$$

$$L_1 = \frac{I_X - I_Y}{4\pi(f_x - f_y)}$$

$$f_s = \frac{f_x + f_y}{2} + (f_x - f_y) * \frac{I_X + I_Y}{I_Y - I_X}$$

$$C_1 = \frac{1}{(2\pi f_s)^{2*} L_1}$$

$$Q = \frac{2\pi f_s L_1}{R_1} \qquad Q_{eff} = Q * \frac{R_1}{R_1 + ABS(T)}$$

Correction of C_p due to the vicinity of f_c to f_s for $f_N >= 30$ MHz :

second calculations for R_1, L_1, f_s, C_1, Q and Q_{eff} with

$$C_p = C_p - \frac{C_1}{1 - \left(\frac{f_p}{f_s}\right)^2}$$

Condition for this measurement :

$$4 * \frac{Q_{eff}}{\pi f_s} < TIME \text{ AND } f_x - f_y < \frac{f_s}{Q_{eff}} \text{ AND}$$

$$[ABS(f_s - f_x) < f_x - f_y \text{ OR } ABS(f_s - f_y) < f_x - f_y]$$

If this condition is not satisfied, a further measurement is done with :

$$TIME = 6 * \frac{Q_{eff}}{\pi f_s}$$

$$f_x = f_s + \frac{f_s}{2Q_{eff}}$$

$$f_y = f_s - \frac{f_s}{2Q_{eff}}$$

5.5 Measurement of load resonance : frequency f_L and resistance R_L

Definition of the load resonance resistance R_L :

$$R_L = R_1 * \left(1 + \frac{C_p}{C_L}\right)^2$$

Definition of the load resonance frequency f_L :

At f_L the imaginary part of the impedance of the quartz crystal is oppositely equal to the reactance of the load capacitor C_L :

$$Z_L = R_L + j\frac{1}{2\pi f_L C_L}$$

The load resonance frequency can be calculated in a first approximation with :

$$f_L = f_s * \left(1 + \frac{C_1}{2*(C_L + C_p)}\right)$$

With this new approach a physical load capacitor is not used, f_L is determined by an amplitude search of Z_L.

In order to get an accuracy of 1 ppm for f_L of fundamental mode crystals, the amplitudes must be measured with an accuracy of 0.1 %, which requires a good measurement circuitry and careful calibration.

As the load resonance frequency f_L is far from the series resonance frequency f_S, the crystal impedance is very high and mainly reactive. Therefore the output power of the generator must be increased in order to give the required drive level P_N at f_L.

If the generator cannot deliver the necessary high output power for f_L, the measurements of both, f_S and f_L, must be repeated at an accordingly lower drive level, which can be realized at f_L . The difference between these two frequencies must then be added to the series resonance frequency at the nominal drive level in order to get the correct load resonance frequency. With this method, the effects of drive level dependence (DLD) of the crystal is encountered.

The output power of the generator is calculated as follows :
For series resonance :

- current through series circuit : $I_S = \sqrt{\frac{P_N}{R_1}}$
- voltage across crystal : $U_Q = I_S * R_1$
- current through C_p : $I_C = U_Q * j2\pi f_S C_p$
- current through crystal : $I_Q = I_S + I_P$
- voltage across C_c : $U_P = U_Q + I_Q * j2\pi f_S L_S$
- current through C_c : $I_P = U_P * j2\pi f_S C_C$
- current through PI-network : $I_T = I_Q + I_P$
- voltage across PI-network : $U_S = U_P + I_T * T$

For load resonance :

- current through crystal : $I_Q = \sqrt{\frac{P_N}{R_L}}$
- voltage across crystal : $U_Q = I_S * \left[R_L + j\frac{1}{2\pi f_L C_L}\right]$
- voltage across C_c : $U_P = U_Q + I_Q * j2\pi f_L L_S$
- current through C_c : $I_P = U_P * j2\pi f_S C_C$
- current through PI-network: $I_T = I_Q + I_P$
- voltage across PI-network : $U_L = U_P + I_T * T$

$$Delta_dBm = 20 * lg\left[ABS\left(\frac{U_L}{U_S}\right)\right]$$

6. Experimental results

To verify the theoretical evaluations, and to proof the expected accuracy improvement, 70 quartz crystal units at 14 different frequencies in the range 1 MHz to 150 MHz were measured at two different places (at the facilities of KVG and at those of TELE QUARZ [TQG]) with different measurement equipment and test fixtures, and with different software packages, which used different algorithms.

The KVG equipment used a network analyzer HP 3577, a coaxial PI-network with the traditional construction as depicted in IEC 444-1.
The PI-network set-up was modelled as described in clause 3 of this paper, i.e. with modelled series inductance L_S and crosstalk capacitance C_C as in Fig.1.

The TQG equipment consisted of a generator (PTS 160), a vector voltmeter (Rohde & Schwarz ZPV) and a PI-network in hybrid construction. The PI-network was modelled by a complex termination impedance R_T instead of modelling discrete circuit elements (as above). The crosstalk capacitance C_C is assumed to be directly across Z.

The results were compared and are depicted in the following graphs:

Fig. 3 shows the comparison of the measured resonance resistance values R_r (resp.the series resonance resistance R_1).

Fig.4 shows the deviation of the measured dynamic capacitance C_1.

Fig.5 shows the deviation of the static capacitance C_0.

Fig. 6 shows the differences in the measured (series) resonance frequency

The results of all these data shows an agreement in the order of +-1ppm for frequencies and a few percent for the other parameters.

The measurement of the load resonance frequency was performed in two different ways:
- measurement with physical load capacitors according to IEC report 444-4, where KVG and TQG were using their own capacitive elements, both calibrated according to their in-house standard procedures; and
- measurement with the "mathematical approach" given in clause 5.5

In Fig. 7 the results for a load capacitance of 30 pF ist shown in three different curves:
- comparison "phys-phys", i.e. using physical load capacitors at both places,
- comparison "math-math", i.e. using the mathematical methods at both places;, and
- comparison "phys-math", i.e. difference between "physical" and mathematical" measurement at one location (KVG).

Fig. 8 the same results are shown for a load capacitance of 10pF, which is extremely sensitive, especially for high fundamental mode frequency crystals.

7. Conclusion

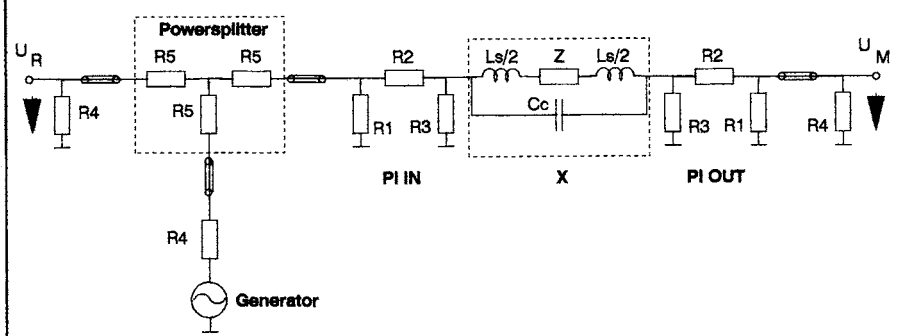
The graphs show quite clearly, that rather large frequency differences occur with physical load capacitors, which is an every-day experience of every crystal manufacturer and user. Contradictory to that, the frequency differences are reduces by a factor of 5 to 10, if the mathematical method is used.

The residual error resp. uncertainity with the mathematical methods comes down to a few ppm, only slightly more than the differences seen at the (series) resonance measurements.

These excellent results are achieved inspite of the fact, that both measurement methods were not identical in every detail, both, hardware and software.

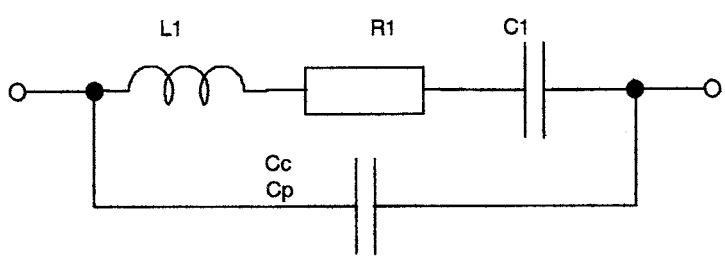
These results encourage us to propose the new method for implementation into the IEC standards. Regarding the good results of load resonance measurement, the new method will - and has to - replace the old IEC 444-4 report within short time.

Acknowledgements:
The authors wish to acklowledge R. Zimmermann (TQG) and H. Dimmler (KVG), who have done the experimental work and the evaluation of the measurement data.



R1: 159 Ohm, R2: 66.2 Ohm, R3: 14.2 Ohm, R4: 50 Ohm, R5: 50/3 Ohm
UR: Complex voltage in the reference channel
UM: Complex voltage in the measurement channel
V: Complex voltage ratio; V = UM / UR with resistance X
Reference voltage ratio; K = UM / UR with resistance X = 0 Ohm, short over PI - network ports

Figure 1: Measurement circuit



fN: nominal frequency of the quartz crystal unit
fs: series resonance frequency, fs = 1 / 2 / Pi / (L1 * C1) ^ 0.5
Q: Q - value, Q = 2 * Pi * fs * L1 / R1 = 1 / 2 / Pi / fs / C1 / R1
Qeff: effective Q - value, Qeff = Q * R1 /(R1+ABS(T))

Figure 2: Equivalent twoport circuit of a quartz crystal unit

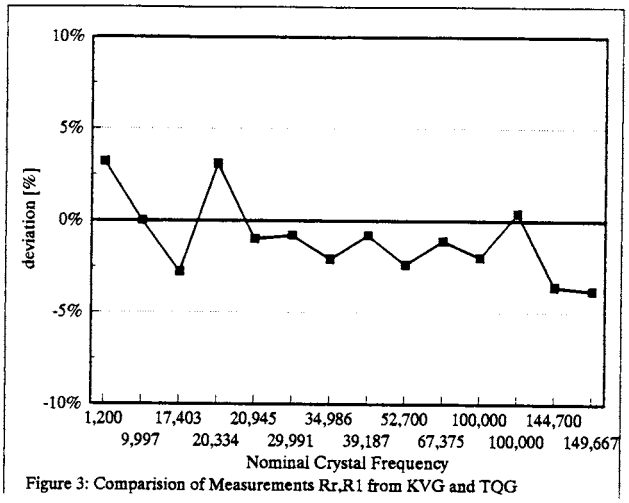


Figure 3: Comparision of Measurements Rr,R1 from KVG and TQG

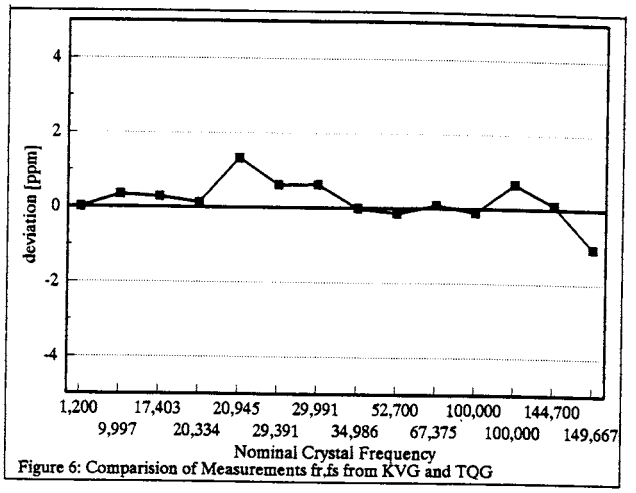


Figure 6: Comparision of Measurements fr,fs from KVG and TQG

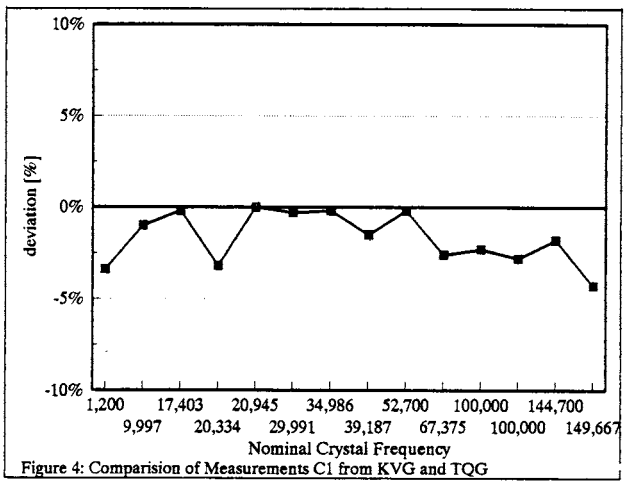


Figure 4: Comparision of Measurements C1 from KVG and TQG

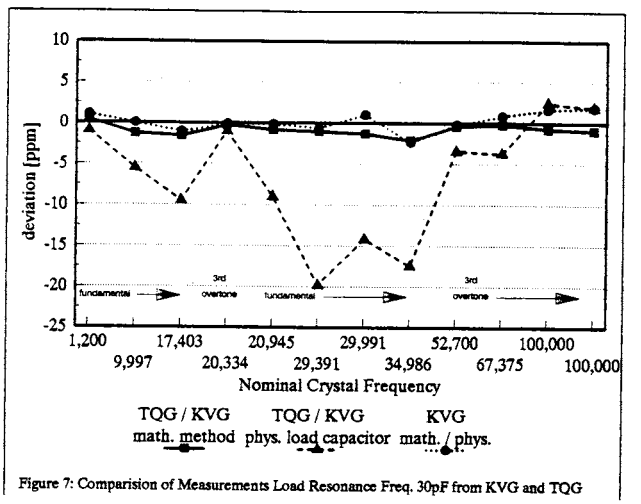


Figure 7: Comparision of Measurements Load Resonance Freq. 30pF from KVG and TQG

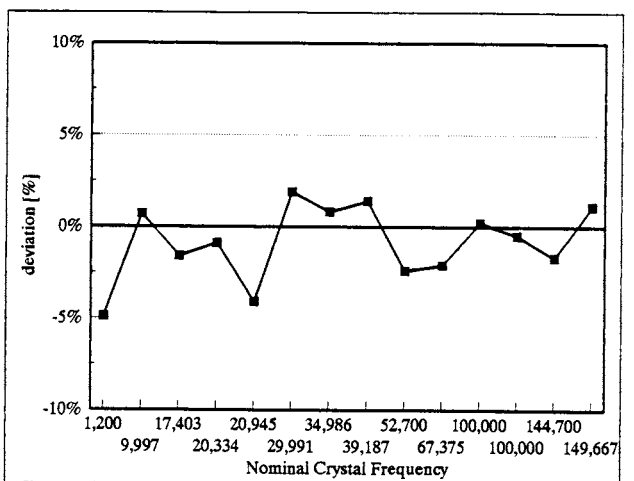


Figure 5: Comparision of Measurements C0 from KVG and TQG

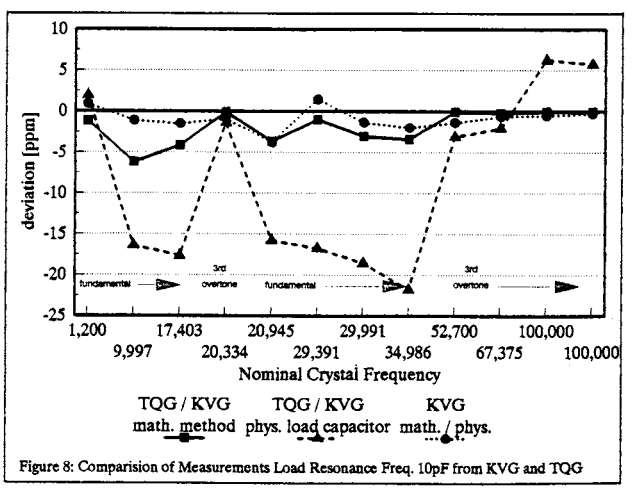


Figure 8: Comparision of Measurements Load Resonance Freq. 10pF from KVG and TQG

THE MEASUREMENT OF THE RESONANT FREQUENCY OF QUARTZ CRYSTAL UNITS AN ONGOING "SAGA"

ERIC KENTLEY

THE RODITI INTERNATIONAL CORPORATION LIMITED

'ABSTRACT/RESUME'

The measurement of the resonant frequency of a quartz crystal unit has always presented problems due to the fact that a quartz crystal unit is not an absolute frequency determining device. The frequency of oscillation of a quartz crystal unit in an oscillator circuit is a function of the crystal unit and the maintaining oscillator circuit.

Since the introduction of zero phase type measurements in a passive network by the I.E.C. in the early 1970's many of these measurement problems have been eliminated or much reduced, but there are still some problems area's. Using the computer aided test methods now available, automatic crystal measurements can be made. These can cause what might be termed computer aided correlation errors due to slightly different measuring proceedures and methods of computation.

This paper will briefly trace the history of practical methods of measurements used by industry, leading to the current situation. The types of problem still being encountered will be examined and possible approaches to eliminate or minimise the current problems will be suggested.

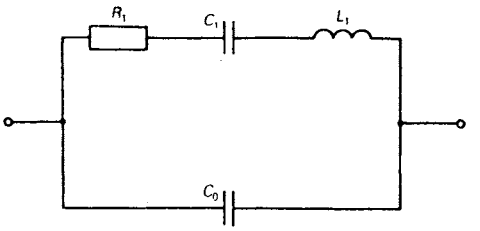
1) INTRODUCTION

Despite the fact that frequency is the most accurately measurable physical property there have always been problems associated with the measurement of the "resonant" frequency of quartz crystal units. This is because quartz crystal units are not absolute frequency determining devices, the "resonant" frequency is a function of the crystal unit and the maintaining oscillator circuitry. Many of the problems once associated with the measurement of the frequency of quartz crystals have been eliminated or much reduced since the advent of zero phase measurements in a pi network, standardised by the I.E.C. in the early 1970's. However some problems still exist and solutions to these problems are required to prevent disagreements arising between users and manufacturers.

2) THE EQUIVALENT CIRCUIT AND THE "RESONANT FREQUENCIES" OF A QUARTZ CRYSTAL UNIT

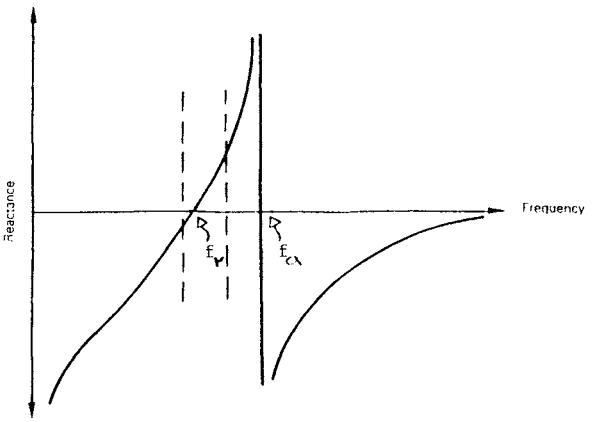
The simple equivalent circuit of a quartz crystal unit is shown in Fig 1 and the reactance versus frequency characteristics for this network is shown in Fig 2. From this it can be seen that there are two "resonant" frequencies where the reactance is zero and these are shown as fr and fa on Fig 2.

When used in an oscillator circuit the loop phase shift should be $2n\pi$ in order to maintain oscillation and if the oscillator circuitry introduces an undesired phase shift then the quartz crystal will operate at a slightly different frequency in order to introduce an equal but opposite phase shift to maintain the loop phase requirements.



Simplified equivalent circuit of a quartz crystal unit.

FIG. No 1



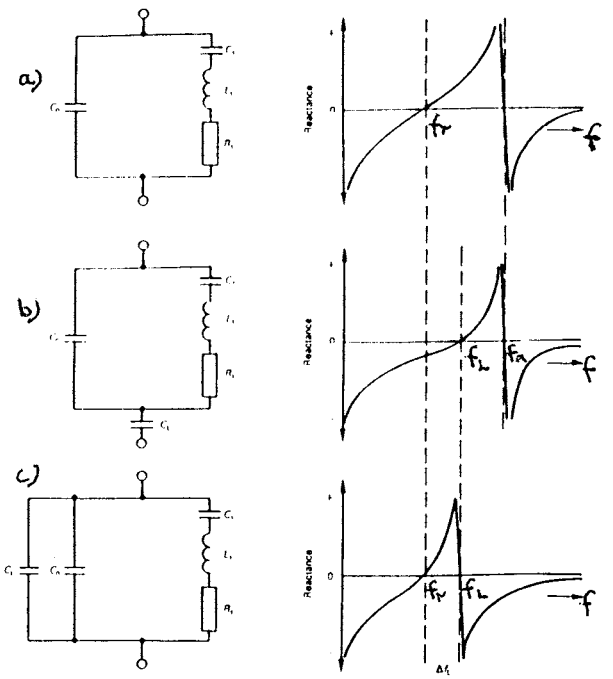
Reactance/frequency variation characteristic in the vicinity of resonance.

FIG No 2

In many oscillator applications the crystal is operated in conjunction with a load capacitance C_L and the resultant frequency f_L is known as the load resonant frequency (see Fig 3).

Actually the question of the frequencies associated with a quartz crystal is much more complex than shown above and is dealt with fully in IEEE 177 and IEC 302 (see Fig 4). From this it can be seen that there are 3 frequencies fm, fs & fr near the lower of the 2 resonances. Of these frequencies only fs is independent of the value of Co. Both fm and fr both depend on the value of Co and will generally be slightly different when the crystal unit is connected in a different fixture or circuit. The differences between fm & fs or fs & fr are a function of $r = C_0/C_1$ and Q.

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).



a). - Quartz crystal unit alone
b). - Quartz crystal unit with a series connected load capacitance C_L.
c). - Quartz crystal unit with a parallel connected load capacitance C_L

FIG. No 3

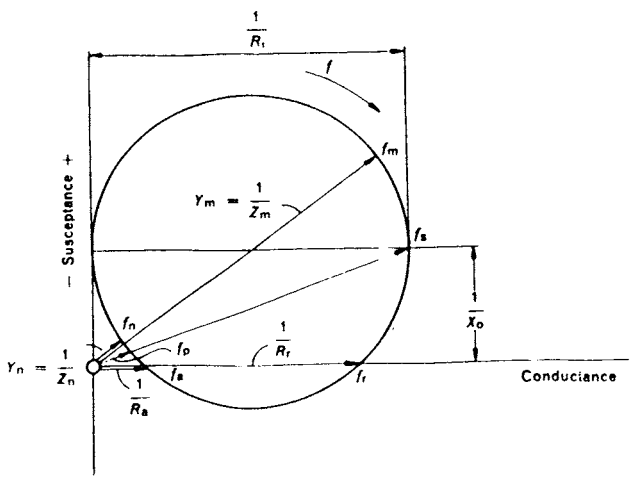


Fig. 4

Admittance diagram of a piezoelectric resonator.

For a typical fundamental mode crystal with r = 200 & Q = 50k, f = 4 x 10^{-8} which is of no consequence, but for an overtone crystal with r = 5000 & Q = 30k then f = 2.8 x 10^{-6}.

There are also 3 frequencies fa, fp & fn at the higher of the resonant frequencies but these frequencies are rarely if ever specified and are therefore unimportant.

Whilst quartz crystal units are now widely used components, they have tended to be ignored by much of the technical literature and are usually only dealt with very briefly in a typical electronic engineering degree course of studies. Because of this there has always been confusion and misunderstandings by many engineers regarding their useage. This state of affairs is not helped by the advent of "digital electronic engineers" who only seem to recognise two circuit states either "on" or "off" and to whom the concept of phase shift seems to be an unknown quantity.

3) BACKGROUND INFORMATION

When quartz crystal units were first marketed as electronic components, they were high priced, custom made devices and were finally calibrated to frequency in a replica of the user oscillator circuit.

The first large scale useage of quartz crystals was in the telephone industry in the mid to late 1930's and most of these crystals were used as "frequency selection" devices in crystal filters as opposed to frequency determining devices in crystal oscillators. The measurement of the resonant frequency was made in a passive network in which the crystal unit acted as a filter and the frequency measured was the frequency of minimum impedance fm. This method of measurement was later standardised by the I.R.E. in the U.S.A. as I.R.E. standard 14.51 (later IEEE standard 177 & IEC 302). This method of measurement remained the reference method until the advent of IEC 444 in 1972, but it was really a laboratory type measurement and hardly suitable as a production type measurement. (see Fig 5).

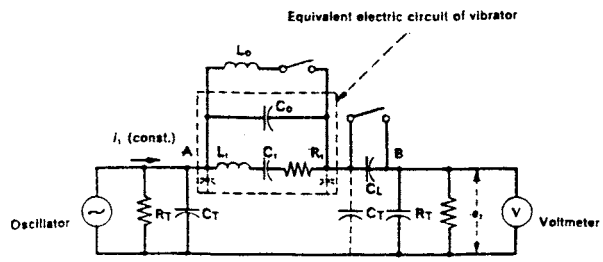


Fig. 5 Schematic of transmission circuit method.

For oscillator crystals, the major increase in useage occurred with the military requirements of W.W.2. and initially replica's of the end use oscillators were used in which to "calibrate" the crystals to the required frequency.

It became apparent to the military authorities both in the U.S.A. and the U.K. that this was an undesirable state of affairs and this lead to the development of standard crystal test sets known as crystal impedance meters (U.S.A) or crystal activity meters (U.K). (see refs 1 and 2). The "test sets" developed by the U.K. Military Authorities operated the crystal units at its load resonant frequency f_L, at a number of standard values (switchable) of load capacitance C_L. Those developed by the U.S. Military Authorities operated the crystal at its series resonant frequencies fs but a load capacitance C_L could be switched in series with the crystal in order to measure the load resonant frequency f_L.

In both types of test set the equivalent resistance of the crystal could be measured by substition of the crystal by a resistance or,in the case of the U.K. type of test set, a variable loss network.

If the crystal in a crystal oscillator can be replaced by a resistor which gives the same frequency & amplitude of oscillation, then the value of this resistor is equal to the value of the equivalent resistance of the crystal.

The U.K. developed test sets covered the frequency range from 8 KHz to 20 MHz and the U.S. test sets covered the frequency range 8 KHz to 250 MHz. These test sets could be used to give "frequency readings" accurate to +/-5 to +/-10 ppm and this degree of accuracy was adequate when the specified frequency tolerances were +/-50 or +/-100 ppm, as they were for the military type crystals being specified in the 1950's. All of the U.S. designed test sets were manually tuned to frequency and as such were prone to operator set up errors.

By the mid 1960's increasing numbers of quartz crystals were being used in civil vhf mobile radio applications and with reduced channel spacings, increased "accuracy" was required in the oscillators used in the transmitter and receiver sections of these radios. The situation was even more acute in the H.F. S.S.B radio systems then being used in marine and land communication systems. In these applications it became normal to specify a frequency tolerance at the reference temperature as well as a separate frequency tolerance over the operating temperature range. The frequency tolerance at the reference

temperature was usually trimmed out by the use of a trimming capacitor which was adjusted during the alignment of the equipment during manufacture. Adjustment tolerances of +/-20 ppm or less were regarded as normal for these applications and the +/-5 to +/-10 ppm errors possible using the military type test sets became a problem.

At about this time civil requirements for quartz crystal units started to become larger than the military requirements, and standardisation activities were taken up internationally by technical committee TC49 of the International Electro Technical Commission (I.E.C). A number of proposals were considered which included locked loop versions of the classical "pi network transmission system" according to IEEE 177, and bridge methods of measurement. Various bridge methods of measurements had long been used for laboratory measurements on quartz crystal units but these methods were unsuitable for production type measurements as the balancing of the bridge involved tedious adjustments by relatively skilled operators.

It has long been realised that a phase sensitive methods of measurement offered increased accuracy over amplitude sensitive methods as the slope of phase change/frequency is much greater than the slope of impedance change/frequency near the resonant frequency (see Fig 6). A.C bridges are in fact phase sensitve and this accounts for their approved accuracy compared to other methods of measurement.

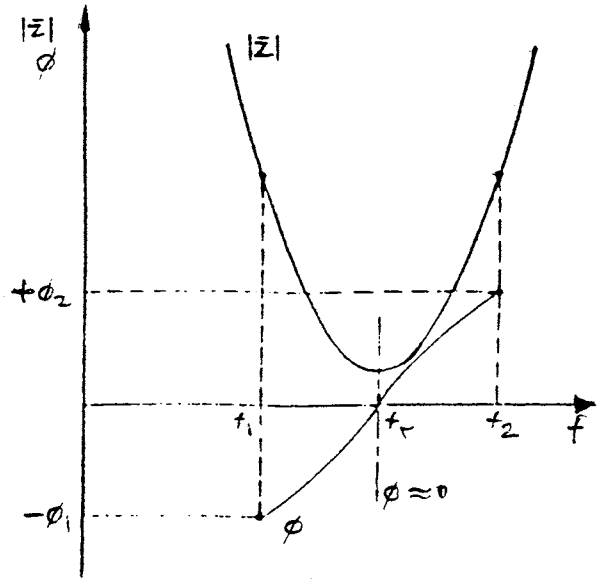
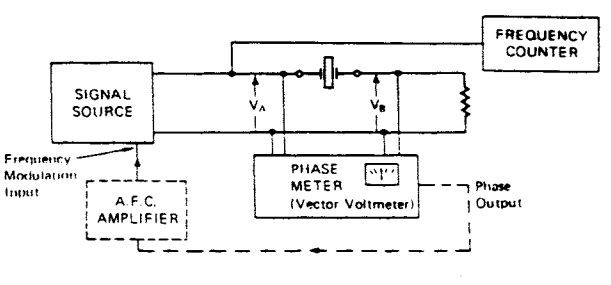


Fig 6
Phase & impedance change in the region of the resonant frequency fr for a quartz crystal unit.

Measuring systems using a low impedance pi network with zero phase detection were in use for low frequency crystals (below 1 mhz) using a oscilloscope as the phase detector. The appearance of a commercially available Vector Voltmeter covering the range 1-1000 MHz with phase measuring capability of +/-0.1° in the late 1960's was quickly used by the IEC TC49 to formulate a standard measurement method for crystal unit and this was published in 1973 as I.E.C. Publication 444. (see Fig 7).



A Simple Phase Zero Measuring System

FIG. No 7

This standard was quickly adopted in Europe, and was incorporated into a number of European National Standards (eg. B.S. & DIN). This did not occur in the U.S.A where the military standardistion authorities were pursuing the development of an "automatic bridge" crystal measurement system. This equipment never became available commercially.

The original issue of IEC Publication 444 claimed an accuracy of +/- 1 $x10^{-7}$ over the frequency range 1-125 MHz for the measurement of fr. In practice the accuracy was nearer to 1 x 10^{-6} but was a considerable improvement over the previously used measurement methods and has proved adequate for most applications.

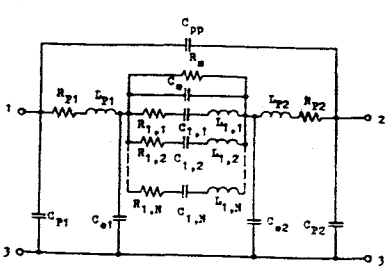
Also in the early 1970's a "new generation" of solid state C.I Meters became available commercially and these covered the frequency range 8 KHz to 250 MHz and carefully used, could correlate to about +/-3 ppm.

4) THE CURRENT SITUATION

The initial version of IEC 444 covered only the measurement of fr but a later version (IEC 444-4 1988) covered the measurement of f_L. This used a load capacitance "insert", which was inserted in the pi network, but these inserts never became available commercially and so this method of measurement has never become a useable international standard. IEC 444 basically covers the frequency range up to 125 MHz but with additionas to cancel the effects of Co, it can be used up to 200 MHz. With the advent of low priced personnel computers and bus programmable instruments, fully automatic measuring systems became available in the early 1980's. These systems can measure fr and Co and calculate the equivalent circuit values for C_1 or L_1. From this the value for f_L for any value of C_L can be calculated. Because the IEC 444 measurement uses a special measuring fixture (the pi network) there have always been some problems concerning traceability. Due to the high frequency limitations and the lack of traceability, work was conducted in the U.S.A. and the U.K. on a measuring method using an automatic network analyser and "S" parameter measurements. This method has been published as E.I.A 512 and work is currently underway in the IEC to produce an international version of this method of measurement. At present this method of measurement is only suitable for the measurement of fr but f_L can be calculated. This method can be used up to 1 GHz.

The major measurement problems encountered today are associated with the measurement of f_L. There are measurement methods (both C.I. Meter and pi network type) which use a physical load capacitance and all of these methods suffer from the "spread" of the actual value of the load capacitor as compared to its stated value. It is difficult to maintain these spreads to less than +/-0.25 pf and this can lead to frequency spreads in the order of +/-5 ppm. The currently used computer aided measuring systems which can calculate f_L suffer from 2 problems. These are that different systems use slightly different measuring methods and computations and these can produce differences in the computed frequency in the range of a few ppm.

A much more serious problem with these measuring systems is due to the fact the all calculated values for f_L assume that the crystal is a linear device with an equivalent electrical circuit as given in Fig 1. Unfortunately many crystal units exhibit a non linear behaviour as a function of crystal drive level and the equivalent electrical circuit is more complex, due to the presence of unwanted responses and coupled modes (see Fig 8). Differences of up to 45 ppm have been observed between the measured value of f_L (using a physical load capacitance) and the calculated value for f_L for crystal having unwanted responses (see ref 3.)



FIG, No. 8

The above problems are causing industrial problems between suppliers and users and an early solution to these problems could probably save industry millions of dollars per annum.

5) POSSIBLE APPROACHES TO IMPROVEMENTS

1) There has been a tendency over the years for the international standards bodies to develop "all encompassing" measuring systems (i.e. one method of measurement to cover the whole range of frequencies covered commercially by quartz crystal units). More progress might occur if particular attention is given to problem areas. For instance the problem of the measurment of f_L is in practice limited to the frequency range say 1 - 35 MHz.

2) Crystal users and manufacturers do not help themselves in minimising these problems by continuing to specify crystals in an unrealistic manner. Crystals are still being specified (& measured) with unrealistic drive levels of 500 uWatts (or higher) instead of a more realistic level of say 100 uWatts. These higher drive levels can cause problems by exciting non linear behaviour in the crystal. (see ref 4).

3) There has been a tendancy for the value of C_L being used in industry to "reduce" over the years. In the days when the military requirements for quartz crystals dominated the market, values for C_L were chose from a list of standard values. Values for C_L were standardised at 20pf and 30pf (32pf in the U.S.A) and this did not impose any undue problems on the circuit designers.

These values have now reduced steadily until we now commonly see values of 10pf or less.

In the case of the crystal controlled clock oscillator on a microprocessor chip the low value for C_L occurs because the crystal oscillator uses on board digital gate circuits which are not ideal for use as crystal oscillators. If the requirements for good circuit design for quartz crystal oscillators were appreciated by the I.C designers then a clock oscillator section of the I.C chip could be incorporated onto the chip at the design stage.

There are a number of applications for quartz crystals where it is necessary to achieve a wide pulling range. Examples of this type of application are T.C.X.O's and V.C.X.O's. In order to achieve a wide pulling range some circuit designers are using low centre values for the input capacitance. A better approach would be to specify the crystal unit for operation a the resonant frequency fr and to use a series LC circuit for pulling purposes. With this method by varying C the frequency can be pulled above and below fr. (see ref 5).

4) As was mentioned earlier there has always been confusion and misunderstandings regarding quartz crystal units. Individual quartz crystal manufacturers have invested much time and money in producing useful applications data in their commercial catalogues and a number of national and international guidance documents have been produced. These documents are aimed at the crystal user who was at one time also the circuit designer. Often this is no longer the case, as the circuit designer is now the I.C. manufacturer and laison between quartz crystal manufacturers and I.C. manufacturers is lacking.

There are internationally a number of companies which include both I.C. and quartz crystal manufacture in their activities and the required degree of laison should be relatively easy to implement in these cases but there are few indications that this has occurred so far.

5) For a number of reasons it appears that crystals which are to be used in a circuit with a load capacitance should be measured under the same conditions. This implies either the use of a physical load capacitance or the measurement of the crystal at the frequency where its reactance is equal to $-1/wC_L$ This method of measurement is mentioned in refs 6 & 7.

6) Whilst in practice there will always be a spread (ie. a tolerance) on the actual value of C_L present in particular samples of a given oscillator or between different examples of measuring equipment, it has never been a common practice to specify a tolerance on C_L. The load capacitance C_L has always been specified as an absolute value and the only tolerance has been placed on the frequency of the crystal unit. Some of this tolerance is required by the crystal manufacturer to cover tolerances dictated by the accurancy of manufacturing processes. Crystal manufacturers should be more realistic in apportioning specified frequency tolerances between those they require and those available to cover "spreads" of C_L.

The possibility of specifying crystals, for us in circuits where they operate with a load capacitance, by specifying an absolute frequency with the only tolerance being placed on the value of C_L should be examined closely.

6) CONCLUSIONS

There are still problems involved in the measurement of the frequency of quartz crystals and these cause disagreements between users and manufacturers.

These problems should be recognised by the international standards authorities and by discussions and agreements, methods to reduce the areas of disagreement should be sought.

7) ACKNOWLEDGEMENT

I would like to thank Northern Telecomm Europe Ltd, Quartz Crystals for their help in providing the transparancies and illustrations used in this article.

8) REFERENCES

1) Prichard A.C and M. Bernstein. "Crystal Impedance Meters" Electronics May 1953.

2) Biggs A.J. and Wells G.M. The measurement of the activity of quartz oscillator crystals. Proc IEE Part III,93,29 (1946).

3) Rose.D. "Load Resonant Measurements of Quartz Crystals" Proc 45th A.F.C.S 1991.

4) Saunders.J. "Drive level effects on transmission measurements of AT overtone crystals. Proc 6th Annual Quartz Crystal Conference. E.I.A. 1984.

5) Neubig, B. Technical Information:Design of Crystal Oscillator Circuits. Special edition of VHF Communications,3,4, 1979.

6) Horton,W., Payne, T. Smyth, R. and Symonds D. "Comparison of Methods for Measurements of Quartz Crystal Resonators with load capacitance. Proc 35th A.F.R.S. 1981.

7) Neubig, B. Pi-Network Measurement of the load resonance frequency without load capacitor. 4th E.F.T.F Neuchatel 1990.

Session 4B

Time and Frequency: Space Applications

Chairman: P. Hartl

DORIS PRECISE ORBIT DETERMINATION AND LOCATION SYSTEM PERFORMANCES OF ULTRA STABLE OSCILLATORS

M. BRUNET

CENTRE NATIONAL D'ETUDES SPATIALES (CNES)
18 Avenue E.Belin
31055 TOULOUSE, FRANCE

ABSTRACT

The Centre National d'Etudes Spatiales (CNES), the French Space Agency, has designed and developed the DORIS system, to meet new needs in precision orbit determination and high accuracy beacon location. DORIS payload has been integrated on three french SPOT satellites, and on the TOPEX-POSEIDON NASA satellite. The first model DORIS-SPOT 2 is operating since 22 January 1990. The fundamental measurement precision depends strongly on the stability of ultra stable oscillators (U.S.O.) which are used in the on-board receiver, and in ground location beacons.

The paper describes elements of the system, the DORIS-SPOT 2 location performances, and performances of DORIS USO s.

1. SYSTEM DESCRIPTION AND CHARACTERISTICS

The DORIS system (figure 1) is based on the measurement, by a satellite-borne receiver, of the spacecraft velocity relative to various beacons: ODBs (Orbit Determination Beacons) and GLBs (Ground Location Beacons). Every beacons transmits very stable frequencies: 2.03625 GHz and 401.25 MHz. The last one is also used to transmit housekeeping and meteorological data. Omnidirectional antennas are used for the beacons and the receiver. The receiver processes one beacon: it measures, every 10 seconds, the Doppler shifts on both frequencies, and also collects the transmitted data. This data is stored in the receiver telemetry memory and dumped twice a day to the ground.

A control centre preprocesses the telemetry, monitors the behaviour of the receivers and the beacons. It also computes the "work program" for the following days, to tell the receiver which beacon is to be received, and when. This program is sent every day to the receiver through the Master Beacon, linked to the control centre. The Master Beacon is also slaved to the international atomic time (TAI) which gives the possibility to determine the time-tagging for all measurements in the same time scale.

The ODB network is 40-50 beacons (figure 2). Their coordinates were known with a variable accuracy, depending on the possibilities of colocation with other precise positioning systems. They have been adjusted to 10 cm by processing of the first DORIS data. ODBs operate in a "laboratory" environment.

GLBs are functionally identical to ODBs. Each GLB transmits for 10 s, 20 s, or 30 s every minute, but only while the satellite is within range. GLBs are packaged for field use in scaled housing with their own power (generators, solar panels, battery, etc.).

2. DORIS-SPOT 2 PERFORMANCES

2.1 Instrument performance

It is specified in term of noise (rms on 10 s measurements) and medium term (mainly the slope during a satellite pass, up to 10 minutes, i.e. 60 measurements). The long term stability does not matter, since the average value of frequencies is identified during processing.

Noise: the typical value of measurement noise is 0.3 mm/s or 1.10^{-12}. Theoretical values, from on-board and ground oscillators, thermal and spurious noises, quantization and other instrumental errors were computed to 0.2 to 0.9 mm/s, depending on satellite elevation and spurious level. This noise is easy to evidence in the orbit residuals, since it has only high frequencies, and may be isolated by filtering. Some very quiet areas show a 0.25 mm/s rms value on one pass. The average value for a 2-day arc is 0,4 - 0,5 mm/s.

Medium term instrument errors are due mainly to the natural noise of oscillators (Allan variances of DORIS USO vary from 5.10^{-13} over 10 s to 1.10^{-12} over 10 minutes) and to tropospheric propagation mismodelling during the satellite pass. Vertical tropospheric values are about 2.4 meters, estimated with an accuracy of 2-5 cm from the three meteorological ground sensors. Globally, medium term error was assumed to be, at first order, a slope, rms 5.10^{-13}/min. This slope is random from pass to pass. It is not possible, at present, to confirm this value by analysis of orbit residuals, because medium term errors look like the orbital non-modellized signal, and are included in it. Trials to identify the slopes give value in the -30 to +30.10^{-13}/min range, but may be still correlated with the orbit determination conditions ; one may only conclude that the instrument contribution is not significantly worse than the expected value, but it is probably better.

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

Figure 3 from [3] gives the short term noise on individual 10 seconds measurements over 7 minutes pass, using the state of the art "GUIER" filtering technique which removes the best "orbit-like" Doppler curve. The solid line shows the typical rms value of 0.3 mm/s or 2 mHz on the 2GHz channel.

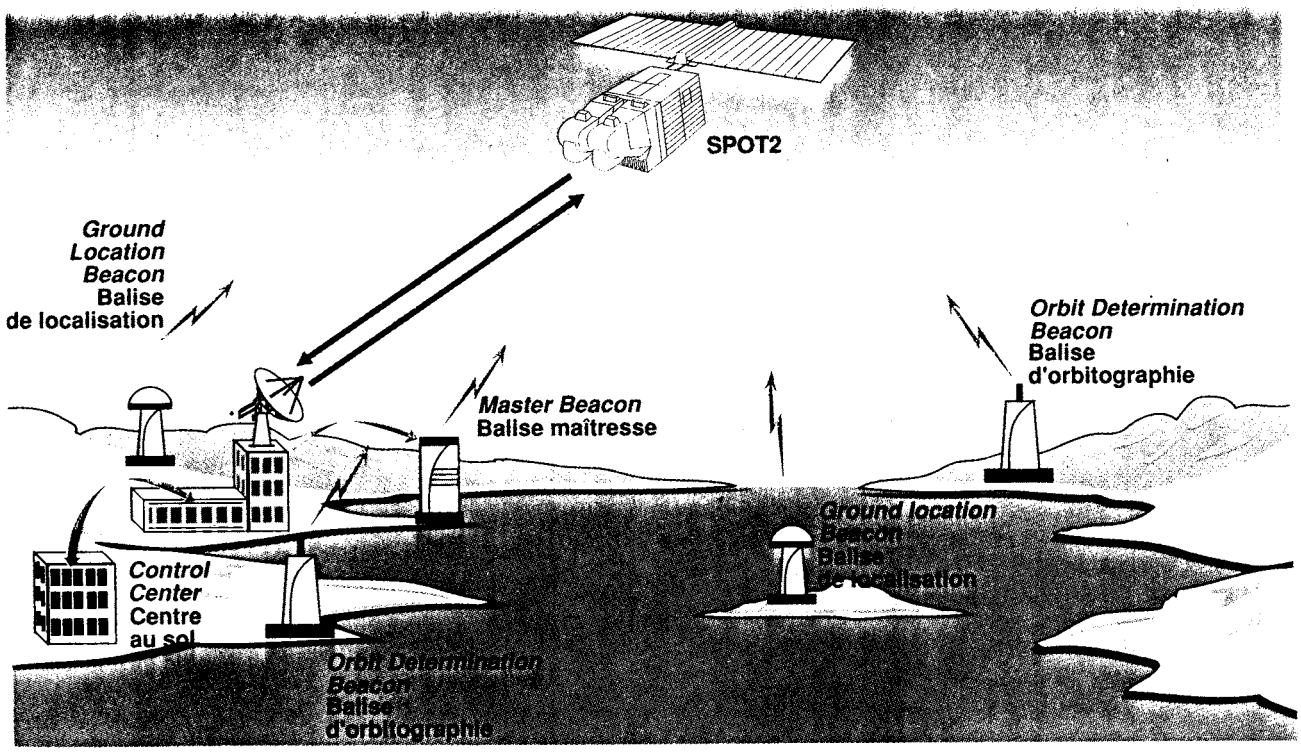


Figure 1. The DORIS system

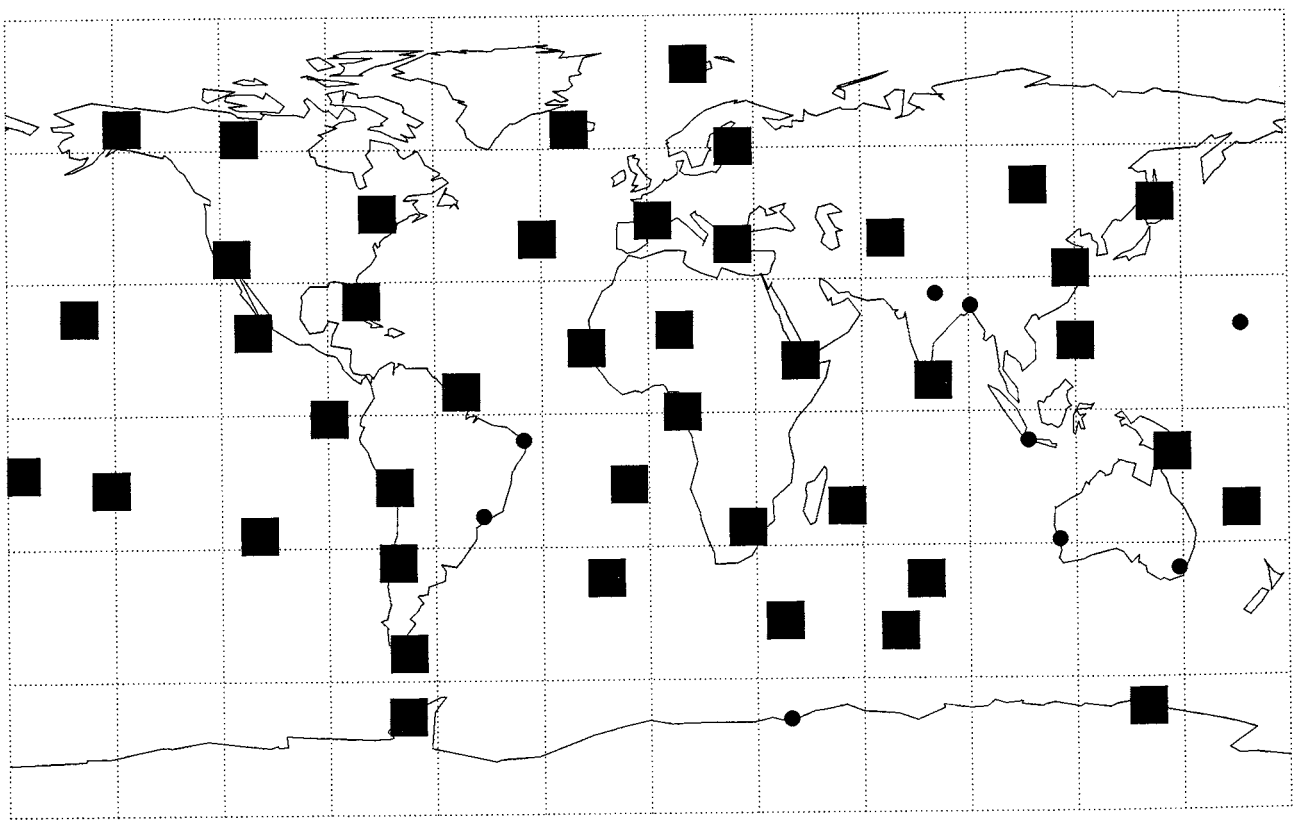


Figure 2. The network of DORIS orbit determination beacons

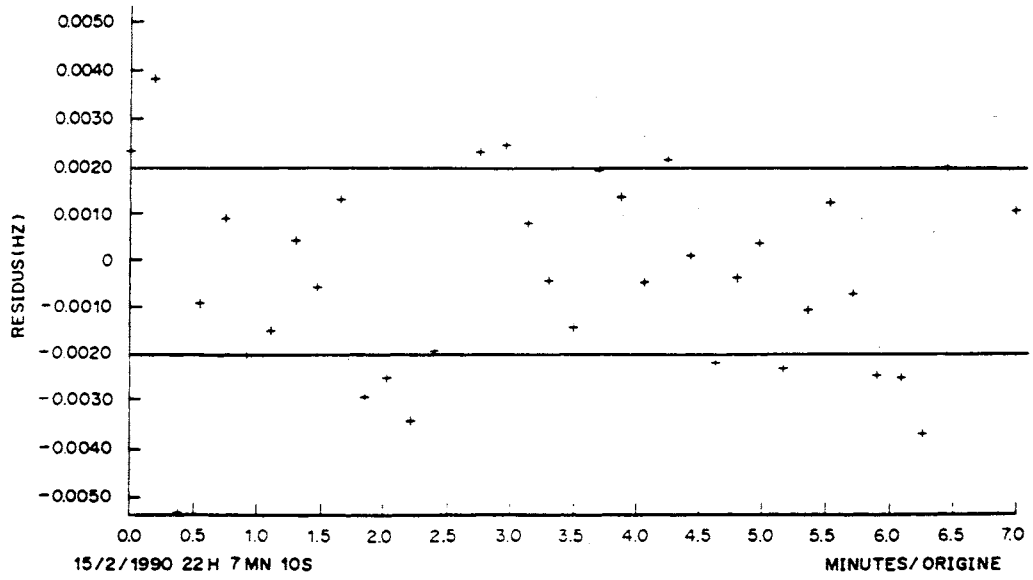


Figure 3. Results of "Guier" filtering: short term noise.

2.2 DORIS orbit determination performances

Orbit determination was actived by the CNES-Orbit Meteorology Department with the help of GRGS as soon as the system started up, and the first meter accuracy orbit was achieved two weeks later. The quality criteria were based on the residual measurements errors. This requires calculation of the mean least-squares discrepancy between the validated and theoretical measurements, i.e. those which would have been obtained from the as-calculated orbit. Note, as a guide, that a residual error of 2,5 mm/s is equivalent to 20 cm accuracy on the satellite-station distance.

As expected in the initial phase, the error is mainly due to uncertainty on force model. We are not merely concerned with measurement noise: part of the signal is still to be modelled. Figure 4 from [1] shows the evaluation of the residual errors according to the state of advancement of the project. Along-track residuals form another quality criteria. The complete set of measurement for a given pass can be compressed onto a particular point, and the orbit advance or delay calculated relative to it. The average value is 0.5 m.

Evaluating the accuracy obtained in other directions is more difficult. However, using overlap techniques, the mean accuracy of the satellite altitude can be estimated to 10 or 20 cm. The progress expected from the final gravity field models, plus better modelling of surface forces on the satellite, will ensure that 10 cm accuracy performance is achieved.

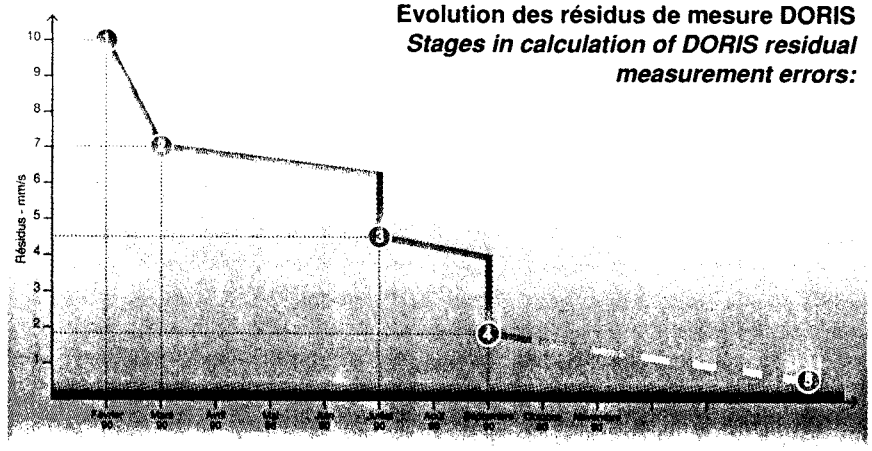


Figure 4. Stages in calculation of DORIS residual measurement errors.

(1) first DORIS computations (with GEM-T2 gravity field model)
(2) better modelisation of air drag
(3) with use of a new gravity field model (PGS3520)
(4) with use of a preliminary model, the GRIM4, calculated in August 1990, including three months of DORIS measurements. A set of preliminary recalculated coordinates is also used.
(5) estimated limit of residual errors: cumulative effect of measurement noise and medium-term errors (oscillator, propagation).

2.3 DORIS ground location performance

Although DORIS was originally designed to accurately locate satellites, on their orbits, it also lends itself perfectly to the accurate location of ground beacons. Terrestrial location has two purposes:

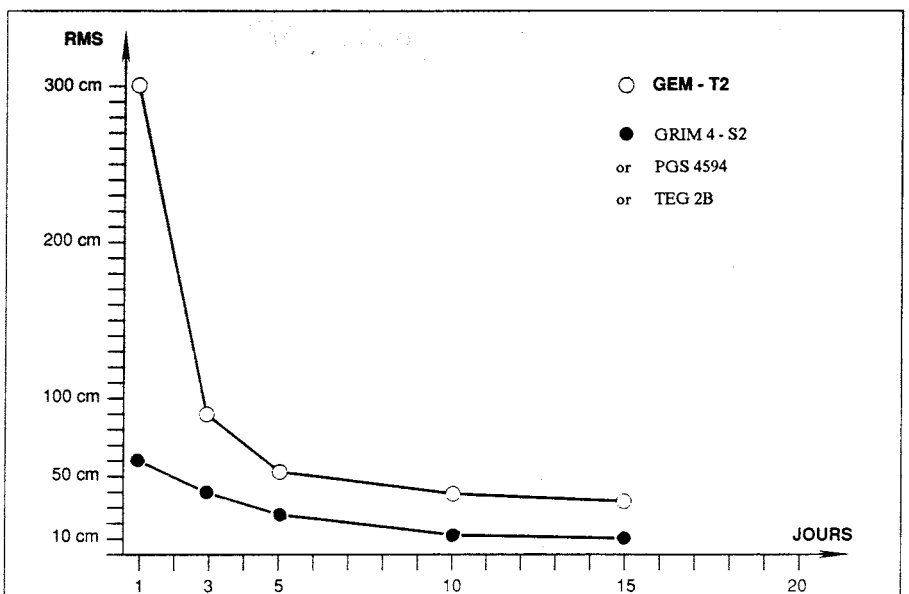
- accurately determine the coordinates of orbit determination beacons (O.D.B.s) and measure variation versus time,
- locate users' Ground Location Beacons (G.L.B.s), individual or networked, for geodetic applications.

Location results can be compared and merged with those from other tracking system such as laser system or GPS.

G.L.B.s are used in various scientific and technological applications of absolute and differential location, whether ad hoc or ongoing. See references [1] and [2]. To qualify location performance for the mission of DORIS-SPOT 2, pilot experiments are being conducted in Djibouti, Hawaï, and Saint-Etienne de Tinee (France).

Figure 5 from [2] shows the variation in absolute precision, according to gravity model and number of days of measurements acquired.

Differential precision is estimated from the baselines obtained in the Djibouti and Hawaï experiments. It varies according to the distance and observation period: figure 6 from [2].



Evolution de la précision de répétitivité de la localisation absolue en fonction du modèle de champ de gravité et en fonction de la durée de cumul des mesures.

Trend in absolute location precision according to gravity model and number of days of measurements acquired.

6

Figure 5. Trend in absolute location precision

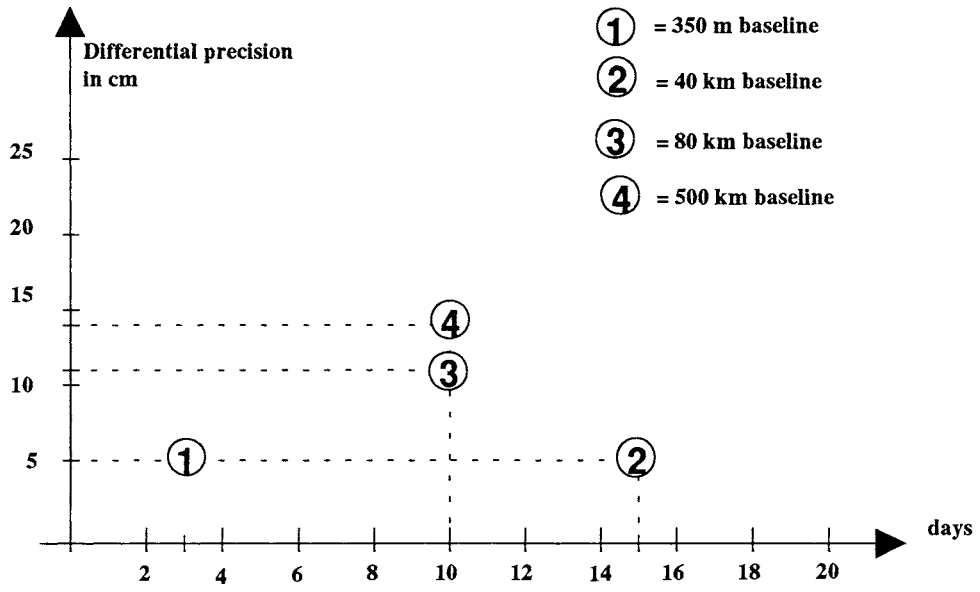


Figure 6. DORIS differential location precision

Absolute accuracy was evaluated by comparing twelve DORIS beacons, colocated with laser or VLBI, and IERS instrumentation. The RMS residuals after a seven-parameter transform are less than 15 cm. Differential accuracy has so far been calculated infrequently, and only over short distances. Over 350 m in Toulouse, the distance calculated over two days is within a centimetre of the terrestrial geodetic measurement. A 15 km baseline at Djibouti was compared with a direct laser geodimeter measurement. The results from two weeks of data match to within 2 centimeters. Table 1, from reference [2], shows absolute location accuracy comparison between several positioning systems, and DORIS. The DORIS system, especially for absolute positioning, can be favourably compared to other systems. It also provides an end-to-end processing and quality control.

LOCALISATION ABSOLUE
ABSOLUTE LOCATION ACCURACY

SYSTEME - *SYSTEM*	EXACTITUDE *ACCURACY* (cm)	DUREE D'OBSERVATION (jours) MEASUREMENT DURATION (days)
GPS Navigation avec SA *GPS Standard Positioning Service with Selective Availability*	10 000	0.1
GPS Navigation *GPS Standard Positioning Service*	2000	0.1
Transit operationnel *Transit Operational Positioning*	200	15
GPS Précis *GPS Precise Positioning Service*	100	8
Transit précis *Transit Precise Positioning*	80	8
DORIS Eclair *DORIS Fast-Delivery Product*	50 - 100	1
SLR 2ème génération *2nd-generation Satellite Laser Ranging*	20	90
DORIS *DORIS Off-Line Product*	10	15
VLBI S/X M2 *Very Long Baseline Interferometry S/X M2*	5	10
GPS Ultra-précis *GPS Ultra-Precise positioning*	5	8
SLR 3ème génération *3rd-generation SLR*	3	60
VLBI S/X M3 *VLBI S/X M3*	< 2	3

Table 1. Absolute location accuracy comparison between several positioning systems and DORIS

3. DORIS USO SPECIFICATIONS AND PERFORMANCES

3.1 Doris Uso Specifications

The main characteristics required for the on-board and ground DORIS USOs are defined in the following Table 2.

Parameters	O.B. USO	O.D.B. USO	G.L.B. USO
Frequency	10 MHz	5 MHz	5 MHz
D.C. Supply variation	$\pm 1.10^{-11}$ (1 %)	$\pm 5.10^{-11}$ (10 %)	$\pm 5.10^{-11}$ (10 %)
Load variation	$\pm 1.10^{-10}$ (10 %)	$\pm 2.10^{-11}$ (10 %)	$\pm 2.10^{-11}$ (10 %)
Pressure variation	$\pm 1.10^{-8}$/bar	$\pm 1.10^{-12}$/mb	$\pm 1.10^{-12}$/mb
Acceleration variation	$\pm 1.10^{-9}$/g	10^{-9}/g	10^{-9}/g
Vibrations (sine, random)	$\pm 5.10^{-9}$	-	-
Magnetic variation	$\pm 5.10^{-11}$/gauss	-	-
Temperature variation	$\pm 2.10^{-10}$ (air, vacuum 0 to 40°C)	$\pm 2.10^{-10}$ (0 to 60°C)	$\pm 1.10^{-10}$ (-20 to 60°C)
Short term stability $\tau = 1$ s	5.10^{-13}	5.10^{-13}	5.10^{-13}
(ALLAN variance) $\tau = 10$ s	5.10^{-13}	5.10^{-13}	5.10^{-13}
DORIS mid Term Stability with $\Theta = 0.1°$ C/mn	0 to 40° C	0 to 60° C	- 20 to 60° C
R.M.S. SLOPE	$\pm 4.10^{-13}$/mn	$\pm 4.10^{-13}$/mn	$\pm 4.10^{-13}$/mn
Residual (1 σ)	1.10^{-12}	1.10^{-12}	1.10^{-12}
Long term Aging after 1 month			
1 day	$\pm 1.10^{-10}$	$\pm 3.10^{-11}$	$\pm 3.10^{-11}$
1 month	$\pm 1.10^{-9}$	$\pm 5.10^{-10}$	$\pm 5.10^{-10}$
Phase Noise dBC/Hz			
10 Hz	-120	-130	-130
1 kHz	-140	-145	-145
D.C. Power			
air, 0° C	6 W	6 W	6 W
Vacuum, 0° C	3,5 W	-	-
Power supply	20 VDC	24 VDC	24 VDC
Random vibration	18 grms	-	-
Sine vibration	20 g - 10 to 100 Hz	2 mm 5 - 25 Hz	2 mm 5 - 25 Hz
Radiations	TOPEX	-	-

Table 2. DORIS USO SPECIFICATIONS

USO frequencies are 10 MHz for space and 5 MHz for beacons. Frequency stabilities versus DC supply, load, pressure, acceleration are very classical. Frequency stability versus magnetic field is necessary for space USO, due to two components:

- the magnetic field generated by the Earth, when the satellite is in low orbit (SPOT, TOPEX...),
- the magnetic field generated by magnetic torquers of the spacecraft, depending on the distance between USO and torquers (2 to 4 gauss peak to peak in the case of SPOT).

The very important characteristics of USO for DORIS are:

- The short term stability empressed in Allan variance: $\leq 5.10^{-13}$ over $\tau = 10$ s, duration of the on board Doppler count.
- The DORIS Medium Term Stability, defined by the Root Mean Squared Slopes (R.M.S.S.) and Residuals at 1 sigma, of 90 samples of frequency measurements made over 10 seconds, every 10 seconds. (R.M.S.S.) and R(1σ) are calculated by the least squared method, and must be less than 4.10^{-13}/mm and 1.10^{-12} respectively. Each USO must be compliant to these specifications over all environmental conditions: temperature, magnetic field, radiation in orbit, and temperature, pressure, humidity on the ground.

The long term frequency specification is very classical for the O.B. USO (1.10^{-10}/day) ; but we need for beacons a better performance (1 to 3.10^{-11}/day), due to the fact that the beacons transmit their messages to satellite in time-sharing.

Of course, O.B. USOs must satisfy all environmental space conditions (vibration, acceleration, vaccum, radiation, EMC...) and space assurance quality (Hi-Rel parts and materials, high quality fabrication, Reliability, Analyses...).

3.2 Doris Uso Performances Before Installation

The DORIS USOs were procured from several manufacturers: FEI (USA) for SPOT2, then THOMSON-CEPE for SPOT3, TOPEX-POSEIDON, and SPOT4. For O.D.B.s and G.L.B.s, 76 oscillators have been bought from Oscilloquartz. These are the state of the art products, though they underwent extra screening during manufacturing to ensure they met the stringent DORIS specs.

All the oscillators were tested in detail, first at the suppliers, then at the CNES Time and Frequency Lab, and during all satellite tests in near real life conditions. These included vibrations, thermal vacuum, thermal gradients to access performances before installation and launching.

Dedicated test sets had to be studied and produced by L.P.M.O.-Besançon, in order to have very stable references better than USOs under test. Large amount of computer data were obtained during 5 years, and are available for the project to do additional processing or make comparisons with frequency and stability data generated in the orbital calculations.

Several non-conformances were found, in the DORIS Medium stability versus temperature variation Reference [11]. Several USOs were returned to suppliers.

3.2.1 On Board USOs performances

DORIS-SPOT 2 FEI USO use "classical" SC PQ swept 5 MHz crystal, two ovens for thermal control, and glass dewar technique. THOMSON-CEPE USOs use QAS SC PQ swept 10 MHz crystal, one oven for thermal control, and stacking boxes.

CNES and other organizations developing this technology in France have conducted many studies in industry and labs. As a result, a satellite USO has been space qualified at CEPE for use on DORIS and other telecomunication projects.

Table 3 gives typical performances, measured by suppliers and CNES. See also reference [6].

Figure 8 illustrates the DORIS medium term stability (Mean Slopes and Residuals) of CEPE EQM model n°4.

Parameter	FEI FM USO	CEPE FM USO
Short term stability		
$\tau = 1$ s	-	5.10^{-13}
$\tau = 10$ s	3.10^{-13}	4.10^{-13}
DORIS Mid term Stability with $\Theta = 0.1$ °C	2.10^{-13}/mn	$1.5.10^{-13}$/mn
Slopes		
Residuals (1 σ)	1.10^{-12}	1.10^{-12}
Long term aging after 1 month	5.10^{-11}/day	5.10^{-11}/day
Phase Noise dBc/Hz		
10 Hz	-128	-140
1 kHz	-150	-155
Temperature Sensitivity		
0 - 40 °C	$1.5.10^{-10}$	1.10^{-10}
in air	1.10^{-10}	7.10^{-11}
in vacuum		
Magnetic sensitivity	5.10^{-12}/gauss	5.10^{-12}/gauss
g - sensitivity	-	2.10^{-10}/g
Pressure sensitivity	2.10^{-9}/bar	8.10^{-9}/bar
DC power		
air	2,7 Watts	6,5 Watts
at 0 °C vacuum	1,8 Watts	2 Watts

Table 3. Typical O.B. DORIS USO performances

USO magnetic sensitivity

USO magnetic sensitivity was "discovered" during DORIS-SPOT 2 tests with FEI USO. The frequency was shifted to several 10^{-11} each time magneto-torquers were on. Figure 8.

The solution was to put USO inside a double magnetic shielding. At the same time, several studies were conducted in France by F. DEYZAC-ONERA - Reference [7] and R. BRENDEL-LPMO - Reference [8]. Table 4 from [8] summarizes the sensitivities of various components used in space USOs. Ferrite core inductors can be replace by air core inductors. Residual sensitivity (5.10^{-12}/gauss) given by QAS crystals leads CNES to maintain the shielding solution for SPOT 3 and SPOT 4.

Component	Relative magnetic sensitivity ppm/gauss	Induced oscillator magnetic Sensitivity in 10^{-12}/gauss
Ferrite core inductor	57	200
Air core inductor	0.028	0.1
SMT capacitor	0.24	0.8
VVC diode	0.003	0.0001
Transistor	-	4
Regular resonator	-	12
BVA resonator	-	4.4

Table 4. Magnetic sensitivities of various components

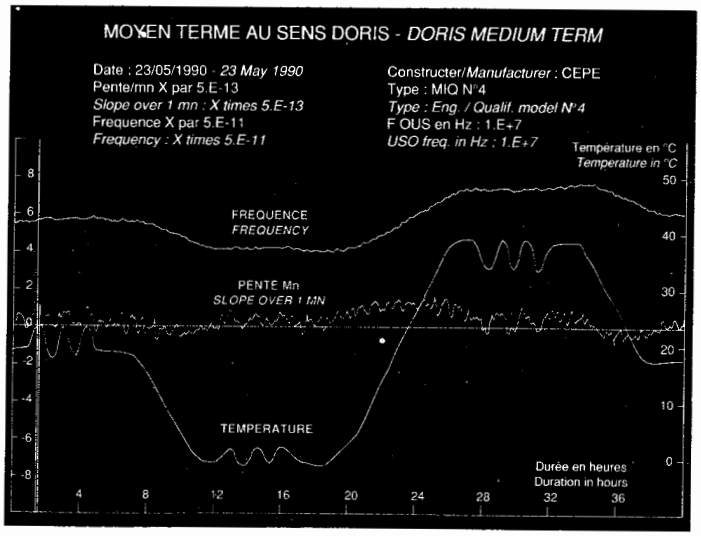


Figure 7. DORIS Medium Term Stability

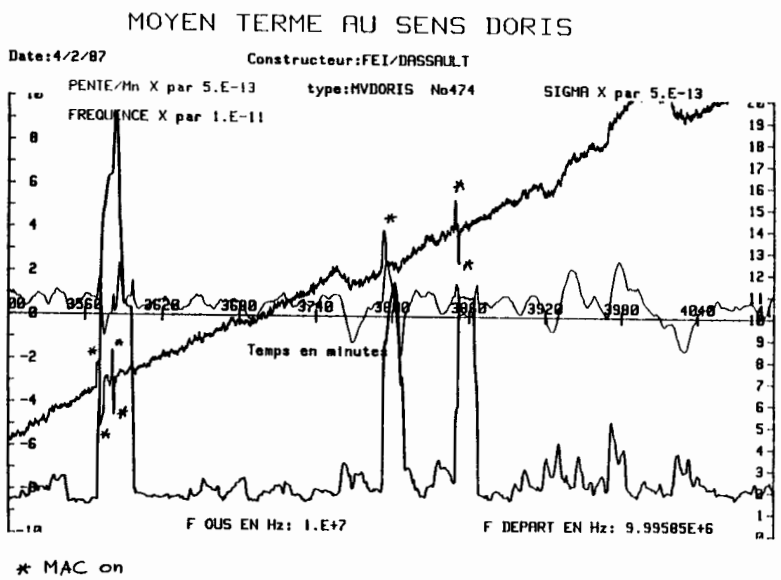


Figure 8. Effect on magnetic field over DORIS-SPOT 2 USO frequency

3.2.2 O.D.B. and G.L.B. USOs performances

76 USOs have been bought from Oscilloquartz, 66 for O.D.B. and 10 for G.L.B. All OSA USO are 8600 type 5 MHz BVA AT crystal using Dewar glass technique. ODB USOs have single oven, GLB USOs have double ovens.

Table 5 gives typical performances measured by CNES.

In order to respect the stringent DORIS specs, all oscillators were selected by OSA. Figure 10 and 11 give typical DORIS stability performances of ODB and GLB USOs respectively.

Parameter	O.D.B. USO	G.L.B. USO
Short term stability τ = 1 s τ=10 s	 5.10^{-13} 5.10^{-13}	 5.10^{-13} 5.10^{-13}
DORIS Medium term Stability with Gradient RMS Slopes	$\pm 5.10^{-13}$/mn	$\pm 4.10^{-13}$/mn
Residuals (1σ)	1.10^{-12}	1.10^{-12}
Long term aging 1 day	$\leq 2.10^{-11}$/j	$\leq 2.10^{-11}$/j
Phase Noise dBc/Hz		
10 Hz	-142	-142
1 kHz	-154	-154
Temperature Sensitivity	$\pm 1.5\ 10^{-10}$ 0 to 45 °C	$\pm 2.5\ 10^{-11}$ -20 to +55 °C
DC power in air	0°: 4.6 W 20°: 4.0 W 40°: 3.5 W	-20°: 3.1 W +20°: 2.4 W +55°: 1.9 W

Table 5. Typical DORIS ground USO performances

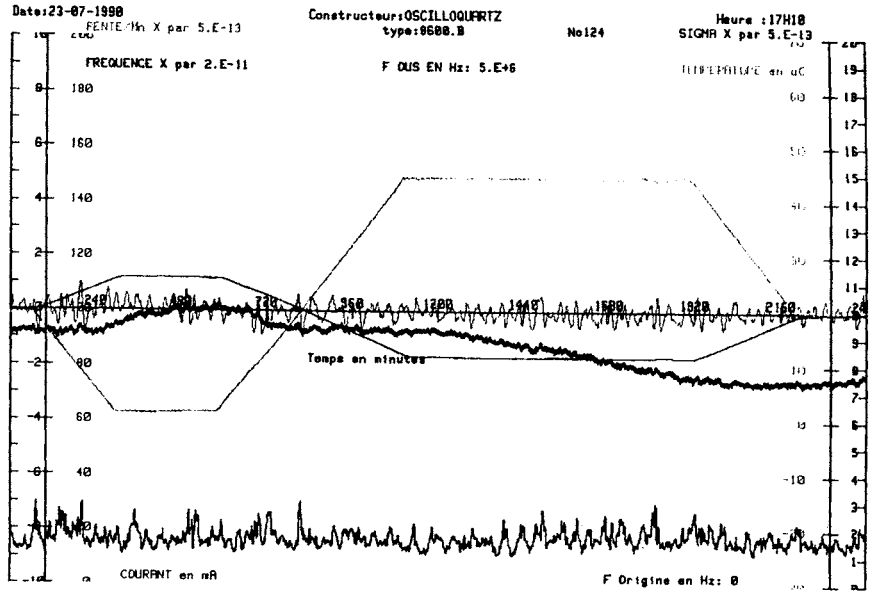


Figure 9. DORIS mid term ODB USO performance

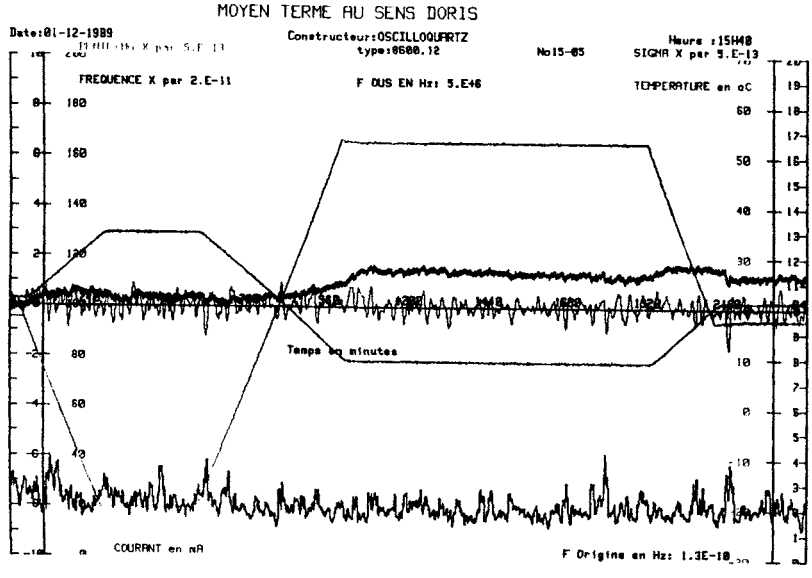


Figure 10. DORIS mid term GLB USO performance

3.3 Doris Uso Performances Via Orbit Measurements

Orbit determination beacons were installed during the 86-91 period. DORIS-SPOT2 was launched by ARIANE on January 22, 1990. The CNES ORBIT METROLOGY DEPARTMENT has processed on-board data using the ZOOM software - References [4,5].

Several parameters are adjusted:

- orbital elements
- dynamic parameters such as ionosphere, earth gravity...
- frequency deviation between on-board USO and ODB USOs, for each pass when the satellite is within range a beacon. We can calculate the absolute deviation when the satellite is over CNES-Toulouse, because we have Master Beacon slaved to an atomic clock with a well known frequency ($+7.10^{-13} \pm 1.10^{-13}$).

Several passes per day of DORIS-SPOT2 over Toulouse permit to know the absolute frequency of the on-board USO. After that, using SAMOUS software, the ORBIT Metrology Department calculates the absolute frequency of each ODB USOs - Reference [9].

In order to test accuracy of SAMOUS software, one experiment was conducted during one month in France, between two separate beacons: CNES-Toulouse and Observatoire de la Côte d'Azur-OCA-Grasse, 800 km apart from each other. The two beacons were located very accurately using laser OCA reference and GPS measurements. The Grasse USO frequency was measured each 10 seconds, with an accuracy of $\pm 1.10^{-13}$, using an hydrogen maser as reference. 77 common views were collected, and experimental measurements were compared to the computed values. Using "SAMOUS" software and "GUIER" filter, they agree within 4.10^{-12} (1 sigma) - Reference [10].

Using this very accurate frequency determination, we are able to study in real conditions, the medium and long term performances of each USO.

3.3.1 In orbit DORIS-SPOT2 USO performance

Figure 12 gives the drift in orbit. Results are $1{,}1.10^{-11}$/day after one month, and 6.10^{-12}/day after one year. This is a better result than lab measurements.

3.3.2 O.D.B. USOs performances

Table 6 gives the mean drift of 38 O.D.B. USOs, calculated over the last month, expressed in 10^{-10} per month. Overall meandrifts are excellent. 36 units have a drift less than 2.10^{-10} per month and 17 units have a drift less than 6.10^{-11} per month.

However, several USOs have frequency jumps of 1.10^{-11} to 6.10^{-10}. Figures 13 and 14 give examples of very low drift with and without frequency jumps. Theses frequency jumps were observed at CNES lab over few units tested during 10 months. The main important point for DORIS is that these jumps can be "absorbed" by Zoom software, and do not affect, on the first order, the DORIS location accuracy. ODBs are on stable environmental conditions (temperature, vibrations).

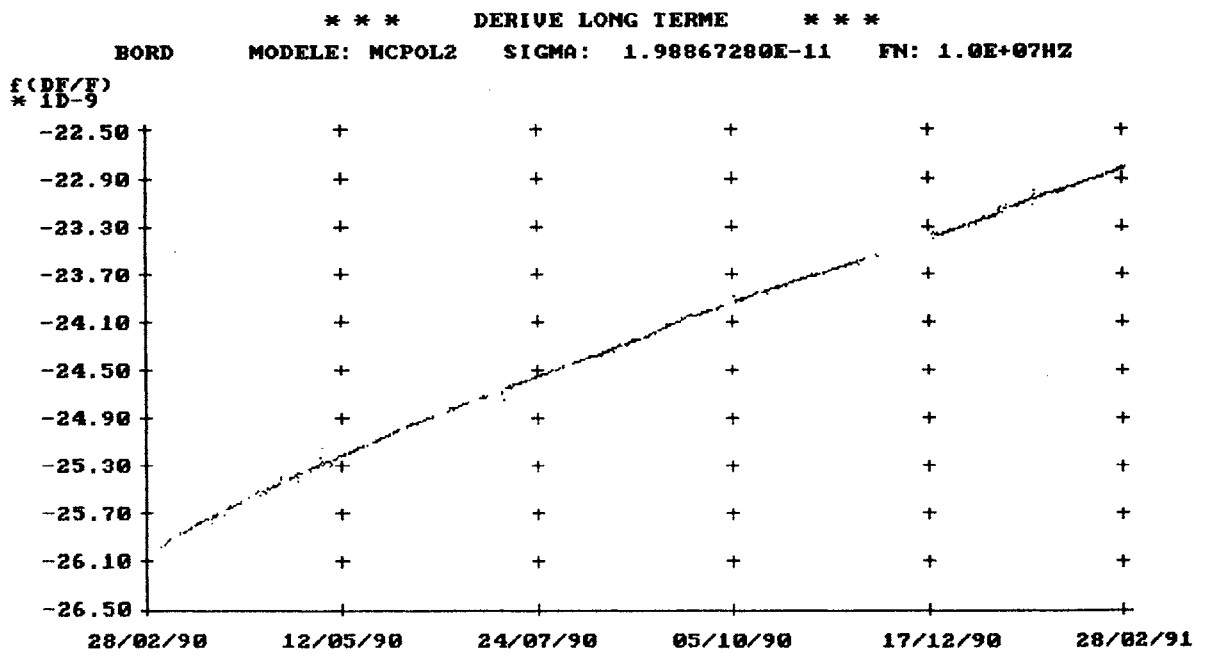


Figure 11. DORIS-SPOT2 in orbit USO DRIFT

Name of site	Country	Mean drift/month
Adélie	France	$+1.4\ 10^{-10}$
Arequipa	Peru	$<6\ 10^{-11}$
Dakar	Senegal	$<6\ 10^{-11}$
Djibouti	Djibouti	$-1\ 10^{-10}$
Fairbanks	Alaska (USA)	$+8\ 10^{-11}$
Goldstone	California (USA)	$+1.6\ 10^{-10}$
Ste Hélène	United Kingdom	$+8\ 10^{-11}$
Kerguelen	France	$<6\ 10^{-11}$
Kourou	France	$+7\ 10^{-11}$
Marion Island	South Africa	$<6\ 10^{-11}$
Moresby	Papua	$+6\ 10^{-11}$
Ottawa	Canada	$<6\ 10^{-11}$
Réunion	France	$-7\ 10^{-11}$
Richmond	Virginia (USA)	$+1\ 10^{-10}$
Santiago	Chile	$<6\ 10^{-11}$
Socorro	Mexico	$<6\ 10^{-11}$
Waimea	French Polynesia	$-6\ 10^{-11}$
Yellowknife	Canada	$<6.10^{-11}$
Tristan da Cuha	United Kingdom	$+8\ 10^{-11}$

Name of site	Country	Mean drift/month
Amsterdam	France	$-1.6\ 10^{-10}$
Arlit	Niger	$+9\ 10^{-11}$
Dionysos	Greece	$-1.5\ 10^{-10}$
Easter Island	Chile	$+5.2\ 10^{-10}$
Flores	Açores (Portugal)	$<6\ 10^{-11}$
Hartebeeshoek	South Africa	$<6\ 10^{-11}$
Hawaï (Observatory)	Hawaï (USA)	$-1.3\ 10^{-10}$
Kokee Park	Hawaï (USA)	$+2.4\ 10^{-10}$
Libreville	Gabon	$<6\ 10^{-11}$
Metsahovi	Finland	$+1.1\ 10^{-10}$
Nouméa	France	$<6\ 10^{-11}$
Purple Mountain	China	$-1.8\ 10^{-10}$
Reykjavik	Iceland	$<6\ 10^{-11}$
Rio Grande	Argentina	$<6\ 10^{-11}$
Signy Island	United Kingdom	$1.2\ 10^{-10}$
Spitzberg	Norway	$+8\ 10^{-11}$
Wallis	French Polynesia	$<6\ 10^{-11}$
Huahine	French Polynesia	$+2\ 10^{-10}$
Tromsoe	Norway	$+1.7\ 10^{-10}$

Table 6. Mean Drift of ODB USO using orbit DORIS data

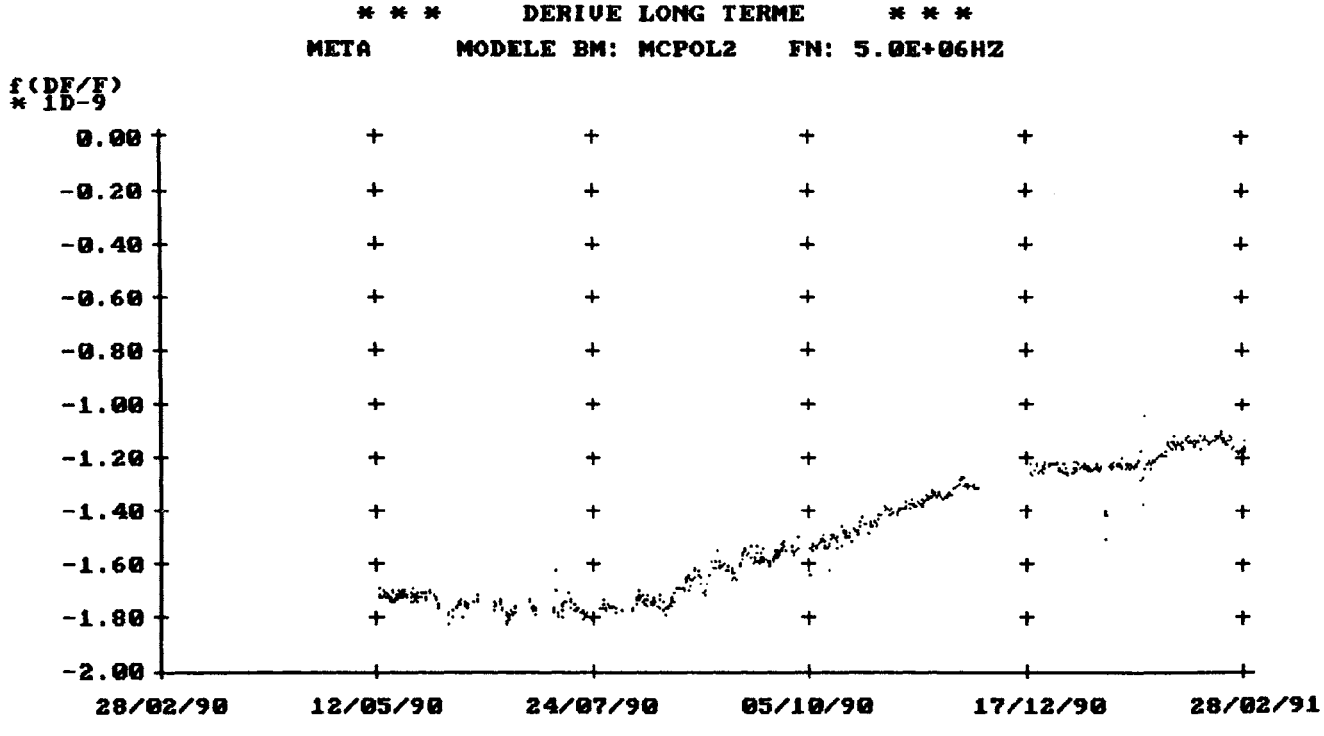


Figure 12. Metsahovi (Finland) ODB USO aging

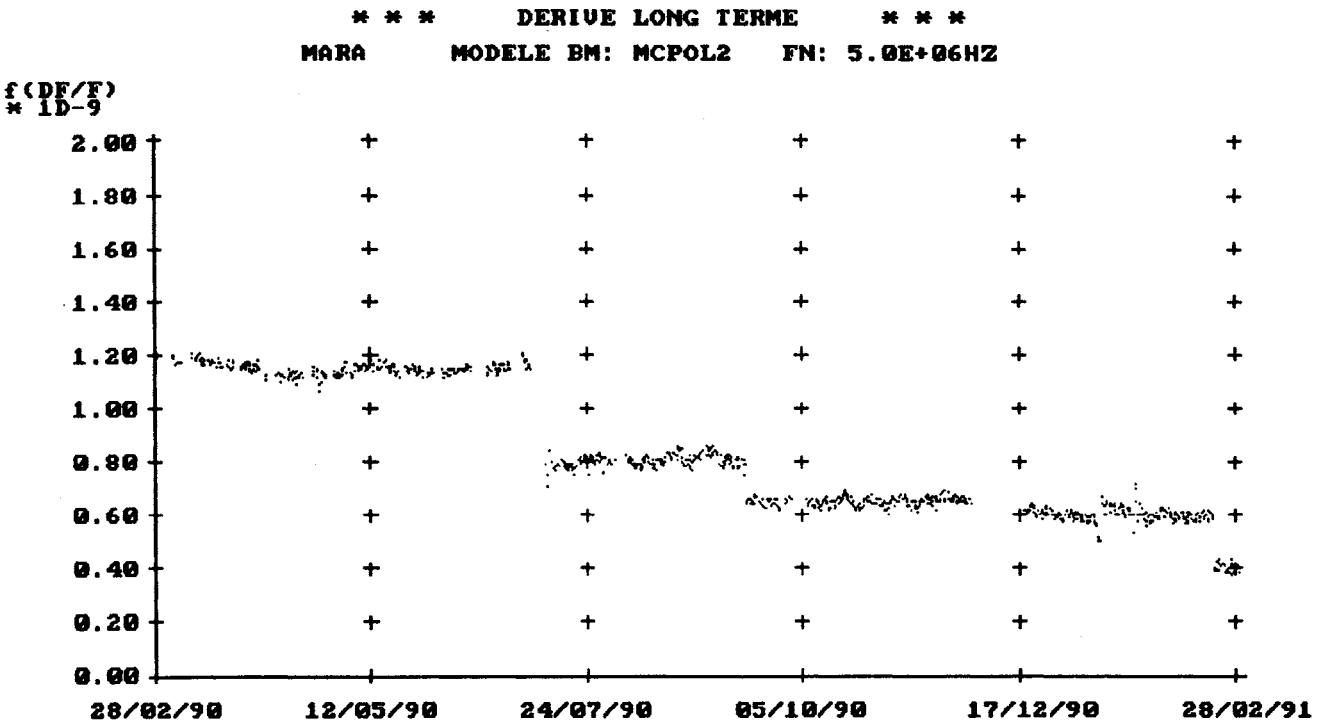


Figure 13. Marion Island (South Africa) ODB USO aging

4. CONCLUSIONS

CNES has designed and developed the Doppler Orbitography and Radiopositioning DORIS system, for several applications such as orbit determination and ground location. DORIS on-board payloads have been integrated on French SPOT satellites, and on the TOPEX-POSEIDON NASA satellite. After two years of data, and using state of the art modeling techniques, the DORIS residual measurement errors are 0,3 mm/s, equivalent to 5 cm accuracy on the satellite-station distance. Absolute ground location accuracy is 10 cm and few centimeters in some kilometer baseline.

The fundamental measurement precision in the DORIS one-way Doppler technique depends strongly on the short term and medium term stabilities of USO used on board receiver and ground beacons. Several USOs have been supplied from state of the art technology. The DORIS project has led to in-depth work on the different techniques contributing to quartz crystal USO stability. About 50 USOs have been carefully tested in the time-frequency lab. at CNES and are now regularly monitored through the DORIS system.

Several works in hand to increase performance further is expected to provide gains of an order of magnitude in the next five to ten years, with short term stability of a few parts in 10^{-14}, and long term stability of a few parts in 10^{-12} a day.

ACKNOWLEDGMENTS

The author is grateful to many collegues, especially M. DORRER and A. AURIOL (DORIS project), F. NOUEL, C. PIERRET, A. PIUZZI, C. JAYLES (ORBIT METROLOGY DEPARTMENT), J.F. DUTREY, J.B. LAPORTE, E. ROBERT (TF Department) for all results and valuable discussions.

REFERENCES

[1] DORIS NEWSLETTER N° 1, NOV 90.

[2] DORIS NEWSLETTER N° 2, JUL 91, published by CLS-ARGOS, 18 Avenue E. Belin, 31055 TOULOUSE, FRANCE.

[3] M. DORRER, B. LABORDE, P. DESCHAMPS - CNES
DORIS: System Assessment Results With Doris on SPOT2
IAF-90-366-Dresden Germany 6-12 Oct 1990.

[4] F. NOUEL and al - CNES
Precise orbit determination of SPOT platform with DORIS
AAS-AIAA conference, paper 91-409.

[5] C. JAYLES, H. RENAULT - CNES
On-board low-earth orbit determination
AAS-AIAA conf. August 91, DURANGO, (U.S.A.), 91-362.

[6] E. GERARD, R. MOLLE - THOMSON-CEPE
Design and qualification of an ultra stable oscillator for space applications
5th EFTF, 1991, pp. 347-356.

[7] F. DEYZAC - ONERA
Magnetic sensitivity of quartz oscillators
4th EFTF, March 1990, pp. 255-258.

[8] R. BRENDEL et al - LPMO
Magnetic sensitivity of oscillator components
5th EFTF, March 1991, pp. 390-395.

[9] C. PIERRET - CNES
Bilan de travail accompli sur le logiciel SAMOUS: comportement à long terme de l'OUS bord et des OUS balises du réseau DORIS.
CNES Technical Report TE/IS/MS/MO 91-260 (1991).

[10] C. PIERRET - CNES
Résultats de l'expérience DORIS-Grasse
CNES Technical Report TE/IS/MS/MO 91-300 (1991).

[11] J.B. LAPORTE - CNES
Specifications and performances of ultra stable oscillators in the DORIS orbitography and positioning system.
2nd EFTF, March 1988, pp. 163-183.

PRECISE DOPPLER TRACKING FROM THE MEDICINA VLBI STATION

R. Ambrosini [1], G. Comoretto [2], L. Iess [3], A. Messeri [3]

[1] Istituto di Radioastronomia, CNR, Bologna; [2] Osservatorio Astrofisico di Arcetri, Firenze;
[3] Istituto Fisica Spazio Interplanetario, Frascati

ABSTRACT

The successful first opposition test of Doppler tracking the Ulysses spacecraft from the Medicina VLBI station has proved its capability to perform a systematic search for gravitational waves. In house and JPL data analysis have shown that the target Allan Variance of 3×10^{-14} at 1000 s, planned for the DSN antennas, has also been achieved from our station.

The main observation campaign during the second opposition phase, will last for thirty continuous nights, from mid February to mid March 1992.

Present article will describe the main hardware and software features developed for this application, together with some results of the first opposition test.

Keywords: Gravitational wave detection, Doppler tracking, very low phase noise receivers.

1. THE SCIENTIFIC RATIONALE

Searching for the experimental evidence of Gravitational Waves has started since 1960 but, until now, without positive success. This has been mainly due to the extremely high sensitivity required to detect very tiny strains on the sampled masses.

Most of the experiments have been arranged in order to detect gravitational bursts with periods in the KHz region with resonant sensors. They are in general very massive aluminum cylinders levitating over liquid helium.

More recently, advances in precise Doppler tracking of interplanetary spacecrafts have reached a precision comparable to the expected spatial strain amplitudes of gravitational waves of very low frequency (from 10^{-1} to 10^{-4} Hz). In this case the gravitational burst should produce a relative displacement between the earth and the spacecraft.

This will be seen as a peculiar signature in the relative frequency comparison, made at the receiving site, between the station H-Maser frequency standard and the received downlink radio carrier from the spacecraft, which is a microwave coherent replica of the uplink signal, generated in turn from the H-Maser of the transmitting antenna (Doppler observable). This can be seen in simplified form in Fig. 1.

Once all the instrumental, propagation media and orbital effects would be removed from the data records collected at the receiving site, three events should be observed. These are the effects of the gravitational burst, acting on the signal frequency transmitted from ground, the one transponded from the spacecraft and finally on the receiving frequency reference, which is seen, in this case, as a pulse with an opposite sign. This time sequence will appear in a reversed order on the data stream as determined by the respective propagation delays from the spacecraft (downlink only) and from the transmitting ground station (both up and downlinks).

Possible astrophysical sources in this low frequency band are: collapse and collisions of black holes in the nuclei of galaxies and quasars; periodic wave trains emanating from massive orbiting binary black holes; or an isotropic background of cosmological origin.

Because Doppler tracking is an inherently broad band gravitational wave detector it is most sensitive to waves whose characteristic time scale is shorter than the round-trip-light-time (RTLT), needed by the radio signal to travel across the up and downlinks. Furthermore since they are supposed to be stronger at lower frequencies, it is important to perform measurements with spacecrafts as far as possible from the earth.

The Ulysses mission is particularly suitable for this purpose because it will reach a distance of about 3 AU from the earth and its radio link configuration will allow some cancellation of the noise sources due to the propagation media.

2. THE ULYSSES EXPERIMENT

The hardware setup needed for a gravitational wave detection experiment is made up of the standard radio link configuration used to track the spacecraft during the mission.

The NASA/JPL stations in California, Australia and Spain transmit and receive to and from the

spacecraft in turns, to overcome the effect of earth rotation. To keep the system phase noise as low as possible, the S-band uplink signal is obtained starting from an atomic, H-Maser, frequency standard. The transponder, on board of the spacecraft, generates two down-link signals, one still at S-band and the other at X-band. They are both coherent with the uplink signal. This arrangement allows to calibrate the dispersive contributions along the downlink path originated from the terrestrial ionosphere and the interplanetary scintillations, because their frequency dependence is theoretically known. Indeed the interplanetary plasma is expected to be the dominant noise source on the present gravitational wave experiment, also because the Ulysses configuration does not allow a direct calibration of the uplink path. To reduce at a minimum this contribution the measurement campaign will be performed during the cruise phase when the earth itself shields from the solar wind the radio path toward the spacecraft. This is called the *opposition* phase and happened the first time in January 1991 and will be repeated in February 1992 and 1993.

In this paper we will deal of the results obtained during the first opposition. This measurement campaign has been performed as a complete test of the system, both hardware and software, in view of the main observations, of the second and third oppositions, when the higher RTLT will give better chances to detect a gravitational event.

By using a separate receiving system with its own frequency standard and with an uncorrelated troposphere and ionosphere (if it is far enough from the transmitting station), it should be possible to discriminate between different noise sources.

This *three way* operation, respect to the case of a single transmitting and receiving antenna called *two way* operation, will be useful not only to validate the detection of a gravitational wave event, but also during the test experiments, to isolate the single components from the total system noise.

The Medicina VLBI station is participating to the GWE experiment under this last configuration, utilizing an original implementation (Ref. 1) of an open loop tracking receiver, called Digital Tone Extractor (DTE).

Main feature of this instrument is its ability to extract amplitude and phase of a tone of known frequency, also when it is buried in wideband noise and the nominal frequency of the carrier is drifting. This capability is exactly what is needed to precisely track the Ulysses's downlinks.

3. THE MEDICINA VLBI STATION IN THE GWE

As shown in the block diagram of Fig. 2, the main reflecting surface of the antenna is a parabolic dish with a diameter of 32 meters, illuminated from the prime focus with a dual S/X feed. The aperture efficiency and system temperature at both frequencies are 45 % and 110 K respectively. The front end is made of uncooled Hemts. The first LO is a PLL driven by the station H-Maser at 5 MHz, with two outputs: one at 2020 MHz for the S-Band and the other at 8080 MHz, obtained by straight multiplication of the first. A comb generator, driven by a separate cable, injects at both receiver inputs phase reference tones made of rails spaced 1 MHz apart and coherent with the H-Maser. All these components are located in the thermostatically controlled box at the prime focus of the antenna. This means that during S/X operation the subreflector, normally used in the Cassegrain configuration, has to be physically removed from its position.

The two IF bands, each containing a spacecraft downlink signal and the calibration tones, are processed by the standard VLBI Mark III terminal. They are first downconverted to video band and low pass filtered to 2 MHz bandwidth, then sampled at the Nyquist rate with one bit resolution and finally formatted in frames with very accurate timing codes, derived from the H-Maser.

From the Mark III terminal two pairs of signal/clock data streams are sent to two DTE crates. Each of these is dedicated to a frequency band (X or S) and contains two tracking channels, one for the spacecraft signal and the other for the calibration tone. So a total of four DTE channels have been used for the Ulysses tracking. This allows to extract the X and S band information, together with a real time calibration of the actual electrical length of the receiver chains, utilized in both bands. By this procedure it is possible to wipe out all phase instabilities due to the station hardware, except for those due to the H-Maser and to the antenna structural deformations (these last are not relevant for the present level of sensitivity of the experiment).

4. DTE OPERATIONAL CHARACTERISTICS

In each DTE crate a dedicated microprocessor digitally multiplies the deformatted data stream, with the in phase and in quadrature outputs of a synthesizer, whose frequency can be controlled by the software running on the main station computer, an HP1000F. This also performs the real time integration of the two products to reduce the measurement noise bandwidth to typically a few Hz, in order to precisely extract the amplitude and phase of the incoming signal.

These quantities, together with time code information, are temporary stored on the system disk, then permanently transferred to magnetic tapes for full off-line analysis.

For dual frequency observations, as is the case for Ulysses, the control program tracks the signal frequency only in one channel (usually the stronger X band). Then the other band is blindly recorded around the frequency derived from the ratio used on board of the spacecraft to generate the two coherent downlinks.

Despite its narrow band, a DTE channel can reliably follow a signal with a large frequency drift. Indeed the host computer software is capable to update in real time the synthesizer frequency according to a low pass filtered combination of the phase residual measured in the previous integration, with a predicted

value *a priori* known. This is usually the case of interplanetary probes, where the orbital parameters are known.

In this way, the DTE and the control program act together as a high precision phase locked loop, with a large tracking capability under poor signal to noise receiving conditions and a low output data rate. This feature allows an easy data handling both at the acquisition time and during the off-line data analysis.

On the other hand, before every three way experiment, the ASCII files (called MDA TEXT PREDIX) containing frequency predictions for each DSN tracking station, have to be processed to compute the sky frequencies as observable from the geographic location of Medicina. These values are then fitted using a six parameter interpolation, in such a way that just few quantities have to be updated in the tracking program before every pass.

If the signal frequency drift is known, it exists a maximum integration time able to handle it. This condition happens when the input phase will vary more than 180 degrees over a single integration time. If this limit is not exceeded, the input phase can be retrieved unambiguously over different contiguous cycles. This means that a DTE channel can track frequency variations up to $B^2/2$ Hz/s, where B is the instantaneous measurement bandwidth, which is approximately equal to the reciprocal of the integration time.

If it is not known the frequency behavior of the input signal and its S/N ratio is high enough, the tracking capability has to be reduced to $B^2/6$ Hz/s.

Because in the Ulysses experiment the maximum rate of change of the Doppler effect, mainly due to the earth rotation, amounts to about 1.5 Hz/s at X-band, the minimum measurement bandwidth should be 2 Hz. This would correspond to a maximum integration time of 0.5 seconds.

During the actual experiment we have used a bandwidth of 2.5 Hz (0.4 s), as the best compromise between reliability in signal locking and amount of stored data.

We mention also that in the present system configuration, the computer speed limits the integration time to be greater than 0.25 s.

Finally the main tracking program provides, as an initial acquisition aid, a searching routine over a selectable bandwidth. This routine is automatically restarted if, for any reason during an observation, the locked condition is lost. In this case a narrower search can be selected to speed-up the search around the last known tracking frequency.

5. NOISE ANALYSIS BUDGET

All media trough which the Doppler signal must pass are potential sources of noise to the experiment. A list of these are: the transmitting H-Maser, electronics and antenna, the uplink atmosphere (wet and dry), the uplink interplanetary plasma, the spacecraft antenna and transponder, the downlink plasma, the downlink troposphere (wet and dry), the receiving antenna, electronics and H-Maser.

From Fig. 3 it is possible to compare the expected relative importance of the single contributions. The interplanetary plasma is considered to be the most relevant noise source, over the 100 s - 1000 s range, relevant for this experiment.

An Allan Variance plot of the X-band DTE tone, accounting for most of the instrumentation noise measured during a typical pass of Ulysses at the Medicina VLBI station is reported in Fig 4. Almost equal performance has been obtained at S-band, because both receivers are quite similar and exposed to the same environment.

Calibration of the signal data according to these tone measurements is not needed under the present sensitivity of the GWE with Ulysses. However for future experiments with higher sensitivity, like the ones planned for the Cassini spacecraft, the station electronics will have to be improved to match the phase stability target of few parts 10^{-16}. This of course will also imply to upgrade the present performance of the H-maser, utilized as the station frequency standard.

6. SOME RESULTS OF THE FIRST OPPOSITION

A detailed analysis of the data from the first opposition experiment ca be found in (Ref. 2, 3). Here we want to show the measured level of system noise achieved over the nine nights dedicated to this type of observations.

Fig. 5 shows the measured Allan Variance at 1000 s, from most of the stations involved in the experiment and with different types of data handling. The target level of 3×10^{-14} has not been reached all times, but the reliability of the overall system has been proved to be quite good.

In particular the DTE operation has been successful and its data analysis has given comparable results with the one obtained at JPL.
Just as a crude example of the usefulness of the three way operation, the Medicina data allowed to immediately recognize that just the transmitting section at the Canberra complex had some degraded performance during the experiment. This was evident in a few data records collected from the Medicina station, at the hand-over between that DSN complex and the other two.

7. CONCLUSIONS AND FUTURE WORK

Three way Doppler tracking of the Ulysses spacecraft from the Medicina VLBI station during the first opposition phase has been successful. The DTE, developed for this type of application, acted as a reliable and accurate tracking receiver, with a very low output data rate (2.5 Hz).

These features come from the fact that the DTE combines the tuning capability of both open loop receivers (frequency predictions) and closed loop receivers (real time measurements of signal frequency residuals).

Both in terms of reliability and quality of the data, performances comparable to the DSN tracking stations have been achieved. A complete analysis of three way data has given a deeper insight in the

Doppler observable. In particular it helped to recognize the relative importance between different noise sources.

The expected goal of sensitivity for the Ulysses Gravitational Wave Experiment has been met, but to cope with future missions like Cassini, the phase stability performance of the station instrumentation has to be improved. This will imply both an upgrade of the master frequency reference (H-Maser) and of the phase stability of the microwave part of the receiver chains.

The successful first opposition test let us afford with rational confidence the main observation campaign of the second opposition phase. The experiment will consists in the Doppler tracking the Ulysses spacecraft for fourteen hours centered over midnight, every night, from the 18/2/92 up to the 18/3/92. This month period has resulted to be the maximum available time of the DSN and Medicina stations for this experiment.

8. REFERENCES

1. G. Comoretto, S. Cattani, A. Gallerani, S. Montebugnoli, September 1991, *Digital Tone Extractor Hardware Technical Manual*, IRA Internal report 151/91.

2. B. Bertotti, R. Ambrosini, S. W. Asmar, J.P. Brenkle, G. Comoretto, G. Giampieri, L. Iess, A. Messeri, A. Vecchio, H. D. Wahlquist, November 1991, *The Gravitational Wave Experiment*, Astron. Astrophys. Suppl. Ser. **92**, 431-440, (1992).

3. B. Bertotti, R. Ambrosini, S. W. Asmar, J.P. Brenkle, G. Comoretto, G. Giampieri, L. Iess, A. Messeri, A. Vecchio, H. D. Wahlquist *Ulysses' Gravitational Wave Experiment. Report on the first opposition*, IFSI, 91-25 Internal Report, December 1991.

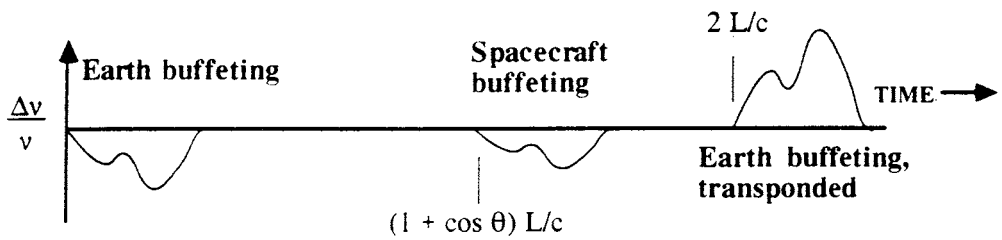


Fig. 1: Data record pattern of a Gravitational wave detected by the Doppler tracking method.

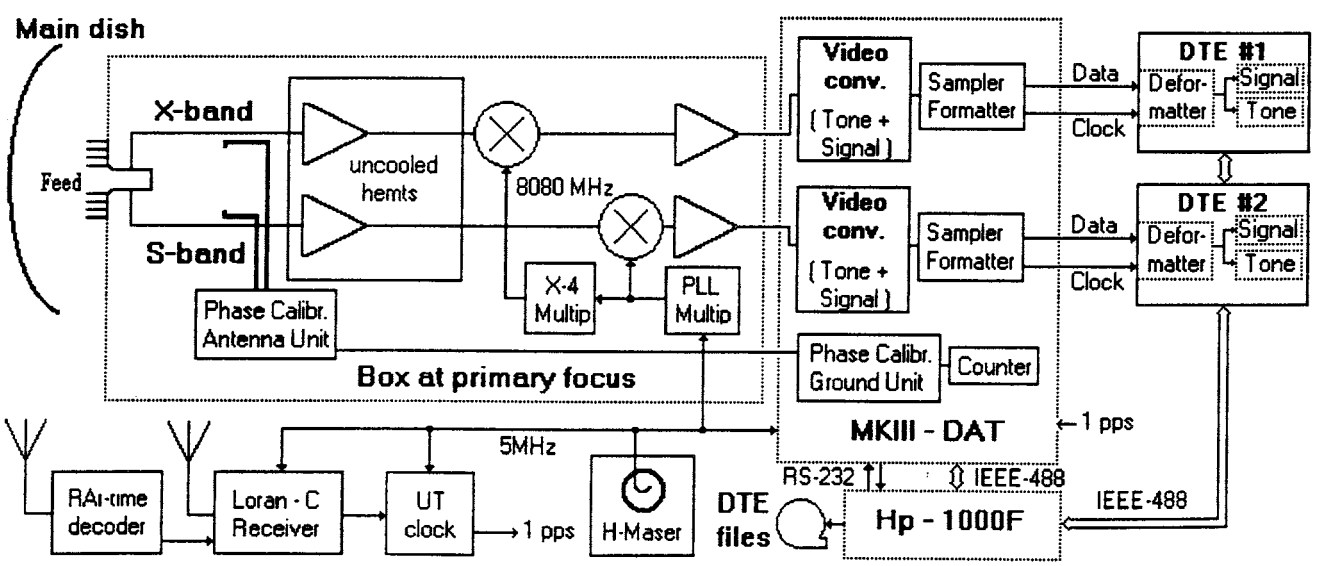


Fig. 2: Block diagram of the Medicina data acquisition system

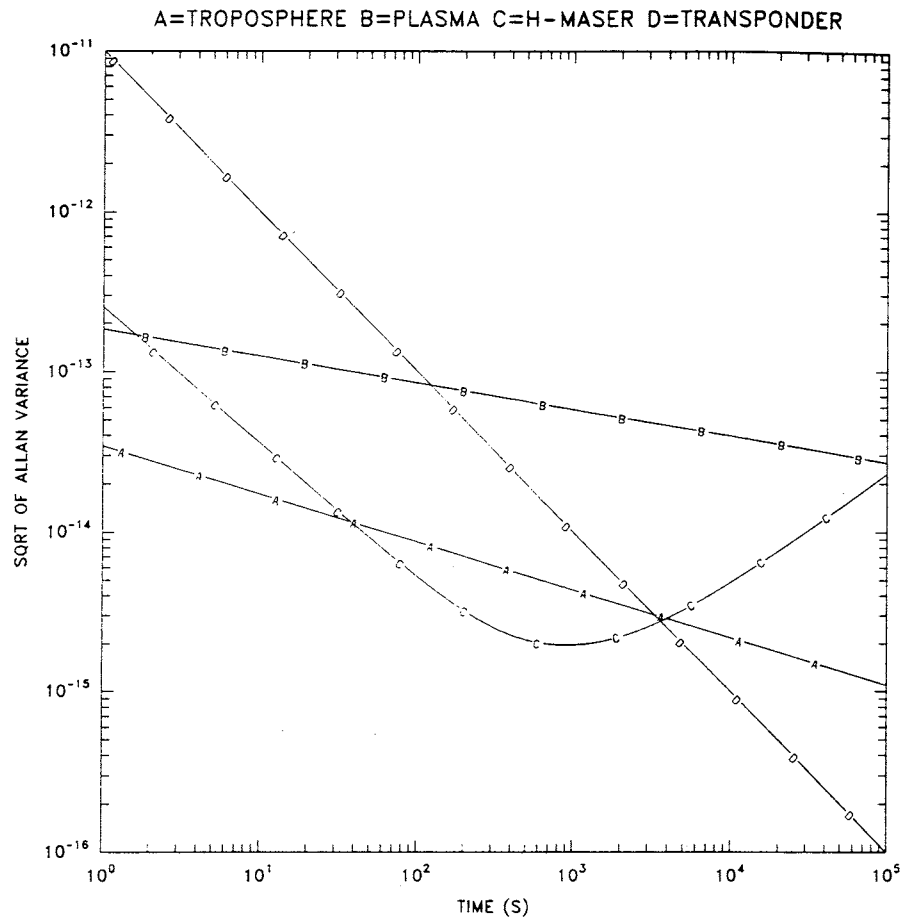


Fig. 3: Expected noise spectra for the first opposition experiment.

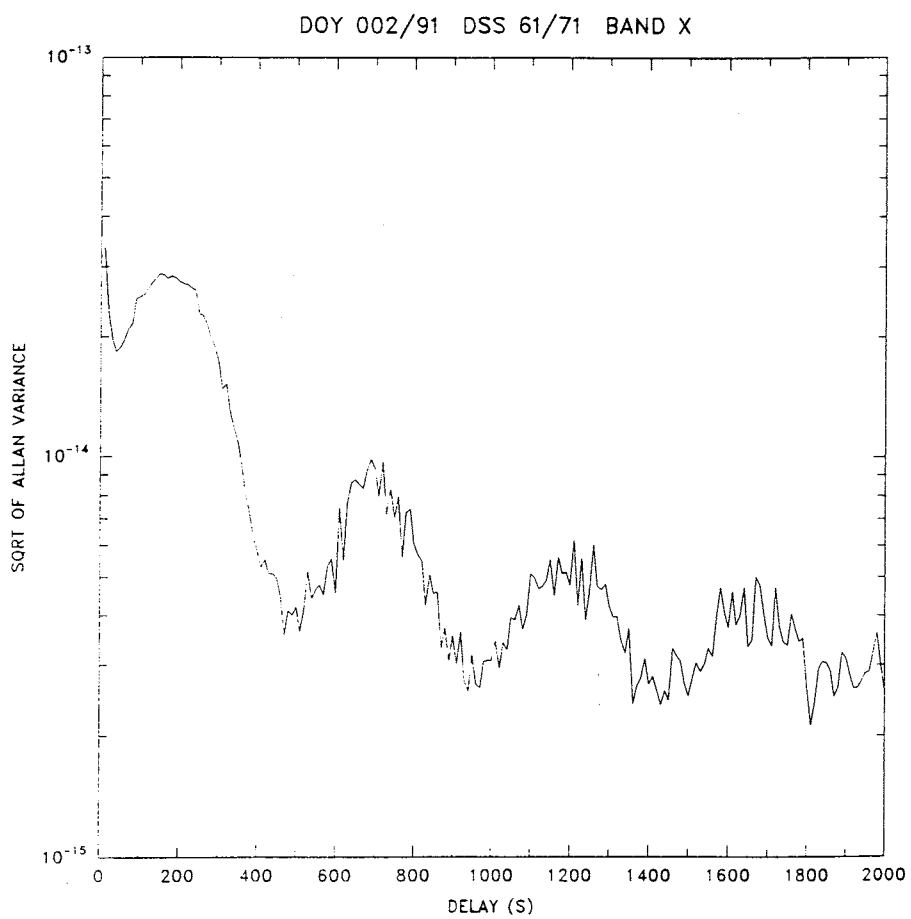


Fig. 4: Phase stability measurement of the whole X-band receiver at the Medicina station.

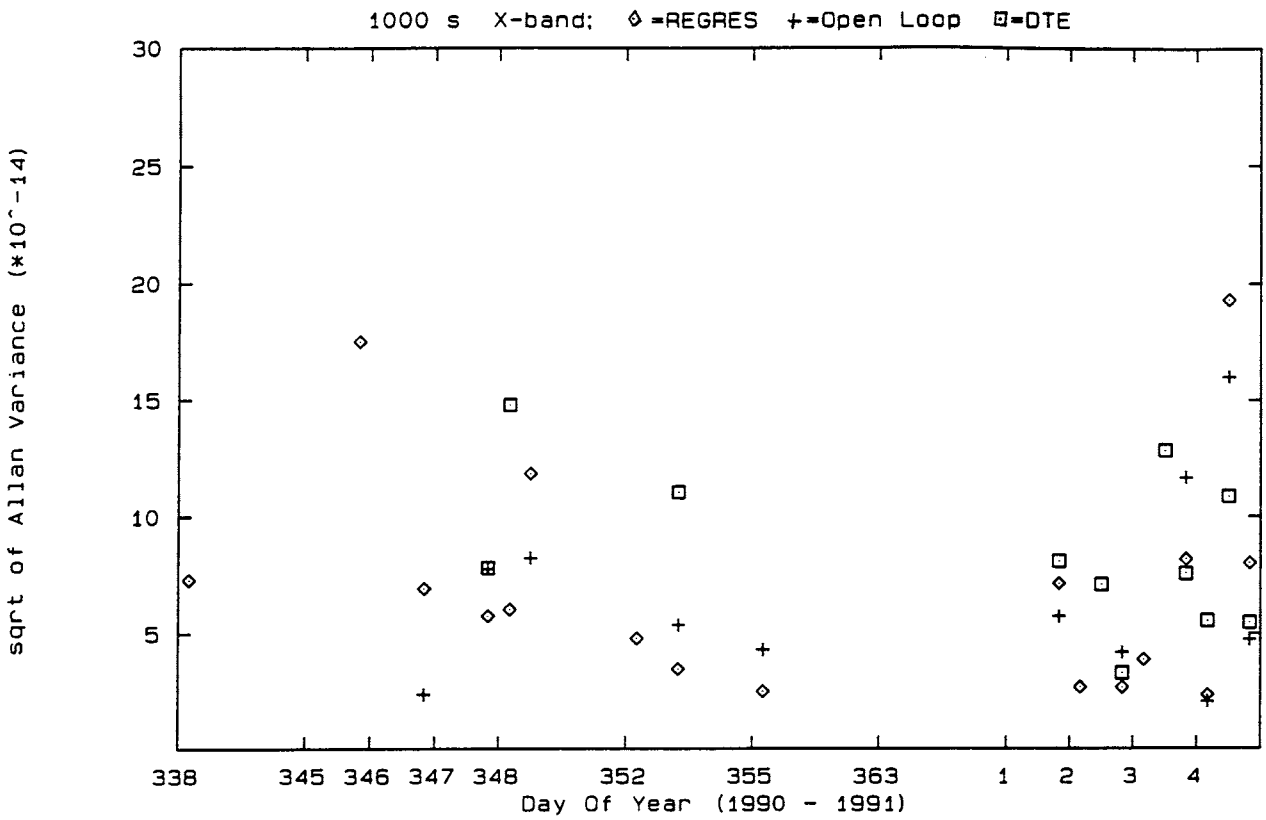


Fig. 5: General summary of 1000 seconds, X-band Allan variances during the first opposition experiment.

A NEW MODEM FOR MICROWAVE TIME SYNCHRONISATION VIA GEOSYNCHRONOUS TELECOMMUNICATION SATELLITES

Dipl. - Ing. Michael Dienert

TimeTech GmbH - Technologiezentrum Nobelstraße 15 - D-7000 Stuttgart 80 - Germany

ABSTRACT

The two way time transfer technique via geosynchronous telecom-satellites is at the moment one of the most accurate methods for synchronisation and comparison of remote clocks. Most of the unknown propagation delays can be eliminated by the two way principle. The use of a coherent spread-spectrum technique with a truncated pseudo-noise-code offers a resolution better than 30 ps of the measured time interval. The receiver is built around a delay locked loop (**DLL**), which correlates the received signal with the known PN-Sequence to derive the control signal of the loop. In the locked state both PN-sequences are synchronous and tracking errors of less than 30 ps are possible. Results showing the accuracy of the modem depending on signal-to-noise-ratio and variation of total input power levels are presented herein and show that the expected improvement of the jitter of the internal delay by an increase of the chip-rate is possible.

1. Time-Transfer

Time-transfer is needed for fast continuous comparison of atomic clocks to build up national and world-wide time networks and references also to meet demands for national and intercontinental time distribution. These applications demand a reliable time transfer system with an accuracy, far better than one nanosecond to keep up with the continuously increasing accuracy of atomic time standards. Links via satellites promise the best results for long-distance transfer.

Unlike clock comparison using the Global Positioning System the method described in this paper is not susceptible to ionospheric errors and military secrecy aims. Furthermore there is no need to know the precise geodetic co-ordinates of the receiver- and transmitter-sites or the satellite's orbit and accuracies below 20 ps can be expected within this decade.

2. Two-way time-transfer

A great many unknown delays which are susceptible to variation over time can occur along a propagation path between two ground stations and a geosynchronous satellite. By using a two-way time transfer technique, delays can be kept to a minimum, provided that the said delays are not dependent on transmission from ground station one to ground station two or vice versa. The symmetry of the link cannot of course, be perfect but there are several criteria for ensuring optimal results. For example, links via transponders using the same antennas for transmission and reception are definitely preferable.

Figures 1a and 1b illustrate this principle.

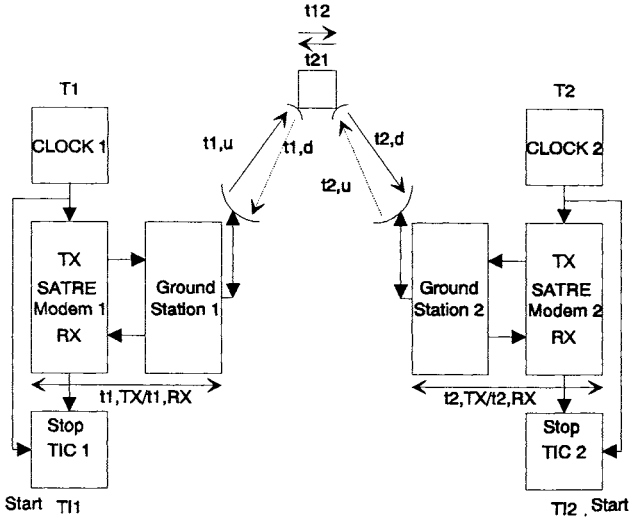


Figure 1a: A two-way time-transfer set-up, showing the different propagation delays

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

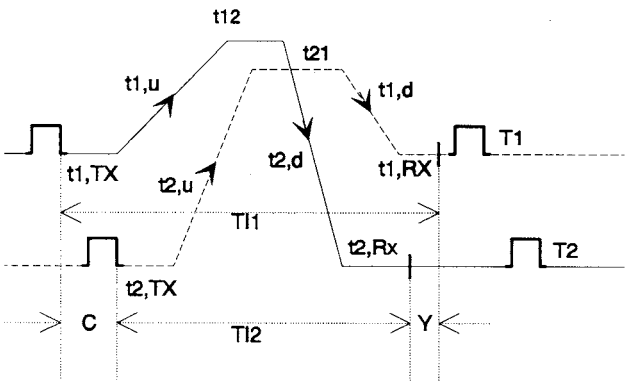


Figure 1b: Time schedule of the propagation delays

Equations:

$$C = T_1 - T_2$$

Ti_1 = time difference at Station 1

Ti_2 = time difference at Station 2

$$Ti_1 = t_{1,TX} + t_{1,u} + t_{12} + t_{2,d} + t_{2,RX} + Y$$

$$Ti_2 = t_{2,TX} + t_{2,u} + t_{21} + t_{1,d} + t_{1,RX} - Y$$

$$C = Ti_1 - Ti_2 - Y$$

$$C = \frac{1}{2}(Ti_1 - Ti_2) + \frac{1}{2}[(t_{1,TX} - t_{2,TX}) + (t_{1,u} - t_{2,u}) + (t_{12} - t_{21}) + (t_{2,d} - t_{1,d}) + (t_{2,RX} - t_{1,RX})]$$

$$C = \frac{1}{2}(Ti_1 - Ti_2) + \frac{1}{2}E$$

Any asymmetry between the two transfer directions results in the remaining error term E.

3. Spread Spectrum Techniques

In order to achieve a good interference margin whilst achieving a high accuracy, the timing signals of the clocks are transmitted using a spread spectrum technique. This means that redundancy is added to the timing data to be transmitted thus making the bandwidth B of the resulting signal much larger than the bandwidth b of the data stream. At the receiver site the signal is de-spread by removing the redundancy and lowering all interferences of the link by the spreading-factor $S = \frac{B}{b}$. Inter-modulation caused by non-linearity within the signal path is also kept at a low because the signal power is low in any given bandwidth.

3.1. Pseudo-Noise Sequences

In general spread-spectrum applications, the redundancy which is added to the data is a binary pseudo-noise sequence. In spread-spectrum applications, a PN-period like this is called a **'Bit'** and each element of the sequence is called a **'Chip'**.

This technique affords certain advantages and is of particular interest to the most important time-transfer-applications in that the phase shift of the received PN-Sequence can be measured to a very high resolution by using correlation techniques. The variance of the phase can be smaller than 1/1000....1/10000 of the period of one element of the sequence even with negative signal-to-noise-ratios at the receiver input.

The clock rate of the sequence can be easily regenerated at the receiver site. This is an important condition for coherent receiver concepts.

3.1.1 Modulation

The user data, i.e. timing-signals, time-code information or measured results can be modulated to the PN-Sequence in two different ways.

The first method would be to transmit the sequence as it is or invert it, depending on whether transmission of a binary "0" or "1" is required.

The second method is to shift a sequence for $+\frac{1}{2}T_{chip}$ and the next sequence for $-\frac{1}{2}T_{chip}$, to transmit a "1" and vice versa to transmit a "0"

In our implementation a combination of both methods is used: the first is used to transmit user data at a rate of 2 kBit/s the second to transmit the time-code signal.

The modulated PN-Sequence is itself modulated to an RF-Carrier by using (B)PSK (**B**inary **P**hase **S**hift **K**ey) or QPSK (**Q**uadrature **P**hase **S**hift **K**ey) modulation.

The required bandwidth is approx. 30% higher than the chiprate.

3.2. Generation of PN-Sequences

PN-Sequences can be easily generated using feed-back shift-registers. With a shift-register of n-stages, PN-codes with a period of 2^n-1 elements can be generated.

Figure 2 shows a simple example of how to generate a sequence with a maximal length of $2^3 - 1 = 7$ chips.

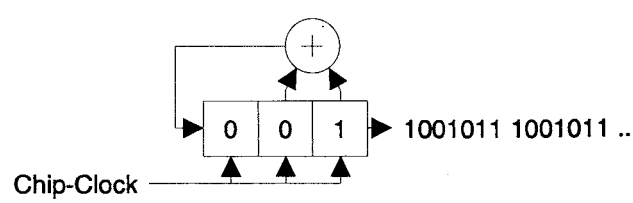


Figure 2: PN-generator with a period of $P = 2^3 - 1 = 7$ chips.

It can be shown, that a sequence of a maximal length of 2^n-1, generated with a shift register, is a pseudo-

random noise sequence. A PN-Sequence has, in general, a two-valued auto-correlation function ACF(τ) with a distinct maximum at ACF(0). The auto-correlation function ACF(τ) is calculated with the convolution of the sequence pn(τ) with the sequence pn(t-τ):

ACF(τ) = pn(τ) * pn(t-τ) (* : convolution operator)

Figure 3 shows the ACF of a PN-Sequence.

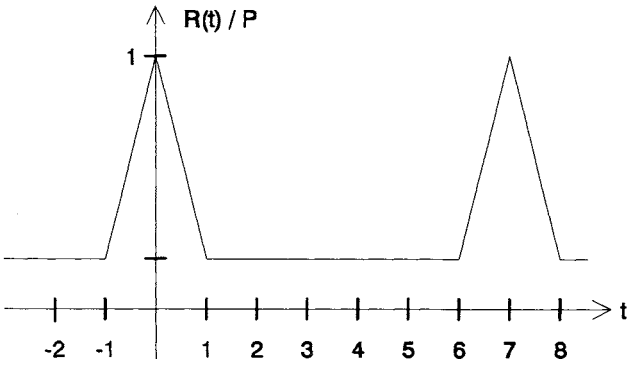


Figure 3: Auto-correlation-function of the PN-Sequence generated by the PN-Generator shown in Fig. 2

3.3 Receiver Concept

The correlation features of PN-sequences make it possible to detect the transmitted PN-Sequence and synchronise it exactly to the PN-Sequence generated at the receiver. This synchronisation is possible even with negative Signal-to-Noise-Ratios in the RF bandwidth.

The transmitter modulates a PN-Sequence, that itself is modulated with user- and timing-data, using PSK onto a sub-carrier. For example the widely used 70 MHz.

This PN-Sequence has to be de-spread at the receiver-site, demodulating the data and synchronising it to an identical PN-Sequence generated by the receiver.

Figure 4 shows a very simplified schematic of a tracking loop receiving and de-spreading the signal.

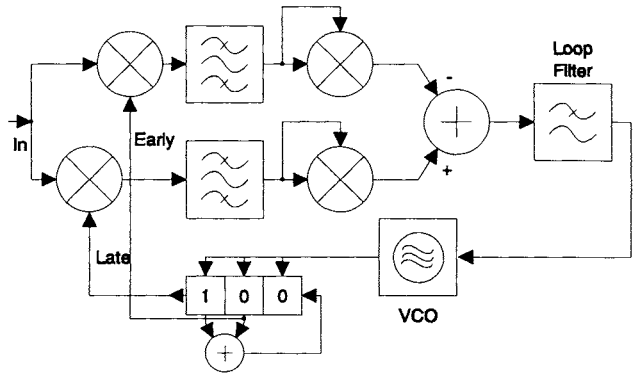


Figure 4: Delay Lock Loop (DLL).

The input signal is processed in two different signal paths. Within the first stage, de-spreading is taken with PN-sequences with a phase difference of one chip-period. The de-spread signal is passed through an intermediate-frequency section and is demodulated by squaring.

The control voltage of the VCO is derived from calculating the difference between the ACF's of the early and the late PN-Sequence and the received sequence. The system is a special type of phase-locked-loop system, called Delay Lock Loop (DLL).

This is depicted in figure 5.

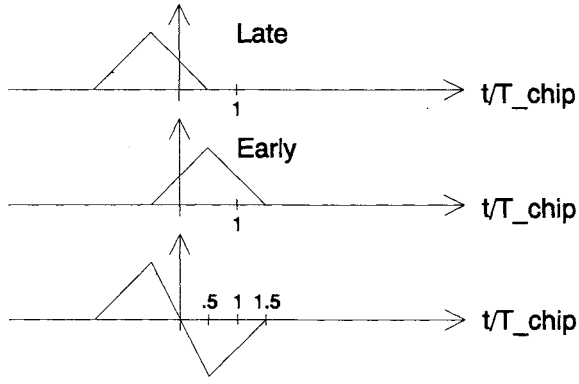


Figure 5: Control-Characteristic of a DLL

In a locked state the phase of the received PN-Sequence lies exactly between the phases of the early and the late sequence. This means the phase differences are $+/- \frac{1}{2} T_{chip}$.

Thus the original PN-Sequence generated at the transmitter-site is now available as a clean duplicate but free of noise and with a very low jitter value.

The precision of this tracking loop facilitates a phase resolution of the PN-Sequence better than 1/1000 of one chip period. With proper digital post-processing, even higher resolutions of up to 1/10000 and more may be reached.

The tracking errors, resulting from the second method for modulating a PN-Sequence described above are removed by interrupting the time interval measuring process during the short transmission period (16 ms) of the time-code signal.

4. Two-Way Time-Transfer with MITREX and SATRE

In 1983 the University of Berlin developed and manufactured a time-transfer-modem using the spread spectrum technique. The MITREX-Modem (Microwave Time and Range Experiment) was designed to operate with Telecom-Satellites and ground-stations using a carrier frequency of 70 MHz (see Fig. 1).

The MITREX-Modem uses PN-Generators with a chip-rate of 2,5 MHz and a truncated PN-Code of 10000 chips. The truncation of a PN-Code of maximal length causes a slight degradation of the auto- and cross-correlation features but reduction of the signal structure by the power of 10 greatly enhances the circuit design.

The typical variation of the internal delay of the MITREX-Modem is 120 ps with an accuracy better than one nano-second.

The SATRE (**SA**tellite **T**ime and **R**ange **E**quipment)-Modem introduces great improvements in accuracy and usage and can be run optionally at reduced chip-rates in a MITREX compatible mode to communicate with MITREX-Modems.

The improvements in resolution and accuracy are achieved by increasing the chip rate to 10 MHz and lowering the jitter of the internal delay four times to 30 ps. The accuracy is approx. 500 ps, provided that the system is carefully calibrated . The chip-length is still kept at 10 000 chips.

This theoretical improvement based on the four times higher chip rate, was proved in practice.

The correlator is a completely analogous design and is based on the so called "non-coherent" delay lock-loop principle, which is not susceptible to frequency and phase instabilities of the RF-Link.

The insusceptibility to phase errors makes it possible to modulate the PN-Sequence at the receiver by inverting or not inverting it, depending on the next digital bit of the user data stream.

The higher chip rate and a re-designed code-search circuitry result in an improved code-acquisition time of about 30 s maximum.

The increase in bandwidth is no disadvantage because the signal is 20 dB below the thermal noise.

The system automatically tunes itself to the carrier frequency by the use of a PLL, locking the LO of the intermediate section to the received carrier frequency.

The frequency scheme of the receiver is carefully designed. Integral multiples of the chip rate frequency have been strictly avoided to keep errors due to beat frequencies as low as possible. Another improvement to the operation of the modem along with the fast lock-in-time is the possibility to transmit user data (i.e. real-time measurement results) at the same time as the time-code is transmitted. This greatly simplifies the coordination of the two partner stations and exchange of measurement data.

The time interval counter is now an integral part of the unit. Its results can be read by any desk-top computer.

5. Results

Figure 6 shows the test set-up. The modulated carrier is attenuated, noise is added and the signal is fed to the input of the receiver. To test the instrument with varying Signal-to-Noise density ratios ($\frac{C}{N_0}$), the power of the noise source can be changed over a wide range with a second attenuator.

Care must be taken when changes of the delay on the pico-second level are measured. The delay within the variable attenuator in the signal path is not constant over the attenuation range! Changes of up to 100 ps occurred. Errors at this level have to be carefully resolved by using different types and combinations of attenuators.

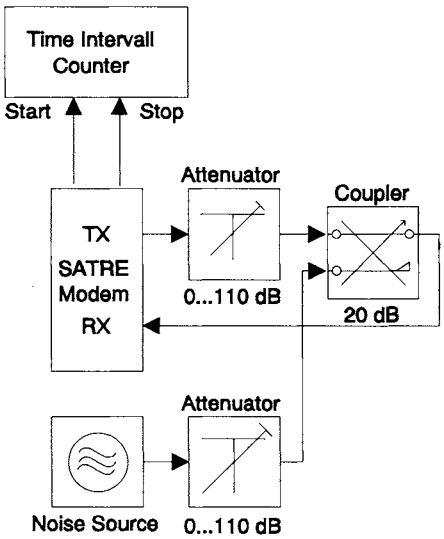


Figure 6: Test set-up

The test set-up should be done very carefully as any echoes and crosstalks distort the results.

Figures 7a, b, c, and 8 show the performance of the modem under varying input power and $\frac{C}{N_0}$ ratios.

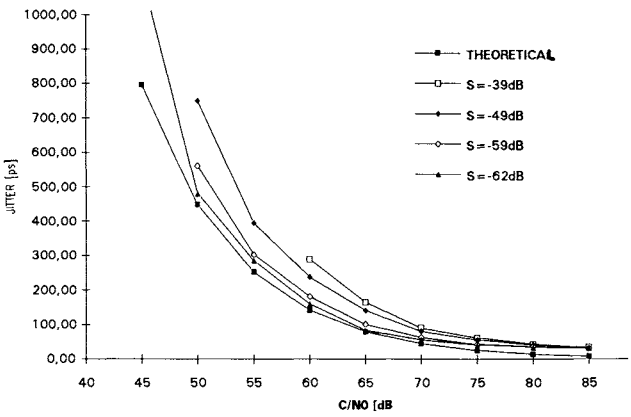


Figure 7a: Instrument-Jitter

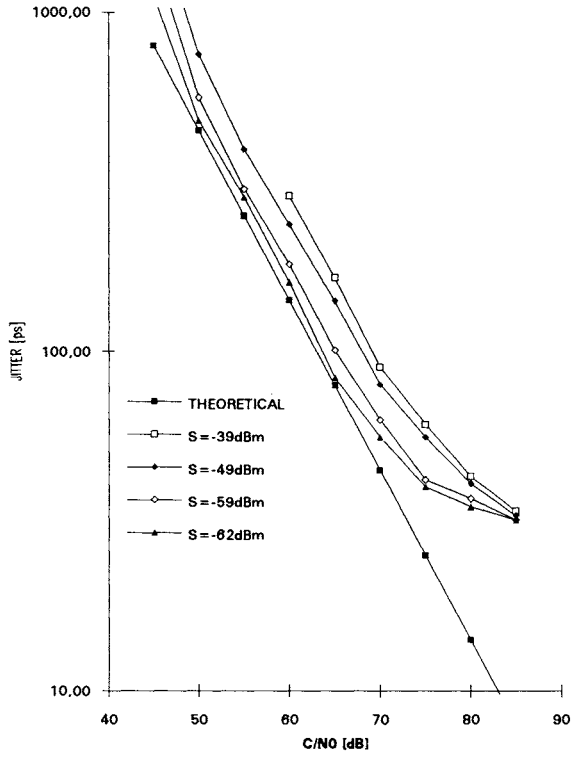


Figure 7b: Instrument-Jitter, logarithmic scaling

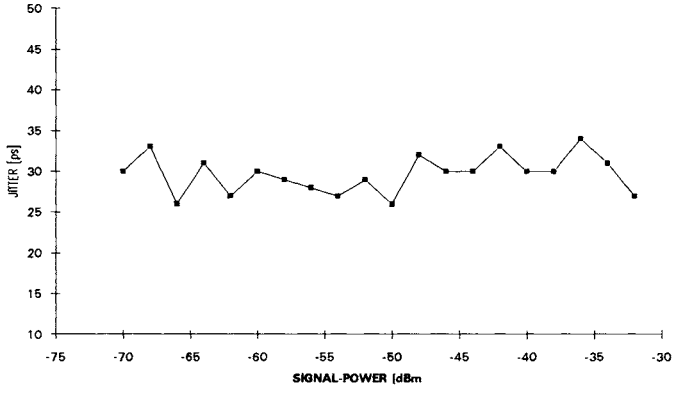


Figure 7c: Instrument-Jitter with infinite $\frac{C}{N_0}$

In Fig. 7b the results are shown in a logarithmic scale. Due to the internal noise of the modem and the time-interval counter the jitter of the modem with a noise-free signal is about 30ps (Fig. 7c)

At low signal-to-noise-density ratios clipping within the tracking-loop causes an increase of the jitter far exceeding the theoretical values. This situation however is easily avoided by proper selection of signal amplitudes.

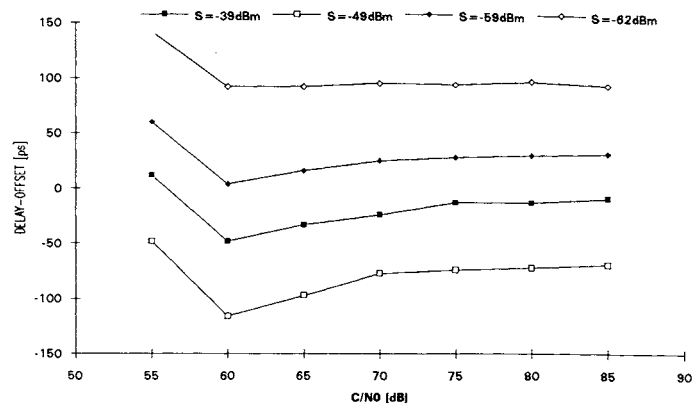


Figure 8: Delay-variation depending on signal-power level and noise-density. The signal-power was adjusted with the variable attenuator described in the text. This causes delay-changes of ca. 50 ps. The noise-density-ratio was only changed by different attenuations of the noise-power, therefore that there are no delay-changes within each curve.

6. Conclusions

Further improvements to the jitter of the internal delay can be expected in the near future.

A further spreading by using a chip rate of up to 80 MHz would theoretically lower the variance to the measured time interval to about 4 ps.

Digital signal processing within the instrument promises further improvements so that an accuracy of 20 ps without an increase in the chip rate should easily be possible within this decade.

REFERENCES

/1/ Jack K. Holmes, Coherent Spread Spectrum Systems, John Wiley & Sons, 1982.

/2/ Karl-Martin Müssener, "Optimierung von Tracking Loops hinsichtlich der Anwendung im PRARE", Thesis, Berlin 1983.

/3/ Philipp Hartl, Wolfgang Schäfer, "Application of Spread Spectrum Techniques in Combined Use for Navigation and Telecommunication in Satellite Systems", FREQUENZ No. 41 & 42, 1987/88.

/4/ Institute for Navigation University Stuttgart, "A Modem for Microwave Time and Ranging Experiments via Telecomunication Satellites", Stuttgart 1985.

SYNCHRONISATION, ACQUISITION AND TRACKING FOR TELEMETRY AND DATA RECEPTION

A. VAN DONINCK

ALCATEL BELL – SPACE GROUND SYSTEMS

Abstract :

Synchronisation, acquisition and tracking are functions vital for the reconstruction of the transmitted information at the reception side. For certain in the Space environment, with receiving conditions of signals which are often corrupted with noise and interfering signals, the quality of the above functions will reflect undoubtedly on the total performance of the system.

In order to secure an error free transmission (or close to the theoretical value) of the useful data to the communication centres, all three functions have their typical parameters to guarantee the overall performances.

We will address these important parameters while highlighting each of the functions separately. The followed sequence is such as the functions occur in the system in time and that for the type of data to be received. Their will be a distinction between telemetry and data reception, between direct carrier modulation or the use of a subcarrier, between deep space and normal reception.

For the telemetry reception the acquisition is described taken into account the difference in performances as we distinguish geostationary or polar orbits and the dependencies on the different doppler offsets and rates. The paper will cover the related functions and parameters and summarize the specifications of an average receiver. Further the synchronisation of the valid data is described with a distinction for data directly modulated or via a subcarrier, the type of modulation and bitrate. Again the relevant functions and parameters of the average receiver/demodulator will be summarized. As a third function the tracking of the signal in the course of the operational phase is described and relevant parameters of an actual system will be presented.

As a following step the reception of real data is handled and as for the first part again a sequence of acquisition, synchronisation and tracking is applied. Here higher bitrates and direct modulation schemes play an important role. Again on the market equipment with the relevant parameters are discussed.

The last part covers the three functions in cases were deep space reception is needed. Here the high performance receiver/demodulator functions are explained and how the acquisition, synchronisation and tracking is handled in such application.

1. INTRODUCTION

In order to perform satellite data acquisition within a satellite groundstation different tracking, acquisition and synchronisation sequences are to be executed.

This paper covers the different sequences as they are performed in time and this for signals with different characteristics as there are telemetry (housekeeping) and real data transmission.

In general the different sequences to be performed can be distinguished in antenna tracking, carrier acquisition and tracking and bitsynchronisation.

Mainly the carrier acquisition and tracking function and bitsynchronisation function will be highlighted.

2. ANTENNA TRACKING

In order to acquire the satellite data, the antenna has first to be pointed towards the satellite. Once the satellite signal is received, an automatic antenna pointing mechanism has to assure tracking on the satellite signal.

In general within a ground station a tracking receive chain is used from which the antenna pointing errors are derived. The antenna pointing errors are converted to error voltages to pilot the antenna servo and drive mechanism.

A difference can be made between geostationary satellites and polar satellites.
For geostationary satellites, where the satellite is at a fixed and predetermined position, the antenna can be pointed by a programm track mechanism.
However for polar satellites, being in constant motion with regard to the groundstation, antenna tracking requires a more complex acquisition scheme. First a prepass positioning has to be performed using coordinates derived from an orbit generator. Second after acquisition of the satellite signal the antenna groundstation tracking and servo system has to assure the antenna to cover the trajectory followed by the satellite during the pass. For this the antenna pointing error signals are to be derived.

One method to derive the error signals is to use a monopulse autotrack system. In this system four antenna feeds are used and the signals from the four feeds are combined in order to deliver three error signals, a sum signal (sum of the four feeds) and two difference signals (ΔEL and ΔAZ). It is clear that optimum satellite tracking is achieved when the sum signal is maximum and the difference signals minimum. The three error signals are down converted to a IF frequency and applied to a 3-channel tracking receiver that generates elevation and azimuth error signals for the servo and drive mechanism.

As a variant on this monopulse tracking system one could use an antenna feed from which the phase of the output signal can be shifted in order to change the antenna beam direction and to become the antenna pointing errors.

Once antenna tracking is performed the data acquisition sequence can be started.

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

3. CARRIER ACQUISITION AND TRACKING

3.1 TELEMETRY CARRIER ACQUISITION AND TRACKING.

3.1.1 General properties of a Telemetry signal.

Most often the telemetry housekeeping data is PM/PSK/PCM modulated on a RF carrier signal of 2 GHz. This means that the PCM telemetry data modulates a subcarrier in PSK, which on his turn phase modulates the RF carrier.
For telemetry signals we will restrict our considerations within this paper to the carrier tracking and acquisition properties of a PM/PSK/PCM modulated signal.

In order to have an idea of the characteristics of the Telemetry signal it should be noted that subcarrier frequencies used for telemetry housekeeping data are limited up to 1.2 MHz and bitrates are situated in the range of 40 bps up to 1.2 Mbps. Moreover a fix division relationship exits between the subcarrier frequency and the bitrate.

3.1.2 Carrier acquisition sequence.

In order to demodulate a PM/PSK/PCM modulated signal a phase demodulation of the received carrier has to be performed. The appropriate circuitry to perform this function is a Phase Locked Loop.

Telemetry carrier acquisition of the received Telemetry RF carrier by the Phase Locked Loop circuitry requires for frequency and phase acquisition. This function is performed by generating a frequency sweep on a voltage controlled oscillator.

Within the ground station the received telemetry carrier is first amplified by means of a Low Noise Amplifier and further downconverted to an IF frequency of 70 MHz.

Carrier acquisition is performed at 70 MHz using a Voltage controlled oscillator generating a frequency of 60 MHz, a mixer, a phase detector at 10 MHz, a selectable loop filter and the appropriate acquisition circuitry.

The received and down converted 70 MHz telemetry signal is mixed with a 60 MHz sinewave generated by a VCXO. The resulting 10 MHz signal is applied to a phase detector receiving also a 10 MHz signal derived from the station frequency reference. The resulting phase error signal is first filtered within the loop filter and then applied as an error voltage to the 60 MHz VCXO. Hence a 10 MHz PM/PSK/PCM signal is obtained coherent with the station frequency reference.

During the acquisition sequence the VCXO is controlled by both the phase detector output and a sweep signal that brings the received signal within the pull-in range of the loop. When frequency acquisition occurs and after some decision time, the sweep signal is switched off, this command being derived from a correlation detector.
Hence the sweep generation is stopped without altering the actual VCXO voltage. From this moment the VCXO is only controlled by the phase detector output in order to obtain phase acquisition and to guarantee phase coherency (tracking) of the received signal with the station frequency reference.

3.1.3 Carrier acquisition performance

It has to be noted that in general a Telemetry receiver is a multipurpose receiver, making the receiver suitable for use in a multipurpose groundstation and being able to acquire signals from different satellites by simple reconfiguration of its operational parameters. This means that the receiver has to be capable to work in a very extensive environment with respect to signal to noise ratios, doppler shifts, doppler rates, input levels, modulation indices, etc.
Hence a number of parameters of the receiver must be configurable in order to obtain optimal performance during the reception of signals with different characteristics.

Some of the parameters have an impact on the carrier acquisition performance of the receiver.

- First of all there is the received signal to noise ratio which determines maximal loopbandwith one can select in order to guarantee the phase locked loop to lock and to remain locked with the required probability.
  A rule of dump tells us that a minimum signal to noise ratio of 10 dB, within the loop bandwidth, is a necessary condition to guarantee a lock probability of 99 %. Hence for optimisation of the acquisition and tracking performance a multipurpose Telemetry receiver will be equipped with a number (in general 4) of selectable loopfilters. Very often a 30 Hz, 100 Hz, 300 Hz and 1000 Hz loopfilter is provided within the Telemetry receiver.

- Acquisition of the Telemetry carrier is obtained by performing a sweep on the 60 MHz voltage controlled oscillator.

  The higher the sweeprate, the faster the received frequency is found and the faster the acquisition is performed. So fast frequency acquisition requires a high sweep rate. However when a frequency ramp is applied on a signal within a phase locked loop, the phase locked loop follows the phase variations, but a phase error will be generated within the loop. It is important to keep this error within proportion in order not to unlock the loop. In general a maximum phase error of 1 radian is supported during the sweep.

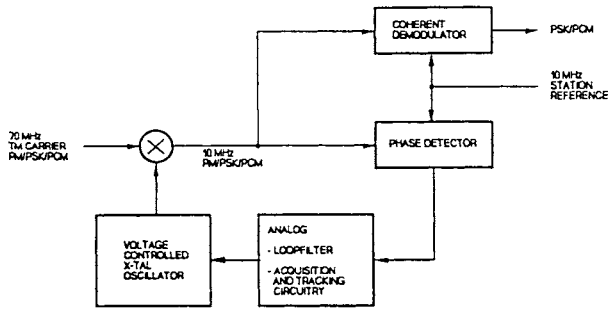
  This makes clear that each of the selectable loop bandwidths need to have its appropriate sweep rate. This in order to obtain an optimal acquisition time for the actual configuration of the receiver.

  As an example a 1000 Hz loopfilter permits a frequency sweeprate of 80 kHz/s, while 30 Hz loopfilter only permits a sweep rate of 80 Hz/s.

- We can resume the impact of the signal to noise and sweep rate on the acquisition performance of the Telemetry receiver as follows : as the signal to noise ratio decreases, a smaller loop filter has to be selected in order to guarantee lock-in of the phase locked loop and hence the acquisition time becomes quite large. Therefore two properties are implemented within the Telemetry receiver in order to gain up the acquisition time; predoppler offset and selectable sweepwidths.

First during prepass orbit calculations a prediction can be made of the expected doppler shift at the moment acquisition is started. This offset can be programmed on the VCXO in order to shorten the sweeptime during frequency acquisition.
Moreover as the expected dopplershift is calculated and programmed on the VCXO, only a frequency sweep over a limited sweepwidth is required to perform frequency acquisition. Therefore selectable sweepwidths are provided. Commonly used sweepwidths are ±12.5 kHz, ±50 kHz and ±150 kHz.

- Finally attention has to be paid to avoid antiside band acquisition. As the Telemetry carrier is phase modulated, sidebands next to the carrier exist on which the phase locked loop can lock. In order to prevent the phase locked loop to remain locked in this false lock condition, it is important to detect the false lock condition and to continue the frequency sweep.



Tracking and Acquisition circuitry of an analog TM receiver.

3.1.4 Carrier acquisition performance of digital Telemetry receivers.

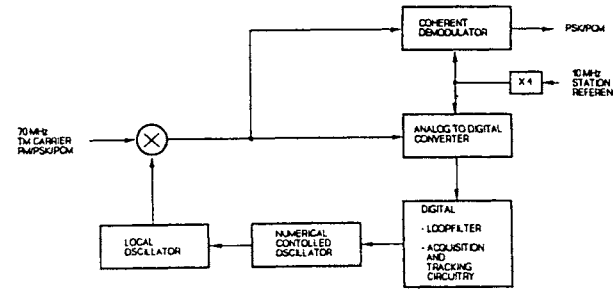
With respect to the first generation of Telemetry receivers, which were mainly concipiated using analog circuitry, the increase of speed and complexity of logic integrated circuits made it possible to develop a new generation of Telemetry receivers, using digital hardware implementations for the phasedetector, the acquisition and tracking circuitry and the local oscillator.

Instead of a classical analog phasedetector a high speed analogue to digital converter, clocked with a 40 MHz signal derived from the station reference, is used to sample the received signal at 10 MHz. Hence at the output of the analog to digital converter the phase error information is available in digital format.
This information is used within a LCA (Logic Configurable Array) containing the digital implementation of the acquisition (sweep generation) and tracking (loopfilter) functions. The LCA delivers at his output an error word of 32 bits, which is used to update the phase accumulator register of a numerical controlled oscillator (NCO).

The 60 MHz Voltage controlled oscillator is replaced by a combination of a fixed local oscillator and a the numerical controlled oscillator.

In order to guarantee acquisition and tracking of the phase of the received signal the 32 bit phase accumulator value of the numerically controlled oscillator is permanently updated using the 32 bit information word generated by the acquisition and tracking circuitry within the LCA.

As explained earlier in the paper a lot of operational parameters are to be selected in the telemetry receiver in order to obtain optimal performance of the receiver in function of the characteristics of the received signal (signal to noise ratio, loopbandwidth, sweeprate, sweeprange, etc).
Therefore within an analog receiver a limited number of bandwidths, sweeprates and sweepranges is selectable.
The great advantage of the digital ranging receiver with respect to the analog receivers, besides the gain in alignment effort, is the full programmability of the operational parameters (loopbandwidth, sweepranges, sweeprates).



Tracking and Acquisition circuitry of a digital TM receiver.

3.1.5 Future developments

Within the next future further implementations of digital signal processing techniques, such as Fast Fournier Transformation, could be envisaged in order to perform carrier acquisition and antiside band lock detection.
Performing a Fast Fournier analysis on the received spectrum, makes a frequency prediction of the received carrier frequency possible. Hence no sweep, or only a sweep with limited sweepwidth, has to be performed. Moreover correct prediction of the received carrier frequency makes sideband acquisition inexistant.

3.1.6 Acquisition and tracking specifications of an average digital Telemetry receiver.

. Input Frequency : 70 MHz, ±1 MHz tunable

. Impedance : 50 Ohms, VSWR : 1.5

. Level
  - Signal : -20 dBm to -100 dBm
  - Noise : -125 dBm/Hz maximum

. AGC
  - coherent carrier demodulation
  - range 80 dB
  - time constant programmable 1 to 1024 msec (in steps of 2^n with n: 0 to 10).

. Functions

  - mode : Phase locked loop (PM)

  - Automatic acquisition
    It is possible to search the signal around a frequency tunable from frontpanel or remotely over ±1 MHz.

  - Automatic tracking range : ±350 kHz.

- Loop bandwidth

  Programmable from 30 to 1000 Hz (in steps of $30x2^n$ with n=0 to 6)

. Threshold

  10 dB carrier to noise ratio within the loop bandwidth.

. Automatic search range

  Programmable up to $\pm$350 kHz with a resolution of 1 kHz.

. Search rate

  depends on the selected loop bandwidth.
  e.g. 80 kHz/s for 1000 Hz
  0.08 kHz/s for 30 Hz

. Acquisition probability : > 95%

. Probability to keep locking : > 99%

. Anti side band lock

  Threshold C/N $\geq$ 42 dBHz for either the carrier or the telemetry subcarrier.

3.2 REAL DATA CARRIER ACQUISITION

3.2.1 General properties of a Real data signal.

Besides the Telemetry housekeeping data a satellite also transmits real data obtained from on board measurement instruments.

With respect to Telemetry data Real data transmission contains a larger amount of information. Higher bitrates, going from 4 Mbps up to more than 100 Mbps are applicable for this data transmission and direct carrier modulation schemes, such as QPSK, UQPSK, BPSK are used.
As a result from the direct carrier modulation scheme the spectrum of the transmitted real data does not contain any carrier information. Hence no classic Phase locked loop circuitry can be used to perform acquisition and tracking.

Due to the higher bitrates real data receivers are not multipurpose, but mission defined.

3.2.2 Carrier acquisition sequence

As for the reception of the Telemetry data, the modulated real data carrier is first amplified in a Low Noise Amplifier and then further downconverted to an appropriated IF frequency.

It has to be noted that due to the higher bitrates and accordingly the larger bandwidth occupation of the real data signal, the RF carrier frequency and also the IF frequency are situated in a higher frequency band than for the Telemetry signals. As Rf carrier very often an 8 GHz frequency is used, while the IF Frequency at which carrier acquisition and demodulation is performed, is situated in the 300 MHz to 1 GHz band. Very frequently used IF frequencies are 375 MHz, 620 MHz and 720 MHz.

Due to the absence of any carrier information within the spectrum of a directly modulated signal, acquisition and tracking of a the real data signal has to be performed using carrier recovery techniques, instead of a classic phase locked loop.

The most commonly used and most appropriate carrier recovery technique for high rate real data carrier recovery is the Baseband Multiplication Algorythm. The operation of this algorythm is very similar to a COSTAS loop, exept that an analog multiplication is performed on the baseband data instead of a decommutation.

Carrier acquisition is performed at the IF frequency, using a voltage controlled surface acoustic wave oscillator (VCSO), demodulation circuitry containing a power splitter, a 90 degree phase shifter and two double balanced mixers, baseband multiplication circuitry and tracking and acquisition circuitry (loopfilter, sweepgenerator, lock in indication).

The received and down converted real data carrier is splitted in two by a 0 degree 2-way power divider. The resulting two equal in phase signals are applied to two double balanced mixer. The local oscillator signal derived from a voltage controlled SAW oscillator is split, with a 90 degree phase difference. The inphase and inquadrature signals of the VCSO are also applied to the double balanced mixers.
As a result the double balanced mixers generate within the inphase and inquadrature branch a baseband signal which is filtered out using a low pass filter.
In order to derive the phase information of the recovered carrier two baseband multiplications are performed. The first baseband multiplication is the result of the multiplication of the signal in the inphase channel multiplied with the signal in the inquadrature channel. The second baseband multiplication is the result of the multiplication of the signal in the inquadrature channel multiplied with the signal of the inphase channel. Both multiplier outputs are subtracted and hence a phase error voltage is derived. The voltage is applied to a loop filter and used as control voltage of the Voltage controlled SAW oscillator.

During the acquisition sequence the VCSO is controlled by both the phase detector output and a sweep signal that brings the VCSO signal within the pull-in range of the loop. When frequency acquisition occurs and after some decision time, the sweep signal is switched off, this command being derived from the bitsynchronizer lock indication.

Hence the sweep generation is stopped without altering the actual VCXO voltage. From this moment the VCXO is only controlled by the phase detector output in order to maintain phase acquisition and to guarantee phase coherency (tracking) of the recovered carrier signal with the received signal.

3.2.3 Carrier acquisition performance

As real data receivers are concipiated for high bitrates and hence mission specific, a number of features as described for the multipurpose telemetry receiver are not implemented within a high rate data receiver.

For specific missions, the satellite to ground station link budgets are fixed within well known limits, max doppler rates and max doppler shifts can be calculated. Due to the knowledge of these parameters the receiver loopbandwidth and sweeprate are well defined and fixed.

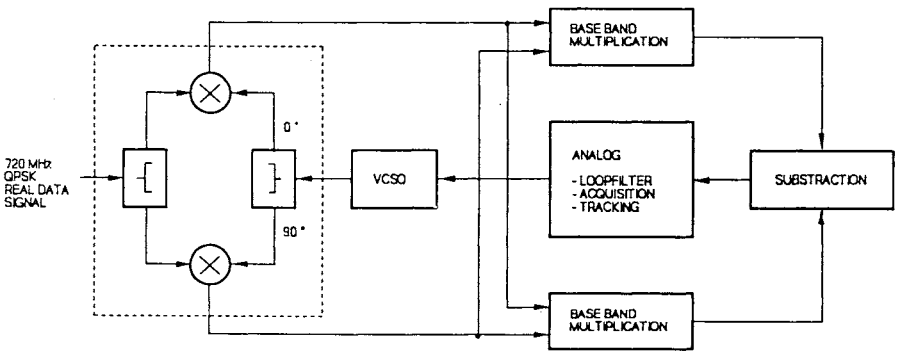
In order to speed up acquisition time, just as for the multipurpose telemetry receiver, predoppler offset and selectable seepwidths are implemented within the high rate receiver.

Due to the direct carrier modulation no modulation sidebands will be present within the real data spectrum and antisideband lock circuitry is not needed within the real data receiver.

3.2.4 Acquisition and tracking specifications of a high data rate receiver.

ERS-1 High rate RECEIVER

.	Input frequency	: 720 MHz ± 800 kHz
.	Dynamic range	: -10 to -50 dBm
.	Modulation	: QPSK
.	Bitrate TM	: 2 x 52.5 Mbps
.	BER degradation	: < 2 dB, differential decoding loss included.
.	Threshold	: E/No = 6 dB
.	Acquisition time	: less than 300 ms
.	Threshold	: Eb/No ≥ 6 dB
.	Search range	: 4 possible search ranges can be selected : ±800 kHz, ±200 kHz, ± 50 kHz, manual



Tracking and Acquisition circuitry for a QPSK real data receiver.

4. DATA ACQUISITION AND SYNCHRONISATION

Once carrier acquisition has been performed first the modulated carrier has to be demodulated followed by the data acquisition and bitsynchronisation.

Due to the space environment the received data bits are corrupted with white Gaussian noise. To obtain a bittransmission performance as close as possible to the theoretical performance a matched filtering is performed on the received bits. The matched filtering, being equivalent to an integration of the received noise corrupted data during one bitperiod, results in a maximum signal to noise ratio at the end bit moment.

As there is an obvious relationship between bittransmission performance and the decision moment, a bitclock restitution (recovery of the bitclock out the received data stream) followed by bitclock acquisition and tracking are to be performed.

4.1 Telemetry data acquisition

For Telemetry data where in general a PM/PSK/PCM modulation scheme is used two consecutive demodulations are to be performed.

The consecutive demodulations can be executed in two possible ways.

Either the classic way is followed ; first a phase demodulation of the PM modulated carrier is effectuated within the Telemetry receiver, second a PSK demodulator performs the subcarrier acquisition, tracking and demodulation, finally a bitsynchronizer reconstructs the bitclock, executes bitclock acquisition and tracking and bitdecision.
Or a more complex but also more performant way can be followed, which consists in a subcarrier demodulation before carrier demodulation. Hence a carrier is obtained, modulated in BPSK with the telemetry data. Further PSK carrier demodulation and bitsynchronisation are to be carried out.

In both cases the final step is a PSK demodulation, requiring a (sub)carrier acquisition and tracking, followed by a bitsynchronizer performing a bitclock restitution, bitsynchronisation and bitdecision.

It is clear that different solutions and algorithms exist to perform sub(carrier) recovery and bitclock restitution.
Within this paragraph it is the intention to highlight the algorithm and associated specifications of a High Performance Demodulator. Although this demodulator being multipurpose, it was specially designed for the reception of the Telemetry data during the deep space Giotto mission.

In order to become an optimal tracking, acquisition and BER degradation, in this receiver the bitsynchronizer functions (bitclock restitution, matched filtering and bitdecision) are embedded within the (sub)carrier recovery circuitry.

4.1.1 High Performance demodulator

4.1.1.1 Matched Filter

The bitrates used for Telemetry data transmission situated within the range of 40 bps up to 1.2 Mbps make it possible to implement the matched filter function in digital hardware.

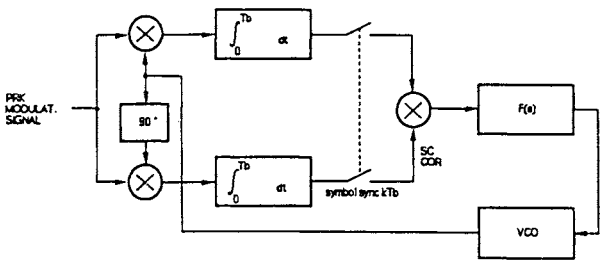
In order to perform the matched filtering a digital integrate and dump filter is implemented. The transmitted data are sampled with a clockfrequency of 40 MHz. The sampled bits are accumulated during the bitperiod and at the end sample moment the accumulator register is read out and reset to zero.

4.1.1.2 Subcarrier Carrier acquisition and tracking

Within the High performance demodulator the subcarrier synchronisation is realised with a Costas loop with active arm filters. Functionally the working of this Costas loop is identical to the QPSK costas loop described in par 3.2 for the acquisition and tracking of the high rate real data transmission.

Note that only one baseband multiplication is performed because the modulation scheme is PSK instead of QPSK. Moreover in order to gain in performance machted filtering is performed within the Subcarrier loop and baseband multiplication is performed on the data sampled at the optimum sample moment.

In order to perform the reception of telemetry data within a multipurpose Telemetry chain the subcarrier frequency and the digital subcarrier loopfilter are programmable.

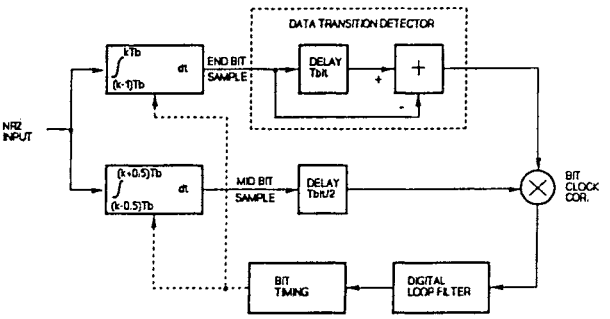


Costas Loop with active arm filters.

4.1.1.3 Bitclock synchronisation

Bitclock restitution is performed within the High Performance demodulator using a Linear Multiply tracking loop, which differs slightly from a Data transition tracking loop in the absence of a bitdecision element before the data transition detector.

Two integrate and dump filters are implemented.
The first one makes an integration from midbit to midbit. If a data transition occurs and if perfect tracking of the bitclock is performed the output of this integration will be zero. A phase error between the bitclock of the received datastream and the internal generated bitclock will generate an offset voltage.
The second integrate and dump filter performs an end bit to end bit integration. In order to prevent that the phase of the bitclock would be updated with this error voltage it is necessary to detect whether there has been a datatransition or not. This is realised by comparison of the actual and the previous output of the end bit to end bit integrate and dump filter.
To guarantee multipurpose operation bitclock frequency and bitclock loopbandwidth are programmable.



LMTL & DTTL Bit Synchronizers.

4.1.1.4 Specifications of the High Performance Demodulator

. INPUT SIGNAL :

- carrier frequency : 10 MHz
- carrier power : - 66..-60dBW
- C / No : ≤ 27 dBHz
- Eb / No : ≥ -3 dB

. MODULATION :

- PCM(NRZ)/PSK/PM :
  - Subcarr.freq. : 2kHz..1.2 MHz
  - Waveform : sine & square
  - Bitrate : 40 bps..1.2 Mbps (NS = 1..1024)
- PCM(SPL)/PSK/PM :
  - Subcarr.freq. : 2kHz..1.2 MHz
  - Waveform : sine & square
  - Bitrate : 40 bps..1.2 Mbps (NS = 2..1024)
- PCM(SPL)/PM :
  - Bitrate : 40 bps..600 Kbps

MODULATION INDEX :

- Sine wave : 0.3..1.7 rad
- Square wave : 0.3..1.1 rad

PERFORMANCE :

- bitrates < 300 bps
  - . Eb/No < 1dB : ≤ 0.4dB
  - . Eb/No <11dB : ≤ 0.2dB
  - . Eb/No >11dB : BER $<10^{-6}$

- bitrates > 300 bps : ≤ 0.5dB

THRESHOLD :

the acquisition threshold is ≤ -3 dB (optionally -9 dB for restricted bitrates)

4.2. Real data acquisition

Having the COSTAS loop in mind described within par 3.2 'Real data carrier acquisition' it is clear that after carrier acquisition has been performed, we find in the inphase and inquadrature branches the noise corrupted baseband data.

An error free, or close to theory, bitstream is obtained within a bitsynchronizer performing bitclock restitution and bitsynchronisation. Therefore a bitsynchroniser is inserted in the inphase and inquadrature branch.

Within the bitsynchroniser the received data stream is first matched filtered.
Due to the higher bitrates used for real data transmission a digital implementation using integrate and dump matched filters is no more possible. However an analog equivalent of the digital integrate and dump filter can be implemented. Such a matched filter, having a roll off matched to the spectral shape of the received data, is bitrate dependant, making the real data receiver mission specific.

For the moment off the shelf available receivers use a Delay and Multiply algorithm for bit clock restitution. This algorithm is selected for its implementation simplicity, even on high data rates.

The Delay and Multiply algorithm consists of multiplying the input data with a version of the input data delayed by half a bitperiod. As a result of the multiplication a spectral line is generated at the bitrate. Further a phase locked loop is implemented within the bitsynchronizer to lock the internally generated Voltage controlled oscillator clock on the reconstructed bitclock. Hence a phase coherency is realised between the internally generated bitclock of the VCO and the bitclock of the received data stream. The reconstructed bitclock is then used for bitdecision.

4.2.1 Future developments

Due to the availability of logic circuitry (ECLIPS) working at frequencies above 100 MHz it becomes now possible to use the previously described DTTL loop in order to perform bitclock extraction within the high rate data receivers.

5. CONCLUSION

As highlighted within the paper a number of acquisition, tracking and synchronisation sequences are to be executed within a groundstation in order to perform acquisition of the data received from a satellite.

Within the large area of tracking, acquisition and synchronisation a few points were picked out in order to make clear that each of the functions has its own requirements with respect to the modulation scheme, the signal to noise ratio, the acquisition time, and maximum dopplershift.

Moreover a short view was made of acquisition sequences as they can be performed within the next future using high speed logic and digital signal processing techniques.

Session 5A

Quartz: Oscillators

Chairman: E. Graf

A MICROCOMPUTER COMPENSATED CRYSTAL OSCILLATOR MCXO

P. COLLET G. PUTOIS F. VIGNAUD

COMPOSANTS QUARTZ ET ELECTRONIQUE C.Q.E.
B.P.16 , rue Keller , Zone industrielle
10150 PONT SAINTE MARIE

FRANCE

FAX N° (33) 25.80.34.57

ABSTRACT

Modern communications systems demand a more stable frequency than currently available with a TCXO and consuming considerably less power than a OCXO.

The MCXO ideally fills this requirement, frequency stability of +/ -20 ppb over the military temperature range consuming less than 50 mWatt being achievable.

INTRODUCTION:

This paper describes a microcomputer compensated crystal oscillator (MCXO) using a resonator operating in a dual mode [1], [2], [3]. The crystal oscillator functions at two frequencies simultaneously, the fundamental Fc mode and the third overtone of the Fc mode. The difference between the fundamental frequency and the third overtone frequency divided by three provides the temperature information to the compensation system. For a SC-CUT resonator operating at 10 MHz, the slope of frequency with temperature of the beat frequency is of the order of 15 Hz/°C.

The output frequency of the dual mode oscillator is compensated through a Direct Digital Synthesizer. An initial calibration is necessary, this is done by determining the relationship between the temperature and the beat frequency, this information is then used to programme the DDS (The coefficients are stored in the EPROM of the microcomputer).

The upper frequency of the system is obtained using a phase locked loop (PPL) inherently this gives good phase noise performance.

We have obtained the best results to date using four forth order polynomials. A frequency temperature stability of better than +/- 20 ppb over the temperature range of -30°C to +85°C has been achieved.

The target specification of the system is as follow:

OUTPUT FREQUENCY	1Mhz, 5Mhz or 10 MHz
LONG TERM STABILITY	5.10-9 / month
TEMP. STABILITY	-30°C to 85°C < 2.10-8 -55°C to 125°C < 1.10-7
SUPPLY VOLTAGE	3 volts<Va<5 volts
VOLTAGE STABILITY	for Va +/-10% < 2.10^{-8}
AVERAGE CONSUMPTION	< 50 mW

This paper will give the current status and future timescales of the project.

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

The figure 1 gives the MCXO block diagram developed by TEKELEC/CQE.

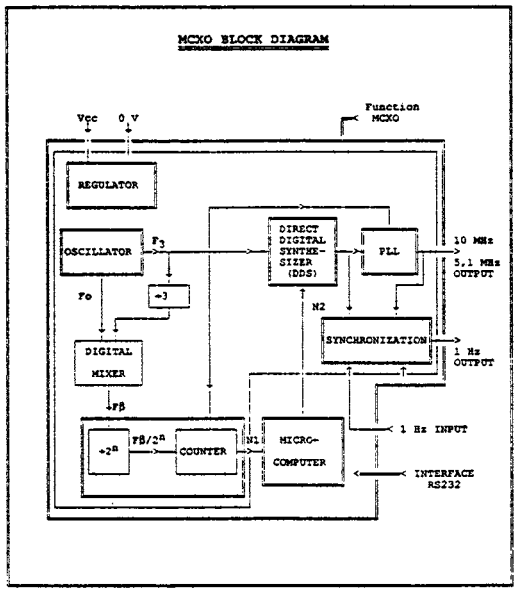


FIGURE 1

The oscillator and compensated system can be divided into eight separate functions.

- Power regulator
- Crystal oscillator
- Frequency Mixer
- Frequency Counter
- Microprocessor
- Direct Digital Synthetizer
- Phase locked loop
- Synchronisation interface

POWER REGULATOR

Power regulator permits input voltages between 3.3 and 7 Volts.

The frequency stability with the input power variation is less than 20 ppb.

CRYSTAL OSCILLATOR

The oscillator is a dual mode system using a T bridge network for the third overtone frequency and a Colpitts circuit for a fundamental mode. Both of theses oscillators are analysed and optimized using CODA Software. [4]

FREQUENCY MIXER

The digital output of the overtone mode divided by three is mixed with the digital output of the fundamental mode by using an exclusive or gate. The beat frequency is obtained by filtering the output of the exclusive or gate by a low pass active filter.

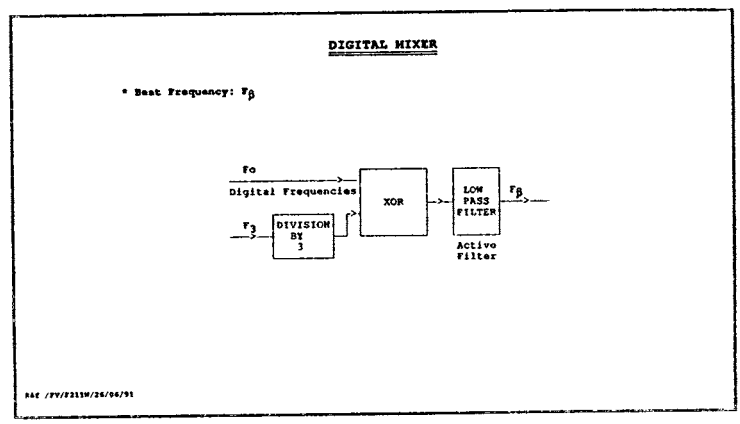


FIGURE 2

FREQUENCY COUNTER

A free running counter is used for this application, it gives a numerical value representing temperature.

The numerical value is proportional to the beat frequency, so this value is proportional to the quartz crystal temperature.

MICROPROCESSOR

The microprocessor stores the frequency counter data and calculates the variation frequency of the third overtone mode with the respect to temperature. It then computes a numerical value which is used to control the direct digital synthetizer.

The microcomputer is used for other applications, Keeping time, 1 pps synchronisation and frequency calibration.

DIRECT DIGITAL SYNTHESIZER AND A PHASE LOCKED LOOP

The phase locked loop with a voltage controlled crystal oscillator gives the MCXO output frequency.

The output frequency is the sum of the third mode of the quartz crystal resonator and the output frequency of the direct digital synthetizer.

SYNCHRONISATION INTERFACE

The network gives one pulse per second, this signal can be synchronised with an input signal.

Figure 3 gives the calibration results for differents polynomials. It can be seen that the peak value for polynomials of the fourth order is smaller than 0.91 Hz at 10 MHz between - 55° and -10 degrees.

Nb of Polys	Temp (°C)	POLYNOMIAL ORDER 1 MQD	1 P V	2 MQD	2 P V	3 MQD	3 P V	4 MQD	4 P V
1	-55/125	253	711	81.6	222.5	9.2	29	1.4	4.8
2	-55/35	112	266	12.6	33	0.9	2.5	0.19	0.8
	35/125	15	48	9	22.6	0.29	1	0.15	0.36
3	-55/5	58	131	4.1	10.4	0.3	0.7	0.21	0.95
	5/65	28.2	65	3.1	7.5	0.11	0.2	0.03	0.09
	65/125	1.5	10.7	2.8	6.5	0.075	0.23	0.06	0.18
4	-55/-10	35.3	78	1.86	4.46	0.26	0.82	0.25	0.91
	-10/35	22.2	49	1.46	3.43	0.05	0.14	0.014	0.04
	35/80	9.76	22.6	1.28	3	0.035	0.092	0.035	0.092
	80/125	2.8	9.8	1.27	2.9	0.066	0.18	0.066	0.18
		5		6		7		8	
1	-55/125	0.23	0.78	0.18	0.84	0.14	0.97	0.14	0.92
2	-55/35	0.18	0.9	0.18	0.9	0.17	0.9	0.17	0.9
	35/125	0.06	0.18	0.06	0.18	0.05	0.18	0.05	0.18
3	-55/5	0.21	0.95	0.21	0.95	0.2	0.94	0.2	0.92
	5/65	0.03	0.09	0.03	0.09	0.03	0.09	0.03	0.09
	65/125	0.06	0.18	0.06	0.18	0.06	0.18	0.057	0.175
4	-55/-10	0.24	0.95	0.24	0.93	0.235	0.9	0.235	0.93
	-10/35	0.014	0.04	0.014	0.04	0.013	0.038	0.013	0.035
	35/80	0.035	0.092	0.035	0.092	0.035	0.087	0.035	0.087
	80/125	0.066	0.18	0.066	0.18	0.061	0.17	0.061	0.17

P V = Peak to Peak value

FIGURE 3

Typical values of the calibration is given by the figure 4.

TEMPERATURE (°C)	CRYSTAL N° : SC61			
	FUNDAMENTAL		THIRD OVERTONE	
	MQD	P V	MQD	P V
-55 à -10	1.6 E-9	4.1 E-9	4.4 E-9	1.3 E-8
-10 à +35	1.8 E-9	3.4 E-9	8.0 E-10	1.8 E-9
+35 à +80	3.4 E-9	8.1 E-9	3.2 E-9	1.4 E-8
+80 à +125	6.6 E-9	1.6 E-8	5.6 E-9	1.3 E-8

FIGURE 4

Figure 5 and 6 give the calibration configuration and flow chart.

At intervals of one degree Celcius, the PC Computer searches for the numerical value which corresponds to the required frequency correction.

At the end of the calibration, the PC optimizes the compensating functions; then, the polynomial coefficients are stored in the microcomputer memory.

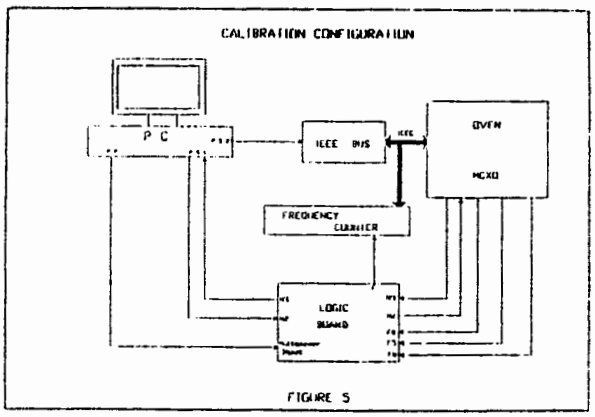


FIGURE 5

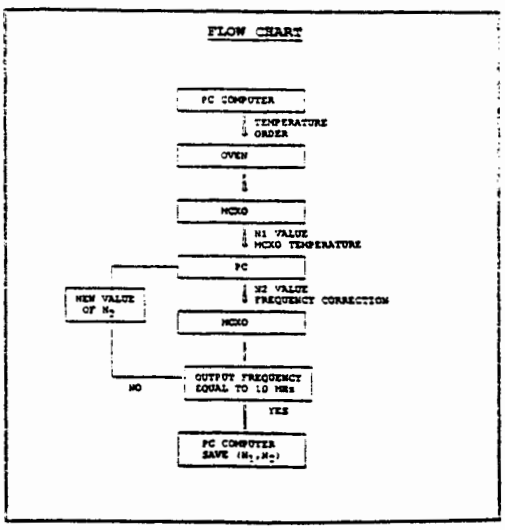


FIGURE 6

Figure 7 gives the MCXO output frequency for a temperature variation of 1 degree per minute between -55° and + 130 degrees Celcius.

The peak to peak frequency variation is less than ± 10 ppb.

The frequency step during the temperature variation is due to the electronic lag.

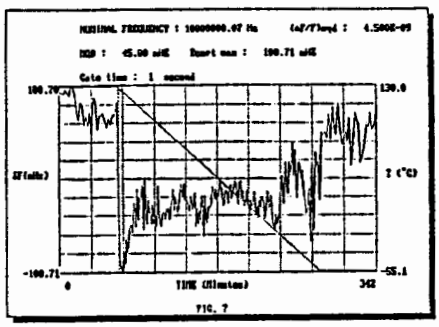


Figure 8 gives the MCXO output frequency at constant ambiant temperature for a period of 6 hours. The peak to peak frequency variation is less than ± 2 ppb.

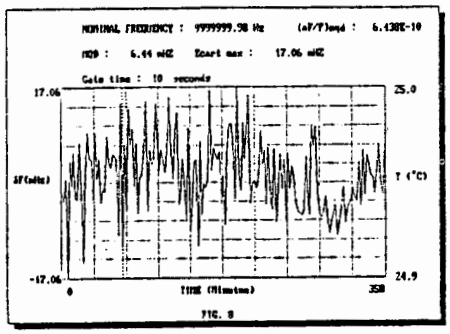


Figure 9 shows the comparaison of the MCXO output with and without compensation.

The frequency stability improvment is of order of 5000.

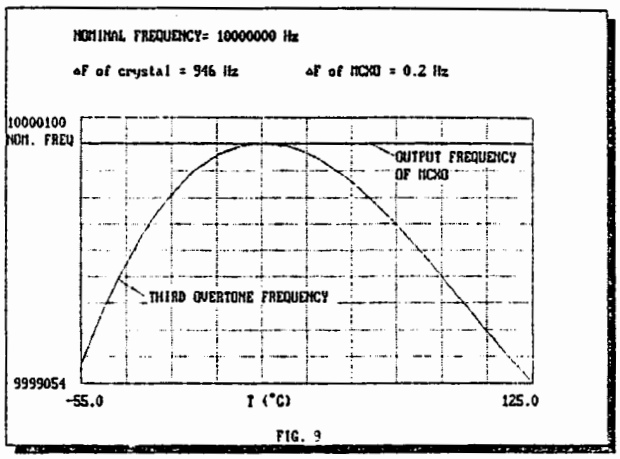


FIG. 9

FUTURE DEVELOPMENTS

The final objective of the MCXO is to obtain a very stable frequency with low power consumption in a small volume.

We propose to develop a mixed ASIC incorporating all of electronic functions.

The first ASIC prototype is scheduled for November 1992 and the pre-production MCXO function for July 1993 in a final holder (40 x 40 x 10 mm).

Additional work is required on the calibration ageing performance and on hysteresis characteristics.

ACKNOWLEDGE

We acknowledge Mike HODGKINS for helpful discussions.

The autors wish to aknowledge the support of the French Army (STTE).

REFERENCES

[1] STANLEY - S. SCHODOWSKI

Resonator Self - Temperature - sensing using a dual - Harmonic - mode crystal oscillator 43rd Symposium on frequency Control 1989.

[2] RAYMOND L.FILLER and JOHN R.VIG

Resonator for the microcomputer compensated crystal oscillator.
43rd Symposium on Frequency Control 1989.

[3] A. BENJAMINSON - S.C. STALLINGS

A microcomputer compensated crystal oscillator using a dual-mode resonator.
43rd Symposium on frequency Control 1989.

[4] The CODA Software program US Army Contract DDA L01-88-C-0804

OSCCAR
OSCILLATOR COMPENSATED WITH A CAPACITANCE ARRAY

W. Daniau, G. Marianneau

Laboratoire de Physique et Métrologie des Oscillateurs du CNRS
associé à l'Université de Franche-Comté-Besançon,
32 av. de l'Observatoire, 25000 Besançon, France

ABSTRACT

The study that we are presenting here is based on the realisation of a capacitance switch array (CSA), usable on a Digitally Temperature Compensated Crystal Oscillator. An original conception method of the CSA has been proposed, and a prototype has been realised, and also an experimental DTCXO using the CSA. We present the conception of the CSA and the results obtained with the DTCXO. Our first prototype has got a frequency stability of $\pm 2.2\ 10^{-7}$ within the range of temperature 0 to 80°C, with a power consumption of 25 mW and a supply voltage of 5V.

INTRODUCTION

The present study is based on the realisation of a capacitance switch array (CSA), usable on a Digitally Temperature Compensated Crystal Oscillator. An original conception method of the CSA has been proposed, and a prototype has been realised, and also an experimental DTCXO using the CSA. The conception of the CSA and the results obtained with the OSCCAR are now discussed.

The original idea of this work was shown in 1983 by Takehiko Uno and Yoshio Shimoda (ref. 1), who had realised such a system with discrete elements.

As it can be seen on Fig. 1, a conventionnal DTCXO, independently of the compensation system, is based as all TCXO's on the modification of the load capacitance of the crystal, by application of a tension beyond a varactor. So there is a digital treatment with in fact an analogical action, it means that it needs a Digital to Analog Converter. The method using a CSA is more natural, we've got a digital treatment and a digital action, and so we need neither a converter nor a varactor.

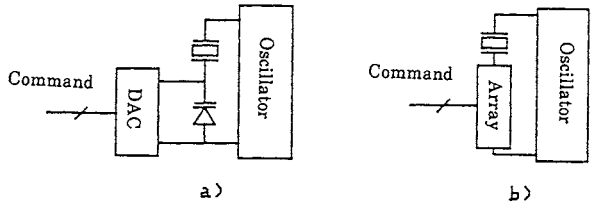


Fig. 1
a) Conventional system b) CSA system

The problem caused by the converter is essentially a problem of consumption; the problem caused by the varactor is that if we want a wide band of capacitance variation, we need a wide band tension, and so a great supply tension (usually 10V).

So, the advantages of our system are essentially the elimination of these two problems.

The fonctionnal scheme of the CSA is shown on Fig. 2, it is composed of capacitors in parallel, each placed in line with a digitally controlled analog switch (CMOS). If N is the command word, the resultant capacitance will be NC_o.

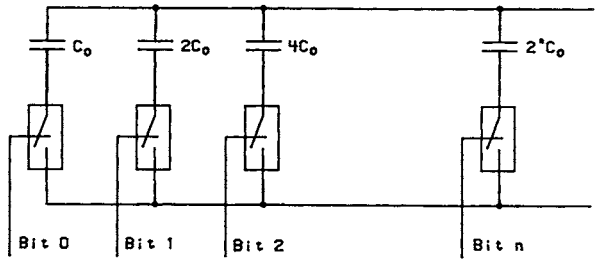


Fig. 2 : *Functional scheme of the CSA*

Theoretically, the ideal curve Capacitance-Command Word will be the one as shown on Fig. 3, which is a staircase with a constant "*step*" of C_o.

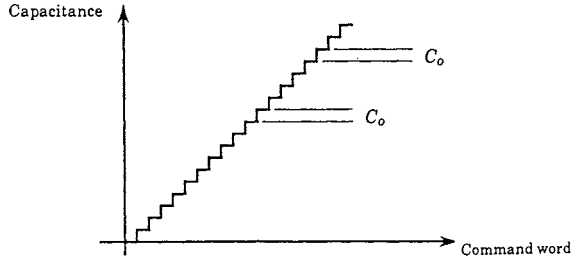


Fig. 3 : *Ideal curve capacitance-command word*

In fact, independently of the problems of parasitic capacitance, the tolerance on the values of the capacitors, will induce important error on the values of the steps.

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

For a N bits system, there are N different ways to realise the capacitance incrementation, so there are N different capacitance steps, for example, these numeric incrementations induce the same capacitance step :

state n	0	100	110	110100011010
state n+1	1	101	111	110100011011

and these too :

state n	1011	1111011	10011
state n+1	1100	1111100	10100

S_0 = bit 0 : 0 to 1
S_1 = bit1 bit 0 : 01 to 10
S_2 = bit2 bit1 bit0 : 011 to 100
etc

in general for the i^{th} step S_i :

$$S_i = C_i - \sum_{k=0}^{i-1} C_k \quad (1)$$

where C_i represent the capacitance associated with the bit i.

To simplify the calculation, we will assume that all the capacitors have the same relative tolerance :

$$\forall i, \quad \frac{\Delta C_i}{C_i} = K \quad (2)$$

from (1) and (2) :

$$\frac{\Delta S_i}{C_o} = K(2^{i+1} - 1)$$

This result means that the tolerance on the step is increased exponentially with the number of bits. A more important fact is that if all the capacitors are discrete elements, with no correlation, the tolerance can be either positive or negative and so the curve won't be monoton.

PROPOSED SOLUTION (ref. 2)

The proposed solution resolves the problem of monotony : all the capacitors will be realised with an addition of a certain number of basic elements, as it is shown in Fig. 4. Thus, all the tolerances will be correlated.

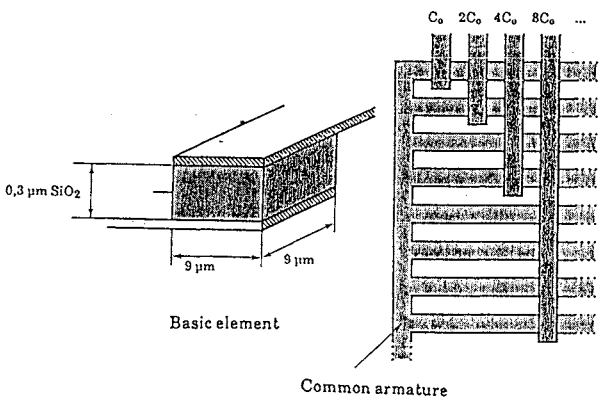


Fig. 4 : *Proposed solution*

The problem of the step values will still be present but not the monotony problem. So the system will just needs to be corrected by suppressing of some basic elements.

A prototype in thin film hybrid technology has been realised. This prototype include the capacitors and their associated CMOS switches as rapported monolithic circuits. It is realised in "lift-off" technic, by sputtering of two aluminium levels and one S_iO_2 level as the dielectric.

RESULTS

To measure the capacitance variations versus command word, the circuit has been placed in an oscillator and the frequency variations has been measured, the results are presented in Fig. 5.

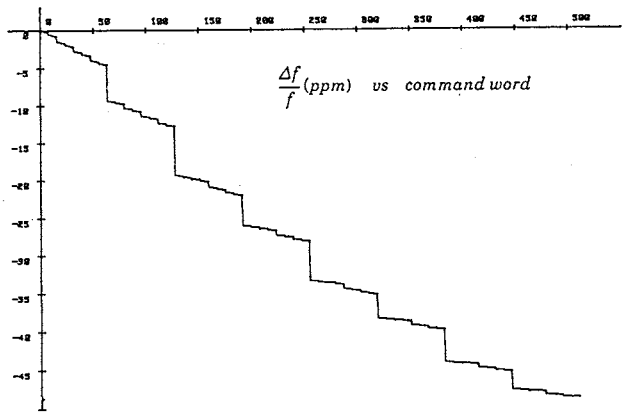


Fig. 5 : *First measurement*

As it has been anticipated, the monotony of the curve has been respected, but the values of certain steps are too high, so after correcting the value of certain capacitors, the curve on figure 6 was obtained.

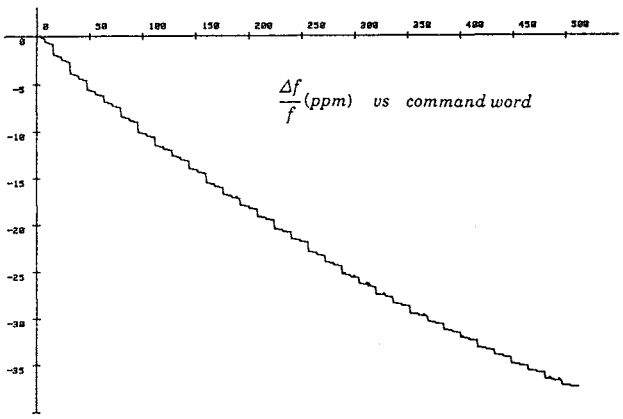


Fig. 6 :
Results after correction of the values of the capacitors

OSCCAR

To valid our CSA, a DTCXO which used the CSA has been built, these are the results :

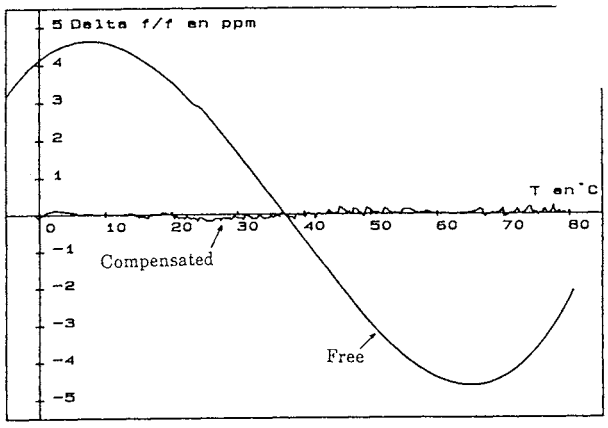


Fig. 7

Stability	:	$\pm 2.2.10^{-7}$ within the temperature range 0 to 80°C
Supply tension	:	5V
Consumption	:	~25mW

CONCLUSION

The CSA system still needs some improvments, to envisage an industrial fabrication, but the tests are very encouraging, the OSCCAR, has got a low consumption, a low supply voltage, and a good stability.

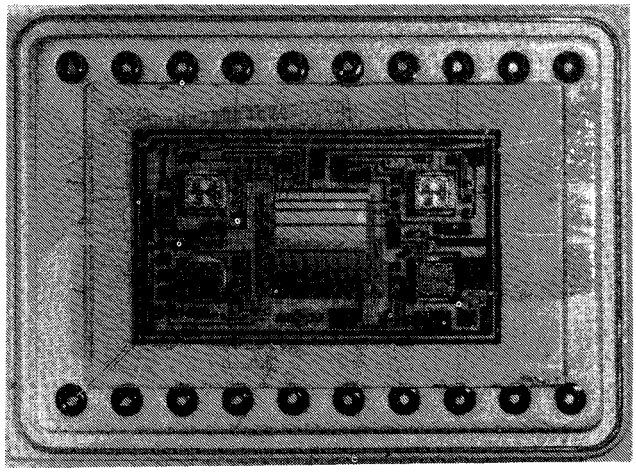


Fig. 8 : *Photo of the circuit*

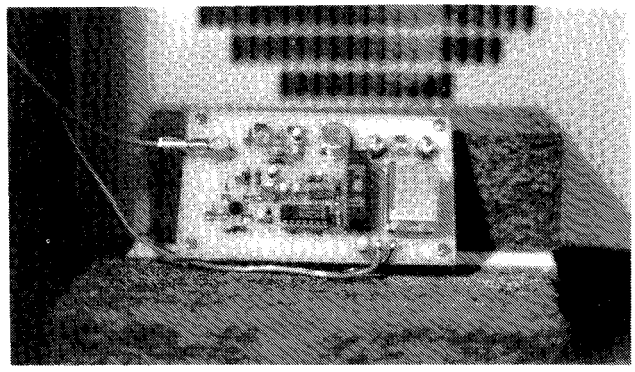


Fig. 9 : *Photo of the OSCCAR*

REFERENCES

1. Takehiko Uno, Yoshio Shimoda, "A new digital TCXO circuit using a capacitor switch array"; Proc. 37th AFCS pp 434-441, 1983

2. William Daniau, "TCXO Numérique à réseau de capacité programmable", Thèse de docteur de l'Université de Franche-Comté, Faculté des Sciences et des Techniques - Université de Besançon, 1991.

DESIGN OF A LOW PHASE NOISE UHF OSCILLATOR USING A PARTLY METALLIZED DIELECTRIC RESONATOR

R.E.ELCHEIKH , M.OLIVIER

Laboratoire de Physique et Métrologie des Oscillateurs du C.N.R.S
associé à l'université de Franche-Comté - Besançon
32, Avenue de l'Observatoire - 25000 Besançon - France

ABSTRACT

In this paper, we present a new type of oscillator, composed of a dielectric resonator (DR) which is part of a microwave integrated circuit (MIC). For this oscillator, the oscillating characteristics can be controlled by several parameters, such as the dimensions, the position of the DR in its environment and the bias voltage of the oscillator.

With this technique, a frequency stability $\sigma_y(\tau)$ of less than 5.10^{-9} for $\tau = 1$ s, an output power of 17 dBm and phase noise as low as -106 dBc at 1 kHz from the carrier in the 800-1000 MHz frequency band are obtained.

1. INTRODUCTION

In the past few years, dielectric resonators have become increasingly important in microwave filters and microwave oscillators. Here, we restrict our concern to those DR's which are partly metallized and suited for use in the UHF band.

Our results show that the dielectric resonator leads to the design of a compact and stable low noise oscillator. The resonator must be mounted in a stable manner to prevent variations in frequency with temperature and time.

Measurements were performed using commercially available $(Zr,Sn)TiO_4$ DR materials, developed by Thomson-CSF Corbeville Laboratories (Ref. 1). They offer an adjustable temperature coefficient of frequency within 1 ppm/°c between -20 and +20 ppm/°c, low microwave losses and high permittivity (ε_r = 36-37) .

In the frequency band below 1 GHz, some size reduction is achieved by using quarter-wave rectangular dielectric resonators with three metallic faces. We report herein the performance of new dielectric resonator oscillator (DRO's) (Refs. 2 and 3) fabricated using partly-metallized DR's.

2. DESIGN CONSIDERATIONS

The design circuit employs the resonator as a parallel feedback element. This configuration has been found to possess numerous advantages over others, such as ease of mechanical and electrical tuning (Refs. 3 and 4).

The basic configuration of the DR feedback network, which has been employed in the oscillator circuit, is shown in Fig. 1a and Fig. 1b ; and it is called a four-port DR configuration. It consists of a DR, placed between two 50 Ω microstrip lines.

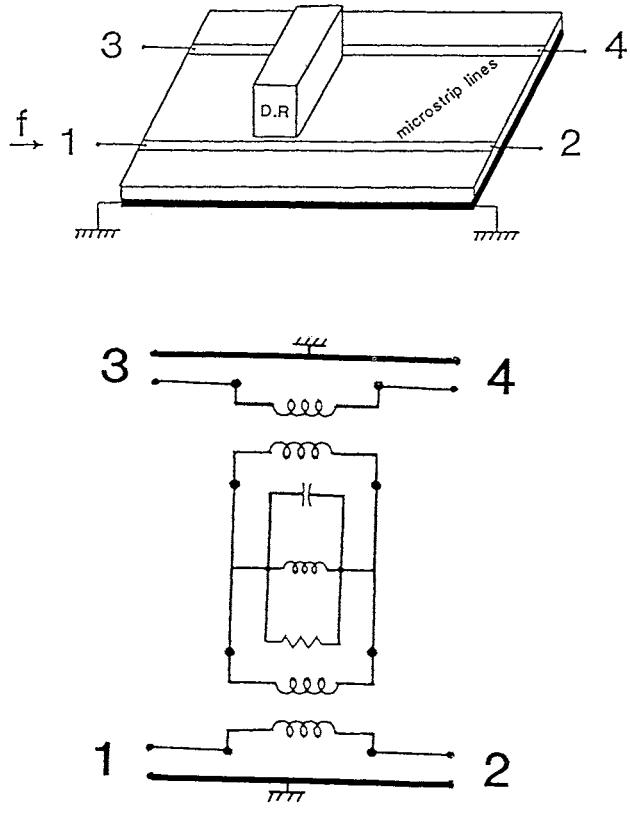
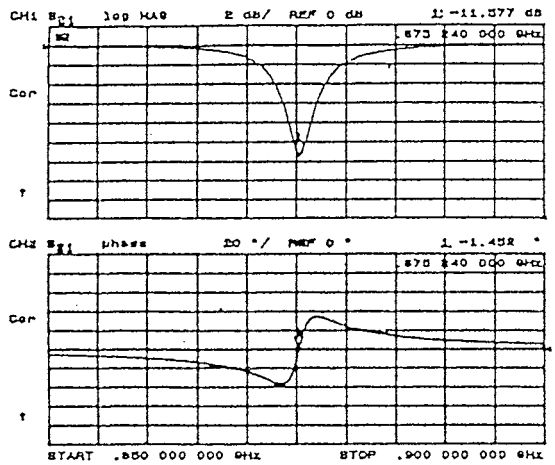


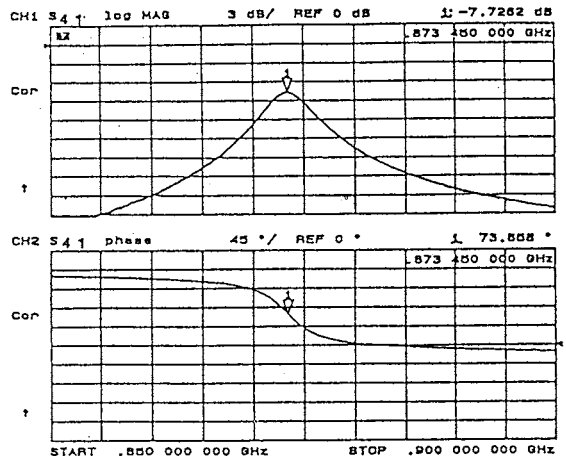
Fig. 1 : *DR configuration*
(a) Four-port DR (b) Equivalent circuit

In this arrangement, there is magnetic coupling between the microstrip lines by way of the DR. This circuit, implemented in microstrip, provides a bandstop response between ports 1 and 2 and a bandpass response between ports 1-3 or 1-4 (Fig. 2).

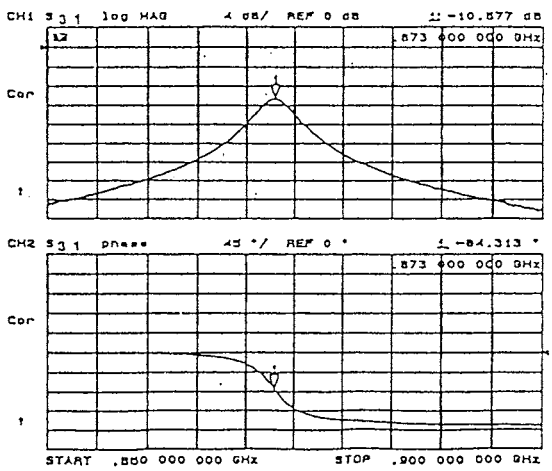
Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).



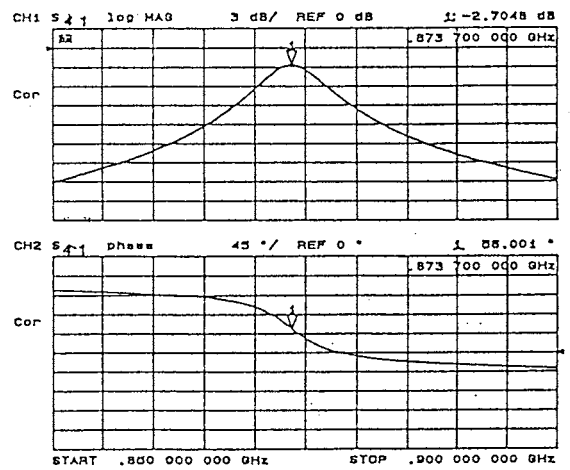
$Z_3 = Z_4 = 50\ \Omega$



$Z_2 = 50\Omega\ ,\ Z_3 = \infty$



$Z_2 = Z_4 = 50\ \Omega$



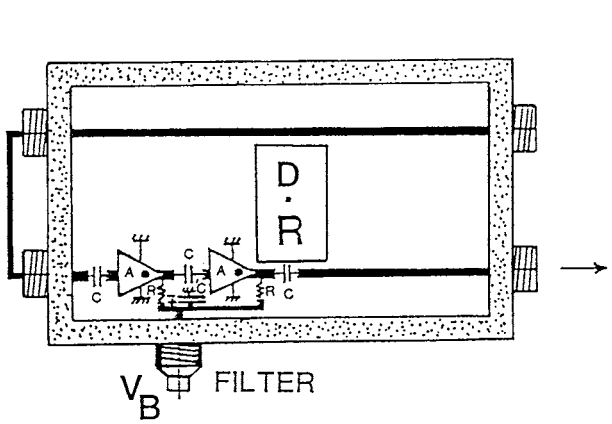
$Z_2 = Z_3 = \infty$

Fig. 2 : *S_{21}, S_{31} and S_{41} - parameters of a four - port DR*

3. UHF DIELECTRIC RESONATOR OSCILLATOR (DRO)

The DRO is fabricated on a dielectric substrate (RT Duroid 6010) of relative dielectric constant $\varepsilon_s = 10 \pm 0.25$. It can be divided into two parts, the first part includes the DR, the support material and the metallic enclosure. The second part includes the oscillator electronics. The amplifier stage is a broad-band, monolithic, cascadable amplifier (21 dB gain, noise figure : 5.5 dB, 50Ω input/output impedance).

The oscillator delivers an output power of 17 dBm in the frequency band 800-1000 MHz . A basic configuration of the DR feedback circuit is shown in Figs. 3a and 3b.



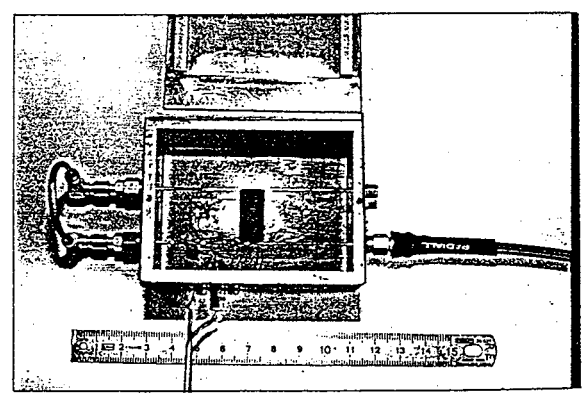


Fig. 3 : *DRO configuration*

(a) Four-port DRO *(b) Photograph of the assembly*

The conditions must be met, corresponding to the following diagram (Fig. 4) :
$|G|\,|S_{ij}| \geq 1$ closed loop gain
$\Phi a + \Phi_{DR} + \Phi_l = 2\Pi N$; $(N = 0,1,2,...)$ phase of the closed loop gain
where G is the loop amlifier gain, S_{ij} the scattering parameter of the DR ; Φa, Φ_{DR} and Φ_l are the respective insertion phases of the amplifier, DR and microstrip line.

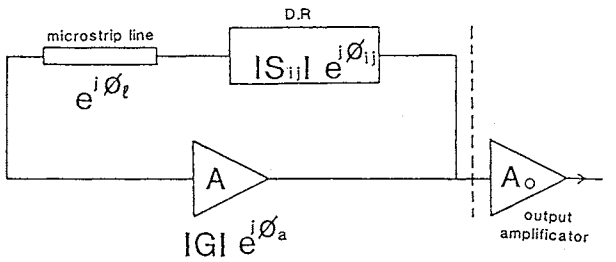


Fig. 4 : *Circuit diagram of the DRO*

The performance of the oscillator depends on parameters such as :
(a) position and distance of the resonator from the microstripline
(b) distance between the resonator and the substrate
(c) polarisation voltage
(d) temperature.

The influence of these parameters on the frequency of the oscillation have been studied and optimized (Ref. 3).

4. TIME-DOMAIN STABILITY AND SPECTRAL PURITY

Two identical DRO's have been fabricated. The Allan variance $\sigma_y(\tau)$ of each oscillator has been measured using a frequency counter whose reference is a very stable 5 MHz oscillator. $\sigma_y(\tau)$ is proportional to $\tau^{-1/2}$ (Fig. 5) for $\tau > 10^{-1}$ s, and this corresponds to a F^{-4} phase spectrum probably due to temperature fluctuations.

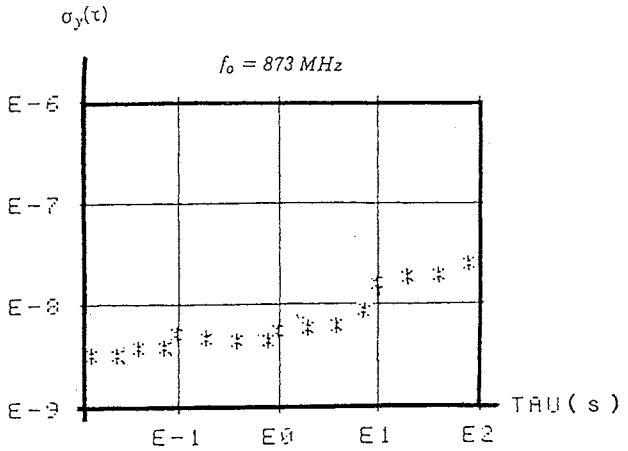


Fig. 5 : *Time domain stability : Allan variance*

The phase-locking method is used to measure the phase noise spectrum of the oscillators. We have utilized a phase noise measurement system developed at the LPMO. It gives, after low noise amplification and filtering, the beat signal between the oscillator under test and a reference oscillator. This reference oscillator can be a synthesizer working at 873 MHz whose frequency is phase-locked to the frequency of the DRO. The spectral purity of the beat signal is limited by the noise floor of the synthesizer, which is of the order of -130 dB for Fourier frequencies above 1 kHz.

We have used two identical DRO's whose frequency difference is of the order of 40 MHz, and a divided frequency synthesizer phase-locked to the beat frequency of the oscillators under test (Fig. 6).

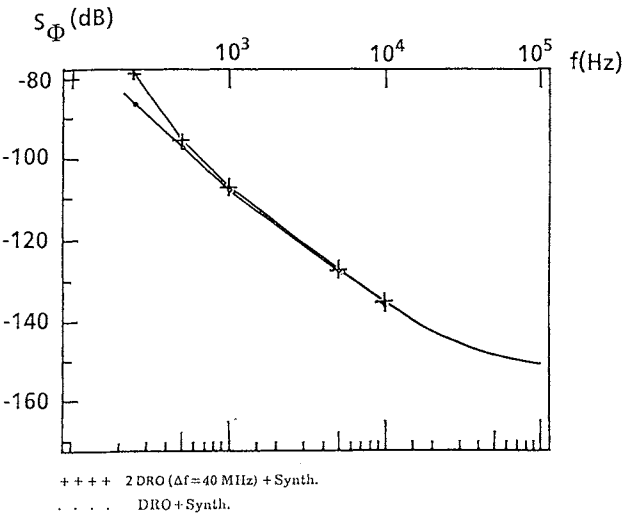


Fig. 6 : *Phase noise spectrum of a DR oscillator*

The phase noise spectrum of the DR oscillator shows a F^{-3} behaviour for Fourier frequencies higher than 250 Hz, and a F^{-4} spectrum in the low frequency range. This spectrum can be as low as -106 dB for F = 1 kHz.

CONCLUSION

The results of the experiments show the feasibility of a compact, light weight and stable dielectric oscillator with a high spectral purity at frequencies below 1 GHz.

REFERENCES

1. J.C. Mage, B. Marcilhac 1987, Amélioration de la stabilité en température des matériaux pour résonateurs diélectriques, *Actes des 5e Journées Nationales Microondes*, Nice, pp.41-43.

2. R.E. Elcheikh, M. Valentin 1989, Temperature dependence of an UHF oscillator using a dielectric resonator, *Proc. 3rd E.F.T.F.*, Besançon.

3. R.E. Elcheikh 1991, Contribution à l'étude de résonateurs diélectriques: Application des résonateurs partiellement métallisés à la réalisation d'oscillateurs UHF, *Thèse de Doctorat* - Université de Franche-Comté Besançon.

4. M. Valentin, R.E. Elcheikh 1988, Microwave oscillator using partly metallized dielectrique resonator, *Proc. 2nd E.F.T.F.*, Neuchâtel.

MINIATURIZATION AND STABILITY : A NEW GENERATION OF OVENISED OSCILLATORS

JP AUBRY, G WAGNER and V CANDELIER
CEPE Compagnie d'Electronique et Piézoélectricité

44 Avenue de la GLACIERE 95100 ARGENTEUIL FRANCE

Miniaturisation is now one of the main target oscillator manufacturers have to meet in new programs.

For example, analog radiotelephone systems ask for "standard" OCXOs but the GSM (digital radiotelephone) program requires specific ovenised oscillators to be installed in the ground stations, exhibiting high performances, regarding mostly "phase jitter".

CEPE has a hudge experience in design and production of clock, temperature compensated and ovenised oscillators, for commercial and military markets, in a wide range of stability.

Our current production of professionnal ovenised oscillators is mainly based on two families, PMT and PMT P which exhibits size of 51*41*25 and 67*60*40 mm^3. In order to propose smaller sizes, CEPE has developped new oscillators.

What is miniaturization and stability ?
This question can be answered on two ways :

* which stability can be achieved in a miniature package ?

* how small can be an oscillator meeting a given stability ?

We will explore the second way in this paper.

In that way, miniaturization and stability mean that we have to achieve simultaneously :

- low ageing, low phase noise
- high temperature stability, fast warm up,
- small size
- low cost

Technologies available in CEPE, such as doubly rotated quartz crystal cut, SMD, monolithic thermal structure, PID thermal loop..., have to be used to meet such requirements.

We present in this paper the design procedure and the electrical performances obtained with very small ovenised oscillators, for which sizes have been reduced to 40*30*19 mm^3, 40*30*16 mm^3 and 36*26*19 mm^3.

The resonators are SC cut 3rd overtone in the range of 10 to 20 MHz in HC 37 can for the PMT and 40*30*19 oscillators, and either AT or SC cut fundamental for 40*30*16 and 36*26*19 mm^3.
Depending of the resonator, ageing performances around 1.5 10^{-8} to 3 10^{-9} /month, along with frequency-temperature stability from 10^{-8} to 5 10^{-9} within -10,+70 °C can be achieved.

Warm up time as low as 1 10^{-6} / 1' or 1 10^{-7} / 2' is also reached for the smallest units (using fundamental mode resonators), and as low as 5 10^{-9} / 5' for the larger units using 3rd overtone resonators.

I- DESIGN PROCEDURE

I-a Block diagram

Figure 1 gives the general structure of an ovenised oscillator. There are 6 main parts to be considered :

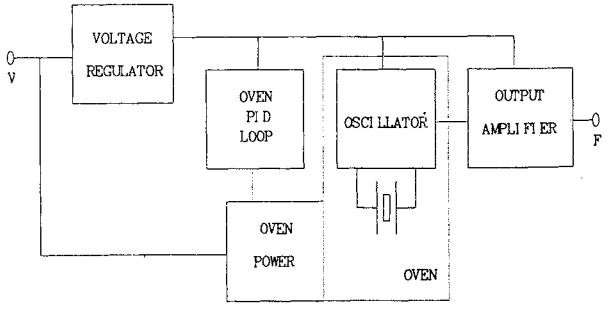


Figure 1 : Oscillator block diagram

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

* the oscillator itself, including the crystal resonator, which is the real "signal generator",

* the output power stage acting as an isolator stage between the oscillator and the load, and giving the final shape to the signal,

* the oven control network whose role is to drive the power heating elements,

* the power elements, which are dissipating heat power to the metallic oven,

* the oven,

* the voltage regulator.

All of these elements have to be optimized individually and the overall structure have also to be worked on because of the "inter-element" mutual reaction of their performances.

I-b Stability criteria

The following table gives the sensitivity criteria between the oscillator structure and the key points of frequency stability :

	Xtal	Osc	OA	OL	OV	VR
F vs temperature	***	**		**	**	***
F vs load		**	***			
F vs supply voltage		**	*	**		***
F vs time (ageing	***	***		*		**
Warm up	***	*		***	***	
Power consumption	**			***	*	
Phase noise	***	***	**	**	**	**
Size	***		*		**	*

where OA stands for Output Amplifier, OL for oven loop, OV for oven, VR for voltage regulator, Osc for oscillator, Xtal for resonator.

The number of stars is an indication of the degree of criticity of the sub-fonction (colomn) versus the sub-electrical specification (line).

One can see in this table the inter dependance of the various design parameters for the overall specification.

From this table, the choise of the crystal resonator acts on the frequency (F) vs temperature (T) behavior, on the warm up characteristics, on the ageing, on the phase noise and indirectly (through its size) on the power consumption.

I-c Resonator selection

The basic design approach of a resonator relates to the crystallographic cut (AT, SC ...) and to the overtone.

When a crystal resonator is placed in an oven, one tries to give to the resonator a F vs T behavior showing a "turn over" temperature around the oven chamber temperature.

In an oven adjusted to work around 80 °C, it is well known that the "radius of curvature" of the F vs T curve of an SC cut is about 10 times flatter that the one of an AT cut at the same temperature.

Then, frequency vs temperature will show, with the same oven, a better stability with an SC cut than with an AT cut, or the SC cut will ask for a simplified oven to get the same f(T) behavior.

Moreover, because of the so called "thermal dynamic behavior" of a resonator, which gives the relationship between frequency and the temperature variation with time (df/f (t,T) ~ ã * dT/dt), the warm up characteristic will be significantly better with an SC cut rather than with an AT cut.

The equivalent ,network of resonators is also a key for the righ selection.

It is well-known that the piezoelectric constant "k" is related to the electrical behavior (C_1/C_0) through :

$$k^2 = 1/8 \; n^2 \pi^2 \; C_1/C_0,$$

where "n" is the overtone mode.

The motional inductance of a resonator can be determined, as a first approach, from :

$$L_1 = 1 / C_1 \; \omega^2$$

$$L_1 = K . (n^3 / 32 \, k^2 \varepsilon) . 1/SF^3,$$

where K is the frequency constant, ε is the dielectric constant, S the electrode area and F the resonator frequency ($\omega = 2 \pi F$).

One can see that it is possible to get quite the same value of motional inductance in SC cut on the third overtone than in AT cut on the fifth overtone. In others words, we can get the same equivalent network in a smaller package using a SC cut rather than an AT cut.

Such a network, mainly a high inductance value is a key to get low ageing and low sensitivity (ageing and thermal behavior) to electronic frequency adjustment.

For all of these three reasons, SC cut has been selected for most of the miniature ovenised oscillators we are dealing with in this paper.

Ageing is strongly related to the resonator either through the technology (clean process, know how...) or through the choise of the overtone mode.

Depending of the exact ageing specification, fundamental mode or 3rd overtone can be used. Because of the size of the resonator, fundamental mode have to be choosen for the smallest oscillator we are presenting.

I-d Electronic design

The oscillator must be very accurately designed. The main criteria we have to take into account are :

loaded Q factor related to phase noise close to the carrier
phase sensitivity vs temperature and time
supply voltage sensitivity
drive level sensitivity
harmonic generation

Non linear analysis was used to optimise the oscillator network.

The output amplifier have to be designed along with specific criteria such as :

good protection against load variation
low thermal sensitivity
high selectivity
low noise

Classical networks such as common emitter or common base single transistor stages can be used.

Best compromise is to be adjusted between the kind, number of stages, surface on the PCB and cost. This optimisation found different solutions in our miniature ovenised oscillators.

The last critical part in the electronic design deals with the voltage regulator. Usually, the oscillator must exhibit performances in various DC external voltage supply variations. Then the oscillator must be protected against these external perturbation. The "filter" is the voltage regulator used to supply the internal oscillator. Natural sensitivity of the "internal" oscillator is in the range of 10^{-9} per volts. To get a frequency variation better than 10^{-9} per volt, the voltage regulator must exhibit a DC output stability better than 10 mV.

I-e Oven design

The oven construction needs very sophisticated thermal analysis. First of all, the equivalent thermal network parameters have to be determined as an input to the determination of the PID loop, and the thermal behavior have to be optimised regarding thermal properties such as temperature distribution in the structure, thermal leak ways, temperature drift between sensor and resonator..., within extreme conditions of external temperature and heat dissipated on the oven.

Finite element thermal analysis was used to determine basic rules regarding the design of small size / low power ovens.

The network used to drive the power heating transistors from the thermal response of the sensor (generally a thermistor), must control either gain and phase of the correction signal. Various networks can be used to achieve this fonction. The best compromise we found, regarding accuracy, size and cost, is a PID (Proportionnal, Integrator, Derivator) thermal loop.

II- EXPERIMENTAL RESULTS

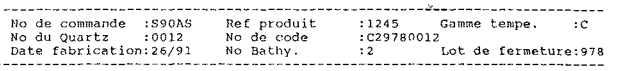
Rather than giving a complete list of all the performances we can get with all of the main "miniature oscillators" we are dealing with (the list would include 4 packages x 2 cuts x 2 overtones...), we will show the typical results we can get in the main areas of stability.

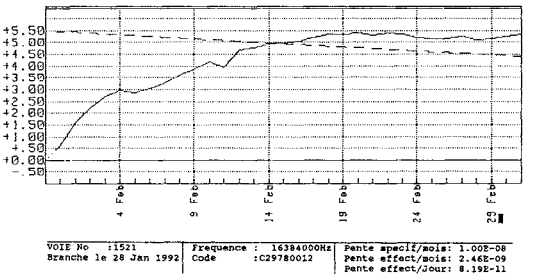
II-a Ageing

The following table gives the typical ageing specifications we can meet with these oscillators :

Osc type	Resonator	Monthly ageing	Yearly ageing
PMT	SC 3rd	$1\ 10^{-8}$	$5\ 10^{-8}$
40*30*19	SC 3rd	$8\ 10^{-9}$	$4\ 10^{-8}$
40*30*16	SC F	$3\ 10^{-8}$	$1.5\ 10^{-7}$
36*26*19	SC F	$\pm\ 1.3\ 10^{-8}$	$1.5\ 10^{-7}$
36*26*19	SC F	$\pm\ 3\ 10^{-8}$	$\pm\ 1.5\ 10^{-7}$

Figure 2-a gives, for example the frequency vs time behavior of a PMT 16.384 MHz, and figure 2-b gives the ageing curve of a 40*30*19.
These results show that ageing characteristic as low as 4 10^{-9} per month can be acheived with a SC cut 3rd overtone in HC 37 can.





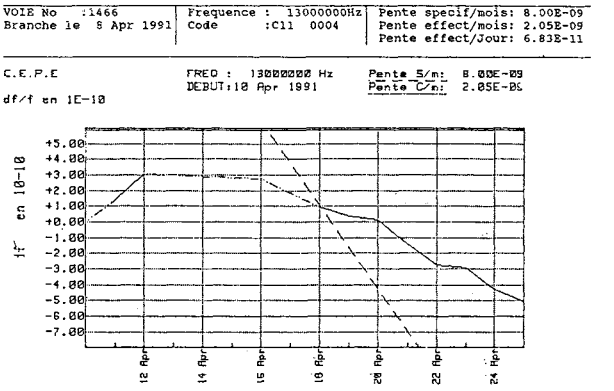
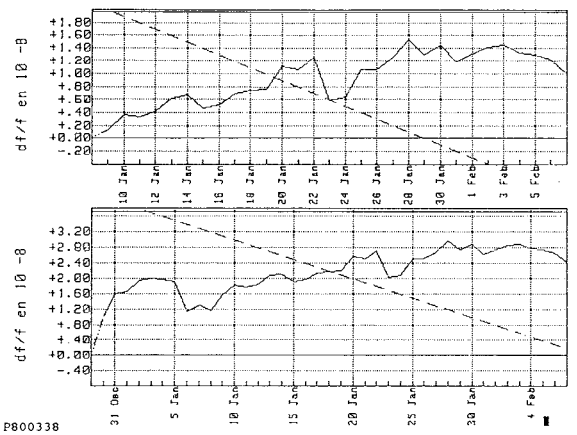


Figure 2-c gives the frequency vs time behavior of a 40*30*16 oscillator, at 10 MHz. Frequency drift is better than 7 10^{-9} per month.



II-b Frequency vs temperature stability

The following table gives the typical f(T) behavior. We can get with these oscillators :

Osc type	Resonator	Temp range	F/F ( )
PMT	SC 3rd	- 30, 70	1 10^{-8}
40*30*19	SC 3rd	- 10, 70	1 10^{-8}
40*30*16	SC F	- 10, 70	1 10^{-7}
36*26*19	SC F	0, 70	1 10^{-7}
36*26*19	AT F		

Figure 3-a gives the frequency vs temperature of the oscillator PMT (Fo = 16.384 MHz). The use of SC cut and of a sophisticated oven gives results (± 1 10^{-9}) by far better than the specification.

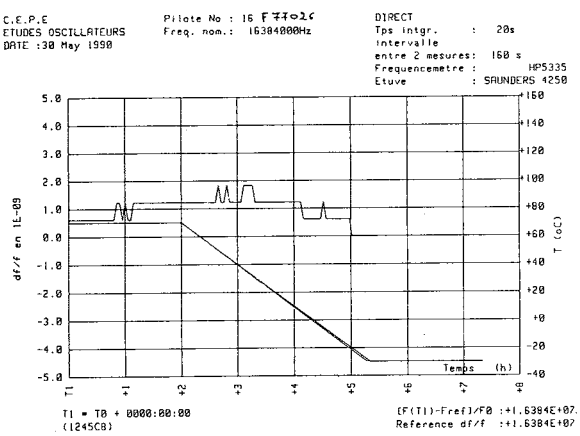


Figure 3-b shows that the f(T) behavior of oscillator 40*30*19 is better than 1 10^{-8} peak to peak.

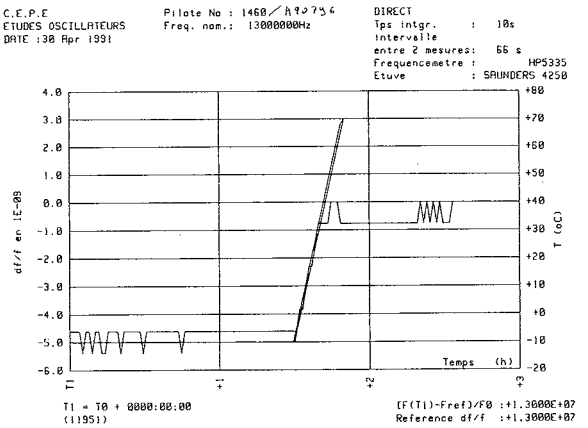
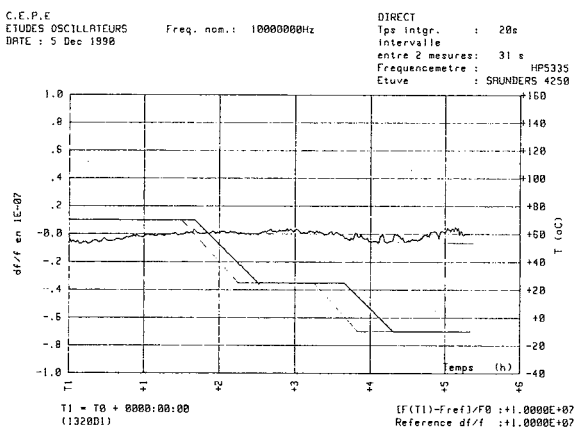


Figure 3-c shows that the f(T) behavior of the oscillator 40*30*16 (using a SC fundamental resonator) is better than 3 10^{-8}.



II-c Phase noise

The following table gives the typical phase noise L(f) results we can get with these oscillators

Osc type	Resonator	L(f) @ 100Hz	L(f) @ 1KHz	L(f) floor
PMT	SC 3rd	- 140 dBc/Hz	- 150 dBc/Hz	- 155 dBc/Hz
40*30*19	SC 3rd	- 130 dBc/Hz	- 140 dBc/Hz	- 145 dBc/Hz
40*30*16	SC F	- 120 dBc/Hz	- 135 dBc/Hz	- 140 dBc/Hz
36*26*19	SC F	- 125 dBc/Hz	- 140 dBc/Hz	- 140 dBc/Hz
36*26*19	AT F	- 100 dbc/Hz	- 130 dBc/Hz	- 140 dBc/Hz

Figure 4a gives the phase noise spectrum of the PMT oscillator.

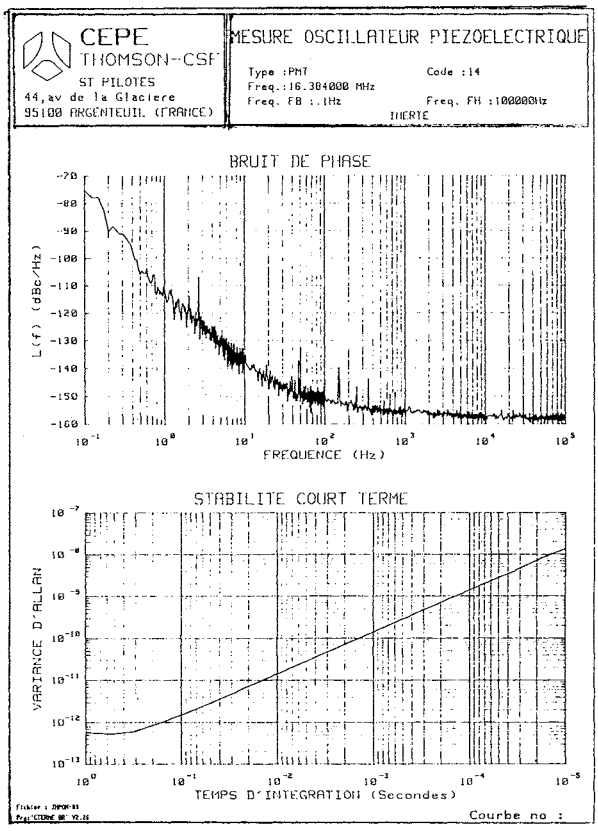


Figure 4-b gives the phase noise spectrum of the oscillator 40*30*19. Noise figure is better than -150 dBc/Hz for frequency greater than 50 Hz offset from the carrier.

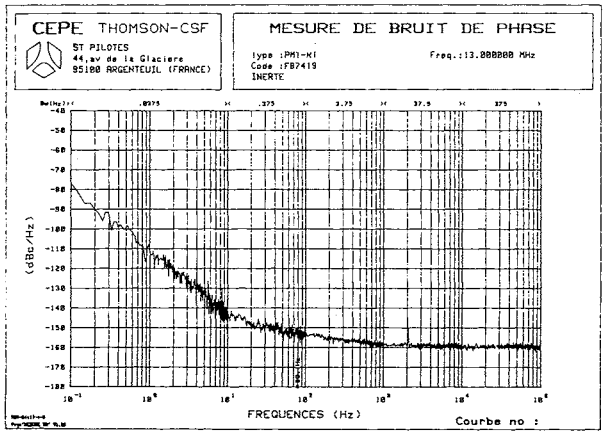
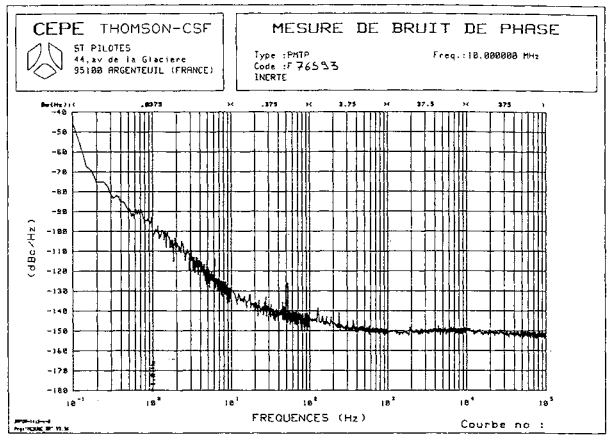


Figure 4-c gives the one of oscillator 40*30*16. Noise floor better than -150 dBC/Hz is reached.
One can see that these results are by far better than their specification.

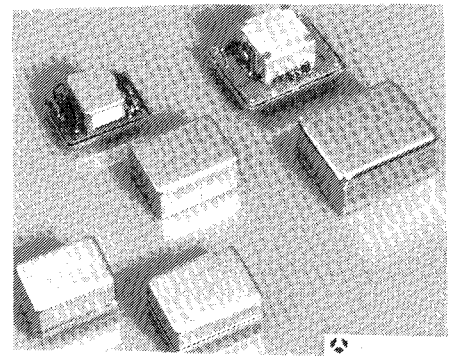


II-d Warm-up time

The following table gives the warm-up characteristics. We can get with these oscillators :

Osc type	Qtz						
PMT	SC 3rd			15'/-30°C	$5\ 10^{-8}$	60'/-30°C	$5\ 10^{-9}$
40*30*19	SC 3rd						
40*30*16	SC F	2'/25°C	$1\ 10^{-8}$	5'/25°C	$2.6\ 10^{-7}$	5'/-40°C	$2.6\ 10^{-7}$
36*26*19	SC F						
36*26*19	AT F						

The experimental warm-up behavior of the oscillator 40*30*16, gives results ($4.3\ 10^{-8}$ at 2 min and $1.3\ 10^{-8}$ at 5 min) by far better than the specification.



CONCLUSIONS

Miniaturization and stability are "relative" concepts and must be accurately defined.

The oscillators designs we have shown offer "intrinsic" high performances regarding either ageing, frequency stability or phase noise, in the way that these performances can be compared to the one we can get from actual standard oscillators, designed in larger package.

One of the key in getting ovenised oscillators in small package is the resonator definition. The use of SC cut allow good performances in ageing, thermal stability and phase noise, in smaller size than an equivallent AT cut.

Miniaturised and stable oscillator specifications can be met by using state of the art of oscillators design and manufacture concepts.

The experimental results show that tight requirements are obtained with very good oscillators yields.

Session 5B

Time and Frequency: Dissemination

Chairman: D. Kirchner

A SYSTEM DESIGN FOR LASER TIME SYNCHRONIZATION VIA GEODETIC SATELLITE AJISAI

H. KUNIMORI[1], F. TAKAHASHI[1], and A. YAMAMOTO[2]

[1] Communications Research Laboratory, [2] Maritime Safety Academy, JAPAN

ABSTRACT

The Communications Research Laboratory (CRL) has been developing a laser time synchronization system based on the CRL satellite laser ranging system developed in 1990. The target satellite is the Japanese geodetic satellite AJISAI, which is equipped with corner cubes and mirrors. Feasibility of the experiment is studied including the time transfer equation, the observation system, the link budget and distribution of reflected beam. Simulation study indicates that two stations thousands of km from each other can be linked by a laser beam with a signal strength of more than 10 photons. The images of the laser beam from the AJISAI mirrors are uniformly distributed on the ground and two stations have many chances to link in a single AJISAI pass.

Keywords: Laser, SLR, time synchronization, two-way time transfer.

1. INTRODUCTION

Time transfer using optical pulses has recently been suggested as potential for giving precision of sub-nano or picosecond time synchronization to overcome the maximum bandwidth of radio waves. The requirement of such an extremely high precision exists in applications in space geodesy, relativity physics and in astronomy.

Laser time synchronization is an application of satellite laser ranging (SLR) which has made progress in the field of precise geodetic measurement. SLR measures the time for laser pulses from the telescope to get to a satellite and back. The satellite has corner cube retro-reflectors so the light is reflected exactly back to the transmitting station. Some SLR stations measure the range to as precise as 30 ps (Refs. 1,2). Since the SLR system is essentially a highly precise epoch recorder transmitting and receiving of optical pulses, a combination of two or more SLR systems has potential as a highly precise time comparison and synchronization system.

There are two configurations to be considered for laser time synchronization using a satellite (See Fig.1). The first is called the one-way up-link configuration. Both SLR stations transmit a laser pulse to the time interval unit (TIU) on the satellite and the TIU measures the interval between their arrival times. Such an optical time comparison has been performed in Europe in the LASSO experiment (Refs. 3,4). In the LASSO experiment of European phase, laser echoes were successfully received at Grasse, France and at other European laser ranging sites. However, time comparison results using on-board TIU have not yet been reported.

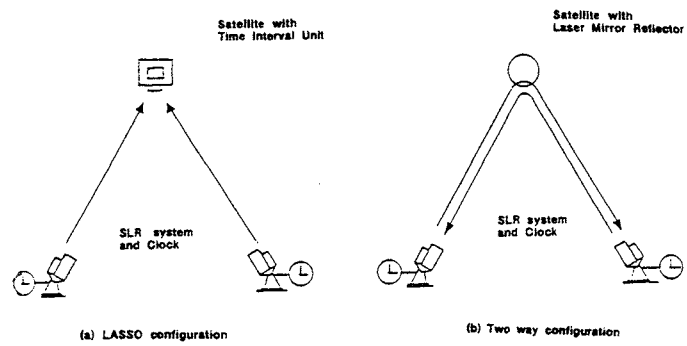


Fig.1 Configuration of Laser time synchronization

The second is the two-way configuration. Each station transmits a laser pulse to the other station via a satellite which has an optical transponder or simply mirror reflectors. The laser pulses are transferred via transponder or mirror from one station to another during the satellite is up on the horizon at both stations. The satellite may also have corner cube reflectors for conventional SLR, and they are used to determine the distance from each station to the satellite to calibrate the time synchronization results.

This paper examinines the two-way configuration using the Japanese SLR satellite AJISAI. The system performance is based on the CRL SLR system (Refs. 5,6), including additions to be made in the future for laser time synchronization.

2. TARGET SATELLITE AJISAI

AJISAI is the Japanese geodetic satellite launched in 1986 by the National Space Development Agency. The major specifications of AJISAI are listed in Table 1 (Ref. 7). It is completly passive satellite and a hollow sphere 2.15 m in diameter and weighs 685 kg. The surface is covered with mirrors as well as corner cube reflectors. The orbit is a circular with an inclination of 50 degrees and an altitude of 1500 km. The mirrors were originally designed for observing flashing solar light reflected by separate mirrors while the satellite is spinning at about 40 rpm. Each mirror is a piece of a surface with a radius of curvature of 8.5 m. Three elements from the same latitude of the 8.5-m sphere at longitudes of 0, 120 and 240 degrees are put on the surface of a 1-m radius satellite retaining the phase angle of the original sphere (See Fig.2). The latitudes from which every set of three is taken are uniformly selected from the original surface so that a distant observer can observe the light reflected from the front surface flashing three times per rotation.

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

Table 1 : Major specifications of AJISAI

Launch Date	:	August, 1986
Configuration	:	Polyhedron inscribed in a sphere of diameter 2.15 m
Weight	:	685kg
Corner cubes	:	1436 pieces
Effective area	:	91.2 cm^2
Number of mirrors	:	318 pieces
Curvature of mirrors	:	8.4-9 m
Reflective efficiency	:	0.85
Brightness	:	1.5-3.5 star mag.
Duration of a flash	:	approx. 5 msec
Rate of flashing	:	approx. 2 Hz
Spin rate	:	40 rpm
Launch orbit	:	Altitude 1500 km Inclination 50 deg Eccentricity 0.001

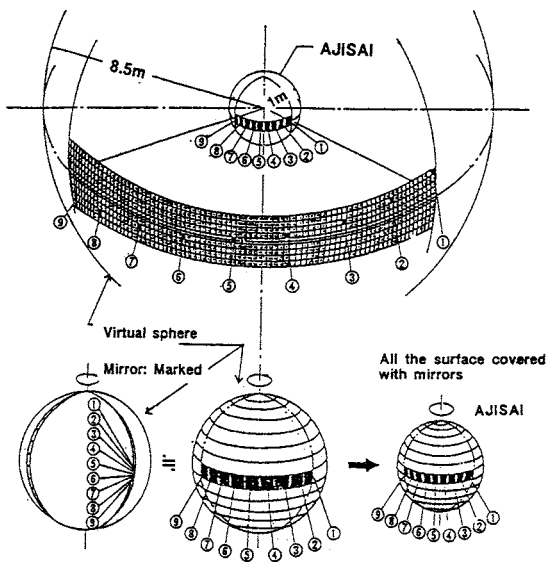


Fig.2 Arrangement of mirrors on AJISAI
Parts of an 8.5m radius virtual sphere (mirror) is placed on the surface of the 1-m radius AJISAI maintaining their latitude angles.

3. TWO WAY TIME TRANSFER

The equation for two-way time transfer via laser pulse is in principle the same as those being developed in radio frequency band (Ref. 8). The time difference d_{12} (: positive when clock 1 is ahead of clock 2) between the clocks of stations 1 and 2 is given by:

$$d_{12} = (t_{21}-t_{12}+t_1-t_2)/2 + (R_{12}-R_{21})/2c \qquad (1),$$

where : $(i,j=1, 2)$,

t_i : Epoch of pulse departure at station i measured by station clock i,

t_{ij}: (i .NE.j) Epoch of station i 's pulse reflected by a satellite mirror arriving at station j measured by station clock j,

R_{ij}: (i .NE.j) One-way distance of a laser pulse to travel from station i to station j via the satellite, and

c: speed of light.

The last term $(R_{12}-R_{21})/2c$ is expected to be nearly zero because of the good symmetry of travelling paths, but not negligible due to the effect on motions of both satellite and stations. It is related to the difference between the satellite position at the times the two laser pulses arrive. It approaches zero if the two pulses arrive at nearly the same time on the satellite. This is done by controlling the firing timing at both stations. Timing control within the precision of 1-ms is necessary to obtain a station to station link by a mirror of AJISAI with 20cm x 20cm size spinning at 40 rpm. If the laser fire timing is controllable to 1μs, $(R_{12}-R_{21})/2c$ is down to a few ps on the assumption that both the predicted position of the satellite and the clock synchronization are known with the accuracy of less than 1 μs (300 m in distance) before the experiment. The μs firing control can be performed by a fully active mode-locked laser operation .

4. OBSERVATION SYSTEM

4.1 Laser ranging system

We use an active-passive mode-locked Nd:YAG laser to generate 532 nm wavelength optical pulses 100 ps wide. The passive mode-locking is performed by saturable dye. The energy is 100 mJ per pulse and the nominal repetition rate is 10 pulses per second. The repetition rate can be controlled from 7 to 14 pps by real-time software. If the the passive mode-locking by a saturable dye is replaced with an active component, the synchronization of firing timing can be controlled on nanosecond level (Ref.9) .

The receiving telescope aperture has diamter of 1.5 m. A micro-channel plate photo multiplier (MCP-PMT) is used for detector. It has a 300-ps rise time, 8% quantum efficiency and the transit time jitter of less than 30 ps with a constant fraction discriminator. The MCP detector can be also gated temporarily by the prediction of photons arrival in 20-ns time steps.

The timing system consists of a high performance quartz oscillator, a GPS receiver, and a time interval unit (TIU). Either a GPS timing receiver or cesium clock of UTC (CRL) can keep the reference to UTC. Timing epoch is measured for up to 4 stop events with a resolution of 40-ps.

4.2 Calibration system

In the actual experiment, we must consider that the reference signal of the atomic clock is transmitted to the SLR system via cables and electronics. Furthermore, the laser transmitting and receiving point is not at the telescope reference point.

The cables and electronics delay between the atomic clock and the TIU is monitored by an independent measurement system. Figure 3 is a block diagram of the monitoring system used in the experiment. The combined 5-MHz+1-pps signal from the cesium clock is transmitted via a 500-m optical fiber. Output signal (50 MHz) of the quartz oscillator is phase-locked by the 5-MHz and used as the reference frequency in TIU. The 1 pps signal is used for the epoch

reference of UTC. The 5-MHz and 1-pps signals are combined again at SLR site and are sent back to the clock room. Then delay and phase difference with respect to the original signals are continuously monitored at the clock room. It has been installed to get the characteristics of the long-term stability of the system.

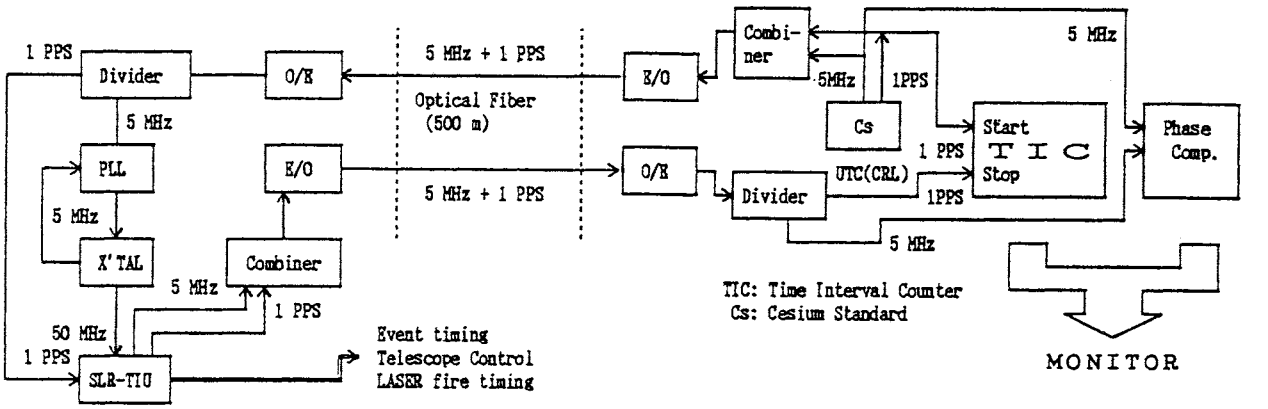


Fig.3 Reference signal monitoring system for laser time synchronization.

The epoch latch point at each laser pulse start and stop must be treated as if these events occurred at the telescope reference point (intersection of two rotation axes). The calibration target is put on the telescope moving axis, as it takes partial reflection of transmitting beam and its physical distance from reference point must be measured precisely.

The time difference between clocks including calibration term (d'_{12}) is given by:

$$d'_{12} = (t_{21}-t_{12}+t_1-t_2)/2 + (R_{12}-R_{21})/2c + \{ (\tau_{1,a}-\tau_{1,b})-(\tau_{2,a}-\tau_{2,b})+(\tau_{2,x}+\tau_{2,y}) -(t_{1,x}+t_{1,y}) \}/2 \quad (2),$$

where : (i=1,2), in addition to the parameter defined in Eq.1,

$\tau_{i,a}$: Optical path delay from firing (laser input) point to telescope reference point at station i,

$\tau_{i,b}$: Optical path delay from telescope reference point to receiver point at station i,

$\tau_{i,x}$: Electronic delay from firing point to start-epoch latch gate at station i, and

$\tau_{i,y}$: Electronic delay from receiver point to stop-epoch latch gate at station i.

In order to calibrate the time difference between clock 1 and 2, each value of parameter in { } in the Eq.2 must be evaluated. The optical delay can be measured by a distance meter or by scales, and it will be stable in time unless optical design is changed. To cancel out the electronic delay, the common portable receiver can be collocated at two stations. This measurement would still have time-varient errors coming from the dependence on signal strength and temperature.

5. LINK BUDGET

In the laser time synchronization experiment, the number of photons (N_p) detected at the remote ground station are given by :

$$N_p=(E\lambda/hc)\cdot(r_1r_2r_3r_4^2) \cdot(16A_sA_r/\pi^2R_1^2R_2^2q_t^2q_s^2) \quad (3),$$

where

E : Energy in a pulse,

λ : Laser wavelength,

h : Planck's constant,

c : Speed of light,

r_1 : Transmission beam efficiency,

r_2 : Receiving beam efficiency,

r_3 : Target reflector reflectivity,

r_4 : Atmospheric transmission efficiency (one-way),

A_s : Total reflector surface area,

A_r : Received aperture surface area,

R_1 : Station 1 to satellite distance,

R_2 : Station 2 to satellite distance,

q_t : Transmission beam divergence, and

q_s : Reflectors beam divergence for incident light return.

Figure 4 shows the expected number of photons for AJISAI mirror reflection and the satellite elevation from stations, assuming the parameters have the values listed in Table 2. The position of AJISAI is roughly in the middle of the two station's common sky. From stations 2500 km apart, AJISAI is observed at above 40 degrees and more than 10 photons are expected.

Table 2 : Parameters for estimating the number of photons

Satellite Altitude	: 1500km (AJISAI)
Energy of pulse (E)	: 100 mJ
Wavelength (λ)	: 532 nm
Transmission efficiency (r_1)	: 0.6
Receiving efficiency (r_2)	: 0.3
Target reflectivity (r_3)	: 0.9
Atmospheric transmission efficiency (r_4)	: 0.6
Reflector surface area (A_S)	: 0.04 m^2
Received surface area (A_r)	: 0.27 m^2 (60cm aperture)
Transmission beam divergence (q_t)	: 5 arcsec
Reflectors beam divergence (q_r)	:2.5 deg (mirror of AJISAI)

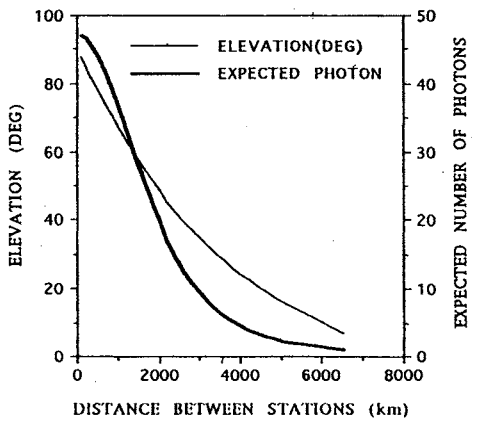


Fig.4 Elevation and Expected number of photon v.s. distance between stations in laser time transfer via AJISAI

6. SPACIAL DISTRIBUTION OF LASER REFLECTION

We simulate the spatial distribution of laser reflection from AJISAI to study the possibility of an optical link for time synchronization between stations. We calculate first the rotation phase and the incident angle of AJISAI for a given transmitting station and epoch, then project the image onto the ground from all mirrors visible to both stations. The calculation continues according to the given time step.

In Fig.5, each rectangle is the image of the laser reflection on the ground for an instantaneous laser shot and the number by the rectangle shows the reflector number on AJISAI. In Fig.6, three steps of 50 msec are illustrated successively.

A several percent of East Asia can locate an image, but there is a chance for many stations to get the link every AJISAI pass. If the timing of every shot is as accurate as the prediction, time comparison can be performed on average once every few seconds by simultaneously determining the rotation phase of AJISAI.

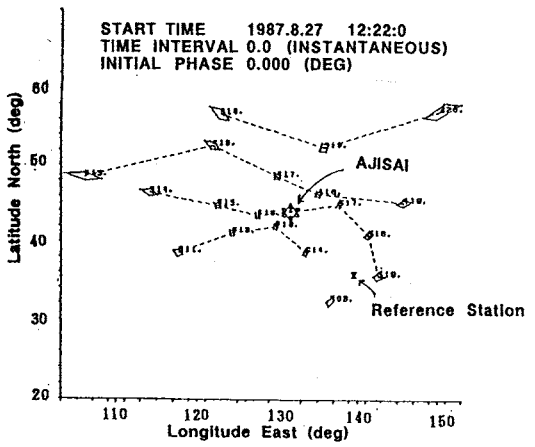


Fig.5 Instantaneous spatial distribution of reflected images on the ground via AJISAI mirror reflector. (The number next to each image indicates the mirror number on AJISAI)

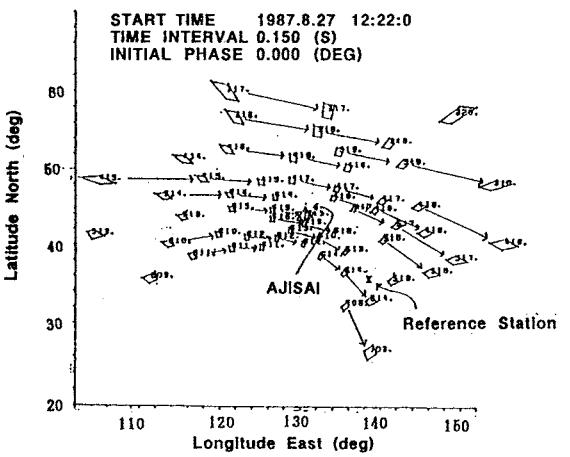


Fig.6 Spatial distribution of reflected images in three 50 msec steps.
(The number next to each image indicates the mirror number on AJISAI)

7. CONCLUSION

We have studied the feasibility of the sub-nanosecond precision laser time synchronization system with a target satellite of AJISAI based on CRL laser ranging system, whose timing system is connected with UTC(CRL).

Simulation study indicates that two stations at thousands of km distance from each other can be linked by laser beam with signal strength of more than 10 photons. The images of laser beam from AJISAI mirrors is uniformly distributed on the ground and two stations have many chances to link each other during a single AJISAI pass. It requires the μsec control of laser fire timing, however, it bring on the precise information of the orbit as well by using the self-returned satellite range data.

The system has been operating in geodetic mode since CRL started the SLR observations to the

major geodetic satellites from 1990. While global position of the SLR station has been determined with the precision of several centimeters, we are re-designing our system so that current system has also for time synchronization mode operation.

REFERENCES

1. Degnan J.J., 1985, Satellite Laser Ranging : Current Status and Future Prospects, IEEE Trans. Geosci. Remote Sensing, GE-23, 4, pp.398-4 13.
2. Pearlman M., & al. 1989 , Ground-based laser ranging measurement techniques and technology, Report of CDP-COOLFONT, section2.1, pp.1-26.
3. Serene B. and Albertinoli P. 1979, The LASSO Experiment, Proc.PTTI 11th conference, pp.145-167.
4. Veillet C. 1989, LASSO - the European Phase Aug.88-Sept.89, Proc.7th International Workshop on Laser Ranging Instrumentation, Matera 2-6 .
5. Kunimori H. & al. 1991, "Satellite Laser ranging system", Journal of CRL, Vol.38, No.2 ,pp.303-317 .
6. Greene B.A. 1984, Epoch timing for laser ranging, Proc. 5th International Workshop on Laser Ranging Instrumentation (Herstmoncex) , pp.247-250.
7. Sasaki M.and Hashimoto H. 1987, Launch and observation program of the experimental geodetic satellite of Japan, IEEE Trans.Geo.Rem., GE-25, 5, pp.526-533.
8. Kirchner D., 1991, Two-way time transfer via communication satellites, Proc.IEEE, Vol.79., No.7.
9. Electro Optic Systems 1991, EOS-V1 Active/Active Laser Characterization .

THE ACCURACY OF SYLT- AND RNS-BASED TIME SCALES AT MTI AND SNT DURING 1990-1991

Tapio Mansten[1], Kalevi Kalliomäki[1,2], Bengt Selin[3]

[1] Technical Research Centre of Finland, Metrology Research Institute (MTI)
[2] University of Oulu, Department of Electrical Engineering
[3] Swedish National Time and Frequency Laboratory (SNT)

ABSTRACT

The phases of transmissions from Norwegian Sea Loran Sylt and Soviet Radionavigation system (RNS) Bryansk stations were measured at MTI (Finland) and SNT (Sweden) daily for two years to find out the uncertainty of time scales based on them. The phases were traced back to UTC by using BIPM time circulars and appropriate time information letters from other laboratories (PTB, VNIIFTRI, USNO). The reference time scale at MTI is a GPS receiver. The reference time scale at SNT is based on one of three Cs clocks which contribute to UTC. The results show that by common view Sylt measurements a Loran-C-based time scale can be maintained at 100 ns level/10 days. By using RNS common view measurements a time scale at about 500 ns/10 day level is achieveable.

Keywords: Loran-C, RNS, time scale

1. INTRODUCTION

The phase of Norwegian Sea Loran-C Sylt transmitter is monitored and reported by several time and frequency laboratories, for instance SNT and PTB. The master station, Ejde, situates on Faeroe Islands and is monitored and reported by USNO.

Russia has a radionavigation system, RNS, that closely resembles the Loran-C system. The master station is equipped with a hydrogen maser and situates at Bryansk. The National Scientific and Research Institute for Physical, Technical and Radiotechnical Measurements, VNIIFTRI, near Moscow monitors the phase of the master station and publishes the results in time information letters.

The Metrology Research Institute (MTI) of the Technical Research Centre of Finland situates near Helsinki. It contains, amongst others, the National Standards Laboratory of Time and Frequency. MTI has no Cs clocks. The time and frequency of MTI are based on USNO MC time via a GPS Satellite receiver. The time and frequency are traceable to BIPM via atomic clock transports with SU and SNT and via time reports of BIPM. MTI also monitors the Loran-C Sylt transmitter as a secondary means of traceability and compares the results with those of SNT and PTB. The Ejde Measurement results of USNO are used, too.
In addition MTI monitors the RNS Bryansk master station phase together with SNT and Russia and uses the results as an alternate secondary means of traceability for time and frequency.

Swedish National Time and Frequency laboratory (SNT) maintains the time and frequency of Sweden. The reference time scale at SNT is based on one of three Cs clocks which contribute to UTC and traced to UTC via the GPS common view measurement program coordinated by BIPM, Paris. SNT monitors and distributes the phases of Loran-C Sylt and Bryansk RNS transmitters with MTI in an agreed Euromet project (A 089).

The positions of transmitters and receivers concerned can be seen in the map, figure 1. Some of the distances are represented in table 1.

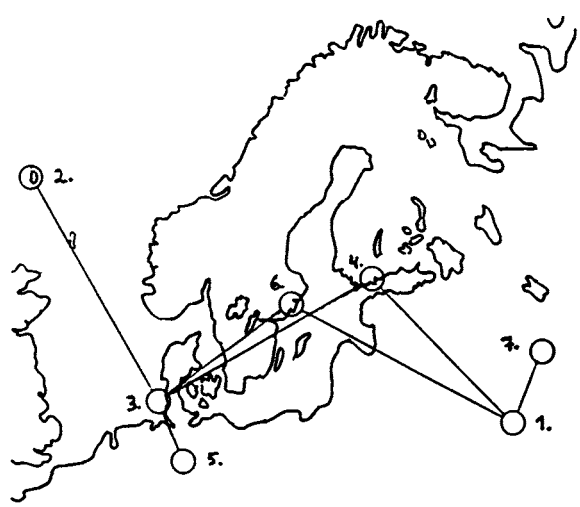


Figure 1:The locations of transmitters and laboratories
Transmitters: 1. Bryansk, 2. Ejde, 3. Sylt
Laboratories: 4. MTI, 5. PTB, 6. SNT 7. VNIIFTRI

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

Table 1:Approximate distances between the transmitters and receivers concerned

Stations	Distance, km	Over sea, %
Sylt-SNT	770	30
Sylt-MTI	1155	60
Sylt-PTB	330	30
Bryansk-SNT	1200	35
Bryansk-MTI	970	6
Bryansk-VNIIFTRI	360	0

2. MEASUREMENTS

We have compared continuously the Sylt and Byansk measurement data for almost two years (Sylt from 10. Jan 1990 to 31. Dec.1991, Bryansk from 10. Jan. 1990 to 30. Sept. 1991). The comparison was done by calculating all the daily measurement results back to UTC by using appropriate time information letters according to table 2. Linear interpolation was used for BIPM "circular T" reports.

Table 2. Time information calculation paths

	BIPM circular T	Time information letters
MTI:	UTC(MTI) = UTC(USNO)	+ c0
UTC-UTC(Rx1Sylt)	= UTC-UTC(USNO)	+ UTC(MTI)-UTC(Rx1Sylt) + c1
UTC-UTC(Rx1RNS)	= UTC-UTC(SU)	+ UTC(MTI)-UTC(Rx1RNS) + c2
SNT:		
UTC-UTC(Rx2Sylt)	= UTC-UTC(SNT)	+ UTC(SNT)-UTC(Rx2Sylt) + c3
UTC-UTC(Rx2RNS)	= UTC-UTC(SNT)	+ UTC(SNT)-UTC(Rx2RNS) + c4
PTB:		
UTC-UTC(Rx3Sylt)	= UTC-UTC(PTB)	+ UTC(PTB)-UTC(Rx3Sylt) + c5
USNO:		
UTC-UTC(Rx4Ejde)	= UTC-UTC(USNO)	+ UTC(USNO)-UTC(Rx4Ejde) + c6
VNIIFTRI:		
UTC-UTC(Rx5RNS)	= UTC-UTC(SU)	+ UTC(SU)-UTC(Rx5RNS) + c7

c1 = +0.5 µs, c2 = -70 µs, c3 = +1.25 µs, c4 = -4000 µs
c5 = -0.9 µs, c6 = +0.6 µs, c7 = -1202 µs

The arbitrary constants c1..c7 were fixed in the beginning of the comparison period. Their purpose is to keep the average values of the compared results near zero for convinience in graphs etc. SNT commercial receiver ±n * 5 µs phase slips were also corrected.
The results were averaged in 11 day groups around 9-ending MJD's. Only the days that had been reported by all of the participating laboratories were included. The 11-day average values were then compared to each other by calculating their mutual differences.

3. RESULTS

In figure 2 are the results of Loran-C Sylt measurements with UTC as reference. The constants of table 2 have been added to time information letter values.

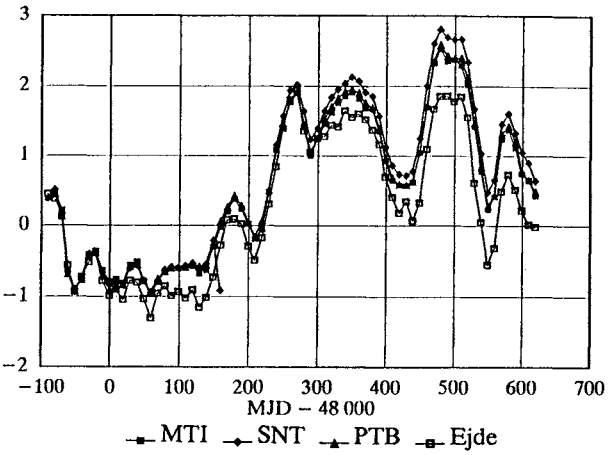


Figure 2: The overall behaviour of Sylt measurements compared to UTC. Y-axis is in µs.

In Figure 3 are the deviations of three different Loran-C Sylt timescales from UTC(MTI).

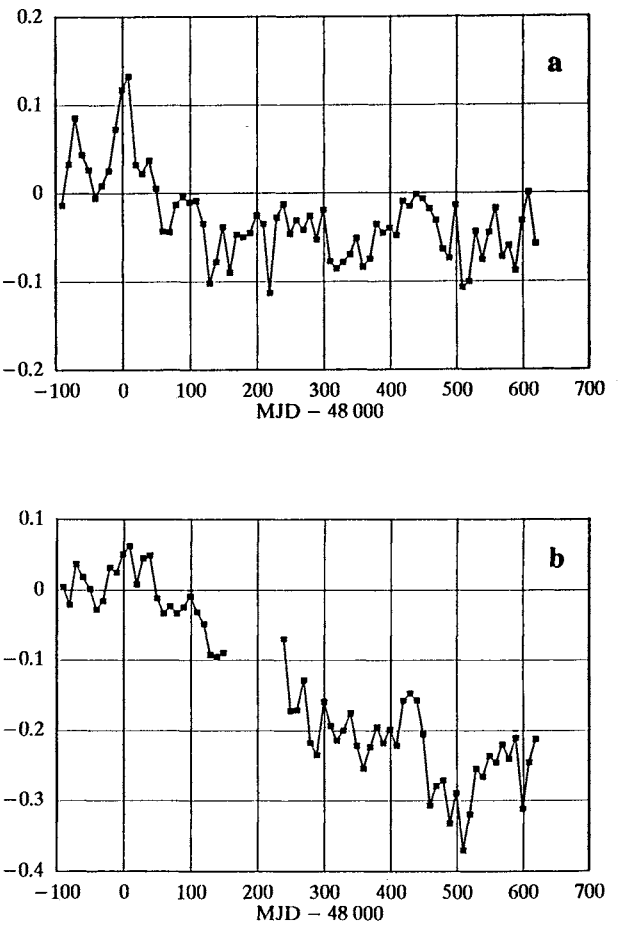


Figure 3:
a) UTC(MTI) - Loran-C time scale via PTB, µs
b) UTC(MTI) - Loran-C time scale via SNT, µs

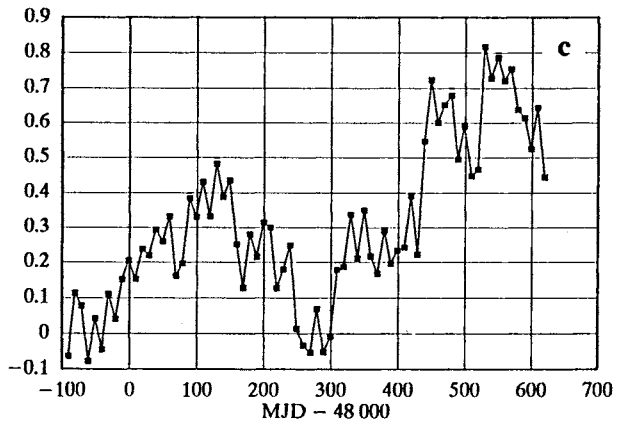


Figure 3:
c) UTC(MTI) - Loran-C time scale via USNO, μs

In figure 4 can be seen the deviations of two different Loran-C Sylt timescales from UTC(SNT).

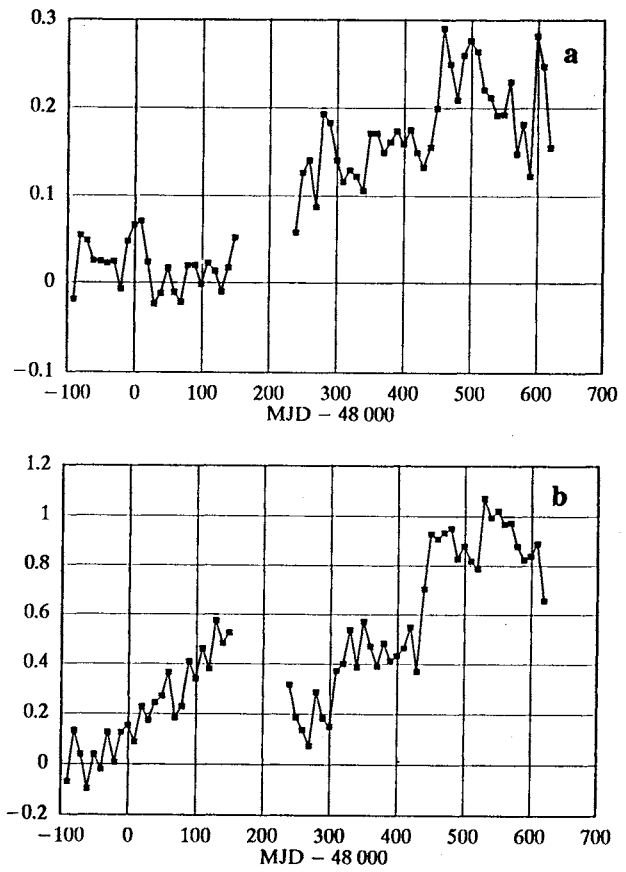


Figure 4:
a) UTC(SNT) - Loran-C time scale via PTB, μs
b) UTC(SNT) - Loran-C time scale via USNO, μs

The statistics of Sylt results are in table 3. At MTI we had at first difficulties with our GPS receiver, so the results have been divided in two parts, except in the case of Ejde, which is "noisy". SNT had a receiver failure and after repairment their results differed from the previous ones, so their results are in two parts, too.

Table 3a: Loran-C Sylt time scales at MTI:

	MJD's	AVG, ns	SD, ns	Trend, ns/day	t
Via PTB:	47900-48049	+ 40	40	-	-
	48059-48620	- 50	30	-	-
	47900-48620	- 30	50	-0.13	- 5.6
Via SNT:	47900-48149	- 9	45	-0.38	- 3.9
	48239-48620	-225	60	-0.30	- 4.3
	47900-48620	-140	120	-0.50	-18.0
Via Ejde:	47900-48620	+310	230	+0.77	+ 8.0

Table 3b: Loran-C Sylt time scales at SNT:

	MJD's	AVG, ns	SD, ns	Trend, ns/day	t
Via PTB:	47900-48149	+ 20	25	-	-
	48239-48620	+180	55	+0.29	+ 4.4
	48900-48620	+115	90	+0.36	+14.1
Via Ejde:	47900-48149	+215	185	+2.35	+11.3
	48239-48620	+615	290	+2.25	+10.7
	47900-48620	+460	320	+1.26	+13.1

In figure 5 are the results of RNS Brjansk master measurements with UTC as reference. The constants of table 2 have been added to reported values.

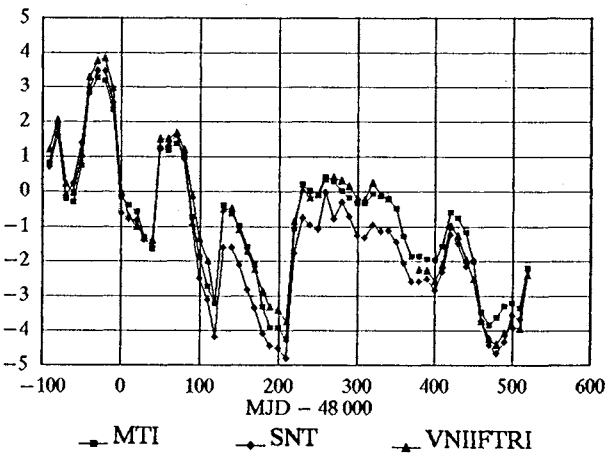


Figure 5: The overall behaviour of RNS Bryansk measurements compared to UTC. Y-axis is in μs.

Figure 6 represents the deviations of two different Bryansk RNS timescales from UTC(MTI).
Figure 7 represents the deviation of Bryansk RNS timescale from UTC(SNT).

The statistics of RNS results are in table 4. Around MJD 48070 the SNT receiver experienced an unexplained phase jump. Between MJDs 48338..376 the VNIIFTRI RNS receiver was faulty and after recovery its phase had jumped, too. These jumps have been taken into account in table 4.

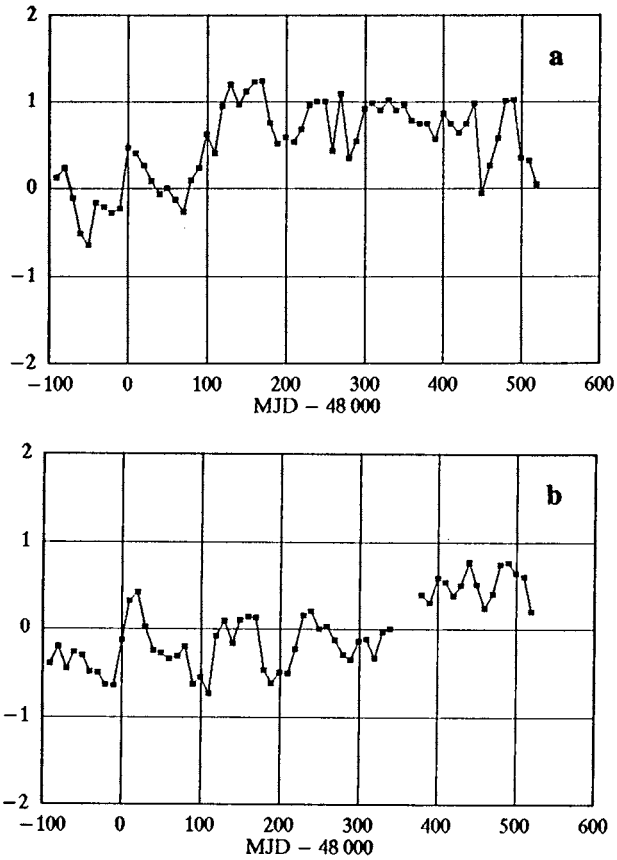


Figure 6:
a) UTC(MTI) - RNS time scale via SNT, μs
b) UTC(MTI) - RNS time scale via VNIIFTRI, μs

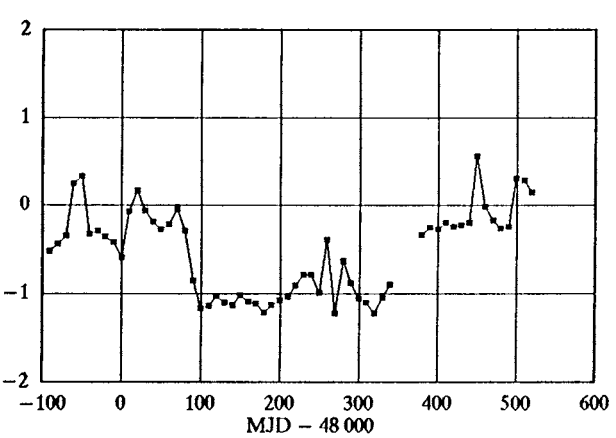


Figure 7:
UTC(SNT) - RNS time scale via VNIIFTRI, μs

Table 4a: RNS time scales at MTI:

	MJD's	AVG, ns	SD, ns	Trend, ns/day	t
Via SNT:	47900-48079	- 50	290	-	-
	48089-48520	+740	310	- 0.71	- 2.0
	47900-48520	+510	470	+1.34	+ 4.6
Via VNIIFTRI:	47900-48339	-215	280	-	-
	48379-48520	+515	180	-	-
	47900-48520	- 35	410	+1.58	+ 7.4

Table 4b: RNS time scale at SNT:

	MJD's	AVG, ns	SD, ns	Trend, ns/day	t
Via VNIIFTRI:	47900-48079	-200	250	-	-
	48089-48339	-1000	190	-	-
	48379-48520	- 80	270	+3.46	+ 2.5
	47900-48520	- 520	490	-	-

4. CONCLUSIONS

Loran-C Sylt time scales at MTI provide a low noise, low trend secondary path for traceability. Both PTB- and SNT-data allows for ±100ns uncertainty/10 days. Because PTB master clock behaves smoothly and the Sylt reports of PTB come fast (usually in two..three weeks), MTI can relatively quickly verify the national time maintained via our GPS receiver. Loran-C Sylt time scale at SNT via PTB is somewhat worse, but allows for ±100 ns uncertainty/ 10 days if receiver failure phase step is eliminated.

USNO reports of Ejde have considerable trends and high scatter compared to Sylt. By utilizing USNO/Ejde results a time scale can be maintained at ±1μs level/10 days only.

Sylt is phase-locked to Ejde master station, so one would not expect the observed long-time deviation trends between the two stations. Probably they are due to some slow weather patterns in the North Atlantic propagation path.

RNS time scale at MTI offers another low trend secondary path, but the achievable noise level is much higher than with Sylt, about ±500ns/10days via VNIIFTRI if the receiver failure step at VNIIFTRI is eliminated, and somewhat worse via SNT. The RNS time scale at SNT achieves about the same level.

Compared to Sylt the RNS signal phase is noisy and subject to frequent phase jumps and relatively fast drift (typically $6*10^{-13}$/day). Because of high phase scatter no long-time trends can be seen. The signal is strong in the whole Finland, so it can be used for time-keeping purposes far from the coast. The noise in RNS result differences is about a decade higher than in the Sylt result differences although the distances are about the same.

By portable clock visits the constants of table 2 can be corrected. The last H-maser visit from VNIIFTRI to MTI in January 1992 caused a 300 ns correction to UTC(MTI).

Russia measures Sylt, too, but the results obtained at Pulkova, near Petrograd are very noisy, and were therefore not included into the comparison.

Reference: Tapio Mansten, Kalevi Kalliomäki and Bengt Selin; Comparison of Soviet radionavigation system and Sylt-Loran with GPS as a reference, EFTF 91, Besancon

CONTROL OF ENVIRONMENTAL AND CLOCK PARAMETERS IN TIME AND FREQUENCY STANDARDS LABORATORY

Kalevi Kalliomäki [1,2] , Tapio Mansten[2]

[1]University of Oulu, Department of Electrical Engineering
2 [2]Technical Research Centre of Finland, Metrology Research Institute (MTI)

ABSTRACT

Environmental parameters and frequencies of rubidium and precision crystal oscillators have been measured during 1990..1991 in two time and frequency laboratories in Finland. The effects of ambient fluctuations to oscillator frequencies were then studied using multiple linear regression. Laboratory temperature seem to be most significant factor even in air conditioned rooms. Additional thermally controlled cabinets have been built to reduce temperature effects.

Key Words: Environmental effects, Aging, Temperature stability.

INTRODUCTION

Precision time and frequency standards are subject to environmental effects such as temperature, pressure, humidity and power line voltage. Diurnal and annual variation of ambient temperature affects inside temperature even in air-condioned laboratory spaces leading to syncronous drift behaviour of frequency standards. In addition instruments generate temperature patterns around them affecting other clocks nearby.
In small countries like Finland time and frequency standards laboratories cannot be big and expensive cesium-atomic clocks are rare. To manage with cheap (rubidium, crystal) clocks in limited spaces one has to study and largely eliminate disturbing environmental effects extensively.
In this paper environmental factors and their effects into time and frequency standards are studied, relatively small laboratory in mind. A computer controlled data acquisition system was built to monitor relevant factors. The performance of our standards is now improved by about one decade.

Materials

Computer controlled environmental monitoring has been applied in two time and frequency standards laboratory, 1) in the Metrology Research Institute of Technical Research Centre of Finland (VTT) and 2) in the laboratory of Electronic Measurements at the University of Oulu. Both laboratories have several (3..5) temperature monitoring channels and one pressure channel. Outside and inside temperatures are monitored. The first laboratory has a special air-conditioned room for clocks, the second laboratory uses a usual room. To stabilize the temperature of primary clocks as carefully as possible, they are placed into closed, thermally controlled cabinet. The operating voltages of the clocks are monitored and externally stabilized to keep power dissipations fixed. A fan circulates the air inside clock cabinet to keep thermal gradients low.
In this study the monitored clocks are as follows. Two R&S XSRB rubidium oscillators (3 and 15 years old), two EFRATOM FRS-C rubidium oscillators (1 year old) and three 5 MHz crystal oscillators, HP 105A (17 years old) , OSA 8601 BVA (10 years old) and R&S XSF (20 years old). As a time and frequency reference there are several Loran-C receivers and one GPS-receiver.

Methods

The main clock is a GPS-receiver with disciplined rubidium oscillator (KINEMETRICS DC & EFRATOM FRS). Analog (12 bit resolution) and digital channels (1 ns resolution) are sampled every ten minutes and a data set called Q-data (QUICK) is formed. This relatively data base is saved for one year only. Every two hours Q-data is automatically compressed in 1 to 12 ratio into a new data set called S-data containing two hour averages for every channel. The next compression step, M-data, gives daily averages, trends and scatter.
When studying environmental effects into clocks, daily averages are normally used. Only in case of transient phenomena or delay time determinations by cross-correlation method, S- or Q-data is used.

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

Because environmental effects primarily affect the frequency of oscillators, frequency changes are studied as a function of environmental parameters. Multiple regression analysis with elapsed time, temperatures and pressure as variables has been applied. Lotus 123 spreadsheet program is used for final data analysis and presentation.
During the visit of a transportable hydrogen maser "Saphire" from VNIIFTRI, Moscow, the short term fluctuations of our clocks were also studied.

Results

The annual variation of average daily outside temperature in Helsinki was 35 °C and 40 °C in Oulu, recpectively. Typical daily variation is 5 °C, highest in springtime. Inside normal laboratory rooms maximum variation is up to 10 °C, temperature can be as high as 30 °C during hot summer days. The "shielding factor" (outside/inside variation) is around 7 in normal laboratory room. In air condioned laboratory rooms temperature variation is well inside +/- 1 °C and the shielding factor is about 30. Daily outside fluctuation is clearly visible inside but six hours (90 degrees) late. The thermal time constants of buildings seem to be 1..2 days. The problems are associated with service periods of air conditioner. After service the temperature may have changed.

To build a thermally stable clock cabinet is not an easy task. Due to limited space several clocks has to be put into the same cabinet. In our designs the shielding factor ranges from 10 to 30. The laboratory and outside temperatures still correlate with cabinet clock frequencies indicating that the design is not perfect.

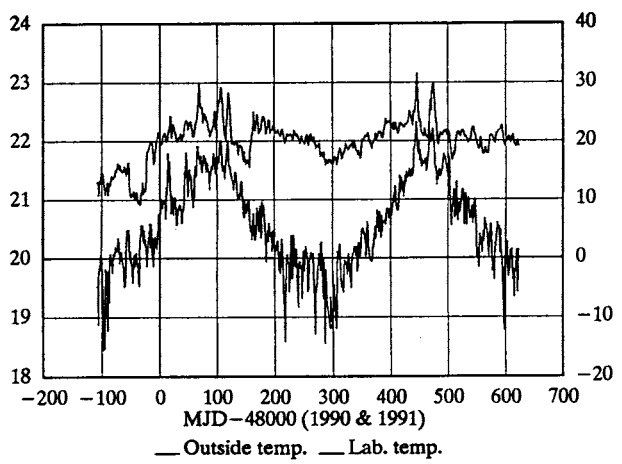


Fig 1 Temperature variation outside and inside the laboratory of VTT during 1990..1991.

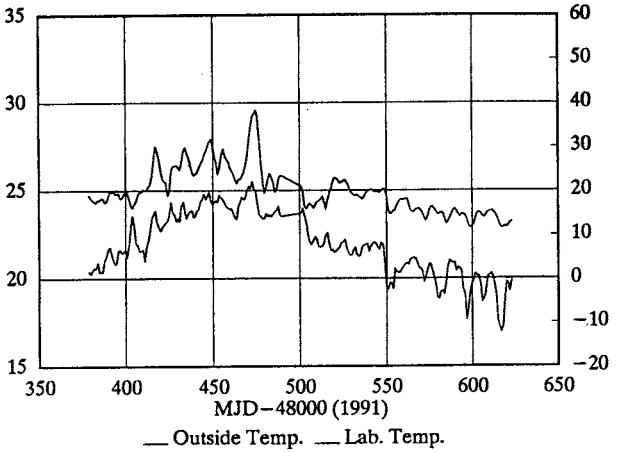


Fig 2 Temperature variation outside and inside the laboratory of Oulu during 1991.

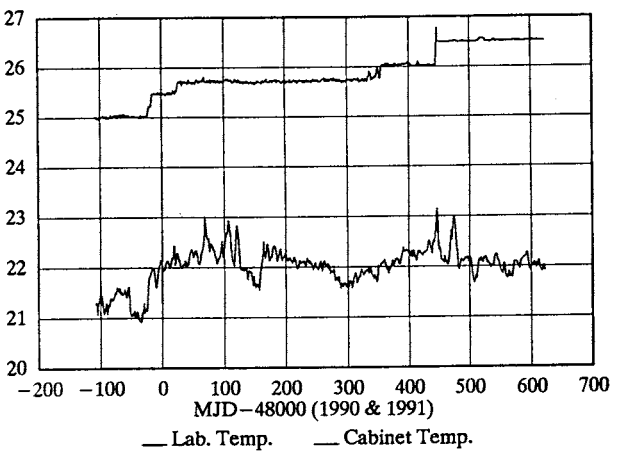


Fig 2 Temperature variation outside and inside the laboratory of Oulu during 1991.

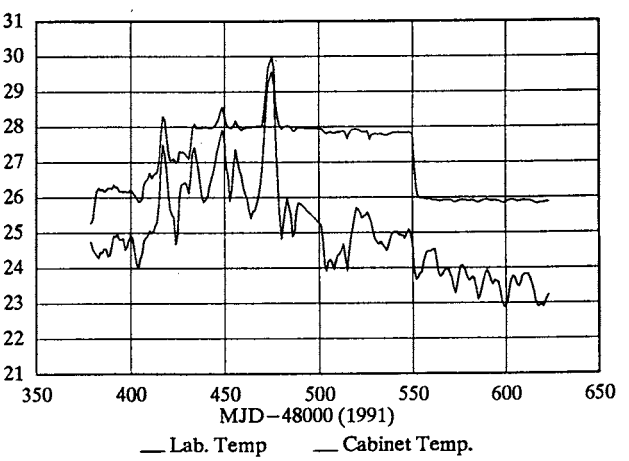


Fig 4 Temperature variation inside the laboratory and clock cabinet in Oulu

All selected variables (time, temperatures and pressure) were as a rule significant. Table 1 shows the coefficients, significance levels (t-value) and coefficients of determination for crystal oscillators in study.

Table 1. Aging, temperature and pressure coefficients of OXCO:s, unit exp-10/factor

	HP105	OSA 8601	R&S XSF
Aging (1/d)	-0.05	0.03	-0.03
Temperature	2	5	1
Pressure	ns	-0.1	ns

Next table (table 2) shows multiple regression coefficients and significance levels for rubidium oscillators in study

Table 2. Aging, temperature and pressure coefficients of rubidium oscillators, unit exp-11/factor

	R&S1 XSRB	R&S2	EFRATOM FRS-_C
Aging (1/m)	-0.13	0.24	-1.2
Temperature	-0.5	0.06	-2
Pressure	0.003	-0.004	-0.015

The estimated and observed frequencies for different oscillators are plotted in figures 3..5

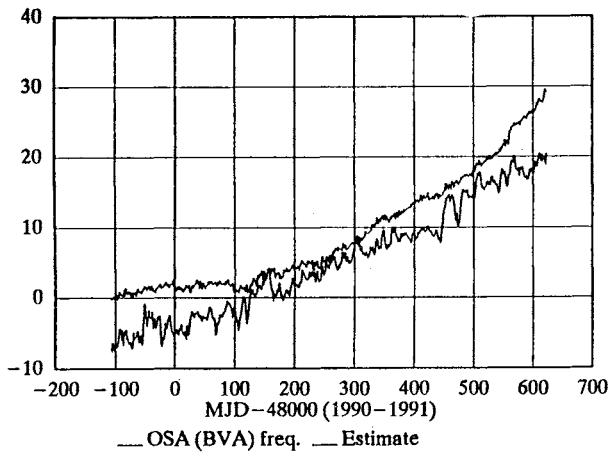


Fig 5 The estimated and observed frecuency of a OSA 8601

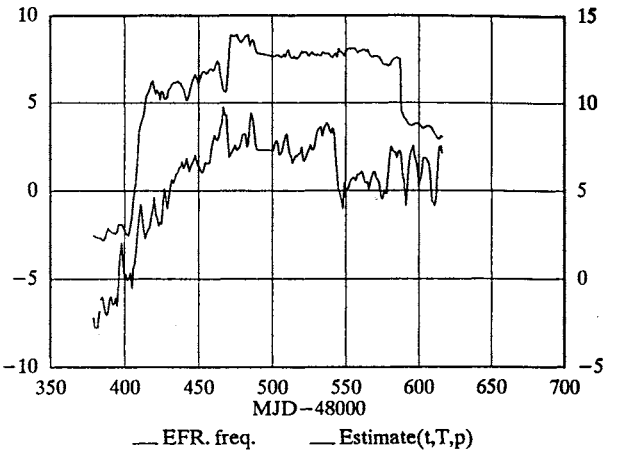


Fig 6 The estimated and observed frecuency of a rubidium oscillator (Efrat., FRS-C)

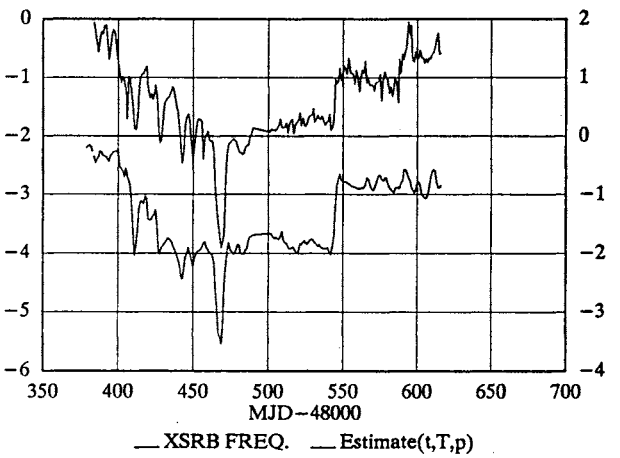


Fig 7 The estimated and observed frecuency of a rubidium oscillator (R & S, XSRB)

Discussion

As mentioned above, the number of studied precicion clocks is limited and most expensive atomic clocks are beyond the scope of this paper. Hewlett-Packard, Oscilloquarz and Rohde & Schwarz are well-known high quality ovenized crystal oscillator makers, however. Rohde & Schwarz and Efratom rubidium oscillators are quite common at least in Europe. Therefore even a small number of samples is sufficient.
Due to limited resolution (1 ns) of time interval counter and lack of extremely stable reference like hydrogen maser, the short term effects are not extensively studied. We have concentrated into diurnal and slower effects.

In spite of good thermal insulation in Finnish buildinds and extra insulation in our air-conditioned time and frequency laboratory, the outside temperature has a significant influence on the frequency of the clocks. The time laboratory in VTT/MTI is a "building inside building" ie. it does not have a wall touching outside air. The outside temperature variation seems to penetrate the most "holy", thermally controlled clock cabinet, too. Later we realized, that the circulation of air inside the cabinet is not effective and homogeneous enough and the shielding factor is quite bad expecially at the top of the cabinet. On the other hand our experience tells that forced ventilation may deteriorate the performance of precision oscillators.
The temperature coefficients mentioned in Tables 1 and 2 are not very exact, because different combinations of ambient factors give diverse results.
The pressure effects on frequency seem to be quite low compared to thermal effects. Thus the compensation for pressure is not necessary as a first approximation. Only extremities in pressure have to be taken into account.

In case of R&S rubidium oscillator the estimated and measured frequencies match well, the coefficient of determination is 0.900 ! Efratom rubidium seems to suffer from unexpected frequency changes.
The frequency drifts of different oscillators followed quite well linear model. Only HP105 and FRS-C displayd some additional random walk behaviour.
Temperature gradients inside laboratory could be significant too. Therefore the measurement of laboratory temperature is a difficult task in fact. We have observed several frequency changes of precision oscillators when the position of nearby instruments has changed.

Reference: K.Kalliomäki et al 1990 A computerized time and frequency standards monitoring system, EFTF 90, Neuchatel, 177..180

ANALYTICAL TECHNIQUES FOR IMPROVED TIME TRANSFER USING COMMON VIEW OF DIRECT BROADCASTING SATELLITES

John A Davis and Peter R Pearce.

National Physical Laboratory,
Queens Road, Teddington, Middlesex, TW11 0LW, UK.

ABSTRACT

Time transfer using TV signals from direct broadcasting satellites is an accurate and inexpensive method of comparing atomic clocks at distant locations. The major source of error arises from the residual motion of the geostationary satellite. Previous work [1] has shown that it is possible to determine the components of the satellite motion by observing two widely spaced geostationary satellites simultaneously from three locations. In this paper, a method is developed where the common view measurements are corrected for the variation of the positions of the satellites. In contrast to previous work there is no requirement to calculate explicitly the co-ordinates of the satellites. Time transfers with uncertainties of a few parts in 10^{14} have been obtained.

Keywords: time measurement, direct broadcasting by satellite, distance measurement.

1. INTRODUCTION

The use of television signals from direct broadcasting satellites (DBS) for common view time comparison experiments has received attention recently within Europe. This is due to the presence both of high powered DBS and medium powered television distribution satellites, the signals from which may be received with small antennas [1,2,3]. Several methods of improving the accuracy of time transfers by determining the position of the geostationary satellite have been studied. These include satellite position calibration by GPS receivers [2], the use of satellite ranging data [2,4], and pseudo-ranging techniques [1,2,4,5].

In this paper improved methods for correcting the common view measurements for variations of satellite position are described. Their advantages over those given in the previous paper [1] are discussed, along with the limiting errors. An investigation of their accuracy over distances of up to 400km has been carried out. The satellite motion is composed of sinusoidal variations with periods of approximately 12 and 24 hours and longer term drifts. The determination of these components provides a valuable check on the accuracy achieved in the time transfer results.

The experimental arrangement has been described previously [1]. Four organizations observed television signals from both the Astra (19.2°E) and Marco Polo (Britsat) (31.0°W) satellites. The collaborating organizations were the National Physical Laboratory (referred to in this paper as Teddington), Rugby Radio Station, Hillmorton, Rugby, CV23 OAS, (Rugby), Hewlett Packard Limited, Winnersh Triangle, Wokingham, Berkshire, RG11 5DZ, (Winnersh), and Radiocode Clocks Ltd, Radiocode House, Jennings Road, Kernick Road, Penryn, Cornwall, TR10 9LY, (Penryn). The atomic clocks at Teddington, Rugby, Winnersh and Penryn were HP5061A Option 004, HP5061A, and HP5061A Option 004 caesium clocks, and a rubidium standard respectively.

2. CORRECTION OF THE COMMON VIEW MEASUREMENTS

2.1 Origins of satellite position variations

The residual motion of a geostationary satellite observed in a geocentric co-ordinate system fixed with respect to the earth's surface has several components [6]. A non zero satellite orbit eccentricity causes both the radius r and longitude φ of the satellite to vary with period of one sidereal day, but with a phase difference of $\pi/2$, and amplitudes er and $2e$ respectively. Inclination of the plane of the satellite orbit away from the plane of the equator produces variations in the latitude θ with a period of one sidereal day. Movement with a period of half a day is caused by tidal forces on the satellite from the sun and moon. Perturbations of the satellite orbit produce a slow change in both amplitude and phase of each of the periodic variations. In addition, station-keeping orbit corrections result in discrete changes in the amplitude and phase of the variations of period one sidereal day.

In this work, the non-periodic variations in the position of the satellites are assumed to be a function of the longitude φ only. The mean latitude and radius averaged over a sidereal day are assumed to be zero and 42,164.2 km respectively. This approximation is limited by the longer term satellite excursions in the radius, typically of the order of only a few km when frequent longitude orbit corrections occur [6]. From geometry considerations, this will introduce an error of a few ns into a long distance time transfer within the UK.

2.2 Composite measurement method

Previous work [1] has demonstrated that by observing two satellites from three widely spaced locations it is possible to determine the residual longitudinal drifts of both satellites, and to correct the common view measurements for long term

satellite drifts. An analytical method is developed here to correct the measurements of the clock differences without the need to explicitly calculate the satellite longitude.

Consider common view measurements between three laboratories. The measured time differences M_{abi}, between laboratories a and b via satellite i are given by

$$M_{abi} = D_{abi} + T_a - T_b \quad (1)$$

where D_{abi} is the delay difference between the transmitted signal detected at the two locations, and $T_a - T_b$ is the corrected time difference measurement between the clocks at location *a* and *b*.

D_{abi} has a linear dependence on small perturbations $\delta\phi_i$ of the satellite longitude where

$$\delta D_{abi} = K_{abi}\delta\phi_i \quad (2)$$

and K_{abi} the constant of proportionality is calculated from geometry and the speed of light *c*. The following linear combinations C_i of time difference measurements (composite measurements) may be formed from the common view measurements between three locations (a, b, c) and a single satellite (i),

$$C_i = \alpha_i(M_{abi} - N_{abi}) + \beta_i(M_{aci} - N_{aci}) \quad (3)$$

where N_{abi} and N_{aci} are the reference delay differences D_{abi} and D_{aci} calculated using the nominal values of satellite longitude, latitude and radius. Considering the variation of the delay differences to be a function of satellite longitude only,

$$\delta C_i = \alpha_i(K_{abi}\delta\phi_i + T_a - T_b) + \beta_i(K_{aci}\delta\phi_i + T_a - T_c) \quad (4)$$

If the constants α_i and β_i are chosen such that $\alpha_i K_{abi} + \beta_i K_{aci} = 0$, the composite measurements are independent of the variations of satellite longitude. When the satellites are separated along the geostationary arc and the three ground stations form a triangle on the earth's surface the constants K_{abi} and hence the constants α_i and β_i are different for each satellite.

The corrected time difference measurements $T_a - T_b$ and $T_a - T_c$ may now be determined in a longitude-independent form as

$$T_a - T_b = (\beta_2 C_1 - \beta_1 C_2)/(\beta_2\alpha_1 - \beta_1\alpha_2) \quad (5)$$

$$T_a - T_c = (\alpha_2 C_1 - \alpha_1 C_2)/(\alpha_2\beta_1 - \alpha_1\beta_2) \quad (6)$$

The corrected measurements $T_a - T_b$ and $T_a - T_c$ contain residual periodic components originating from variations in the satellite radius and latitude.

In comparison with the previous analytical method [1] the above formula are simple to program. The requirement to explicitly determine the satellite longitude made the previous method considerably more complex. In the previous method data was averaged over a sidereal day, thus making the technique sensitive to missing data. This contrasts with the present method where the clock difference measurements are corrected for longitudinal drift without any requirement to average the data. Periodic variations of the satellite longitude, and longitudinal components of the satellite orbit corrections are also eliminated from the clock difference measurements by the composite method.

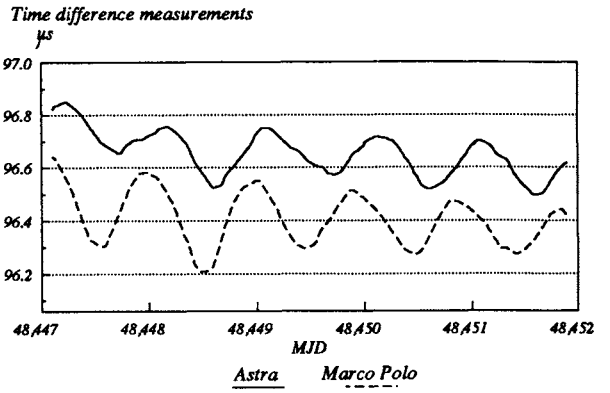


Figure 1. Common view measurements (Rugby - Teddington) via the Astra and Marco Polo satellites

Examples of measurements of the clock differences between Rugby and Teddington are shown in Figures 1 and 2. The uncorrected values M_{abi} are shown in Figure 1 with the fixed delay difference N_{abi} subtracted as in all examples shown. The measurements corrected by the composite method (Figure 2), have nearly identical daily periodic variations originating from the satellite motion, with either Penryn or Winnersh used as the third (calibrating) station. The 100ns difference between the two curves is due to errors in determining the ground station positions and from delays in the satellite receivers. A variation of less than 10ns was observed in the shape of these two curves, over a period of 35 days, demonstrating that in many cases the method was not sensitive to the choice of the third ground station. No long term delay variations were observed, while the short term noise was attributed to the satellite receivers.

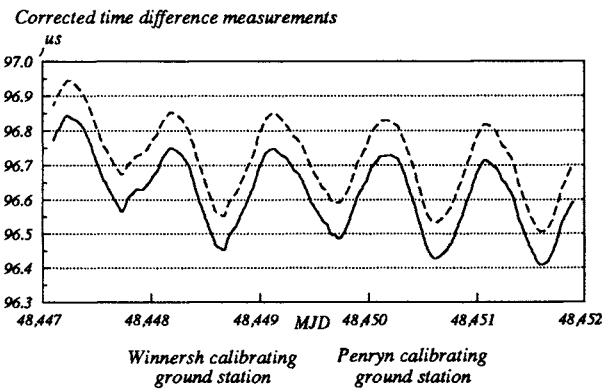


Figure 2. Common view measurements (Rugby - Teddington) corrected by the composite measurement method

2.3 Digital filtering techniques

A digital filtering method was used to remove the periodic variations of satellite origin from the corrected measurements. The filtering was applied after the composite measurement method. A notch filter was used, with reject frequencies at both half and one day [7]. Data was processed in discrete sections, a un-broken data set being required within each section. Missing data of periods up to one day was interpolated, longer gaps resulted in the data being processed in separate sections. Data either side of a satellite orbit correction was also processed in separate sections.

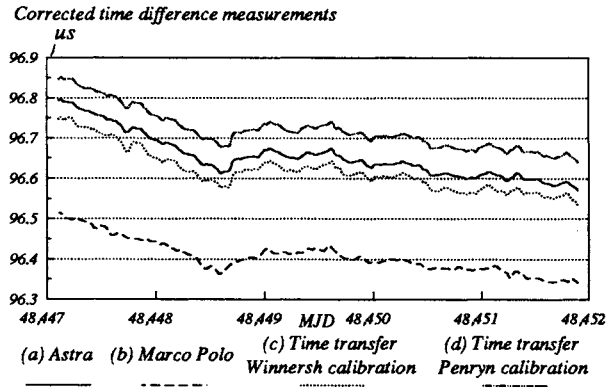


Figure 3. Time transfers (Rugby - Teddington)

2.4 Results

Plots of the time transfer results after digital filtering are shown in Figure 3. No satellite orbit corrections occurred during the period shown. In curves 3a and 3b the periodic motion of the satellite but not residual drifts have been removed, the common view measurements being made via the single satellite indicated. The differences in the shapes of the two curves may be accounted for by receiver delay variations and satellite longitude drifts. The offset between the two curves is due to errors in determining the earth station co-ordinates, in receiver delay differences, and from longitudinal drift of the satellite. The longitudinal drifts contributed errors of up to 200ns. Time transfers (curves 3c and 3d) were determined by the composite measurement method followed by digital filtering. The offset was less than 100ns, while the shapes of the two curves agreed to within 10ns.

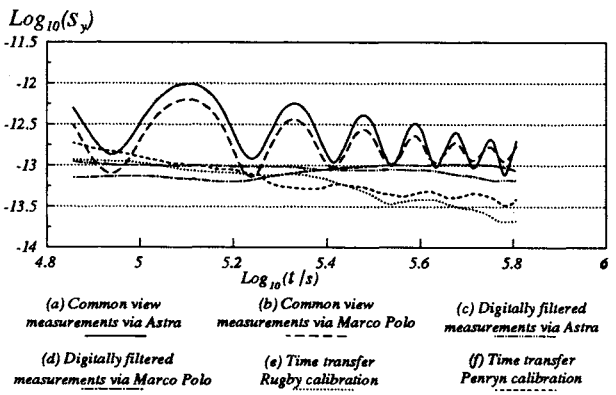


Figure 4. Plots of $Log_{10}\sigma_y$ *against* $Log_{10}T$ *for Winnersh - Teddington measurement of clock differences*

Plots of the Allan variance against the averaging time T are shown on Log - Log plots in Figures 4 and 5. Approximately 1000 measurements were averaged to produce each point. These points are then used to calculate the σ_y values, from which the curves of Figures 4 and 5 are drawn. Averaging reduces the random noise contribution to the σ_y values, introduced via the satellite receivers.

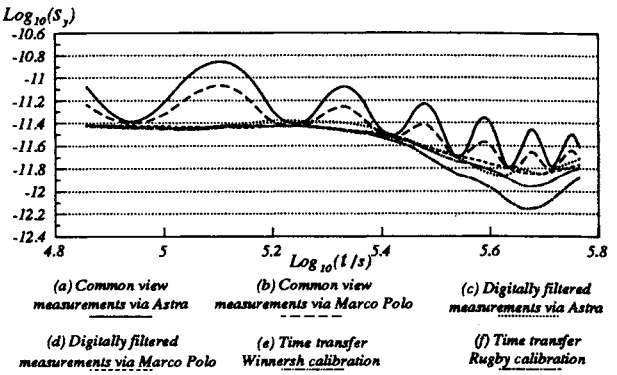


Figure 5. Plots of $Log_{10}\sigma_y$ *against* $Log_{10}T$ *for Penryn - Teddington measurement of clock differences*

Curves of $Log_{10}\sigma_y$ against $Log_{10}T$ after filtering (Figures 4 and 5) do not possess minima when T is a multiple of a day, demonstrating the effectiveness of the digital filtering. This contrasts with the previous curve fitting method [1], where only partial elimination of the periodic variations was achieved.

The application of the composite method, followed by filtering (curves 4e and 4f) resulted in a lower variance for time transfer σ_y compared with filtering the common view measurements from a single satellite (curves 4c and 4d). Variance of time transfer of a few parts in 10^{14} were obtained (curve 4e). The composite measurement method is more effective when a triangular arrangement of ground stations is used where all the perpendicular bisectors are relatively long. The ground station triangle Penryn - Winnersh - Teddington has a perpendicular bisector of only 17km. Using Rugby (curves 4e and 5f) rather than Penryn or Winnersh (curves 4f and 5e) as calibrating station reduced the uncertainties in the time transfer results. Both time transfers have better accuracies than obtained by filtering the common view measurements (Figures 4 and 5) made via a single satellite, showing that even a poor arrangement of ground stations did correct the common view measurements for satellite position variation.

The σ_y values are higher for the Penryn - Teddington measurements, due to the rubidium standard at Penryn (Figure 5). The arrangement of ground stations is also less favourable, a 365km Penryn - Teddington link being calibrated by a 120km Rugby - Teddington link. Time transfers obtained by using the composite measurement method employing a good triangle of ground stations (curve 5f) gave the lowest variance for time transfer σ_y. The usefulness of the time transfer method over the longer links has thus been demonstrated.

2.5 Sources of error

Errors in correcting the common view measurements for the long term drifts of satellite position are principally due to differential delay changes in the satellite receivers, a typical value being 10ns [1]. The resulting error depends on the locations of the three earth stations, the nominal positions of the geostationary satellites, and the validity of the initial assumptions (2.1). Sources of errors in the digital filtering method arise from edge effects at the boundary of the data set. Filtering short sections of data due to presence of frequent orbit corrections, or missing data, increases the resulting uncertainties. Errors may also occur if the clock rates possess variations of one day or half day period (3.1).

3. DETERMINATION OF SATELLITE POSITIONS

3.1 Introduction

There has been interest recently in using common view measurements of geostationary satellites to determine satellite longitude, for example in a joint European experiment at present in progress [8]. The satellite motion is considered to be composed of periodic variations of both a sidereal day and half day period, and a residual longitudinal drift. The magnitude of each of these components are determined here by separating them with an appropriate digital filter.

The reproducibility of the results obtained using different arrangements of ground stations provides a useful check on the accuracy of the time transfer results. Errors in determining the satellite longitudinal drift have the same origin as those present in the composite measurement method. If the clock rates possess variations of one day or half day period the digital filtering process will extract the combined periodic variations of both clock and satellite origin. The clock rates may be perturbed by daily variation in the ambient temperature or other environmental factor. Errors then occur when the satellite longitude and latitude periodic variations are determined from different triangles of ground stations (Figure 8).

3.2 Determination of satellite longitudinal drift

The longitudinal drift is determined by forming composite measurements from a triangular arrangement of ground stations. The differences between common view measurements made via the two satellites are used to form composite measurements. The factors α_i and β_i are chosen to make the composite measurements C_i insensitive to the longitude variation of one of the satellite, their long term variations are then proportional to the longitude of the other satellite. Periodic variations are removed by digital filtering.

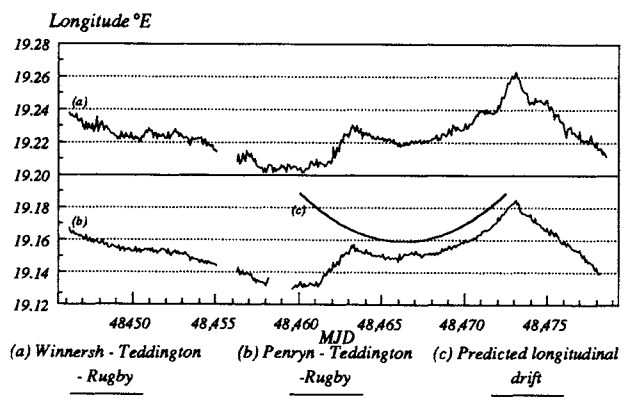


Figure 6. Satellite longitudinal drift - Astra

The longitudinal drifts are calculated from the Winnersh - Teddington - Rugby triangle of ground stations. These results (curves 6a and 7a) which included the Winnersh - Teddington link of only 37km, agreed well with those obtained from a larger triangle of ground stations (curves 6b and 7b). Determination of the longitudinal drift from the Winnersh - Teddington - Penryn triangle, with its short perpendicular bisector, was poor.

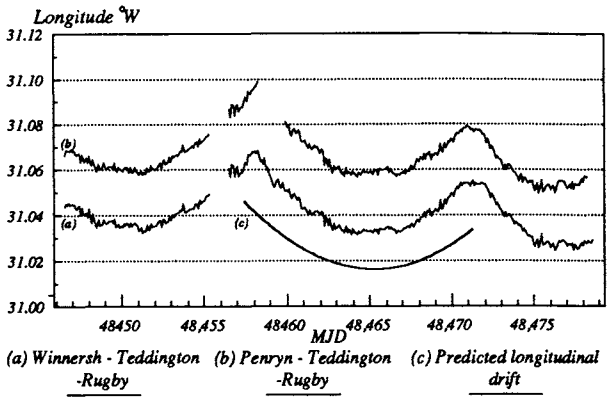


Figure 7. Satellite longitudinal drift - Marco Polo

Discrepancies in the absolute determination of mean longitude (excluding periodic variations) appear as a linear offset between the curves (Figures 6 and 7). These offsets are due to errors in determining ground station positions and receiver delay differences. The determination of the mean longitude variation agree well, the maximum differences between the shape of the curves being approximately 0.005 degrees and 0.003 degrees for the Astra and Marco Polo satellites respectively.

A satellite in geostationary orbit will experience a small uniform force in its longitudinal direction, the magnitude of which depends on the satellite position [6]. Good agreement is obtained between the predicted and measured satellite motions, shown in curves 6c and 7c.

3.3 Determination of the sidereal day periodic variations of satellite longitude and latitude

The sidereal day variations of both the satellite latitude and longitude were determined from composite measurements using suitable bandpass filters. These composites are independent of satellite longitude and latitude changes respectively. The resulting periodic variations contained components due to changes in orbital radius and longitude. The ratio of the amplitudes and phase differences ($\pi/2$) of these two components was determined from geometry, enabling longitude and radius components to be separated. The daily variations of satellite latitude were then determined by subtracting the radius variations from the relevant composite measurement after filtering.

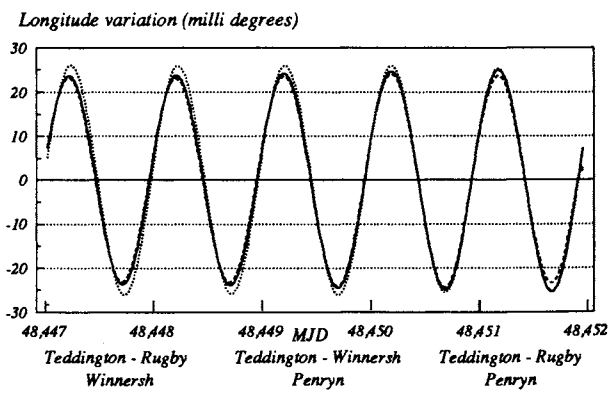


Figure 8. Sidereal day periodic variations of satellite longitude - Astra

Discrepancies were observed in the daily variations of satellite latitude and longitude when calculated from different triangles of ground stations (Figure 8). The main source of error is due to fluctuations of the underlying clock rates. The amplitude of the periodic variations observed were approximately equal to the clock fluctuations present.

3.4 Determination of the half day periodic variations of satellite longitude and latitude

Half-daily variations of longitude and latitude may be obtained by a similar method to the daily variations. The half-daily variations of orbital radius are small [6] and produce a negligible effect on the clock measurements compared with the similar longitude variations. To a first approximation, these radius variations are negligible.

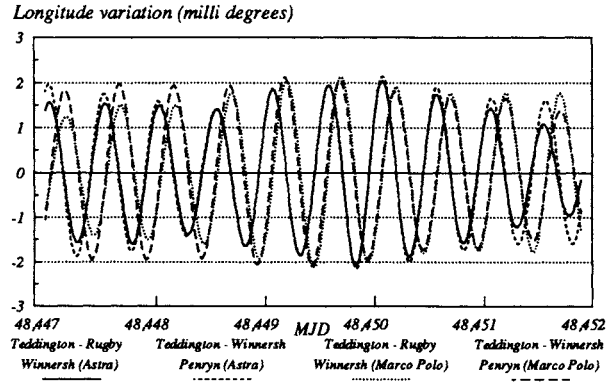


Figure 9. Half day periodic variations of satellite longitude - Astra and Marco Polo

Half-daily variations of longitude were determined from different triangles of ground stations (Figure 9). Curves for Astra and Marco Polo satellites have similar amplitudes, with a 100° phase difference between them. This is consistent with the two satellites experiencing forces of similar amplitude but at different times due to their separation in the geostationary arc. Plots of latitude variation (Figure 10) display a phase reversal on MJD 48449. This may be accounted for by relative changes of the sun and moon locations. The amplitude of the periodic variations observed were between one and two orders of magnitude less than clock fluctuations present, demonstrating the effectiveness of the digital filtering technique.

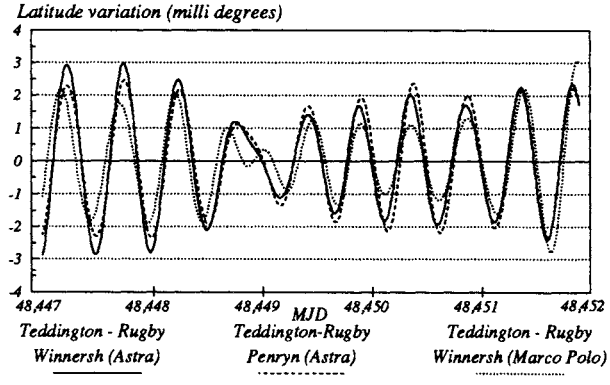


Figure 10. Half day periodic variations of satellite latitude - Astra and Marco Polo

3.5 Use of multiple composite measurements

A composite measurement was formed with contributions from all four ground stations. This measurement was independent of all clock and satellite longitude variations, and was almost insensitive to changes of satellite latitude and radius. No long term variation in the composite measurement were observed over 35 days, the short term variations of a few hours duration being due to satellite receiver noise, with a maximum variation of approximately 20ns. This suggested that for periods up to 35 days the longer term delay variations within the satellite receivers, atmosphere, and ionosphere were less than the short term receiver noise.

4. CONCLUSIONS

An analytical method to correct time measurements for the position variation of the geostationary satellite has been demonstrated. Time transfers with uncertainties of a few parts in 10^{14} have been obtained. The longitudinal drift and periodic variations of the satellite positions have also been determined.

5. ACKNOWLEDGEMENTS

The authors thank M.Hutchins (Hewlett Packard Ltd) and M.Sagin (Radiocode Clocks Ltd) for the provision of experimental facilities, P.B.Coates, B.R Swabey, R.Jefferies, D.J.E.Knight, and D.R.Allen for valuable advice given during the course of the work, P.Menary, C.Fleming, K.Roke and R.Cookman for the monitoring of equipment and collection of data. This work has been supported under the UK National Measurement System Programme for Time and Frequency.

6. REFERENCES

1. Davis J, Pearce P R, Menary P, Hutchins M, Sagin M: Time Transfer By Passive Use Of Direct Broadcasting Satellites, *Proc. Fifth EFTF*, Besancon, 12-14 March 1991, 74-81.

2. Meyer F, Brunet M, Granveaud M, Vernotte F, Vincent M: Time transfer using a geostationary direct TV satellite, *Proc. Fifth EFTF*, Besancon, 12-14 March 1991,

3. Pearce P R, Davis J A, Allen D R, Swabey B R, Jefferies R, Knight D J E, Menary P, : Time transfer in the United Kingdom using a direct broadcasting satellite, *Proc. Fourth EFTF*, Neuchatel, 13th - 15th Mar 1990, 643-650.

4. Buzek O, Cermak J, Vondrak J, Cordara F, Galliano P G, Pettiti V, Tavella P, : Synchronization of time scales by television method using ECS satellites, preliminary results, *Proc. Third EFTF*, Besancon 1989, 204-214.

5. Cordara F, Pettiti V, Cenci A, Fermi M, Sciarretta C : Time and Frequency comparisons in Europe by means of ECS 5 geostationary satellite, *Proc. 22nd PTTI*, Vienna, Virginia 1990.

6. Soop M E, : *Introduction to Geostationary Orbits*, ESA SP-1053, European Space Agency 8-10 rue Mario-Nikis, 75738 PARIS CEDEX 15, France Nov. 1983.

7. Press W H, Flannery B P, Teukolsky S A, Vetterling W T , : *Numerical Recipes, The Art of Scientific Computing*, Cambridge University Press 1986 436-443

8. European collaborative experiment being organised by Professor P.Paquet of the Royal Observatory of Belgium.

Session 6A

Quartz: Filters and Sensors

Chairman: B. Neubig

ELASTIC PLATE MODE SENSITIVITIES TO MASS LOADING APPLICATIONS TO GAS SENSORS

M. Hoummady, D. Hauden

Laboratoire de Physique et Métrologie des Oscillateurs du C.N.R.S.
associé à l'Université de Franche-Comté-Besançon
32 avenue de l'Observatoire - 25000 Besançon - France

ABSTRACT

Elastic waves are generated and detected by means of IDT transducers plated on piezoelectric films or substrates. It has been shown that both surface and bulk elastic waves can be launched by IDT's. The type of elastic wave excited is related to the piezoelectric material properties and its crystallographic orientation. In a thin piezoelectric plate, surface and bulk vibrations propagate in the volume of the plate. Thus, propagating plate modes are stationary in the plate thickness. Two kinds of modes can be distinguished : Shear Horizontal modes (S.H.) and Lamb modes. To adapt a plate mode device to a gas sensor, it is necessary to deposite a selectively-sensitive film which can absorb the gas to be detected. Then, the velocity of the wave is modified because the induced variation of the mass loading of the film changes the boundary conditions. The theoretical study of the mass loading sensitivity of both SH and Lamb modes is presented in this paper. For instance, expected sensitivity for the SH mode of a gas sensor is in the order of 1 ng of the gas for 1 cm^2 of sensitive film.

INTRODUCTION

Elastic waves are used in sensors to measure mass loading variations both in quartz crystal balances resonating in shear thickness mode (Ref. 1) or in SAW gas sensors where Rayleigh wave velocity variations depend on the gas mass absorbed (Ref. 2).

In this paper we propose to analyze theoretically the possibility of measuring mass loading variations by means of bulk wave modes generated with interdigital transducers (IDT).

First, different mode excitations are analyzed with regard to the crystallographic orientation, the plate thickness and the IDT position in the crystal axis.

Either Shear Horizontal (SH) waves in thin plates propagating as guided waves, or Lamb waves can be sensitive to mass loading.

The simulation of the gas-mass-absorbed sensitivities of SH and Lamb waves is made taking into account the dispersion properties and the geometry of the devices.

Finally, comparisons of sensitivity are presented between the different waves and modes with SAW experimental results obtained with an NO_2-sensor. It is shown that these plate waves are more sensitive than Rayleigh waves, and the resolution is better with the same frequency stability of the device (10^{-8}/mn)

1. ELASTIC WAVES GENERATED BY INTERDIGITAL TRANSDUCERS

In piezoelectric materials, electric fields applied by interdigital transducers (IDT's) induced mechanical displacements within the crystal. Different elastic waves at different polarizations are generated : surface acoustic waves (SAW) or bulk acoustic waves (BAW).

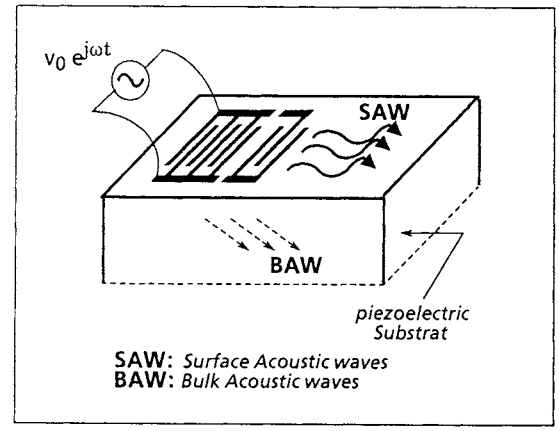


Fig. 1 : *Elastic waves generated by IDT*

The excitation of one particular elastic wave depends on the crystal symmetries and on the position of the IDT with respect to the crystallographic axis. The IDT position defines the electric field components and the crystal symmetries impose the non-zero piezoelectric constants. Both lead to the stress-tensor which defines the polarization of the wave.

1.1 Quartz-crystal propagations

When an IDT is metallized on the surface of a quartz crystal plate, different elastic waves are generated following its position with regard to the crystallographic X-axis (parallel or perpendicular).

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

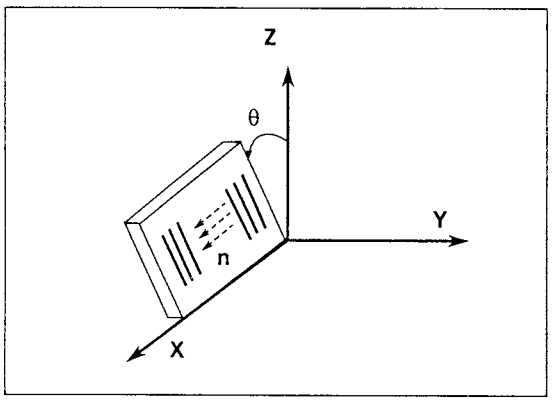


Fig. 2 : *IDT perpendicular to X-axis*

If the IDT is perpendicular to the X-axis, electric field components E_i are (E_1, E_2, E_3) ; then the piezoelectric tensor of the quartz crystal demands that the non-zero stress components are :
$T_{11} = e_{11} E_1$, $T_{22} = e_{12} E_1$, $T_{23} = e_{14} E_1$, $T_{13} = e_{25} E_2$ and $T_{12} = e_{26} E_2$.

Combining the effects of T_{11}, T_{13}, T_{12}, a surface wave is generated at the surface of the quartz plate which propagates in the X-direction with an elliptic polarization ; it is a Rayleigh wave.

Stress components T_{22} and T_{23} produce bulk elastic waves.

When an IDT is deposited parallel to the X-axis, only the electric field components E_2 and E_3 are non-zero. The induced stress components are T_{13} and T_{12}. They result in bulk waves with a rectilinear polarization parallel to the X-axis, with 2 displacement components along the Y and Z axes (horizontal polarization).

1.2 Propagation in a thin quartz plate

The different waves generated by an IDT must propagate also in a thin plate. Then they are able to propagate as plate modes with a particular velocity distribution. Two kinds of plate modes can be distinguished according to their polarizations : shear horizontal waves (SH), and Lamb waves.

2. PLATE MODE PROPERTIES

In thin plates of crystalline quartz, SH waves propagate as guided waves. They have the properties of shear waves with in-plane transverse polarization. They are generated when the IDT is parallel to the X-axis (propagation is in the Z-direction).

Lamb waves are generated when the IDT is perpendicular to the X-axis. They are generated with different frequencies from the electric field signal.

Their polarization is elliptic because it results from the combination of a longitudinal and a transverse mode. Two types of Lamb waves propagate in a thin plate of quartz : symmetric modes and antisymmetric modes (that is also the flexure vibration mode).

Velocities of SH-waves and Lamb waves depend on the ratio of thickness to wavelength. They are dispersive waves. Figure 3 presents the dispersion curves of the two first SH-mode of a Y-cut quartz crystal when they are propagating in the Z-direction.

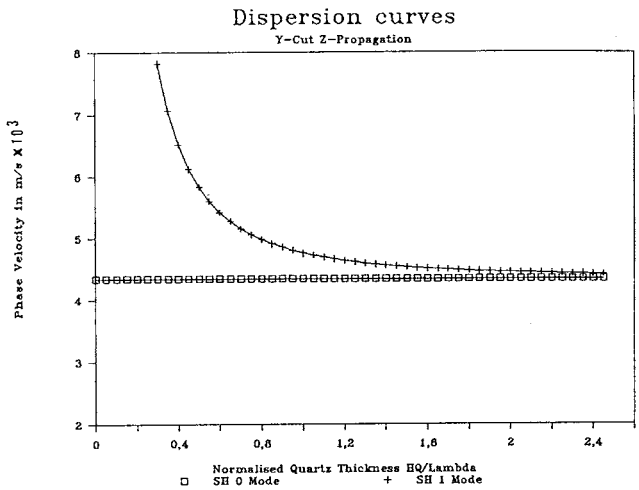


Fig. 3 : *Dispersion curves of SH-waves*

Dispersion curves of A_0-antisymmetric Lamb waves and S_0-Symmetric Lamb waves are plotted in Figure 4. A_0-Lamb waves have a very low velocity for very thin quartz crystal plates.

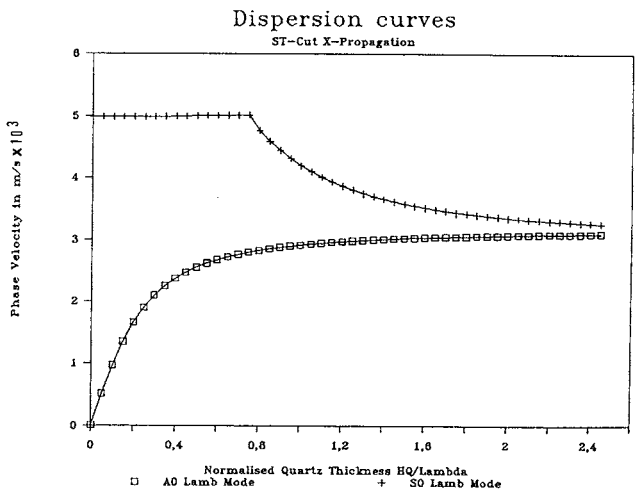


Fig. 4 : *Dispersion curves of Lamb waves A_0 and S_0*

3. MASS LOADING SENSITIVITY

For gas or liquid analysis, specific sensors can be prepared from a quartz device vibrating in plate modes, exhibiting either SH-waves or Lamb waves. In both cases, a selective, sensitive thin film deposited at one surface of the plate will absorb a specific species contained in the analyzed fluid. This leads to a variation in the mass loading of the thin film and then to a consequent change in the velocity of the wave (Ref. 3).

Figure 5 summarizes the principle of the sensor.

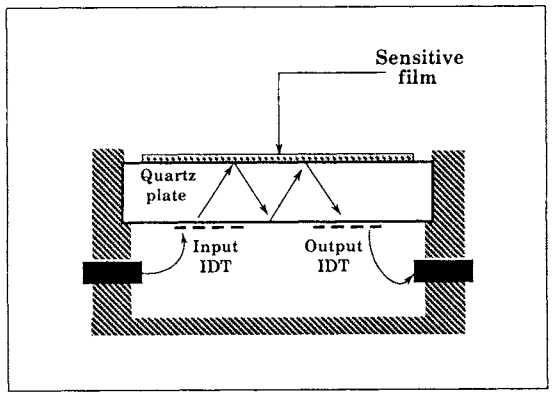


Fig. 5 :
Principle of the gas (or liquid) sensor using plate mode

Theoretical results obtained in the case of NO_2 gas detection diluted in dry nitrogen gas are presented. The selective film is a sputtered Copper phthalocyanine film of 1 000 Å thickness.

Previously, mechanical properties of the thin film were determined experimentally using surface acoustic wave propagation coupled with a perturbation method (Ref. 4). By introducing Young's modulus, Poisson coefficients and specific mass (ρ'), in a numerical calculation for a multilayer structure where SH waves and Lamb waves are propagating, it is possible to calculate the wave velocities before gas absorption. This calculation is based on a matrix method to solve the propagation equation and successive interface boundary conditions taking into account the quartz and thin film constants (Ref. 5).

4. MASS LOADING SENSITIVITIES OF SH-MODES

For the two first SH-modes, sensitivities are calculated in terms of relative variations of velocity versus specific mass variations ($\Delta v/\Delta \rho'$). The hypothesis made is that the gas absorbed does not increase the volume of the film.

Figure 6 presents the theoretical sensitivity of the SH_0-mode for successively-increasing specific masses of 5 %, 10 % and 15 % versus the film thickness to wavelength ratio (H'/λ).

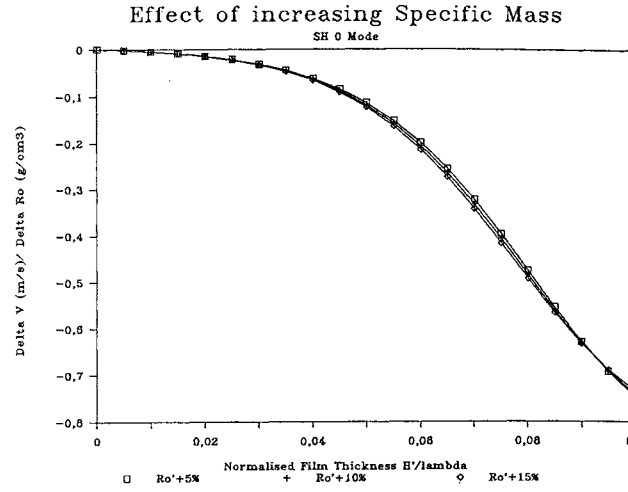


Fig. 6 : *Mass loading sensitivity of the SH_0-mode*

Figure 7 presents the same results for the SH_1 mode.

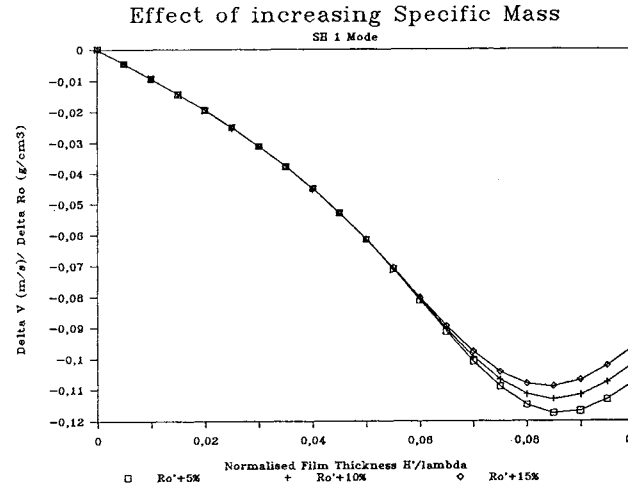


Fig. 7 : *Mass loading sensitivity of the SH1-mode*

In both cases, sensitivities to mass loading depend on the thickness of the film and on the kind of SH-mode. One can see also that for very thin films (small thickness), the sensitivities are linearly dependent on the mass loading. On the contrary, when the thickness of the film is larger than 6 % of the wavelength, its stiffness becomes important and its effect is opposite in sign to the mass loading effect (Ref. 6).

Expressed in term of relative velocity variation, one can write :

$$\frac{\Delta v}{v_o} \simeq A f H' \rho' - B f H'$$

where $AfH'\rho'$ represents the mass loading and BfH' the stiffness. These two terms reinforce one another at lower values of H'/λ but at higher values, the second is preponderant, producing the maximum as shown in Fig. 7.

A detailed analysis of the case of very thin films of Copper Phthalocyanine absorbing NO_2 gas shows that the sensitivity is strongly dependent on the SH-mode, which increases with wave order. Results are presented in Fig. 8.

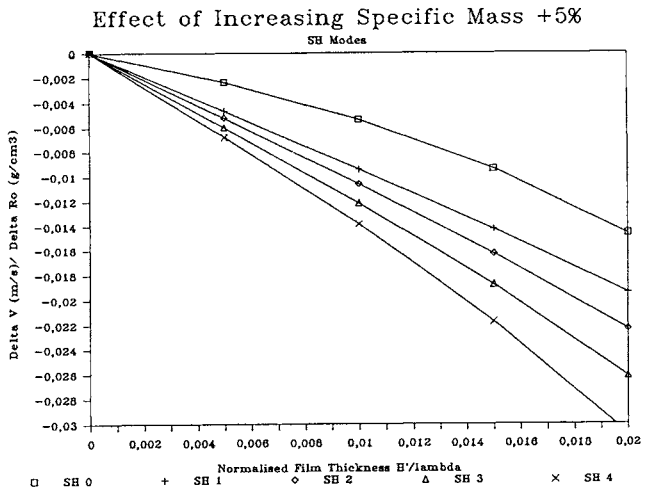


Fig. 8 : *SH-mode order influence on the mass loading sensitivity*

However, it is important to come back to the relative velocity variation (or frequency shift) versus gas mass absorbed by surface unit ($H' \Delta\rho'$) to determine the minimum mass of gas detectable. If the analysis is based on experimental results, both the mass loading effect and the film thickness variation are combined in the data.

Figure 9 presents the relative frequency shift as a function of the mass per surface unit, and one can observe the strong effect of the SH-mode on the linearity.

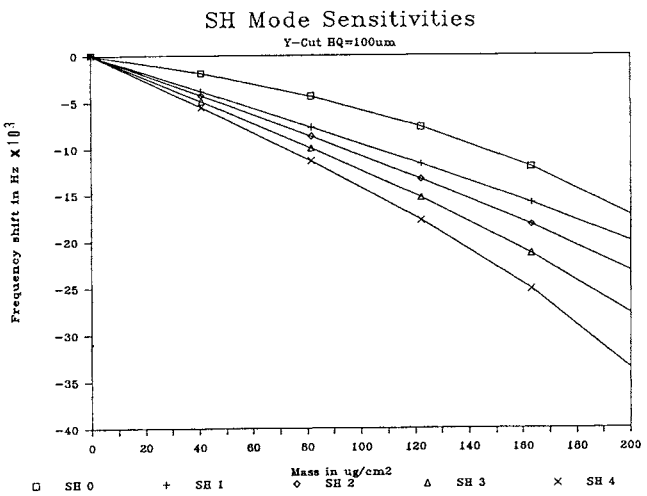


Fig. 9 : *SH-mode sensitivities in terms of relative frequency shift versus gas mass in µg/cm²*

The minimum gas detectable is determined by the frequency stability (over the measuring time of a few seconds) and by the sensitivity. Also, for an average stability, (around 10^{-8} for a measurement time of one minute), the detectable mass variation is about 1 nanogram for the SH_0 mode.

5. MASS LOADING SENSITIVITIES OF LAMB MODES

Lamb waves also present a high sensitivity to mass loading. Theoretical results are plotted for the lowest modes A_0 and S_0 (antisymmetrical and symmetrical). Figure 10 shows the sensitivity in terms of the ratio (velocity to specific mass variations) versus the normalized thickness (H'/λ).

The symmetric mode seems more sensitive to gas absorption than the antisymmetrical mode.

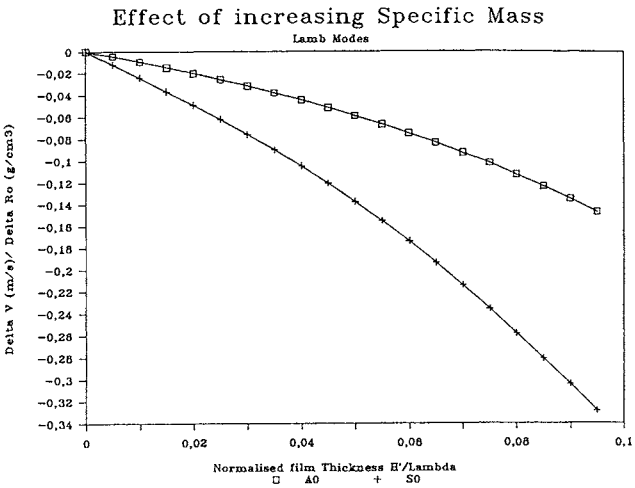


Fig. 10 : *Relative velocity variation of A_0 and S_0 Lamb waves versus the normalized thickness*

6. COMPARISON OF SENSITIVITIES TO MASS LOADING OF LAMB AND SH WAVES

A simulation of the sensitivities calculated under the same conditions is made for Lamb and SH waves. Plate thickness is equal to 100 µm and the mass per surface unit varies between zero and 800 µg/cm². The respective frequencies are 29 MHz for the A_0-Lamb wave, 42 MHz for the S_0-Lamb wave and between 43 and 71 MHz for the SH-modes.

Figure 11 shows the very linear relative frequency shifts as a function of the gas mass absorbed for low quantities (below 300 µg/cm²). In this range, Lamb modes are several times more sensitive than SH modes. This is generally the case for weak concentration of NO_2 in air. But in the case of high gas concentrations, SH-modes can be more sensitive than Lamb modes.

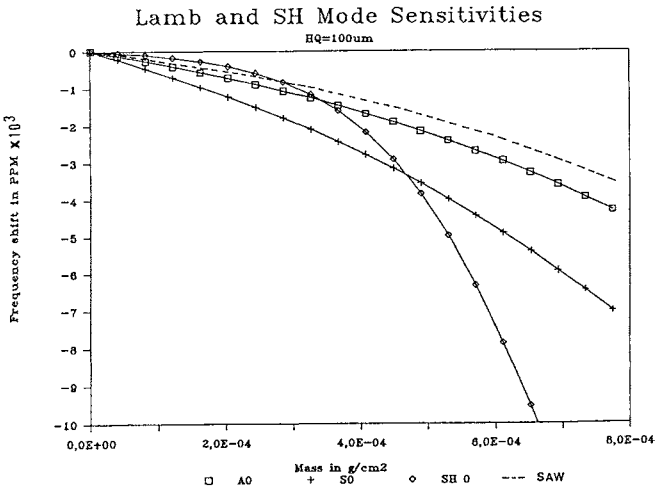


Fig. 11 : *Sensitivity comparison of SH and Lamb waves*

Experimental values were obtained previously for a SAW device which exhibits a sensitivity of 4 ppm of NO_2 gas concentration per Hz at 100 MHz (Ref. 7). Using SH modes or Lamb modes can increase the sensitivity by an order of magnitude even for the same frequency stability of the Rayleigh wave devices.

In this table, comparison between the different waves is made for two quartz plate thicknesses (H_Q) and three normalized sensitive film thicknesses (H').

In addition, as opposed to the SAW device which needs a selective film on the propagating surface, requiring delicate techniques for plating, Lamb and SH waves have a surface free of IDT which can be used to accommodate the selective film.

These theoretical results are applicable whatever the kind of gas if the mechanical properties of the film are first determined by regular acoustic methods. Generally today, selective films are organometallic or polymer materials where stiffness properties are very different but are weak, and our model will be well-adapted to calculating their properties.

REFERENCES

1. C. Lu and A.W. Czauderna 1984, Applications of piezoelectric quartz crystal microbalances, *Methods and Phenomena*, vol. 7, Elsevier, Amsterdam.

2. H. Wohltjen 1984, *Sensors and Actuators*, 5, 307.

3. M. Hoummady, D. Rebière, D. Hauden, J. Pistré, P. Cunin, R. Planade 1991, *Proc. 5th European Frequency and Time Forum*, Besançon, France.

4. M. Hoummady, D. Rebière, D. Hauden, J. Pistré, P. Cunin, R. Planade 1990, *3th Int. Meeting on Chemical Sensors*, Cleveland, USA.

5. E.L. Adler, G.W. Farnell, J. Slaboszewicz and C.K. Jen 1988, *Proc. IEEE Ultrasonic Symp.* n° 88CH2578-3, Chicago, USA.

6. B.A. Auld 1973, *Acoustic fields and waves in solids*, John Wiley & Sons, New York, vol. II, chap. 12.

7. D. Rebière, M. Hoummady, D. Hauden, J.L. Aucouturier, J. Pistré, P. Cunin, R. Planade, 1991, *Transducers'91, International conference on Solid-State*, San Francisco, USA.

Modes	Frequency in MHz		Sensitivities in ng / cm² Stability (S = 10^{-8} / mn)					
			$H_Q = \lambda_o = 100$ µm			$H_Q = \lambda_o = 50$ µm		
	$H_Q = \lambda_o = 100$ µm	$H_Q = \lambda_o = 50$ µm	$H'/\lambda_o = 2\%$	$H'/\lambda_o = 6\%$	$H'/\lambda_o = 8\%$	$H'/\lambda_o = 2\%$	$H'/\lambda_o = 6\%$	$H'/\lambda_o = 8\%$
SH_o	43.38	86.76	6	1.5	0.5	3	0.75	0.25
SH_1	48.92	97.84	6	3.5	3	2	1.75	1.5
SH_2	55.07	110.14	5.0	2.5	2*	2.5	1.25	1*
SH_3	62.67	127.34	5.0	2	2*	2.5	1	1*
SH_4	71.26	142.52	4.5	1.2	2*	2.25	0.6	1*
A_o	29.12	58.24	3	2.5	2	1.5	1.25	1
S_o	42.87	85.74	2	1.5	1	1	0.75	0.5
	$\lambda_o = 100$ µm	$\lambda_o = 50$ µm	$\lambda_o = 100$ µm			$\lambda_o = 50$ µm		
			$H'/\lambda_o = 2\%$	$H'/\lambda_o = 6\%$	$H'/\lambda_o = 8\%$	$H'/\lambda_o = 2\%$	$H'/\lambda_o = 6\%$	$H'/\lambda_o = 8\%$
Rayleigh (quartz thick plate)	31.57	63.14	5	4	3	2.5	2	1.5

H_Q : quartz plate thickness
H' : sensitive film thickness
λ_o : acoustic wave length
* Effect of stiffness limitation

Table 1 : *Comparison of wave sensitivities to mass loading variations as a function of film thickness and quartz plate thickness*

A STATE-OF-THE-ART CRYSTAL FILTER DESIGN PACKAGE USING WINDOWS 3.0 TECHNIQUES

Markus Frey and Bernd W. Neubig

TELE QUARZ GROUP
Landstrasse, D-6924 Neckarbischofsheim, Germany

1. Abstract

Tele Quarz WINFIL is a company-owned software package for professional crystal filter engineering, which enables the competent user to design crystal filters systematically for any sensible requirement. To achieve the high development standard necessary to realize also rather difficult filter specifications, TQ WINFIL offers access to several levels of the filter synthesis, i.e. the pole-zero plane, the normalized lowpass domain and the synthesis of the final crystal filter circuit. At each synthesis level an analysis module as well as an optimization module are included. This allows the user to analyze the theoretical models as well as the real crystal filter circuit including losses, component tolerances and non-ideal characteristics such as spurious responses of the crystals. The optimization modules make it possible to search iteratively for a network solution from a starting set of values in order to meet special requirements. Frequency-unsymmetric responses can be computed by a frequency-transformation module, which effects the bi-linear transformation of all circuit components of the filter. Any parameter changes in the synthesis and any action of the optimization module not only cause changes of the desired response, they also influence the other characteristics of the filter. The WINDOWS graphical user interface (GUI) allows the simultaneous display of all interesting parameters and results during the whole cycle of synthesis - analysis - optimization -transfomation.
To put all these features together, including the complex mathematic routines for fast synthesis, analysis and optimization, the comfortable user interface etc., a new programming technique was used. The software package TQWINFIL is built with the programming tools of "Microsoft QuickC for Windows", which allows the direct computer-assisted development of WINDOWS applications.TQWINFIL can be run simultaneously together with other WINDOWS applications , e.g. the spread-sheet program Microsoft EXCEL which can be used for further evaluation of the analysis data either in tables or graphically.

2. TQWINFIL Building blocks

2.1 Survey

The complete software package consists of three groups:
(1) synthesis, analysis and optimization in the pole-zero plane of the transmission function
(2) synthesis, analysis and optimization in the normalized lowpass domain
(3) synthesis, analysis and optimization of crystal filter circuits

Each group contains
an **input module for synthesis** to input directly design parameters, such as
poles and zeros of the transmission function,
normalized lowpass element values, filter design data (nominal frequency, bandwidth etc.),
or alternatively
the desired standard transmission function, which may be
(a) high selectivity characteristics e.g. Butterworth, Chebyshev, Inverse Chebyshev, Cauer (i.e. Elliptical) etc.
(b) linear phase characteristics, e.g. Gauss, Bessel, Ulbrich-Piloty, Feistel-Unbehauen etc.
(c) transitional, i.e. selective linear phase characteristics, e.g. 6dB-Gaussian, 12dB-Gaussian, Rakovich etc.
(d) characteristics with prescribed time domain response, e.g Dirac, Optimum response, Stehle etc.
(e) special functions collected from literature and experience,
and
the filter specification defined by
(a) selectivity, which may be described by
centre frequency,
pass bandwidth at x dB,
pass band ripple,
max. transducer attenuation
stop bandwidth at y_1, y_2 ,... dB,
ultimate attenuation etc.,
(b) phase linearity demands,
(c) group delay constraints,
(d) time domain requirements
(e) others

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

a **synthesis module** to compute for a given characteristics
- the poles and zeros of the transmission function,
- the lowpass element values,
- the crystal filter element values.

Furthermore it allows
- to synthesize the lowpass elements from a set of the poles and zeros,
- to calculate the crystal filter circuit elements from an arbitrary set of lowpass elements.

an **input module for analysis and optimization**, with which circuit data such as poles and zeros, lowpass element values, or filter circuit elements,can be inserted from the appropriate synthesis resp. optimization program, or manually

an **analysis module** to analyze
- in the complex frequency domain (pole-zero analysis),
- in the normalized frequency domain (normalized lowpass analysis)
- in the "real" frequency domain (crystal filter analysis)

or
- in the normalized time domain (pole-zero and normalized lowpass analysis)
- in the "real" time domain (crystal filter analysis).

a **transformation module** to perform a bilinear frequency transformation in order to derive a crystal filter circuit with a non-frequency- symmetric response.

an **optimization module** which allows to vary a starting set-up of values of
- pole-zero locations, or
- lowpass element values, or
- crystal filter element values

in order to fulfill the requirements of the given filter specification in an optimum manner.

2.2 Program group "Pole-zero plane"

The transmission function H(p) of a two-port is defined as

$$H(p) = \frac{U_0(p)}{2*U_2(p)}\sqrt{\frac{R_1}{R_2}}$$

which is related to the characteristic function K(p) by

$K(p)*K(-p)=H(p)*H(-p) - 1$

For the filters treated here, the transmission function H(p) and the characteristic function K(p) are real rational functions of the complex frequency variable p= s + jW, and can be written as a quotient of two polynomials:

$$H(p) = C_H * \frac{E(p)}{G(p)}$$

and

$$K(p) = C_K * \frac{F(p)}{G(p)}$$

where E(p) is a Hurwitz polynomial and F(p) and G(p) are even or odd polynomials; C_H and C_K are constants.

The roots of E(p) are the poles, the roots of G(p) are the zeros of transmission.
The frequency and time responses of a filter can be derived from H(p) for p=jΩ . E.g.:
effective attenuation a_B:

$$a_B(\Omega) = 20 * \lg |H(j\Omega)|$$

phase angle bB:

$$b_B(\Omega) = \arg \{H(j\Omega)\}$$

normalized group delay t_g:

$$t_g(\Omega) = \frac{db_B(\Omega)}{d\Omega}.$$

as well as the filter characteristics in the time domain.
The program group "pole-zero plane" computes all transmission characteristics from the location of the poles and zeros of H(p) in the complex p-plane (analysis part). The locations can be calculated analytically for some standard characteristics such as Butterworth, Chebyshev, inverse Chebysheff, Cauer ("elliptical filters") or can be inputted arbitrarily (synthesis part).
The optimization part varies the coordinates of the poles and zeros from a starting set of values until a given target specification is met or approached in a minimum error sense. Details about the algorithm used are described in the program group "crystal filter circuits", see clause 2.4.5.
Optimization moves in the pole-zero plane are easy and fast to perform, however there is no control that the poles and zeros found will result in a realizable lowpass or crystal filter circuit.

Figure 1 shows
- the options to choose standard characteristics and
- the belonging window to enter the effective parameter, in order to synthesize the poles and zeros,
- the synthesized poles and zeros,
- the alternative option to enter the poles and zeros arbitrarily,
- the menu item for optimization,
- the menu item for analysis and the resulting analysis datas in tables and graphics, of the group "Pole-zero plane".

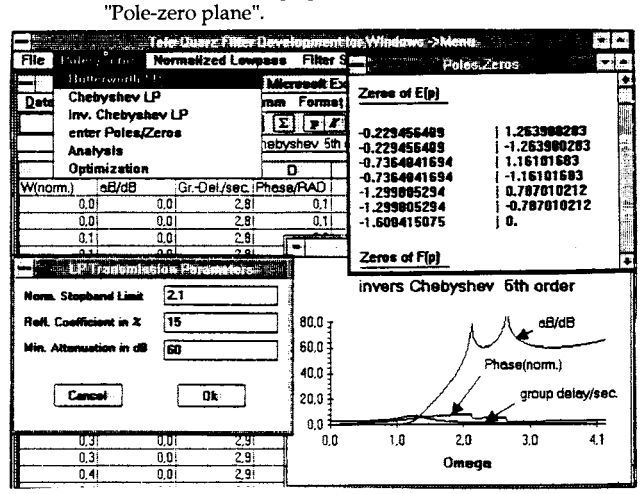


Figure 1: Operations in the "Pole-zero plane"

2.3 Program group "Normalized lowpass"

Because crystal filters in general have very narrow relative bandwidths, the elementary filter synthesis directly as band-pass or band-stop filter would lead to extremely high computional errors. Therefore modern crystal filter synthesis is based on a suitable normalized lowpass, from which the final characteristic is derived by a couple of transformations.
The normalized lowpass can be synthesized from the polynomials E(p), F(p) and G(p) of the transmission function and the characteristic function. The technique therefore is well established. To realize a ladder lowpass network, a series of successive complete and partial-pole removals from the open- and short-circuit input and output reactances is performed. The procedure yields a set of multiple solutions, whose number depends on the filter order. All the solutions with positive lowpass element values are selected. The crystal filter synthesis finally shows, which one of these sets yields the most convinient element values in the final circuit, especially regarding the realizability of the crystal resonators.

The program group "normalized lowpass" synthesizes the normalized lowpass circuit elements from the poles and zeros of H(p) and K(p). Alternatively the element values can be calculated analytically for some standard characteristics such as Butterworth, Chebyshev, inverse Chebyshev, Cauer ("elliptical filters") or can be inputted arbitrarily. The analysis part allows a fast network analysis for a given set of LC-lowpass elements.
The optimization part varies the element values from a starting set of values until a given target specification is met or approached in a minimum error sense. Details see below (clause 2.4.5). Optimization runs in the lowpass domain are fairly fast and simple, but there is no guarantee, that the found set of lowpass elements is a "minimum phase" or a "minimum insertion loss" solution.

Figure 2 displays
- the option to synthesize the lowpass elements using the poles and zeros of the "Pole-zero plane",
- the calculated lowpass elements normalized,
- the alternative option to enter the normalized lowpass elements arbitrarily,
- the menu item for optimization,
- the menu item for analysis and the resulting analysis datas in tables, of the group "Normalized lowpass domain".

Any evaluations with EXCEL are also possible in this group.

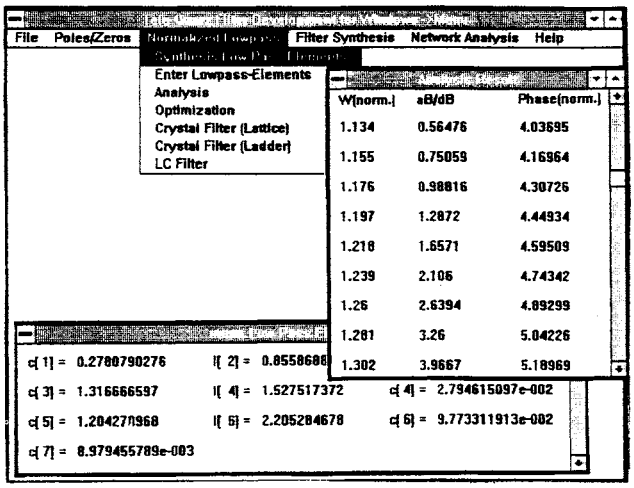


Figure 2: Operations in the "Normalized lowpass domain"

2.4 Program group "crystal filter circuits"

A large variety of different crystal filter circuits can be synthesized, analyzed and optimized in this program group.

2.4.1 Crystal filter synthesis

Actually the synthesis covers the following types of crystal filter circuits:

2.4.1.1 BANDPASS FILTERS

synthesis based on the effective parameter method

2.4.1.1.1 lattice filters, narrow and intermediate bandwidth,"all pole", i.e. all attenuation poles at infinity)
2.4.1.1.2 lattice filters, narrow and intermediate bandwidth, finite attenuation poles (e.g. Caue resp. "elliptical" filters and others)
2.4.1.1.3 lattice filters, wide bandwidth, "all pole"
2.4.1.1.4 ladder filters, narrow and intermediate bandwidth (LSB and USB)
2.4.1.1.5 monolithic and polylithic filters, narrow and intermediate bandwidth, "all pole"
2.4.1.1.6 special filter structures, e.g. "Lerner" filters

synthesis based on the image parameter method

2.4.1.1.11 lattice filters, narrow, intermediate, and wide bandwidth
2.4.1.1.12 ladder filters, narrow and intermediate bandwidth

2.4.1.2 BANDSTOP FILTERS

synthesis based on the effective parameter method

2.4.1.2.1 lattice filters and equivalent circuits, e.g. bridged-T
2.4.1.2.2 lowpass ladder filters with crystals in the parallel branch
2.4.1.2.3 special filter structures

synthesis based on the image parameter method

2.4.1.2.11 lattice filters and equivalent circuits, e.g. bridged-T, narrow, intermediate and wide bandwidth

2.4.1.3 ALL PASS FILTERS, GROUP DELAY EQUALIZERS
2.4.1.3.1 lattice filters, effective parameter method
2.4.1.3.2 lattice filters, image parameter method

It is impossible to describe all these synthesis programs in this presentation. Therefore only a few of them will be described shortly:

2.4.1.1.1 LATTICE FILTERS, NARROW AND INTERMEDIATE BANDWIDTH,"ALL POLE"
This is the most widely used synthesis program, as the (half)lattice allows the most flexible design regarding realizability. The filter consists of at least one half lattice section, where usually each branch contains one crystal. The half lattices are usually realized by tuned hybrid transformers
If several sections are cascaded, a coupling network is inserted in between them, which at least consists of one capacitor C_k. In the intermediate bandwidth design, the (negative) coupling capacitor is realized by a parallel resonant circuit L_k, C_k (see Fig.3).
The synthesis is derived from the normalized lowpass by a sequence of transformations as follows (see Fig.4) :

- → perfom a lowpass-to-bandpass transformation: the Ls convert to series resonant circuits, the Cs convert to parallel resonatn circuits. Choose a suitable termination impedance level for realizable motional parameters for the crystal elements.
- → transform the parallel resonance circuits to series resonant circuits by insertion of impedance inverters $-C_T$,C_T,$-C_T$ at both sides.
- → absorb the negative capacitors into the adjacent series or parallel capacitors.
- → apply the Bartlett bi-section theorem, which transforms symmetric P- or T-networks to a lattice.
- → remove admittance/impedance parts, which are common inboth lattice branches, until the branches contain realizable crystal elements by the extraction method

The synthesis can be performed with the constraint of equal motional parameters for all crystals in the cascade, or with unequal motionals, however a simpler filter circuit.

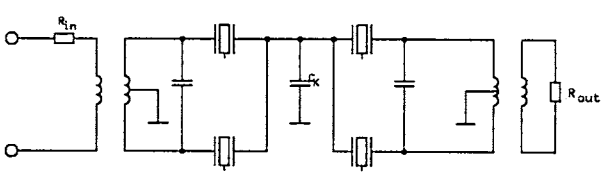


Figure 3: two cascaded half-lattice crystal sections

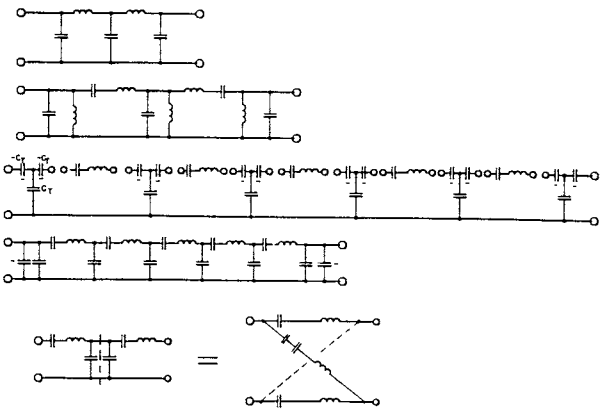


Figure 4: transformation sequence from the normalized lowpass to a crystal lattice

2.4.1.1.2 LATTICE FILTERS, NARROW AND INTERMEDIATE BANDWIDTH, FINITE ATTENUATION POLES (E.G. CAUER RESP. "ELLIPTICAL" FILTERS)
The synthesis is based on transformations from the normalized lowpass prototype by using the "Holt & Gray transformation. It results in the same basic circuit as the all-pole half lattice (Fig.3), but the motional and the parallel capacitances are not identical in both branches of the lattice.

2.4.1.1.3 LATTICE FILTERS, WIDE BANDWIDTH, "ALL POLE"
The broadband design realizes two "poles", i.e. two crystal resonators, in a parallel tuned circuit parallel to the crystal. Applying a simple circuit transformation moves this circuit to the input and output port of the half lattice.
The basic design is a four-pole filter using two crystals (one in each branch). The six-, eight-pole etc. filter has two, three etc. crystals in each lattice arm. The whole filter has only one lattice section (see Fig. 5 for an eight-pole filter). Alternatively an eight-pole can be realized by a cascade of two sections with totally four crystals, a twelve-pole consists of 3 cascaded sections etc.

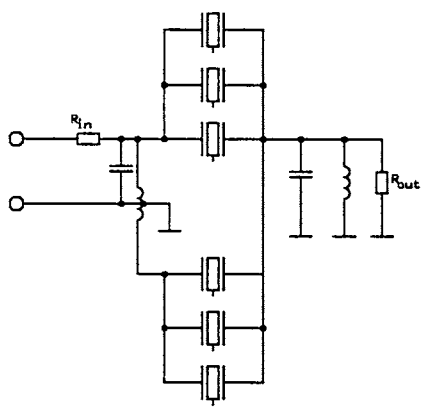


Figure 5: 8-pole broadband crystal filter

2.4.1.1.4 LADDER FILTERS, NARROW AND INTERMEDIATE BANDWIDTH (LSB AND USB)

Crystal filters in ladder configuration allow only a very limited design freedom. No allpole characteristics can be realized due to the inevitable two resonances [(series)-resonance and anti-resonance)] of crystal resonators. Usual applications are very narrowband filters for single sideband (SSB) applications.

With the use of bi-linear transformation (see below) and the concept of pseudo-reactances, lower-sideband (LSB-) filters can be synthesized with the crystals in the series arms of a ladder network. The multiple attenuation pole (transformed from infinity) is then the antiresonance frequency of the crystal (plus parallel capacitor).

Accordingly upper-sideband (USB)-filters can be made with the crystals in the parallel arm of the ladder. The attenuation pole is given by the (series-) resonance frequency of the crystals (see Fig. 6).

With this method, the filter design can be derived from the normalized lowpass. However - due to the distortion caused by the transformation - the filter performance is different from that known from normal lowpass-to-bandpass transformations. Only for very narrow bandwidth filters the response is - to a certain extend - usual.

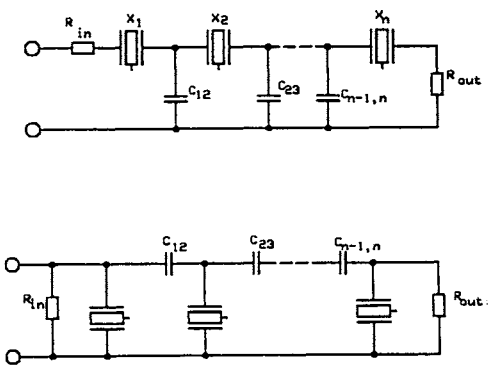


Figure 6: LSB- and USB ladder filters

2.4.1.1.5 MONOLITHIC AND POLYLITHIC FILTERS, NARROW AND INTERMEDIATE BANDWIDTH, "ALL POLE"

The design is the same as for the half lattice all pole filters, as the lowpass-to-bandpass transformation directly yields a circuit equivalent to that of monolithic resonators.

2.4.1.1.11 IMAGE PARAMETER DESIGN: LATTICE FILTERS, NARROW, INTERMEDIATE, AND WIDE BANDWIDTH

The image parameter filter design is based on explizite design formulae for a single-section half lattice.

The transmission poles and zeros can be placed arbitrarily to meet the required specification.

The program allows a rapid interactive optimization of the location of poles and zeros.

Higher selectivity is achieved by cascading multiple sections, however, the final filter performance is not exactly as predicted by image design. With the use of the circuit analysis programs described below. This disadvantage can be overcome.

The filter design by the modern effective parameter method is nowadays much simpler and faster, so that the image parameter design method is used only occassionally.

A survey of the branch elements covered by the program is depicted in Fig.7.

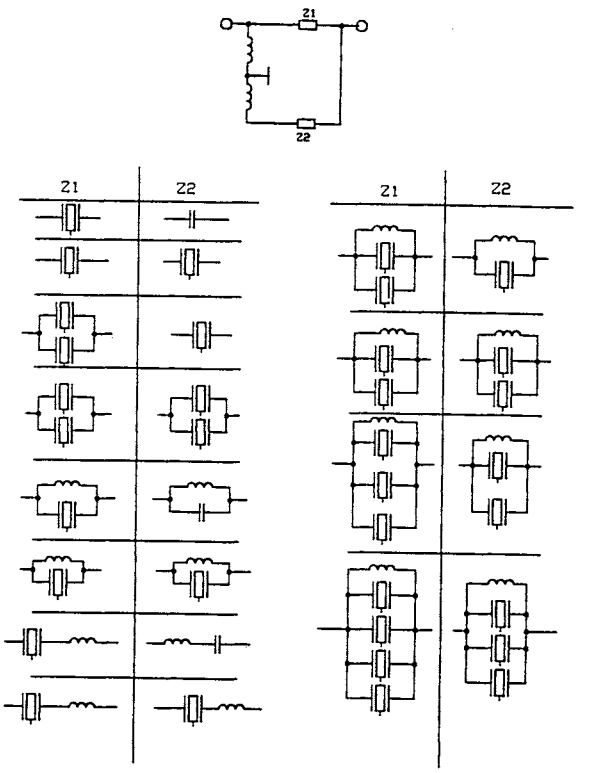


Figure 7: lattice branches of image parameter crystal filter

2.4.2 Crystal filter analysis

Special circuit analysis programs have been written in order to perform a very fas complex circuit analysis without tedious input work.

The network analysis is carried out by the use of the chain matrice $| | A_{ik} | |$

Cascading of two-ports is equivalent to multiplication of their chain matrices. Bridging element across one two-port can be considered as parallel connection of two two-ports. Thus is not necessary to describe nodes and meshes, and the computation is rather fast.

All filter circuits are assumed to consist of the following stages:

→ Input impedance transforming section
→ Input circuit
→ crystal section # 1

→ coupling section # 1
(repeated until)....
→ last crystal section
→ Output circuit
→ Output impedance transforming section

The number of coupling sections is equal the number of crystal sections minus 1.

2.4.2.1 INPUT AND OUTPUT IMPEDANCE TRANSFORMING CIRCUIT
The filter termination at the input and output is described by a resistance R_{in}, R_{out}, and a capacitance to ground C_{in},C_{out}. It can be matched to the required internal termination impedance by the folowing means:
→ impedance transformer
→ C-L- or L-C- transformation (including losses)
→ series, and/or parallel resistor
→ or combination thereof.

2.4.2.2 INPUT AND OUTPUT CIRCUIT
This is a capacitor, a lossy inductor or both connected in parallel.

2.4.2.3 CRYSTAL SECTION
The crystal section may either be a half lattice circuit, a ladder circuit, a bridged-T circuit, or a monolithic two-pole segment (dual).
For descrete (i.e. not monolithic) crystal filters each branch in a (half) lattice, in a ladder network or in a bridged-T has the following universal general structure (see Fig. 8).
The crystal(s) and their spurious modes are modeled by an arbitrary number of series resonant circuits L_i, C_i, R_i. the static capacitance(s) of the resonator(s) and possible additional parallel capacitors are summarized in C_{par}. An inductor L_{par} with Q-factor Q_{par} can be connected in parallel to C_{par}. In series to the whole "resonator block" may be a series capacitor C_{ser}, or a series resonant circuit C_{ser}, L_{ser} with Q-factor Q_{ser}. Every component - including the crystal(s) may be present or not. During the computation the overall branch impedance Z_{br} is determined at the considered frequency.

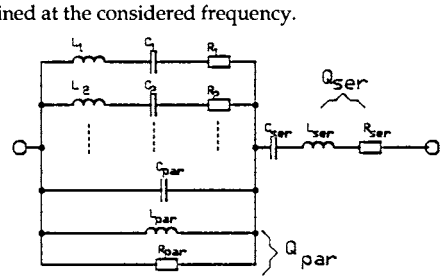


Figure 8: universal general structure of one branch in a lattice or ladder filter

2.4.2.3.1 HALF LATTICE CIRCUIT
The impedances of the two adjacent branches Z_{br1} and Z_{br2} are included in the chin matrix for a half lattice stage:

$$a_{11} = a_{22} = \frac{Z_{br1}+Z_{br2}}{Z_{br2}-Z_{br1}}$$

$$a_{12} = \frac{Z_{br1}*Z_{br2}}{Z_{br2}-Z_{br1}}$$

$$a_{21} = \frac{4}{Z_{br2}-Z_{br1}}$$

2.4.2.3.2 LADDER CIRCUIT
The chain matrix of a series element Z_{br} of a ladder circuit is:

$a_{11} = a_{22} = 1$
$a_{12} = Z_{br}$
$a_{21} = 0$

An element to groud is given by

$a_{11} = a_{22} = 1$
$a_{12} = 0$

$$a_{21} = \frac{1}{Z_{br}}$$

2.4.2.3.3 BRIDGED-T CIRCUIT
The bridged-T section has a matrix similar to that of the half lattice, only the branch impedances have to divided resp. multiplied by two.

2.4.2.3.4 MONOLITHIC AND POLYLITHIC FILTER CIRCUIT
The crystal section is a monolithic two-pole resonator ("bilithic" or "dual"), whose equivalent circuit is shown in Fig. 9. The two resonator halfs are represented by series resonant circuit L_j,C_j,R_j. Their mechanical coupling factor k is described by an "impedance inverter" -L_k,L_k,L_k, where - for a symmetric device -

$L_k = k*L_j$

At the ports of the device are the static capacitances C_0.
Alternatively the dual resonator can also be characterized by the two characteristic frequencies f_{sym} an f_{asym}, which are connected with the coupling factor k by

$$k = \frac{(f_{asym}-f_{sym})^{*2}}{f_{asym}+f_{sym}}$$

2.4.2.4 COUPLING SECTIONS
The coupling section between two adjacent crystal section may be
→ a coupling capacitor (to ground)
→ a tuned lossy parallel resonant circuit L_kC_k with Q factor Q_k
→ a P-arrangement of capacitors
→ a P-arrangement of two parallel tuned circuits with a series capacitor,
→ a resistive P- or T-network.
The matrix elements can be derived elementarily.

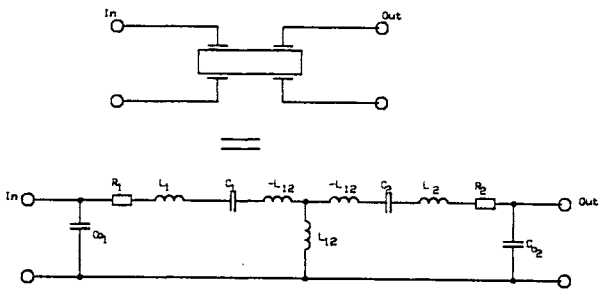


Figure 9: equivalent electrical diagram
of a monolithic twopole resonator (dual)

2.4.2.5 OUTPUT FORMATS

All analysis programs have the same output formats, which are

→ frequency, attenuation a_B, phase b_B and group delay t_g
as standard output,

→ input and output impedance/admittance, reflection VSWR as optional outputs,

→ chain matrice elements $||a_{ik}||$ vs. frequency for further data processing e.g. time domain analysis, by Fourier transformation, or bi-linear frequency transformation

All data can be stored on ASCII data files, and/or displayed graphically. Standard grafic output formats are implemented for fast evaluations. The data files are structured such, that they can be read by the spread-sheet programm MS EXCEL. A specially suited EXCEL macro programm allows comfortable graphical presentations in any desired format.

2.4.2.6 SPECIAL FEATURES

Two additional options are common to all analysis programs:

→ analysis of spurious responses of the crystals, and

→ Monte-Carlo analysis for sensitivity studies.

For the analysis of the influence of spurious modes of the crystals an arbitrary number of spuri can be added as additional series resonant circuits in parallel to the crystal resonator. Each spurious can be defined by its resonance frequency and its motional resistance.
In the Monte-Carlo analysis, a maximum tolerance can be associated with each component value.

In a series of cyclic analysis runs the actual component values are varied in a random manner within the given tolerance. The random function can be selected to be either equally distributed ("white") or "Gaussian" with selectable widths of the Gaussian distribution function.

2.4.3 Crystal filter circuit transformation

Circuit transformations are an important tool for effective and cost-saving crystal filter synthesis. The program contains frequency transformation features and means for circuit transformations.

2.4.3.1 BI-LINEAR TRANSFORMATION

The bi-linear transformation transforms the frequency variable df to a transformed variable Df by the following formula

$$Df = \frac{df}{1-K*df}$$

where

K is the transformation (skewing) factor, while
$df = f - f_0$ before, and
$Df = f - f_0$ after transformation

The bi-linear transformation allows the transformation of a frequency symmentric response to an asymmetrical response. For K<0 the frequency f= ∞ of the symmetric case is transformed to a finite frequency above f_0, which generates a "lower sideband" (LSB) amplitude response. For K>0 the frequency f = ∞ is transformed to a finite frequency below f_0 according to an "upper sideband" (USB) amplitude response.
The program "bi-linear transformation" performs the transformation of all circuit elements of a crystal lattice filter structure according to the chosen K value.
Therefore, also frequency-unsymmetric responses can be synthesized using the normalized lowpass and the conventional symmetric lowpass to bandpass transformation as a starting point.
The advantage of this "indirect" method is, that no accuracy problems occur, which would be unevitable in direct synthesis of crystal filters in the bandpass domain.

2.4.3.2 OTHER CIRCUT TRANSFORMATIONS

There are some useful circuit transformations when dealing with crystal filters, which are published in a different papers, e.g. those of Poschenrieder, Colin, Szentirmai etc.The implementation of these transformations into the program package are anticipated, but not yet realized.

2.4.4 Crystal filter analysis in the time domain

The analysis of the filter response in the time domain is done by Fourier transformation of the response in the frequency domain.
The program "time domain analysis" uses the analysis data from the frequency domain network analysis programs described in clause 2.4.2. in the format of the complex chain matix elements $||a_{ik}||$ vs. frequency. It computes the step response at the filter output for an input signal in the form of a switched carrier, from whom the carrier frequency and the switching frequency can be chosen.
If a(k) are the coefficients of the Fourier series of the input signal, the output signal $U_2(t)$ of a filter with the transmission function at the technical frequency jw

$|H(j\omega)| * e^{j(j\omega)}$

can be expressed as

$$U_2(t) = \sum_{k=0}^{\infty} \frac{a(k)}{2*H(k*\omega)} * e^{j[k*\omega t - j(k*\omega)]}$$

The analysis data in the frequency domain must include the whole filter passband, and parts of the stop band, at least down to a relative attenuation of 50dB or more (dependent on the required accuracy).

The number of analysis points in the frequency domain determines the allowed carrier switching frequencies.

2.4.5 Crystal filter optimization

As in the pole-zero domain and in the lowpass domain, an optimization routine also can be effected with the final filter circuit.
The optimization needs a starting circuit with given element values, the characteristic of which should be as close as possible to the requirements of the final filter.
In the optimization run, the value of selected elements can be changed in order to meet a given specification. Therefore severel elements should be coupled, i.e.

- Þ in a tuned circuit an increasing C should automatically decrease the adjacent L,
- Þ in an "all-pole" lattice filter the motional capacitances and the overall parallel capacitances in the opposite arms should vary in parallel
- Þ in a geometrically symmetric circuit, it is sometimes reasonable to keep symmetry, etc.etc.

The optimization strategy is the "Least p-th" method, a generalization of the "least-squares" method. It minimizes the error function

$$\varepsilon = \sum_{i=1}^{n} [w_i(F(x,i)-D(i))]P$$

in which F(x,i) is the i-th frequency sample of the response function, D(i) is the specification limit at this point, and w_i are the weighting factors, which are chosen according to the width of the acceptable tolerance of the desired value D(i). The optimization strategy to find the minimum of the error function is the so-called pattern search algorithm.

Figure 10 shows

- the options of several crystal filter circuit structures and the according characteristics to be synthesized,
- the window to enter filter spezification and required crystal datas,
- the input window to perform a bi-linear transformation,
- the menue item to synthesize crystal stop band filters,
- the option to perform synthesis of crystal filters by using image parameters, of the group "Crystal filter circuits"

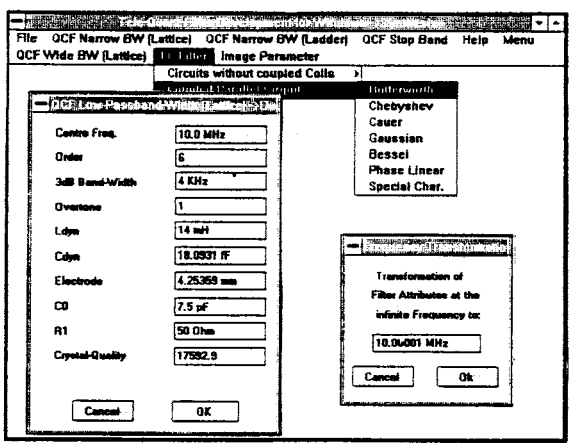


Figure 10: input operations in the "Crystal filter circuit" domain

Figure 11 displays

- the synthesized crystal filter circuit datas bi-linear transformed,
- the according analysis datas in tables,
- the selectivity response and the transformed selectivity response graphically, of the group "Cristal filter circuits".

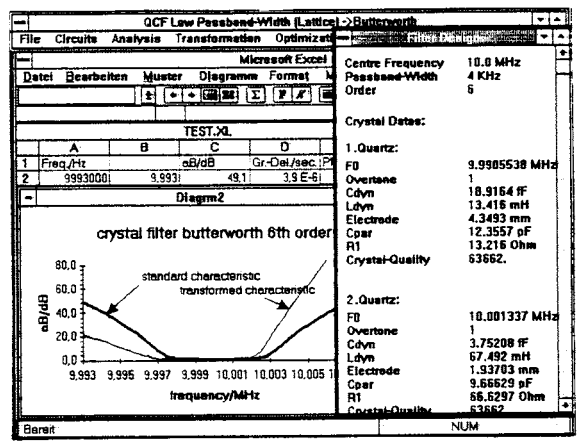


Figure 11: Output operations in the "Cristal filter circuit" domain

3.Summary

Figure 12 finally shows an overview of the interfacing management between the TQ WINFIL Building blocks and their modules.

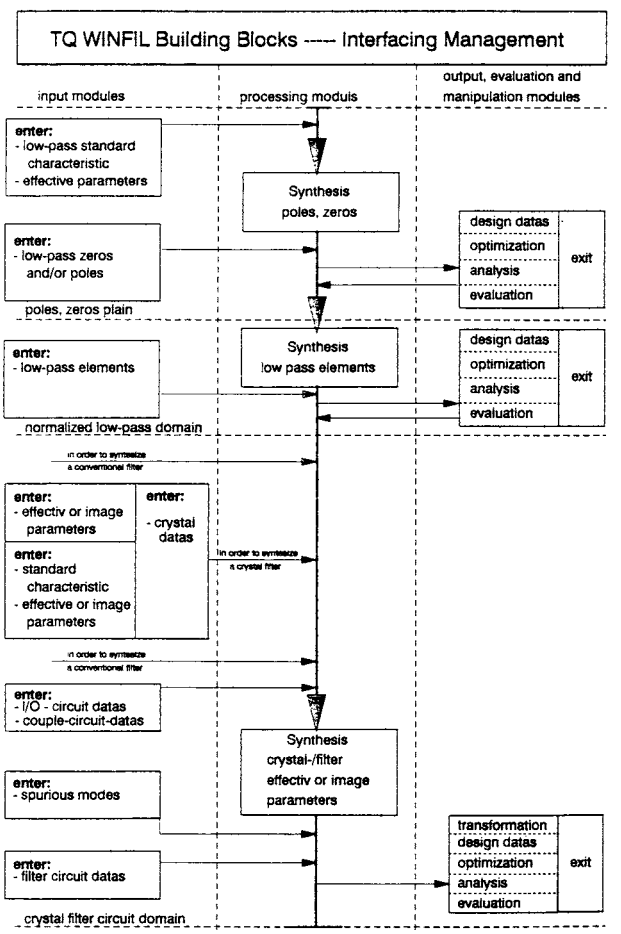


Figure 12: Overview

MONOLITHIC QUARTZ VIBRATING BEAM ACCELEROMETER

F. Deyzac, D. Janiaud, S. Muller

Office National d'Etudes et de Recherches Aérospatiales
BP 72 - 92322 Chatillon Cedex - France

1. SUMMARY

Previous results of this study, performed at ONERA under DRET contract, have been presented at the 4 th EFTF (Neuchâtel, march 1990).
The objective is machining, by chemical etching process, miniature accelerometers suitable for various applications, like guidance systems and vehicle dynamic control.
The monolithic quartz active part includes two vibrating beams, the proof mass and its articulation on a base.
It is obtained by a process similar to the one used in watch crystal manufacturing, so the potential cost is very low.
Moreover, the simplicity of piezoelectric excitation and the numerical output (frequency signal) are key features to reduce dimensions and cost.
The expected accuracy of this sensor is 10^{-3}in harsh environment.
First experimental results are presented.

2. INTRODUCTION

In the field of accelerometry for guidance applications, pendular accelerometers are still baseline devices.
A second type of accelerometer, so called Vibrating Beam Accelerometer, is the subject of intensive studies aiming at improved characteristics.
That is the case of the Monolithic Quartz Accelerometer developed at ONERA since a few years.
The principle of operation of this accelerometer is based on the flexure vibrations of two beams; these vibrations are maintained with piezoelectric excitation by means of electrodes and oscillator circuits. When submitted to an acceleration, the proof mass applies, on vibrating beams, axial forces which modify their resonance frequencies.
So, the variations of the beams frequencies leads to the measurement of applied acceleration.
The reduction of the accelerometer thermal sensitivity is obtained by the use of the two beams in a differential configuration.

3. MONOLITHIC ACCELEROMETER CONFIGURATION

The configuration of the Monolithic Quartz Accelerometer, patented by ONERA, is presented in figure 1.

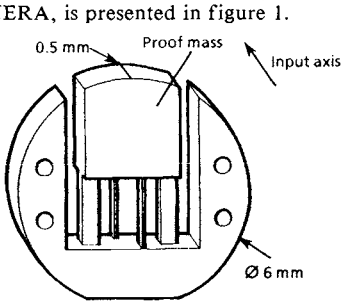


Figure 1. Monolithic Quartz Accelerometer with the two vibrating beams at the center between the two articulation beams.

This geometry has been first defined through analytical studies for 50 G range and 0-500 Hz measurement bandwidth.
The sizes have been optimized with a finite element model by use of SAMCEF software.

4. CHEMICAL ETCHING OF THE STRUCTURE

The original feature of the ONERA's sensor rests on its monolithic design well adapted to the chemical etching of quartz crystal: Wafer machining through photolithographic technics (see figure 2).
The accuracy of chemical etching leads to high quality miniature devices.
This machining process is performed at ONERA on commercially available wafers used in watch industry.

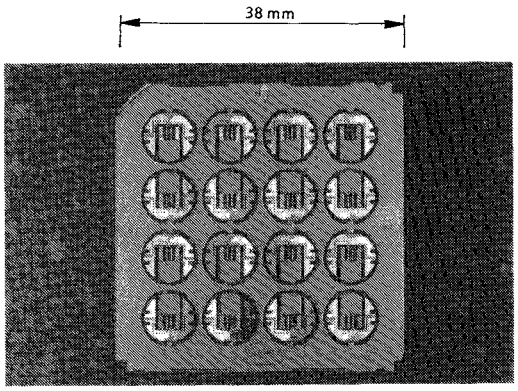


Figure 2. Quartz wafer (16 accelerometers)

5. FIRST EXPERIMENTAL RESULTS

Presently, the main effort is devoted to the integration of the crystal into a small tight case (0,5 cm^3) (see figure 3).

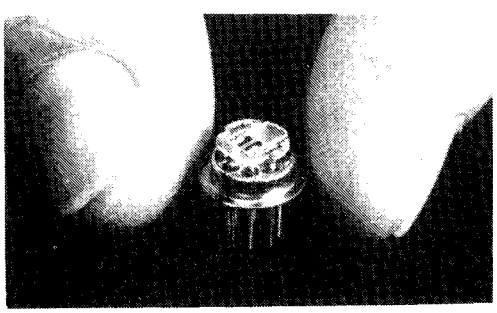


Figure 3. Monolithic accelerometer fixed on a TO 5 base (TO 5 is a nominal capsule for transistors).

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

This first evaluation of prototypes performed this year has exhibited encouraging results:

- Operating frequencies 100 kHz
- scale factor 50 Hz/G
- resolution 10 μG
- residual errors between ± G: 50 μG
  (after modeling)

6. CONCLUSION

The faisability of this Monolithic Quartz Vibrating Beam Accelerometer is now established and several prototypes have been made in ONERA.
The presently developed version has been optimized in view of a low cost miniaturized accelerometer.
Next studies will be devoted to the development of new versions in order to investigate the high potential performances of this concept of monolithic sensor.

SURFACE TRANSVERSE WAVE (STW) QUARTZ RESONATORS IN THE GHz RANGE

E.Bigler and E.Gavignet (*), B.A.Auld (**), E.Ritz and E.Sang (***)

(*) Laboratoire de Physique et Métrologie des Oscillateurs du CNRS
associé à l'Université de Franche-Comté-Besançon
32, avenue de l'Observatoire - 25000 Besançon - France - phone (33) 81 66 69 54

(**) Ginzton Laboratory, Stanford University, Stanford, CA94305 (USA) phone (415) 723 0264

(***) Tektronix, SAW technology, MS 50-240, P.O.Box 500, Beaverton, OR 97077 (USA)

ABSTRACT

Surface Transverse Wave (STW) resonators offer several advantages over conventional Rayleigh wave SAW resonators, one of the most interesting advantages being higher operating frequencies (+ 60 %) for given geometries of the transducers and reflectors. A model is presented to compute the resonant frequency of a 3-grating STW resonator, and the influence of important design parameters : cavity length, metal thickness, grating periodicities and finger-to-gap ratio, is discussed. Starting from first experimental results obtained on metal strip (37°-Y-rotated) quartz STW resonators at 500 MHz, new devices at 1 GHz have been fabricated and tested. New experimental results are presented and discussed.

1. INTRODUCTION

Surface transverse wave (STW) resonators were introduced several years ago (Refs. 1, 2, 3). The theory of these devices can be considered as well established (Refs. 4-9), and very attractive examples of devices in the GHz range have recently been presented (Refs. 10, 11).

But only few experimental results have been reported on the influence of design parameters on fundamental properties : resonant frequency, Q-factor, insertion losses ... Only recently in an experimental work by Avramov (Ref. 12) a systematic investigation of the influence of metal thickness on fundamental properties of metal strip STW resonators was reported.

The scope of this paper is to present a model predicting the influence of various design parameters : geometrical dimensions, metal thickness or groove depth, on the resonant frequency of STW resonators. This model is based on the theory (Refs. 4, 9, 13) of STW propagation at the surface of an anisotropic substrate covered with a periodic array of grooves or metal strips, in the nonpiezoelectric approximation.

Theoretical results are presented in the case of a 3-grating configuration (mirror-cavity-mirror) for both cases of metal strips or grooves. The influence of cavity length, metal thickness, grating periodicities and finger-to-gap ratio is demonstrated, and predicted results are compared with already published experimental data (Refs. 3, 12). Results for metal strips obtained on (37°-Y rotated) quartz STW resonators at 500 MHz and 1 GHz are presented and discussed.

2. DETERMINATION OF THE RESONANT FREQUENCY OF A 3-GRATING STW RESONATOR

The model used to predict the resonant frequency of a STW resonator is based on an expansion of the acoustic field in coupled modes for transverse horizontal waves propagating under a periodic grating (Fig. 1).

In the case of small h/Λ ratios, and in the near stopband approximation, two coupled modes only can be taken into account and analytical formulae can be derived for STW propagation on a singly-rotated quartz plate (Fig. 1).

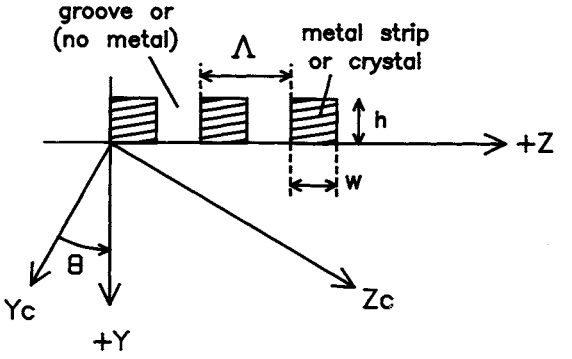


Fig. 1 : *STW propagation under an array of grooves or metal strips*
h = grooved depth / strip height Λ *= periodicity of the array*
Xc Yc Zc crystallographic axes
X = Xc, Y, Z singly-rotated axes
w = strip width w/Λ *= finger-to-gap ratio*

Basic formulae for STW propagation under metal strips or grooves have already been published in Ref. 14 and will be used here to model a 3-grating resonator (Fig. 2).

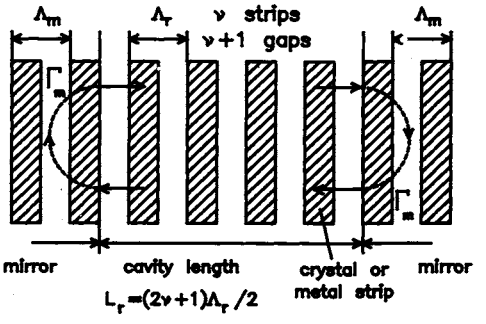


Fig. 2 : *Model of a 3-grating resonator*

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

A cavity, made of v strips and $(v+1)$ gaps (or $(v+1)$ grooves), with a periodicity Λ_r, total length $Lr = (2v+1)\Lambda_r/2$ is placed in the middle of two reflectors made of a high number of strips or grooves with a slightly different periodicity Λ_m. A resonance can take place if the following condition is fullfilled

$$\Gamma_m^2 = e^{2i\beta_r L_r} \qquad (1)$$

where Γ_m is the reflection coefficient of the mirror, operating in the stopband of the periodic structure (Λ_m), and β_r is the propagation constant of the wave inside the cavity operating in the passband of the periodic structure (Λ_r) with $\Lambda_r < \Lambda_m$. Equation (1) is a simplified resonant condition, considering inside the cavity that propagation of waves is the same as it would be either under a continuous layer of effective thickness h (w/Λ_r) (Love wave limit for metal strip gratings) or under a non-grooved medium (SSBW limit).

The reflection coefficient Γ_m is obtained in the limit of an infinite array length as (Ref. 14)

$$\Gamma_m = \sin\Theta - i\cos\Theta = e^{i(\Theta-\pi/2)} \qquad (2)$$

where the phase angle Θ is a function of ω, and is given by a derivation not presented here, similar to Ref. 14, but now taking into account a given value of the finger-to-gap ratio w/Λ_m

$$\cos\theta = -\frac{q_o\, W/\Lambda_m}{q\sqrt{K}\sin(\pi W/\Lambda_m)} + \left(\frac{1-\frac{\delta\omega}{V_s}/[K q^2 \sin^2(\pi W/\Lambda_m)]}{2} - \frac{q_o^2 \frac{w^2}{\Lambda_m^2}}{K q^2 \sin^2(\frac{\pi W}{\Lambda_m})}\right)^{1/2} \qquad (3)$$

where
V_s = velocity of shear horizontal waves (SSBW) in an ungrooved medium

$$\delta\omega = \omega - \frac{\pi V_s}{\Lambda_m} \qquad (4)$$

q_o, q are combinations of material constants (crystal + metal) and K is the characteristic coupling constant of the periodic array as given in Ref. 14, those expressions being valid for a finger-to-gap ratio w/Λ_m. For grooved arrays, eq. (3) takes a simpler form with the limit values $q = 1$, $q_o = 0$.

For a 50/50 finger-to-gap ratio another simplification takes place with $w/\Lambda_m = 0.5$ and $\sin(\pi\, w/\Lambda_m) = 1$.

The strong dependence of the phase angle Θ in eq. (3), with respect to the variations of angular frequency $\delta\omega$ within the mirror stopband is a new phenomenon, very different from what happens with conventional Rayleigh (SAW) resonators.

The algorithm used to find the resonant frequency consists in computing the values of β_r as a function of ω (using the Love-wave limit or the SSBW limit) and checking for the phase matching in eq. (1) using eq. (3) to compute the phase angle of the reflection coefficient. This can be done using a graphical determination (Ref. 14)

3. SIMULATION RESULTS

Simulations using the algorithm described in Sec. 2 will be presented here, obtained with the following nominal parameters

- nonpiezoelectric approximation for quartz, Y + 36° cut (convention on cut angles according to Ref. 15)
- operating frequency close to 500 MHz
- metal thickness or groove depth : 1 000 Å
- periodicity of mirrors : 5.1 µm

constants of Y + 36° quartz : ρ = 2 650 kg/m^3
$C_{66} = 2{,}895\times10^{10}$ Pa
$C_{eff} = 6.830\times10^{10}$ Pa V_s = 5 070 m/s
metal Aℓ + 4%Cu ρ' = 2 695 kg/m3 V_m = 3 046 m/s
(all constants defined as in Ref. 14)

3.1 Influence of the number of elements inside the cavity

Fig. 3 shows the dependence of the resonant frequency vs number of strips v inside the cavity (see Fig. 2) in the case of a metal strip resonator. Since maximum of energy trapping effects are obtained at the bottom of the stopband, best results are expected in both cases in the vicinity of 25 and 48 strips inside the cavity. Based on these simulations, metal strip resonators have been fabricated and tested with 25, 31 and 37 strips on (Y + 37°) quartz. Results are presented below in sec. 4.

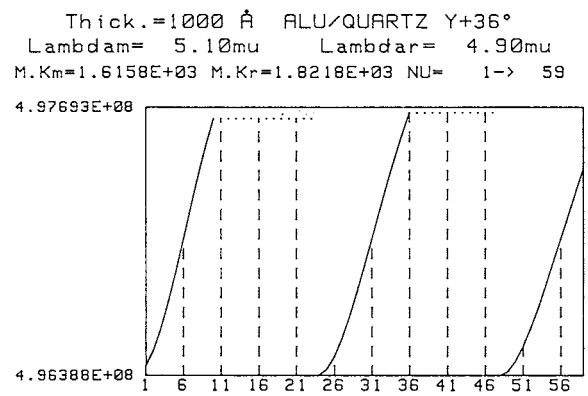


Fig. 3 : Metal strip resonator
Influence of the number of elements inside the cavity on the resonant frequency. Resonant frequencies in the upper part of stopband are not of practical use due to radiation losses of the device

3.2 Influence of the finger-to-gap ratio

Equation (3) shows the dependence of the reflection coefficient Γ_m versus the finger-to-gap ratio w/Λ_m. Figure 4 shows a simulation of the influence of this factor on a 500 MHz design, with the following parameters : cut angle Θ = 37°, number of strips inside the cavity : v = 31, aluminum thickness = 1 100 Å, other parameters same as above. Starting from a 50/50 ratio, the simulation shows that 1 % of variation in w/Λ will induce a frequency change of ~ 16 ppm.

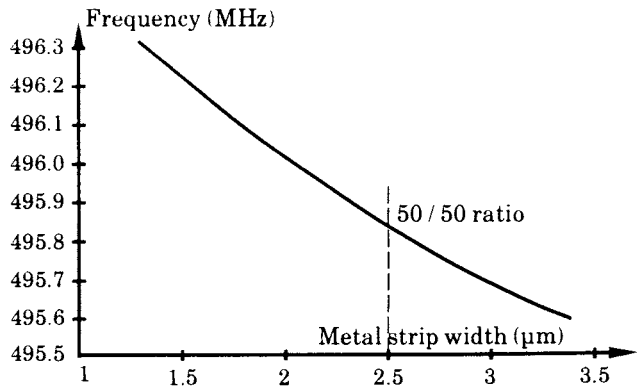


Fig. 4 : *Influence of the the finger-to-gap ratio on a 500 MHz design* (Λ = 5 µm)

3.3 Influence of different design parameters

Based on nominal parameters given before, a metal strip resonator design with 25 grooves inside cavity exhibits sensitivity to different other parameters.

* a scaling factor applied to both geometrical dimensions Λ_r and Λ_m will scale the resonant frequency by the same factor.
* metal thickness changing from 1 000 Å to 800 Å will increase the resonant frequency by 900 ppm (see below sec. 4.2).
* changing the cut angle from 36° to 37° decreases the resonant frequency by 2 000 ppm, but as mentioned in Refs. 3 and 6 the most important effect of changing the cut angle is changing the turnover temperature of the device.

4. COMPARISON WITH ALREDAY PUBLISHED DATA. NEW EXPERIMENTAL RESULTS

4.1 Corrections to the nonpiezoelectric approximation

Comparison of frequency ranges predicted by the model (Fig. 3) for a 500 MHz device with already published data (Ref. 3) shows that predicted operating frequencies are lower than observed experimentally. A reason for this may be found in the nonpiezoelectric approximation used here, predicting SSBW velocity V_s = 5 070 m/s for Y + 36° quartz cut, while experimental value given by Lewis (Ref. 16) is 5 099 m/s, very close to the value given by classical piezoelectric bulk mode analysis. To take into account this increased value (by 0.5 %) for V_s, an attempt was made by simply changing value of V_s and $C_{eff} = \rho V_s^2$ in all formulae used in the simulation, and a better agreement between theory and experiment was found for new metal strip resonators (see below sec. 4.3).

4.2 Influence of metal thickness : comparison with Avramov's results (Ref. 12)

A systematic experimental investigation of the influence of metal thickness is presented by Avramov (Ref. 12) on 620 MHz metal strip STW resonators on (Y + 36°) quartz. To simulate the behavior or such a device, the following inputs were used in our simulation program.
mirror periodicity Λ_m = 4.1 µm
acoustic wavelength λ_a = 8.2 µm
resonator periodicity Λ_r = 4.05 µm
number of strips inside cavity : 11
quartz SSBW velocity : 5 099 m/s

Figure 5 compares experimental values of frequency shifts with predicted results according to our model. the agreement is fairly good.

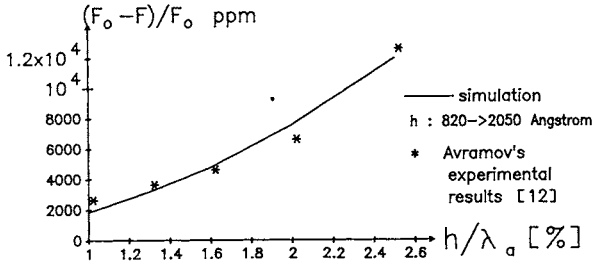


Fig. 5 : *Influence of metal thickness on a 620 MHz metal strip resonator.*
Reference frequency F_0 corresponds to h/Λ_a = 0.4 %

4.3 Results obtained on 500 MHz metal strip resonators

Metal strip resonators were fabricated and tested according to the design curve of Fig. 3. Three sets of devices were fabricated on Y + 37° quartz with respectively 25, 31 and 37 strips inside cavity. Results obtained on the first prototypes with 25 and 31 strips are listed on Table 1. Aluminum thickness was 1 100 ± 50 Å, Λ_m = 5.1 µm Λ_r = 4.9 µm.

Table 1 :
Comparison of predicted values of frequency with experimental results on metal strip STW resonators
Mean value and standard deviation of experimental data are obtained from a set of four devices

Number of strips inside cavity	25	31
Predicted frequency (non piezo) Vs = 5 066 Vm = 3 046 (MHz)	495.386	496.010
Corrected value of predicted frequency V_s=5 092 V_m=3 046 (MHz)	497.679	498.449
Experimental frequency (MHz)	497.737 ± 0.014	498.037 ± 0.010
Unloaded Q factor (experimental)	6.90 ±0.61 × 10^3	2.98±0.08 × 10^3
Insertion losses (experimental) (dB)	9.1 ± 0.4	11.4 ± 0.2

Frequency responses of a resonator with cavity length of 25, 31 and 37 strips are shown on Fig. 6.

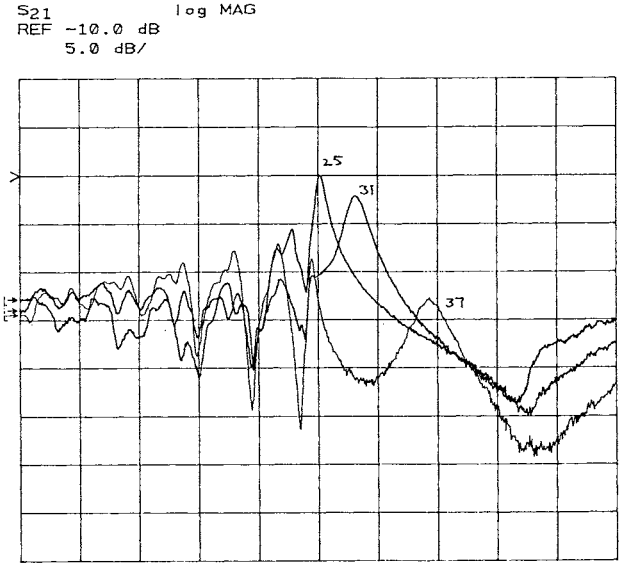


Fig. 6 :
Response curve at 500 MHz for 25, 31 and 37 strips
Frequency scale : 500 kHz/div center : 497.675 MHz
log amplitude scale : 5 dB/div

These results clearly show the influence of a correct position of the resonant frequency inside the stopband, best values of Q factor and insertion losses being obtained at lower edge of the stopband. Corrections to the nonpiezoelectric model by simply changing values of V_s give a better agreement between theory and experiment ; but the question is to know a correct value for ρ' and V_m, metal properties depending on the deposition process.

4.4 Results obtained at 1 GHz

Starting from the previous design, devices at 1 GHz were fabricated and tested. Preliminary results are shown on Fig. 7 for a resonator with a cavity length of 25 strips, with aluminum thickness between 550 and 600 Å. The unloaded Q-factor is close to 4 800.

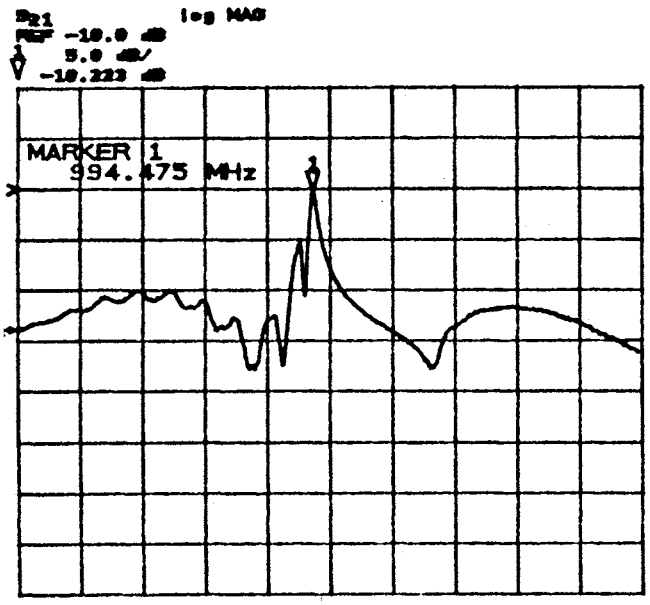


Fig. 7 :
Frequency response of a 1 GHz metal strip STW resonator cacity length : 25 strips - frequency scale : 2 MHz/div marker : 994.47 MHz / -10.2 dB log amplitude scale : 5 dB/div

4.5 Discussion

The general shapes of the frequency responses show a very smooth edge on the higher frequency side of the resonant peak, while numerous secondary peaks appear on the lower frequency side. Variability in the position and relative height of these sidelobes could be due to the simultaneous excitation of untrapped SSBW modes in conjugation with the STW modes. SSBW modes are not eigen modes of the corrugated surface, but they are radiated by the transducers and the relative coupling efficiency of the transducer to both modes could be controlled by a proper choice of the number of finger pairs in the transducers as well as adjusting the metal thickness (Refs. 10 and 17). A better understanding of this phenomenon will be required to define a stable design with a high peak to sidelobe ratio.

5. CONCLUSION

The analytical model presented here allows with simple computations a systematic investigation of the basic design parameters of a STW resonator. One of the most important point consists in operating at the lower edge of the mirror stopband to ensure high Q and low loss, and this can be achieved using the model in its present form.

It has been shown how a basic design can be scaled to higher frequencies. However several phenomena like coupling between SSBW and STW modes probably induce some variability in the shapes of the response curves.

The model can be extended to any combination of groove/metal strip gratings that could be required for a better control temperature effects in a STW resonator.

ACKNOWLEDGEMENTS

The authors wish to thank William S. Drummond and Robert A. Mueller for their valuable assistance with device fabrication and measurement, and also Walter D. Fields for his assistance with the measurements.

REFERENCES

1. V.J. Anisimkin, M.A. Magomedov, V.I. Fedorov 1985, "Resonators on surface transverse waves", *Soviet Physics : J. Techn. Phys.* (letters), 11, 1, 56-60.

2. A. Ronnekleiv 1986, "High Q resonators based on surface transverse waves", *Ultrasonics Symp. Proc.*, IEEE cat. n° 86CH2375-4, pp. 257-280.

3. T.L. Bagwell and R.C. Bray 1987, "Novel surface transverse wave resonators with low loss and high Q", *Ultrasonics Symp. Proc.*, IEEE cat. n° 87CH2492-7, pp. 319-324.

4. A. Renard, J. Henaff, B.A. Auld 1981, "SH surface wave propagation on corrugated surfaces of rotated Y-cut quartz and berlinite crystals", *Ultrasonics Symp. Proc.*, IEEE cat. 81CH1689-9, pp. 123-127.

5. D.F. Thompson, B.A. Auld 1986, "Surface Transverse Wave propagation under metal strip gratings", *Ultrasonics Symp. Proc.*, IEEE cat. n° 86CH2375-4, pp. 261-266.

6. B.A. Auld and D.F. Thompson 1987, "Temperature compensation of Surface Transverse Waves for stable oscillator applications", *Ultrasonics Symp. Proc.*, IEEE cat. n° 87CH2492-7, pp. 305-312.

7. C.A. Flory and R.L. Baer 1987, "Surface transverse wave mode analysis and coupling to interdigital transducers", *Ultrasonics Symp. Proc.*, IEEE cat. n° 87CH2492-7, pp. 313-318.

8. A. Ronnekleiv 1987, "Anisotropy in surface transverse waves resonators", *Ultrasonics Symp. Proc.*, IEEE cat. n° 87CH2492-7, pp. 325-329.

9. B.A. Auld 1989, "The STW resonator, a new control element for high stability oscillators", *Proc. 3rd European Frequency and Time Forum*, Besançon, France, pp. 152-158.

10. I.D. Avramov 1991, "1 GHz low loss coupled resonator filter using surface skimming bulk waves and Bleustein-Gulyaev waves", *Elect. Lett.*, 5, 27, 414-415.

11. D.F. Thompson and A.R. Northam 1991, "High frequency, low loss, fundamental mode STW resonators used for stable oscillator design", *Ultrasonics Symp.*, Lake Buena Vista, Florida.

12. I.D. Avramov 1990, "High Q metal strip SSBW resonators using a SAW design", *IEEE Trans. Ultr. Ferro. Freq. Contr.*, UFFC-37, 6, 530-534.

13. P.S. Cross and R.V. Schmidt 1977, "Coupled surface acoustic wave resonators", *Bell System Technical Journal*, 56, 1447-1482.

14. E. Bigler, B.A. Auld, E. Ritz and E. Sang 1991, "An analysis of the influence of design parameters on the resonant frequency and Q-factor of surface transverse wave (STW) resonators", *Proc. 45th Ann. Freq. Cont. Symp.*, Los Angeles.

15. IEEE Standard on Piezoelectricity 176-1949, *Proc. IRE*, 37, 1378-1395.

16. M. Lewis 1977, "Surface skimming bulk waves, SSBW", *Ultrasonics Symp. Proc.*, IEEE cat. 77CH1264-1SU, pp. 744-752.

17. D.L. Lee 1980, "S-band SSBW delay lines for oscillator application", *Ultrasonics Symp. Proc.*, IEEE cat. n° 80 CH 1602-2, pp. 245-250.

COMPARISON OF ENERGY TRAPPING RESONATORS AND FILTERS USING THE AT CUT OF QUARTZ, BERLINITE AND GALLIUM PHOSPHATE.

J.Détaint*, J.Schwartzel*, C.Joly*, B.Capelle**,
A.Zarka**, Y.Zheng**, Y.Toudic***, E.Philippot****.

*CNET dept.PAB/BAG/MCT/CMM 196,Av.H.Ravera 92220 Bagneux France.
**L.M.C.P. Université Paris VI 4,place Jussieu 75005 Paris France.
***CNET dept. LAB/OCM/MPA route de Tregastel 22301 Lannion France.
****L.P.C.M.S. Université du Languedoc Place E.Bataillon Montpellier France.

Abstract:

Quartz and its analogues are the only known materials that lead to devices having a very high thermal frequency stability in a large temperature range together with a small acoustic dissipation. The square of the coupling coefficient is approximatively doubled from AT quartz to AT berlinite and again doubled from AT berlinite to AT GaPO4 so that it is possible to achieve a large range of applications in the fields of frequency generation or filtering, using these three materials.

In a first part of this communication, the properties of the AT cuts of these three materials are compared (velocity of the shear waves, coupling coefficient, and thermal behaviour). Then a more precise evaluation of the properties of AT cut resonators using the quartz analogues is made in considering the computed properties of energy trapped resonators. The computed properties of resonators are obtained using a three dimensional model based upon the theory of the essentially thickness modes (H.F.Tiersten). Using computed equivalent schemes, the bandwidth achievable at 20 MHz are calculated for 4 poles Jaumann filters having a Chebishev response with 1 dB ripple and a maximal termination impedance of 2 kohm. The obtained results indicate that berlinite and gallium phosphate permit to product filter with much larger bandwidth than quartz. They suggest also that the use of new designs (resonators with elliptical electrodes and/or grooved electrodes, integrated filters etc...) will permit to lower further the equivalent inductance and/or the influence of the stray capacitances. Such designs would lead with the high coupling quartz analogues to much better characteristics than the usual designs using round electrodes.

Experimental results (equivalent scheme of resonators, observed mode shape etc..) are successfully compared with computed ones to validate the model use for the computations.

1 Introduction:

Since 15 years many works considering the obtainement, the properties and the applications of several quartz like materials having enhanced piezoelectric properties were reported. Berlinite was the first analogue of quartz to be synthesized [1][2][3][4] and evaluated [5]. Berlinite crystals of good quality and sufficient size are now available [6] to permit detailed investigations of devices and the obtention of industrial prototypes. The obtention of gallium phosphate crystals having significative dimensions and perfection was reported in the recent years [7] together with the measured constants of this material and the first evaluations of devices [8] [9]. In these proceedings, a major advance in the obtention of large GaPO4 crystals having an improved quality is reported [10], together with new results of characterization of devices.

To determine precisely the respective potentialities of the quartz analogues in view of an application to I.F. filtering in G.S.M. sets, an investigation was made of the properties of filter type resonators using these materials. In this communication we report some of the results obtained that permit a comparison of the properties of AT resonators and filters using quartz berlinite and GaPO4.

2 Properties of the AT cuts of Quartz, AlPO4 and GaPO4.

In table I the computed or observed basic properties of the piezoelectrically excited slow shear modes existing in the AT cut of quartz, aluminium phosphate and gallium phosphate are presented. The values given for quartz were computed from the constants of James [11] by a one dimensional model. For aluminium and gallium phosphate, they are the results of recent experiments made at CNET. The most remarkable facts to notice in table I are the decrease of the shear velocity from quartz to gallium phosphate and the increase of the dielectric permittivity and mostly of the coupling coefficient. The exact angular positions of the AT cuts in berlinite [12] and GaPO4 [7] [10] are still to be determined as accurately as in quartz, but, since the quantities reported in table I have slow angular variations, they are believed to be accurate to about 5% for berlinite while they represent for GaPO4 the most accurate evaluation presently available. In this latter case, the comparison of computed and experimental values of resonator properties will give further informations on their accuracy (see § 5).

Table I properties of the AT cut of quartz like materials.

	QUARTZ	AlPO4	GaPO4
V (m/s)	3321.	2942.*	2543.*
$\bar{C}$ (10^{10}N.m^{-2})	2.921	2.285*	2.310*
Eps_r^S	4.499	4.917*	6.119*
K (%)	8.8	11.*	16.*

The most important property of quartz like materials is to possess a cut giving a very nearly pure third order thermal behaviour ("true" AT cut). This means that for a certain cut and a certain reference temperature, the 1st and the 2nd order temperature coefficients of the frequency vanish simultaneously. With a good approximation, the variations of the resonance frequencies can then be described by the relation:

$\frac{fr(T)-fr(T_{inf})}{fr(T_{inf})} = Tf^{(3)}_{inf}.(T-T_{inf})^3$ where: $Tf^{(3)}_{inf}$ is the third order T.C. referenced to the inflexion temperature T_{inf}.

In table II we compare the third order temperature coefficients and the inflexion temperatures for these three materials. The values given for GaPO4 are still very tentative since they were obtained from measurements made for two near AT cuts with the assumption that as in quartz and berlinite, the 3rd order temperature coefficient has small angular variations in the vicinity of the AT cut. They are to be taken as a first evaluation that will be perfected in the near future. The results in table II seem to indicate a very similar thermal behaviour of quartz and of its analogues but centred at different temperatures. This can

be already verified for berlinite (figure 1) by a direct comparison to the variations known for AT quartz.

Table II: 3rd order temperature coefficients of the resonance frequency.

	Quartz	Berlinite	GaPO4
$Tf^{(3)}$(°C^{-3})	$9-10.10^{-11}$	$8-12.10^{-11}$	$4-16.10^{-11}$?*
T_{inf}(°C)	25-28°	62-67°	40-50°? *

*1st estimations from near AT results

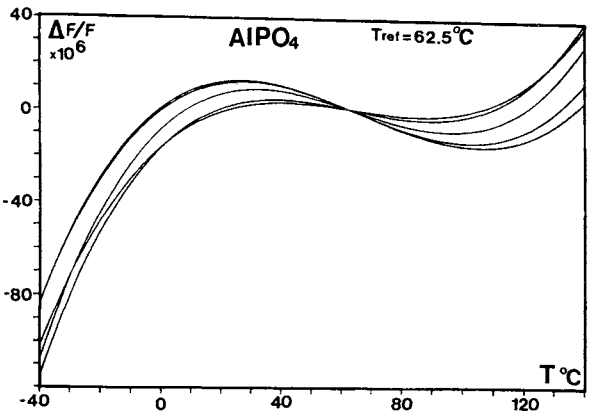


Figure 1: Thermal variations of fr near the AT cut of berlinite.

3 Comparison of energy trapping properties and of energy trapping devices:

Since the properties of real resonators cannot be well described by the one dimensional theory, we have further compared these three materials by an evaluation of the properties of energy trapping resonators made with them. For this purpose a model [12] based upon the theory of H.F.Tiersten [13] was used. This theory that was experimentally verified in many instances, postulate that the thickness variations of the displacements are essentially similar to that of the one dimensional modes whereas the lateral variations are governed by the equation:

$$M_n\frac{\partial^2\tilde{u}_1^n}{\partial x_1^2}+P_n\frac{\partial^2\tilde{u}_1^n}{\partial x_3^2}-\frac{n^2\pi^2C^*}{4h^2}\tilde{u}_1^n+\rho\omega^2\tilde{u}_1^n=\rho\omega^2(-1)^{(n-1)/2}\frac{e_{26}4V_0e^{j\omega t}}{c^{(1)}n^2\pi^2} \qquad \{1\}$$

In this equation, $\tilde{u}_1(x_1,x_3,t)$ is the lateral dependence of $\hat{u}_1^n(x_1,x_2,x_3,t)=\tilde{u}_1^n\sin(n\pi x_2/2h)$.. $\hat{u}_1$ arises from a transformation of the usual component of the displacement u_1 and of ϕ the potential , made to replace the inhomogeneous boundary conditions $\phi=\pm V_0e^{j\omega t}/2at\,x_2=\pm h$ by the homogeneous conditions $\hat{\phi}=0$. The inhomogeneous term of the equation results from this transformation. c* is either $\bar{c}^{(1)}$ for the unelectroded part of the resonator or $\hat{c}^{(1)}$ for the electroded part. $\bar{c}^{(1)}$ Is the stiffened elastic constant relative to the corresponding one dimensional mode (Eigen-value of the Christoffel Matrix see Table I). $\hat{c}^{(1)}=\bar{c}^{(1)}(1-8k^{(1)2}/n^2\pi^2-2\hat{R})$ is a constant that includes the electrical effect of the metallization. $4k^{(1)2}/n^2\ \pi^2$ is approximatively the relative frequency lowering due to the electrical effect of the metallization, $\hat{R}$ is the mass loading. $c^{(1)}=\bar{c}^{(1)}(1-k^{(1)})^2$ Is a "pseudo ordinary elastic constant" which is equal to C_{66} for the Y rotated cuts. $k^{(1)}$ is the coupling coefficient of the corresponding one dimensional mode, this quantity is equal to k_{26} (noted k in Table I) for the Y rotated cuts . M , P are functions of the material constants and of the plate orientation that can be obtained as described in reference [13].

The eigen modes (at V=0) of the resonator are obtained by the simultaneous solution of the equations for the electroded and unelectroded regions taking into account the continuity conditions at the electrode edge. The forced modes are computed as a linear combination of the eigen modes verifying the inhomogeneous equation. At V=0 the homogeneous form of the approximate equation for the two regions reduces to :

$$M_nu_{1,11}^n+P_nu_{1,33}^n+4\pi^2\rho(f^2-f^{*2})u_1^n=0 \qquad \{2\}$$

with : $f^*=f_{ce1}=\frac{n}{4h}\sqrt{\frac{\hat{c}^{(1)}}{\rho}}$ for the electroded region. Or:

$f^*=f_{cl}=\frac{n}{4h}\sqrt{\frac{\bar{c}^{(1)}}{\rho}}$ for the unelectroded region.

Using coordinate transformations :

$X_1=\left(\frac{C^*}{M_n}\right)^{1/2}.\gamma.x_1;X_3=\left(\frac{C^*}{P_n}\right)^{1/2}.\gamma.x_3$; $X_1=r\cos t;X_3=r\sin t$ with:

$\gamma^4=\frac{n^2\pi^2}{(2h)^4}$ where: $c^*=\hat{c}^{(1)}$ for the electroded region and $c^*=\bar{c}^{(1)}$ for the unelectroded region. We obtain the equation :

$$u_{1,rr}+\frac{1}{r}u_{1,r}+\frac{1}{r^2}u_{1,tt}+Au_1=0$$

$$Where:A=\frac{n\pi}{4}\left(\frac{f^2-f^{*2}}{f^{*2}}\right) \qquad \{3\}$$

that can be separated as $u_1(r,t)=R(r).T(t)$ and solved as described in [12] to obtain the eigen modes from which the forced modes are computed.

The important parameters of equations {1} to {3} for the fundamental mode of the AT cut of the quartz like materials were computed from the published constants corrected for berlinite and GaPO4 by the results given in table I. They are given in table III:

table III Energy trapping constants for the AT cuts of the quartz like materials (in 10^{10} N.m^{-2}

	M1	P1	$\bar{C}$	$\sqrt{M_1/P_1}$
Quartz	11.008	6.922	2.921	1.261
AlPO4	8.718	5.084	2.285	1.309
GaPO4	7.663	3.805	2.310	1.419

Beside the "elastic constant" like variations of the coefficients between quartz and GaPO4, two important observations are to be made about the data presented in table III: -First, the ratio $\sqrt{\frac{M_n}{P_n}}$ that was shown [14] to represent the "lateral anisotropy" of the plate increases notably from quartz to GaPO4. -The second fact to be noticed is that the interval (f_{cl}-f_{ce}) in which trapped mode can propagate increase also greatly from quartz to GaPO4, due to the large increase of the electrical frequency lowering $\left(\frac{4k^2}{\pi^2}\right)$. So that at equal geometry, and mass loading, more and more modes are excited as the coupling coefficient increases. This fact leads to diminish the electrode dimensions from quartz to GaPO4, if it is necessary to have a response free from anharmonics or without too many anharmonics.

Using the model [12] we have computed the equivalent scheme of 20.0 MHz quartz, AlPO4 and GaPO4 resonators with circular electrodes having in each case a diameter determined so that the response has no symmetrical anharmonics (filter type resonators) with a mass loading of 0.5% . The computed diameter were rounded down to the nearest tenth of mm, and the equivalent schemes of the resonators were computed for these dimensions. The results are presented in table IV. The corresponding computed modes are represented at the same scale in figure 2. The results in table IV and in figure 1 indicate (as previously said) that smaller and smaller electrodes dimensions are required for AlPO4 and then GaPO4 than for quartz. The inductance decreases from quartz to GaPO4, the value of C_0 decreases also due to the reduction of the electrode area, although the dielectric permittivity increases. As expected, the C_0/C_1 capacitance ratio decreases strongly from quartz to GaPO4.

On table V the bandwidth of 4 poles Jaumann filters having a 1.dB ripple Chebishev response achievable with equal termination resistances of 2000 Ohm and the resonators of table

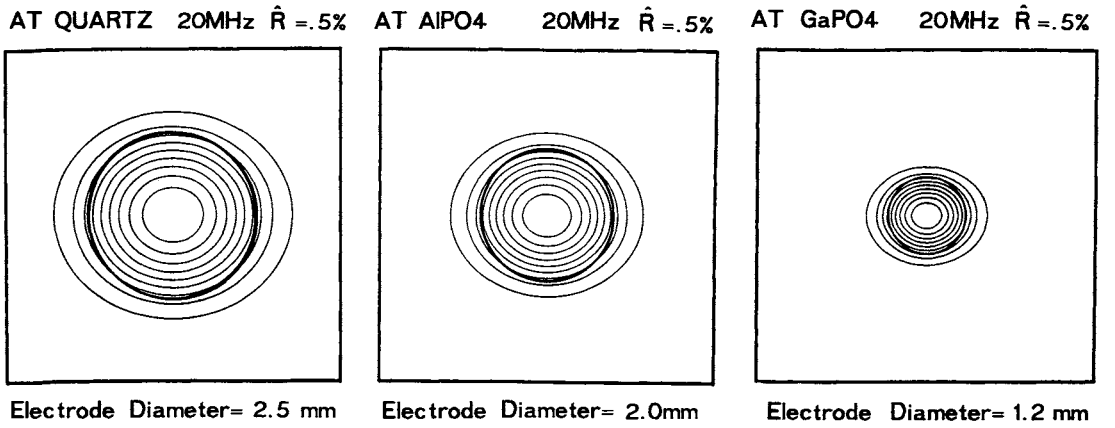


Figure 2: Computed modes of the resonators considered in table 1.

IV are given. As expected, the larger bandwidth is achieved with GaPO4, while berlinite give also an interesting value. The corresponding computed responses are represented in figure 3.

Table IV: Equivalent scheme of resonators designed to have no symmetrical anharmonics.

(max. diameter to have no symmetrical anharmonics with R=0.5%)

material diameter mm	motional inductance L (mH)	motional capacitance C1 (fF)	static capacitance Co (pF)	Co/C1
Quartz 2.5	5.045	12.537	2.367	188
AlPO4 2.	4.123	15.366	1.862	121
GaPO4 1.2	3.761	16.833	0.973	57.8

Table V: Achievable bandwidth for 4 pole Jaumann filters.

Max B.W. with suppression of the symmetrical anharm.

	AT Quartz	AT Berlinite	AT GaPO4
B.W. (kHz)	107.6	131.7	145.

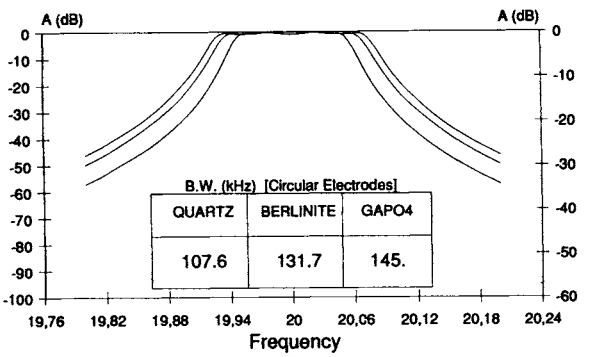


Figure 3: Computed response curve of 4 poles Jaumann filters made using the resonators considered in table IV.

Particularly at V.H.F. frequencies, the small imperfections that are inherent to a practical elaboration process (electrodes misalignment or asymmetry, slight defect of parallelism etc...) can induce the excitation of antisymmetrical modes that are normally unexcited (eigen modes non electrically active by charge compensation on the electrodes with perfect symmetry conditions). These modes can have an as detrimental effect as those of the usual symmetrical anharmonics on a filter response.

In figure 4a, we have represented the computed eigen frequencies of the two antisymmetrical modes that exist in the resonators previously considered. It can be easily observed (figure 4b) that if they are electrically active, they can impair greatly the response curve of the filters.

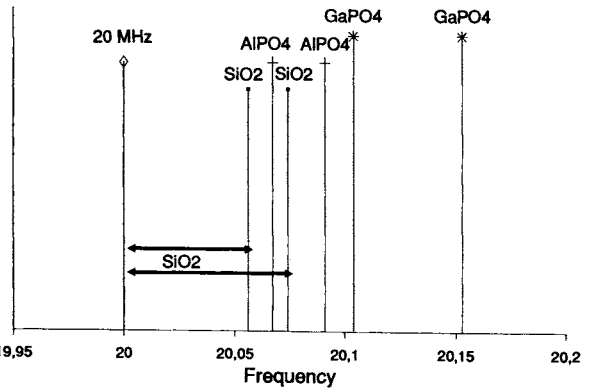


Figure 4a: Eigen frequencies of the antisymmetrical anharmonics.

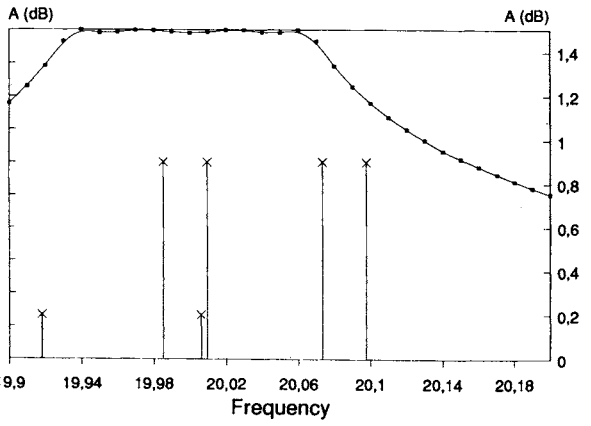


Figure 4b: Position of the anharmonics in the response (berlinite filter).

It was experimentally verified [14] that for small imperfections they appear at frequencies very close to the corresponding eigen frequencies computed for a perfect geometry. So it can be interesting in some practical applications to consider resonators designed so that no antisymmetrical modes can propagate. We have computed for quartz, AlPO4 and GaPO4 20.0 MHz resonators (mass loading= .5%) the electrode dimensions below which they do not appear. For these dimensions rounded down to the nearest 0.1mm (or .05mm for GaPO4) the equivalent scheme where computed and are displayed in table VI, where, we can observe that in these conditions, AT berlinite

seems to give the lowest inductance (see end of §4). The C_0/C_1 capacitance ratio of GaPO4 is still from far the lowest although somehow degraded. Due to the electrode dimensions, the static capacitance takes smaller and smaller values from quartz to GaPO4. These computed values indicate the possibility of a greater sensitivity to stray capacitances for the high coupling quartz analogues in these conditions. The achievable bandwidth computed with the same specifications as previously considered are given in table VII. They are lower than those of table V, as expected, but still sufficient for many applications particularly if a transposition is made to V.H.F. frequencies where such a "conservative" design can be very useful.

Table VI: Equivalent scheme of 20.0 MHz resonators designed so that the no antisymmetrical nor symmetrical anharmonics can propagate.

(max. diameter to suppress also the antisymmetrical anharmonics)

material / diameter mm	motional inductance L (mH)	motional capacitance C1 (fF)	static capacitance Co (pF)	Co/C1
Quartz 1.6	13.66	4.612	0.969	210
AlPO4 1.3	10.72	5.879	0.786	133
GaPO4 0.75	10.89	5.766	0.380	65.9

Table VII: Achievable bandwidth for 4 poles Jaumann filters using the resonators of table VI.

Max B.W. with suppression of the antisym. anharm.

	AT Quartz	AT Berlinite	AT GaPO4
B.W. (kHz)	39.75	50.60	50.1

4 Comparison of computed and experimental results:

The good accuracy of the model was previously verified for quartz for many resonator designs [12], [14]. For berlinite, without correction of the constants, the agreement between computed and experiment results was observed to be reasonably good for the frequencies and only fair for the equivalent scheme [12]. A much better agreement is obtained using the values of table I in the computations, the other rotated constants included in M and P being computed from the published data. In Figure 5a the response curve of a 20.523 MHz berlinite resonator having a thickness of 71.4 microns (+-.1) silver electrodes of 1.8mmm diameter and a mass loading of approximatively .5% (+.2%,-1%) is represented. The experimental inductance is 4.56 mH (the mean experimental value for 4 similar resonators is:4.59 mH); the computed frequency is 20.543 MHz and the computed inductance is 4.73mH. In figure 5b the response curve of a resonator very similar to those of figure 5a is represented. In this case, a condition (very slight unparallelism) exist permitting the excitation of antisymmetrical anharmonics.

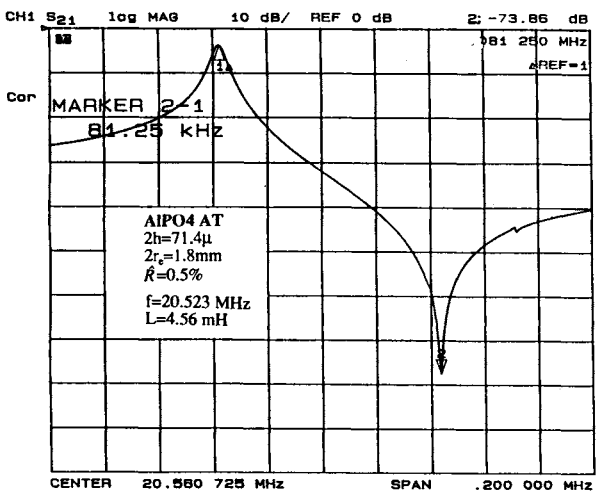


Figure 5a: Response curve of a 20MHz berlinite resonator.

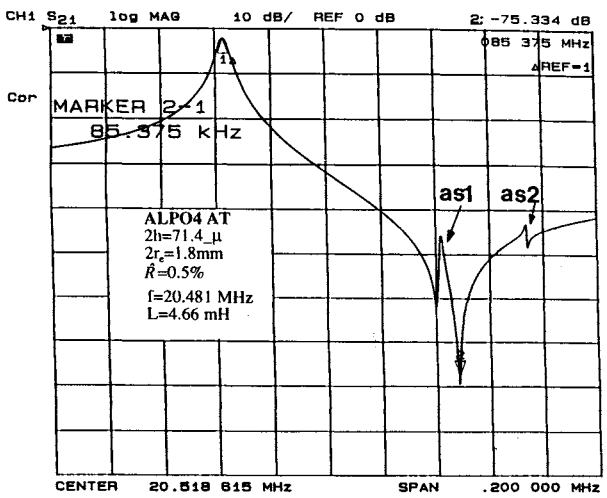


Figure 5b: Response curve of a 20 MHz berlinite resonator with antisymmetrical anharmonics.

For GaPO4 only resonators having fundamental frequencies in the range 5.0 to 7 MHz were studied. The response curve of a near AT GaPO4 resonator is represented in figure 6. The experimental parameters are: thickness= 208.97 microns, electrode diameter : 3.5mm, mass loading: 0.9%, resonance frequency: 5.9952 MHz, measured inductance: 12.79 mH (the mean experimental inductance measured with 5 similar samples is 13.16 mH). The computed frequency is 6.0035 MHz and the computed inductance is 16.329 mH. In the calculation we have corrected the constants by the measured values of table I (without correction the computed frequency is 5.927158 MHz and the computed inductance is 16.26 mH). The slight discrepancy between the experimental and computed values of the inductance indicates that at least one parameter of table I and III is not accurate enough (we suspect that the coupling coefficient k_{26} may still be somehow larger than the value presently measured and given in table I). This difference indicates also that the computed values above given underestimate some how the actual possibilities of AT GaPO4 resonators.Therefore,to obtain a more accurate evaluation of the possibilities of these resonators, the computed values, previously given of the inductance and bandwidth are to be divided (respectively multiplied) by a factor of about 1.2 .

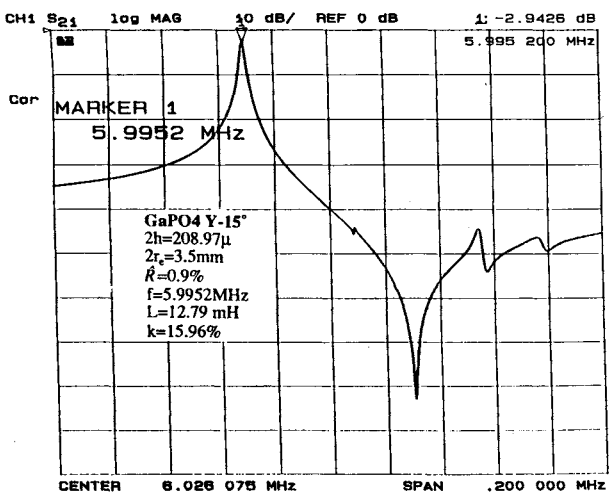


Figure 6: Experimental response of a near AT GaPO4 resonator.

5 Discussion

The first observation resulting of the computations (corrected as described above for GaPO4) is the fact that the quartz like materials have very interesting properties for filtering and

frequency generation applications (thermal stabilities, achievable bandwidth or shift of oscillators). With the first design considered in § III , the AT resonators using GaPO4 would lead to filters having very important bandwidth and to oscillator with large shifts.

The second observation is to notice that, as previously shown [12] [15] for berlinite, much more interesting designs than those using circular electrodes can be found for the devices that lower the inductance values, allows larger bandwidth for the filters and reduce the influence of stray capacitance in circuits. One of them, the use of electrode geometries respecting the lateral anisotropy is particularly efficient with materials presenting a large value of $\sqrt{\frac{M_n}{P_n}}$ and is very simple to apply. The principle of such new designs is recalled in figure 7.

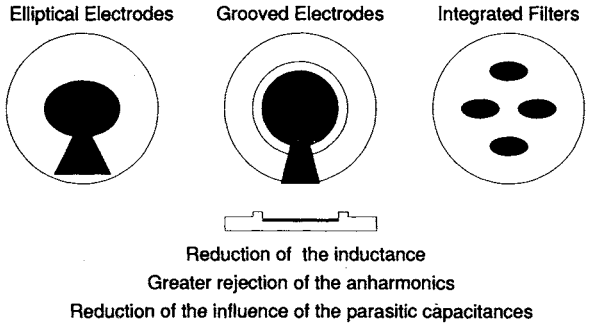


Figure 7: Principle of new designs for resonators and filters.

Conclusion

The properties of AT resonators made with quartz, berlinite and GaPO4 were compared using calculations made with a 3 dimensional model that takes in account the energy trapping phenomena existing in real resonators. These computations, reasonably verified by the experiments, have shown that the high coupling quartz like materials lead to very interesting properties in terms of inductance and capacitance ratio. The AT resonators using berlinite and GaPO4 permits to obtain much larger filter bandwidth or oscillator shifts than AT quartz. Berlinite and even more GaPO4 can lead to devices with much smaller dimensions than quartz. This is very favourable for miniaturization. However, in many cases it is probably wise to use new types of designs permitting to lower the inductance and the sensibility to stray capacitances in increasing somehow the electrode area.

For several applications, (for example: VHF filters with suppresion of the antisymmetrical anharmonics) the use of berlinite appears to constitute an interesting compromise. For the applications requesting the largest bandwidth or oscillators shifts, the use of gallium phosphate appears to be very promising. The three materials appears to lead to very favourable and similar 3rd order thermal variations of the resonance frequencies that are centered at different inflexion temperatures.

Acknowledgements: The authors acknowledge the assistance of MM. Caumont, LeThuault and Mrs Daniel for the elaboration of the studied samples. They are indebted to M.P.Amstutz for his synthesis programs and to M. Pélissié for helpful contribution.

References:

[1] E.D.Kolb, R.A. Laudise, J. Crystal Growth vol.43 p133 (1978).
[2] B.H.T. Chai, M.L.Shand, E.Bueler, M.Gileo Proc. IEEE Ultrasonics Symp. p.577 (1979).
[3] H.Poignant, L. Le Mareschal, Y.Toudic, Mat.Res.Bull. vol. 14 p.603-612 (1979).
[4] E.Philippot, A.Goiffon, M.Maurin, J.Détaint, J.Schwartzel, B.Capelle, A.Zarka. J.Crystal Growth v.104 p.713-726 (1990).
[5] J.Détaint, A.Zarka, B.Capelle, Y.Toudic, J.Schwartzel, E.Philippot, J.C.Jumas, J.C.Doukhan proc. 40th Annual Frequency Control Symp. p.101,114 (1986).
[6] X.Buisson, A.Zarka, R.Arnaud, B.Capelle, Y.Zheng, J.Schwartzel, J.Détaint, Y.Toudic proc. 6th European Frequency and Time Forum (these proceedings) (1992).
[7] G.F.Engel, P.W.Krempl, J.Stadler Proc. 3rd European Frequency and Time Forum p.50-56 (1989).
[8] P.Krempl, J.Stadler, W.Wallnöfer, W.Ellmeyer, R.Selic. Proc. 5th E.F.T.F. p.143,147 (1991).
[9] F.Huard These Université de Montpellier 27 Sept 1985
[10] E.Philippot, AIbanez, A.Goiffon, B.Capelle, A.Zarka, J.Schwartzel, J.Detaint. Proc. 6th EFTF (these proceedings) (1992).
[11] B. James Proc.42nd Annual Frequency Control Symp. p.146-154 (1988)
[12] J.Detaint, H.Carru, J.Schwartzel, C.Joly, E.Philippot, Proc. 5th.European Frequency and Time forum p.30-40 (1991)
[13] D.S Stevens, H.F.Tiersten J.Acoust.Soc. Amer.vol.79 n°6 p.1455 (1976)
[14] J.Détaint, J.Schwartzel, C.Joly, E.Philippot Proc. 41st Annual Frequency Control Symposium p.314-324 (1987)
[15] J.Détaint, H.Carru, J.Schwartzel, C.Joly, B.Capelle, A.Zarka. Proc. 43rd Annual frequency Control Symposium p.563-573 (1989).

Session 6B

Time and Frequency: GPS and GLONASS Activities

Chairman: R. Lucas

RUBIDIUM ATOMIC FREQUENCY STANDARDS FOR GPS BLOCK IIR

W. J. Riley
EG&G Frequency Products
Salem, MA 01970 USA

ABSTRACT

EG&G, Inc. has been selected to provide the Rubidium Atomic Frequency Standards (RAFS) for the GPS Block IIR NAVSTAR satellites. These satellites will replenish and upgrade the space segment of the Global Positioning System in the mid 1990s. The EG&G GPS RAFS is the latest generation of these high-performance rubidium frequency standards. They 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 Block IIR version of these devices will have improved performance, higher reliability, smaller size, and greater radiation hardness.

The GPS Block IIR atomic clocks have a "natural frequency" configuration whereby they output a frequency of about 13.4 MHz that is a submultiple of the atomic resonance. The RAFS operates at a low, fixed C-field for increased stability. The unit has been repackaged into a smaller 4.6″ x 8.5″ x 5.8″ outline, but is somewhat heavier (12 lbs.) because of additional radiation shielding. Elimination of the ground tuning logic and the secondary loop synthesizer (with its ovenized crystal oscillator) has reduced the RAFS complexity and improved its reliability to 0.80 for the 7.5-year mission. The RAFS power consumption is only 13 W at +20°C in vacuum.

This paper describes the GPS Block IIR RAFS design, including the changes and improvements made, the test results obtained, and the progress being made as this program enters the production phase.

Keywords: Rubidium, frequency standard, atomic clock, GPS.

1 INTRODUCTION

The GPS Block IIR NAVSTAR satellites will replenish the Rockwell Block II satellites currently being deployed. General Electric Astro Space Division is the prime USAF contractor for this program, ITT Aerospace/Communications Division is the navigation payload subcontractor, and EG&G has been selected to supply the rubidium atomic frequency standards. A major new feature of the Block IIR satellites is crosslink ranging that can be used to exchange clock correction data between satellites without ground segment control, thus providing a high degree of autonomy. Information regarding the EG&G GPS RFS design and performance has been previously reported[1,2,3,4]. This paper updates that information with the Block IIR changes and recent test results[7].

2 REQUIREMENTS

The requirements for the GPS Block IIR Atomic Frequency Standards are given in the USAF Space Division Technical Requirements Document[5] and the ITT Specification[6]. A summary of those requirements is shown in Table 1. The navigation performance of the GPS system depends critically on the stability of the atomic clocks on board the NAVSTAR satellites. The rubidium frequency standard chosen for the GPS Block IIR application is the high performance unit developed at EG&G especially for this application. This clock has the highest stability of any available device, while also having the practical size, weight, power, reliability, and life advantages of an RFS.

The stability requirements for the Block IIR atomic clocks are shown in Table 2, along with the range error predicted from the measured frequency stability of the EG&G GPS RFS. This high-performance rubidium clock is able to meet the stringent 1- and 14-day autonomous timing requirements.

3 BLOCK DIAGRAM

The most important new design feature of the Block IIR RAFS is the "natural frequency" approach to integrating the clock and secondary loop synthesizer into a Total Navigation Package. The same synthesis hardware that provides frequency adjustments and selective availability also converts the Cs or Rb AFS frequency to 10.23 MHz for the navigation payload. This concept eliminates the RFS secondary loop and ground tuning sections, reducing size, weight, power, and complexity, while improving performance, radiation hardness, and reliability.

A block diagram of the RAFS is shown in Figure 1. A voltage controlled crystal oscillator (VCXO) provides the output via an amplifier and crystal filter. The VCXO also excites the Rb physics package via an rf multiplier chain. A servo amplifier processes the discriminator signal from the physics package to lock the VCXO to the Rb resonance. A lamp exciter, C-field source, and temperature controllers support the physics package operation. A power supply operates the unit, while monitor signals indicate the RAFS status. This arrangement is a simplified "natural frequency" configuration that eliminates the secondary loop synthesizer and ground tuning interface by operating the unit at fixed, low C-field and outputting an exact submultiple of the Rb atomic frequency.

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

4 PHYSICS PACKAGE

The GPS Block IIR RAFS physics package design is based on classical rubidium gas cell frequency standard principles. This approach gives the highest performance consistent with allowable size. The discrete isotopic filter cell gives zero light shift (ZLS) over a range of light intensity and a high S/N ratio. This permits operation at a relatively low light level, thus reducing temperature and rf power dependencies. The characteristics of the lamp, filter, and absorption cells are described in Reference 1. The lamp operates in the Kr-Rb mixed mode and is excited with 0.45 W of rf power at 105 MHz. The lamp output is free from self-reversal. Lamp life is assured by tight heat sinking, alkali resistant glass, high vacuum processing, small envelope surface area, and calorimetric measurement of rubidium fill.

The filter cell operates in a separate oven whose temperature is adjusted for ZLS. The filter cell buffer gas and pressure are chosen to give optimum signal and lowest temperature coefficient (TC).

The absorption cell buffer gases are chosen for narrow linewidth and low TC. The mix ratio is adjusted for a slightly positive TC to partially compensate for the filter cell TC. The nominal fill pressure is determined by the 13.40134393 MHz output frequency, and is sufficient to reduce the wall relaxation rate without excessive buffer gas collisional broadening. The absorption cell length is optimized for maximum signal at the chosen light intensity and temperature, the latter being the coolest practical for operation at a baseplate temperature of +50°C.

A cross-sectional view of the GPS Block IIR RAFS physics package is shown in Figure 2. The three ovens are supported by a fiberglass structure that makes effective use of the vacuum environment for thermal insulation. The low thermal conductivity mounting and leads, and low emissivity ovens, reduce heat losses so that the entire physics package requires only 1.8 W of oven power under normal operating conditions.

A lens is used to collimate the lamp output and provide a uniform light distribution. A two-section C-field coil configuration is used to provide a uniform magnetic bias field. The absorption cell is well-bonded to its oven to avoid temperature gradients, and sapphire oven windows are used for high thermal conductivity. These features help to ensure a homogeneous absorption line. Precisely registered, thin, double-layer foil oven heaters are used to provide a low residual magnetic field.

The microwave cavity is excited with an E-probe and has slotted end covers that support the desired TE_{111} mode while allowing light transmission without significant microwave leakage. The SRD multiplier is mounted against the cavity and shares its stabilized thermal environment. The physics package includes two 0.025 inch Hipernom cylindrical magnetic shields.

5 ELECTRONICS

The RAFS electronic circuits determine, to a large extent, the overall frequency stability that is achieved.

5.1 Temperature Controllers

The temperature controllers are dc thermistor bridges and dissipative regulators, with static thermal gains of about 2000. The oven demand power is determined by the 1-hour warmup requirement. Vacuum conditions not only reduce oven losses, but also raise the thermal gains and servo stability margins.

5.2 Lamp Exciter

The lamp exciter is a Clapp rf power oscillator with the Rb lamp located inside the series-tuned coil. The lamp network presents a range of loads as a function of lamp mode, and the circuit has been characterized under these various conditions to ensure proper operation. Starting may require the exciter to redistribute condensed rubidium inside the lamp by rf induction heating. Ignition takes place when sufficient voltage exists across the lamp coil. The running condition is stabilized against environmental changes and supply ripple by a current regulator circuit.

5.3 Preamplifier

The photodetector preamplifier is a one-stage configuration with dc and ac transimpedances of 100 kΩ and 5 MΩ, respectively. The ac gain is broadly peaked at the fundamental modulation frequency. The signal-to-noise ratio (S/N) and stability of the RAFS are determined primarily by the strength of the Rb discriminator signal and the level of white frequency noise. The latter is predominantly due to shot noise at the photodetector, and VCXO phase noise at twice the servo modulation frequency (see below). The dominant source of preamplifier noise is op amp input voltage noise flowing through the shunt resistance of the photodetector, and, because of the low leakage of the EG&G photodiode, this noise contribution is small.

S/N data for the GPS Block IIR RAFS are as follows:

DC Photocurrent	100 μA
Shot Noise	5.7 pA/$\sqrt{\text{Hz}}$
Op Amp Noise	0.9 pA/$\sqrt{\text{Hz}}$
Total Preamp Noise	5.8 pA/$\sqrt{\text{Hz}}$
Preamp Noise PSD	$5.8\text{x}10^{-24}\text{Hz}^{-1}$
VCXO Noise PSD	$5.8\text{x}10^{-24}\text{Hz}^{-1}$
Discriminator Signal	240 pA per $1\text{x}10^{-10}$
Predicted Stability	$2.4\text{x}10^{-12}\tau^{-1/2}$
Measured Stability (Including reference noise)	$2.8\text{x}10^{-12}\tau^{-1/2}$

5.4 Magnetic Bias Supply

The magnetic bias supply is an active current source that operates the Rb physics package at a low, fixed C-field to minimize its magnetic dependence. The reference voltage for the C-field supply is derived from a precision low-TC, rad-hard zener diode.

5.5 RF Chain

The rf multiplier chain consists of a 13.40134393 MHz VCXO, a phase modulator, a diode tripler, a push-push doubler, and a x85 step recovery diode (SRD) multiplier. The low noise of the Rb reference permits a tight lock loop that allows the use of a low-complexity, nonovenized crystal oscillator. An ALC loop is used to maintain constant drive to the SRD multiplier. This approach gives high stability and spectral purity with minimum complexity. Particular care is taken to avoid frequency offsets caused by AM and PM modulation distortion. A pure modulation waveform is generated by passive integration of a precision squarewave, and highly linear phase modulation is obtained by applying small excursions to a hyperabrupt varactor diode in an all-pass phase modulator configuration.

VCXO phase noise at $2 \cdot f_{mod}$ from the carrier (where f_{mod} is the servo modulation frequency), 146 Hz, can degrade the overall RAFS stability[8]. The measured value of the VCXO phase noise, $\mathcal{L}(292\text{Hz}) = -142$ dBc/Hz contributes a non-negligible but tolerable $1.2\text{x}10^{-12}\tau^{-1/2}$ to the RAFS instability.

A new feature of the Block IIR design is the inclusion of a high-Q helical resonator rf bandpass filter ahead of the SRD multiplier. This greatly improves the spectral purity and helps to avoid frequency offsets due to spurious components on the microwave spectrum. The SRD multiplier uses a low capacitance diode in a shunt mode with an external dc bias resistor that is used to adjust the microwave power level to the optimum value.

5.6 RF Output Amplifier

The rf output amplifier must not only provide the +18 dBm, low-distortion RAFS output, but must also maintain a phase-continuous output under transient radiation. During transient radiation, the active devices of the VCXO and output amplifier are subject to upset. Short-term ($\leq$1 μsec) flywheeling is provided by the LC output tank. Medium-term ($\leq$100 μsec) flywheeling is provided by a crystal bandpass filter ahead of the output stage. The output stage recovers before the energy in its tank is dissipated. The VCXO crystal resonator itself keeps vibrating during the radiation, and the VCXO circuit recovers within the ringing interval of the crystal filter.

5.7 Power Supply

The power supply consists of an input filter, a dc/dc converter, and three linear regulators. The dc/dc converter uses a single-ended flyback configuration that provides both dc isolation and regulation. The +28 V output is used directly as the heater supply while the +5 and $\pm$15 V supplies are further stabilized by precision linear regulators before supplying the logic and analog circuits. Under normal operating conditions in vacuum, the RAFS circuits consume under 10 W. The total dc input is 33 W demand during warmup and 13 W steady-state at +20°C baseplate.

6 MECHANICAL PACKAGING

The RAFS outline is shown in Figure 3. The height was reduced by about one inch by eliminating the secondary loop and ground tuning circuitry, and particularly by the removal of the large ovenized secondary loop crystal oscillator. The length and width were increased slightly because of the addition of tantalum covers on the four sides for radiation shielding. The number of mounting feet was increased from four to eight, and the command/monitor connector was changed to a smaller size.

6.1 Packaging Layout

The RAFS packaging is shown in Figure 4. The outer magnetic enclosure and the side and top covers have been removed in the right photograph to expose the electronic boards and other subassemblies in cavities of the chassis. The primary objectives in this packaging design are ruggedness, thermal transfer, access for assembly and inspection, EMI shielding, radiation shielding, and light weight.

The main chassis provides the thermal and structural interface for the various assemblies and printed circuit boards (PCBs). It is a one-piece, machined aluminum box-like structure with a solid baseplate and four vertical walls. The walls have integral stiffening ribs/partitions that form compartments on the outer surfaces. There are ten PCBs mounted in these compartments. The boards are located around the structure to minimize the length of wiring runs and crosstalk. The partitions between the compartments stiffen the wall, provide EMI shielding between the PCBs and provide a mounting surface for heat sinks for components with high power dissipation.

The four walls of the chassis form a central cavity that contains several assemblies. The physics package is mounted at the bottom on the baseplate of the structure. The lamp exciter and helical resonator assemblies are attached to the physics package before installation into the structure. The electrical connections are made with nonmagnetic semi-rigid coaxial cables. The rf output connector is located on the structure wall so that it protrudes directly into the output amplifier compartment.

Filter boxes are mounted at the power and monitor connectors to eliminate radiation to or from exposed leads. EMI gaskets are used under the connector flanges. All leads pass through feedthrough filter capacitors, and the assembly is located near the power supply to minimize exposed lead length. There are also feedthrough filter capacitors in partitions in the chassis where needed for EMI suppression. The electronic compartments have aluminum covers fastened against the chassis with closely spaced screws for containment of electromagnetic radiation. The four sides and the top also have covers of tantalum for nuclear radiation shielding. An outer magnetic enclosure is used to provide additional magnetic shielding.

7 RADIATION HARDENING

Radiation hardening is a critical aspect of the GPS Block IIR RAFS requirements, and one that has received much attention during the design effort. The RAFS design has undergone detailed radiation hardening analysis by GE, with the assistance of EG&G, GE, ITT, and JAYCOR. Critical parts and subassemblies have also been subjected to radiation testing. In particular, the Rb physics package has been subjected to total gamma dose testing, and the complete RAFS has been subjected to transient radiation testing. The analysis of the RAFS response to transient radiation used the Kenschaft model for the Rb physics package[9]. This analysis and test have shown that the design can meet the stringent clock performance requirements under both the natural and manmade radiation environments.

8 RELIABILITY

The predicted MTBF of the overall GPS Block IIR RAFS is 298,000 hours, and 348,000 hours excluding those components (such as monitors) that do not contribute to a mission failure. This corresponds to reliability values of 0.80 and 0.83, respectively, for a single Rb clock over the 7.5-year mission duration. Each satellite is expected to have one active and one hot-standby rubidium clock, and one cesium clock as a cold-standby unit.

9 TEST RESULTS

Two more GPS RFS prototype units (S/Ns 3 and 4) were built as part of the GE proposal effort. These units have the original mechanical configuration, but incorporate the changes associated with the 13.4 MHz "natural frequency" approach. In addition, a Brassboard model of the Block IIR design (see Figure 4) was built and subjected to qualification-level testing. All the EG&G GPS RFS/RAFS units show excellent performance that exceeds the Block IIR requirements.

9.1 Stability

All of these RFS units have consistently shown drift-corrected Allan variance stabilities at or below $\sigma_y(\tau) = 3\text{x}10^{-12}\tau^{-1/2}+1.5\text{x}10^{-14}$ when measured against the NRL hydrogen maser. One-day stabilities in the high 10^{-15} range are typical. An example of a recent NRL 107-day stability run at constant temperature for S/N 3 is shown in Figures 5–7[10]. The stability is about 10^{-14} between 10^5 and 10^6 seconds. The higher level of white FM at shorter averaging times is due to the measuring system.

9.2 Aging

The GPS RFS units have shown aging $\leq 5\text{x}10^{-14}$/day. The aging is always negative, tends to decrease gradually with undisturbed operating time, and is consistently smooth and highly modelable.

9.3 Temperature Stability

The average TCs of S/Ns 3 and 4 were measured at EG&G to be about $-0.7\text{x}10^{-13}/°\text{C}$ and $+0.4\text{x}10^{-13}/°\text{C}$, respectively, over a baseplate temperature range from +20°C to +45°C. S/N 4 was then re-optimized for operation from +15°C to +20°C, where it displayed an average TC of about $+0.6\text{x}10^{-13}/°\text{C}$. The RAFS Brassboard unit has a TC of $-0.8\text{x}10^{-13}/°\text{C}$. The steady-state TC is always smooth, hysteresis-free, and quite modelable.

A test was conducted to measure the clock error caused by orbital temperature variations. S/N 4 was subjected to a +15°C to +20°C triangular temperature profile having a 12-hour period. The resulting time error was about ±2 nsec, well within the operational requirements.

9.4 Power

The steady-state power versus temperature characteristic of EG&G GPS RFS S/N 4 was measured as a function of baseplate temperature from −15°C to +35°C. The power was 12.6 W at +20°C and varied by about −50 mW/°C.

9.5 Retrace

A series of power on/off retrace tests were conducted on S/N 1 at NRL. This 58-day test included 4 retrace cycles, with off periods from about 2 to 10 days. This retrace record is unique because it is the only such record known for an RFS of such high stability against such a high stability reference. The measurement noise is therefore very low, even in relation to the low retrace values. The retrace record is also interesting in that the RFS frequency seems to "freeze" during the off period. This indicates that the dominant aging mechanism depends on elevated internal operating temperature. The most likely such mechanism is redistribution of rubidium within the lamp. This process would indeed literally "freeze" while the unit was off.

The retrace behavior of the unit was excellent, ranging from essentially zero (for 2 days off) to about $3\text{x}10^{-13}$ (for 10 days off). These results are consistent with similar retraces conducted on GPS RFS S/N 2 as part of a total gamma dose physics package radiation test (see below). This RAFS design easily complies with the $5\text{x}10^{-12}$ retrace requirement, and the warm-up time required to achieve this retrace is less than the 3 days specified.

9.6 Qualification-Level Tests

A series of qualification-level tests, including shock, random vibration, temperature, temperature cycling and EMI, were successfully conducted on the RAFS brassboard unit. All requirements were met[11].

9.7 Radiation

A total gamma dose radiation test was conducted on the RAFS physics package. The objective of the test, conducted jointly by EG&G Frequency Products and GE Astro Space, was to determine the effects of total dose gamma radiation on the frequency, photodetector current, and other characteristics of the high-performance Rb physics package. Most of all, the test was run to rule out any unexpected effects. Measurements were made before and after doses of 7, 21, and 70 krads(Si) of gamma radiation from a Co^{60} cell. No significant frequency shifts or other changes in performance were observed within the limit of about $1\text{x}10^{-11}$ set by the drift and retrace effects

(primarily due to the disassembly and reassembly of the RAFS before and after each irradiation). Since these exposures represent x1, x3, and x10 of the total dose expected from natural radiation during the 7.5-year life of the GPS satellite, during which the RAFS can be expected to have a total frequency shift of about $1x10^{-10}$, this test has confirmed that the effect of the radiation on the Rb physics package is negligible.

The Block IIR RAFS brassboard unit was subjected to a series of transient radiation tests at the GE Flash X-Ray (FXR) facility. The silicon photodiode in the Rb physics package is a sensitive radiation detector which, when exposed to a transient radiation pulse, can overload the servo amplifier circuits and cause a large disturbance to the VCXO. This behavior can be simulated by exposing the photodetector to a light pulse from a xenon flashlamp. A means was devised to disable the servo circuits until they recovered, thus avoiding any significant time error. The wide servo bandwidth ($\approx$ 45 Hz) assures rapid correction of the VCXO frequency error. Excellent agreement was obtained between the analysis, flashlamp simulations, and the actual FXR tests.

10 PROGRAM STATUS

The development phase of the GPS Block IIR RAFS program is complete, and the program is entering its production phase. Brassboard testing is complete, and fabrication of an Engineering Development Model is underway. GPS Block IIR Rb clock deliveries are scheduled to begin in July 1993.

11 ACKNOWLEDGMENTS

The recent work on this EG&G high performance rubidium frequency standard has benefited from the contributions of many persons both within and outside the EG&G organization. Prominent among the former are S. Goldberg, T.J. Lynch, K.D. Lyon, and J.R. Vaccaro. Among the latter, the author wishes to acknowledge radiation hardening consultation by D.H. Swant of GE, and systems engineering by A.P. Baker of ITT. We also wish to thank the NRL Space Applications group for their continued test support.

12 REFERENCES

1. W.J. Riley, "A Rubidium Clock For GPS," *Proceedings Of The 13th Precise Time And Time Interval (PTTI) Applications And Planning Meeting*, pp. 609-630, December 1981.
2. T.J. Lynch and W.J. Riley, "Test Results For GPS Rubidium Clocks," *Proceedings Of The 15th Precise Time And Time Interval (PTTI) Applications And Planning Meeting*, pp. 269-280, December 1983.
3. S. Goldberg, T.J. Lynch and W.J. Riley, "Further Test Results For GPS Rubidium Clocks," *Proceedings Of The 17th Precise Time And Time Interval (PTTI) Applications And Planning Meeting*, pp. 145-155, December 1985.
4. F. Danzy and W. Riley, "Stability Test Results For GPS Rubidium Clocks," *Proceedings Of The 19th Precise Time And Time Interval (PTTI) Applications And Planning Meeting*, pp. 267-274, December 1987.
5. "GPS Block IIR Space Segment Technical Requirements Document," CWNZ-87-500, 12/31/87.
6. "Performance Specification, Atomic Frequency Standard (AFS)," ITT Aerospace/Communications Division Dwg. No. 1527179, Rev. C, 8/28/91.
7. W.J. Riley, "Rubidium Atomic Frequency Standards For GPS Block IIR," *Proceedings of the 22th Precise Time and Time Interval (PTTI) Applications and Planning Meeting*, pp. 221-230, December 1990.
8. C. Audoin, V. Candelier and N. Dimarcq, "A Limit to the Frequency Stability of Passive Frequency Standards Due to an Intermodulation Effect," *IEEE Transactions on Instrumentation and Measurement*, Vol. 40, No. 2, pp. 121-125, April 1991.
9. R.P. Kenschaft, *Response of Quantum-Mechanical Resonances to Modulated Radio-Frequency Signals*, Ph. D. Thesis, University of Pennsylvania, 1969.
10. E.D. Powers, Jr. and F. Danzy, "Interim Results from the Characterization Testing of the Engineering Development (EDM) Rubidium Clocks for Satellite Applications", *Proceedings of the 22th Precise Time and Time Interval (PTTI) Applications and Planning Meeting*, pp. 325-330, December 1990.
11. T.J. Lynch, K.D. Lyon, W.J. Riley and J.R. Vaccaro, "Qualification-Level Test Report, GPS Block IIR RAFS," EG&G Frequency Products FP0043, June 1991.

TABLE 1 KEY REQUIREMENTS GPS BLOCK IIR RUBIDIUM ATOMIC FREQUENCY STANDARD (RAFS)	
PARAMETER	REQUIREMENT
Frequency	13.40134393 MHz (Nominal)
Frequency Offset	$\pm 1 \times 10^{-9}$ (Initial Tolerance Limits)
Frequency Adjustment	None (Fixed C-Field)
Frequency Drift	$\pm 5 \times 10^{-14}$/Day (Maximum after 30 Days)
Frequency Stability	$3 \times 10^{-12} \tau^{-1/2} + 5 \times 10^{-14}$ (Maximum Sigma)
Temperature Stability	$\pm 2 \times 10^{-13}$/°C (Average over 0 to +25°C)
Output Power	+18 dBm ±1.5 dB
Harmonics	−50 dBc
Phase Noise	$\mathcal{L}$(1 Hz) = −85 dBc/Hz $\mathcal{L}$(1 kHz) = −90 dBc/Hz
DC Power	50 Watts (Maximum during Warmup) 14 Watts (Maximum SS at +20°C in Vacuum)
Size	5.0″ x 8.5″ x 7.0″ (Maximum)
Weight	12 lbs. (Maximum)
Reliability	0.755 for 7.5 years (Minimum)
Life	7.5 years (Minimum)
Survivability	± 2 nsec (Maximum Prompt Time Error) −2 dB (Maximum Prompt Output Level Change) 2.6 meters (Maximum URE for 24 Hours) 36.3 meters (Maximum URE for 14 Days)
Note: The detailed requirements for the GPS Block IIR RAFS are given in ITT A/CD Dwg. No. 1527179.	

TABLE 2 GPS BLOCK IIR CLOCK STABILITY REQUIREMENTS		
INTERVAL	CLOCK ERROR, nsec	
Days	REQUIREMENT	PREDICTED
1	8.7	2.5
14	121.0	33.0
180	8.7	Dependent On Crosslink Ranging
Note: Predicted time error based on measured drift-corrected RAFS stability		

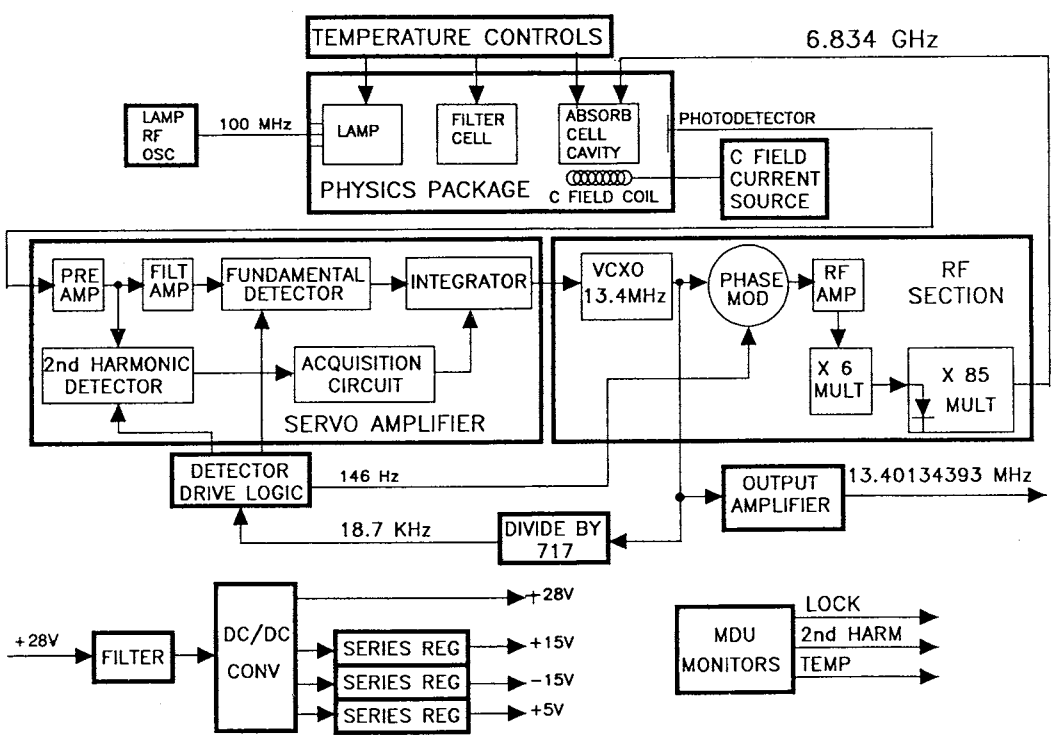


Figure 1. Block Diagram of GPS Block IIR RAFS.

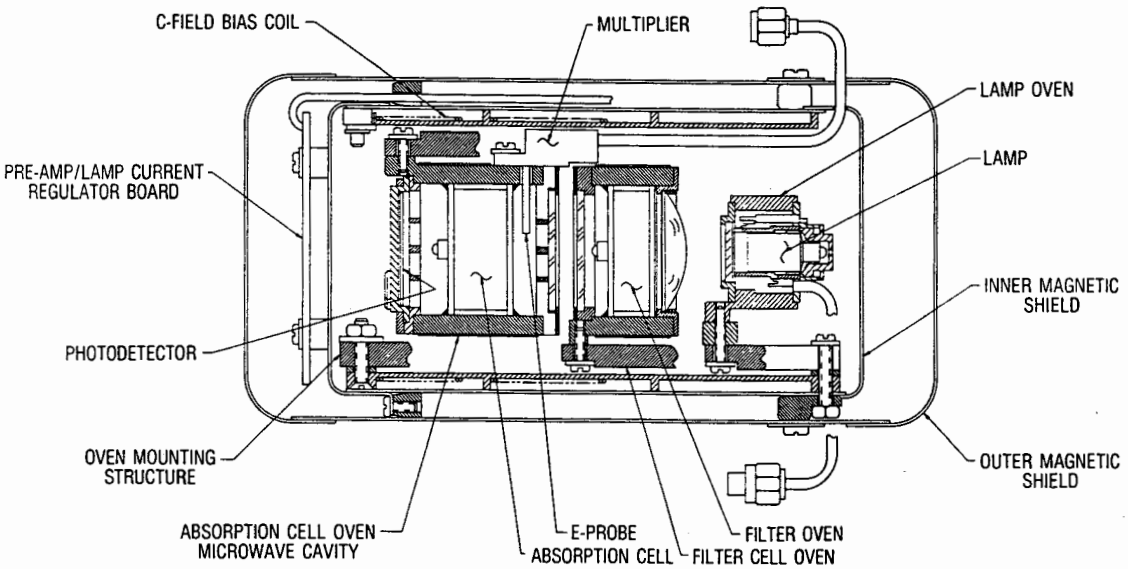


Figure 2. Physics Package Cross-Section.

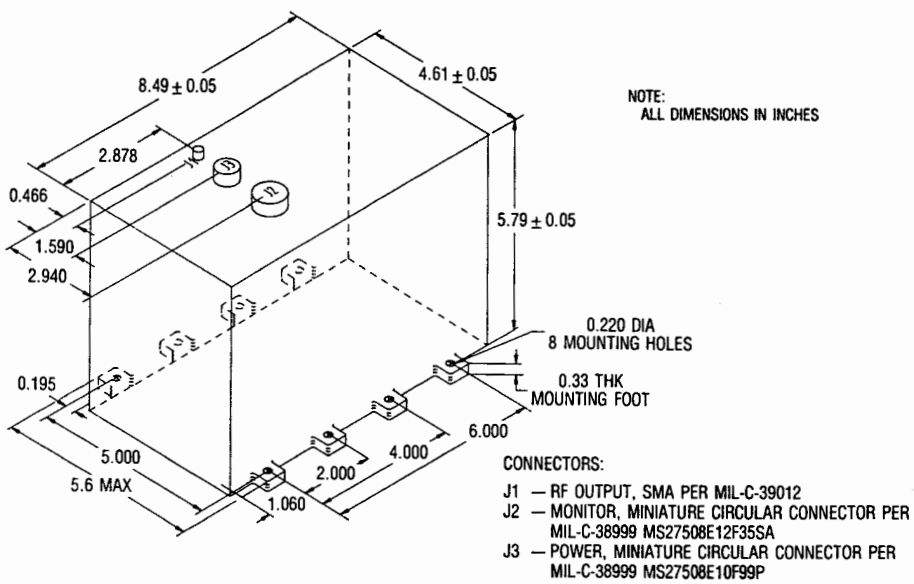


Figure 3. RAFS Outline.

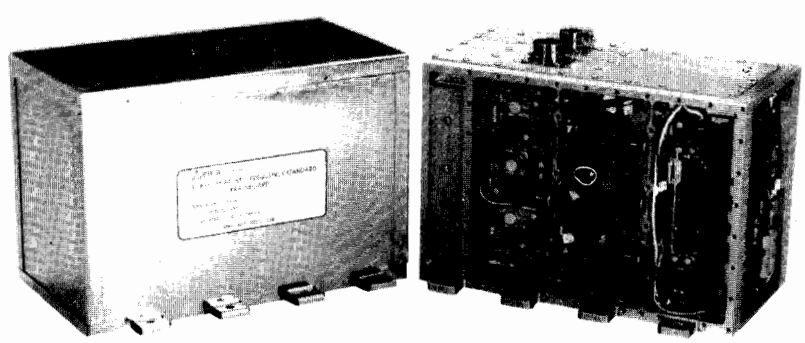


Figure 4. Photographs of GPS Block IIR RAFS.

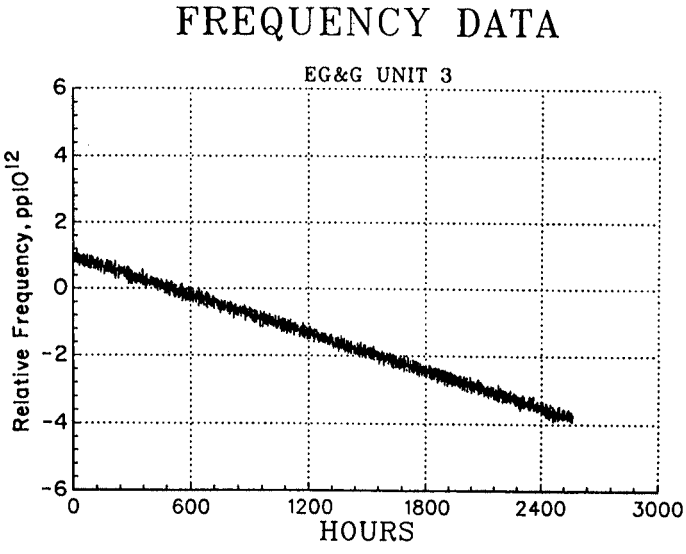


Figure 5. S/N 3 Frequency Record.

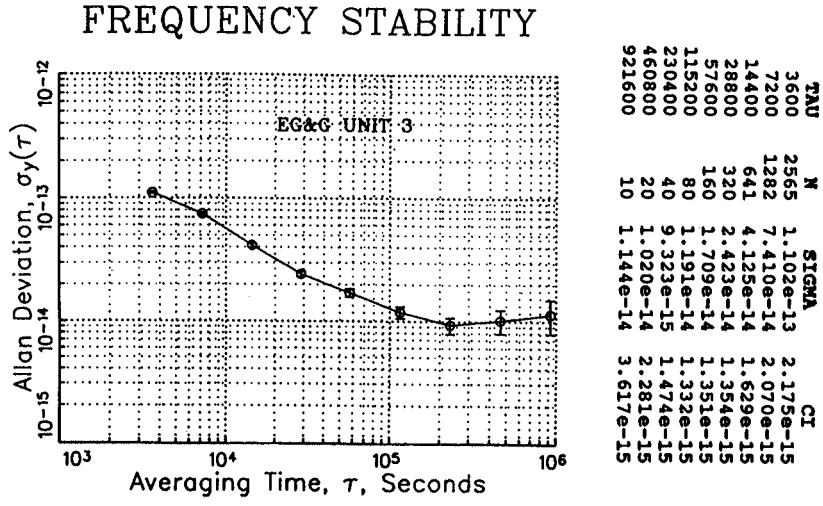


Figure 6. S/N 3 Stability.

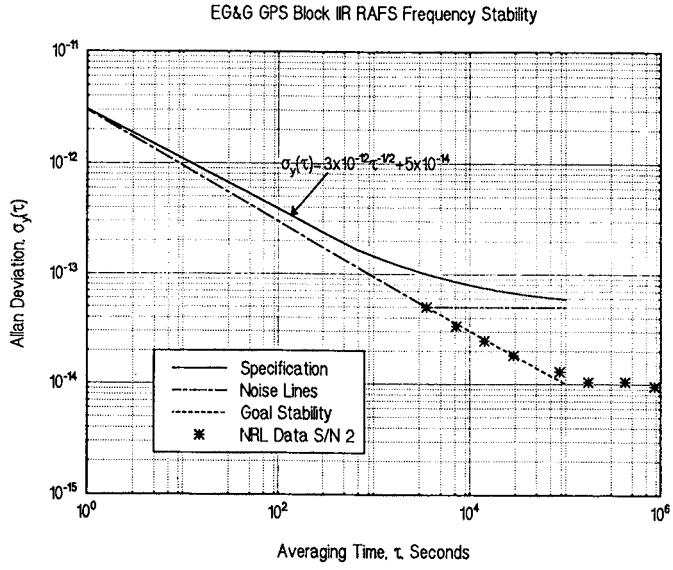


Figure 7. RAFS Stability Requirements.

MIXED CONSTELLATION SYNCHRONIZATION OF GPS TIME

A. L. Satin, C. H. Yinger, W. A. Feess

The Aerospace Corporation
El Segundo, California

ABSTRACT

The Global Positioning System (GPS) will consist of a mixed constellation of Block II/IIA ground-navigated and Block IIR autonomously navigated (Autonav) satellites for many years starting in 1996. The GPS time references for these two types of sub-constellations will have to be synchronized to within a few nanoseconds if user navigation accuracy is to be preserved. A technique for achieving Root Mean Square (RMS) phase and rate synchronization of 2.9 ns and 0.13 ns/hr, by steering the ground system to Autonav, is described and validated by simulation. Also, coincidental with synchronization steering, it is shown that Autonav ensemble time (for a subconstellation of as few as four satellites) can be maintained to within 20.7 ns of Universal Time Coordinated (UTC) using a drift-rate steering control law.

1. INTRODUCTION

1.1 Constellation of Block II/IIAs

The current deployment of Block II and IIA Global Positioning System (GPS) space vehicles (SVs) will culminate in 1993 with the full operational capability of a constellation of 21 SVs and 3 spares. This constellation will be maintained by the Operational Control Segment (OCS), which consists of three co-located monitor stations (MSs) and ground antennas (GAs) at Kwajalein, Diego Garcia, and Ascension Islands; a GA at Cape Canaveral Air Force Station, Florida; two additional MSs at Keane Point, Oahu and Colorado Springs, Colorado; and a Master Control Station (MCS), also in Colorado Springs. MSs track L-band P-code ranging signals and provide pseudorange measurements to the MCS. There, a sub-optimal Kalman filter processes the measurement data to estimate ephemeris, clock phase, frequency, aging, and solar pressure states for each of the 24 SVs and clock phase and frequency states for each of the 5 MS frequency standards. Clock states are represented as offsets to GPS system time or GPS time, for short. For the Block II/IIA constellation, the GPS-time construct is that of an ensemble -- an ensemble of all the clocks in the system (i.e., SVs and MSs) (Ref. 1). This is a direct result of the fact that the filter solves for states for all clocks in the system, i.e., there is no single master clock that is not solved for (Ref. 2).

Periodically, usually once per day in staggered fashion, the MCS builds a navigation data load for each of the SVs and sends them to the GAs for transmission to the SVs. Onboard the SV, the navigation data, consisting primarily of predictions of ephemeris and clock states, is transformed into a serial navigation message, modulated into the L-band signal, and transmitted to the users. From the information in the message and their own pseudorange measurements, users can navigate their position and determine receiver clock offsets from GPS time.

GPS system specifications require that the user achieve 16-meter Spherical Error Probable (SEP) position accuracy and that system time be maintained within 1.0 microsecond of Universal Time Coordinated, UTC (USNO), with the offset known to 100 nanoseconds (3.0 sigma). The position accuracy specification will be met (for all users averaged over all time and space) if the RMS user range error (URE), for each SV over all time, is 6 meters (1.0 sigma). The URE at time t, for an individual SV, is computed as the projection of the navigation message position error on the user line of sight minus the navigation message clock phase error. Error components are computed in the MCS by differencing navigation message position and clock phase from the current Kalman filter estimates of these quantities. The requirement to maintain GPS time to within 1.0 microsecond of UTC necessitates the steering of GPS time to UTC. In this process, a drift-rate steer (or drift-rate command) is computed daily (using a bang-bang control law) from GPS-UTC measurements provided by the USNO. It is applied to the Kalman filter states of all clocks, and so, as it integrates into frequency and phase, it slowly alters the states and effectively moves the ground ensemble to UTC. Since the drift-rate term takes effect at the same time for all clock states (i.e., steering is synchronous) the process is transparent to filter operation (i.e., pseudorange measurement residuals are unaffected). The GPS-UTC requirement is currently being met with a large margin, since the RMS offset is only 10 ns (1.0 sigma). The 100-ns knowledge requirement, routinely monitored by the USNO, is also being met and the RMS of this quantity is also approximately 10.0 ns (Ref. 3).

1.2 Mixed Constellation of Block II/IIA and IIRs

The next-generation GPS SV, the so-called Block IIR, where the R stands for replenishment, will be capable of meeting or exceeding the 6-meter URE requirement while performing onboard autonomous navigation (Autonav) for 180 days without ground intervention (Ref. 4). Autonav simulations indicate that an RMS URE of 2.7m or better may be achievable. The Autonav system allows each SV to determine its own ephemeris and clock states through the application of a UHF Crosslink Transponder Data Unit (CTDU) that supports inter-SV data exchanges and pseudoranging.

Initially, when only a few IIRs are on orbit, they will be operated like IIAs, i.e., with the MCS providing daily uploads. For this situation, the onboard computer is programmed to allow Autonav to run in the background while the SV is being operated more or less like a IIA. In this test mode, daily dumps of Autonav data are used by the MCS to verify/evaluate Autonav performance.

Once a sufficient number of Block IIRs are in orbit (as few as six when the replenishment geometry is favorable), Autonav will meet the 6.0-meter-URE, 180-day requirement and will be declared operational. For some time thereafter, until all Block II/IIAs are replaced by Block IIRs, the 24-SV constellation will contain both types of SVs. During this period, the OCS will track L-band and compute states for all SVs with the clock states being referenced to MCS Kalman filter (or simply ground) GPS time while onboard the Block IIRs, Autonav will compute clock states (from UHF crosslink pseudorange measurements) that are referenced to Autonav GPS time. The latter is an ensemble-time much like the former except it is based only on the Block IIR clocks. User navigation will degrade if ground and Autonav GPS times are not synchronized to within a few (3.0) nanoseconds. Finding ways to accomplish this is the central aim of this work. Section II below discusses simulation results for two techniques (steering and correction) that produce

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

synchronization to the required accuracy. Section III completes the presentation of simulation results by showing how constellation steering of GPS time to UTC may be accomplished coincident with mixed constellation synchronization steering. Section IV summarizes and concludes the paper.

2. MIXED CONSTELLATION SYNCHRONIZATION

Individual phase and frequency synchronization errors are defined for each IIR as the difference between the Autonav phase/frequency estimates and the ground phase/frequency estimates. Control error, used for both the "steering" and "correction" techniques, is the average synchronization error over the IIR constellation. It may be reduced by steering (or correcting) either the ground states to Autonav (i.e., Autonav steering ground) or the Autonav states to ground (i.e., ground steering Autonav).

A six-SV constellation simulation test bed was used to evaluate the performance of the steering and correction techniques. Thirty days of simulated L-band tracking data for five MSs, and UHF pseudorange data for the CTDUs was generated and processed, respectively, by the TRACE/SRIF program (Ref. 5) to generate MCS Kalman filter states, and by the Autonav simulation program to generate Autonav clock states. Both sets of states were input into a program that simulated the synchronization techniques and computed performance parameters and plots. The ground clock in these simulations was assumed to be a master clock with stability of 3.0E-14 and the IIR clocks had a stability of 1.0E-13.

2.1 Synchronization Using the Steering Technique

Case 1 simulates the steering of ground states hourly to the Autonav ensemble using a drift control law as illustrated in Fig. 1. In the figure, ϕ_A and f_A are the Autonav phase and frequency states and similarly for the ground states (ϕ_G and f_G). N is the number of Block IIR SVs. Control gains K_1 and K_2 are functions of the system time constant τ and damping ratio (0.7). As Table 1 indicates, the RMS over time of the phase control error ($\hat{\phi}_\varepsilon$), when the system reaches steady state, decreased at the expense of RMS frequency error ($\hat{f}_\varepsilon$) as the system time constant is decreased. An RMS synchronization error of 2.9 ns can be achieved with a 24-hour time constant. Although steering directly controls only the average synchronization errors were found to be small (3.5 ns RMS). The ground should be steered to Autonav only when there are enough Block IIRs on orbit so that the Autonav ensemble stability is comparable to the ground ensemble stability. Since the stability specification of the Block IIR natural frequency Rubidium oscillators is 5.0E-14 at one day (compared with 2.0E-13 for the Block IIs), this condition can occur with as few as six SVs on orbit. The steering of the ground ensemble to the Autonav ensemble, with a drift-rate control law, was also simulated but gave worse results (4.3-ns error as compared with 2.9).

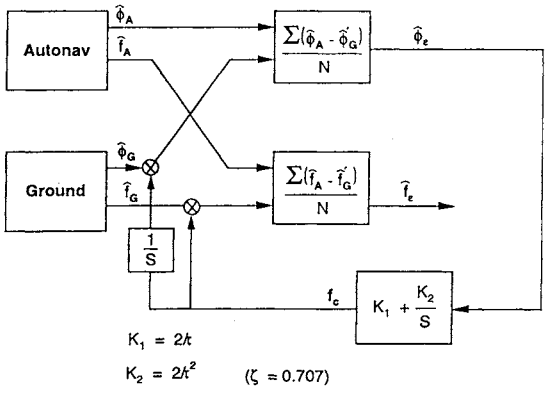


Fig. 1: Case 1, Autonav steers ground, drift control law.

Table 1: Case 1, Drift-control performance (Steer ground, 1-hr delay).

τ (hrs)	RMS ($\hat{\phi}_\varepsilon$) (ns)	RMS ($\hat{f}_\varepsilon$) (ns/hr)	RMS (f_c) (ns/hr)
6	1.2	0.46	0.57
12	2.1	0.30	0.45
24	2.9	0.13	0.32
36	3.4	0.08	0.27
48	4.0	0.08	0.24

Case 2 simulates steering of Autonav states to the ground with a drift control law. In this case, a drift command that must take effect at the same time in all Block IIR SVs (so that Autonav is not perturbed), is uploaded to the SVs along with the time the command is to take effect, i.e., the so-called effectivity time. Since uploads by their very nature are staggered in time, the effectivity time must be far enough in the future to allow time for all the SVs to be uploaded. Therefore there is a delay between the time that commands are computed and when they take effect. This sequence of events is shown in Fig. 2. In the figure, "compute steering command" denotes times when commands are computed and "effectivity time" denotes the times they take effect. The delay between these two times allows for uploading of the SVs. Once in the SVs, the commanded drift will be activated simultaneously at the effectivity time. As Table 2 indicates, steering Autonav to the ground does not work nearly as well as drift control steering of the ground to Autonav, because of the time delay associated with uploading the SVs. Though 1-hour delays were analyzed, the minimum practical delay is 8 hours.

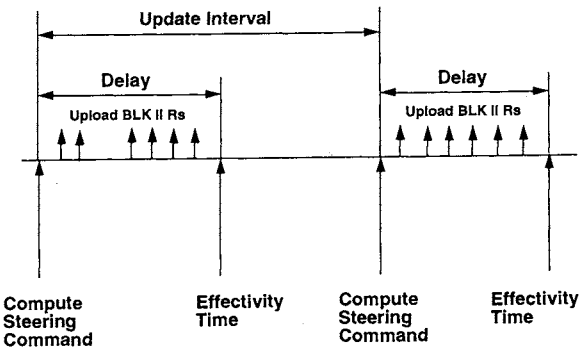


Fig. 2: Autonav steering mode scenario.

Table 2: Case 2, Steer Autonav, drift-control law.

Update Interval (hrs)	Delay (hrs)	τ (hrs)	RMS ($\hat{\phi}_\varepsilon$) (ns)	RMS ($\hat{f}_\varepsilon$) (ns/hr)	RMS (f_c) (ns/hr)
8	1	24	3.9	0.27	0.41
8	8	48	5.8	0.28	0.31
12	1	36	4.4	0.24	0.33
12	8	60	6.2	0.26	0.27
12	12	72	7.1	0.26	0.27
24	1	72	6.5	0.22	0.25
24	8	96	8.4	0.24	0.24
24	12	108	9.3	0.24	0.23
24	24	144	12.2	0.25	0.22

2.2 Synchronization Using the Correction Technique

Case 3 (Fig. 3) applies the correction technique to the Autonav states to achieve synchronization. In this technique, the full synchronization offset is uploaded to all IIRs and subtracted from the Autonav states without being modified by control gains. The sequence of operations for correcting Autonav states to the ground is the same as for steering Autonav states (Fig. 2) except that the word steering is replaced by the word correction. At the "Compute steering (or correction) Command" time, in Fig. 2, synchronization offsets ($\hat{\phi}_\varepsilon$ and $\hat{f}_\varepsilon$) are sampled and predicted forward to the effectivity time. After a delay for the time it takes to upload the Block IIR SVs, the corrections ($\Delta\phi_C$ and Δf_C) are applied to the Autonav states ($\hat{\phi}_A$ and $\hat{f}_A$). The results for this case (Table 3) are better than those obtained for Case 2, e.g., for an update interval of 12 hours and delay of 8 hours this case gives 4.8 ns, while Case 2 gives 6.2 ns. It has been found that the correction technique generally gives better results than steering.

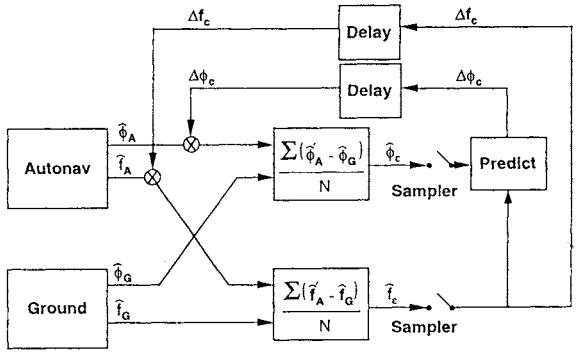


Fig. 3: Autonav correction mode.

Table 3: Case 3, Autonav correction mode.

Phase Only Corrections				
Update	Delay	RMS ($\hat{\phi}_\varepsilon$) (ns)	RMS ($\hat{f}_\varepsilon$) (ns/hr)	RMS ($\Delta\phi_C$) (ns)
8	1	2.3	0.23	3.3
8	8	4.3	0.23	3.3
12	1	3.1	0.23	4.6
12	8	4.8	0.23	4.6
12	12	5.6	0.23	4.6
24	1	4.4	0.23	6.6
24	8	5.7	0.23	6.6
24	12	6.3	0.23	6.6
24	24	7.9	0.23	6.6

3. SIMULTANEOUS CONSTELLATION SYNCHRONIZATION AND UTC STEERING

A program was developed to simulate the simultaneous correcting of ground states for constellation synchronization and drift-rate steering of Autonav to UTC. Any mix of IIA and IIR SVs (up to a total of 24) can be simulated. Fig. 4 is a schematic of the system simulation. The outer loop is the GPS-UTC drift-rate control loop. At the bottom of the diagram is the box labeled "USNO." This is where USNO measurements of GPS-UTC, for individual SVs, are generated. These measurements are then input to the least-squares-fit box labeled "USNO LS FIT" that produces best-fit estimates of GPS-UTC phase ($\hat{\phi}_\varepsilon$) and frequency ($\hat{f}_\varepsilon$). The box is sampled daily and estimates are used to compute drift-rate commands (f_C) as specified by the "steering law" box. These commands are sampled before each Block IIR upload sequence just as was shown in Fig. 2 earlier. After a delay for uploading, the effectivity time will be reached and Autonav clock states as well as ground clock states will begin to be steered in unison. Autonav states are used to steer the ground just as was discussed in case 1. Block IIA states, uploaded from the ground, reflect the effects of both constellation synchronization steering and UTC steering. The magnitude of the drift-rate command may be limited (to 1.5 ns/day/day) by program input. Cases were run with the delay and with and without a limit, and then some cases were run without a delay or limit. These results are shown in Table 4 where N_1 is the number of IIRs, N_2 is the number of IIAs, $\hat{\phi}_\varepsilon$ is the GPS-UTC phase synchronization error, and $\hat{f}_\varepsilon$ is the GPS-UTC frequency synchronization error. As expected, the best performance is obtained for the three cases run "without delay and limits." A zero delay, however, can only be realized with a crosslink upload capability, which is currently not part of the Block IIR baseline design. An RMS UTC synchronization offset of 13.2 ns can be achieved in this case with as few as 4 IIRs. Without the 1.5 ns/day/day limit but with the delay, an RMS error of 20.7 ns is possible with 4 SVs. This kind of performance is well within specification. Without the limit and with 12 or more IIRs, the RMS error is comparable to what is currently being experienced (i.e., approximately 10 to 20 ns).

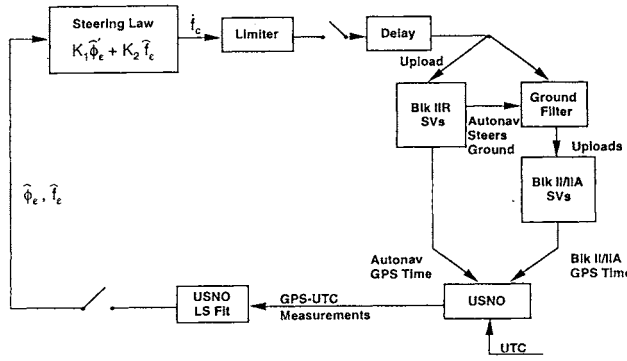


Fig. 4: UTC steering (Autonav steering ground for constellation synchronization).

Table 4: UTC steering performance (RMS in ns and ns/day).

N_1	N_2	RMS $\hat{\phi}_\varepsilon$	RMS $\hat{f}_\varepsilon$	Remarks
20	4	16.0	2.72	With 1.5 ns/day/day limit
12	12	19.8	4.12	
4	20	31.6	6.24	
20	4	9.0	2.79	Without limit
12	12	13.4	3.97	
4	20	20.7	6.69	
20	4	6.2	1.80	Without delay and limit
12	12	8.8	2.36	
4	20	13.2	4.14	

4. SUMMARY AND CONCLUSIONS

With the introduction of the Block IIR Autonav capability and the desire to go operational with a sub-constellation of IIRs imbedded in the Block II/IIA constellation, the need arises to synchronize Autonav GPS time and the Block II/IIA (or ground) GPS time. Using simulation techniques, it was found that steering ground states to Autonav with a drift control law produces RMS synchronization of 2.9 ns, and judging from the results for Autonav correction mode, even better performance may be possible. This additional error has little impact on the overall URE error budget of 6.0 meters (1.0 sigma).

The simultaneous synchronization of ground-to-Autonav and Autonav-to-UTC was also simulated. Constellations with various combinations of Block II/IIA and Block IIR SVs were examined with and without a drift-rate limit of 1.5 ns/day/day and/or a delay of one day or so. All results are well within the current system specification of 1.0 microsecond.

ACKNOWLEDGEMENTS

The authors wish to gratefully acknowledge Dr. Eric Stroud for generating the Autonav state data and Dr. Howard Holtz for his timely support with simulation development. We also appreciate the efforts of Dr. Henry Fliegel and Harold Bernstein, who helped review the paper.

REFERENCES

1. K. R. Brown, Jr. "The Theory of the GPS Composite Clock." Proceedings of the ION GPS-91 Meeting, Albuquerque, New Mexico (1991).

2. A. L. Satin and C. T. Leondes, "Ensembling Clocks of the Global Positioning System," IEEE Trans. on Aerospace and Electronics 26 (January 1990).

3. M. P. Ananda, et al., "Global Positioning System Autonomous Navigation." Paper presented at IEEE Plans Meeting, Las Vegas, Nevada. March 1990.

4. W. A. Feess et al., "Evaluation of GPS/UTC Steering Performance." Paper presented at Precise Time and Time Interval Meeting, Pasadena, California. December 1991.

5. L. A. Campbell, TRACE Trajectory Analysis and Orbit Determination Program, Vol. 13: Square Root Information Filtering and Smoothing, Report No. SSD-TR-91-07, The Aerospace Corporation. 15 March 1991.

GPS STANDARDIZATION FOR THE NEEDS OF TIME TRANSFER

W. Lewandowski, G. Petit and C. Thomas

Bureau International des Poids et Mesures
Pavillon de Breteuil, Sèvres, France

ABSTRACT

It has been clear for some time that a desirable and necessary step for improvement of the accuracy of GPS time comparisons is to establish GPS standards which may be adopted by receiver designers and users. For this reason, a formal body, the CCDS Group on GPS Time Transfer Standards (CGGTTS), was created in 1991. It operates under the auspices of the permanent CCDS Working Group on TAI, with the objective of recommending procedures and models for operational time transfer by the GPS common-view method. It works in close cooperation with the Subcommittee on Time of the Civil GPS Service Interface Committee.
This paper illustrates the need to develop standards in GPS receiver software and hardware, in particular with the implementation of Selective Availability. It also reports the most recent activities of the CGGTTS.

Keywords: GPS time receiver, standardization, Selective Availability

1. INTRODUCTION

In recent years the operational GPS worldwide time transfer by C/A code receivers in common-view mode (Ref. 1) has seen significant progress in both precision and accuracy (Ref. 2). The accuracy of GPS time links within continents now approaches two nanoseconds on an operational basis. Between continents, the accuracy of operational links is between 10 and 20 nanoseconds. Some recent studies, however, have shown that, when using an accurate homogeneous reference frame for antenna coordinates, ionospheric measurements and post-processed precise satellite ephemerides, these long-distance time comparisons can be made with an accuracy of a few nanoseconds (Ref. 3).

But when approaching such a level of accuracy other problems arise. These are mainly due to the lack of standardization in the software and hardware of commercial receivers which, for example, process raw data differently or treat the input signal to the antenna in different ways. There is, in addition, a need to remove the effects of SA degradation of GPS signals.

The first section of this paper presents an analysis of the effects of SA with the example of its implementation since November 1991. We show here the absolute necessity of strict common views and of an unified treatment of short-term data in GPS time receivers.
The second section deals with some other deficiencies of GPS time transfer which could be reduced in the framework of an international standardization. Several examples are given and are illustrated by the diversity of GPS time receiver types now in operation at the BIPM and national centers.
The third section of this paper briefly reports on the actions of the formal bodies concerned with GPS standardization: the CGSIC Subcommittee on Time and the CCDS Group on GPS Time Transfer Standards.

2. IMPLEMENTATION OF SA

The GPS system was designed with optional facilities to degrade the GPS signals available to non-cleared users: 'Selective Availability' (SA) and Anti-Spoofing (AS). The activation of SA and AS makes the GPS Precise Positioning Service (PPS), which contains the full accuracy of GPS, inaccessible to those without encryption keys (authorized users). However, the degraded Clear Access (C/A) 1,023 Mhz code on L1 frequency remains available for all users and provides the GPS Standard Positioning Service (SPS). In case of SA, the SPS has a stated 95% accuracy in two dimensions of 100 meters in position and 167 ns in time (Ref. 4).
According to the information available to the civil community, the degradation brought about by SA concerns only Block II satellites and consists of a phase dither in the satellite clock and a changeable bias in the broadcast ephemerides. For the civil users, SA causes peak-to-peak inaccuracies of several hundreds of nanoseconds in the direct extraction of GPS time from Block II satellites (Ref. 5).
This has been observed since 15 November 1991 with only small periods of interruption (for instance from 22 to 24 February 1992). Figure 1 shows an example of raw GPS data taken from a one-channel C/A code time receiver located at US Naval Observatory (Washington DC, USA): a time modulation as high as 250ns is added to the usual noise affecting the GPS data.

2.1 Computation of strict common views

It is possible to eliminate on-board clock dither in the comparison of remote clocks on the Earth. This supposes that strict common-view observations are available (Refs. 1, 5), that is observations having:

* the same track length (780s),
* the same start time (within 1s).

These conditions on timing express the need to monitor tracks with a unique reference time scale. This is not always the case at present: the BIPM international schedule refers to UTC time, but some commercial receivers refer to GPS time or even mix UTC time and GPS time. It is also wise to follow, strictly, the international GPS common-view schedules as issued by the BIPM.

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

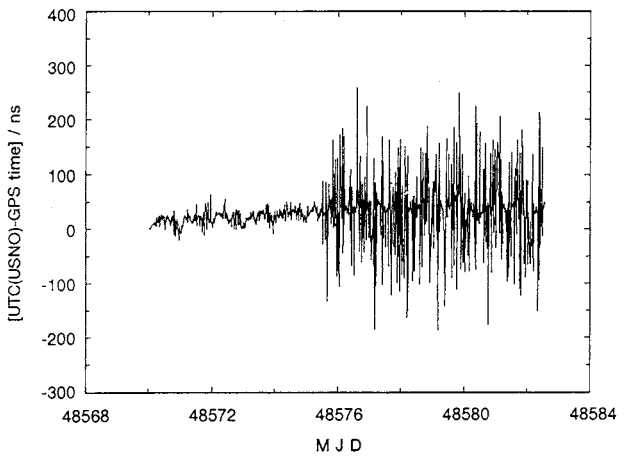


Figure 1: Raw GPS data [UTC(USNO)-GPS time] taken at the US Naval Observatory from Block I and Block II satellites in November 1991. SA was implemented on MJD 48575 (15 November 1991).

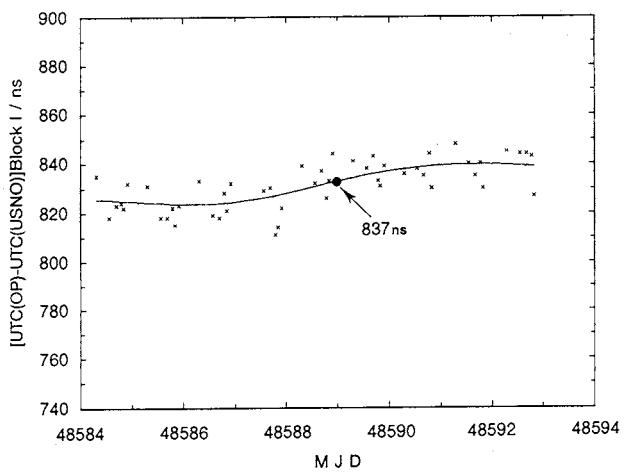


Figure 2.a: Time link [UTC(OP) - UTC(USNO)] computed using strict common views from Block I satellites only,
x raw values from individual common views,
— smoothed values.

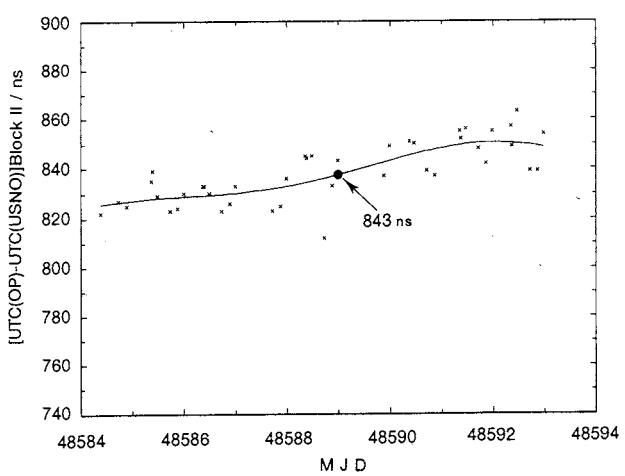


Figure 2.b: Time link [UTC(OP) - UTC(USNO)] computed using strict common views from Block II satellites only,
x raw values from individual common views,
— smoothed values.

To overcome clock degradation during SA, the long-distance time comparisons used for TAI calculation are regularly computed with strict common views by the BIPM. An example is given here for the time link between Paris Observatory (Paris, France) and US Naval Observatory (Washington DC, USA). The values [UTC(OP) - UTC(USNO)], obtained from Block I satellites only, are shown on figure 2.a and, from Block II satellites only, on figure 2.b.
The standard deviations of the residuals to smoothed values, obtained through a Vondrak smoothing with a 6-day cut-off period (Ref. 6), are nearly equal when Block I or Block II satellites only are involved. The noise brought about by SA then becomes indiscernible with the use of strict common views, which shows that only clock dither is implemented in this case.
For the BIPM standard date MJD 48589 at 0h00 UTC, the smoothed value [UTC(OP) - UTC(USNO)], obtained from Block I satellites only, is equal to 837ns while it is equal to 843ns when obtained from Block II satellites only. There are probably several reasons for such a discrepancy, in particular the use of modelled ionospheric delays and broadcast satellite ephemerides. Another explanation lies in the fact that the two GPS time receivers in operation at OP and USNO do not come from the same maker. The short-term data which are taken on both sites, and which can be very noisy because of SA, do not correspond to the same sampling interval and are not treated identically. This lack of standardization of GPS time receivers brings then inaccuracy in time transfer, even if SA is not implemented. It also could make it difficult to define the actual start times of tracks when strict common views are necessary.

To observe the effect of SA in more detail, we have plotted the short-term data obtained at the BIPM from observation of one Block I satellite (figure 3.a) and of one Block II satellite (figure 3.b). They correspond to 15-second sampling, each satellite being observed for the usual 13-minute track.
For the Block I satellite, the data is well-suited to treatment by a linear fit. It gives a precisely known slope, which is very little affected by the track start time or by the track length.
For the Block II satellite, the 15-second points show a systematic non-linear variation corresponding to periods of order a few minutes. The time deviation from one 15-second point to the next one can reach several tens of nanoseconds. Obviously in this case a linear-fit treatment is not suitable. In particular its slope drastically depends upon the length of the track. In such a case it is essential, for time transfer purposes, to strictly define the start time of the track and its length. It is also necessary to decide upon the sampling duration and to design a unified treatment for short-term data. In addition, it is important to report the value [REF - GPS] at the mid-point of the track, where it is known more precisely, rather than at its start time.

2.2 Post-processed precise ephemerides

The implementation of SA can also bring a changeable bias in the broadcast satellite ephemerides. This effect can be overcome if precise post-processed ephemerides are available. Such precise ephemerides are produced by the Defence Mapping Agency (DMA) and the National Geodetic Survey (NGS) (Ref. 7) and are accessible within several weeks. Their estimated accuracy is of order 3m.
In practice, computations with precise ephemerides require knowledge of the broadcast ephemerides used by the receiver software in order to apply differential corrections (Ref. 8), so it is necessary to collect GPS broadcast ephemerides regularly, at least at some sites in the world. Another difficulty is the possible change of ephemeris parameters during the usual 13-minute tracking period. This makes it necessary to modify the software of current GPS receivers in order to retain a single set of ephemeris parameters for the full duration of the track.
Examples of computation of long-distance time comparison, using precise satellite ephemerides when SA is on, are given in Refs. 7, 9 and 10.

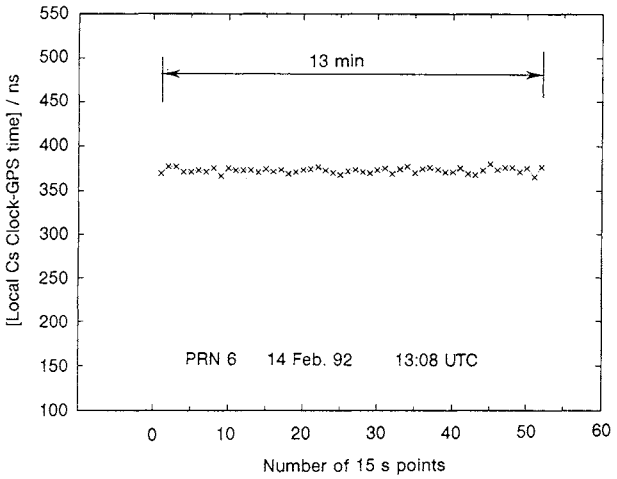


Figure 3.a: Short-term GPS data [Local Cs Clock - GPS time] taken every 15 seconds at the BIPM, for a 13-minute track of one Block I satellite (PRN 6) on 14 February 1992.

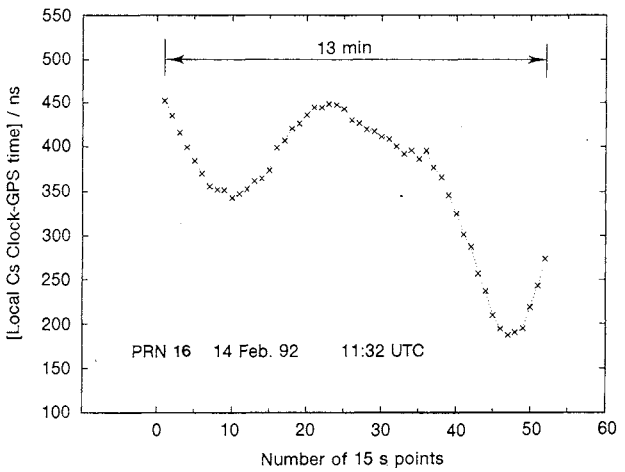


Figure 3.b: Short-term GPS data [Local Cs Clock - GPS time] taken every 15 seconds at the BIPM, for a 13-minute track of one Block II satellite (PRN 16) on 14 February 1992.

3. EXAMPLES OF DEFICIENCIES CAUSED BY A LACK OF STANDARDIZATION

In this section we give some examples of deficiencies in GPS time transfer caused by a lack of standardization. We report briefly on recent progress and underline residual difficulties. Some of these points have already been discussed in Ref. 10, so we have chosen here to focus on topics of more recent concern.

3.1 Homogenization of antenna coordinates

At one time, errors in antenna coordinates contributed one of the largest terms to the global error budgets of GPS time links (Ref. 2). These errors were first reduced for continental links (Ref. 11, 12), then, in 1990, a global homogenization of coordinates was realized in one of the most accurate reference frames, the ITRF (IERS Terrestrial Reference Frame) (Ref. 13). The BIPM continues this effort of homogenization by providing accurate coordinates in ITRF (uncertainties range from 10cm to 1m) to new laboratories equipped with GPS time receivers and contributing to TAI. Recently the BIPM has determined the position of the VNIIFTRI GPS antenna near Moscow: it is the first ITRF point in Russian Federation, the uncertainty of the determination is 1m (Ref. 14).

3.2 Receiver hardware

A study reported in Ref. 15 showed one particular GPS time receiver type to be sensitive to external temperature, in particular for long antenna cables (100m). Another typical example of such a sensitivity has been demonstrated very recently at the BIPM. The receiver involved was compared with another unit, considered as a reference, for a 100-day period. The difference of their readings is shown in figure 4.a, together with a recording of the external temperature for the same period (figure 4.b). These two series of values are strongly correlated. The GPS data provided by the receiver under study are smaller than the reference ones by about 2ns/°C. This error greatly exceeds other instrumental errors and limits the potential of GPS.

At present time, the reasons for this effect are not completely understood, but it is already clear that standardization in hardware design may be necessary.

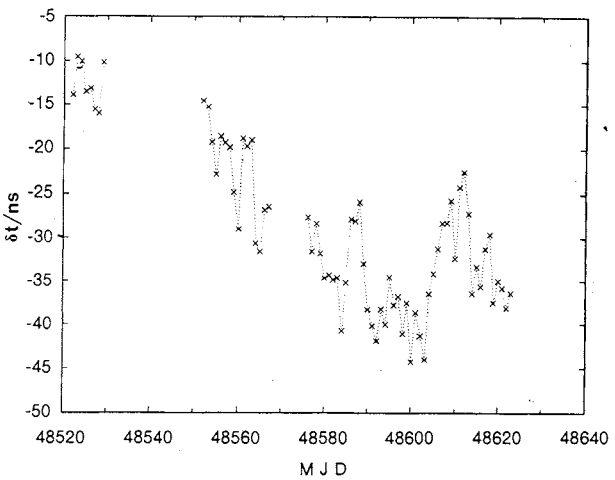


Figure 4.a: Comparison between GPS data taken at the BIPM from two different GPS time receivers: Rec 1 and Rec 2. The reported values are:
δt = *[Local Cs Clock - GPS time]*$_{Rec\ 1}$
- [Local Cs Clock - GPS time]$_{Rec\ 2}$

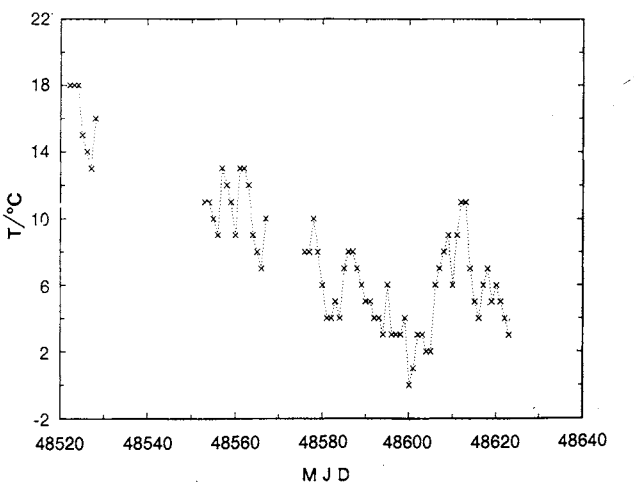


Figure 4.b: External temperature recorded at the BIPM for the period corresponding to figure 4.a.

3.3 Receiver software

Several deficiencies of GPS time receiver software, which needs an agreed standard, have already been outlined in section 2; they mainly concern:

* the reference scale for track monitoring,
* the definition of the track start time,
* the determination of the track length,
* the sampling of short-term data,
* the treatment of short-term data,
* the 13-minute freezing of broadcast ephemerides.

Typical examples concerning these points are given in Ref. 10. In this section, we focus rather on the choice of an unified tropospheric model and on the need to change the GPS data format.

3.3.1 Tropospheric model

To reach an accuracy of several nanoseconds in worldwide GPS time transfer, it is essential to choose common models for estimating signal delays due to atmospheric refraction.

If for the estimation of ionospheric delays, most receiver software uses the model proposed in the STANAG document (Ref. 4), such an agreement is not yet met for the estimation of tropospheric delays. On the contary, different models are used. They give the modelled tropospheric delay MDTR of the GPS signal as a function of the elevation El of the GPS satellite and of the altitude h of the observation site, referenced to mean sea level.

The three models which are usually implemented are:

* the 'NBS model', implemented in 'NBS type' receivers (Ref. 16), with:

$$MDTR = 7{,}255x(1+X+X^2/2)/\sin El$$

where $X = -1{,}4x10^{-4}h$, h is expressed in m and MDTR in ns (the 'NBS model' is divergent for a nul elevation angle);

* the 'STI model', implemented in receivers manufactured by Stanford Telecommunication Inc. (Ref. 17), with:

$$MDTR = 8{,}081xEXP(Y)/(\sin El + 0.026),$$

where $Y = -1.335x10^{-4}h$, h is expressed in m and MDTR in ns;

* the 'STANAG model' proposed in the STANAG document, implemented for instance in the receivers manufactured by SERCEL (France).

In this case, the analytical tropospheric correction is written in several equations depending on the altitude range, and is not reported here. It can be found in Ref. 4.

The three expressions used to model the troposheric delay are slightly different. They converge for high elevation angles but are not well adapted to low elevation angles.

Reference 18 contains one of those models which allow prediction of troposperic delays based on daily surface weather measurements (pressure, temperature and relative humidity). Evaluated against balloon measurements, the surface model of Ref. 18 was found to be able to predict the zenith tropospheric range delays with an accuracy at the subnanosecond level if the daily surface measurement is made near local noon (Ref. 19). A semi-empirical equation was then derived for lower elevation angles: the total correction is the sum of the dry and wet (water vapour) components, functions of the local pressure P, temperature T and relative humidity H.

We have computed the troposheric delays given by the three usual models for an altitude h equal to 125m and for elevation angles increasing from 10deg to 90deg by step of 1deg. We have then compared these values to the tropospheric corrections given by the semi-empirical model of Ref. 18, for the same elevation angles and for two different situations:

* summer afternoon in a temperate region ($P = 10^5$Pa, T = 25°C, H = 40%) in figure 5.a,
* tropical conditions ($P = 10^5$Pa, T = 35°C, H = 100%) in figure 5.b.

For the first case (figure 5.a), the four models agree within 1,5ns for elevation angles greater than 40°. At low elevation angles, the 'STANAG model' is in good agreement with the semi-empirical model and shows discrepancies of several nanoseconds when compared to the 'NBS and STI models'.

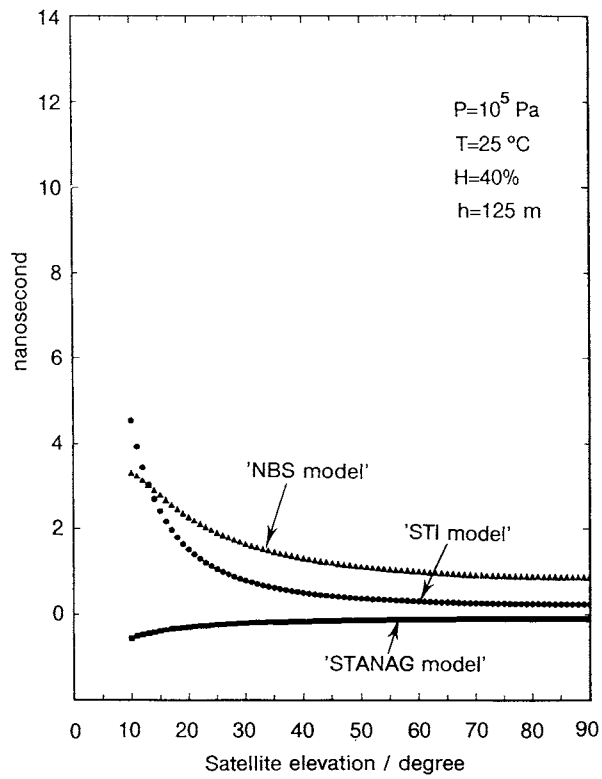


Figure 5.a: Comparison of troposheric delays obtained from the semi-empirical model based on weather measurements of Ref. 18 and from the usual models ('NBS, STI & STANAG models') implemented in GPS time receivers, for all satellite elevations. The local parameters are: pressure: 10^5 Pa, temperature: 25°C, relative humidity: 40% and altitude: 125m.

For the second case (figure 5.b), the three usual models are consistent within 1ns at high elevation angles, but give smaller values than the semi-empirical model. This is so because they do not take into account the wet component which is not negligible here. For low elevation angles, the discrepancy between the different estimations of the tropospheric delay is always of several nanoseconds and can even reach 13ns.

Some GPS users have at their disposal weather measurements, so that they could provide good estimations of tropospheric delays even at rather small elevations. In this case it is absolutely necessary to know which of the different models has been used in the receiver in order to apply differential corrections.

In addition, for sake of standardization, it would be wise to recommend a unique model for use in time receivers. The model proposed in the STANAG document seems to be a good candidate, mainly because it does not diverge for a nul elevation angle (Ref. 4).

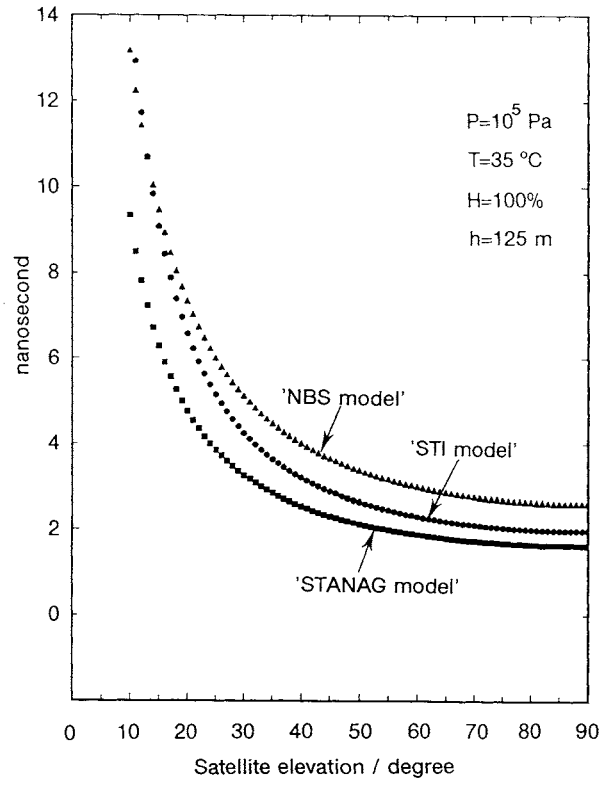


Figure 5.b: Comparison of tropospheric delays obtained from the semi-empirical model based on weather measurements of Ref. 18 and from the usual models ('NBS, STI & STANAG models') implemented in GPS time receivers, for all satellite elevations. The local parameters are: pressure: 10^5 Pa, temperature: 35°C, relative humidity: 100% and altitude: 125m.

3.3.2 GPS data format

At present most GPS receivers use the so-called 'NBS format' initially developed for 'NBS type' receivers in 1983. Untill early 1990, this standard format has fully played its role. The problem of defining a new format for GPS data files arose when ionospheric measurement systems began to operate in tandem with current time receivers. The automatic correction of GPS data by ionospheric measurements raised several questions, in particular the need to provide additional data columns for ionospheric measurements and the corresponding statistical parameters. The values of the modelled tropospheric delays could also be included in the output files.
A separate issue is that there is an incoherence in the usual data format: the quantities distributed do not refer to the same instant of the track. 'START TIME' is given for the beginning of the track while 'ELEVATION' and 'AZIMUTH' of the satellite refer to its position at the end of the track. Values of 'REF-GPS' are referenced to the beginning, the mid-point or the end of the track, depending upon the receiver. As, statistically, the mid-point value is the most reliable, all the values representing physical quantities should be referenced to the mid-point of the track.
The choice of unit is also questionable. It has become necessary to specify time values in tenths of nanoseconds rather than in nanoseconds. The elevation and azimuth of the satellite should also be given in tenths of degrees.

Most discussion about the GPS data format is now open. Receiver manufacturers and the staff of the national time laboratories are invited to give their opinions and suggestions. Agreement on the format may constitute the first concrete output of the official bodies set up to deal with the problem of GPS standardization.

4. FORMAL BODIES FOR GPS STANDARDIZATION

Two formal bodies are concerned with GPS coordination and standardization, the CGSIC Subcommittee on Time and the CCDS Group on GPS Time Transfer Standards (Refs. 10, 20).

The Subcommittee on Time of the Civil GPS Service Interface Committee (CGSIC) is mainly a forum for the exchange of information between military and civilian elements. It cannot undertake formal decisions. On the one hand, the Subcommittee provides up to date information to the civil timing community, as presently reports on progress in the computation of precise satellite ephemerides and their availability. On the other hand, it promotes the needs of the civil community, especially with respect to SA, during general meetings of the CGSIC.

The CCDS Group on GPS Time Transfer Standards (CGGTTS) operates under the auspices of the permanent Working Group on TAI of the Comité Consultatif pour la Définition de la Seconde (CCDS). This Group can initiate formal procedures as the CCDS could choose to submit its recommendations and standards to the approbation of the Comité International des Poids et Mesures (CIPM) and then to the Conférence Générale des Poids et Mesures (CGPM).

The CGSIC Subcommittee on Time and the CCDS Group on GPS Time Transfer Standards are indispensable and are complementary.

Note: The Group on GPS Time Transfer Standards was set up during the summer of 1991. This group, chaired by David W. Allan, NIST Boulder CO, includes 12 experts from national timing laboratories and from the BIPM. The BIPM also ensures its secretariat.
The CGGTTS organized an open forum on 2 December 1991, just before the 23rd PTTI meeting in Pasadena, California, and held its first formal meeting soon after. The minutes of these meetings are available on request from the BIPM Time Section. These minutes include, in particular, two annexes concerning a standardized tropospheric model and a proposed GPS data format, and a letter from the Chairman inviting feedback on these propositions.
The first formal action of the CGGTTS may be the recommendation of a new data format for implementation in GPS time receivers before the end of 1992. This will be discussed during the next formal meeting scheduled for spring 1992.

5. CONCLUSIONS

Accuracy of a few nanoseconds in GPS time transfer is now possible even for long-distance links using post-processed corrections. Further improvements are feasible through international coordination and standardization of receiver hardware and software. Joint action is required to overcome the SA degradation of GPS signals. Two complementary formal bodies are concerned with these matters, the CGSIC Subcommitee on Time and the CCDS Group on GPS Time Transfer Standards. At the beginning of 1992, the prime activities of these two committees are, respectively, providing information on SA to the civil timing community and initiating a widespread debate on GPS data format.

6. ACKNOWLEDGMENTS

The authors are grateful to the members of the CGGTTS for helpful discussions.

7. REFERENCES

1. Allan D W & Weiss M A 1980, Accurate Time and Frequency Transfer during Common-View of a GPS Satellite, *Proc. 34th Annual Symposium on Frequency Control*, 334-346.

2. Lewandowski W & Thomas C 1991, GPS Time Transfer, *IEEE Special Issue on Time and Frequency*, **79**, 991-1000.

3. Lewandowski W, Petit G & Thomas C 1991, Accuracy of GPS Time Transfer Verified by Closure around the World, accepted in *Proc. 23rd Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting*.

4. NATO Standardization Agreement (STANAG) 4294 (1 August 1990), Arinc Research Corporation, 2551 Riva Road, Annapolis, MD 21401, United States, Publication 3659-01-01-4296.

5. Allan D W, Granveaud M, Klepczynski W J & Lewandowski W 1990, GPS Time Transfer with Implementation of Selective Availability, *Proc. 22nd Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting*, 145-156.

6. Vondrak J 1969, A Contribution to the Problem of Smoothing Observational Data, *Bull. Astron. Inst. Czechoslovakia*, **20**, 349-355.

7. Lewandowski W, Petit G & Thomas C 1991, Precise GPS Ephemerides from DMA and NGS Tested by Time Transfer, accepted in *Proc. 23rd Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting*.

8. Lewandowski W & Weiss M A 1989, The Use of Precise Ephemerides for GPS Time Transfer, *Proc. 21st Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting*, 95-106.

9. Lewandowski W, Petit G, Thomas C & Weiss M A 1990, The Use of Precise Ephemerides, Ionospheric Data and Corrected Antenna Coordinates in a long-distance GPS Time Transfer, *Proc. 22nd Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting*, 547-558.

10. Lewandowski W, Petit G & Thomas C 1991, The Need for GPS Standardization, accepted in *Proc. 23rd Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting*.

11. Guinot B & Lewandowski W 1989, Improvement of the GPS Time Comparisons by Simultaneous Relative Positioning of the Receiver Antennas, *Bulletin Géodésique*, **63**, 371-386.

12. Lewandowski W, Douglas R J, Klepczynski W J, Strange W, Suter J & Weiss M A 1989, Positioning of GPS Antennas in Time-Keeping Laboratories of North America, *Proc. 43rd Annual Symposium on Frequency Control*, 218-224.

13. Lewandowski W 1991, High Accuracy Ground-Antenna Coordinates for GPS Time Transfer, *Proc. IAG Symposium G2-Permanent Satellite Tracking Networks for Geodesy and Geodynamics*, Vienna, in press.

14. Lewandowski W, Petit G & Thomas C 1991, Comparison of GLONASS and GPS Time Transfers between University of Leeds and VNIIFTRI, accepted in *Proc. 23rd Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting*.

15. Lewandowski W & Tourde R 1990, Sensitivity to the External Temperature of some GPS Time Receivers, *Proc. 22nd Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting*, 307-316.

16. Weiss M A (NIST, Boulder, CO, USA) 1989, personal communication.

17. Kirchner D (TUG, Graz, Austria) 1991, personal communication.

18. Madrid G A, Chao C C *et al* 1974, Tracking System Analytic Calibration Activities for the Mariner Mars 1971 Mission, JPL Technical Report 32-1587.

19. Fliegel H F (The Aerospace Corporation, El Segundo, CA, USA) 1992, personal communication.

20. Lewandowski W, Thomas C & Allan D W 1991, CGSIC Subcommittee on Time and CCDS Group of Experts on GPS Standardization, in *Proc. ION GPS-91 4th International Technical Meeting*, in press.

COMPARISON OF GLONASS AND GPS TIME TRANSFERS

P. Daly[1], N.B. Koshelyaevsky[2], W. Lewandowski[3], G. Petit[3], C. Thomas[3]

[1]University of Leeds, Leeds, United Kingdom
[2]VNIIFTRI, Mendeleevo, Moscow region, Russian Federation
[3]Bureau International des Poids et Mesures, Sèvres, France

ABSTRACT

Russian global space navigation system GLONASS could provide a technique similar to GPS for international time comparison. The main limitation to its use for time transfer lies in the lack of commercially available time receivers. The University of Leeds built a GPS/GLONASS receiver about five years ago and since then has provided continuous information about *GLONASS time* and its comparison with *GPS time*. For the last two years VNIIFTRI and some other Russian time laboratories have used Russian-built GLONASS navigation receivers for time comparisons. Since June 1991, VNIIFTRI has been operating a GPS time receiver on loan from the BIPM. This offered, for the first time, an opportunity for direct comparison of time transfers using GPS and GLONASS. This 7-month experiment shows that even with relatively imprecise data recording and processing, in terms of time metrology, GLONASS can provide continental time transfer at a level of several tens of nanoseconds.

Keywords: time transfer, GPS, GLONASS

1. INTRODUCTION

Two global space navigation systems, the American GPS and the Soviet GLONASS, are at the about same stage in the development of their space segments, but they are unequally used for international time comparisons. GPS, with a large range of time-specialized receivers, has for many years been exploited worldwide for accurate time transfer (Ref. 1), while GLONASS is still used on an experimental basis by only a few laboratories (Ref. 2). Although at present GPS time transfer fully satisfies the needs of time metrology, it is the sole method which is operationally effective and the lack of redundancy is felt. There is also a growing concern about GPS degradation by Selective Availability. In this context GLONASS is of increasing interest as an excellent additional source.

For about five years the University of Leeds (Leeds, United Kingdom) has operated a GPS/GLONASS timing receiver built in-house, and provided continuous information about *GLONASS time* and its comparison with *GPS time*. For the last two years VNIIFTRI (Mendeleevo, Moscow Region, Russian Federation) and some other Russian time laboratories have used Russian-built GLONASS navigation receivers for time comparisons. Since June 1991, VNIIFTRI has been operating a commercial GPS time receiver on loan from the BIPM. This offers an opportunity for direct comparison of time transfers by GPS and GLONASS between laboratories of West Europe and Russian Federation. Values of [UTC(LDS) - UTC(SU)] and [UTC(OP) - UTC(SU)], as provided by GPS and by GLONASS, are reported together with estimates of the errors involved.

This experiment covers the period from 4 July 1991 to 31 January 1992 and corresponds to the BIPM international common view schedules n°17 and 18. It is the continuation of the 2-month experiment reported in Ref. 3. The GPS Block II satellites have generally been used for GPS common-view treatment as Selective Availability, which has been currently implemented since 15 November 1991, consists only in clock dither (Ref. 4). Nevertheless, the GPS Block II satellites have not been used for a 5-day period at the beginning of July 1991, when SA also presented a degradation of satellite broadcast ephemerides (Ref. 5).

2. GLONASS AND GPS OBSERVATIONS AT THE UNIVERSITY OF LEEDS

The time laboratory of the University of Leeds is equipped with the following time receivers:

- TRIMBLE 5000A and AOA TTR6.

The TRIMBLE 5000A GPS time receiver was programmed with the BIPM international common view schedule n°17 and refers to UTC(LDS). It stopped operation on 29 November 1991 and was replaced by an ALLEN OSBORNE ASSOCIATES TTR6 GPS time receiver on 11 January 1992 (last weeks of the experiment). The AOA TTR6 is programmed with the BIPM international common view schedule n°18 and also refers to UTC(LDS).

- TRIMBLE 4000A.

This receiver generates a 1 PPS synchronised on UTC(USNO) as broadcast by GPS satellites with an uncertainty of 100ns (Ref. 1). This signal is the external reference of the Leeds University GPS/GLONASS receiver.

- Leeds University GPS/GLONASS receiver.

This receiver performs the measurements of [UTC(USNO) - *GPS time*] and [UTC(USNO) - *GLONASS time*]. It is noticeable that there is no direct access to [UTC(LDS) - *GLONASS time*].

These receivers are connected to the one omni-directional antenna and use the same set of coordinates expressed in the WGS 84 reference frame with an uncertainty of 3 m.

In the past GLONASS observations by GPS/GLONASS Leeds University receiver provided a series of navigational solutions expressed in the 'Soviet reference frame' SGS 85 used by GLONASS satellites. These results agreed with WGS 84 coordinates within 5 meters.

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

The delays of the receivers are known approximately within 20ns. The schema of the whole installation is illustrated in figure 1.

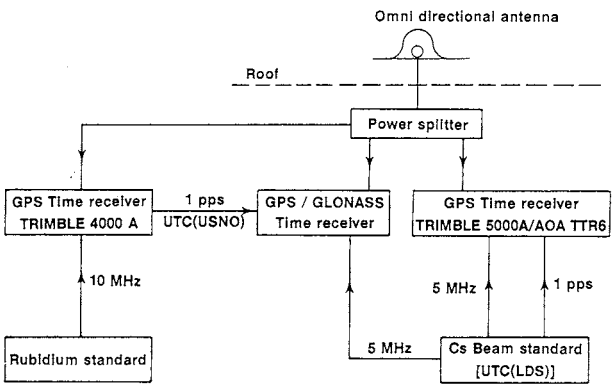


Figure1: Schema of installation of time receivers at the University of Leeds.

The GPS/GLONASS receiver performs the measurements almost continuously, observing all available GLONASS satellites and a selection of GPS satellites (all Block I and two Block II satellites). For a given satellite, the receiver starts measurements of [UTC(USNO)- *GPS time*], or [UTC(USNO)- *GLONASS time*], and does this once a second for 180s. This data is stored, filtered to remove outliers and averaged. The typical standard deviation for 180s averages is 50ns. During one day the receiver performs roughly 340 tracks of 180s, 75% of them corresponding to GLONASS satellites. Daily averages of tracks have a standard deviation around 50ns for both GPS and GLONASS data. These results are not corrected for ionospheric or tropospheric delays.

The GPS observations produced by the GPS TRIMBLE 5000A receiver, and then by the AOA TTR6 receiver, are used in this study for common-view time transfer with VNIIFTRI. Previous analysis of common-view time transfer between the University of Leeds and the Paris Observatory showed that the TRIMBLE 5000A data is affected by a noise which limits the uncertainty of such a time transfer to 10-15ns. This noise is partly due to uncertainty in the coordinates. Several attempts have been made to improve the Leeds antenna coordinates by the BIPM method (Ref. 6). All of them produced the coordinates with uncertainties of several meters which indicates other than geometrical error sources. Since 11 January 1992, the use of the AOA TTR6 receiver has reduced the uncertainty of GPS common-view time transfer to about 7ns but more accurate GPS coordinates have not yet been determined.

The TRIMBLE 5000A and the AOA TTR6 GPS time receivers were programmed during this experiment with the usual 13-minute tracks of the BIPM international tracking schedules including all Block I and Block II satellites. About 25 tracks were available each day for this experiment.

3. GLONASS AND GPS OBSERVATIONS AT THE VNIIFTRI

The VNIIFTRI is located in Mendeleevo, near Moscow and is responsible for the maintenance of the Russian national time reference UTC(SU). An ensemble of high-quality atomic clocks, mostly hydrogen-masers, is operated.

GLONASS time observations at VNIIFTRI have been carried out, since 1989, using a Russian-built commercially available receiver A-724 designed for aircraft navigation. The receiver is supplied with 1 PPS and 5 MHz provided by the VNIIFTRI master clock. The readings of the master clock are corrected 'a posteriori', to transform them into UTC(SU). The receiver uses a fixed model of the ionospheric delay and does not correct observations for tropospheric delay. The antenna coordinates are expressed in SGS 85, with estimated uncertainty of order 5m, provided by a series of navigational solutions.

There are three to five observations per day of [UTC(SU) - *GLONASS time*]. All of them are performed at low elevations in the direction of East. With the limited model of ionosphere and the lack of tropospheric correction this particular configuration of observations can produce a bias.

As an estimated uncertainty for the daily averages of [UTC(SU) - *GLONASS time*], we adopted the value of 50ns as already deduced for the observations at the University of Leeds.

Since June 1991 VNIIFTRI has operated a commercial GPS time receiver on loan from the BIPM, referenced to the local master clock. The delay of the GPS receiver was determined by the comparison with the Paris Observatory 'on line' GPS receiver (Ref. 7). The coordinates of its antenna were determined by the BIPM method (Ref. 6) and expressed in the ITRF reference frame (Ref. 8) with an uncertainty of 1m. The receiver is programmed with the usual 13-minute daily tracks according to the BIPM tracking schedules n°17 and 18. About 35 tracks were available each day for this study. The GPS installation at the VNIIFTRI allows the comparison of the UTC(SU) in common-view mode to the West European time laboratories with an accuracy of a few ns.

4. GLONASS TIME TRANSFER BETWEEN MENDELEEVO AND WESTERN EUROPE

To realise time transfer between Leeds and Mendeleevo via GLONASS we use the measurements of [UTC(SU) - *GLONASS time*], [UTC(LDS) - *GPS time*], and [*GLONASS time* - *GPS time*]. The combination of these three values gives [UTC(LDS)-UTC(SU)].

VNIIFTRI provides the daily values of [UTC(SU) - *GLONASS time*] with uncertainty of 50ns and Leeds provides the measurements of [UTC(LDS) - *GPS time*] realized by the TRIMBLE 5000A receiver. The 25 or so available daily measurements of [UTC(LDS) - *GPS time*] are smoothed at the BIPM to provide daily values at 0h UTC with an uncertainty of 15ns. Since the implementation of the AOA TTR6 on 11 January 1992, this uncertainty has been reduced to about 10ns.

To obtain the values of [*GLONASS time* - *GPS time*] we use the daily averages of [UTC(USNO) - *GLONASS time*] and [UTC(USNO) - *GPS time*] provided by Leeds with uncertainties of 50ns. We believe that these two measurements are affected partly by the same systematic biases (uncertainty of the UTC(USNO) as locally reconstituted, ionospheric delay, coordinates,...). For this reason when we remove UTC(USNO) by differencing the above measurements to obtain the daily values of [*GLONASS time* - *GPS time*], we adopt the uncertainty of 50ns for this difference.

For the final values of [UTC(LDS) - UTC(SU)] obtained from this process we adopt an uncertainty of 70ns, which is derived from the quadratic combination of the involved uncertainties.

We have also realized the comparison of UTC(SU) with the Paris Observatory time scale UTC(OP) using both systems GLONASS and GPS. Paris Observatory operates a commercial GPS time receiver connected to UTC(OP). The comparison of UTC(SU) with UTC(OP) via GLONASS was realized in a similar way to that described above, the values [UTC(LDS) - *GPS time*] being replaced by [UTC(OP) - *GPS time*]. The 35 or so available daily measurements of [UTC(OP) - *GPS time*] are smoothed at the BIPM to provide daily values at 0h UTC with an uncertainty of 7ns. The final values of [UTC(OP) - UTC(SU)] via GLONASS are provided with an estimated uncertainty of 70ns.

5. GPS COMMON-VIEW TIME TRANSFER BETWEEN MENDELEEVO AND WESTERN EUROPE

The common-view time transfer between Mendeleevo and Leeds was realised with about 25 tracks available, and between Mendeleevo and Paris with about 35 daily tracks available. In both cases a Vondrak smoothing (Ref. 9), which acts as a low-pass filter with a cut-off period of about 1 day, was performed on the common-view values. For this experiment, the smoothed values are interpolated for 0h UTC of each day. The precision of both time links is estimated from the residuals of the smoothed values. Over the period of this study, the residuals ranged from 3 to 4ns for the [UTC(OP) - UTC(SU)]. They ranged from 12 to 15ns for the [UTC(LDS) - UTC(SU)] through the TRIMBLE 5000A receiver and fell to 7 to 8ns through the AOA TTR6 receiver.

6. COMPARISON OF GLONASS AND GPS

Time transfer via GLONASS between Mendeleevo and two West European laboratories was realized with an estimated uncertainty of 70ns. The GPS common-view time transfer provided a time link between Mendeleevo and Leeds with an uncertainty of about 15ns, for the most part of the experiment, and between Mendeleevo and Paris with an uncertainty of about 4ns. Accordingly the comparison of GLONASS with GPS for the Leeds-Mendeleevo link has an estimated uncertainty of 80ns and for Paris-Mendeleevo, 70ns.

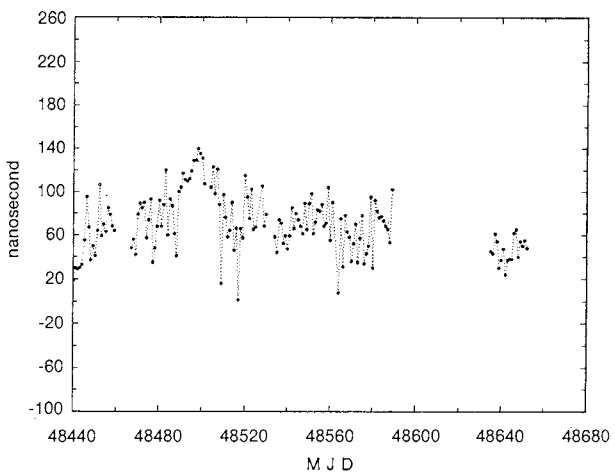


Figure 2: [UTC(LDS) - UTC(SU)] as obtained by GPS minus [UTC(LDS) - UTC(SU)] as obtained by GLONASS+GPS. Daily values interpolated for 0h UTC.

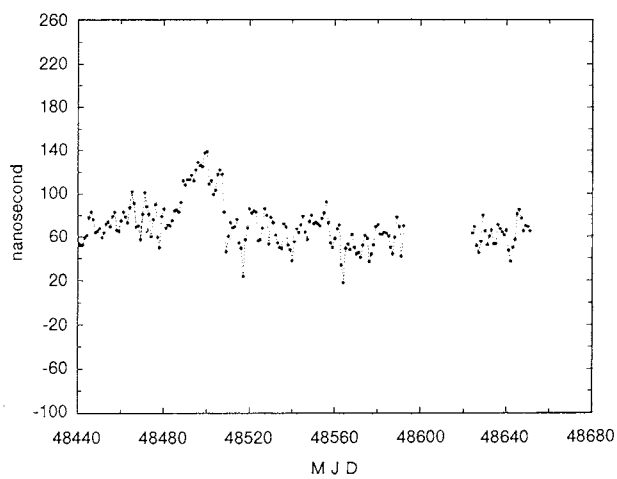


Figure 3: [UTC(OP) - UTC(SU)] as obtained by GPS minus [UTC(OP) - UTC(SU)] as obtained by GLONASS+GPS. Daily values interpolated for 0h UTC.

The daily differences between the two methods are presented in figures 2 and 3. Use of the modified Allan variance (figures 4 and 5) allows us to characterize the noise affecting the values of comparison for each of the two links. For both links the data exhibit white phase noise up to an averaging time of about 4 days. This justifies computation of mean values for periods of duration up to 4 days and corresponding standard deviations of the mean. The results are given in table 1 for a 4-day averaging period.

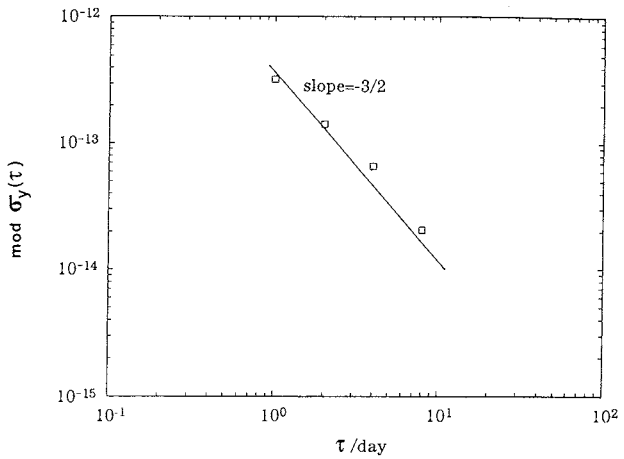


Figure 4: Square root of the modified Allan variance of the differences presented in Figure 2.

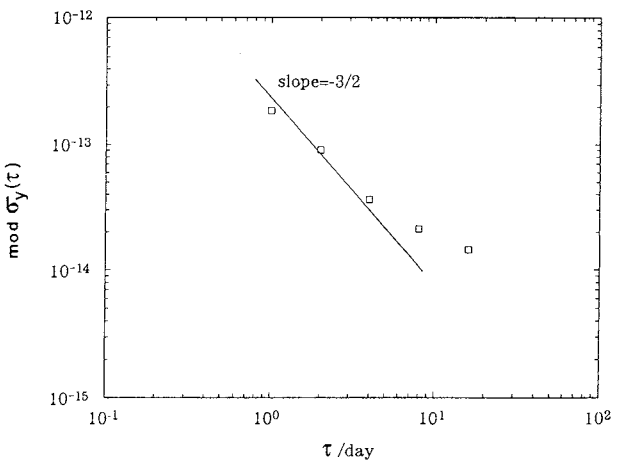


Figure 5: Square root of the modified Allan variance of the differences presented in Figure 3.

Figures 2 and 3, and table 1, show a fairly constant bias of about 65ns between the two techniques for nearly the whole period of the experiment and for the two links. It presents also a shift of 50ns, which suddenly appears on MJD 48487 and desappears on MJD 48507. This can be explained by a delay wrongly introduced between the 1PPS of the master clock and the GLONASS receiver at VNIIFTRI on MJD 48487, and then deleted on MJD 48507.

One can also observe, on figures 2 and 3, a lower noise level for the link between Paris and Mendeleevo than that between Leeds and Mendeleevo for the period when the TRIMBLE 5000A receiver was in operation in Leeds. For the month of January 1992, corresponding to the operation of the AOA TTR6 receiver in Leeds, the noise is at the same level on both of the figures.

We have computed the means of the daily values of the comparison between the two techniques, for each link, and for the whole duration of the experiment (with exclusion of the 10-day period MJD 48487 to 48507). As the noise is not white for averaging intervals longer then 4 days, we adopt

the root mean square of the residuals to the average as an estimation of the confidence of the mean. The averaged value for the link Paris-Mendeleevo is 65ns with estimated confidence 14ns and for the link Leeds-Mendeleevo respectively 64ns and 23ns. Both estimates of confidence for the period of this study are significantly lower that for our estimation of the uncertainties of the involved measurements (70ns and 80ns). This indicates that our estimates are too conservative and that there is a consistency between the daily comparison values. The observed 65ns bias should be removed by a differential calibration of the equipment.

Table 1: [UTC(OP) - UTC(SU)] as obtained by GPS minus [UTC(OP) - UTC(SU)] as obtained by GLONASS+GPS; mean values, in ns, for 4-day averaging period (column 2) and standard deviation, in ns, (column 3).
[UTC(LDS) - UTC(SU)] as obtained by GPS minus [UTC(LDS) - UTC(SU)] as obtained by GLONASS+GPS; mean values, in ns, for 4-day averaging period (column 4) and standard deviation, in ns, (column 5).
The 4-day averaging periods are given in column 1.

48440-48443	54	2	27	2
48444-48447	74	5	62	13
48448-48451	64	2	48	6
48452-48455	70	2	74	10
48456-48459	73	5	74	5
48460-48463	77	2		
48464-48467	87	7		
48468-48471	77	9	66	10
48472-48475	76	6	76	7
48476-48479	70	9	61	12
48480-48483	74	4	92	10
48484-48487	82	3	75	8
48508-48511	66	8	69	18
48512-48515	67	5	64	9
48516-48519	50	9	48	16
48520-48523	84	1	96	9
48524-48527	67	7	66	1
48528-48531	71	6	84	10
48532-48535	54	3	51	5
48536-48539	57	5	64	5
48540-48543	56	6	64	8
48544-48547	68	5	71	4
48548-48551	75	2	85	7
48552-48555	75	3	74	5
48556-48559	67	9	82	8
48560-48563	58	8	72	12
48564-48567	42	8	48	17
48568-48571	50	4	52	6
48572-48575	53	4	60	9
48576-48579	50	7	55	13
48580-48583	65	2	70	13
48584-48587	54	4	71	3
48588-48591	62	8	78	17
48624-48627	57	5		
48628-48631	63	6		
48632-48635	58	3		
48636-48639	66	2	47	7
48640-48643	50	6	36	5
48644-48647	75	6	51	7
48648-48651	67	1	50	3

7. CONCLUSION

This study shows that even with relatively imprecise data recording and processing, in terms of time metrology, GLONASS can provide continental time transfer at a level of several tens of nanoseconds. By introducing common-views, this performance should be significantly improved. Further possible improvements are the, at least partial, removal of ionospheric and tropospheric delays. More precise determination of GLONASS antenna coordinates in the SGS 85 reference frame is another task. The development of automatic GLONASS receivers dedicated especially for time transfer would be a most significant breakthrough.

8. REFERENCES

1. Lewandowski W & Thomas C 1991, GPS time transfer, *IEEE Special Issue on Time and Frequency*, **79**, 991-1000.

2. Daly P, Cherenkov G T, Koshelyaevsky N B & Pushkin S 1991, Satellite time transfer between UTC(USNO) and UTC(SU) using Navstar GPS and GLONASS, *Proc. 4th Institute of Navigation Technical Meeting*, 199-206.

3. Daly P, Koshelyaevsky N B, Lewandowski W, Petit G & Thomas C 1991, Comparison of GLONASS and GPS Time Transfers between two West European Time Laboratories and VNIIFTRI, accepted in *Proc. 23rd Annual Precise Time and Time Internal (PTTI) Applications and Planning Meeting.*

4. Lewandowski W, Petit G & Thomas C 1992, GPS Standardization for the Needs of Time Transfer, accepted in these Proceedings.

5. Lewandowski W, Petit G & Thomas C 1991, The Need for GPS Standardization, accepted in *Proc. 23rd Annual Precise Time and Time Internal (PTTI) Applications and Planning Meeting.*

6. Guinot B & Lewandowski W 1989, Improvement of the GPS Time Comparisons by Simultaneous Relative Positioning of the Receiver Antennas, *Bulletin Géodésique*, **63**, 371-386.

7. Lewandowski W 1991, Determination of differential time corrections between the GPS time receivers located at the Observatoire de Paris, the Observatoire de la Côte d'Azur and the Technical University of Graz, *Rapport BIPM-91/6.*

8. Lewandowski W 1991, High Accuracy Ground-Antenna Coordinates for GPS Time Transfer, *Proc. IAG Symposium G2-Permanent Satellite Tracking Networks for Geodesy and Geodynamics*, in press.

9. Vondrak J 1969, A Contribution to the Problem of Smoothing Observational Data, in *Bull. Astron. Inst. Czechoslovakia*, **20**, 349-355.

AN EXPERIMENTAL HEALTH MONITORING UNIT FOR GPS AND GLONASS*

B.Eissfeller[1], M.Haunschild[2], A.Jansche[1], N.Niklasch[2]

[1] Kayser-Threde GmbH, Munich; [2] MAN Technologie AG, Munich

ABSTRACT

An Experimental Health Monitoring Unit (HMU) for GPS and GLONASS is developed in contract to the European Space Agency. The purpose of the HMU is to provide timely warnings to the navigation users if the satellite navigation systems should not be used for navigation. The monitoring unit is a key element in the integrity channel concept of a civil overlay NAVSAT (Refs. 1, 2) to GPS and GLONASS. The most stringent requirements are those of civil aviation: In the nonprecision approach phase horizontal position errors exceeding the 600m SPE threshold have to be reported within 10 seconds. The concept of the HMU is to use a GPS receiver as well as a GLONASS receiver in a known location in order to detect large pseudorange errors and to identify the malfunctioning satellites. This means that the station is able to work autonomous. But it will also provide pseudorange errors compatible with the american RTCA SC-159 approach and for a future wide area DGPS system.

Keywords: Integrity monitoring, NAVSAT, GPS and GLONASS

1. INTRODUCTION

GPS and GLONASS are very accurate satellite navigation systems which nearly fulfil all navigation user requirements. Concerning integrity one has to note that both systems are military systems designed to enhance the effectiveness of military missions. This means that only military integrity requirements and risk levels are applied to derive the health of the navigation service with respect to warning time and alarm limits. It is known that the time between the occurance of a satellite malfunction and a corrective action of the ground control segment can be as large as 1.0 to 1.5 hours for GPS and 18 hours for GLONASS. Thus, both GPS and GLONASS are deficient in integrity.

An important user group of GPS and GLONASS will be the civil aviation, which has very high integrity requirements in comparison to military users and all other civil users. For aviation users time-to-alarm of one hour and more can not be accepted in all phases of flight, even in the oceanic en route phase. The most stringent requirement is the 10 sec time-to-alarm during nonprecision approach.

1.1 Definition of integrity

Integrity of a navigation system is defined by the RTCA SC-159 working group in the following way (Ref. 3): "Integrity is the ability of the system to provide timely warnings to users when the system should not be used for navigation". Of course this definition is user dependent because the integrity levels to be checked depend on the phase of flight (en route, terminal area, nonprecision approach) of an aircraft.

1.2 Integrity channel requirements

One possibility to solve the integrity problem in satellite navigation is to establish a so called GPS or GLONASS integrity channel (GIC), which is based on a network of earth-fixed monitoring stations in known locations and a data transmission channel.

The general requirements on the integrity channel are the following:

- The GIC shall provide protection against large horizontal signal derived position errors (SPE).
- A positive indication of satellite usability shall be provided for each healthy satellite in view.
- The GIC shall provide an integrity alarm if a satellite fails which is indicated as healthy in the navigation message or if the monitoring system itself has a failure.
- The GIC is not allowed to set satellites healthy which are set unhealthy by the GPS and GLONASS Operational Control Segments (OCS).
- For sole means satellite navigation the reliability of the GIC has to be 100% an the miss probability 0%.
- For supplementary navigation the requirement has to be close to 100% and close to 0% respectively.
- For aviation the update rate of integrity information has to be smaller as 10 seconds.
- The GIC has to apply the same data reduction algorithms on the raw measurements as the user receiver does.

For aviation users the following requirements and goals for the time-to-alarm and the alarm limits are recommended by the RTCA SC-159:

Phase of flight	Oceanic en Route	Domestic en Route	Terminal Area	Nonprecision approach
Requirements:				
Alarm limit	23 km	3 km	2 km	0.6 km
Time to alarm	120 s	60 s	15 s	10 s
Goals:				
Alarm limit	5 km	1 km	0.5 km	0.1 km
Time to alarm	30 s	30 s	10 s	6 s

Table 1 : Requirements and goals for alarm limits and time-to-alarm (RTCA SC-159)

Note again that the alarm limits in Table 1 are horizontal errors thresholds.

1.3 Failure mechanisms and satellite errors

The GPS and GLONASS have a lot of built-in features and operating procedures to ensure the integrity of the navigation service. These include equipment redundancy, communication error detection codes, estimation and prediction consistency checks and human operator qualification verification. But like in every technical system, also in GPS and GLONASS

* This paper is a summary of the work performed under ESA contract No.: 9281/91/F/RD(SC)

failures are possible which are unknown to the GPS and GLONASS Operational Control Segments (OCS) in short term. Although of a satellite failure the satellite may be declared as healthy in the navigation message. There are known and unknown failure mechanisms in both systems, which lead to different type of error. On side of known failure mechanisms the following malfunctions can occur:

- Interruption of the onboard cesium standard feed-back loop. In this case the clock is driven by its crystall oscillator, which leads to a large pseudorange drift.
- Human mistakes or computer errors can lead to the up-load of wrong ephemeris and wrong system time.
- Not a real failure case is the combination of bad satellite geometry with unusual large SA and ionospheric effects.

These types of failure are to be detected by the integrity monitoring stations and may be classified as soft failures. In contrary dramatic events - like the satellite stops to transmit a signal - are detected by the user receivers itself. A further classification of errors leads to line-of-sight dependent and line-of-sight independent errors. For example ephemeris errors have vectorial character and are only detected by a monitoring station if the error vector is not orthogonal to the line-of-sight. Timing errors are typical line-of-sight independent errors.

2. DESCRIPTION OF THE EXPERIMENTAL MONITORING UNIT

The experimental monitoring unit was defined in such a way to show the feasibility of integrity monitoring with low cost receiver equipment and to gain experience with monitoring algorithms including satellite failure simulation. Because a network of monitoring stations does not exist at the moment the health monitoring unit (HMU) is autonomously able to check the health of GPS and GLONASS satellites. The operational version in a future global integrity monitoring network will perhaps use more sophisticated hardware and will work on data sampling only. The health decission is then derived in the regional center where the data of the entire network is received and processed.

2.1 Hardware built-up of the monitoring unit

The built-up of the experimental health monitoring station is shown in Figure 1:

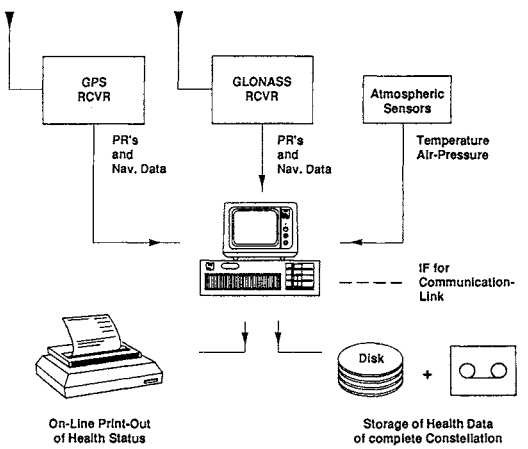


Figure 1 : H/W built-up of the experimental health-monitoring unit

The HMU is built-up by a GPS receiving unit, a GLONASS receiving unit, digital atmospheric sensors and a personal computer including a printer and a disk for storage of satellite and integrity data. The main requirement on the receiving units is that an output of the complete raw pseudorange, phase and doppler data as well as the complete navigation message (almanach, ephemeris) is possible.

In the GPS receiving unit two Magnavox MX 4200 receivers are used in order to provide an all-in-view solution when GPS is operational in near future. Only one GPS antenna is connected to both receivers by using a power divider. Each MX 4200 has a power supply, a raw data port and a control port. All ports are connected to the PC using RS 422 interfaces. The tropospheric data (air-pressure and temperature) is input into the PC via an RS 232 interface.

The PC is AT compatible, has a 80386 (25 MHz) processor, a 80387 coprocessor, a 4 MByte RAM, a 1.44 MByte floppy and a 80 Mbyte hard disk. In summary 4 RS 422 and 8 RS 232 as well as an ARINC-429 for GLONASS on interface level are foreseen. Furtheron a VGA video card (1024 x 786), a VGA color monitor and an Epson needle printer are used in the HMU.

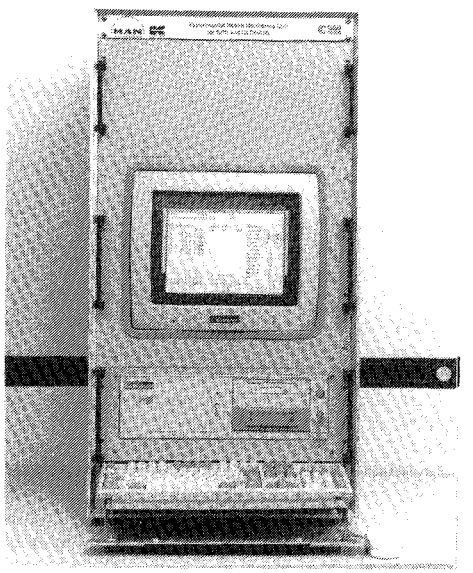


Figure 2 : Experimental monitoring unit

The PC components are integrated mechanically in a racktype of box (Figure 2) and the MX 4200 GPS receivers are integrated in a seperate box near the antenna. The RS 422 data cables between receiver unit and PC box have a maximum length of 50m.

The complete integration of the PC and the GPS receiver unit was performend in 1991 and is now under test at TU Graz (Institut für Nachrichtentechnik und Wellenausbreitung) and at ESTEC.

Because at the start of the project in early 1991 GLONASS equipment was not available, the GLONASS option was not integrated in the first development step. In the mean time russian GLONASS receivers as well as american GLONASS ASICS are available on western market. Therefore it is foreseen in this year to incorporate GLONASS into the HMU. Concerning GLONASS it is planned to integrate a russian receiver (ASN-16 or SKIPPER) as well as an american poduct into the HMU.

2.2 Software functions

A basic flow-chart of software operations to be performed by the HMU is outlined in figure 3.

With the input operations the raw pseudorange, navigation message and atmospheric data is read-in the PC. In the first computational step the geometric ranges between the HMU and all satellites in view are predicted. In the second step the measured code phase data is scaled and corrected to metric pseudoranges. Afterwards the difference between predicted range and pseudorange is computed (pseudorange error or residual). Because the pseudorange residual is biased by the unknown clock off-set of the receiver, it has to be elimated by averaging over all visible satellites. The so obtained range errors are then tested for outliers and compared with the pseudorange thresholds derived from the horizontal alarm limits. This means that range errors out of bound are detected and the individual malfunctioning satellites are indentified. Based on this detection and identification procedure the health decision is performed by keying it with respect to the different phase of flight alarm thresholds. The monitoring algorithm is described later in more detail.

2.3 Required time budget for integrity monitoring

In the integrity monitoring approach it has to be verified that the monitoring unit fulfils the 10 second time-to-alarm requirement for warning the user. A realistic estimate for all the time delays in the monitoring system is the following:

Operations and Delays	300 bit/sec	9600 bit /sec
Monitoring stations:		
- receiver data update rate	1 s	1 s
- data transfer to computer	<< 1 ms	<< 1 ms
- processing time	2 s	2 s
Regional center and uplink:		
- monitoring unit/regional center	0.24 s +1 s	0.24 s + 0.03 s
- processing time	2.5 s	2.5 s
- regional center/uplink station	tbd	tbd
- uplink station/user receiver	0.24 s + 1 s	0.24 s + 0.03 s
- data detection at user	1 s	1 s
	Σ 8.98 s	Σ 7.04 s

Table 2 : Time budget for integrity monitoring

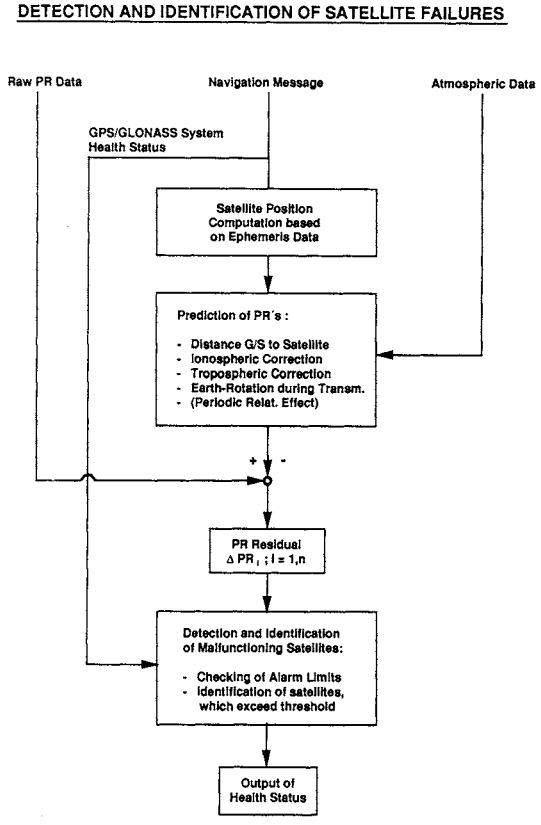


Figure 3 : Basic software flow-chart

The share of processing time budgets of the local HMU with respect to the regional center depends on the S/W tasks attached to both types of stations. The transmission delay between regional center and uplink station is TBD because it depends largely on the design of the whole monitoring network and its relation to the INMARSAT coast earth uplink stations. Of course the transmission delays depend on the transmission rates used. It is seen from Table 2 that there is still some margin for the latter transmission delay. All the values in Table 2 are only estimates and have to be verified.

3. MONITORING ALGORITHM

The integrity monitoring algorithm programmed on the HMU employes in its basic version the measured code phase or pseudorange data and applies several physical propagation delay corrections. Of course the strongly time varying but known geometric distance between HMU and all visible satellites is subtracted. After this processing step the so-called pseudorange error is obtained, which is the basic input to the monitoring algorithm. Note further that health monitoring is always a two-step procedure. The one step consists of the detection part and the second step of the identification part. The first step is usually performed by a global error criteria which checks only the presence of a failure. Only if the detection criteria is violated the second step is started and the malfunctioning satellites are identified.

3.1 Detection and identification of satellite failures

The starting point of the monitoring algorithm is the pseudorange measurement equation (Ref. 4) to a satellite i

$$PR(i) = | \underline{r}(i) - \underline{r} | + \tau + \Delta_{CORR}(i) + \varepsilon(i) \qquad (1)$$

where

$\underline{r}(i)$	satellite position vector
$\underline{r}$	monitoring unit position vector
τ	receiver clock offset
Δ_{CORR}	propagation delay corrections
ε	range error.

The physical propagation delay corrections Δ_{CORR} are decomposed of the following components:

$c\,\Delta t(i)$	satellite clock offset
c	speed of light
Δ_{ION}	ionospheric delay
Δ_{TROP}	tropospheric delay
Δ_{TRANS}	earth-rotation during signal transmission
Δ_{REL}	periodic general relativity correction

The physical corrections are without limitation predictable based on suitable models. Based on this fundamental equation the pseudorange errors are computed:

$$\Delta PR(i) = PR(i) - | \underline{r}(i) - \underline{r} | - \Delta_{CORR}(i) \qquad (2)$$

$$\Delta PR(i) = \tau + \varepsilon(i) \qquad (3)$$

If measurements to several satellites are available the receiver clock offset can be estimated by averaging over the visible satellites:

$$\tau = \sum_{i=1}^{n} \frac{\Delta PR(i)}{n} \qquad (4)$$

The range errors are then defined by:

$$\varepsilon(i) = \Delta PR(i) - \tau \qquad (5)$$

For failure detection we evaluate the variance of the above computed clock offset estimate and compare it with a fixed threshold, e.g.: $\sigma > \sigma_0 = 3$ UERE or we perform a statistical test (Ref. 5) based on the τ-criterion. In order to avoid the miss of a common bias in addition the difference $| \tau - \tau^* | > \delta_0$ is tested. The latter threshold comparison implies the use of a receiver clock model, τ^*.

The receiver clock model is also used for satellite failure identification. In case that the detection criteria are violated the identification procedure is entered and the clock model predict τ^* is used in a first step in order to avoid biasing of the clock offset estimation.

The last step is the comparison of the actual range errors with a predefined range threshold ε_0 derived from the horizontal SPE in Table 1:

$$\varepsilon(i) > PR_{max} = \varepsilon_0 \qquad (6)$$

The satellite for which the above equation is true is set to unhealthy for the specific phase of flight and is removed before the receiver clock model is updated.

The position computation for GPS and GLONASS satellites is done based on the ICD-GPS-200 (Ref. 6) and based on the GLONASS ICD (Ref. 7). The ionospheric correction supported only in GPS is computed based on the

Klobuchar model (Ref. 6). For tropospheric correction the model of Hopfield (Refs. 8, 9) and the model of Black may be used (Ref. 10).

3.2 Alarm limit algorithms

In integrity monitoring literature (Ref. 11) three types of algorithms for the computation of alarm thresholds are published:

- PR_{max} algorithm.
- HASE (Horizontal Absolute Sum Error) algorithm.
- SPE (Signal Derived Positioning Error) algorithm.

All algorithms have advantages and draw-backs concerning probability of false alarm and miss.

3.2.1 PR_{max} algorithm

The PR_{max} algorithm is the simplest way of transforming the horizontal SPE alarm thresholds to pseudorange level. Under the assumption of an allowed $HDOP_{max}$ value the following error bound for the horizontal SPE is found:

$$\text{Horizontal SPE} < 2.57\ HDOP_{max}\ PR_{max} \qquad (7)$$

By inversion of this formula the following pseudorange alarm limits are obtained:

Phase of flight	Reqiurement		Goal	
	SPE [m]	PR_{max} [m]	SPE [m]	PR_{max} [m]
Nonprecision approach	600 (10 s)	78	100 (6 s)	13
Terminal area	2000 (15 s)	259	500 (10 s)	65
Domestic en route	3000 (60 s)	389	1000 (30 s)	130
Oceanic en route	23000 (120 s)	2983	5000 (30 s)	648

Table 3 : PR alarm limits for $HDOP_{max} = 3.0$

In summary the PR_{max} algorithm to be followed by the monitoring station consists in the following steps:

1. Reject all satellites where $\varepsilon(i) > PR_{max} = \varepsilon_0$.
2. Alarm if actual HDOP exceeds the assumed $HDOP_{max}$.

The advantage of the PR_{max} method is that it is very simple. Its draw-back is that always a worst case assumption on the horizontal satellite geometry is used. This could lead to a higher false alarm probability.

3.2.2 HASE algorithm

The HASE algorithm is not working with fixed pseudorange alarm thresholds but it takes into account the actual satellite geometry. Thus, the HASE leads to variable alarm thresholds (Ref. 11) based on the following expression:

$$PR_{max} = \varepsilon_0 = \text{horizontal SPE / HASE} \qquad (8)$$

The HASE or Horizontal Absolute Sum Error can be derived based on the navigation solution in horizontal coordinates on a reference ellipsoid :

$$\text{North SPE} = b_{11}\,\Delta PR_1 + \ldots + b_{1n}\,\Delta PR_n \qquad (8\text{-}1)$$

$$\text{East SPE} = b_{21}\,\Delta PR_1 + \ldots + b_{2n}\,\Delta PR_n \qquad (8\text{-}2)$$

The HASE is then defined by the sum of the squares of the matrix row norms:

$$HASE = (|b_{11}| + \ldots + |b_{1n}|)^2 + (|b_{21}| + \ldots + |b_{2n}|)^2 \qquad (9)$$

In summary the integrity monitoring algorithm using HASE has to perform the following tasks:

1. Compute HASE for optimum satellites selected for navigation.
2. Determine instantaneous threshold ε_0.
3. If $\varepsilon(i)$ is exceeding the threshold the satellite is rejected and set unhealthy.

The HASE algorithm has an advantage over the PR_{max} algorithm in that it takes into account the instantaneous satellite geometry. However HASE is only an estimate for the actual horizontal error situation.

3.2.3 SPE algorithm

In the SPE algorithm the equations (8-1 and 8-2) are directly used for checking the horizontal error level by inserting the actual pseudurange errors $\Delta PR(i)$. It is so far the most straight forward method for satellite failure detection. In summary the SPE algorithm has to perform the following tasks:

1. Compute SPE for satellites selected for navigation.
2. If the SPE exceeds the allowed level for the phase of flight, reject the false satellite for navigation and set it unhealthy.

This algorithm has the disadvantage that malfunctioning satellites have to be identified on horizontal SPE level which is not a trival task.

Finally it should be outlined that the above described rejection algorithms depend on the satellites being tracked and choosen for the navigation solution. This means that although for demonstration purposes this part of algorithm is programmed on the experimental HMU it should be performed by the user in the operational system, if the range errors are transmitted to him via the civil overlay satellites.

3.3 Satellite failure simulation

For purpose of demonstration and verification of the detection and identification algorithm a satellite failure simulator was added to the experimental monitoring unit on software level. The simulated failures, this means metric errors for up to three satellites are added to the measured pseudoranges. Based on the known satellite failure mechanisms described earlier the typical error time history will be errors of ramp type time functions. Therefore the following kind of failure model is used in the software

$$c = c_0 \left[1 + a_0 + a_1 \left(\frac{t - t_0}{\Delta t} \right) \right] \qquad (10)$$

where

t_0 start of failure

$\Delta t = t_{\text{end of failure}} - t_0$.

Via the menue of the PC the parameters in the above equation can be generated by nine different data failure sources:

- e eccentricity
- $d\Omega_0/dt$ rate of right ascension
- $\sqrt{a}$ square root of semi major axis
- Ω_0 right ascension parameter
- ω argument of perigee
- M_0 mean anomaly
- i_0 inclination
- Δn mean motion offset
- satellite clock/channel time

Based on the failure simulation the performance of the algorithms and the time-to-alarm can be checked.

4. CONCLUSIONS

An experimental health monitoring unit is developed in contract to ESA. The unit is completely integrated concerning GPS and is currently under test. The GLONASS option will be integrated throughout this year. In a second ESA contract another monitoring unit is developed by a spanish company (GMV). After intensive testing of both units ESA is planning to

hand the two units over to INMARSAT and EUROCONTROL. ESA will give both institutions the opportunity to work with the units and to gain experience over a longer time. It is also foreseen to establish an experimental monitoring network with data transmission links to a regional center. In a further definition study in contract to ESA the role of the regional center will be investigated.

5. REFERENCES

1. Rosetti, C.: A civil complement to GPS and GLONASS, RIN89, London 17-19, Oct. 1989.
2. MAN Technologie AG, Kayser-Threde GmbH, Institut für Navigation Stuttgart, Contraves Italiana S.p.A, Politecnico di Torino 1991, NAVSAT Phase B1 Study: System Architecture & Satellite Navigation Package, Final Report, ESA Contract No.: 8933/90/F/RD(SC).
3. Brown, K.: Civil Integrity Requirements for the Global Positioning System, Journal of the Institut of Navigation, Vol. 35, No. 1, Spring 1988.
4. Wells, D. et al: Guide to GPS Positioning, Canadian GPS Associates, Fredericton, N.B., 1986.
5. Wolf, H.: Ausgleichungsrechnung, Ferd. Duemmlers Verlag, Bonn.
6. Interface Control Document ICD-GPS-200, Rockwell International Cooperation, Sept. 1984.
7. Global Satellite Navigation System GLONASS, Inerface Control Document, Research- and Production Association of Applied Mechanics, Glavkosmos, Institute of Space Device Engineering, USSR.
8. Hopfield, H.: Two/quartic tropospheric refractivity profile for correcting satellite data, JGR 74, No. 18, August 20, 1969.
9. Hopfield, H.: Tropospheric effect on electromagnetically measured range: Prediction from surface weather data, Radio Science, 6, No. 3 (1971), pg. 357-367.
10. Black, H.D.: An easily implemented algorithm for the tropospheric range correction, JGR 83, B 4 (1978).
11. Brown, K., Foska, C.: Algorithms for the GPS Integrity Channel, US ION Satellite Division Meeting, 21. Sept., 1988.

COLLABORATION IN THE UK ON TIME TRANSFER USING THE GLONASS SYSTEM

P Daly[1], D J E Knight[2], R Jefferies[2] and B R Swabey[2]

[1]Department of Electronic and Electrical Engineering,
University of Leeds, Leeds, LS2 9JT, UK
[2]National Physical Laboratory, Teddington, Middlesex, TW11 0LW, UK

ABSTRACT

This paper describes collaboration between NPL and the University of Leeds on the use of the GLONASS satellite navigation system for time transfer, and the main results obtained. The University of Leeds has investigated GLONASS clock stability and has monitored GLONASS system time since 1988 within an accuracy of about ± 50 ns. Current news of the GLONASS system is given, and measures towards improving accuracy are discussed.

Keywords: time transfer, GLONASS, GPS, UTC

1. INTRODUCTION

The University of Leeds was the first laboratory outside the former Soviet Union to report observation of signals useful for timing from the GLONASS navigation satellite system [1]. At this point a collaboration began with the National Physical Laboratory to investigate the use of GLONASS for time transfer, and in particular towards improving links with Soviet timescales. A Protocol exists, for example, between NPL and the former Soviet national laboratory, VNIIFTRI to improve links between UTC(NPL) and UTC(SU). Also the improvement of links to Soviet timescales was expected to benefit the international atomic timescale TAI and the international practical timescale UTC through a better link to BIPM of the primary caesium standards and hydrogen masers maintained at VNIIFTRI [2].

Timing stability measurements were made on the GLONASS clocks with an NPL caesium clock at Leeds in 1989 [3], and a continuous caesium timescale UTC(LDS) was established there from December 1990 with help from the Royal Greenwich Observatory, following cessation of RGO's own timescale [4].

Since early 1990, and under a contract with NPL, the University of Leeds has made daily observations of GLONASS time against GPS time and has provided the data to BIPM, which in turn has published it in Circular "T". This information has become more valuable to TAI and UTC since VNIIFTRI began supplying the time difference UTC(SU) - GLONASS time to BIPM from June 1991 [5], which completes an accurate link from UTC(SU) to the western European UTC timescales via UTC(USNO) or GPS time. It should be noted that from January 1992 Leeds University is operating the BIPM GPS common-view schedule with a fully-appropriate timing receiver.

The collaboration complements other work in the United Kingdom on satellite time transfer [6]. This paper summarises the scientific benefits of the NPL/Leeds collaboration to date, and looks towards the more general use of the GLONASS system for time-transfer.

2. THE GLONASS SYSTEM

The GLONASS system is in many respects similar to the GPS system, and is indicated finally to have 24 satellites in 3 orbit planes similarly distributed with respect to the earth. GLONASS satellites transmit on different frequencies whereas the GPS satellites transmit on a common frequency but with different pseudo-random-noise (PRN) codes. Both systems have on board caesium clocks and offer similar accuracies for position and time. The time reference for GLONASS is UTC(SU) and the timescale used is UTC which differs from that of GPS because GPS, like Loran-C and Omega uses a timescale which does not introduce leap seconds. Various compared parameters are shown in Figure 1.

GPS/GLONASS SIGNALS & REFERENCES

	GPS	*GLONASS*
Carrier	**L-band**	**L-band**
Bandwidth	**2 MHz**	**1 MHz**
Access	**CDMA**	**FDMA**
Numbers	**unlimited**	**unlimited**
Coverage	**world**	**near-world**
Accuracy (position)	**<100m**	**<100m**
On-board clocks	**2Cs/2Rb**	**3Cs**
Time reference	**UTC(USNO)**	**UTC(SU)**
Time scale	**GPS**	**UTC**
Coordinates	**WGS-84**	**SGS-85**

Fig. 1: GPS and GLONASS: signals and references.

The recent launching history of GLONASS satellites is shown in Figure 2. Only two of the three orbit planes, labelled I and III, are currently populated with satellites, but of these 7 out of 8 slots were filled with working satellites early in 1992. The number of active satellites is shown by date in Figure 2. and the number has steadily increased from 1989 onwards to the present. It can also be seen that no. 39 which ceased transmitting in 1991 resumed transmitting in early 1992,

indicating that the life of the satellites in orbit is improving. Two more launches of up to three satellites each are anticipated this year, so that the system could rapidly become stronger. More details of the GLONASS system are given elsewhere [7].

Because of the number of different frequency channels in the bands near 1600 MHz used by GLONASS, and because of the out-of-band signals generated in connection with the generation of the spread-spectrum code there have been problems of interference with radio astronomy [8]. However on 28 October 1991 there was a report from IUCAF, the Inter-Union Commission on Frequency Allocations for Radio Astronomy and Space Science, that an agreement had been arrived at with the GLONASS administration to study these problems. A meeting agreed to a resolution to set up a coordinating group comprising members of the GLONASS administration, the USSR Scientific Council for Radio Astronomy and IUCAF, and also to modify future GLONASS satellites with filters to remove the interference in the band 1660-1670 MHz by the end of 1994.

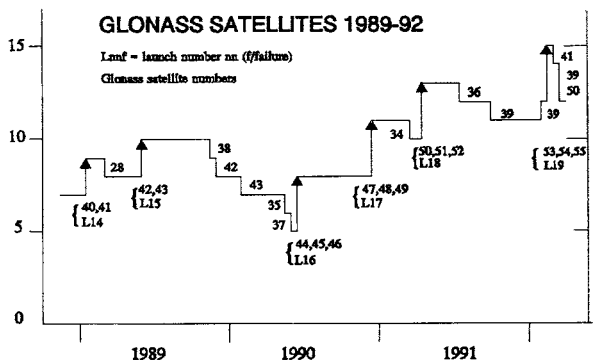


Fig. 2: GLONASS launch history.

3. TIME TRANSFER BY GLONASS AND GPS

Figure 3 shows the equipment arrangement at Leeds University which inter-alia measures the difference between GPS system time and GLONASS system time. There is a receiver of their own construction for receiving the time either from GPS satellites or from GLONASS satellites. This receiver for historical reasons used as its local timescale an output from a Trimble 4000A receiver offering UTC(USNO) with an accuracy of ± 100ns, within which much of the uncertainty arises from a fixed but uncalibrated receiver delay. The position of Leeds University in the timing "chain" via GLONASS and GPS is shown in Figure 4. In the USA the second markers of UTC(USNO) are used as a reference for the GPS timescale and in Russia UTC(SU) is used as a reference timescale for the GLONASS system. In addition UTC(OP) is indicated as a typical western European timescale to which the more-distant UTC(SU) needs to be related in order to acquire good contributions to TAI and UTC from clocks at VNIIFTRI, near Moscow, which operates UTC(SU). The UTC(OP) to UTC(SU) common-view link via GPS is a recent development described in a parallel paper in this session [9]. Our interest here is in the GLONASS system as an alternative to the GPS system, particularly as it appears not to have accuracy-reducing features such as "selective availability".

Observation of the GLONASS satellites from Leeds University is computed as a 24-hour average with the help of BIPM [9] because no common-view arrangements have been possible yet with the VNIIFTRI station. The VNIIFTRI GLONASS measurements have been made in common view with stations to the East of Moscow, which involve satellites not in common-view with Leeds.

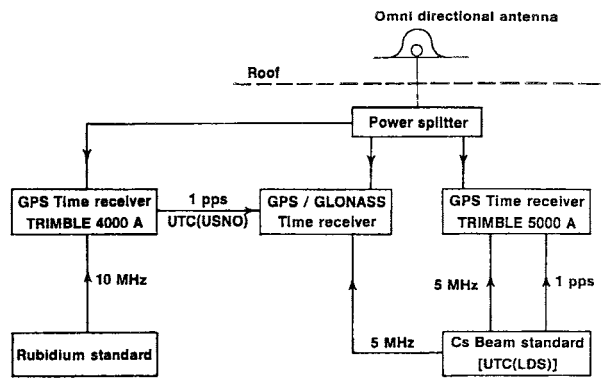


Fig. 3: Time receivers at Leeds.

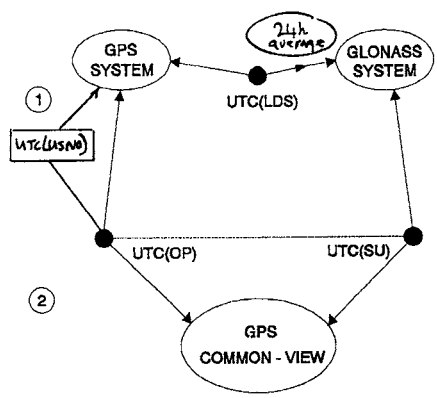


Fig. 4: Time transfer via GPS and GLONASS.

The connection of the timescale UTC(LDS) with other BIPM timescales, including UTC(NPL), is achieved by common view of the GPS system using a separate receiver, see Figure 3. This was a Trimble 5000A, which was replaced in January 1992 by an Allen-Osborne timing receiver, model TTR-6. It will be desirable when the caesium timescale UTC(LDS) is securely operational to relate all Leeds measurements directly to UTC(LDS). The use of UTC(USNO) for GPS and GLONASS measurements goes back to before the installation of the caesium clocks at Leeds.

4. OBSERVATION OF GLONASS SYSTEM TIME

The main contribution of the work at the University of Leeds is from the reporting of the time difference UTC(USNO) - GLONASS system time. This is shown in Figure 5 for the period from about 26 June 1988 to about 15 August 1990. About July 1990 a large step adjustment of about 40 μs was made in GLONASS system time to bring it within about 10 μs of UTC(USNO) and since that time a steady convergence has continued. Since GPS system time is very close to UTC(USNO) this has the effect that the second markers of GLONASS system time and GPS system time have been brought into near coincidence.

The quality of the time links achieved via GLONASS is indicated in Figure 6, which shows the square root of the modified Allan variance as a function of the observation time, using log-log scales. In this the UTC(SU)-GLONASS time data is supplied by VNIIFTRI and UTC(USNO)-GLONASS time is supplied by Leeds University. The satellite observation data has a minimum for observations of about 10 days, beyond which it increases, but that for UTC(USNO)-UTC(SU) continues downwards towards 1 part in 10^{14} at times exceeding about 30 days. The latter is consistent with the use of hydrogen masers at both USNO and VNIIFTRI. The noise in the comparisons is of the order of 3 parts in 10^{13} over 1 day and progressively diminishes out to periods of 100 days.

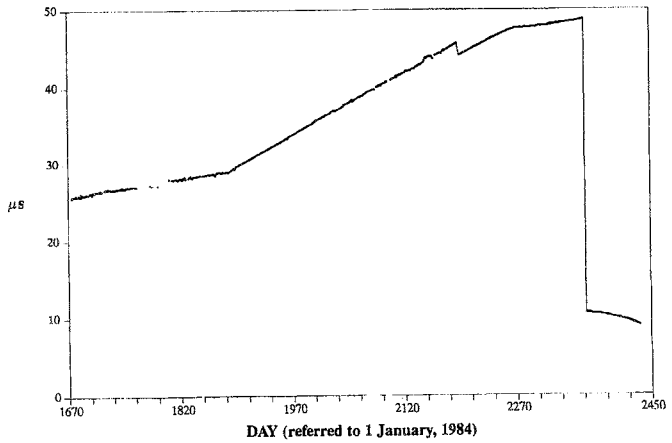


Fig. 5: UTC(USNO) - GLONASS system time, from about 26 June 1988 to about 15 August 1990.

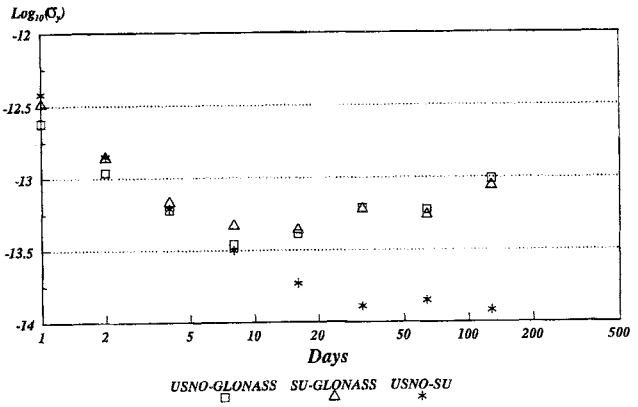


Fig. 6: Square root Allan variance (mod σ_y) of observations.

A long-term check on the accuracy with which the timescale UTC(USNO) is realised at Leeds University via the Trimble 4000A receiver is indicated in Figure 7, which shows the time difference UTC(USNO)-GPS system time as transmitted by satellite (at top), and as measured by Leeds University (at bottom). The average difference of 57 ns between the readings arises from a delay in the Leeds receiver and the standard deviation of this difference averaged over 24h is 26 ns, which is well within the accuracy of 50ns anticipated when this series of measurements was embarked upon.

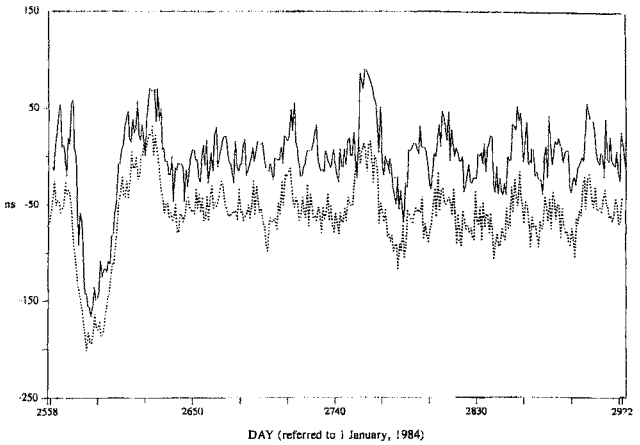


Fig. 7: UTC(USNO) - GPS system time (1991); top: broadcast value, bottom: measured at Leeds.

5. CONCLUSION

Leeds University will continue to observe GLONASS system time against UTC(LDS) etc. and measures are being taken to improve accuracy where possible. A further intention is to participate in common-view time transfer with GLONASS, when other receivers can be made ready, for example in the Russian laboratories, and with NPL using receiver technology from Leeds. GLONASS is expected to get stronger as a system, because the two further launches expected this year are each capable of launching three satellites, and because equipment in the satellites appears to be lasting longer.

ACKNOWLEDGEMENTS

This work was carried out as part of the UK National Measurement System programme for Time and Frequency, & we thank P Raby and S Riley for help in preparing this paper.

REFERENCES

1. P. Daly "Aspects of the Soviet Union's GLONASS satellite navigation system", J. of Navigation, 41 (2) 1988 pp 186-198.

2. VNIIFTRI, State Time and Frequency Service of USSR: Report to CCDS, Bur. Int. Poids Meas. CCDS, 11th session, 1989: Document de travail 89-6, 89-6bis.

3. P. Daly, I D Kitching, D W Allen and T K Peppler "Frequency and time stability of GPS and GLONASS clocks", Int. J. Satellite Communications 9 1991 pp 11-22.

4. RGO, Report to the 11th session of CCDS, Bur. Int. Poids Meas. CCDS, 11th session 1989, Document de travail 89-20.

5. P. Daly, N. Koshelyaevsky, W. Lewandowski, G. Petit and C. Thomas, "Comparison of GLONASS and GPS time transfer between two west European time laboratories and VNIIFTRI", 23rd Annual PTTI meeting, Pasadena CA, Dec 3-5, 1991.

6. J. A. Davis and P. R. Pearce "Analytical techniques for improved time transfer using common view of direct broadcasting satellites", Proc. 6th European Frequency and Time Forum, ESTEC, Noordwijk, Netherlands, 17-19 March 1992.

7. P. Daly "Progress towards joint civil use of GPS and GLONASS", Proc. IEEE PLANS Meeting, Monterey, CA, April 1992.

8. J E B Ponsonby "Spectrum management and the impact of the GLONASS and GPS systems on radio astronomy", J. Inst. Nav. 44 (3) 1991 pp 392-298.

9. P. Daly, N B Koshelyaevsky, W. Lewandowski, G. Petit and C. Thomas", Comparison of GLONASS and GPS time transfers", Proc. 6th European Frequency and Time Forum, ESTEC, Noordwijk, Netherlands, 17-19 March 1992.

Session 7

Plenary Poster Session

DEDICATED OSCILLATOR FOR IMMERSED CRYSTAL SENSOR

G. Bardèche, C. Filiatre and J.-P. Valentin

Equipe de Physique et Electronique des Capteurs
ENSMM - Route de Gray - 25000 BESANCON - FRANCE

ABSTRACT

For chemical applications a quartz oscillator may be successfully implemented. For these applications, the quartz crystal is generally immersed in liquid medium. Therefore, the quality factor and the motionnal resistance are very different compared to usual cases. So, the oscillator circuit must be strongly modified in order to obtain oscillation of the immersed crystal. The characteristics of an immersed quartz crystal are determined and, from the measured values, an electronic design is proposed.

Keywords : quartz sensor, immersed crystal, CCVS quartz oscillator.

I. INTRODUCTION

During the early days of radio communication the frequency quartz oscillator was adjusted by a pencil mark on the crystal. It was the first practical application of Lord Rayleigh's observation, who showed that a small change of the mechanical resonator mass modifies its frequency. At 1957 Sauerbrey [1] suggested the quartz crystal use as a sensor for measuring the thickness of thin vacuum deposited films.

Since the thin film deposition process control by using a quartz crystal oscillator is well known and many extensions were drifted from it : applications to plasma assisted etching, gas adsorption studies on well - characterized films, chemisorption of gases on single element films, contamination studies of spacecraft surfaces, oil contamination of surfaces in vacuum, aerosol mass measurements, and at last analytical chemistry in liquid medium.

In this last area the pioneers have been G. Guilbault, T. Nomura, S. Bruckenstein and K. Kanazawa [2] to [5]. The contribution of these authors was to point out that the oscillation of an immersed crystal resonator was possible, in spite of the general impression in the quartz community which thought that the viscous damping would cause large losses in the quality factor Q leading to cessation of oscillation.

II. SPECIFIC PROBLEMS OF IMMERSED CRYSTALS

Immersion of a quartz resonator introduces a first difficulty which is the holding of the crystal. Standard holders are not convenient because, in the topical case, one crystal face only must be immersed (the explanation of this particularity will be given more far). The crystal may be fixed to the electrochemical cell using silicone rubber sealant or it may be fixed by the means of an O-ring of elastomer set into circular groove in the cell body. Unfortunatly these holdings introduce a residual stress in the quartz resonator and the result is a simultaneous measurement of mass and stress variations during experimentations. Several solutions exist to treate this effect. They would not be presented here.

A second difficulty appears which is the part due to the liquid medium on the immersed crystal plate oscillation. Using crystals oscillations in thickness-shear mode, the surface motion is similar to that on an oscillating flat plate. Hydrodynamic coupling occurs between the vibrating crystal and the liquid. That coupling increases crystal motion away from the electrode edge and excites the surrounding region of the crystal into motion. So, the normal gaussian repartition of the mechanical oscillation amplitude observed in the case of trapped energy resonator is modified. On the other hand, in the liquid medium the amplitude of mechanical oscillation decreases because the liquid viscosity and density increase the shear stress on the crystal surface. Due to the non uniform surface velocity and the continuity equation, vertical liquid motion must occur in addition to the horizontal motion induced by viscous drag. So, immersed crystals need a precise theoretical study.

The third main difficulty is the problem of electrical potentials. If the liquid medium is an organic solution both sides of the crystal may be immersed but in the case of electrolyte solutions, one side only may be immersed for avoid short circuit between the crystal electrodes.

Besides it is suitable to connect the immersed electrode to the grounded potential, in order to impose a fixed reference potential inside the electrolyte. This fact reduces drastically the choice of the oscillator circuit among the usual designs of quartz oscillators i.e. Buttler, Colpitts, Hartley, Miller, Pierce, for only quoting the main designs.

III. MEASURED CHARACTERISTICS FOR IMMERSED CRYSTALS.

The quartz crystal can be represented by the usual circuit constituted by the motional branch in parallel with a capacitance Co. It is connected to a HP 4195 A network analyzer by a special connector and so that one face is in contact with the liquid. We measured the equivalent circuit parameters in four different conditions : in normal atmosphere ; one face in contact with pure water18 MΩ; one face in contact with a saline solution at the 1 mole/l concentration ; the same solution diluted by a factor 5. The results are given in the table 1.

PARAMETERS	NORMAL ATMOSPHERE	WATER 18 MΩ	DILUTED SALINE SOLUTION N/5	SALINE SOLUTION N
Resonant serie frequency (Hz)	5 993 825	5 991 000	5 990 650	5 989 400
R (Ω)	20	480	500	720
L (mH)	37	33	31	16
C (fF)	19	21	23	44
C_o (pF)	4	9	12	16
$Q(\frac{L\omega}{R})$	70 000	2 600	2 300	800

Table 1 : Equivalent circuit parameters measures

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

The frequency decreases when the quartz is in contact with a liquid. Kanazawa [6] proposed an equation to modelize the frequency variations when one face of the quartz crystal is in contact with a liquid. The frequency depends on the solution density and viscosity.

The motional resistance R increases by a factor 20, when one face is in contact with water and by a factor 30 when the liquid is a saline solution. The inductance decreases slightly when one face is in contact with water or a diluted solution and by a factor 2 with the saline solution. The capacitance increases in the same manner. The capacitance C_o increases regularly. At present the parameters variations are only explained for C_o variations, the relative permittivity of C_o capacitor being increased by edge effects through the liquid medium.

The quality coefficient is very affected by the liquid contact and falls to about 1 000. This fact is well shown on the admittance modulus and phase versus frequency curves (figure 1).

They are drawn with the network analyzer for a 6 MHz AT-cut plane quartz crystal.

If we compare the curves when the quartz is in normal atmosphere and when one face is in contact with water, we can make several remarks :

- at the series resonance frequency, the admittance modulus decreases (note that the scale is divided by 10 for the second curve). A bad ratio around 10 appears between the extreme admittance values.

- the curves sharpness are degraded.

- in normal atmosphere the phase varies from + 90° to -90°. In the liquid medium the phase goes through zero and only reaches 20°. The phase margin is small from electronic point of view.

Sometimes the phase does not go through zero when the pure solution is in contact with one face; so not any oscillator is achieved without phase external adjustment.

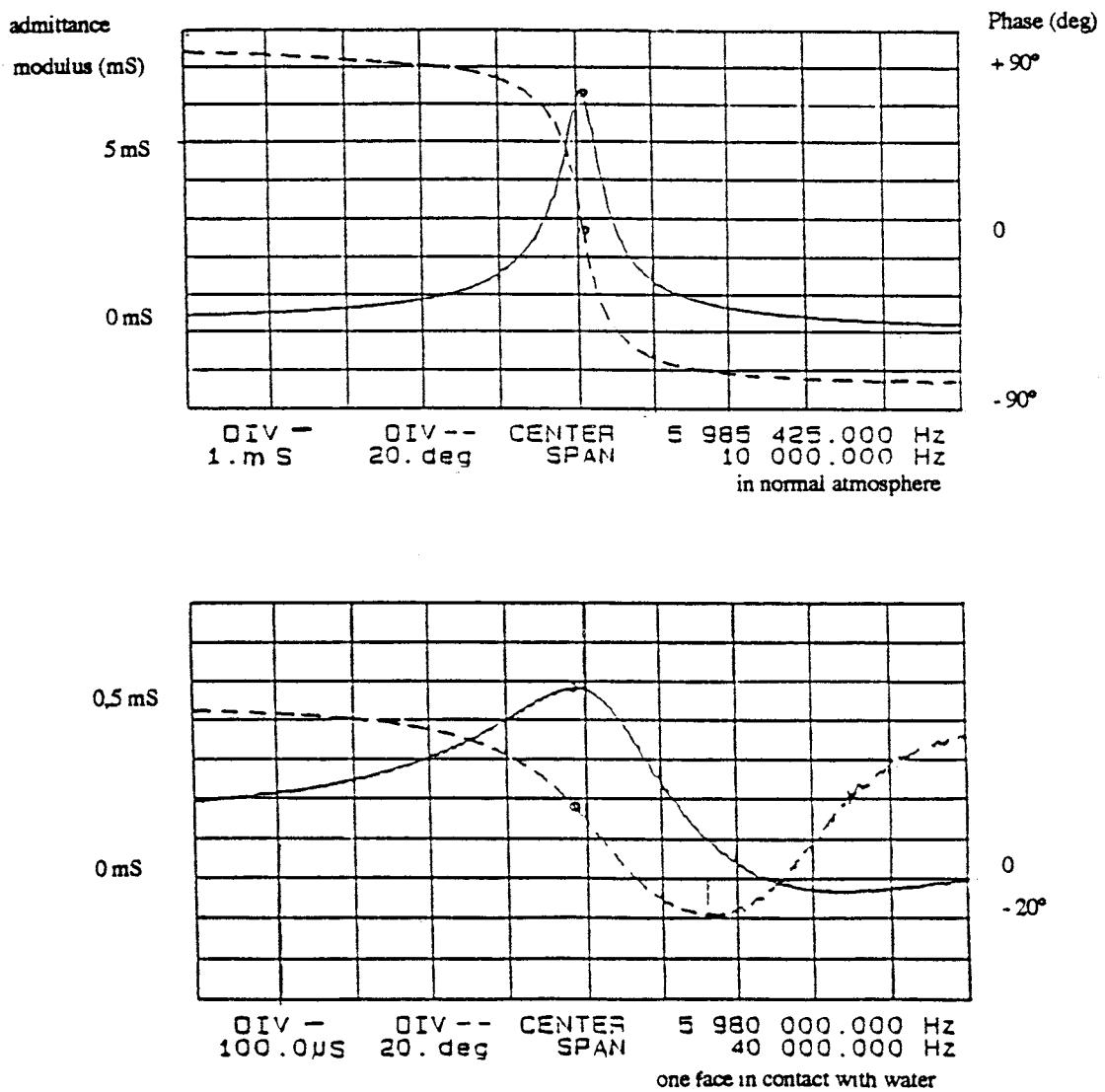


Figure 1 : Admittance modulus (-) and phase (...) versus frequency near the series resonance

IV. OSCILLATOR DESIGN

The use of a quartz crystal to define an oscillator frequency and to measure its variations needs an electronic design to sustain these oscillations.

Among the usual circuits for the metrological applications, we need a circuit with a grounded electrode, like the Colpitts circuit. But in the conditions we described before, it is difficult to realise. The frequency is indeed defined by several complex parameters and the quartz crystal parameters are not necessary the determining factors. Moreover the reproductibility, the stability and the sensitivity are obtained with drastic conditions.

For these reasons we began the study and the realisation of some specific circuits in the aim to control the oscillator running parameters, so to achieve measures with a better accuracy.

The figure 2 shows the principle of the circuit presented in this paper.

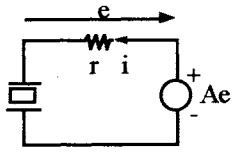


Figure 2 : Principle of the studied oscillator

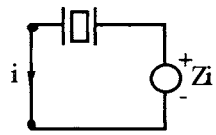


Figure 3 : Principle of the CCVS oscillator

The quartz crystal works in series resonance with a grounded electrode. This circuit can be considered like an approximation of the optimal circuit where the resonant element is loaded by the zero impedance of a current-controlled voltage source (CCVS) (figure 3).

In our case the quartz motional resistance R reaches a value of some hundred ohms. So it is possible to choose the resistance r less than R and at the same time the load (r + R) not too small for the amplifier. For this one, we considered an integrated video amplifier with a gain bandwidth product greater than 100 MHz and a slew rate greater than 100 V/µs. In fact its use may be doubtful because it is impossible to adjust the amplifier characteristics.

The amplifier proposed in figure 4 is a development of the usual differential stage

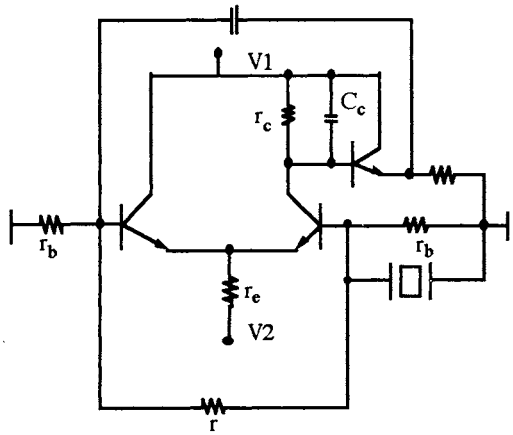


Figure 4 : Oscillator circuit

This circuit is characterized by different particular aspects :

- UHF Darlington transistors circuit minimize the quartz shunt capacitance. Also the low base bias current allows the choice of high resistance value r_b.

- the L.F. dynamic stability and the absence of relaxation oscillations are obtained by giving to the stage a negativ differential gain with a modulus < 1, which implies the

conditions : $0 < gr_c \left[\frac{\alpha_o - f(r_e)}{1 + f(r_e)} \right] < 1$

where $f(r_e) = \frac{gr_e}{1 + gr_e}$ and $\alpha_o = \frac{r_b}{r + r_b}$

g is the transistor transductance.

On the other hand at the quartz frequency where $Z_q \approx R$, we want to obtain

$$gr_c \left[\frac{f(r_e) - \alpha_q}{1 + f(r_e)} \right] > 1$$

where $\alpha_q = \frac{R}{r + R}$

We can notice that these conditions prevent from using a current source use instead of the resistance r_e.

- it follows from the quartz capacitance behavior at high frequency that the gain increases and the phase rotates contrary to the H.F. transistors response. The gain control rules out the H.F. oscillations with the help of a compensating capacitance C_c shunting the resistance r_c.

- a common-collector stage ensures a low output impedance.

This study is performed with the circuit simulation software PSPICE. The figure 5 shows the effect of the compensating capacitor C_c on the open loop transfert function T(f). Under the same conditions the figure 6 details T(f) around the quartz resonance frequency. The response analysis can predict a C_c value close to 50 pF in this case.

Oscillation production, stability control and sensitivity determination can be studied by time simulation. The sensitivity of the oscillation frequency versus various parameters can be obtained. For example the expected frequency variations in function of motional inductance represent 80 % of the theoretical variation.

1) Sauerbrey Z.Phys., 155, p 206-212,1959.

2) J. Hlavay J., Guilbault G.G. Anal. chem., 49, p 1890, 1977.

3) Nomura t., Iijima m.J. Anal chim. acta, 131, p 97-102, 1981.

4) Bruckenstein S., Shay M.J. Electroanal. chem., 188, p 131-136, 1985.

5) Kanazawa K.K., Gordon II J.G. Anal. chem., 57, p 1770-1771, 1985.

6) Kanazawa K.K., Gordon II J.G. Anal. chim. acta., 175, p 99-105, 1985.

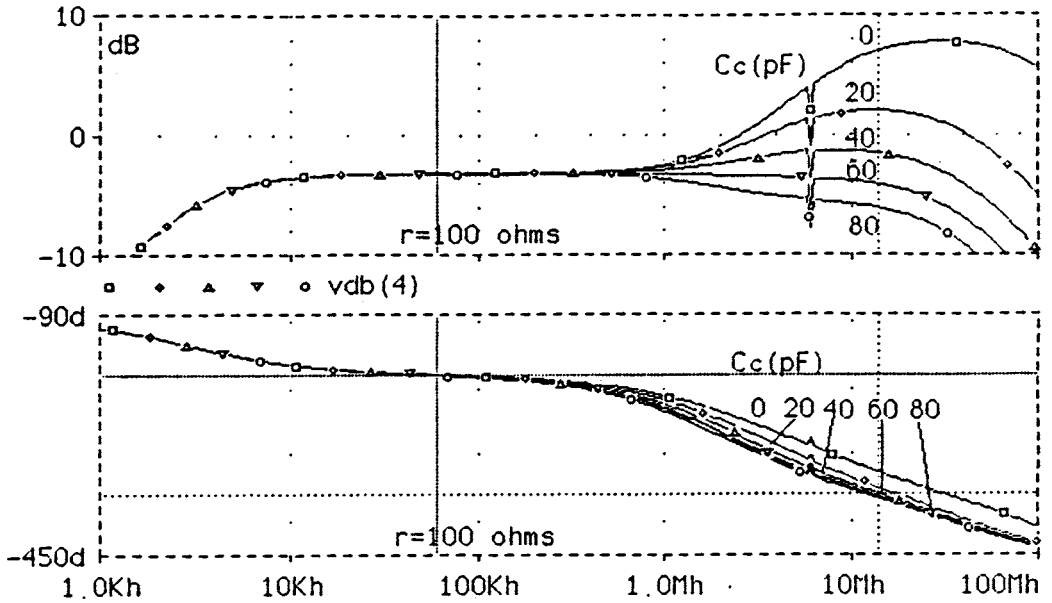


Figure 5 : Broadband response with a compensating capacitance C_c

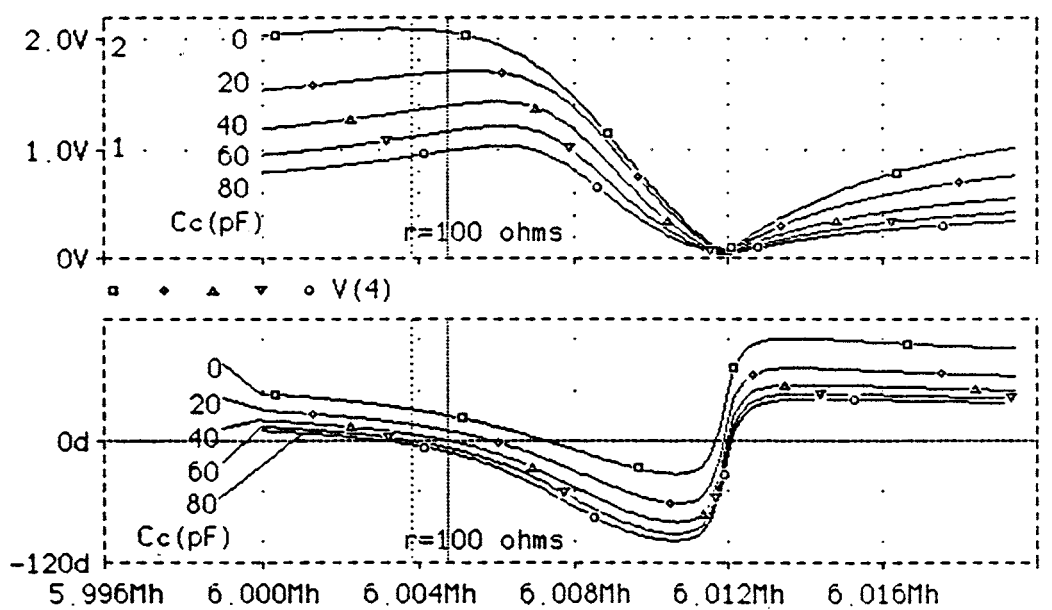


Figure 6 : Response close to the quartz resonances

RADIATION EFFECTS IN GERMANIUM DOPED QUARTZ

Harish Bahadur

National Physical Laboratory, Dr. K.S. Krishnan Road, New Delhi–110012 INDIA

ABSTRACT

Cultured quartz crystals doped heavily with Ge (Ge concentration ~3000 ppm) have been studied for their infrared spectral characteristics as a function of irradiation temperature. The irradiation was done in sequence at 77K, 300K and finally again at 77K. Infrared spectra show that in addition to the growth-defect bands the previously unreported peak at 3656 cm^{-1} is caused due to the presence of Ge in quartz. Intense irradiation at 77K in prior 300K-irradiated samples exhibit the presence of a band at 3400 cm^{-1} which is a different defect center than the growth-defect band at this frequency. Production curve for this band shows a linear dose dependence.

Keywords:

Quartz crystals, infrared spectra, Ge-doping, 77K-irradiation, 300K-irradiation, dose dependence.

1. INTRODUCTION

Quartz oscillator crystals when exposed to ionizing radiation suffer both transient and steady frequency offsets (Ref.1 and those cited in it). These frequency offsets are undesirable in a number of precision aerospace applications. The frequency offsets are known to be due to impurities present in quartz material. Aluminum is the most pervasive impurity that substitutionally replaces Si^{4+} in quartz lattice. Being a trivalent substitutional impurity it needs a charge compensation that is provided by an alkali ion or an electron hole. In some cases such as natural quartz or irradiated quartz the aluminum is also charge compensated with protons. Aluminum thus acts as a hole or alkali trap.

Substitutional aluminum and a group of unidentified hydrogen-containing growth-defect centers are probably the most importan point defects in quartz from the point of view of resonator performance. In the as-grown material, the aluminum is charge compensated by an interstitial alkali forming the Al-Li or Al-Na centers. If an as-grown quartz sample is irradiated above 200K the interstitial alkali is removed from the aluminum site and replaced by a mixture of Al-OH and Al-hole centers (Ref.2,3). In this case, the hydrogen comes from the growth-defect centers (Ref.4). The Al-Na center has a strong acoustic loss peak at 53K and a weaker peak at 135K. (Ref.5). Three acoustic loss peaks, one at 23K, a second at 100K and a third at 135K seem to be associated with the presence of Al-hole center (Ref.6). At low temperatures, the hydrogen-containing growth-defect centers produce infrared absorption bands at 3350,3400, 3437 and 3581 wvenumbers which thermally broaden at room temperature so that only the high frequency band is still resolved.

Germanium (Ge^{4+}) is another impurity which when present in quartz isomorphously replaces Si^{4+} in the SiO_4 tetrahedron. Ge acts as an electron trap. Irradiating a sample produces a large number of electrons and holes which move through the crystal until they recombine or are trapped by Ge and the free holes forms Al-hole center. At irradiation about and above 200K, the alkali ion diffuses from aluminum and stabilizes the Ge-defect site forming in the process the $[Ge(Al)e^-/M^+]^o$ and $[Ge(C)e^-/M^+]^o$ with M^+ =Li or Na. Weil (Ref.7) has given a complete discussion of the paramagnetic centers found in Ge-containing samples.

While the contribution of aluminum in radiation response of quartz has been extensively studied (Ref. 8), the Ge effects are relatively ignored and less studied; possibly due to the fact that Ge effects are relatively small and unrecognized on oscillator frequency. Recently, Euler and Kahan (Ref.9) have reported the production of a radiation-induced anelastic peak at 249K in cultured quartz crystals doped with about 30 ppm of germanium. By comparing the production of this loss peak with the production of paramagnetic Ge-Li centers and the aluminum-hole center, they were able to show that the Ge(C)-Li center is responsible for this 246K anelastic peak. We have observed this peak in two different natural quartz crystal resonators (Ref.10) and not in conventionally grown quartz. Recently, Halliburton, Chen and Tapp (Ref.11) have however found that most commercial quartz material contain Ge impurity of the order of 0.1 ppm. Literature survey has shown that studies on infrared spectral characteristics of Ge doped quartz have not so far been carried out. Euler and Kahan (Ref.9) have found that in the Ge doped quartz the infrared spectra did not show any novel peaks or unusual peak strengths, which could be attributed to Ge defects. These observations nonetheless prompted us to extend our investigations on irradiating temperatures and accumulated doses. In the present work we have examined, using low temperature infrared absorption measurements in 3100-3700 cm^{-1} range two cultured quartz crystals with high Ge concentration. One of the crystal was a Sawyer Ge doped quartz and the other was grown at Bell Laboratories, Murray Hill, NJ, U.S.A.

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

2. EXPERIMENTAL DETAILS

Optical samples were prepared from the Z-growth region and attached to the cold finger of a cryostat capable of reaching 77K. Quartz samples were irradiated in the cryostat itself. The radiation port for the cryostat consisted of a 0.005-in-thick aluminum window. The tail of the cryostat could be rotated by 90^{o} after the irradiation to position the CaF_2 windows in the path of light beam. Near infrared absorption was measured with a Beckmann 4240 double beam spectrophotometer in which the ratio of sample and reference signals was taken electronically, thereby allowing a direct continuous recording of the optical density as a function of wave numbers. The irradiation was carried out by a Van de Graaff electrostatic accelerator. The acceleration voltage was 1.75 MeV and the current density on the sample was 10 $\mu A/cm^2$. Irradiation time was controlled by a Gralab model 900 programmable timer. Crystals were irradiated in a sequence of 77K, 300K and finally again at 77K.

3. RESULTS AND DISCUSSION

It must be mentioned at the outset that both the Ge-doped quartz crystals exhibited 'all' of the growth-defect bands i.e., the bands at 3350, 3400, 3437 and 3581 cm^{-1} which are normally seen in the cultured quartz (Ref. 12-14) were present in the samples studied. Comparative studies on IR spectra and irradiation effects on the two crystals reveal the following systemic similarities and differences in the two types of Ge-doped quartz crystals.

(i) Both crystals exhibited the presence of a previously unreported IR band at 3656 cm^{-1}. This band appears to be Ge-related infrared active defect center and is better resolved in Sawyer crystal which is a slowly grown material. This band appeared in much smaller strength compared to the growth-defect bands. Irradiation at 77K did not remove this band for radiation doses as high as upto 20 Mrad in both crystals. Room temperature irradiation (dose = 2 Mrad) removed this band completely in the fast grown crystal. On the other hand in the Sawyer crystal the room temperature irradiation (dose = 2 Mrad) reduced this band by about 20%.

(ii) A very small band at 3506 cm^{-1} was produced in both the crystals at the first 77K- irradiation and was removed at room temperature irradiation and did not show up on subsequent irradiations at low temperatures.

(iii) A small band at 3525 cm^{-1} was produced at room temperature irradiation (dose = 2 Mrad) in both the samples. Subsequent low temperature irradiations reduced this band but it did not completely vanish even after strong radiation doses tried upto 16 Mrad.

(iv) A band at 3602 cm^{-1} was observed upon strong irradiation (dose = 8 Mrad) at 77K only in the fast grown crystal which was irradiated previously at room temperature. This band increased to its double strength upon further irradiation with similar dose. Room temperature irradiation was a necessary step for the production of this band at subsequent low temperature irradiation.

(v) Final irradiation at 77K produced in both the crystals a band at about 3400 cm^{-1}. Figure 1 shows the spectra of as-received and those obtained upon final 77K-irradiation. It may be mentioned here that this band at 3400 cm^{-1} is different than the 3400 cm^{-1} growth-defect band found in conventionally grown as received cultured quartz. Therefore, it is suggested that a different species of a defect center other than the as-grown-defect center get produced at this stage. It was found that by increasing the radiation dose at 77K the new band produced at 3400 cm^{-1} got strengthened. Figure 2 shows the production curve of the new 3400 cm^{-1} band for the the fast grown quartz. Again, like in the production of 3602 cm^{-1} band at 77K-irradiation, the room temperature irradiation was a necessary step for the production of the 3400 cm^{-1} band also.

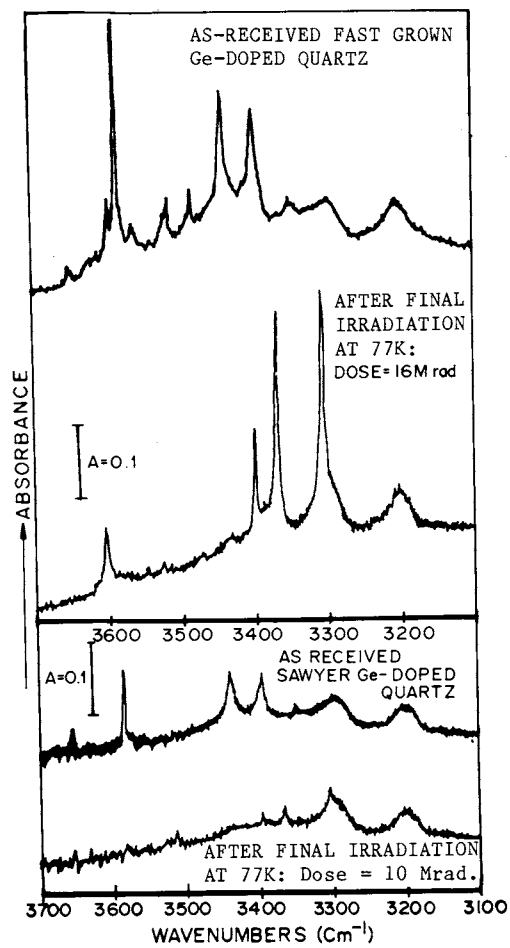


Figure 1. Infrared spectra of Ge-doped quartz crystals in their as-received and irradiated conditions.

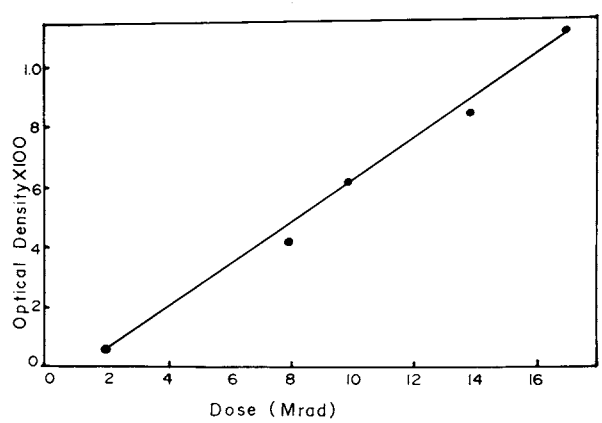


Figure 2. Production curves for radiation-induced 3400 cm^{-1} band in the fast grown crystal.

(vi) Al-OH^- bands at 3306 and 3367 cm^{-1} are produced upon room temperature irradiaton in both the crystals. This fact is in accordance with earlier work (Ref.2-6, 8-14) on cultured crystals. However, the remarkable feature seen here is that at the final irradiation at 77K even with the high radiation dose (~10-15 Mrad) these bands are still observable. Sawyer crystals of this high purity of aluminum without being Ge-doped show that these bands vanish at much lower radiation dose (e.g. Ref.2). Probably, this fact is a characteristic of Ge introduced into the crystals. Hayes and Jenkin (Ref.15) have shown that Ge in quartz acts both as a hole trap as well as an electron trap at low temperature. Thus, the presence of Ge will restrict the occurrence of recombination electrons and holes near the Al-OH^- centers. This will consequently limit the removal of protons from these centers which can normally occur when the electron and hole recombination takes place around Al-OH^-. On the other hand the strength of the two bands becomes equal exhibiting that both the oxygens around the aluminum are equally populated. This corroborates earlier work (Ref. 2,13).

3.1 Isochronal Anneling:

Isochronal anneal, following an intense 77K-irradiation (Figure 3) of Al-OH^- bands show that like in the case of other cultured crystals (e.g. Ref.13) and natural crystals (Ref.12), both the bands recover almost uniformly uptil about 200K and after that a reorientation takes place and the 3367 cm^{-1} (e_2) grows and the 3306 (e_1) reduces. At room temperature the e_2 band attains a value about seven times of e_1. This intensity ratio was also the same in the case of Sawyer Ge-doped quartz. In conventionally grown quartz the e_2 band is about three to five times of e_1 band. It is suggestive from the present studies that this ratio increases due to germanium doping.

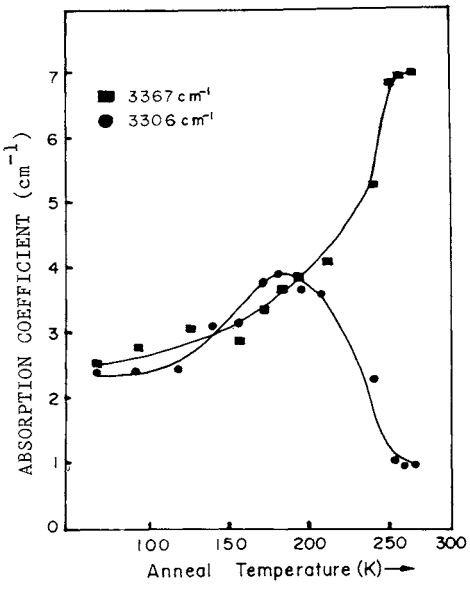


Figure 3: Isochronal anneal of Al-OH^- centers in Ge-doped cultured quartz.

4. SUMMARY AND CONCLUSION

Ge-doped cultured quartz crystals have been investigated for their radiation response in the near infrared region. The crystals were subjected to a sequence of low temperature, room temperature and finally again at low temperature. It has been found that heavily doped with Ge the cultured quartz crystals show that in addition to the growth-defect bands the previously unreported band at 3656 cm^{-1} is caused due to the presence of Ge in quartz. Intense irradiation at 77K in prior 300K-irradiated samples exhibit the presence of a band at 3400 cm^{-1} which is a different defect center than the growth-defect band at this frequency. Production curve for this band shows a linear dependence.

The Al-OH^- bands in Ge-doped quartz are produced at room temperature irradiation as in the case of conventionally grown quartz. However, these bands do not vanish at final low temperature irradiation even for very high radiation doses. Comparison with conventionally grown cultured quartz shows that at this stage these bands completely vanish at relatively much less dose. Isochronal anneal data shows that like the case of conventionally grown cultured quartz both the Al-OH^- bands recover in about equal strength uptil about 200K. Beyond this temperature the long bond OH^- band (3367 cm^{-1}) increase and the sort bond OH^- band decreases and both of these bands saturate at about 300K. The present studies indicate that the ratio of strength of Al_l /Al_s is somewhat more in Ge-doped quartz than the conventionally grown crystals; Al_l is the long bond Al-OH and the Al_s represents the short bond Al-OH^-.

5. ACKNOWLEDGMENTS

This work was done at Oklahoma State University in the laboratory of Professor J.J. Martin and was supported by contracts with Sandia National Laboratories, Albuquerque, NM and Rome Air Development Command, Hanscom, MA under the principal investigator Professor J.J. Martin.

6. REFERENCES

1. King JC 1985, Radiation Effects on Resonators in Precision Frequency Control, Eds. E.A. Gerber and A. Ballato, Academic Press, Orlando, FL , 147-159 .

2. Halliburton LE, Koumvakalis N, Markes ME and Martin JJ 1981, Radiation Effects in Crystalline SiO_2 J.Appl. Phys. 573565-3674.

3. Markes ME & Halliburton LE 1979, Defects in Synthetic Quartz: Radiation-induced Mobility of Interstitial Ions, J. Appl. Phys. 50, 8172-8180.

4. Halliburton LE, Markes M, Martin JJ, Doherty SP, Koumvakalis N, Sibley WA, Armington, AF, & Brown RN 1979, IEEE Trans. Nucl. Sc. NS-26,4851-4856.

5. Martin JJ, Hwang HB & Bahadur H 1985, Radiation Effects in the Acoustic Loss Spectra of AT-Cut Quartz Crystals, Proc. 39th Annual Frequency Control Symposium, Philadelphia, PA, USA, IEEE Document no. 85CH2186-5 266-269.

6. Martin JJ 1984 Aluminum-related Acoustic Loss in AT-cut Quartz Crystals, J. Appl. Phys. 56, 2536-2540.

7. Weil JA 1984 , A Review of Electron Spin Resonance Spectroscopy and Its Applications to the Study of Paramagnetic Defects in Crystalline Quartz, Phys. Chem. Minerals, 10, 149-165.

8. Martin JJ 1990 Radiation-induced Frequency Offsets and Acoustic Loss in AT-cut Quartz Crystals, J. Appl. Phys. 68, 5095-5104.

9. Euler FK & Kahan A 1987, Radiation Effects and Anelastic Loss in Germanium-doped Quartz, Phys. Rev., 35, 4351-4359.

10. Martin JJ, Hwang HB, Bahadur H & Berman GA 1989, Room-temperature Acoustic Loss peaks in Quartz, J. Appl. Phys., 65, 4666-4670.

11. Halliburton, LE, Chen CY & Tapp SD 1985, Radiation Effects in Quartz: Low Doses an dDefect Production Mechanisms, Proc. 39th Annual Frequency Control Symposium, Philadelphia, PA, USA, IEEE Document no. 85CH2186-5 259-265.

12. Bahadur H 1989, Infrared Characterization of Natural and Cultured Quartz: The Effect of Electrodiffusion and Irradiation, J. Appl. Phys., 66, 4973-4982.

13. Sibley WA, Martin JJ, Wintersgill MC & Brown JD 1979, The Effect of Radiation on the OH^- Infrared Absorption of Quartz Crystals, J. Appl.Phys., 50, 5449-5452.

14. Lipson HG & Kahan A 1985 Infrared Characterization of Aluminum and Hydrogen Defect Ceneters in Irradiated Quartz, J. Appl. Phys., 58, 963-970.

15. Hayes W & Jenkin 1986, Charge-trapping Properties of Germanium in Crystalline Quartz, J. Phys. C:Solid State Phys. 19 , 6211-6219.

TRAPPED ENERGY QUARTZ CRYSTAL RESONATOR on WEDGE-SHAPED PLATE.

R . BOURQUIN

Laboratoire de Chronométrie Electronique et Piézoélectricité
Ecole Nationale Supérieure de Mécanique et des Microtechniques

Route de Gray 25030 BESANCON Cedex - FRANCE -

ABSTRACT

Based upon the general relation of dispersion for a quartz plate of doubly rotated cut , an analytical model is presented. It describes the modes shapes of a resonator exhibiting a linear thickness variation along in-plane directions and is only submitted to the following assumptions: small wave numbers along the plate plane and slowly varying thickness. The proposed model allows a direct computation of overtones and anharmonic modes frequencies,as well as the motional elements of the resonator, as functions of the tilt angle

Keywords: Mode shape, Resonant frequencies, Motional parameters.

1. INTRODUCTION

In 1983, Dworsky [1] presented a study of the properties of AT quartz resonators on wedgy plates. Instead of the numerical method used in that work to obtain the distribution of the vibration in the plane of the resonator, we give here the analytical form of the vibration. The starting point is the dispersion relation for a quartz plate, also used in the case of trapped energy or contoured resonator [2]. By using the corresponding differential equation, and taking into account of a slowly linearly varying thickness, the linear differential equation governing the modes of the resonator in the vicinity of a pure thickness mode is obtained. The use of a proper change of variable enables us to obtain the analytical expression of the displacement in the resonator.

By inserting this solution in the continuity conditions at the junctions between the different regions of the trapped energy resonator transcendental equations are obtained, which are numerically solved to obtain the wavenumbers along the main directions in the mean plane of the resonator.

2. THICKNESS MODES IN A WEDGY PLATE

In this section we consider a doubly rotated quartz plate as shown by the fig.1. The X_i-axes are the cristallographic axes, whereas the X'_i-axes are obtained from the previous axes after rotations of φ and θ angles. It is convenient to refer the studied plate to a third coordinate system xi , deduced from the X'_i- axes by a rotation of angle ψ about X'_2. In this new coordinates system, the equation governing a thickness mode vibration of a free plate has been established by Stevens and Tiersten [2] :

$$M^*_n u_{,11} + P^*_n u_{,33} - \left(\frac{n^2 \pi^2 C_e}{4h^2} - \rho\omega^2 \right) u = 0 \qquad (1)$$

where $u(x_1,x_3)$ is the in-plane variation of the mechanical displacement u_M :

$$u_M(x_i, t) = u(x_1, x_3) \sin \frac{n\pi x_2}{2h} e^{j\omega t} \qquad (2)$$

C_e is the effective elastic coefficient corresponding to the propagation of a plane wave in the x_2 direction for the considered mode (A,B or C-mode), obtained by solving the Christoffel's determinant and u_M is the mechanical displacement along the corresponding eigenvector [3] .n is the number of overtone. The coefficients M^*_n and P^*_n and the ψ -angle are given in ref [2].

If the thickness is slowly varying , equation (1) remains valid. In our case, we consider a plate exhibiting a small defect of parallelism of its faces. So we assume that h is a linear function of x_1 , x_3 :

$$h(x_1, x_3) = h_o \left(1 + \frac{\gamma_1}{2} x_1 + \frac{\gamma_3}{2} x_3 \right) \qquad (3)$$

By expanding $1/h^2$ to first order in x_1 and x_3 , eq (1) becomes :

$$M^*_n u_{,11} + P^*_n u_{,33} - \left[\frac{n^2 \pi^2 C_e}{4h_o^2} \left(1 - \gamma_1 x_1 - \gamma_3 x_3 \right) - \rho\omega^2 \right] u = 0$$

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

Substituting

$$u(x_1, x_3) = X(x_1)\,Z(x_3) \tag{5}$$

we obtain two independant equations :

$$X_{,11} + \left(\alpha_M x_1 + m^2\right) X = 0$$

$$Z_{,33} + \left(\beta_M x_3 + p^2\right) Z = 0 \tag{6}$$

where

$$\alpha_n = \frac{n^2 \pi^2 C_e}{4h_o^2 M_n^*}\gamma_1$$

$$\beta_n = \frac{n^2 \pi^2 C_e}{4h_o^2 P_n^*}\gamma_3 \tag{7}$$

The decay or wave numbers m and p are submitted to the relation

$$\frac{n^2 \pi^2 C_e}{4h_o^2} + M_n^* m^2 + P_n^* p^2 = \rho\omega^2 \tag{8}$$

Eq (6) are of the form

$$\frac{d^2 f}{dx^2} + (\alpha x + \upsilon) f = 0 \tag{9}$$

By using the substitution

$$z = \frac{2}{3\alpha}\left(\alpha x + \upsilon^2\right)^{3/2} \tag{10}$$

if z is real (i.e. if $\alpha x + \upsilon^2 > 0$) we obtain the Bessel equation

$$\frac{d^2 v}{dz^2} + \frac{1}{z}\frac{dv}{dz} + \left(1 - \frac{1}{9z^2}\right) = 0 \tag{11}$$

with

$$f(x) = \left(\alpha x + \upsilon^2\right)^{1/2} v(z) \tag{12}$$

So the general solution of (9) is expressed in terms of ordinary Bessel functions of order +-1/3 [4]

$$f(x) = \left(\alpha x + \upsilon^2\right)^{1/2}\left[A J_{+\frac{1}{3}}(z) + B J_{-\frac{1}{3}}(z)\right] \tag{13}$$

Conversely , if $\alpha x + \upsilon^2$ is < 0 , we obtain a modified Bessel equation instead of (11) . The general solution of (9) is then

$$f(x) = \left(-\alpha x - \upsilon^2\right)^{1/2}\left[-A I_{+\frac{1}{3}}(z) + B I_{-\frac{1}{3}}(z)\right] \tag{14}$$

with in this case

$$z = \frac{2}{3\alpha}\left(-\alpha x - \upsilon^2\right)^{3/2} \tag{15}$$

3. TRAPPED ENERGY RESONATOR ON A WEDGY PLATE

The figure 2 depicts a trapped energy resonator with rectangular electrodes of length 2l and width 2b . The thickness and mass density of the electrodes are 2h' and ρ' respectively . The parameters γ_1 and γ_3 are supposed positive , without loss of generality .

For such a trapped energy resonator , we can define nine different regions : one electroded region, denoted (-) , four lateral regions (1's and 3's) and four corner regions (2's) . As usual in this kind of treatment , we can state that the mechanical displacement and its derivatives must be continuous accross the boundaries of adjacent regions.

In the electroded region , the effective elastic coefficient C_e appearing in (1) is replaced by

$$\overline{C_e} = C_e\left(1 - \frac{8e_{26}^2}{n^2\pi^2\varepsilon_{22}C_e} - \frac{4\rho' h'}{\rho_o h_o}\right) \tag{16}$$

Thus , the $\overline{m}$ an $\overline{p}$ wave numbers along x_1 and x_3 axis , respectively , differ from the corresponding quantities m and p in the unelectroded regions. These wave numbers verify the following relations

$$M_n^* \overline{m}^2 + P_n^* \overline{p}^2 + \frac{n^2\pi^2}{4h_o^2}\overline{C}_e = \rho\omega^2$$

$$M_n^* m^2 + P_n^* \overline{p}^2 + \frac{n^2\pi^2}{4h_o^2} C_e = \rho\omega^2 \tag{17}$$

$$M_n^* \overline{m}^2 + P_n^* p^2 + \frac{n^2\pi^2}{4h_o^2} C_e = \rho\omega^2$$

The geometry of the resonator in fig. 2 assumes that both coefficients M^*_n and P^*_n are positive. So the squares of the wave numbers in the electroded region , $\overline{m}^2$ and $\overline{p}^2$ will be positive if the defect of parallelism is small , because $\overline{C_e}$ is less than C_e . In the unelectroded regions , m^2 and p^2 are negative .

The only physically reasonable solutions must tend tozero when x_1 (or x_3) tend to the infinity . This is always verified by the solution given in (13) , but this is not the case of the solution given in (14) , which involves modified Bessel functions . By examining the asymptotic expansion of $I_n(z)$ [4] , we can see that the convergence of the solution (14) when z tend to+ infinity requires that B = A

Since in our discussion we suppose γ_1 and γ_3 to be positive, the coefficients α_n and α_n are also positive . So, it is convenient to write the mechanical displacement in

a trapped energy resonator exhibiting a small defect of parallelism of its faces in the following form [5]

in the electroded region :

$$\bar{u} = \bar{A}\left(\alpha_n x_1 + \bar{m}^2\right)^{1/2}\left(\beta_n x_3 + \bar{p}^2\right)^{1/2}$$

$$\left[J_{+\frac{1}{3}}\left(\bar{z}_1\right) + \bar{K}_1 J_{-\frac{1}{3}}\left(\bar{z}_1\right)\right]\left[J_{+\frac{1}{3}}\left(\bar{z}_3\right) + \bar{K}_3 J_{-\frac{1}{3}}\left(\bar{z}_3\right)\right]$$

$$\text{for } -1 <= x_1 <= +1 \text{ and } -b <= x_3 <= +b$$

in the unelectroded regions :

$$u^{(1)} = A^{(1)}\left(-\alpha_n x_1 - m^2\right)^{1/2}\left(\beta_n x_3 + \bar{p}^2\right)^{1/2}$$

$$\left[I_{+\frac{1}{3}}(z_1) - I_{-\frac{1}{3}}(z_1)\right]\left[J_{+\frac{1}{3}}\left(\bar{z}_3\right) + \bar{K}_3 J_{-\frac{1}{3}}\left(\bar{z}_3\right)\right]$$

$$\text{for } x_1 <= -1 \text{ and } -b <= x_3 <= +b$$

$$u^{(1')} = A^{(1')}\left(-\alpha_n x_1 - m^2\right)^{1/2}\left(\beta_n x_3 + \bar{p}^2\right)^{1/2}$$

$$\left[I_{+\frac{1}{3}}(z_1) - I_{-\frac{1}{3}}(z_1)\right]\left[J_{+\frac{1}{3}}\left(\bar{z}_3\right) + \bar{K}_3 J_{-\frac{1}{3}}\left(\bar{z}_3\right)\right]$$

$$\text{for } x_1 >= +1 \text{ and } -b <= x_3 <= +b \qquad (18)$$

$$u^{(3)} = A^{(3)}\left(\alpha_n x_1 + \bar{m}^2\right)^{1/2}\left(-\beta_n x_3 - p^2\right)^{1/2}$$

$$\left[J_{+\frac{1}{3}}\left(\bar{z}_1\right) + \bar{K}_1 J_{-\frac{1}{3}}\left(\bar{z}_1\right)\right]\left[I_{+\frac{1}{3}}(z_3) + I_{-\frac{1}{3}}(z_3)\right]$$

$$\text{for } x_3 <= -b \text{ and } -1 <= x_1 <= +1$$

$$u^{(3')} = A^{(3')}\left(\alpha_n x_1 + \bar{m}^2\right)^{1/2}\left(-\beta_n x_3 - p^2\right)^{1/2}$$

$$\left[J_{+\frac{1}{3}}\left(\bar{z}_1\right) + \bar{K}_1 J_{-\frac{1}{3}}\left(\bar{z}_1\right)\right]\left[I_{+\frac{1}{3}}(z_3) - I_{-\frac{1}{3}}(z_3)\right]$$

$$\text{for } x_3 >= +b \text{ and } -1 <= x_1 <= +1$$

where

$$\bar{z}_1 = \frac{2}{3\alpha_n}\left(\alpha_n x_1 + \bar{m}^2\right)^{3/2}$$

$$\bar{z}_3 = \frac{2}{3\beta_n}\left(\beta_n x_3 + \bar{p}^2\right)^{3/2}$$

$$z_1 = \frac{2}{3\alpha_n}\left(-\alpha_n x_1 - m^2\right)^{3/2} \qquad (19)$$

$$z_3 = \frac{2}{3\beta_n}\left(-\beta_n x_3 - p^2\right)^{3/2}$$

These above relations are valid if the quantities $\bar{z}_i$ or z_i are real . If this is not true we must switch from the solution (13) to the solution (14) . This mean , by example , that the vibration stop to decrease uniformly after the point of abscissa $x_1 = -m^2/\alpha_n$ and becomes again a quasi periodic function . In some sense , because of the increase of thickness in the $+x_1$ (or $+x_3$) direction , the vibration is de-trapped far outside the electroded region .

The continuity of the mechanical displacement and stress at the interfaces between the electroded and the unelectroded regions gives

$$\bar{u}(-1) = u^{(1)}(-1) \qquad \bar{u}_{,1}(-1) = u^{(1)}_{,1}(-1)$$

$$\bar{u}(+1) = u^{(1')}(+1) \qquad \bar{u}_{,1}(+1) = u^{(1')}_{,1}(+1) \qquad (20)$$

and

$$\bar{u}(-b) = u^{(3)}(-b) \qquad \bar{u}_{,3}(-b) = u^{(3)}_{,3}(-b)$$

$$\bar{u}(+b) = u^{(3')}(+b) \qquad \bar{u}_{,3}(+b) = u^{(3')}_{,3}(+b) \qquad (21)$$

By inserting the solution (18) in (20) , we obtain a set of four linear homogeneous equations in $\bar{A}, \bar{K}_1$, $A^{(1)}$ and $A^{(1')}$ In order to obtain a non-zero solution , the determinant M associated to (20) must be equal to zero. Since $\bar{m}^2$ and m^2 are related by the first two relations (17) , the values of $\bar{m}$ are solutions of

$$\det M\left(\bar{m}^2\right) = 0 \qquad (22)$$

The amplitude ratios $\bar{K}_1$, $A^{(1)}:\bar{A}$ and $A^{(1')}:\bar{A}$ are then obtained from (20) for each value of $\bar{m}$. In the same way, $\bar{p}$ and the ratios $\bar{K}_3$, $A^{(3)}:\bar{A}$ and $A^{(3')}:\bar{A}$ are obtained from (21) . Each set ($\bar{m}$, $\bar{p}$) corresponds to an eigenmode of the resonator , the frequency of which is given by the first relation in (17) .

From a numerical point of view , it is necessary to use asymptotic expansions of Bessel functions when the argument z is large . In particular , it is necessary to use the asymptotic expansion of the difference $I_{+1/3}(z)-I_{-1/3}(z)$ itself since modified Bessel functions I_{+-n} tends to the same

limit for large z [4].

4 EXAMPLE

Although our model is valid for a cut of any orientation, we present here results in the case of an AT-Cut, which is well known. The here considered resonator is a 40 Mhz, third overtone, with the following caracteristics : thickness $2h_0$ = 0.125mm, size of electrodes 2lx2b= 4x4mm, mass loading $R = 2\rho' h'/\rho_0 h_0 = 0.002$. In order to simplify the presentation of the results, we only considere a defect of parallelism in the x_1-direction ($\gamma_3 = 0$). Fig3 shows the change of the in-plane distribution $u(x_1,0)$ of the main mode when the defect of parallelism increase from 0 to $\gamma_1 = 2$ (0.125mrd). Dashed lines indicates the edges of the electrodes. The maximum of vibration is shifted to the side of the resonator where the thickness is the biggest. In the case of the highest value of γ_1, the vibration is de-trapped not far outside the electroded region.

Fig4 is drawn in the case of the first anharmonic (or spurious) mode. When the faces of the plate are perfectly parallel, the mode is antisymmetric. But when a defect of parallelism occurs, the mode becomes distorded and no longer remain purely antisymmetric. In this last case, the mode can be electrically driven, then exhibiting a finite motional resistance. The table I summarizes the motional resistance and inductance for all the modes of the resonator operating in third overtone. Theses quantities are obtained directly from the distribution of the amplitude of vibration in the resonator given in (18) by putting it in the following expressions :

$$R_1 = \frac{n^2 \pi^2 h_o \eta_{66}}{4 e_{26}^2} . F \qquad (23)$$

$$L_1 = \frac{n^2 \pi^2 h_o c_{66}}{4 e_{26}^2 \omega^2} . F \qquad (24)$$

with

$$F = \iint_{S_R} \left[u(x_1, x_3)\right]^2 ds \Big/ \left[\iint_{S_E} u(x_1, x_3)\, ds\right]^2 \qquad (25)$$

where S_R and S_E denotes the whole surface of the resonator and the surface limited by the electrodes. In table I, instead of the absolute value of R_1, which depend of the used value of η_{66}, we have reported the relative value R_1/R_0, where R_0 is the value of the motional resistance of the main mode in the case $\gamma_1 = 0$ [6]

As it may be observed the resistance and inductance of the main mode (n00) increase as the defect of parallelism increase, whereas the same quantities decrease at same time in the case of anharmonic modes (nλ0). The decrease of the frequency of the n00-mode is due to the diminution of $\bar{m}^2$ which can become negative.

REFERENCES

1. L.N.Dworsky _ Properties of AT Quartz Resonators on Wedgy Plates, Proc. 37th An. Symp. on Freq. Cont. ,p232-238, (1983)

2. D.S.Stevens, H.F.Tiersten _ An Analysis of doubly rotated quartz resonators utilizing essentially thickness modes with transverse variation, J. Acoust. Soc. Am. 79 (6) p1811-1826 (1986)

3. B.A.Auld _ Acoustic Fields and Waves in Solids, John Wiley & Sons, New York (1973)

4. E.T.Whittaker & G.N.Watson _ A Course of Modern Analysis, Cambridge University Press, Cambridge (1965)

5. For conciseness the expressions of the mechanical displacement in the corner regions are omitted. They can be easily deduced from the solutions of the other regions in the same way like for the usual trapped energy resonator.

6. If the value η_{66}=3.4 10^{-4} is used (obtained from the data published by Lamb & Richter) then R0 = 9.7 Ohm.

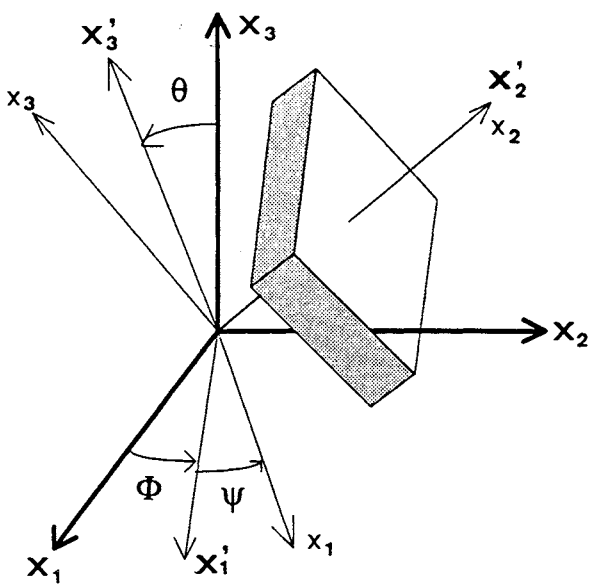


Fig 1 Doubly rotated cut

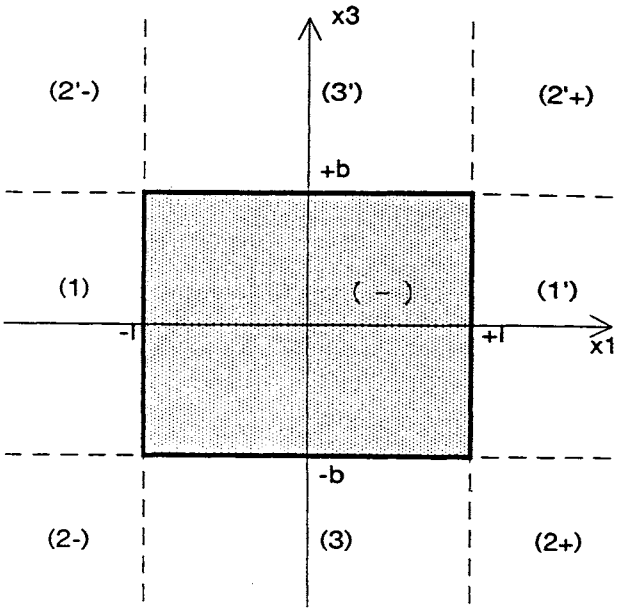


Fig 2 Trapped energy resonator

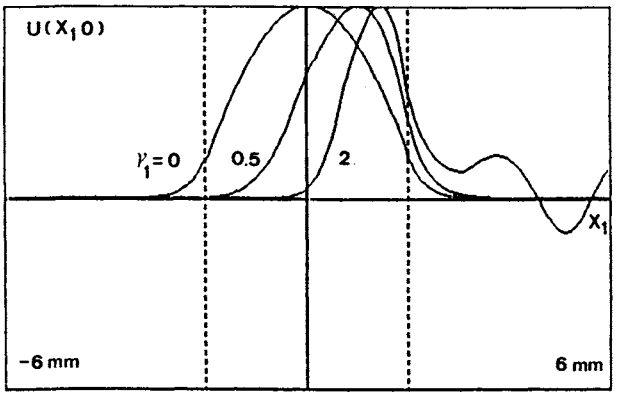


Fig 3 Main mode

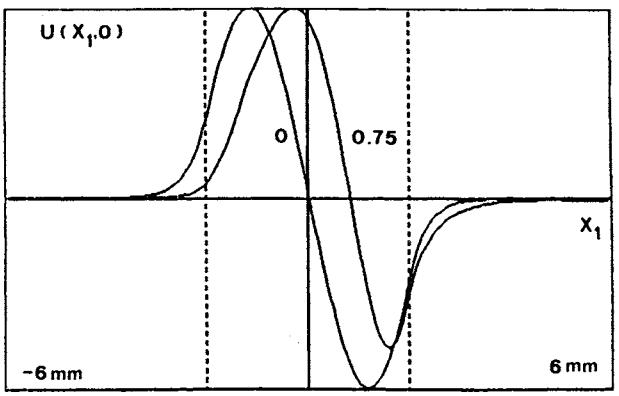


Fig 4 First anharmonic mode

gamma1 = 0				gamma1 = 0.25				gamma1= 0.75			
f(Mhz)	npq	R_1/R_0	L(H)	f(Mhz)	npq	R_1/R_0	L(H)	f(Mhz)	npq	R_1/R_0	L(H)
39.7922	300	1	0.009	39.7909	300	1.14	0.010	39.7838	300	1.41	0.013
39.8223	302	7.9	0.072	39.8210	302	9.0	0.082	39.8139	300	11.1	0.101
39.8756	304	63	0.57	39.8743	304	71	0.647	39.8672	304	88	0.798
39.8046	310			39.8050	310	15.4	0.010	39.8056	310	6.6	0.060
39.8347	312	∞		39.8351	312	122	1.11	39.8366	312	52	0.047
39.8880	314			39.8884	314	968	8.78	39.8889	314	413	3.75
39.8249	320	13.8	0.126	39.8255	320	11.0	0.099	39.8264	320	9.4	0.085
39.8549	322	110	0.994	39.8556	322	87	0.788	39.8565	322	74	5.31
39.9082	324	865	7.84	39.9088	324	687	6.22	39.9097	324	586	5.31
39.8517	330			39.8527	330	375	3.40	-		-	-
39.9350	332	∞		39.9360	332	2967	26.9	-		-	-
39.9350	334			39.9360	334	23500	212	-		-	-

Table I Frequencies and motional elements for various defects of parallelism

TIME AND FREQUENCY REFERENCE FOR A TRAJECTORY MEASUREMENT SYSTEM

G. Boime

TEKELEC SYSTEMES
29, avenue de la Baltique
Z.A. de Courtaboeuf
91953 Les Ulis cedex
FRANCE

ABSTRACT

There is a growing need in modern measurement systems for precise timing and dating of events. In particular, there is a requirement of precise time and frequency referencing in trajectory measurement which allows tracking of long-range missiles. Time and frequency are provided to different systems such as radar, optical sighting and telemetry. The time and frequency system is made up from a main clock reference and secondary clocks synchronized via a local area network. This redundant architecture improves the reliability and allows time and frequency distribution even when the main clock is unavailable.

Keywords : time, synchronization, frequency distribution, LAN, redundancy, reliability.

1 - INTRODUCTION

As part of the renewal of the test support equipment of the DGA (French Ministry of Defence) has decided to build a new ship for measurement and test.

This ship is designed to track long and medium range missiles on test programmes. She can compute the precise trajectory of flying objects, and will cover the requirements well into the 20th century.

The time and frequency reference described in this paper is installed on the ship for dating and synchronization of all the different measurement instruments.

2 - GLOBAL REQUIREMENT OF TRAJECTOGRAPHIC SYSTEM.

During the firing and flight of missiles, the ship is located near the target area. Launch and ascent are tracked from the ground firing base and ground support bases spread along the missile trajectory. The ship tracks with radars, optical and telemetry equipment the descent part of the flight trial. (see figure 1).

We identify three levels in the time and frequency distribution system on board :

- The global entity of tracking and measurement equipment is synchronized from ground bases to the ship.

- Sensors on board the ship synchronize each other.

- Each sensor dates the real time recording and synchronizes its various measurement channels.

Regarding the high costs of such an experiment, it is absolutely necessary to insure a very low failure rate of the system during the whole offshore mission.

In addition, it is necessary to secure failure rate on recording systems as low as possible during the critical phase of the flight test.

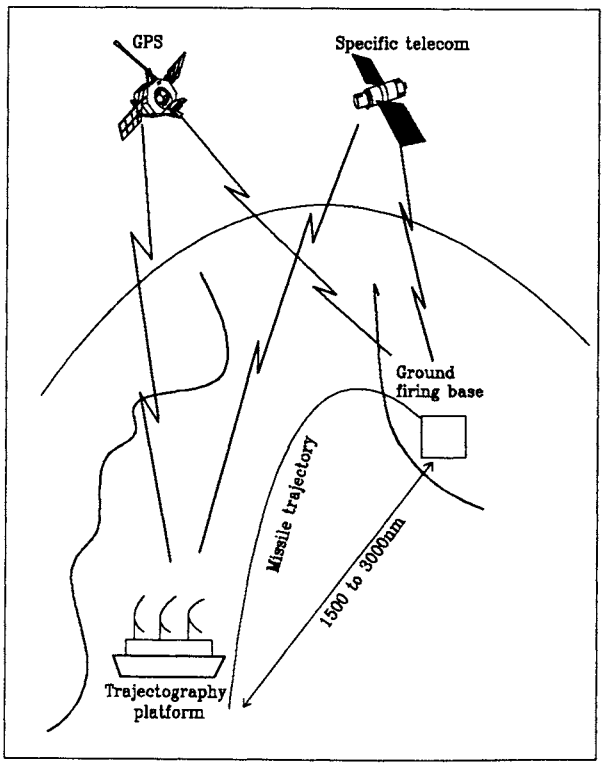


Figure 1. Trajectographic system synchronization.

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

3 - TIME AND FREQUENCY REFERENCE.

3.1 Synchronization of tracking platforms.

For correlation of the whole telemetry and trajectography measurements, each base is synchronized to a common reference source. During operating conditions, the UTC/ USNO is the reference for timing and synchronization. This time is broadcast by the GPS constellation.

On board, a GPS receiver acquires the UTC reference with high precision (about one hundred nanosecond accuracy).

During firing and flight trials, the different actions must be synchronized to ensure the linkage of the measurement sequence on the recording systems.

To meet this goal, the ground firing base initializes a count-sequence. This count-sequence is sent via a satellite telecommunications link. This link uses a secure channel of the Syracuse system or a common voice coding chanel of Inmarsat telecommunications vehicules. A one second accuracy suffices to set up the measurement sequence chain.

3.2 Synchronization on the ship platform.

An autonomous time and frequency reference system is installed on the tracking ship.

A central clock generates the time and frequency primary reference. A dedicated local area network distributes time and synchronization on secondary clocks throughout the platform.

3.2.1 Central clock.

The central clock is based on two rubidium atomic secondary references. A time generator defines time from the frequency references.

Two atomic sources are used for redundancy. The time generator automaticaly chooses the available source and switches in case of failure of one of them.

A time reference option is available on rubidium frequency standards, delivering a one pulse per second reference. These time references can be synchronized ashore to a common source before the beginning of the offshore mission. Regarding the stability of the frequency source, primary initialization is enough to maintain the time drift below 1ms during the average time of an offshore mission.

Through the GPS receiver of the navigation system, the UTC/USNO time reference is sent to the central clock command unit. The distance from the two 1pps rubidium references, the 1pps signal from GPS and the 1pps delivered from the central clock time generator is calculated. The time difference between the time generator and other sources may thus be controlled.

Compensation of delay is done with a 100ns accuracy.

A modular distribution unit can supply up to twenty outputs of synchronous 5 or 10MHz sine-wave signals.

To ensure the autonomy of the central clock, the time generator, the distribution unit and the two atomic references are supplied with backup batteries. They ensure the availability of the master clock during the critical sequence of missile trials in case of main power failure.

3.2.2 Time and Frequency delivery on the platform.

3.2.2.1 Architecture.

In accordance with the huge length of the ship, distribution of time is made to measurement equipment hundreds of meters apart. To ensure the backup of the measurement recording during the sequence of a missile trial, a distributed architecture has been chosen (see figure 2). On each main measurement area, a secondary clock is installed.

Secondary clocks have their own frequency reference based on an ovenized cristal oscillator (OCXO). They are synchronized and tuned to the central clock reference through a dedicated Local Area Network. In case of unavailability of the central clock during the critical sequence of trials, secondary clocks deliver a synchronized time enough precise for the required measurement reliability. The critical sequence duration is about half an hour long and the synchronization between sensors must be kept better than 1ms during this period.

3.2.2.2 Local Area Network.

Each secondary clock, synchronizes itself to the central clock through the time specific local area network.

This network distributes time reference (date generated within the central clock), count information and synchronization. Distribution includes three RS 485 lines.

- One line of data from central clock to secondary clocks terminals.
- One line of data from secondary clocks to central clock.
- One line of synchronization with 1pps.

The central clock sends dating to the terminal equipment each second. It remotely controls the availability of the terminals. It asks selectively the secondary clocks and receives an acknowlegment and wellfare message from each one. Ruggedized protocols of the transport layer protects the distribution of time against corruption such as glitches and transmission errors.

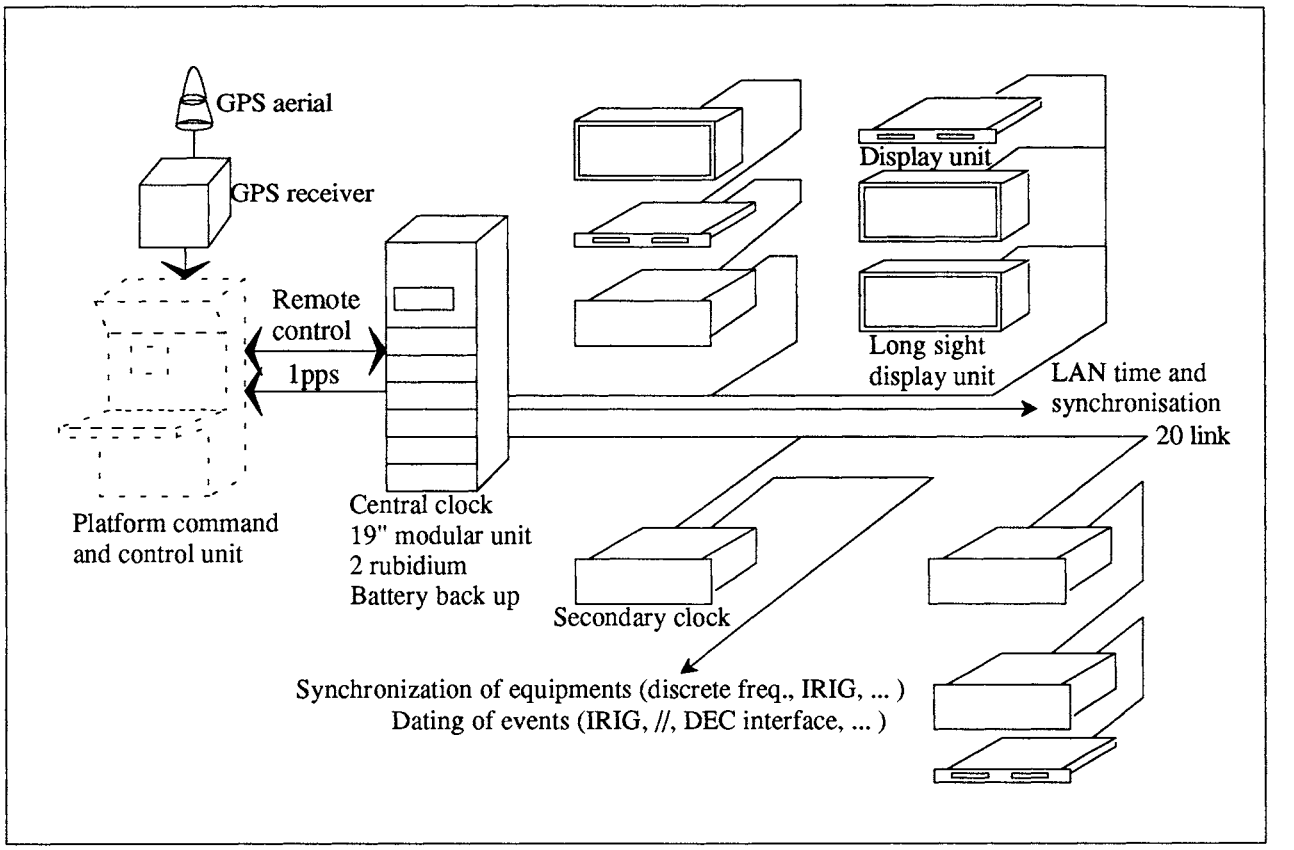


Figure 2. Time and frequency distribution system on the ship platform.

Via these links, data can be sent up to a thousand meters. The central clock distributes on twenty separate links similar information.

3.2.3 Synchronization features.

The synchronization accuracy measured between secondary clocks is better than 20µs. In the case of central reference unavailability, ie: broken link or central clock disabled, the drift measured on secondary clocks after tuning is much better than the required 1ms for half an hour.

Specified figure on the OCXO frequency reference of secondary clocks are :

- Short term frequency ageing : 3.10^{-9} / day.
- Thermal shift from -30°C to +70°C : 1.10^{-7}.
- Electrical tuning of frequency, accuracy about 8.10^{-10}.

These figures meet the free-running stability requirement.

The drift actually noted on a sample is indicated on figure 3. We can check that the drift in normal operating conditions is lower than 250µs for a 12 hour period.

3.3 Synchronization and dating measurement equipment.

Each one of the secondary clocks keeps its own frequency and timing reference. When the central clock is available, the secondary clock automaticaly synchronizes to it. When this is not the case, the secondary clock runs on its own frequency reference and keeps to its own time clock.

The secondary clock delivers to the measurement system multiple frequencies synchronized to the central clock system.

The time broadcast is performed with a modular unit delivering many standard interfaces :

- Serial codes : PDM, IRIG B, ...
- Parallel codes : BCD, binary, DEC standard protocols...

The synchronization of all the frequencies and codes delivered is done on the local time generator with better than 500ns accuracy.

3.4 Time and count display.

In the goal of unification of the time and count acquisition throughout the platform, display units are part of the time reference system.

Two types of units are used :

- Rackable one unit high 19" standard display for insertion in measurement equipment bay,

- Wall-mounted display panel visible from up to 15m for command and control rooms.

Both types can be connected to the time distribution local area network or through a secondary clock. The time displayed is the time generated by the secondary clock. This time is the same as central clock time in normal operating conditions. It can be independant during a training test or when the distribution is interrupted.

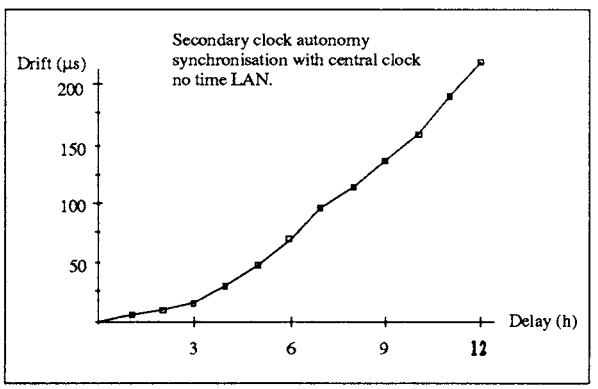


Figure 3. Secondary clock stability. Drift on a secondary clock from the rubidium reference when the oscillator has been synchronized and tuned to the primary frequency reference before time 0.

4. - FIGURES.

When the time and frequency system is fully available and locked to the GPS reference, the synchronization drift between the UTC/USNO and the synchronization references delivered to the measurement systems is less than 15µs.

When the central clock of the system is free-running, if time references (1pps) of the rubidium sources have been synchronized and tuned before the begining of the offshore mission, the drift will stay below than 500µs for the whole platform.

The distributed architecture allows the reduction of the unavailability rate during the critical trial phase. In this case, the synchronization drift within the platform is lower than 100µs from one distribution output to another.

The availability rate during the offshore mission will be better than 99%. The mean time between failures will be better than 10 000 hours.

5. - CONCLUSION.

The use of a centrallized primary clock with secondary clock terminals including low precision references has proved its worth for distribution of time and synchronization of up-to-date measurement systems spread throughout a large platform.

The reliability of the time and frequency system allows the highest availability of the telemetry and trajectography measurement ship.

6. - ACKNOWLEDGMENT.

The author wishes to acknowledge the support of M. DUNAUD I.A., Central exploitation system Manager in the "Centre Technique des Moyens d'Essais", part of the ETCA, French Ministry of Defence which funded the development of the time and frequency system.

BASIC DESIGN OF A SPACEBORNE HYDROGEN MASER

L.G. Bernier

Observatoire de Neuchâtel
Beauregard 3, 2036 Cormondrèche, Switzerland

ABSTRACT

Neuchâtel Observatory is planning the construction of a space qualified atomic hydrogen maser to be flown on EURECA III in a technology demonstration experiment. The paper discusses the basic issues in the design of the physics package at the feasibility level.

1. INTRODUCTION

Neuchâtel Observatory is planning the construction of a space qualified hydrogen maser in preparation of a technology demonstration experiment to be flown on the EURECA III retrievable spacecraft in 1998. The demonstration flight will be a NASA/ESA joint experiment involving one maser built by ON, one maser built by SAO, a microwave Time & Frequency Transfer system based on PRARE and a laser Time Transfer system (Ref. 1,2,3).

This paper is based on a feasibility study performed for ESTEC under contract N° 9099/90 /NL/US(SC), Work Order 03 (Ref. 4). The basic trade-off issues in the design of the physics package are discussed at the feasibility level. The main topics are the choice of the microwave resonator, the choice of the basic operational parameters, the hydrogen storage, distribution and pumping, the magnetic shielding and the storage bulb coating.

2. ENVIRONMENTAL CONDITIONS

2.1 Thermal

On the EURECA platform, the "operational mission phase" sink temperature under the insulation blanket (Ref. 5) is specified to be in the range [0 C°, 40 C°]. As a matter of fact the actual temperature is expected to be stable within a few C°. It is the mean temperature that cannot be predicted since it depends upon the specific power dissipation of spaceborne experiments that have not been selected yet. Nevertheless the maser must be designed in order to be operable at 40 C° if passive cooling is to be used. This is problematic since the regulated temperature of the microwave cavity is normally 45 C°. Under these conditions the use of a connection to the EURECA Freon cooling loop will be necessary. The Freon loop inlet temperature is specified in the range [14 C°, 25 C°].

2.2 Magnetic

The EURECA Instrument Interface Proposal (Ref. 5) specifies that the operation of the magnetic torquers used for attitude and orbit control during the operational phase of the EURECA flight will continuously produce magnetic pulses with field strengths of 10 Oersted close to the torquer while the strength is generally smaller than 3 Oe at the baseplate of the instruments.

The above specification assumes that the instrument can be in any position on the platform. However, according to document ERC-LET-ER-666-86 that shows the exact position of the magnetic torquers on the EURECA platform, sketched on figure 1, it is possible to place the spaceborne maser at about 2 m from the magnetic torquer N°1. This is the maximum distance that can be achieved given the platform configuration. At 2 m of the closest magnetic torquer, the magnetic field pulses have a 100 mOe intensity along the worst case axis.

This is to be compared with the 1 mOe magnetic field variations that ground hydrogen masers encounter under normal operating conditions.

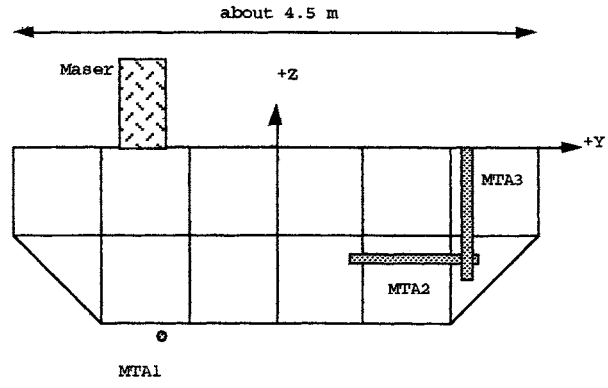


Figure 1
Location of Magnetic Torquers

2.3 Acceleration and Vibration

The vibration test spectrum specified at the level of the experiment support panel has an rms value of 9 g in plane and 13.8 g out of plane. In first approximation a simple rule of thumb for the design of the mechanical structure is to account for 20 g static acceleration in all directions.

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

3. MASER DESIGN REQUIREMENTS

The design requirements that are particular to a spaceborne maser are the following.

The maser must survive the acceleration and shocks of launch. In particular the microwave cavity must be designed in such a way that no detuning occurs.

In the perspective of the EURECA technology demonstration program, the storage bulb coating, the vacuum system and the hydrogen supply must be designed for a 10 years autonomy, i.e. requirements for full scale space applications of the maser, even though the maser will actually run only 6 to 9 months in space on the EURECA flight.

The maser must have state-of-the-art performance despite the weight, power consumption and environmental restrictions imposed by space environment.

4. FIRST ITERATION DESIGN

4.1 Overview

A sketch of our first iteration design is shown on figure 2. The basic features are as follows. We plan to use a TE_{011} cavity and a full size storage bulb similar to the EFOS design (Ref. 6). This is necessary for obtaining state-of-the-art frequency stability. The hydrogen vacuum vessel will be separated from the thermal vacuum vessel. This will allow the use of the ambient space vacuum for thermal insulation. Getter pumping of hydrogen will be used in the hydrogen vacuum vessel in combination with a small ion pump for the residual molecules other than hydrogen. A solid state hydrogen supply will be used at the input of the palladium pressure regulation control loop. Hydrogen dissociation will be done conventionally through a RF plasma discharge. An auxiliary mode automatic cavity tuning system will be used (Ref. 1, 7) for removing the effect of the drift of the cavity on the long term stability. Five magnetic shields will be used. The quadrupole state selector will be similar to the EFOS design. Titanium will be largely used for the vacuum vessel and other mechanical structures.

4.2 Mechanical Structure

The basic structure of a hydrogen maser resumes to a massive microwave cavity suspended within many layers of thin magnetic shields that must be well insulated thermally from the outside shell.

It is clear that the most difficult part of the mechanical design will be to suspend the microwave cavity in a structure that can withstand 20 g static acceleration without requiring large holes in the magnetic shields and without creating a strong thermal path to the base-plate.

4.3 Mass and Power Consumption

The overall physics package is expected to be 700 mm high for a diameter of 400 mm. The goal for the mass of the physics package alone is 60 kg. Assuming 4 magnetic shields of 0.8 mm as in the EFOS design, the total mass of the magnetic shields is 38 kg. Therefore the mass budget is governed mainly by the size of the microwave cavity and by the number and thickness of the magnetic shields.

The power consumption goals are 45 W for the physics package and 30 W for the electronics.

5. MICROWAVE CAVITY

The goal being to obtain state-of-the-art frequency stability, the cavity must be of the TE_{011} circular type. When the different types of cavities used in hydrogen masers are reviewed (Ref. 4) it appears that truly compact cavities have too low a quality factor for active operation of the maser and, on the other hand, impose a small storage bulb i.e. a reduced quality factor of the atomic resonance. Cavities partially loaded with a dielectric do not yield a gain in size that is worth the loss of quality factor.

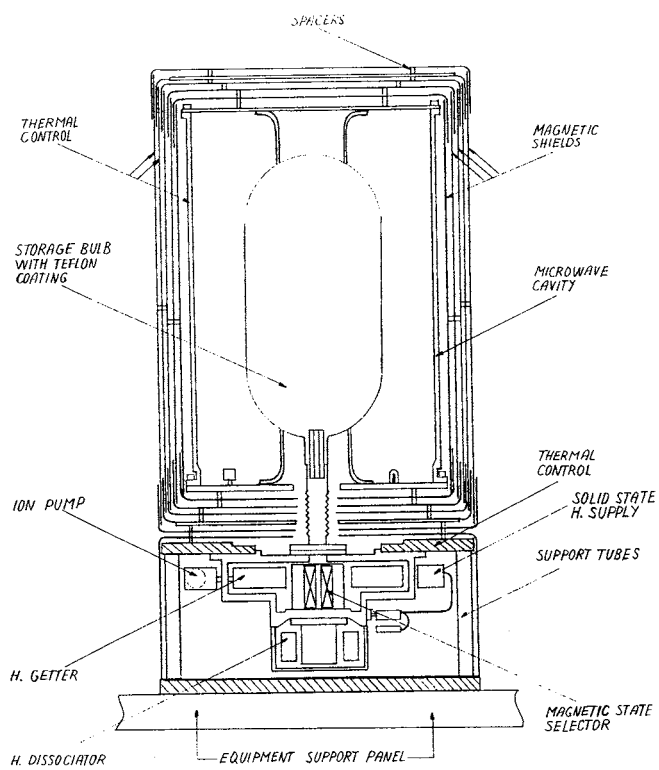


Figure 2
First Sketch of Physics Package

Then the important design decision concerns the choice of the material used for producing the cavity. A low thermal expansion glass-ceramic TE_{011} cavity has a thermal coefficient dominated by the quartz storage bulb. A coefficient of -300 Hz/C° is achievable using a 150 g storage bulb (Ref 8). On the other hand an aluminium cavity similar to the EFOS design cavity has a -30kHz/C° coefficient determined by the thermal expansion coefficient of aluminium.

Our experience with the EFOS masers shows that the drift of the maser frequency due to ageing of the thermal control bridge is less than $1x10^{-14}$/day. In the case of a low thermal expansion cavity the drift is of the same order of magnitude and is produced by mechanical relaxation (Ref. 9, 10). If an automatic cavity tuning system is used the cavity drift is removed anyway. The relevant choice is a compromise between thermal and mechanical considerations.

The glass-ceramic cavity has a low temperature coefficient and high rigidity but it is difficult to secure the covers to the cylindrical body in order to prevent detuning under high vibrations. The cavity can be autotuned only by means of a varactor. On the other hand the metallic cavity can be autotuned by means of small temperature changes. The covers are easily secured to the cylindrical body and it is also easy to maintain a low temperature gradient.

6. Automatic Cavity Tuning

The automatic cavity tuning system is based on the auxiliary mode principle (Ref. 6, 7). By using the TE_{013} mode at 1800 MHz a -70 dBm interrogation power can be used without perturbing the maser signal frequency at the $1x10^{-15}$ level. Our preferred interrogation scheme (Ref. 1) is to use a cavity oscillator locked to the auxiliary mode. The pulling factor between cavity and maser frequencies is about $1x10^{-5}$. If the goal is to stabilize the maser frequency to the $1x10^{-16}$ level, the flicker of frequency of the cavity oscillator must be better than $1x10^{-11}$ at 1 Hz. Assuming a quality factor of 30'000 for the auxiliary mode, the flicker of phase of the electronics must be better than -125 dB ref 1 rad^2/Hz at 1 Hz. This requirement makes critical the circuit design and the selection of the semiconductor parts. On the other hand the cavity oscillator scheme is much simpler than the modulation & detection scheme in terms of complexity of the electronics.

6. MAGNETIC SHIELDING

As mentioned above, the magnetic shields account for a large fraction of the mass budget. Given the density of the shielding material, the only way to spare mass is to reduce either the number of shields or their thickness.

The Russian CH1-75 maser has 5 magnetic shields with a thickness of 0.5 mm each. A dynamic axial shielding factor of 50'000 can be achieved (Ref. 10). This was verified for a ±200 mOe perturbation on a CH1-75 maser now on loan at Neuchâtel Observatory. Assuming a C-field of 400 µOe the resulting frequency coefficient is $3x10^{-14}$/Oe. This result is representative of what can be achieved with 5 magnetic shields.

Assuming in the worst case that 100 mOe pulses parallel to the maser axis are produced by the closest magnetic torquer, the frequency perturbations will have a $3x10^{-15}$ intensity under the assumption of a $3x10^{-14}$/Oe axial coefficient.

Supplementary countermeasures are thus necessary. Using thicker shields would increase the mass considerably. On the other hand it is possible to add an active compensation system (Ref. 10) or to tune the cavity so that the change of pulling effect due to the linewidth modulation compensates for the frequency shift due to the pure shielding effect.

7. HYDROGEN GETTER PUMPING

The SAO 1976 red shift experiment used ST171 getters manufactured by SAES in Italy (Ref. 11). The ST171 getter is based on a ZrC alloy and has a capacity of 2400 mBar.litre for 171 g. 4 of these getters weight 684 g and are expected to yield 7 years of operation at $4.3x10^{-5}$ mBar.litre/s which is our estimation of the hydrogen consumption.

The Russian maser CH1-75 uses 1 kg of Titanium powder as a getter material and claims 5 years of continuous operation (Ref. 10).

In any case a small ion pump is necessary in order to pump residual gases other than hydrogen.

8. HYDROGEN STORAGE

The SAO 1976 red shift experiment used $LiAlH_4$ for the solid-state storage of hydrogen (Ref. 11). When heated, 38 g of this material produce 2 mole of H_2 which represent 30 years of operation at $4.3x10^{-5}$ mBar.litre/s. Russian masers, on the other hand, use $LaNi_5H_x$ (Ref. 10). 250 g of this material yield $18x10^3$

mBar.litre which represent 10 year of operation at 4.3×10^{-5} mBar.litre/s.

9. STORAGE BULB COATING

9.1 Effect of Cosmic Radiation

The long-term effects of cosmic radiation on the storage bulb coating is not known. Most of the literature about the effect of cosmic radiations on Teflon concerns the degradation of mechanical properties of solid PTFE which is used in the production of mechanical parts. One exception is the report by J.J. Sutter, from the Applied Physics Laboratory at John Hopkins University, who observed a slight improvement of the atomic line quality factor after a storage bulb was exposed to 10 krads of 1.25 MeV photons from a gamma ray source.

The most likely damage caused by cosmic radiations is the breaking of the Teflon chains at the weakest point, i.e. the carbon bonds, producing shorter chains with an increased number of end groups. New experimental work is required in order to determine the expected lifetime of the coating in the space environment.

9.2 Estimation of Quality factor

We define the contribution of the coating to the operating line Q factor

$$\frac{1}{Q_{operating}} = \frac{1}{Q_{se}} + \frac{1}{Q_{bulb}} + \frac{1}{Q_{coating}}$$

by considering only the last term of the above equation which separates the contributions of the spin exchange, the bulb time constant and the wall relaxation. A contribution of 1×10^{10} for the coating seems feasible with the present technology. We have obtained 7×10^{9} in recent EFOS masers using FEP120 Teflon. We estimate that the Russian F10 Fluoroplast would yield 9×10^{9} in a bulb of the same size as in the EFOS design. An operating line Q between 3×10^{9} and 4×10^{9} is thus feasible.

10. CONCLUSION

The foreseeable critical issues in the realization of the physics package of a space qualified hydrogen maser are the mechanical, thermal and magnetic designs. The mechanical structure for holding the cavity and storage bulb under the extreme conditions of the space qualification vibration tests and the shielding of the maser against the very high magnetic perturbations of the EURECA environment are especially challenging.

REFERENCES

1. Busca G., Bernier L.G., Starker S., Feltham S.,"In Orbit Demonstration of a H-Maser Clock System," To be published in *Proc. 23rd PTTI Application and Planning Meeting*, Pasadena, December 1991.

2. Vessot R.F.C., Mattison E.M., Decher R., "Test of an Orbiting Hydrogen Maser Clock System Using Laser Time Transfer," To be published in *Proc. 23rd PTTI Application and Planning Meeting*, Pasadena, December 1991.

3. Vessot R.F.C., Mattison E.M., Nystrom G.U., Coyle L.M., Decher R., Feltham S.J., Busca G., Starker S., Leschiutta S., "Hydrogen Masers for Space," To be published in Proc. 6th EFTF, Noordwijk, March 1992.

4. *Space Maser Study*, Work Package N°1, Phase 1 Report, ESTEC Contract N° 9099/90/NL/US (SC), Work Order 03.

5. *EURECA Instrument Interface Proposal (IIP)*, ESA-EURECA DOC. No. ERC-PL-121-0.

6. Bernier L.G., Busca G., "Recent Progress in the Development of the EFOS-B Hydrogen Maser at Neuchâtel Observatory," *Proc. 5th EFTF*, Besançon, March 1991, pp.404-406.

7. Busca G., Johnson L., "A New Automatic Cavity Tuning for Active H Masers," *Proc. 1st European Time and frequency Forum*, Besançon, March 1987, pp.339-341.

8. Vessot R.F.C., Levine M.W., Mattison E.M., *A Report on the Evaluation of the Performance of the Smithsonian Institution Astrophysical Observatory VLG-11 Atomic Hydrogen Masers*, Final Report Contract JPL-954818, Grant NSG-8052.

9. Mattison E.M., "Physics of Systematic Frequency variations in Hydrogen Masers," *Proc. 22nd Annual PTTI Applications and Planning Meeting*, Vienna, Virginia, December 1990, pp. 453-464.

10. Demidov N.A., Uljanov A.A., "Design and Industrial Production of Frequency Standards in the USSR," *Proc. 22nd Annual PTTI Applications and Planning Meeting*, Vienna, Virginia, December 1990, pp. 187-208.

11. Vessot R.F.C., Levine M.W., Mattison E.M., Hoffman T.E., Imbier E.A., Têtu M., Nystrom G., Kelt J.J., Trucks H.F., Vaniman J.L., "Space-Borne Hydrogen Maser Design," *Proc. 8th Annual PTTI Applications and Planning Meeting*, Washington, December 1976, pp. 277-233.

12. Morris D., Report on Special Hydrogen Maser Workshop, *Proc. 22nd PTTI Applications and Planning Meeting*, Vienna, Virginia, December 1990, pp. 349-354.

CRYSTAL GROWTH OF BERLINITE: METHODS FOR GROWING LONGER CRYSTALS AND IMPROVING THE QUALITY.

X. BUISSON*, A. ZARKA**, R. ARNAUD*,B. CAPELLE**,Y. ZHENG**, J. SCHWARTZEL***J. DETAINT*** and Y. TOUDIC****

*S.I.C.N. 4 RUE DU RADAR. 74005. ANNECY.CEDEX. FRANCE

** L.M.C.P. 4 PLACE JUSSIEU. 75252. PARIS CEDEX. FRANCE

***C.N.E.T./PAB/ BAG. 132 AV. H. RAVERA. 92220. BAGNEUX. FRANCE.

****C.N.E.T./LAB/OCM. ROUTE DE TREGASTEL.2230. LANNION. FRANCE.

ABSTRACT

Natural berlinite of large size are not known. Therefore it was necessary to find an hydrothermal method for growing crystals of sufficently large sizes to use as piezoelectric materials. For that purpose studies of the optimal conditions of growth have been performed using extension of different techniques previously applied particularly to quartz crystals. To control the quality of the obtained crystals,X-ray topographic studies were done and typical features of the samples are presented.

Berlinite. Large Crystals. X-ray topography.

INTRODUCTION

The berlinite AlPO4 (1) is a material analogous to the quartz. Its structure is obtained by an alternated substitution of the silicon by phosphorus and aluminium atoms. Its piezoelectric properties are more important than those of quartz. Berlinite is also one of the rare materials with crystalline orientations which gives very few fluctuations of the resonance frequencies in relation to the temperature variations. The researches on the growth of this material and the realisation of piezoelectric devices have been initiated at the end of the seventies in different laboratories (C.N.E.T., Bell Laboratories, Allied Signal, G.E.C.). The great majority of the results concerning the devices using berlinite have been obtained in the begining of the eighties at the laboratoire de Bagneux (C.N.E.T., France)(2).The high quality of material obtained permits the construction of devices with very high surtensions.

PROPERTIES OF THE CRYSTALS

Presently, the most interesting are applications which give piezoelectric features better than those given by the quartz crystals. These characteristics are necessary in order to obtain filters with bandwidth twice greater than those of quartz or oscillators with important frequency bandshift with a high stability of the frequency. Another type of application is the possibility to obtain high frequency resonators by several techniques of chemical etching. Contrary to quartz, the method is very simple and permits a significantly lower price for these devices. The use of berlinite for acoustical waves devices and also in thermostated oscillators leads to important advantages which results in electromechanical coupling coefficients superior to those of quartz. The expected advantages have permitted the construction of resonators with relative shift between resonance and the antiresonance frequency which are twice those of quartz crystals. Numerous resonators with frequency between some MHz to more than 75 MHz have been made at the C.N.E.T. and other french industries. These resonators have, at the same frequency, lower dimensions than

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

those of quartz. Mathematical models have been done in such a way that their properties do not change when the resonators are made. These models have also demonstrated that different geometries could bring very important advantages (3).

GROWTH OF CRYSTALS

In the past ten years, many efforts were made to develop hydrothermal methods that can produce large berlinite crystals of high perfection needed for piezoelectric applications.

Growth cycles were first performed in H_3PO_4 acid at 240°C (4) followed by systematic investigation of the solubility of berlinite which demonstrated that solubility in sulfuric acid is much higher than in all other known solvants (5). It is well known that hydrothermal growth method is based on thermal variation of the solubility of a material in a solvant to obtain a good single crystal. In case of berlinite, sulfuric acid at high temperature (about 220°C) and under high pressure is used. Under such conditions, the solubility is retrograde (i.e. decreases with the temperature). The technique used in industry comprises vertical autoclaves in which a spatial vertical gradient is established (Fig. 1). In the upper part of the autoclave there is a basket with nutriant materials which are small synthetic crystals. These crystals were synthetized from commercial product in a preliminary growth. In the lower part (separated from the higher part by a diaphragm) single crystals seeds are placed. The autoclave is then filled with a known quantity of sulfuric acid solution of aluminium phosphate. The temperature gradient and the correspondant solubility variation creates convection current.

In the lower part, crystallization takes place on the seeds and the solution concentration decreases. Therefore, the solution moves upwards, dissolves nutriant and becomes more concentrated. Consequently, the solution get down in the crystallization zone where excess of concentration leads to a growth on the seeds. This is more favorable for the nucleation.

Presently the S.I.C.N. Company produces about 7500 slices/year of a good quality crystals which permits the realisation of prototypes or little series of piezoelectric devices. The aim is to increase the capacity of its autoclaves and to improve the quality of the crystals.

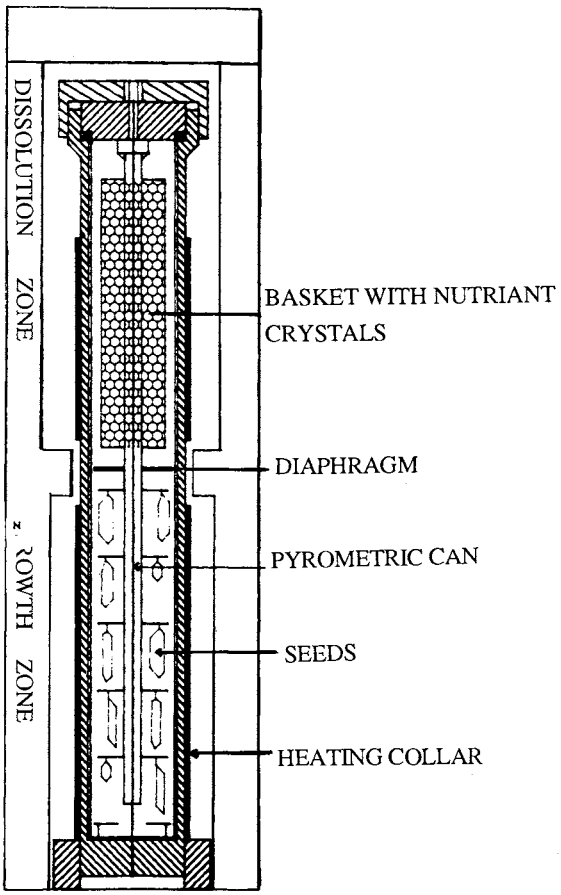


Fig. 1 Technique of growth.

METHODS FOR GROWING LARGER CRYSTALS

The observation of the crystal morphology (Fig.2), their size and the growth parameters have shown that the growth velocity along the Y direction is very low. Therefore, the studies of the optimal conditions to extend the size of the sample have been performed using two particular techniques:

- The first technique (6) is used to elongate the crystal. Its consists in joining together several plates of berlinite having the same orientation. Its remains always a slight

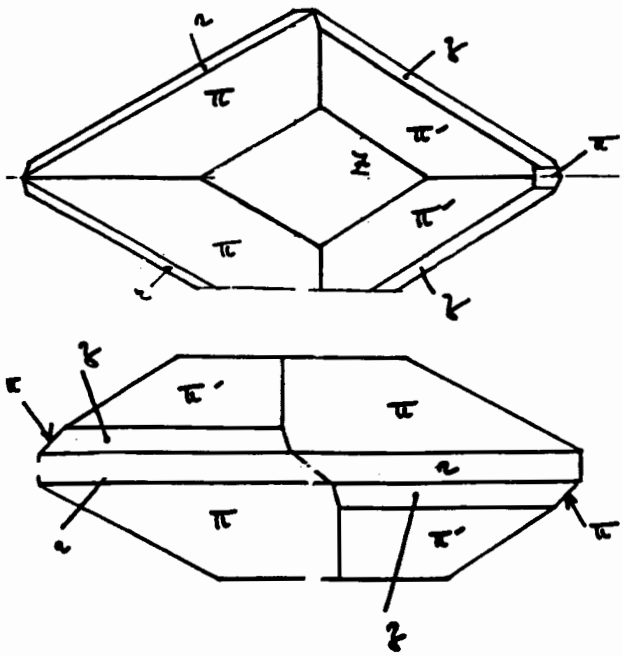


Fig. 2 Morphology of the crystals.

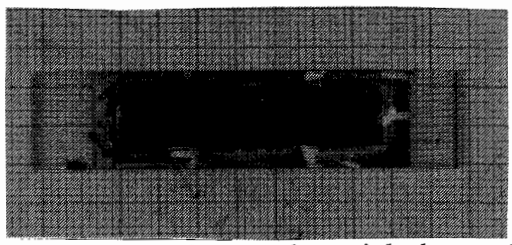


Fig. 3 Mounting used to stick the seed crystals.

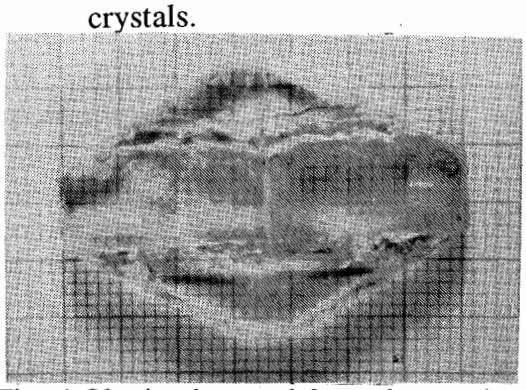


Fig. 4 Obtained crystal from the previous seed.

desorientation between the different parts of this seed and the as-grown crystal is not perfect. In figure 3 is presented a photograph of two crystals which were joined together to form a single seed with the mounting which was employed. In figure 4 is shown the obtained crystal and it may be observed that the quality is not perfect. To avoid or to reduce this type of imperfection one can growth the crystal in several steps as it was done for quartz crystals.

- This corresponds to the second technique which is the so-called "crossed-growth" (7). In the first stage, the seed is choosen and cut in a manner that the different desoriented parts or the zones where the dislocations are numerous are avoided. The growth is performed with such seed and the density of the defects (dislocations) are less numerous. If this not the case, new cycle is done in the same manner. This process may be repeated until the final crystal is of a good quality (Fig.5).

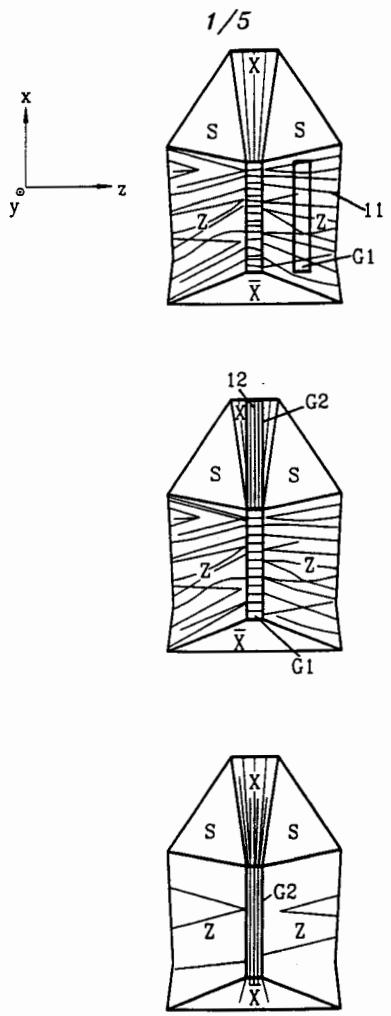


Fig. 5 "Crossed-growth"technique.

METHODS OF STUDY

In order to examine the different crystals obtained by described techniques for growth, experimental methods have been employed. Among them, observations at the optical microcope and X-ray topography. This second technique was used with laboratory equipments and with the Synchrotron radiation at L.U.R.E. (Orsay. France). In this last case we have used the Laue setting (8) which is the simplest. It consists of setting the sample perpendicularly to the white incident beam and recording the diffracted beam on a film. Due to the characteristics of the beam, all the diffracted spots are topographs and from their analysis the features of the defects present in the crystal can be determined.

RESULTS

In figure 6 is shown an optical photograph of a Z-cut slice. It may be observed that the quality is enough good. However, it remains an imperfect zone which corresponds to the joining zones of the seed. This may be also observed in the following figure 7 which is a topograph of the same sample. Numerous dislocations which were created at the level of the seed and above all those which are in the imperfect zones are visible.

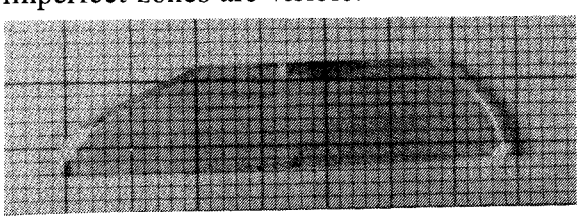


Fig. 6 Optical photograph of a Z-cut slice.



Fig. 7 Topograph of the crystal presented in Fig. 6.

Another example is presented in figure 8. This topograph was obtained from a Z-cut slice (X seed). This crystal presents a good quality and it may be observed the different growth zones. Nevertheless, these different parts of the crystal present numerous dislocations which were created at the start of the growth on the seed, or during the development of the crystal.

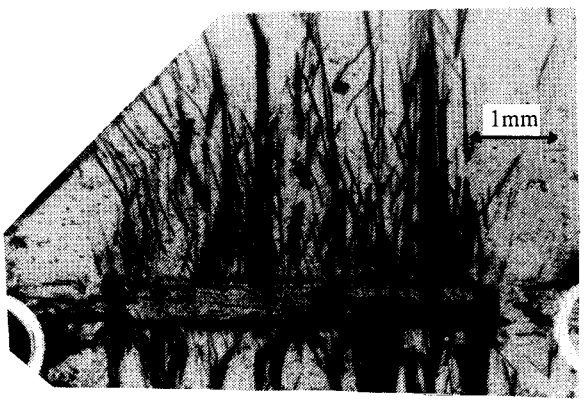


Fig. 8 Topograph of a Z-cut slice.

The topograph presented in figure 9 corresponds to another crystal which was grown in the same conditions as the previous samples. But in this case the different parts of this slice are enough desoriented and this can be observed by the different contrasts obtained. This is due to the misorientations of the lattice introduced at the level of the joining zones.



Fig 9 Topograph of a crystal which presents different desoriented zones.

The last example is given in the figure 10 which represents a topograph of a Z-cut crystal and only the seed and the X zone are visible. As the previous example the main features which are observed are the following. The start of the growth is not absolutly perfect and growth bands parallel to the seed are visible at the begining of the growth. It remains also several desoriented zones which are due to the sticked zones of the seed where bundles of dislocations are located. These defects are less numerous

than in other cases but they are still present and mark the different joined zones of the seed. Interference fringes due to progressive variation of the parameters between the seed and the growth zone may be also observed. These fringes are due to wave fields recreation of the X-rays at the boundary between the two slightly desoriented zones. They are visible all along the Y direction. The study of this sample by plane-waves X-ray topography (9) has shown (Fig. 11) that a desorientation of about one minute of arc exists between the lower part of this slice (seed) and the upper part (as-grown crystal). Therefore, the topograph obtained from these different crystals have shown that quite all the samples have a good quality in spite of the defects and the slight desorientations they present.

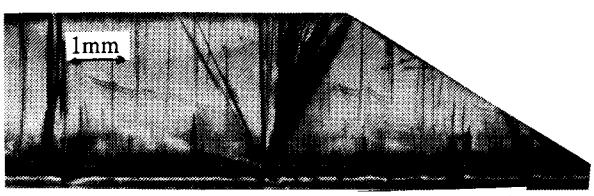


Fig. 10 Topograph of a Z-cut crystal.

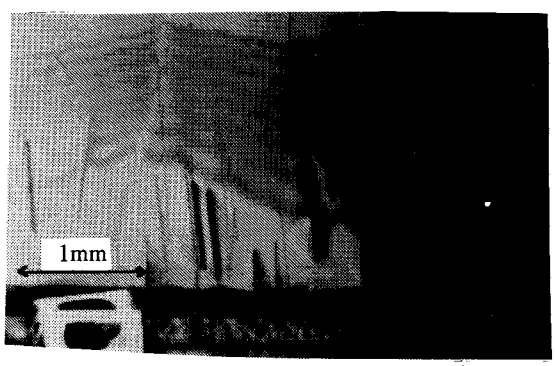


Fig. 11 Plane-wave topograph of the same sample as in Fig. 10.

CONCLUSION

Due to their piezoelectric properties, berlinite crystals of large size are needed for industrial purpose. In this work, we have shown that it is possible, by using different techniques of growth to obtain large crystals of good quality. We have used two techniques of growth. The first one starts by joining several crystals of the same orientation to make a single seed. If this is not sufficient, the second technique which is the "crossed growth" can be used. The obtained X-ray topographs from different samples have confirmed the good quality of the crystals.

REFERENCES

(1) J.H.SHERMAN. Proc. 40th Frequency Control Symposium. 1986. pp. 91-95.

(2) J. DETAINT, M. FELDMAN, J.HENAFF, M. POIGNANT, Y. TOUDIC. Proc. 33th Frequency Control Symposium. 1979. pp. 70-74.

(3).J..DETAINT,A.ZARKA, B.CAPELLE, Y.TOUDIC,J.SCHWARTZEL, E.PHILIPPOT,J.C.JUMAS,A.GOIFFON, J.C.DOUKHAN. Proc. 40th Annual Frequency Control Symposium .1986. pp. 101-104..

(4) E.PHILIPPOT, A.GOIFFON, J.C.JUMAS, C.AVINENS, J.DETAINT, J.SCHWARTZEL, A.ZARKA. Proc. 42st Annual Frequency Control Symposium 1988. pp. 138-145.

(5) B.H.T. CHAI, E. BUEHLER, J.J. FLYM. U.S. Patent. 1982 N° 4. 324773.

(6)J..DETAINT,J.SCHWARTZEL. Y.TOUDIC, E.PHILIPPOT,B.CAPELLE, A.ZARKA, A. GOIFFON, R. ARNAUD. Brevet d'invention. n° 233 072 D. 13606 MD. 24. 05. 1991. France.

(7)A.ZARKA,J..DETAINT, J.SCHWARTZEL, Y.TOUDIC,B.CAPELLE,Y.ZHENG E.PHILIPPOT, X. BUISSON R. ARNAUD. Brevet d'invention. n° 233 191 D. 13719 DL. 14. 11. 1991. France.

(8) T. TUOMI, K. NAUKHARINEN and P. RABE. Phys; Stat. Sol. A. 25. 1974. p. 93.

(9) A. ZARKA, LIU LIN and M. SAUVAGE . J. Cryst. Growth. 62. 1983. pp. 409-423.

TF-4020A: TFL's NEW COMPACT MILITARY RUBIDIUM STANDARD

E. Eltsufin & A. Stern

T.F.L., Time & Frequency Ltd.
P.O.Box 1792, Holon 58117, Israel

ABSTRACT

The paper describes and analyses TFL's new military Rubidium Frequency Standard model TF-4020A. It is the most compact (volume less than 500 cc) military Rubidium Standard available in the market today that meets MIL-E-5400 (airborne electronic equipment) class 2 (-55°C to +71°C, 70,000ft).

Key features of the unit:

a. Compact design.
b. Extremely small Physics Package (87cc) with a special rugged design to withstand strong vibration.
c. High stability over the wide operating temperature range of -55°C to +71°C.
d. Low susceptibility to EMI.
e. Low sensitivity to external dc and ac magnetic field.
f. Excellent short-term-stability of $7x10^{-12}/\sqrt{tau}$ for $0.1s<tau<300s$.
g. Excellent long-term-stability of $1x10^{-11}$ per month.
h. Versatility to adopt versions without size increase.

INTRODUCTION

The successful development and production of a compact industrial Rubidium frequency standard TF-4010A [1] allowed step forward to a Military Frequency Standard TF-4020A. The new model inherits features of the industrial one and in addition meets MIL-E-5400 class 2 requirements for airborne electronic equipment as a result of an intensive modification of a previous model. The unit, shown in Fig 1, is the smallest military Rubidium frequency standard available in the market today and at the same time it has advanced operational characteristics enabling it to compete with more expensive models.

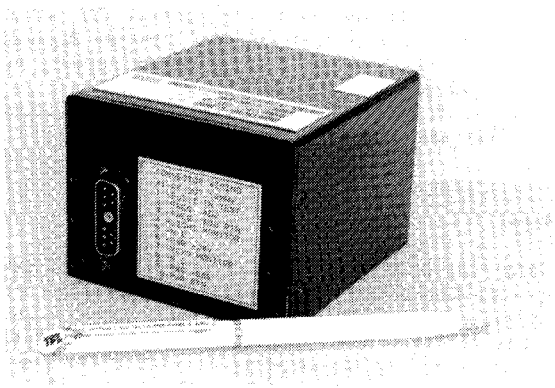


Fig 1 Rubidium Frequency Standard model TF-4020A

GENERAL DESCRIPTION

The functional block diagram is shown in Fig 2. The conventional type of frequency lock servo was chosen and the main goal was to get a **miniaturized, rugged and low cost unit with stable and repeatable characteristics.**
A 10MHz TTL signal from a TFL oven controlled crystal oscillator (OCXO) is multiplied by 9 to obtain 90MHz which together with a synthesized 5.3125MHz are injected into a Step-Recovery-Diode (SRD) multiplier. The SRD generates a Rb clock transition frequency equal to 6.834678500Hz from the 76th harmonics of the 90MHz and its 5.3125MHz side frequency. The clock transition frequency modulated by 444Hz is radiated onto Rb87 atoms in a resonance cell inside a microwave cavity.
A Rb87 lamp emits light filtered by a Rb85 filter cell. The light is transmitted through a Rb87 resonance cell to be detected by a photodiode. The signal which has passed an amplifier, a phase detector and an integrator comes to the OCXO thus closing the frequency-lock-loop.
The unit is fed through an EMI filter and an optional transient protector that could be installed when the unit is subjected by prolonged high voltage transients.
The unit is comprised of 4 modules: the control board, the Power supply board, the Physics-Package and the OCXO. These modules are packed into 93x82.6x65mm box. A general package design is shown in Fig 3.
The main electronics are condensed in 2 boards. The control board (78x88mm) includes the multiplier, synthesizer, preamplifier, frequency-lock-loop and Built-In-Test circuits.
The Power Supply board (78x44mm) includes a low frequency EMI filter, a DC/DC power converter as well as switching temperature controllers for the microwave cavity and the lamp oscillator of the Physics-Package.

The Physics-Package shown in Fig 4 is extremely small (87cc) and comprises lamp and cavity subassembly and a twofold magnetic shield. The oscillator is a miniature oven controlled crystal oscillator with a SC-cut crystal for a low vibration susceptibility.
The main effort during the design process was concentrated on the design of reliable circuits. That means general reliability, stability and low susceptibility to external and internal influences like temperature, vibration and electromagnetic interference (EMI).

DESIGN DESCRIPTION

Lamp Oscillator

The short-term stability of a Rubidium Standard depends on the noise factor of the light received by the photodiode. A low noise RF source is required to produce a low noise light.

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

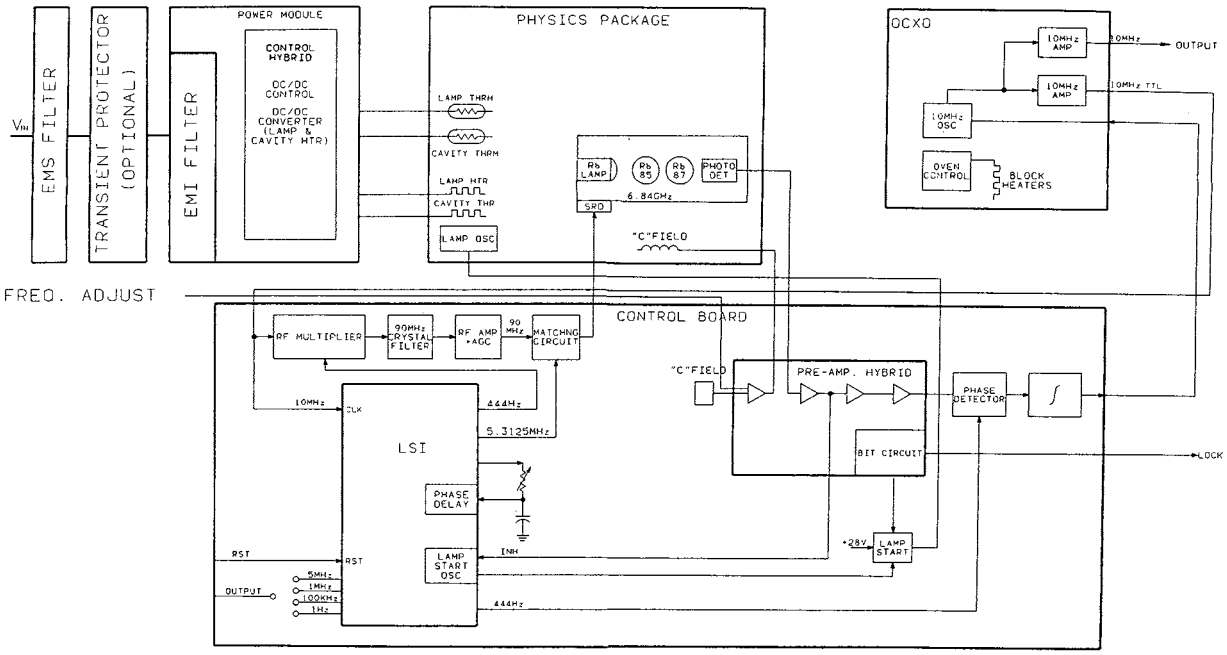


Fig 2 Functional block diagram

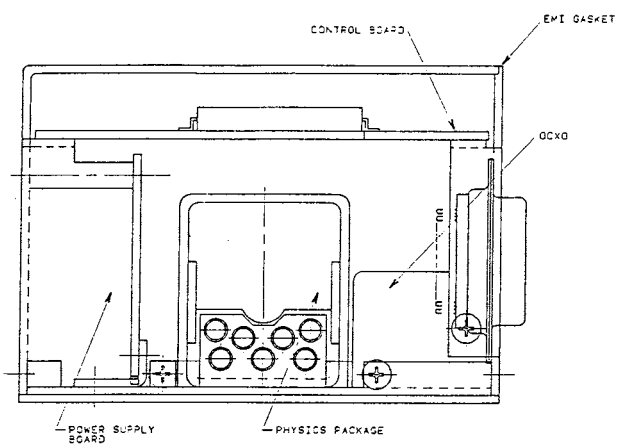


Fig 3 General package design

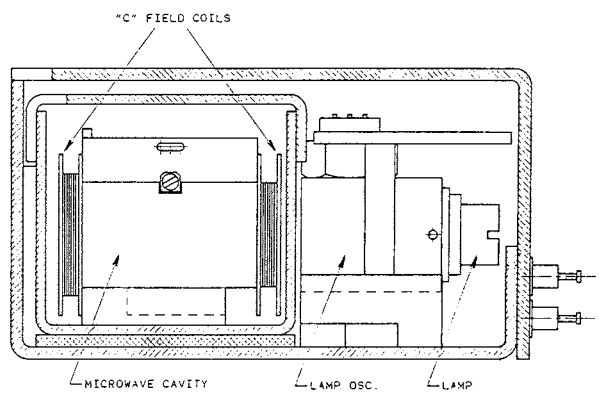


Fig 4 The physics package design

The second requirement is a minimum RF power with a minimum Electromagnetic emission. A very simple circuit was designed that does not increase the phase noise.

Microwave Cavity

A novel compact rectangular-cylindrical cavity partially loaded with dielectric is implemented [2]. It operates in a pseudo TE-101 mode. A high filling factor for the vapor cells together with a homogeneous RF magnetic field provides stable characteristics. The cavity has a low power consumption with quality factor more than 80. Extremely small value of the cavity together with careful thermal design provides a short warm-up time at a relatively small power consumption.

SRD multiplier circuits

The SRD multiplier circuit in a chip and wire form is located on the top of the microwave cavity to form an integral part of it. The circuit provides filtering and matching to a coaxial cable. This construction allows stable characteristics and real 50Ω impedance to preclude reflections and consequent spurious oscillation The actual required input power is about 0.1W which is the known threshold value for the SRD. A nonmagnetic construction drastically improves temperature stability.

LSI Synthesizer

A great deal of miniaturization is achieved due to the design of a custom LSI based on 1.5 CMOS technology. The LSI provides auxiliary and frequency synthesis functions required. It is built as a synchronous counter and includes a BIT circuit used during production phase. The optional TTL outputs of 5MHz, 1MHz, 0.1MHz, 1Hz and reset input are provided for custom applications which may often be utilized in the original package.
Excellent clock stability is achieved by the implementation of a very high speed LSI logic technology.

Hybrids

A large volume of the circuitry is condensed into 2 hybrids. The photo-diode preamplifier, precision current source for "C" field and a BIT circuit are included in a hybrid located on the Control Board. Most of the circuitry of the DC to DC converter as well as the switching temperature controllers are included in a hybrid located on the power Supply Board.

OCXO and the Frequency Lock Loop

A new very compact TFL OCXO in a miniature 20cc hermetically sealed package is used. It has a rugged construction and EMI protectors for better performance at hard environment. A Stress-Compensated (SC-cut) crystal is chosen to withstand vibration and to improve short and long term stability.

Generaly the short-term-stability (STS), usually measured by the "Allan deviation", of a Rubidium Frequency Standard is related to the OCXO for averaging times shorter than the loop-time-constant and STS depends on the Rubidium atoms for averaging times which are longer than the time constant. This partition is also being manifested in the frequency domain where the phase-noise is dependent on the oscillator for frequencies (from carrier) higher than the loop-bandwidth and depends on the Rubidium resonator for frequencies which are lower than the loop-bandwidth. We have chosen to use a loop-time-constant around 17ms and loop-bandwidth of around 100Hz to improve G sensitivity.

However, in order not to degrade the frequency and phase stability in regions that are still dependent on the Rubidium, the OCXO phase-noise at 444Hz from carrier should be kept below certain level (444Hz is our modulation frequency as described above).

Therefore a special effort was made to achieve a low phase-noise around 444Hz and to improve the oscillator. Typical results that were achieved are -140dBc (at 1Hz bandwidth) around 444Hz.

Output Frequency	10MHz 5MHz, 1MHz, 0.1MHz, 1Hz with synchronization
"Allan" deviation	$7x10^{-12}$ at 1sec $9x10^{-13}$ at 100sec
Phase-noise (1Hz BW,10MHz)	-105dBc at 10Hz -135dBc at 100Hz -145dBc at 1kHz
Long Term drift	$1x10^{-11}$/month
Temperature Stability	$1x10^{-10}$ from -55°C to71°C
Warm-up time 25°C, 22V	4 min to $2x10^{-10}$
Steady state power 25°C, 28V	12.5W
Start up power	51W
Loop time constant	17ms
Size	93x82.6x65mm

Table 1: Characteristics and typical results measured for first prototypes.

PERFORMANCE

Table 1 summarizes typical results as were measured for the first prototypes. The Rb standard characteristics are expected to be as repeatable as was proved during the production of the Industrial model which is very simmilar to the military one. Fig 5 show frequency offset versus supply voltage and ambient temperature.

Fig 6 shows frequency offset and power consumption during warm-up time.

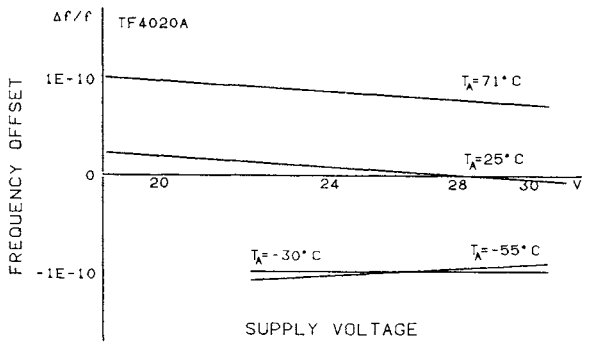


Fig 5 Frequency offset vs. supply voltage and ambient temperature

Fig 7 shows power consumption versus supply voltage and ambient temperature.

Fig 8 shows lamp-start peak current versus supply voltage during start-up.

Fig 9 shows warm-up time versus ambient temperature and supply voltage.

As seen from the table and graphs the exhibited performance is comparable or better than the performance usually stated for a higher priced Rubidium Frequency Standard sold in the market today. In particular we would like to emphasize on the **excellent short-term-stability** of 7 to $8x10^{-12}/\sqrt{tau}$, the **long-term-stability** of about $1x10^{-11}$/month and the **low power consumption** at start-up and steady state. The good short-term-stability is achieved due to the low noise lamp driver, the low noise oven controlled oscillator and the careful design of the multiplier and SRD circuit as explained above. The good long-term-stability is achieved by maintaining the same operating criteria used in bigger and higher priced units e.g. a low light-intensity sensitivity due to the light-shift effect. Low power consumption is achieved due to a careful thermal design of the whole unit.

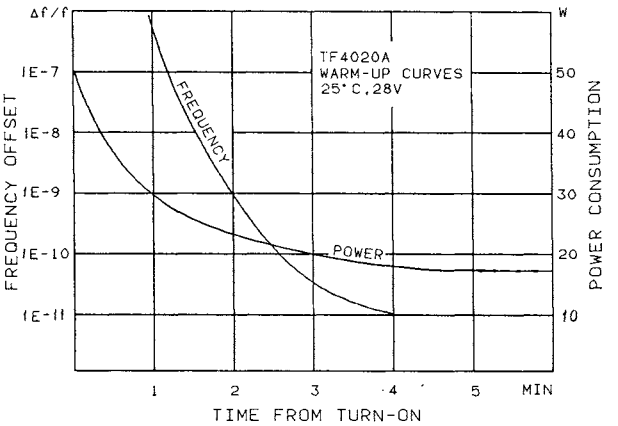


Fig 6 Frequency offset and power consumption during warm-up time

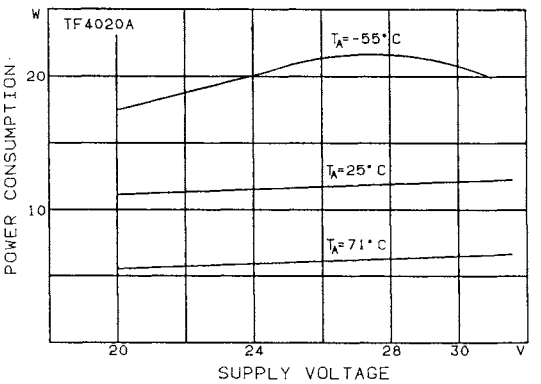


Fig 7 Power consumption vs. supply voltage

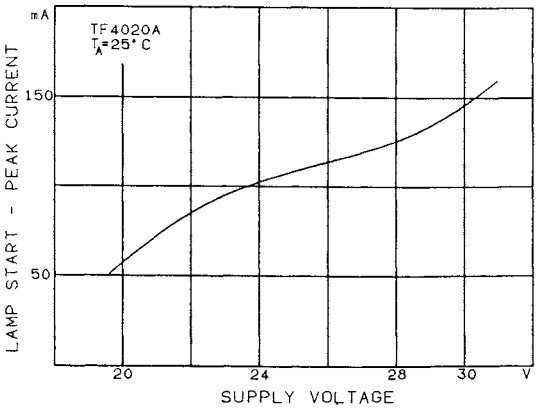


Fig 8 Peak current vs. supply voltage

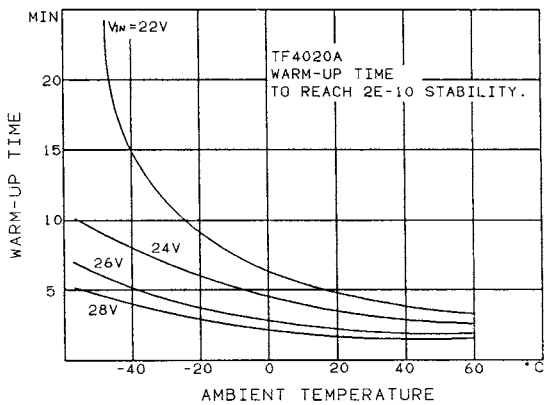


Fig 9 Warm-up time vs. ambient temperature and supply voltage.

MILITARY STANDARD QUALIFICATION

The unit has been qualified according to **MIL-E-5400 class 2** and frequency offset, BIT, DC light and STS have been monitored.
Environmental test has been performed according to MIL-STD-810C as below:

1. Low pressure, Method 500.1
2. High temperature, Method 501.1 proc II, storage at +95°C, operating +71°C.
3. Low temperature, Method 502.1, proc I: storage at -62°C, operating -55°C.
4. Temperature Altitude, Method 504.1 proc I:
   Operating at 70000 feet
   Continuous operation -54°C to 71°C on base plate.
   Intermittent operation +95°C, 30min
   (frequency offset is not specified).
5. Rain, Method 506.1, Proc II.
6. Humidity, Method 507.1, proc II: 94% Humidity temperature cycling.
7. Fungus, Method 508.1, Proc I.
8. Salt Fog, Method 509.1, Proc I.
9. Dust, Method 510.1,
10. Explosive atmosphere, Method 511.1, Proc I.
11. Acceleration, Method 513.2, Proc I.
    Worst case frequency offset $<2\times10^{-10}$ at 15g.
12. Vibration, Method 514.2:
    Random 5.2 GRMS low frequency profile.
    No frequency offset, STS $<4\times10^{-11}$ at 0.3sec
    Sine wave, procedure VIII at 4.2g.
    Frequency offset $<2\times10^{-11}$.
    Momentary frequency offset at the vibration frequency 444Hz is less than 1.2×10^{-10}.
13. Shock, Method 516.2 Proc I, Half sine shock 11ms, 20g:
    Momentary frequency offset $<1\times10^{-9}$.

Power characteristics meet **MIL-STD-704D.**
Normal operation: 22V to 29VDC at the operating temperature range.
Ripple amplitude 1.5V.

The standard was tested from 18V to 31.5VDC to have a specified frequency offset (see Fig 5).

The standard was tested at 2.8 Vrms ripple amplitude at the nominal voltage. Voltage transient protection is optional. It is recommended for implementation at a system level.

EMI test has been peformed according to **MIL-STD-461C.**
Emission tests include:
CE01 and CE03, conductive emission, to power leads, 30Hz to 15kHz and 15kHz to 50MHz. CE07, conducted switching transient inrush current.
RE01, radiated emission 14kHz to 10GHz.
The standard exceeds requirement for airborne equipment and meets Army requirements.
Susceptibility tests include:
CS01 and CS02, conductive susceptibility, dc power leads, 30Hz to 50MHz and 50MHz to 400MHz.
CS06, conductive susceptibility, spikes on power leads.
RS02, radiated susceptibility, magnetic induction field, spikes.
RS03, radiated susceptibility, electric field 14kHz to 10GHz.
External Magnetic field susceptibility:
DC magnetic field: 6×10^{-12} parts/gauss.
AC magnetic field: not susceptible. An STS increase was observed when subjected to 8gauss RMS AC magnetic field at frequencies 444Hz and 888Hz.

RELIABILITY

Most of electronics are incorporated into the LSI and into two hybrids. All parts are QPL level B (JANTX) or equivalent screen with full military derating. No failures were observed during the Qualification and evaluation.

CONCLUSION

We have described a Military model of a family of new Rubidium Frequency Standards made in **TFL.**
High performance, compactness and low cost and weight are the main features of the family.

REFERENCES

1. A. Stern and E. Eltsufin, TF-4010A - TFL's compact Rubidium Frequency Standard,
   Proc 5th European Frequency and Time Forum, 1991.
2. E. Eltsufin, A. Stern and S. Fel, Compact rectangular-cylindrical cavity for Rubidium Frequency Standard, Proc 45th
   Annual Symposium of Frequency Control, 1991.

ULTRA STABLE OSCILLATOR FOR SATELLITE APPLICATIONS

John C. Ho

Frequency Electronics, Inc., USA

Electronic and voice communication through the use of satellites have become common activities in our daily existence. Transmitting and receiving systems are demanding ever more accurate and complex timing to keep track of all the information flow. For this reason, a reliable ultra-stable signal source becomes a key component needed for both ground and satellite systems.

For the past twenty years, Frequency Electronics has developed, manufactured, and delivered hundreds of precision quartz crystal and Rb reference oscillators for ground station and space satellite operations. The most advanced communication satellite, MILSTAR, and others, use FEI's newest 5 MHz reference oscillators. The crystal oscillator package meets the long term (15 year mission) stability and reliability requirements demanded by space applications and the SC cut crystal, fabricated under the most exacting environmental conditions, meets the most challenging specifications.

The state-of-the-art circuit design and physical package provides an output frequency phase noise of −145 dBc/Hz 10 Hz from carrier (see figure 1); short term frequency stability is better than $5pp10^{13}$/second; temperature coefficient is less than $2pp10^{12}$/°C over the operating temperature −20°C to +50°C; vibration "g" sensitivity less than $6pp10^{10}$/g; and long term frequency stability less than $5pp10^{12}$/day in the actual space environment. A typical performance specification is shown in Table I.

These precision oscillators have frequency up-date capability allowing them to be tuned from the ground by command. A built-in 12 bit D/A converter with either parallel or series command codes controls the oscillator frequency. This tuning ability enables the oscillator to remain at an accurate frequency for missions of 10 to 15 years. Radiation-hardening protection is also incorporated as required. A system block diagram is shown in Figure 2.

Table II is a list of satellite programs and companies using FEI precision ultra stable crystal oscillators.

TABLE 1. ULTRA STABLE 5 MHz REFERENCE OSCILLATOR SPECIFICATION

1.0	Output Frequency	5 MHz
2.0	Output RF Power	8 dBm
3.0	Output VSWR	1.25:1
4.0	Frequency Stability:	
	a) Freq.vs B+(+5%)	$5pp10^{11}$/v
	b) Freq.vs Load (50Ω+10%)	$5pp10^{11}$
	c) Freq vs "g" sensitivity	$1pp10^{9}$/g
	d) Freq. vs Magnetic Field	$5pp10^{11}$/gauss
	e) Freq. vs Ambient Temperature	$4pp10^{12}$/°C
	f) Freq. vs Time	$5pp10^{12}$/0.1sec
		$2pp10^{12}$/sec
		$2pp10^{11}$/day
		$1pp10^{7}$/15 yrs
	g) Freq. vs Pressure (Air to Vacuum)	$1pp10^{9}$
	h) Freq. vs Pwr ON/OFF Cycles	$5pp10^{9}$
	i) Freq. vs Storage Time	$2pp10^{9}$/year
5.0	DC Input Voltage	20 volts
	DC Input Power at turn-on warm-up	8 watts

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

TABLE 1. ULTRA STABLE 5 MHz REFERENCE OSCILLATOR SPECIFICATION (CONT'D)

	Operating at +25°C (Air)	2.4 watts
	Operating at +25°C (Vacuum)	2.0 watts
6.0	Stabilization Time	2 hours from power on at +25°C
7.0	Environmental Conditions:	
	a) Random Vibration	12.8 grms
	b) Shock	20 gms
	c) Operating Temperature	-20°C to +50°C
	Storage Temperature	-40°C to +70°C
	d) Radiation	Total dose 100Krad
8.0	Size:	7.65"L × 3.2"H × 2.5"W 19.43cmL × 8.13cmH × 6.35cmW
9.0	Weight:	32 oz.

TABLE II. LIST OF SATELLITES THAT USE ULTRA STABLE CRYSTAL OSCILLATORS SUPPLIED BY FREQUENCY ELECTRONICS, INC.

SATELLITE NAME	ORGANIZATION
FLTSATCOM	TRW
SPACE SHUTTLE	WESTINGHOUSE
HEAO	TRW
APOLLO-SOYUS	RAYTHEON
VOYAGER	JPL
DSCS III	GE
INTELSAT V	FACC
ARGOS	ESD

TABLE II. LIST OF SATELLITES THAT USE ULTRA STABLE CRYSTAL OSCILLATORS SUPPLIED BY FREQUENCY ELECTRONICS, INC. (CONT'D)

SATELLITE NAME	ORGANIZATION
SARGOS	ESD
SARSAT	SPAR
GALILEO	HAC
LANDSAT	TRW/HAC/GE
HS-350	HAC
GPS	USNRL/ROCKWELL
SPACE TELESCOPE	LOCKHEED
INTELSAT VI	HAC
MISSION 34	HAC
DSP	TRW
DMSP	RCA/GE
MILSTAR	HAC
ERBS	GULTON
ACTS	TRW
MARS	GE
TELECOM II	ALCATEL
SPOT 3	ESD
SARP 2	ESD
TOPEX	ESD
UHF FOLLOW ON	HAC
EHF FOLLOW ON	HAC
TIROS	GE
SNR	TI
SPACE STATION	GE/GULTON

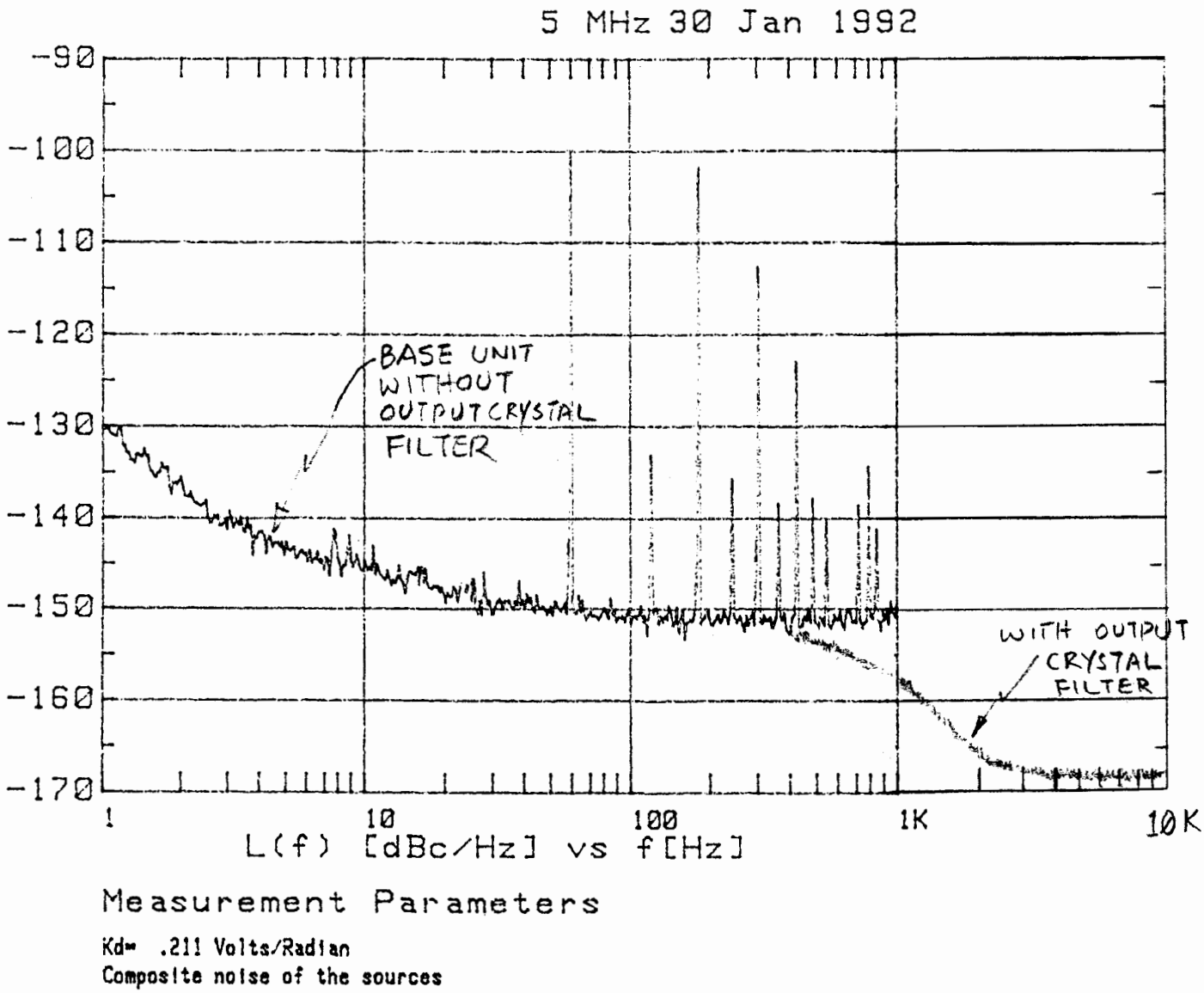

5 MHz 30 Jan 1992
-90
-100
-110
-120
-130
-140
-150
-160
-170
BASE UNIT WITHOUT OUTPUT CRYSTAL FILTER
WITH OUTPUT CRYSTAL FILTER
1
10
100
1K
10K
L(f) [dBc/Hz] vs f[Hz]
Measurement Parameters
Kd= .211 Volts/Radian
Composite noise of the sources

Figure 1.

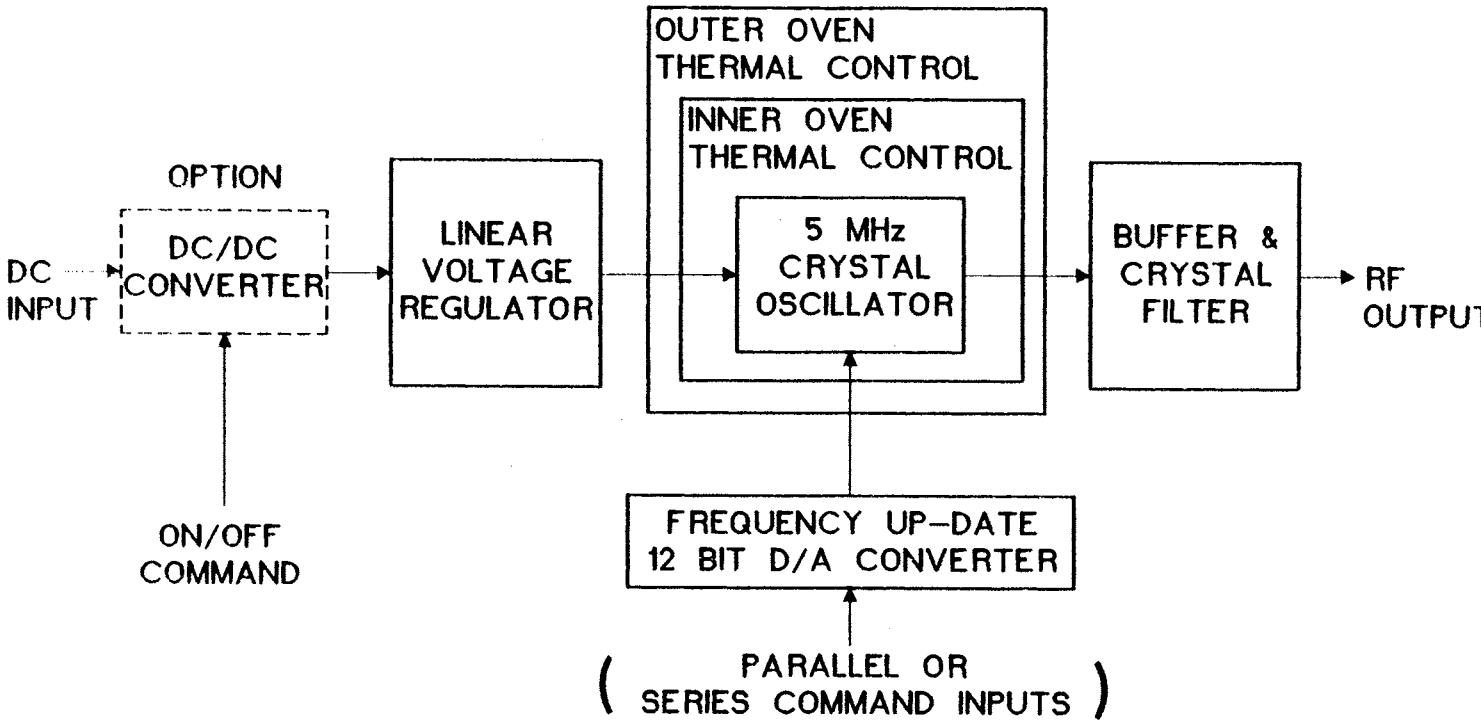


Figure 2. Block Diagram

PRELIMINARY RESULTS OF A NEWLY DESIGNED "MINI OCXO" OSCILLATOR ESPECIALLY DEDICATED TO SPACE APPLICATIONS

by Evelyne GERARD and Gerard WAGNER
C.E.P.E. 44 avenue de la Glacière

95100 ARGENTEUIL (FRANCE)

Abstract

Recent requirements of space equipment designers concern ovenized oscillators, small size, stability in 10^{-8}, to be integrated directly in the equipment. Dimensions and weight are so critical that the CEPE Space USO (Ultra Stable Oscillator) is not suitable for these applications.

So, to answer this identified need, CEPE has under development, a new ovenized oscillator, especially dedicated for Space Applications, called "MINI OCXO". This OCXO, will be available, on the space market, in 1993 for Flight Models, at a nominal frequency between 5 to 20 MHz, in a relatively small holder, (82x50x35 mm) with frequency stabilities in the range $10^{-8}/10^{-7}$.

This paper presents the different steps of the development phase, a short description of the design and the first results obtained on the Breadboard Model of the "MINI OCXO"

I.- INTRCDUCTION

To meet recent needs expressed by Space Equipment designers about ovenized oscillators of small dimensions and weight, intermediate stability (10^{-8}), a new product is already under development at CEPE.

The target is to design in the smaller holder, an oscillator, available in frequency range 5 to 20 MHz, with stability performances in the range $10^{-8}/10^{-7}$, in respect with space requirements, in terms of design, reliability and quality matters.

In that way, miniaturization of this "MINI OCXO" needs the use of new technologies in the Space domain, as SMD components.

II.- TECHNICAL SPECIFICATIONS

The main characteristics required to the MINI OCXO. are defined in the following paragraphs

2.1. Frequency range

5 to 20 MHz

2.2 Frequency accuracy

Setting accuracy : $\pm\ 1.10^{-8}$

The frequency adjustment, by resistor bridge, is internal, but easily accessible.

2.3 Frequency stability

a) The frequency stability is defined versus overall parameters, temperature, DC supply, load variations.

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

temperature range : - 20 to 60 °C
DC supply : ± 5 %
Load : SRWR < 1.5

specified : < ± 5.10^{-8}
objective : < ± 5.10^{-9}

b) Long term stability :
per year : ± 5.10-8
over 10 years : ± 5.10-7

c) The stabilized frequency at ± 1.10-9 is obtained after a warm-up time of 30 minutes.

2.4 Output characteristics

One sine output is available with the performances listed on following table 1.

table 1

CHARACTERISTICS	SPECIFICATIONS
output level	0 to 6 dBm
harmonics	30 dBc
Non harmonics	30 dBc
Phase Noise 10 Hz 100 Hz 1000 Hz >10 KHz	 -120 dBc/Hz -130 dBc/Hz -140 dBc/Hz -150 dBc/Hz
Output SRWR	<1.5 on 50 ohms load

2.5 DC Supply interface

DC voltage : 12 V or 24 V ± 5 %

Ripples on the DC line :

200 mV peak to peak between 30 Hz to 50 MHz , depending on the application.
With such level of ripples, the generated spurious have to meet the non harmonics specification.

Input powers are defined in the table 2.

table 2

INPUT POWER	AMBIANT PRESSURE	UNDER VACUUM
Max	6.0 W	6.0 W
at -20 °C	5.0 W	2.0 W
at +20 °C	4.0 W	1.0 W

2.6 Mechanical Environment

The qualification levels for mechanical environments are based on new requirements of launchers :

a) Random vibrations

20 Hz to 100 Hz + 6 dB/oct
100 Hz to 1000 Hz 0,5 g2/Hz
1000 Hz to 2000 Hz - 6 dB/oct

composite value : 26 g rms
on the 3 axes,
with a duration of 2 mn/axis.

b) Sine vibrations :

5 - 100 Hz : 30 g
2 oct/mn
on the 3 axes

c) Constant acceleration :

30 g
1mn/axe
on the 3 axes

2.7 Thermal environment

a) Operating temperature range :
- 20 to + 60°C

b) Non-Operating temperature range :
- 40 to + 70°C

c) Cold start-up : - 30°C

d) Pressure :

Operating under vacuum and ambient pressure, as well.

2.8 Reliability Requirements :

a) Life time : Minimum 10 years

b) Reliability prediction :
300 fits .

2.9 Mechanical/Electrical Interfaces

a) Dimensions : 82 x 50 x 35 mm

b) **Mounting :** 4 holes diam.2.5 mm

c) Weight : < 150 g

d) Electrical interfaces :

DC supply Input : on by-pass filter
Ground (-) : on pin
RF output : on RIM/SMA coaxial connector

III.- MINI OCXO DEVELOPMENT PHASE

3.1 Design Phase

The first phase of the development of the Mini Ocxo has consisted to define the electronic functions and the thermo-mechanical structure of the oscillator ,taking account all the technical specifications, especially the dimensions requirements, but also the specific space rules concerning parts, materials and technologies used in the design, stresses on parts, and so.

According the background of CEPE in matters of ovenized oscillators, the electronic concept is based on the design used on the new generation of OCXO and adapted according to specific space requirements.

The design of the structure is optimized by both theoritical and experimented approaches, to obtain mechanical resonances outside the spectrum of vibration.

The design will use technologies and concepts, first validated during development phase of the Space Ultra Stable Oscillator. However, to meet the dimensions goal, technology intended for components is Surface Mounting type , previously used on the USO, but at level of capacitors only.

In that way, qualification works first concluded under CNES survey, have to be extended to the new added SMD components, selected when possible, in the PPL, according to specific programs, first approved by the Space agency.

3.2 Verification of the design

3.2.1 Mechanical aspect

To improve the mechanical design, the first step consists to built a mechanical model, with no electronic functions. Tests under vibrations have to be performed in order to determine the mechanical resonance of the structure, and in particular, to verify these frequencies are higher than 2000 Hz, frequency limit of the vibration spectrum.

3.2.2. Electrical aspects

A breadboard model is built and tested to verify the technical specification.

First results obtained on this model show the ability of the design to meet the foreseen technical and environmental requirements.

3.3 Following phases

The phases planned after the demonstration of the feasibility of the product, consist to qualify the design on one or more Qualification Engineering Models according to a qualification test plan.

IV.- TECHNICAL DESCRIPTION

4.1 Thermal and mechanical design

As the Mini Ocxo is integrated in equipments and so, can see high level of vibrations, the design of the mechanical structure is defined as so to avoid mechanical resonance under 2000 Hz.
An other paramater taken account for the structure design is the facility of assembly and the accessibility of the different parts of the oscillator during mounting and demounting steps.

With both considerations, the design consist of one central oven, mounted on two PC boards, the whole integrated in a rigid holder.

- The central oven, a cubic aluminium alloy box, is ovenized by power transistors, welded on 2 faces. The crystal resonator in HC37 holder and the oscillator network are fitted in this oven.

- This oven is fixed by screws between two PC boards, to ensure the mechanical mounting and thermal insulations.

On the first PC board, which is below, are laid out all functions that do not require adjustements, DC power distribution with, filtering and regulation functions, and the heating function.

On the top, on the same PC board, easily accessible for adjustments , are laid out the thermal control function and the output amplifier.

Connections between the different PC boards are today ensured by wires. In the future, flexible PC boards will be developped to simplify the mounting.

- The external holder of the MINI OCXO rigid box in aluminium alloy, ensures the mechanical interface with fixation by four screws M2.5. The electrical interface is realized by by-pass filter and feedthrough pin for DC supply and coaxial connector RIM/SMA type for Output frequency, located on the same face.

4.2 Electrical Design

As classical ovenized oscillator, the electronic diagram contains three functions :

a) the metrologic function consists of the oscillator built as using a resonator, SC cut, third overtone, in PQ swept material, and the amplifier stages.

The adjustment of the frequency is available by a resistors bridge, located on the PC board which is on the top.

For frequencies less than 10 MHz, due to the size of the resonator holder, the oscillator works at the double of the nominal frequency and a divider by two is included in the output stage.

b) To maintain the oven at the turn-on point of the resonator used, with a good accuracy, the thermal control includes a PID in the loop, using the temperature information given by a thermistor, with a good long term stability.

The heating is assured by two power transistors, welded on the oven.

c) The DC power distribution includes a by-pass filter at the head, to avoid the ripples effects on the output signal and is dispatched after regulation, in two

lines, one for the metrologic function , one for the thermal control. The heating is directly driven by the filtered DC supply.

The basic design can accept 24V or 12V DC supply, as well.

V.- PRELIMINARY RESULTS OF THE MINI OCXO

5.1 Mechanical results

A mechanical breadboard model, built with an accelerometer in place of the resonator, was submitted to random vibration. A low level, 1 g rms, was applied, to visualate the mechanical resonance of the structure, in each of the three axes.

Results show that the lower mechanical resonance frequency is higher than 2000 Hz, about 3000 Hz.

Tests under high level are not yet performed, but according to the design and CEPE experience, we can ensure that the model will be successfully submitted to the qualification spectrum .

5.2 Thermal Results

To verify the electrical characteristics, a breadboard unit is built according to the defined design, at 10 MHz with a DC supply of 12 V.

To well identify the thermal behaviour of the Mini OCXO, the first step has consisted to adjust the oscillator on the thermometric mode of the resonator and to test in temperature.

The equivalent temperature variation observed by the resonator is less than 1°C for a temperature range of - 20 to +50 °C, under vacuum and at atmospheric pressure, as well.

5.3 Electrical results

5.3.1 Frequency characteristics

The table 3 gives the comparison between the defined specification and first characteristics obtained on the Breadboard. We can see that the preliminary results are very promising.

On this table, the stability in temperature range is given peak to peak. Typical curves are given in figures 1 and 2.

Table 3 :Frequency stability results

PARAMETER	SPEC	UNIT	BB RESULTS
in temperature range - 20 to 50°C			
under vacuum	NO SPEC	10^{-9}	± 2
at ambient pressure.	NO SPEC	10^{-9}	± 2
DC variation $\pm$ 5 %	NO SPEC	10^{-10}	<1
SRWR < 1.5	NO SPEC	10^{-10}	$< \pm$ 2.5
OVERALL STABILITY	$\pm$ 5	10^{-9}	± 2.25

We see very low DC supply and load sensitivities, showing the efficiency of the electronic design as particularly, the DC supply regulation.

For information, on this breadboard, the frequency drift between ambient pressure and vacuum is about 5.10-9.

Figure 1 : Frequency stability
at ambient pressure

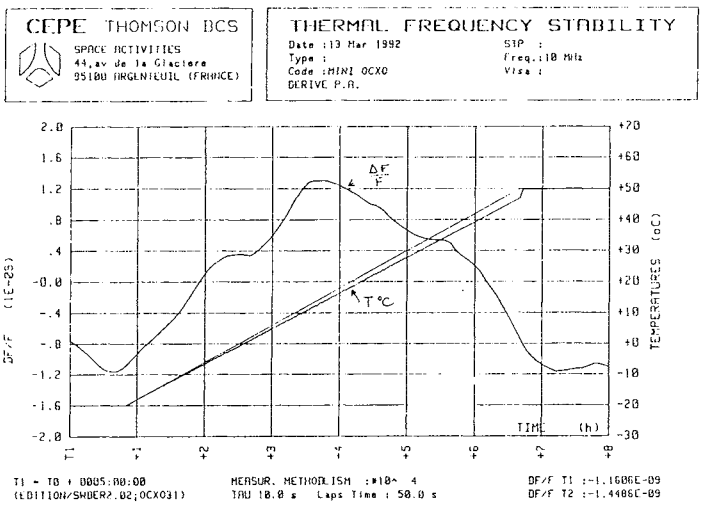


table 4:
Output Signal characteristics

OUTPUT SIGNAL	SPEC	UNIT	BB RESULTS
Level	0 to 6	dBm	5
Output SRWR on 50 ohms	<1.5		1.15
Harmonics	<-30	dBc	< -60
Non Harmonics	<-30	dBc	< -70
Phase Noise			
10 Hz	- 120	dBc/Hz	- 130
100 Hz	- 130	dBc/Hz	- 143
1000 Hz	- 140	dBc/Hz	- 151
> 10 KHz	- 150	dBc/Hz	- 154

Figure 2 : Frequency stability under vacuum

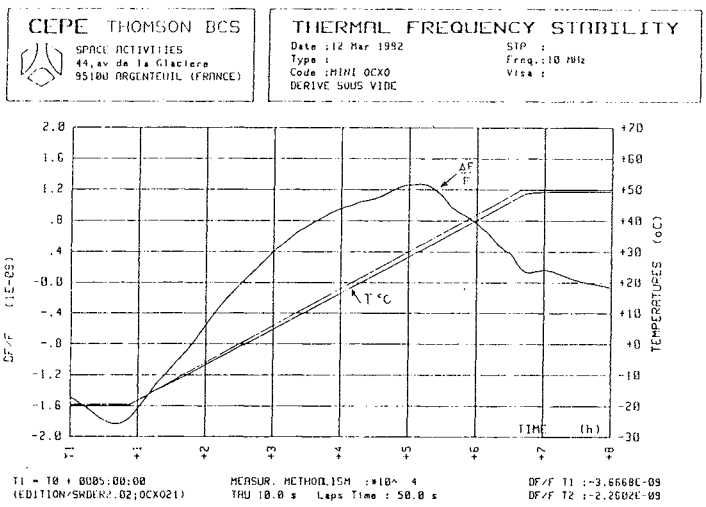
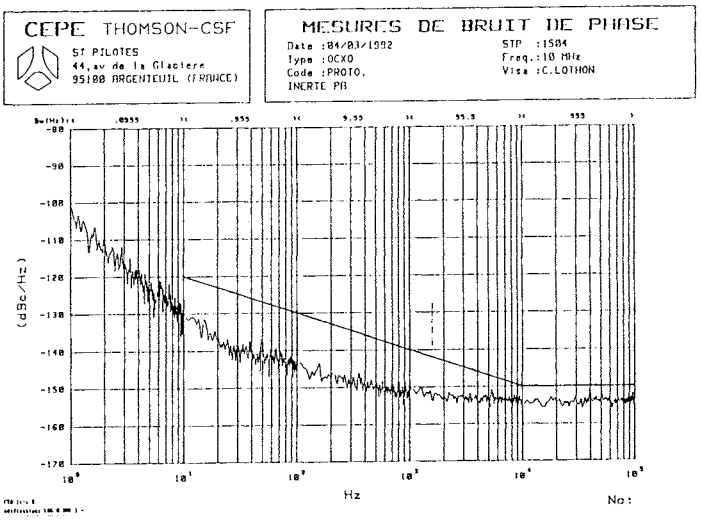


Figure 3 : Phase noise measurement



5.3.2.Output signal characteristics

The table 4 and figure 3 give the characteristics of the output signal.

5.3.3 DC power characteritics

First results about DC power at ambient pressure and under vacuum are given hereafter.

table 5 : DC power characteristics

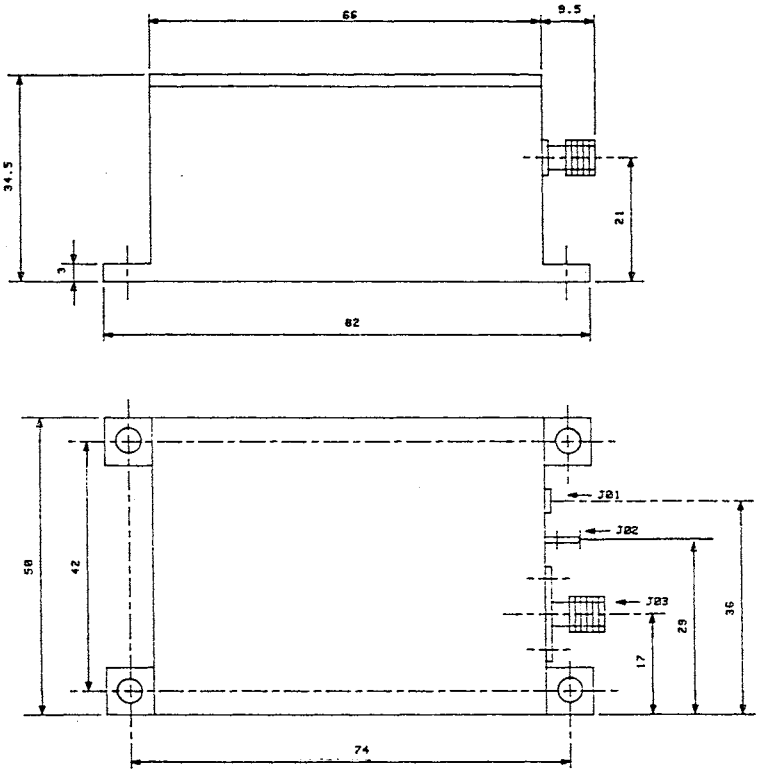
DC Power	spec	unit	BB results
at start-up	6.0	W	6.1
under vacuum			
- 20 °C	2.0	W	1.8
+ 20 °C	1.5	W	1.1
+ 50 °C	1.0	W	0.5
ambient pressure			
- 20 °C	5.0	W	5.0
+ 20 °C	4.0	W	3.4
+ 50 °C	3.0	W	1.3

5.3.4 Weight and Dimensions

The objectif size is yet met as shown by figure 4 that details of outline dimensions (83x50x34.6mm).

About weight, the breadboard is weighter than expected of about 10g.

Figure 4 : MINI OCXO Outlines



VI.- CONCLUSION

To answer a recent need expressed by space equipment designers of ovenized oscillators, intermediate performances, small dimensions and weight, a new product called "MINI OCXO", space devoted, is already under development at CEPE.

The target of the R/D phase is to put on space market in 1993, an OCXO compatible with the identified need of the users, in terms of weight, dimensions and performances.

This new product will be available in frequency range 5 to 20 Mhz, in a relatively small holder (83x50x34.6) low weight, 160g max, with nevertheless, stabilities, for the higher version, in range 10^{-9},10^{-8}.

To fullfil the dimensions criteria with in the same time, respect of space rules about the lay-out of PC boards, new technologies in space activity have to be implemented, as use of SMD components.

Results obtained on the first breadboard, electrical, mechanical and thermal, as well, are very promising : low power under vacuum, less than 2W under 12V DC supply, good stability versus temperature , +-2.10-9 in a relatively wide range good phase noise floor, no mechanical resonance under 2000 Hz.

When the first requirements are weight and size, with nevertheless stabilities awaited for ovenized oscillators, this new product, the "MINI OCXO" is the best solution.

This "MINI OCXO", will increase the selection of CEPE products devoted for Space Applications, including resonators, quartz crystal filters, hybrided clock oscillators OTO16S, miniatured TCXO (50x50x18) and the Ultra Stable Oscillator.

USE OF THE TV-SIGNAL IN STABILIZING A LOCAL RUBIDIUM OSCILLATOR

Reijo Kivelä

University of Oulu, Finland
(Jutel Oy, Oulu, Finland)

ABSTRACT

This paper describes the methods and results of using the 15625 Hz frame synchronizing frequency of a normal TV- broadcast as a reference to stabilize automatically a local rubidium oscillator.

The characteristics of the Finnish TV -network were studied. Consequently, TVCRO (TV Controlled Rubidium Oscillator) microcomputer control algorithms and TVCRO prototype hardware were designed. Prototype measurements showed that the daily frequency can be maintained within 5x10exp-12 (peak to peak). The standard deviation around the mean value is 0.9x10exp-12. During a 40 day test run, the TVCRO output phase was within 0.8 µs of the TV -phase. The long term drift of the TVCRO is in the order of a few parts in 10exp-13 per year when referenced to transmission centre frequency.

Keywords: Standard Frequency, Rubidium Oscillator, TV Frame Syncronizing Pulses.

1. INTRODUCTION

The TV-signal frame synchronizing pulses have been used for precision time comparison and frequency distribution for a long time. The high precision of the method has been demonstrated in many studies.

For remote oscillator synchronization, however, the TV-technique has not widely been used. Other radio frequency signals are more common for this purpose. Today especially the GPS- system have gained the attention. This is partly explained by the special drawback features involved in the normal TV-broadcast. The signal may exist only part of the day. Also, the signal phase may include occasional and irregular jumps caused by link path or video program rerouting.

The advantages of the TV-signal for time/frequency transfer in certain applications are obvious. The signal is easily received. In many countries, the TV- network is uniform. The frequency be nationally controlled and the method offers an independent distributing system.

In Finland, the TV frame synchronizing center is synchronized with a high quality rubidium oscillator. The TV signals have been used for frequency calibration in Finland since 1985. There are commercially available TV- calibrators for this purpose. Plans have been made for locking the TV - frequency with the national standard frequency in the near future.

The microprocessor software technology enables intelligent filtering of the received TV-signal phase. This makes it possible to use theTV-signal for automatic oscillator synchronizing. Studies of the TV-signal performance were performed from this point of view. Methods to use the signal as a reference for oscillator control were evaluated. A prototype of a TVCRO (TV-Controlled Rubidium Oscillator) was constructed and the performance studied.

2. TV -SIGNAL CHARACTERISTICS

2.1 Previous results

Many studies of the TV- signal stability have been reported (Ref. 1-5). However, many of the results can not be directly used to characterize the properties of the TV-signal for automatic oscillator controlling. Many of the important features of the TV- transmission system may vary in different networks. In Finland, previous measurements of the TV frame signal phase stability have been made (Ref. 6, 7).

2.2 Measuring method

Most of the TV measurements were made in Oulu. The TV- signal is routed to this town by 12-13 microwave link paths totalling approx. 600 km from the transmission center in Helsinki. The TV-signal phase was registered by using a normal TV-receiver, a sync separating circuit and a microcomputer controlled automatic measuring system.

The microcomputerized measuring system (Ref. 8) collects data from maximum 16 analog channels and 16 digital channels and stores them on a microcomputer hard disk. The analog channels are used for monitoring environmental data, control voltages etc. The digital channels receive information from the HP-counter, which measures the phase difference of various oscillators with 1 ns resolution against the Main Clock (R&S Rubidium). The Main Clock starts the counter and it is stopped by the next pulse of an individual oscillator.

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

The microcomputerized measuring system measures normally the data values at 10 min. intervals. The user can monitor a selected channel also with a selectable time interval (1-50 s). The 10 min data are stored to Q-data files, 10 days each file. The S-data files (10 days each), which are normally used, are 2 h filtered averages of the Q-data.

The measuring system contains several precision oscillators. In this study, the Loran receiver (internal crystal oven oscillator) and a free running FRS-C rubidium oscillator were used for reference measurement. The temperature effects were not analyzed in this study, because all of the measured frequency standards were either in a temperature controlled cabinet (Main Clock, Loran, FRS-C) or had a separate temperature control (TVCRO).

3. RESULTS OF THE TV -SIGNAL STUDIES

3.1. General properties of the TV-signal

In October 1991 the Finnish Broadcasting Company started to transmit the channel 1 signal 24 hours per day without any regular interrupt.

The TV- signal phase includes jumps caused by occasional link path rerouting. These vary by amplitude and also by appearing time interval in different parts of the country. The typical occuring time interval for these jumps is from a few hours to days and the amplitude can be even several hundred microseconds. The phase shows several clear alternate levels of link path delays as can be seen also in fiqures 4-5.

The daily TV- phase contains also occasional smaller phase jumps of 10 - 40 ns, which may occure at the change of program source linking.

3.2. TV-signal Allan Variance

The TV-signal measurements are summarized in figure 1, which shows the Allan variance (Ref. 9) of each sampling time interval. It must be noted, that the calculated variances also include the effect of the Main Clock (R&S rubidium), which was not eliminated due to the much different behaviour of the oscillators.

The 1s, 10 s, 100 s results are are calculated directly from phase data values measured using the manual mode of the measuring system.The 600 s, 2400 s and 9600 s Allan variances are calculated from 10 minute interval phase sample of 5 days. This clean sample contained no significant phase jumps during the period, which made it ideal for the purpose.

The 1 day, 10 day and drift values of the Allan variance graph represent mostly the long term stability of the TV-synchronizing oscillator itself. These data points were obtained by calculating them from the data received from the Technical Research Center of Finland (VTT). The data contained the reported behaviour of the daily TV-frequency. It is monitored by a receiver in Otaniemi approximately 10 km from the TV transmission Center and compared with the Finnish national standard frequency. This data shows in August 1990 - August 1991 the maximum drift of approximately 0.5exp-11 in 100 days, which may partially be explained also from long term seasonal temperature variation.

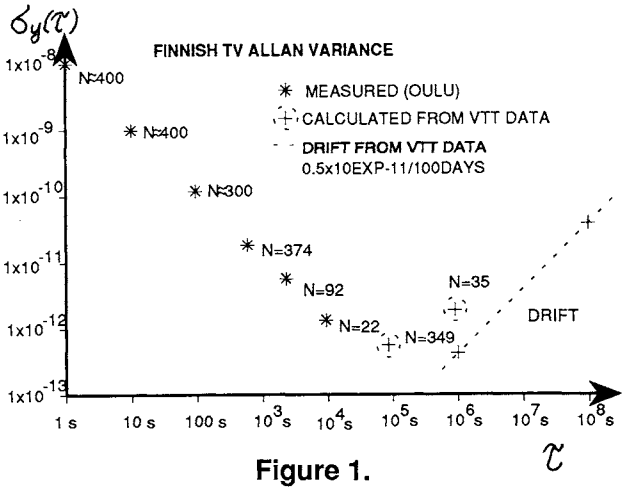


Figure 1.

4. THE TVCRO CONSTRUCTION

4.1. The Control Algorithm

The control time interval was selected to be 1 day, because this was considered straightforward and natural in the case of a TV -signal. Every night (at 24:00) the control algorithm replaces the control voltage with the new value calculated from the last day data.

The control algorithm calculates the frequency deviation during one day and converts this to a change of the control voltage to force the frequency to the proper value. The frequency deviation is determined by measuring the phase drift. The algorithm searches for continuous phase samples, that are clean from phase jumps and uses them for calculations.

The phase is observed in 5 second intervals. Each of these data points represent the median of independent 5 measurements. Because the phase is not continuous, but may include jumps, the software stores 10 most significant phase sample intervals and calculates a weighted frequency deviation value accordingly. The weighting factor of each sample is determined by the sample lenght and inversely to the approximated phase data standard deviation during 20 first and 20 last data values of the sample.

The standard deviation of the phase data values is also used to detect the phase jumps. If 4 successive data values exceed 3 standard deviations in the same direction, this indicates a phase jump and a new measuring interval is started.

The resolution of the digitized phase is 4 ns. Because of natural phase noise, the measuring resolution of the phase is approximately 1 ns (4 ns $\sqrt{20}$).

4.2. The TVCRO hardware

The TVCRO prototype simplified block diagram is shown in figure 2. In addition to the details shown, the TVCRO contains many special diagnostics and monitoring features. The software interface with the user through an RS 232 bus.

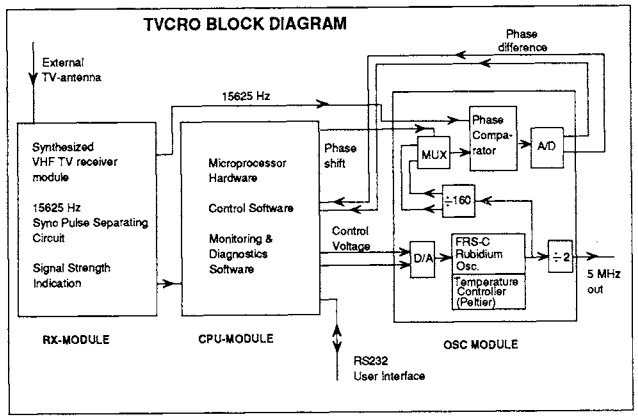


Figure 2.

The receiver module (RX) contains a synthetized VHF TV -tuner and circuit for separating the 15625 Hz horizontal scanning frequency.

The Oscillator (OSC) module contains the FRS-C rubidium oscillator with Peltier element temperature control. The analog phase comparator in the OSC module compares the TV 15625 Hz signal with a 62500 Hz (4x15625 Hz) divided from the 10 MHz rubidium oscillator frequency. The phase of the 62500 Hz signal can be selected to avoid the dead zone of the phase comparator.

The microcomputer module (CPU) receives the digitalized phase difference with a resolution of approx. 4 ns/bit. The software control algorithm produces a digital control word, which is converted to an analog control voltage for the Rubidium oscillator. One bit of the digital control word represents a frequency change of approx. 1.1 * 10exp-12.

5. MEASUREMENTS OF THE TVCRO

Several test periods were run with the TVCRO prototype to evaluate the frequency stability. An external Tv-antenna was connected to the TVCRO and the TVCRO 5 MHz output frequency was connects to the microcomputerized measuring system through a 1/10000 -divider unit. The 2 hour interval phase data was used for the analysis.

5. 1. Phase results

Figure 3 shows the phases of a Loran receiver and an external FRS-C rubidium oscillator referenced to the Main Clock. The FRS-oscillator results are corrected to have zero frequency offset at the start to give only the drift effect. Compared with the TVCRO, this shows the true behaviour of a FRS-C rubidium unit, which has been working for approximately 1 year.

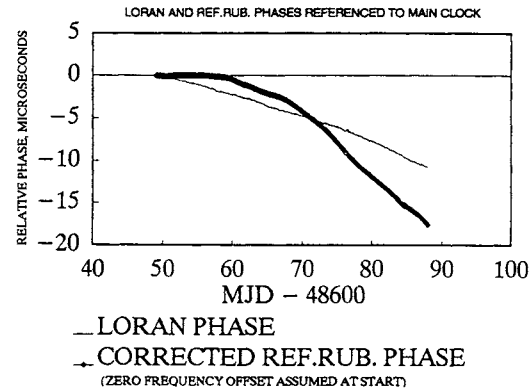


Figure 3.

Figures 4-7 show the method of cleaning and the result of the relative TV -phase. The symboled graph in figure 7 represents the filtered TV- phase referenced to the main clock. The phase jumps have been eliminated.

The TV-phase cleaning has been completed by simple "cut and paste" procedure, which is accomplished by LOTUS 123-software and graphical monitoring as follows:

In figure 4, the value of 8897.1 μs is added to the original measured values below 7000 μs (STEP 1) giving the symboled graph. This data showed a jump at MJD 48670.45, which was removed by adding 30.1 μs to all the values after that time giving the line graph in fig. 5 (STEP 2). This was further cleaned by subtracting 355.67 μs from the values over 12000 μs (STEP 3) giving the symboled graph in fig. 5. In STEP 4, the value of 74 μs was subtracted from all values over 11700 giving figure 6. Finally, after removing values over 11670 μs and some individual values and moving the start to zero, the symboled graph of figure 7 was obtained (STEP 5). This data contains N=437 values (original data N=469 values).

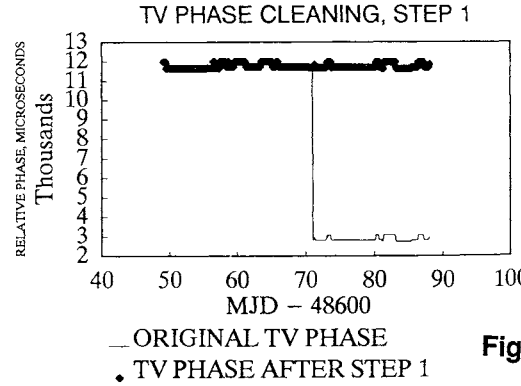


Figure 4.

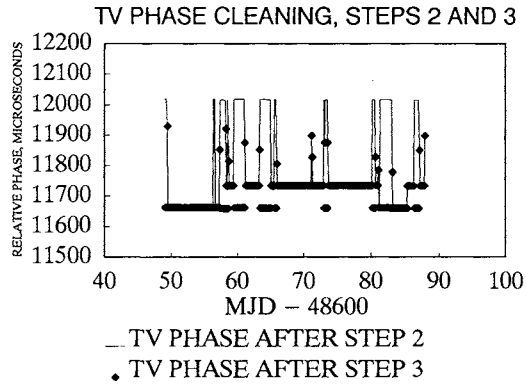


Figure 5.

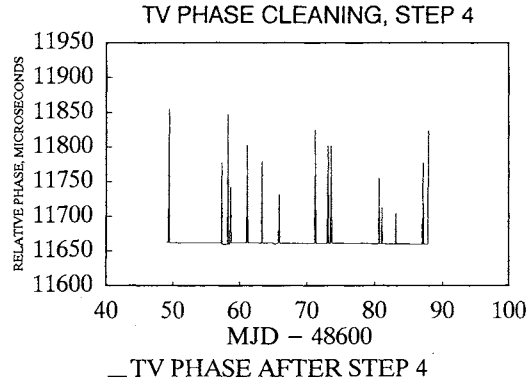


Figure 6.

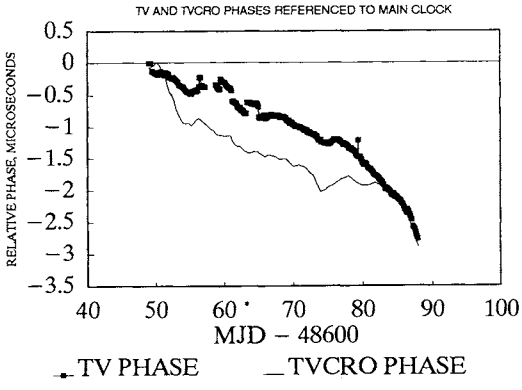


Figure 7.

Figure 7 shows the TVCRO output phase shifted to start from zero as well as the corresponding TV-phase. The maximum difference between TVCRO and TV -phases is 0.8 μs during the test period of 40 days.

5.2. Frequency results

Frequencies referenced to the Main Clock were calculated from the 2 h interval phase values. The frequencies are averaged to eliminate excessive short term noise, which effects Loran and TV phase signals.

Figure 8 shows the Loran, TVCRO and reference FRS-C oscillator averaged frequencies (24 h averaging, Loran 48 h averaging).

Figure 9 shows the TV -frequency and the TVCRO output frequency. The TV-frequency is calculated from the original phase data values and the graph obtained by removing the frequency values exceeding +/- 1exp-11. Before the start of the test run, the TVCRO rubidium frequency offset was in the order of 1.5*10exp-11 (without the TV-control).

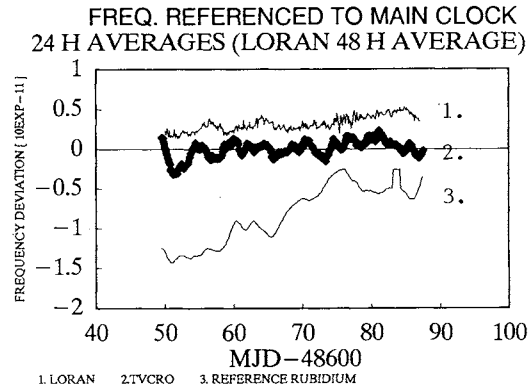


Figure 8.

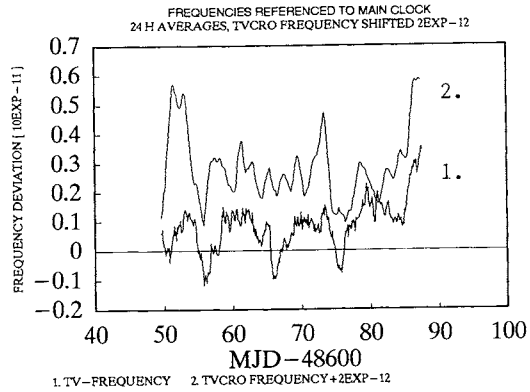


Figure 9.

5.3. TVCRO performance

Figure 10 shows the difference of the averaged TV and TVCRO frequencies with the maximum value of 5x10 exp-12 peak to peak. The standard deviation around the mean is 0.9x10exp-12.

For average frequencies, the calculated trends and standard deviations of residuals are shown in Table 1, linear fit, N=456.

Table 1.

	trend/day	Standard Deviation	Correl.Coef f.
Loran	6.8x10exp-14	0.5x10exp-12	0.80
FRS-C	3.0x10exp-13	1.5x10exp-12	0.91
TVCRO	-	1.1x10exp-12	0.07
TV	3.7x10exp-14	0.7x10exp-12	0.50

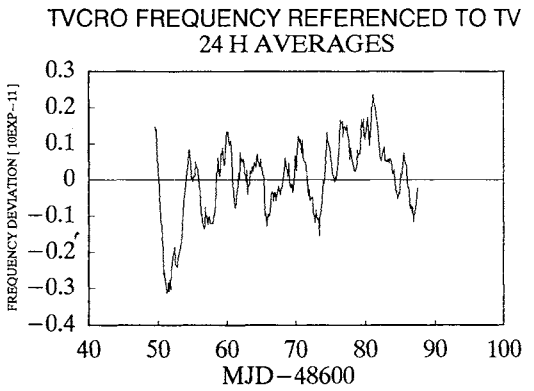


Figure 10.

Figure 11 contains the calculated effect of the monitored daily TVCRO control voltage. This is obtained by multiplying the control voltage value (during the test run values from 824 to 833) with the measured coefficient 1.1exp-12/ control voltage bit. Inverse of this curve approximates, how the frequency of the TVCRO rubidium oscillator would have drifted without the TV - control. The corresponding phase drift would exceeded 8 μs.

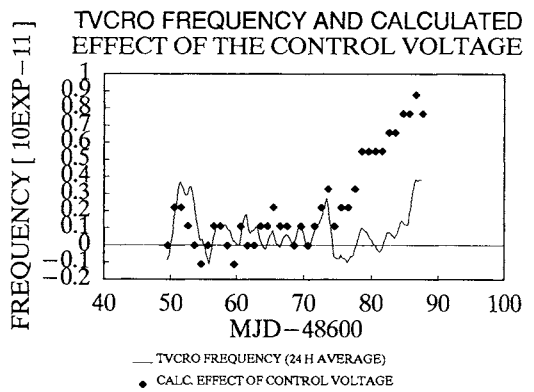


Figure 11.

6. DISCUSSION

In some applications, the TV -network offers important features as a reference frequency distributing system. The signal is easily received and the distributing network is independent and can be nationally controlled.

The TV -frame synchronizing signal of the Finnish TV-network is suitable for automatic controlling of remote rubidium oscillators. The realization of a TVCRO (TV Controlled Rubidium Oscillator) must include software algorithms for handling the special problems introduced by the TV-reference signal control.

The daily stability of TVCRO during the 40 day test run was 5*10exp-12 peak to peak with a standard deviation of 0.9*10exp-12.The phase drift referenced to the TV-phase during the test run was within 0.8 μs.

The long term frequency behaviour indicates, that a frequency stability of a few parts in 10exp-13 per year and a timekeeping ability of 10 μs per year could be expected, if referenced with the TV -transmission center frequency.

The TV -transmission center frame synchronizing frequency stability seems to be the most important factor for the TVCRO long term stability. For best performance, the TV transmission center should be synchronized so, that the long term drift would be eliminated.

7. REFERENCES

1. Tolman J & al 1967, Microsecond Clock Comparison by Means of TV Synchronizing Pulses, IEEE Trans on Instrumentation and Measurement, Vol IM 16, 3, 247-254.

2. Parcelier P 1970, Time Synchronization by Television, IEEE Trans on Instrumentation and Measurement, Vol IM 19, 4, 233-238.

3. Howe D 1972, Nationwide Precise Time and Frequency Distribution Utilizing an Active Code Within Network Television Broadcasts, IEEE Trans on Instrumentation and Measurement, Vol IM 21, 3, 263-276.

4. Leschiutta S 1973, Long-Term Accuracy of Time Comparisons Via TV Signal Along Radio Relay Links, IEEE Trans on Instrumentation and Measurement, Vol IM 22, 1, 84-87.

5. Borisochkin V & al 1981, Centralized Synchronization of the Moscow TV Center: A System for Transmitting Standard Frequency and Time Signals, Measurement Techniques 24 (-81)2, 110-113.

6. Mannermaa J 1990, The Phase Stability of the Synchronization Pulse in a 600km TV Link System, The European Frequency and Time Forum, Neuchatel 1990, 237-238.

7. Mannermaa J& al 1989, On The Stability of TV Transmission delay between Oulu-Helsinki, Den 12. Konferanse Måleteknik og Kalibrering Helsinki 1989.

8. Mannermaa J & al 1989, Studies on The Stability of Frequency Standards Using an Automatic Measuring System, Acta Universitatis Ouluensis C51, Oulu, Finland.

9. Howe D & al 1981, Properties of Signal Sources and Measurement Methods, Proceedings of the 35th Annual Symposium on Frequency Control, 14-60.

EXPERIMENTS ON A DIODE-LASER-PUMPED ^{87}Rb MASER

Deng Jinquan, Tan Yongfang, Xiang Yuanhai and Zhu Xiwen

Laboratory of Magnetic Resonance and Atomic and Molecular Physics,
Wuhan Institute of Physics, Academia Sinica, Wuhan 430071, P.R.China

ABSTRACT

We succeeded in a demonstration of self-sustained oscillations of a diode-laser-pumped ^{87}Rb maser and measured the output powers of the maser in various pumping light intensities and cell temperatures. The largest output power obtained with the diode laser power of 3.6 mW was 1×10^{-10} W at the optimal cell temperature of 58°C.

Keywords: Rb maser, laser pumping.

1. INTRODUCTION

^{87}Rb maser produces a coherent microwave signal at about 6835 MHz corresponding to the 0--0 transition between two hyperfine multiplets (F=1, 2) of the ground state of ^{87}Rb atoms[1]. The maser has outstanding short-term frequency stability when the averaging time is less than 1s[2], which is based mainly on its relatively high output power and its narrow resonance linewidth, compact size and been required for a variety of purposes. And improvements of its performances are essential for the progresses of these applications.

The short-term frequency stability of ^{87}Rb maser is determined by both the output power and the resonance linewidth[3]. It has been known from theoretical and experimental results[4,5] that the output power increased with optical pumping rate and atomic density in thin vapors. There is no experimental results on ^{87}Rb maser operated in thick vapors due to the pumping rate is limited by the weak pumping light from an rf discharge lamp. Recently, there have been considerable interest in and speculation on the use of single-mode diode lasers for atomic frequency standards with optical pumping. Besides the major efforts focused on Cs beam[6] and Rb gas cell[7] frequency standards with the single-mode diode laser, there has also been a growing interest in the area of the diode-laser-pumped ^{87}Rb maser[8-10]. In a theoretical analysis, much more output power was expected to gain from the maser pumped by one or two lasers[8] and some experimental results of the self-oscillating and the relative maser frequency pulling from the laser were presented when a diode laser was used for pumping a rubidium cell[10]. It was also predicted that the narrowed symmetric lineshape and frequency tunability of the laser might drastically reduce the light shift effect[11]. To our knowledge, so far there has been no experimental results to support these expectations and to provide insights into choice of the optimal pumping conditions.

In the present study, experiments were carried out to realize the self-sustained oscillation of the diode-laser-pumped ^{87}Rb maser. The experimental setup is described in Section 2. In Section 3, we observed the stimulated emission signal and measured the output powers of the maser under the conditions of different pumping laser intensities and cell temperatures. And a conclusion is presented in Section 4.

2. EXPERIMENTAL SETUP

Fig.1 shows an experimental setup for a laser-pumped ^{87}Rb maser, in which our laboratory prototype of the ^{87}Rb maser was used. An rf discharge lamp and a ^{85}Rb filter cell were replaced by a 780 nm AlGaAs diode laser (Hitachi HL7801E). We found that the laser wavelength was arrived at the ^{87}Rb--D_2 line when the laser diode's injection current and junction temperature were 78.22 mA and 30.7°C. And its corresponding laser output power was 3.6 mW.

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

As the emission from the diode laser is rapidly divergent, we measured the distribution of laser intensity in a plane which was perpendicular to the laser beam and in the distance of 29 cm from the laser, along two directions, i.e., vertical and parallel to the junction plane and obtained a near-Gaussian distribution. The laser intensity was 0.056 mW/cm^2 in the center, 0.033 mW/cm^2 and 0.023 mW/cm^2 respectively at the points of 5 cm apart from the centre along the vertical and parallel directions and the total incident power was 3.3 mW. Therefore, there was no need of beam-expanding for irradiating the whole resonance cell.

The junction temperature fluctuations of laser was reduced to a value as low as 10^{-3} °C at near 30°C by using a peltier element and a feedback loop. A low noise electronics powered by batteries was used to control the diode laser injection current and the fluctuations of the current was 10 μA. The frequency stability of this free-running laser was about 60 MHz which was sufficient enough for reproducible measurements of the output power of the maser.

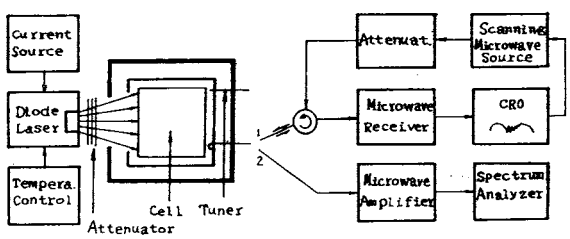


Fig.1: Experimental setup. Position 1 and 2 correspond to the arrangements used to observe the stimulated emission and to measure the output power respectively.

The resonance cell which contains an excess of ^{87}Rb metal and 11 Torr N_2 is located inside a microwave cavity tuned to the hyperfine resonance frequency of ^{87}Rb (This is the 0—0 transition which is insensitive to an applied magnetic field in the first order.). The cavity is of cylindrical cross section and operates in the mode TE_{021}. Resonance cell temperature is controlled by bifilar resistive heater glued onto the outside surface of the cavity. And the temperature is monitored by a platinum resistive sensor on the cavity, which is also used to servo-control the temperature. A temperature stability of better than 0.1°C is obtained over the typical measurement periods. The entire cavity assembly is positioned in a magnetic field of about 250 mG along the optical transmission axis.

3. EXPERIMENTS AND RESULTS

In order to know whether the condition of the maser self-sustained oscillations with the diode-laser-pumping was satisfied or not, we turned on the coupler of the cavity at position 1 in Fig.1 and did the experiments on observing the stimulated emission signal of ^{87}Rb. Due to the optical pumping, an over-population is built up in the F=2 state. When a microwave signal is scanned over the frequency of the 0—0 transition, ^{87}Rb atoms in the F=2 state will be stimulated and decayed into the F=1 state. The beat of this stimulated emission signal and the reflection signal of the scanning microwave by the cavity was fed to a microwave receiver. We observed the unilateral beat signal overlapped on an envelope of an intermediate frequency bandpass of the receiver and an absorption lineshape of the cavity As the diode laser intensity in the present experiments was adequate for the maser to oscillate, we observed the symmetric beat signals, i.e., the signals of self-sustained oscillations when the laser frequency, cavity frequency and cell temperature were set to the correct values.

The output powers of the diode-laser-pumped ^{87}Rb maser under the conditions of different laser intensities and cell temperatures were measured by using a low noise microwave amplifier and a spectrum analyzer in the case where the coupler was at position 2. The light intensity was varied by means of different pieces of film attenuators in the front of the light beam. The measurement error is ∓0.5 dbm and the data obtained are shown in Fig.2, from which the following can be drawn:

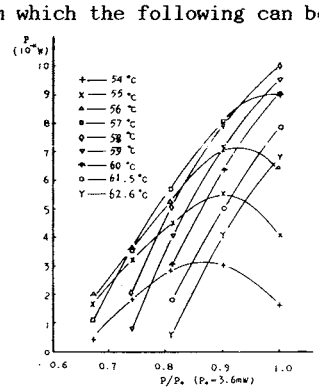


Fig.2: Variations of the output power vs. laser intensity under various cell temperatures.

1. The maser began to oscillate under certain light intensity and the output power increased first and

then decreased to stop oscillation with the increasing of the light intensity when the cell temperature was relatively low.

2. The light intensity under which the maser began to oscillate increased when the cell temperature increased. In those cases, we only observed that the output power increased monotonously with light intensity and the maximum output power was larger than that in the case of lower cell temperature.

3. Increasing the cell temperature continuously, we also observed the monotonic increase of the output power. But the output power decreased progressively at the same light intensity as the case of the lower cell temperature.

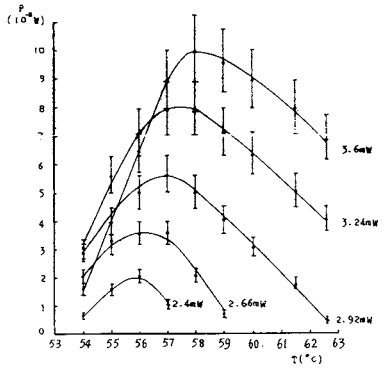


Fig.3: Variations of the output power vs. cell temperature under various laser intensities.

The data in Fig.2 could be redrawn in Fig.3, from which the output power for a given light intensity approached a maximum near a definite temperature and that both the maximum power and the temperature corresponding to that power increased with light intensity. These features were very similar to those in the rf-lamp-pumped ^{87}Rb maser[5] except that the output powers decreased much rapidly in the high temperature range for the latter case, presumably due to the decrease of the efficiency of the ^{85}Rb filter. The largest output power of the diode-laser-pumped ^{87}Rb maser was 1×10^{-10} W and approached the largest output power of an rf-lamp-pumped ^{87}Rb maser currently. Comparing with theoretical results[12], we found that the variations of the output power vs. light intensity and cell temperature stated above were the same as the calculated ones. In consideration of the beam divergence and the transmissivity of the transmitting rings of the cavity and cell walls, the actual light power acted upon rubidium atoms was about 2 mW. The theoretical value of the maximum output power corresponding to incident light power of 1.89 mW is 1.09×10^{-10} W[12], which was in good agreement with the experimental results.

4. CONCLUSION

In this note we reported the self-sustained oscillations of a diode- laser-pumped ^{87}Rb maser and measured the output power of the masers in the different pumping light intensities and cell temperatures. The largest output power obtained was 1×10^{-10} W, which might be increased further by using a diode laser with larger intensity and by optimizing the cell temperature. The investigations of the linewidth and the light shifts of the diode-laser-pumped ^{87}Rb maser with a frequency stabilized and tunable diode laser are under way.

REFERENCES

[1] Davidovits P and Novick R 1966, The optically pumped rubidium maser, *Proc. IEEE*, 54, 155-170.

[2] Tutu M et al 1985, Experimental results on a frequency standard based on a rubidium 87 maser, *Proc. 39th Annual Symposium on Frequency Control*, Philadelphia, 64-71.

[3] Tetu M et al 1973, Short-term frequency Stability of the ^{87}Rb maser, *IEEE Trans. on Instr. Meas.*, IM-22, 250-257.

[4] Tessier M and Vanier J 1971, Theorie du maser au rubidium 87, *Can.J.Phys.*, 49, 2680-2689.

[5] Zhu X, Xiang Y, Chen X, Deng J and Tan Y 1989, Study of the power, linewidth and frequency shifts of ^{87}Rb maser, *Proc. International Symposium on Electromagnetic Metrology*, Beijing, 297-300.

[6] Picque J L 1975, Hyperfine optical pumping of a cesium atomic beam, *Metrologia*, 13, 115-119.

[7] Hashimoto M and Ohisu M 1987, Experiments on a semiconductor laser pumped rubidium atomic clock, *IEEE J.Quan. Electr.*, QE-23, 446-451.

[8] Zhu X 1986, Some Aspects of the Theoretical analysis of a laser-pumped ^{87}Rb maser, *Proc. 2nd National Symposium on Precision Time*, Guangzhou, 101-103, (in Chinese).

[9] Deng J, Tan Y, Xiang Y and Zhu X 1988, Frequency-stabilized laser for atomic frequency standards using saturated absorption by magnetic-modulation, *Proc. 2nd European Frequency and Time Forum*, Neuchatel, 555.

[10] Michaud A et al 1991,Experimental study of the laser diode pumped rubidium maser, *IEEE Trans. on Instr and Meas.*, IM-40, 170-173.

[11] Vanier J et al 1981, Progress and prospects in rubidium frequency standards, *J. Phys.* (Paris) Colloq.Colloq., 42, Suppl.C8 139-150.

[12] Deng J and Zhu X, Theoretical Analysis of the Diode-Laser-Pumped ^{87}Rb Maser (to be published).

OPTICAL DETECTION OF LASER COOLED IONS IN A PENNING TRAP

Hidetsuka Imajo, Shinji Urabe, Kazuhiro Hayasaka, and Masayoshi Watanabe

Kansai Advanced Research Center, Communications Research Laboratory, Kobe, Hyogo JAPAN

ABSTRACT

A Penning trap is used to trap Be^+ ions, which are then cooled by a laser to approximately 300 mK. The eigenmotions of the cooled ions are measured optically by using the laser fluorescence technique. In addition, several kinds of ions are trapped with Be^+ ions and their motional frequencies are measured. A small number of the cooled Be^+ ions are trapped and their fluorescence is also observed.

Keywords: Penning trap, laser cooling, axial frequency, modified cyclotron frequency, magnetron frequency

1. Introduction

Ion traps have been utilized for various applications, including high-resolution spectroscopy, frequency standards and mass spectrometry. Several ion traps have been used. Among them, the Penning trap is advantageous to laser cooling of ions due to the absence of rf heating. Therefore, it is possible to cool more ions in a Penning trap than in a Paul trap. Some kinds of ions have been laser-cooled, however, the number of species that can be laser-cooled directly is limited by the availability of tunable lasers. Be^+ ions are appropriate for laser cooling experiments due to their simple energy levels and because their light source is easily obtained.

It is attractive to apply a Penning trap to frequency standards because laser cooling of ions reduces Doppler effects and long confinement times provide long interaction times between ions and microwave. In addition, sympathetic cooling [1] in a Penning trap provides a means of cooling ion species which cannot be laser cooled directly. It is also interesting to confine a single ion in a Penning trap, because quantum jumps [2] of an ion can be observed and the accuracy of spectroscopic measurements is improved.

In this paper, measurements of laser-cooled Be^+ ions in a Penning trap are reported. Laser-cooling experiments are performed and the motional frequencies of the cooled ions are optically measured. In addition, as a preliminary experiment, trapping and cooling of a small number of ions are demonstrated, and the fluorescence is observed.

2. Experimental arrangement

The construction of a Penning trap is described in detail in [3]. The electrodes are hyperboloids, with parameters of $z_0=r_0/\sqrt{2}=8$ mm. These electrodes are placed in a vacuum chamber that has a pressure of about 10^{-10} Torr. A Be atomic beam is produced by a small oven and is ionized inside the trap by an electron beam emitted from a hot filament. The static magnetic field is created by an electromagnet.

The cooling transition of Be^+ ions is at 313 nm and this light source is obtained by frequency doubling the output of a dye laser. The UV laser beam (power approximately 1 mW) is perpendicular to the magnetic field direction, focused to about a 0.5 mm-diameter at the center of the trap. Fluorescence from the ions is detected by a photomultiplier tube in

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

photon counting mode.

A small rf voltage is applied between the endcaps and the ring electrode to measure the ion motional frequencies. The rf frequency is scanned and, simultaneously, the fluorescence characteristic changes at ion motional resonances are observed.

3. Laser cooling

The efficiency of laser cooling in a Penning trap depends on the position of the laser beam and the laser frequency. [4] Typically the laser frequency is set at 1 GHz below the resonance. Once the laser position and frequency are set and Be^+ ions are produced, fluorescence from the hot ions is observed. This fluorescence signal rises slowly as the ions cool. When the scatter rate becomes constant, the frequency of the laser is scanned to determine the line profile of the cooling transition shown in Fig. 1(a). This profile is asymmetric due to heating when the laser is tuned to the high-frequency side of the transition.

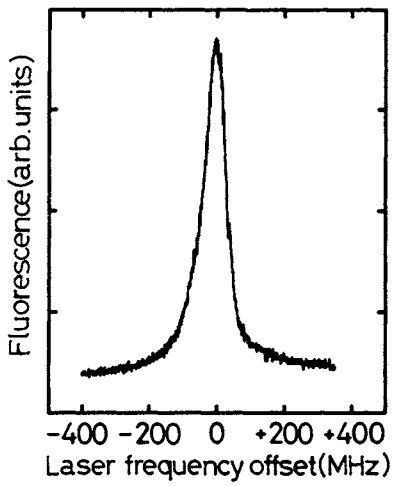


Fig. 1(a). Observed line profile of cooled Be^+ ions when the frequency of the cooling laser is scanned at the rate of 200 MHz/s.

Figure 1(b) shows the line profile when the scan rate of the cooling laser is reduced. In this case, the fluorescence signal suddenly decreases when the laser frequency is at the high-frequency side of the transition due to the diffusion of the ion cloud. The linewidth in Fig. 1(a) corresponds to an ion temperature on the order of 300 mK. To measure ion temperatures more precisely, however, a second probe laser and a smaller laser beam are necessary.

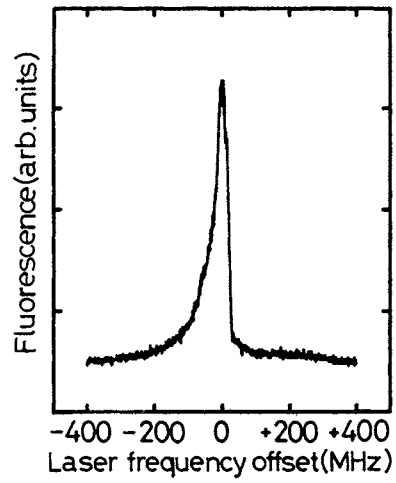


Fig. 1(b). Observed line profile of cooled Be^+ ions when the frequency of the cooling laser is scanned at the rate of 8 MHz/s

4. Measurement of eigenmotions

The eigenmotions of ions in a Penning trap are composed of a harmonic oscillation along the magnetic direction and a superposition of circular motions in the plane perpendicular to the magnetic field direction. The frequencies of these motions are the axial frequency, the modified cyclotron frequency, and the magnetron frequency. These frequencies can be measured by using the laser fluorescence technique.[5] The motions are excited if a small rf voltage is applied between the endcaps and the ring electrode while the ions are being cooled. When the frequency of the applied rf voltage is scanned, the observed fluorescence of the cooled ions changes at their resonances because the velocity distribution of the ions changes. The characteristic changes depend on the detuned frequency of the cooling laser. Using this technique, the frequencies of the eigenmotions are measured as shown in Figs. 2-4. The measured frequencies agree well with the theoretical ones. Figure 5 shows the observed fluorescence when two other kinds of ions are trapped with the Be^+ ions. Signals 1 and 2 correspond to their cyclotron resonances. These ions are heated by the applied rf voltage and they sympathetically heat Be^+ ions by Coulomb interactions. Possible ion candidates in Fig. 5 include a background neutral molecule of H_2O and CO emitted from the filament. These signals can be eliminated by lowering the acceleration

voltage of the electron beam. This technique enables a Penning trap to operate as a mass spectrometer.

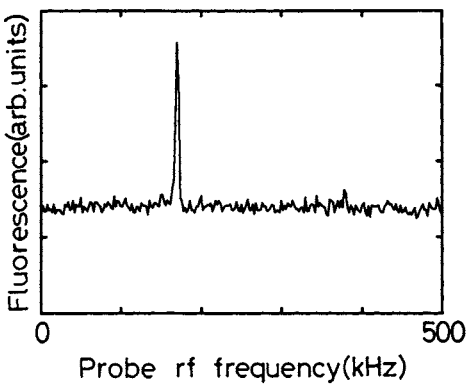


Fig. 2. Observed axial resonance. Trapping voltage: U=8.0 V ; magnetic field :B=0.95 T.

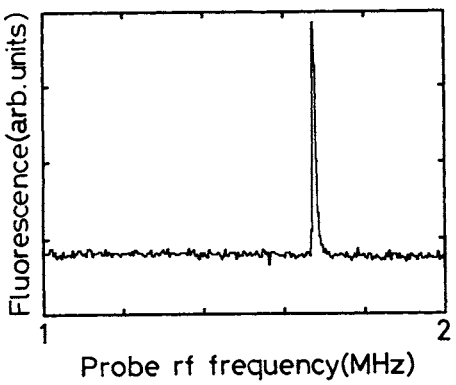


Fig. 3. Observed cyclotron resonance. Trapping voltage: U=8.0 V ; magnetic field :B=0.97 T.

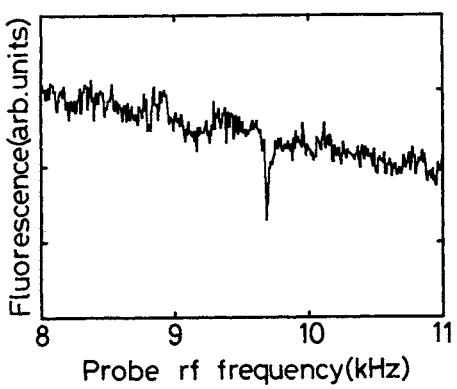


Fig. 4. Observed magnetron resonance. Trapping voltage :U=8.0 V ; magnetic field :B=1.03 T.

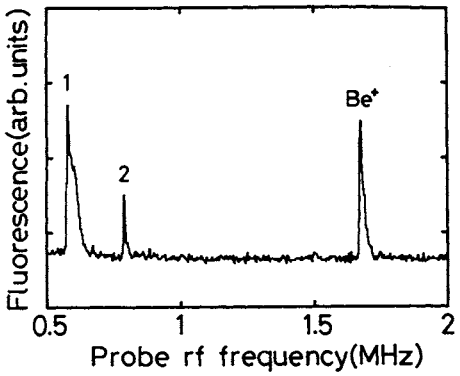


Fig. 5. Observed cyclotron resonances. Two other kinds of ions are trapped with the Be^+ ions (indicated as No.1 and 2).

5. Observation of fluorescence from a small number of cooled ions

To produce a small number of ions, the atomic oven is operated at a lower temperature and the acceleration voltage of the electron beam is applied for a few seconds. The cooling laser frequency is detuned approximately -30 MHz from the resonance. At first, the fluorescence from the hot ions is too weak to be observed, but after a while the signal suddenly rises due to the cooling shown in Fig. 6. It is estimated that the number of trapped ions is less than 10 based on the fluorescence intensity and the detection efficiency. It is expected

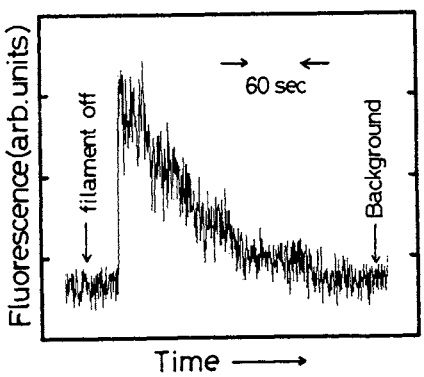


Fig. 6. Fluorescence from a small number of cooled Be^+ ions. Acceleration voltage of the electron beam is applied for 2 s. After approximately 30 s of ion production, the fluorescence signal suddenly increases due to cooling. Cooling laser detuning is approximately -30 MHz.

that if the number of cooled ions decreases one by one, the observed signal decays discretely. In Fig. 6, however, a discrete signal is not clearly apparent. One possible reason is because the direction of the cooling laser is perpendicular to the magnetic field direction and the axial motion of ions cannot be laser-cooled directly. When a large number of ions is trapped, the axial motion is cooled due to Coulomb collisions. For a single ion, however, it will be necessary to tilt the cooling laser beam to the magnetic field direction.

6. Summary

Be^+ ions were stored in a Penning trap and laser-cooled to approximately 300 mK using one laser. The frequencies of the ion eigenmotions, i.e., the axial frequency, the modified cyclotron frequency, and the magnetron frequency were measured by using the laser fluorescence technique. The measured frequencies agree well with the theoretical ones. Trapping and cooling of a small number of ions were demonstrated and the fluorescence was observed. To be able to clearly observe the fluorescence from a single ion, it will be necessary to improve the detection efficiency of scattered photons and to change the cooling laser beam direction.

References

[1] D.J. Larson, J.C. Bergquist, J.J. Bollinger, W.M. Itano, and D.J. Wineland, Phys. Rev. Lett. 57, 70 (1986).

[2] Th. Sauter, W. Neuhauser, R. Blatt, and P.E. Toschek, Phys. Rev. Lett. 57, 1696 (1986).

[3] H.Imajo, S. Urabe, K. Hayasaka, and M. Watanabe, J. Mod. Opt., to be published.

[4] W.M. Itano and D.J. Wineland, Phys. Rev. A25, 35 (1982).

[5] D.J. Wineland, J.J. Bollinger, and W.M. Itano, Phys. Rev. Lett. 50, 628 (1983).

QUARTZ CRYSTAL TWINNING UNDER MECHANICAL STRESS : FINITE ELEMENT ANALYSIS

J.J. Boy*, B. Dulmet*, P. Conraux**

* LCEP / ENSMM - Route de Gray - La Bouloie - 25030 Besançon Cedex - FRANCE
** Framasoft+CSI - 10 rue Juliette Récamier, BP 3083 - 69398 Lyon Cedex 03 - FRANCE

ABSTRACT

We propose some tools to describe the switching process of electrical twinning in various cuts of quartz. In all cases, we apply the principle of minimal energy, together with a threshold condition, to the proper thermodynamic potential to determine which twin will survive the exerted stresses on a quartz sample. We give an analytical expression useful for samples submitted to a known homogeneous compressive stress. For more general stress states, a finite elements approach is chosen to study the propagation of the switching process. It is founded upon a bielastic stress-strain relationship introduced into the plasticity model for non-linear transient analysis in SYSTUS® program.
Keywords: Quartz, bielasticity, electrical twinning, SC-cut, F.E.M., non-linear transient analysis.

1. INTRODUCTION

Sometimes, during cutting or polishing process performed into the current BVA resonators production at L.C.E.P., twins appear on the surface of the blanks. Even if the twin area is very confined, the performances of the quartz resonators can be degraded. Unfortunately, this electrical twinning can be revealed only during the frequency adjustement process when the shape definition of the resonator is already achieved. Though the manufacturing process is settled, local overheatings induced by, for example, the blade of the sawing machine create sometimes high level of stresses which can generate ferrobielastic twinning. Numerous parameters are involved in the twin threshold : temperature, value and direction of the applied force and/or stress, thermal gradients and material quality, ...

This paper does not concern the variations of the "coercive stress" value versus material quality translated in terms of the number of dislocations or impurities contents. Instead, we stick to introducing a mean threshold value of the difference of the proper thermodynamic potential in the twin materials. This potential is either Gibb's free energy or the stored internal energy, depending on the experimental conditions. The threshold value is considered as a data which we guess from previous experiments shortly described below. Then, starting from classical thermodynamic consideration, we calculate the pertinent energy difference for a sample of any orientation, submitted to an homogeneous compressive stress. In this way, we can analytically setup some *twinning* curves for the doubly-rotated SC-cut.

The results simply predicted in this manner seem to be confirmed by a more refined finite elements analysis that became possible after a bielastic behavior was introduced at Framasoft in the plasticity model of SYSTUS® (1) program. The F.E. analysis predicts that the propagation of twins should be far more progressive in a complex structure such as a BVA resonator, than in a parallelepipedic sample submitted to homogeneous stresses.

2. PREVIOUS EXPERIMENTS : A SHORT REVIEW

In contrary of the optical or Brazil twin, electrical or Dauphiné twinning occurs only when Si-O atomic groups are slightly shifted through distances of 0.033 nm for Si atoms and 0.081 for these of Oxygen. The interatomic bonds are not broken and the two possible states so induced can be brought in coïncidence if one of them is rotated through 180 deg. about c- or Z-axis, as shown in figure 1 (2).

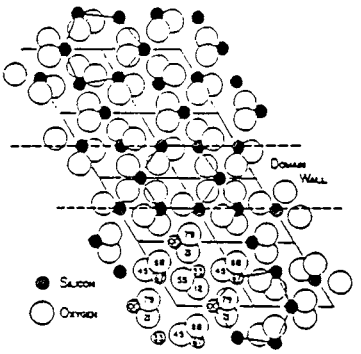


Figure 1

The coercive stress for ferrobielastic twinning in quartz has been measured by T.L. Anderson and J.W. Laughner or E. Bertagnolli et al. on X-cut quartz samples (3), (4), (5). In figure 2, we show the sample and stress orientation with respect to the crystallographic axes. Typical samples dimensions are 5*10*10 cm^3 with the thickness parallel to X-axis (which is labeled Xc, Ys and Zs being the axes-system of the sample) :

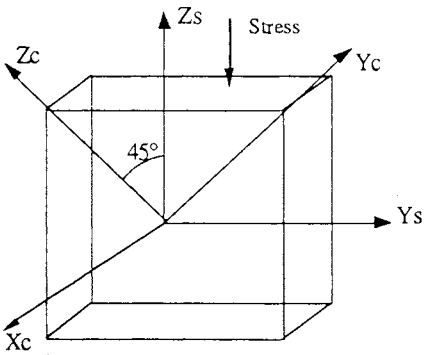


Figure 2

The authors have shown that the stress needed to ferrobielastically twin quartz is on the order of 10^8 N/m^2 (around 3 to 5 Kbar at room temperature). Nevertheless, this

threshold depends on various factors like :

1. Sample origin : Synthetic quartz has a higher "coercive" stress than natural quartz :

In virgin synthetic quartz crystal, twins first appear at stress level at about $5*10^8$ N/m^2 (at room temperature) and the switchover occurs at $7*10^8$ N/m^2. In natural quartz, twin appears at $3*10^8$ N/m^2 and the stress required for switchover is about $4*10^8$ N/m^2 (4 Kbar).

Furthermore, when the pressure is slowly released from a synthetic quartz sample, the crystal spontaneously returns to its original twin state. The process begins at approximately 2 to $3*10^7$ N/m^2 and continues as stress is removed. On the other hand, one saids that natural quartz crystal does not backswitch spontaneously.

2. Temperature : "Coercive stress" decreases when temperature is increased (Ref. 6 and Fig. 3) :

This property cannot be understood only by the temperature variation of compliance coefficients which takes place in the "Gibbs free-energy function". But it is not influenced by structural defects and it is linked to atomic thermal vibrations which increase with temperature. Indeed, when heated, the α-quartz structure looks more closely like this of β-quartz and so the required reorientation force becomes smaller since electrical twinning does not exist in β-quartz.

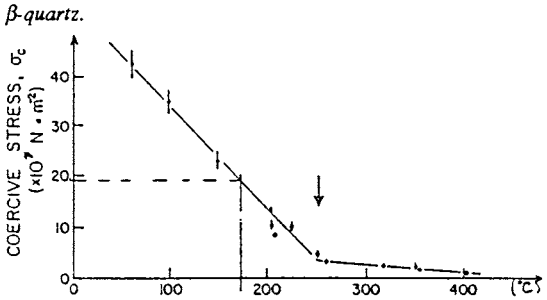


Figure 3 : Coercive stress vs temperature

3. Sample history : "Coercive stress" decreases when a sample is repeatedly twinned. :

Since synthetic crystals revert to the original twin state when the stress is removed, the experiment concerning the effect of repeated switching on twinning behavior is easily repeated. Anderson reports that the first twin appears at $4.4*10^7$ N/m^2 when stress is reapplied, compared to a stress of $5*10^8$ N/m^2 for the first twin seen during the initial stress application. Switchover occurs at about $2*10^8$ N/m^2. Switchback stress is also reduced. The reduction in coercive stress with cycling is not explained by properties of the perfect crystal. Defects are important to this reduction. Defects were seen to hinder ferrobielastic switching. Twin walls can, for instance, change the bias by moving impurities within dislocations.

4. Stressing rate : "Coercive stress" increases when the stressing rate is increased. Stress at switchback decreases when the stress removal rate is increased. :

For example, varying the stress-release rate following switchover changes the twin pattern caused by switchback. When the twinning stress is held slightly below the coercive stress for hours or days, twin growth is seen.

5. Twin walls : Brazil twin walls and accumulated Dauphiné walls inhibit Dauphiné walls motions, thereby increasing coercive stress :

When samples contain Dauphine twins, wall motion occurred at stresses only 10 or 20 % of those needed to produce twin walls in untwinned quartz material. Switchover was a minor event, not so abrupt than in a virgin example. The appearance of ferrobielastic domains can be classified into three categories :

- *Stripe-shaped twins oriented parallel to Z*
- *Wedge-shaped twins oriented parallel to Z, with their width parallel to Y.*
- *Diffuse, curved-wall twins with volumes approching the entire crystal volume as stress is increased.*

The twin shape changes under various conditions. The highest coercive stresses are measured when the twin are stripe-shaped. The lowest coercive stresses are measured when the twins are diffuse. Stripe-shaped twins occupy the least volume and occur only on the X-face of the sample. Diffuse twins occupy the greatest volume, since they slowly grow until the crystal is completely twinned.

3. A SIMPLE ANALYTICAL MODEL

The first law of Thermodynamics is written as :

$$dU = dW + dQ$$

where :
dW and dQ are infinitesimal quantities of *work* and *heat*, dU being the differential form of the internal energy. All quantities are taken per unit volume.

In case of a linear, polarisable, elastic and non-magnetic solid, the internal energy can be written :

$$dU = T{:}dS + E.dP + \theta.d\eta$$

where : θ is the *temperature*, and the three independant variables are :

S : Strain, P : Polarisation, η : Entropy.

If the stress, the electric field and the temperature are the independant variables, instead of strain, polarization and entropy, *Gibbs' free energy function* G must be chosen as state function, instead of U :

$$G = U - T{:}S - P.E - \theta.\eta$$

The differential form is the following :

$$dG = -S{:}dT - P.dE - \eta.d\theta$$

Starting from an initial free state of energy G_0, we perform the following expension of G, restricted to the terms needed for a linear theory :

$$G - G_0 = \frac{1}{2} s_{ijkl}.T_{ij}.T_{kl} - \frac{1}{2}\chi_{ij}.E_i.E_j$$
$$-\frac{1}{2}.C.\left(\theta - \theta_0\right)^2 - d_{kij}.T_{ij}.E_k - a_{ij}.T_{ij}.\theta$$

whichfor the coefficients appear in the following constitutive equations :

$$S_{ij} = s_{ijkl}.T_{kl} + d_{kij}.E_k + a_{ij}.\Delta\theta$$
$$P_i = \chi_{ij}.E_j + d_{ijk}.T_{jk}$$
$$\eta = C.\Delta\theta + a_{ij}.T_{ij}$$

When the conditions of experiment are pseudo-static and isotherm, we can restrict the expansion to :

$$G - G_0 = \frac{1}{2} s_{ijkl}.T_{ij}.T_{kl} - d_{kij}.T_{ij}.E_k - \frac{1}{2}\chi_{ij}.E_i.E_j$$

Due to the symmetries of the quartz crystal, the existence of an electrical twin involves domains the X-axis of which is inverted although the domain actually is made of the same quartz material : only its orientation is changed. So, if we have only one reference frame for the entire block, the two areas, separated by a twin wall, will be characterized by

opposite signs on s_{14} for the elastic compliances and on d_{11} for the piezoelectric coefficients.

Once the cartesian system is chosen, we can arbitrarily decide that the untwin area (called α_1) has its crystallographic X-axis coincident with the X1 of the working frame. So, s_{14} will be negative in the α_1 region for a right handed quartz. Consequently, the twin area α_2 will exhibit a positive value for s_{14}. For either material the developped series expansion of G is the following :

$$G_{1/2} - G_0 = -\frac{1}{2}.\left\{ s_{11}.\left(T_1^2+T_2^2\right)+s_{33}.T_3^2 \right\} - 2.\left\{ s_{44}.\left(T_4^2+T_5^2\right)+s_{66}.T_6^2 \right\} - \left\{ s_{12}.T_1.T_2+s_{13}.\left(T_1.T_3+T_2.T_3\right) \right\} - 2.s_{14}.\left(T_1.T_4-T_2.T_4+4.T_5.T_6\right)$$

If s_{14} denotes the value for the $\alpha 1$ material, the energy difference between the twin configurations is easily deduced from the previous formula :

$$G_2 - G_1 = 4.s_{14}.\left(T_1.T_4-T_2.T_4+4.T_5.T_6\right) + 2.d_{11}.\left(E_1.T_1-E_1.T_2-4.E_2.T_6\right)$$

We consider a sample submitted to an uniaxial homogeneous compressive (or tensile, indeed) stress P, acting along some direction (non specific). In the convenient doubly-rotated frame, we have :

$$T_{X'Y'Z} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -P \end{bmatrix}$$

Or :

$$T_{X'Y'Z} = -P.\left(0,0,1,0,0,0\right)$$

So, we have just to apply the classical tensorial rotation to express the energy difference in the initial reference frame (crystallographic).

$$[T'] = [K][T][K]^{-1}$$

The K matrix results from the successive compositions of rotations Ψ, θ and φ in this order :

$$X'Y''Z'' \xrightarrow{K_{\psi/Y''}} X'Y''Z'$$

$$X'Y''Z' \xrightarrow{K_{\theta/X'}} X'Y'Z$$

$$XY'Z \xrightarrow{K_{\varphi/Z}} XYZ$$

where :

$$K_\psi = \begin{bmatrix} \cos\psi & 0 & \sin\psi \\ 0 & 1 & 0 \\ -\sin\psi & 0 & \cos\psi \end{bmatrix} \quad K_\theta = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{bmatrix}$$

$$K_\varphi = \begin{bmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Using usual tensorial rotations, we obtained the following expressions for the stresses (which must be multiplied by P):

$$T_1 = \sin^2\psi.\cos^2\varphi + \cos^2\psi.\sin^2\varphi.\sin^2\theta + 2.\sin\psi.\cos\psi.\sin\theta.\cos\varphi.\sin\varphi$$

$$T_2 = \sin^2\psi.\sin^2\varphi + \cos^2\psi.\cos^2\varphi.\sin^2\theta - 2.\sin\psi.\cos\psi.\sin\theta.\cos\varphi.\sin\varphi$$

$$T_3 = \cos^2\theta.\cos^2\psi$$

$$T_4 = \cos^2\psi.\sin\theta.\cos\theta.\cos\varphi - \sin\psi.\cos\psi.\cos\theta.\sin\varphi$$

$$T_5 = \cos^2\psi.\sin\theta.\cos\theta.\sin\varphi + \sin\psi.\cos\psi.\cos\theta.\cos\varphi$$

$$T_6 = -\sin^2\psi.\cos\varphi.\sin\varphi + \cos^2\psi.\sin^2\theta.\cos\varphi.\sin\varphi + \sin\psi.\cos\psi.\sin\theta.\left(\cos^2\varphi - \sin^2\varphi\right)$$

So, we eventually obtain the energy difference which determines the possibility of switching :

$$\Delta G = 4.s_{14}P^2.\cos\psi.\cos\theta.\left[\cos\psi.\sin\theta.\cos\varphi - \sin\psi.\sin\varphi\right].\left[3.\sin 2\varphi.\sin\theta.\sin 2\psi + \sin^2\psi.\left(6.\cos^2\varphi - 1\right) + \sin^2\theta.\cos^2\psi.\left(6.\sin^2\varphi - 1\right)\right]$$

Applying the principle of minimal energy together with a threshold condition, we predict the occurence of switching in domains such that :

$$\Delta G < -A$$

From the currently available data in the literature, we found reasonable to choose the threshold value at A = - 0.2 MPa.

The following figures (Figures 4a, 4b and 4c) are polar plots of the switching stress combination ΔG in terms of the angle ψ. The solid lines indicate directions stabilizing α_2 Dauphiné twin state, obtained when ($G_2 - G_1$) is negative. Consequently, dashed lines define areas whichfor stresses directed along ψ will be ineffective. The azimuth is counterclockwise counted from the Z'-axis, leftwards oriented on the drawings. The directions 35, 45, 60, and 75 degs, with $\Delta G<0$ are indicated on the figures.

For an X-cut quartz blank, the angular dependence of ferrobielastic switching stress is obtained for $\varphi = 90°$, $\theta = 0°$ and expressed by :

$$\Delta G = 4.s_{14}.P^2.\sin^3\psi.\cos\psi$$

plotted in fig. 4a.

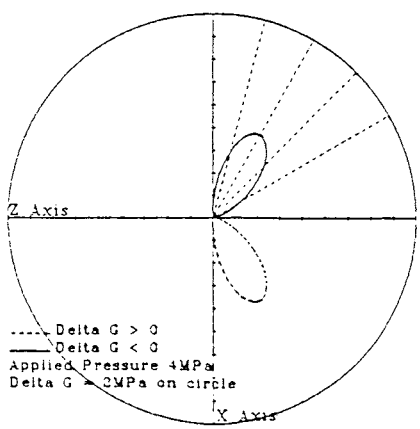


Figure 4a : X-cut

The figures 4b and 4c present the sames curves obtained for AT-cut ($\varphi = 0$, $\theta = 35°$) and SC-cut ($\varphi = 21.93°$, $\theta = 32.93°$) samples.

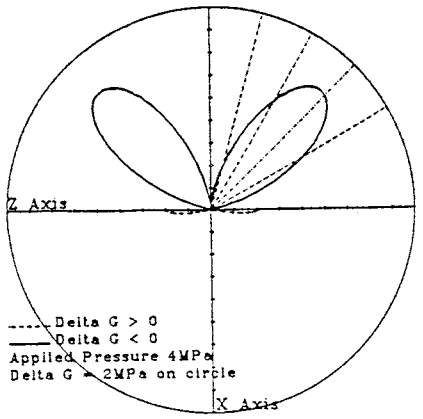


Figure 4b : AT-cut

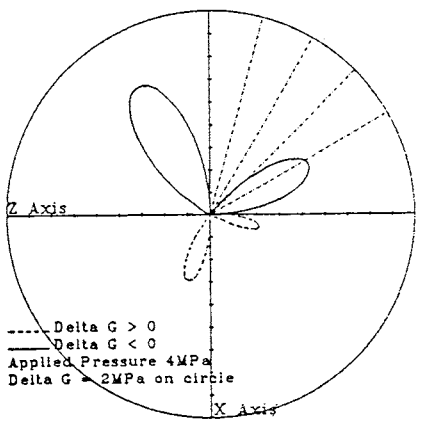


Figure 4c : SC-cut

4. FINITE ELEMENTS MODEL

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 condition 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. But whenever the stresses are not homogeneous and the switching propagates slowly, it should be clear that the pieces of linear behavior of the whole structure are extremely small. Therefore, the succession of linear analyses should be done in such a way that a minimum amount of elements are switched before each run. Thus, numerous runs of the software would be needed, implying the storage of huge amounts of data. In addition, the influence of nonlinearities intrinsic to the deformation is unattainable in this way.

For all these reasons, it is largely better to use an integrated non linear method, despite of the linear material behavior of each element inside the threshold bounds. So a bielastic model was added to plasticity models currently available in SYSTUS program for non linear static transient analysis. The stiffness matrix for linear analysis is computed altogether with the tangent matrixes needed to incrementally build a solution for a given loading. A stress relaxation arises at switching time : it ensures that energy is globally minimal in the structure. Therefore, because FEM is definitely a *displacement method* (7), convergency cannot be granted if the computation is performed at imposed stress. So, we had to impose the displacement in the transient analysis. As a consequence, the thermodynamic potential used in the method could not be Gibbs' free energy, but was the more commonly encountered stored internal energy $U = 1/2\ T : S$, instead.

Because they are not aimed at the same conditions of experiments, the results provided by such an analysis and those predicted by the analytical expressions of the previous section have not to be identical, although a rough agreement can be expected about the possibility of twinning in a given sample.

5. RESULTS

The first example, illustrated by the figure 5, concerns the X-cut sample described above. The chosen dimensions are: 3 * 6 * 6 mm^3 with the thickness parallel to X-axis. The sample being ideally symmetrical, we note that the switchover is instantaneous : the propagation of the twinning process is so fast that the transient analysis converges in a single step to a fully switched state.

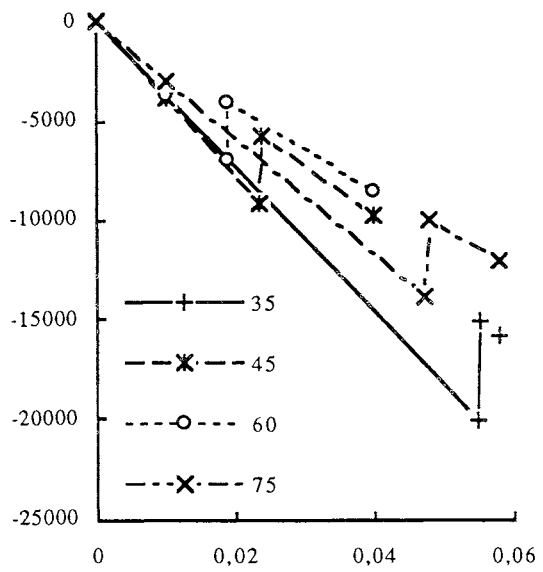


Figure 5 : force (in N) versus displacement during switching process.

The dependency of the switchover threshold versus the azimuth of the applied force roughly agrees with the behavior predicted in Sec. 3.

For $\psi = 35$ deg., the calculated coercive stress is about 1,100 MPa. The *compressive* breaking stress being close to 2,500 MPa, electrical twinning can effectively be observed. It is worth noticing that this will not be true in a sample submitted to tensile stresses, whichfor break will happen at only 100 MPa.

To demonstrate the propagation process in a complex structure, we simulated a SC-cut 4 bridges BVA resonator submitted to a radial imposed displacement oriented along the direction of one pair of bridges :

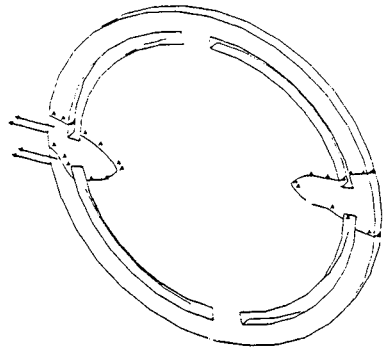


Figure 6a : u=0.07mm, ψ = 30 degs

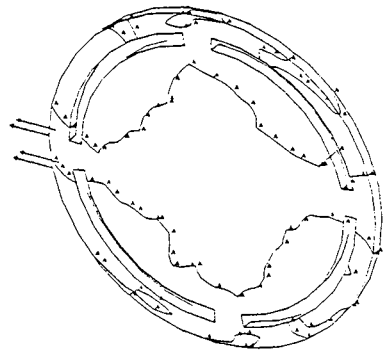


Figure 6b : u=0.33 mm, ψ = 30 degs

According to Fig. 4c, the ψ = 30 and ψ = 120 orientations are expected to facilitate twinning. Transient analyses performed with a regular mesh (706 nodes, 324 linear interpolation elements, 8 integration points per element) predicted a rather progressive twinning versus the imposed displacement u. Figures 6a and 6b show the predicted twin areas for u = 0.07 and u = 0.33 mm respectively, and for the ψ=30 bridges orientation. The ψ=120 orientation provides with similar results but the twinning propagates faster in this case. This is confirmed by fig.7 that gives the number of twinned integration domains versus the imposed displacement for both ψ values. The progressive twinning results in a progressive evolution of the slope of the global force-displacement curves as shown in fig.8. It should be observed that here we did not take care of the breaking stress. The stress diagram obtained during our analyses showed that the twinning process should actually start before breaking occurs, provided the imposed displacement be compressive.

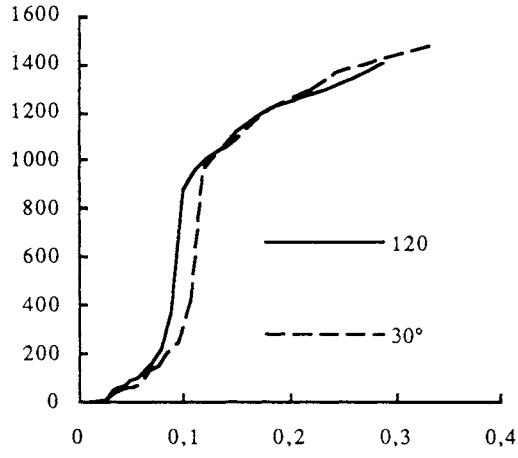


Figure 7 : Twins vs displacement

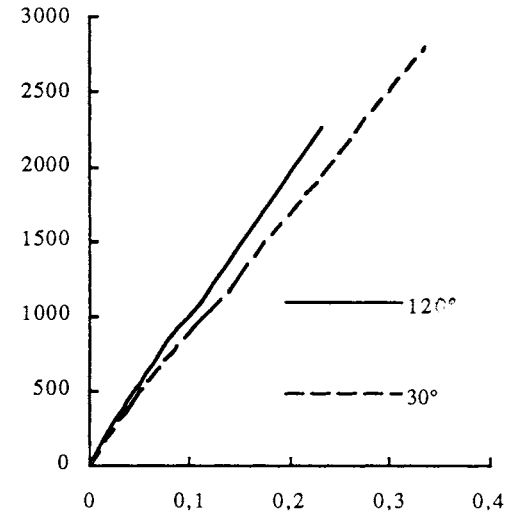


Figure 8 : Reaction force vs disp.

CONCLUSIONS

a) Because twinning can be strongly affected by local stress variations, complementary experiments would require a careful consideration of the mounting system of the samples. Such experimental checkings should be done prior to extending the present features of the software itself.

b) During the pre-lapping stages of its processing, a resonator is subjected to numerous damaging factors, the origin of which is not solely mechanical. For instance, local heating generated by a saw blade or a grinding stone is worth examining. Since our software includes standard thermal loading capabilities, the present work could be extended by taking advantage of them.

NOTES AND REFERENCES

1. SYSTUS is a product of Framasoft+CSI company.

2. T.L. ANDERSON, R.E. NEWNHAM, L.E. CROSS : "Coercive stress for ferrobielastic twinning in quartz", 30th Annual Freq. Control Symp. (1976), pp. 171, 177

3. K.L. BLISNUK, J.G. GUALTIERI, R.A. MURRAY : "A comparison of laser induced twinning of quartz based on characteristics and suppliers", 44th Annual Freq. Control Symp. (1990), pp. 251, 258

4 T.W. CLINE, J.W. LAUGHNER,R.E. NEWNHAM, L.E. CROSS : "Electrical and acoustic emission during ferrobielastic twinning in quartz", 32d Annual Freq. Control Symp. (1978), pp. 43,

5. E. BERTAGNOLLI, E. KITTINGER, J. TICHY : "Ferrobielastic hysteresis in α-quartz", J. Appl. Phys. 50(10) (Oct. 1979), pp. 6267, 6271.

6.J.W. LAUGHNER : "Induced Dauphiné twinning in quartz", Thesis in Solid State Science (Pennsylvania St Uni.) (1982).

7. Strain at integration points is computed from nodal displacements, the stress being afterwards deduced from the strain through the material behavior law.

FIRST EXPERIENCE WITH THE ASN-16 GLONASS RECEIVER

B. Eissfeller, N. Balteas, A. Jansche, V. Liebig

Kayser-Threde GmbH, Munich, F.R.G.

ABSTRACT

The first ASN-16 GLONASS receiver was supplied to Kayser-Threde in June 1991. The ASN-16 is a slow sequencing C/A-code receiver working on one channel. It is basically designed for airborne applications but can also be used in terrestrial and maritime applications. It consists of a navigation computer, a RF-unit, a control and display unit as well as antenna and preamplifier.

An ARINC 429 compatible interface enables data transfer between receiver and external units (input of inertial aiding data, output of position, velocity and time). A software upgrade also provides access to ephemeris, almanac, pseudorange and pseudodoppler measurements.

This paper describes the ASN-16 equipment on hardware and software level. The first results of test measurements and material about the actual GLONASS constellation is presented. During the test period horizontal position errors of up to 30m were measured within 24 hours in stationary mode.

Keywords : GLONASS, ASN-16 Receiver

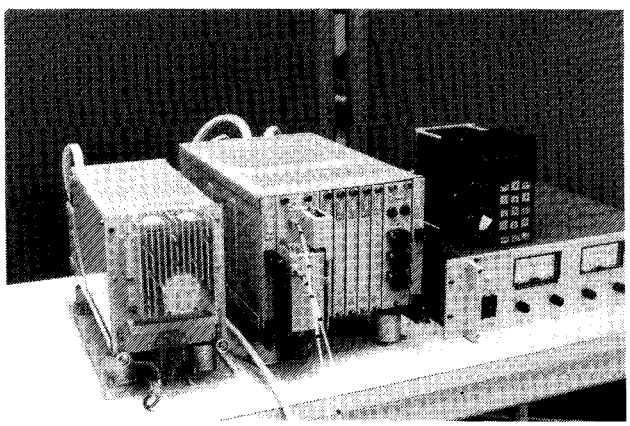


Figure 1: GLONASS ASN-16 receiver

1. INTRODUCTION

In order to provide GLONASS equipment to the western market Kayser-Threde has started in early 1991 discussions with the former Leningrad Scientific Radio Research Institute (LSRRI) and a delegation of LSRRI visited Kayser-Threde in February. Since summer 1991 Kayser-Threde began purchasing receivers directly from one of the Ukrainian manufacturers of the ASN-16 receiver.

As a result the first ASN-16 was delivered in June. This receiver however was not capable to provide the complete raw data output to the user. Therefore two additional memory boards with upgraded S/W were ordered. This modified S/W gives access to the GLONASS raw data via the 2nd port of an ARINC 429 interface.

2. TECHNICAL DESCRIPTION

The Russian ASN-16 shown in Figure 1 is a one channel slow sequencing GLONASS receiver designed by the St. Petersburg institute. The ASN-16 was constructed as navigation equipment on board of civil aircrafts.

A blockdiagram of the receiver including interfaces is shown in Figure 2.

As shown in Figure 2 the receiver consists of several subunits:
- antenna unit
- preamplifier
- receiving unit (RF-unit/SPU)
- navigation computer
- control display unit

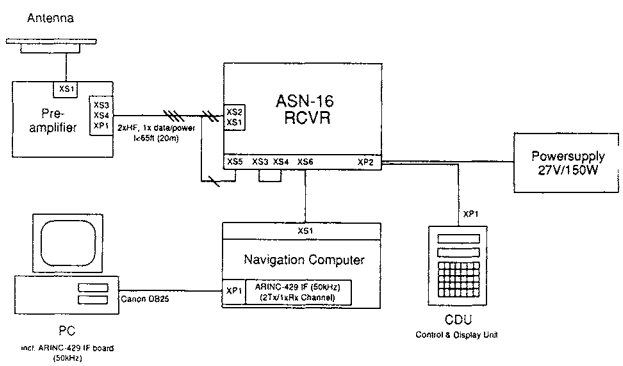


Figure 2: ASN-16 interface cabling blockdiagram

The performance of the ASN-16 is specified by the following data:

Receiver architecture:	one channel sequential
Physical data:	
Power supply:	27 V
Power consumption:	≤ 180 W
Temperature range:	- 60°C to + 55°C
Weight:	25 kg
Data output:	
Interfaces:	ARINC 429 and MIL 1553 compatible
Types of raw data:	pseudorange, pseudodoppler, almanac, ephemeris parameters
Time:	1 pps

Operational modes:
1 self-test
2 almanac mode
3 not used
4 navigation 3d/2d
5 time mode
6 planning mode
7 sophisticated self-test

Accuracy (2 s):

Geographical coordinates:	≤ ± 100 m
Altitude:	≤ ± 150 m
Ground speed:	≤ ± 0.5 m/sec
Current time:	≤ ± 2 µsec

Additional features are a changeable minimum elevation angle, the definition of up to 9 waypoints and an input of IRU (Inertial Reference Unit) position and velocity data. This means that the hybrid processing of GLONASS and inertial data is already foreseen in the navigation software.

3. RAW DATA OUTPUT

In the baseline version of the receiver the ARINC 429 interface provides Moscow time, status information, latitude, longitude, altitude and velocity. By using the two modified memory boards the following additional raw data is available:

- pseudorange
- pseudodoppler
- status of measurement
- time of measurement parameters
- ephemeris and almanac parameters

In Figure 3 the typical measurement scheme of a sequential receiver for pseudorange and pseudodoppler measurements is outlined.

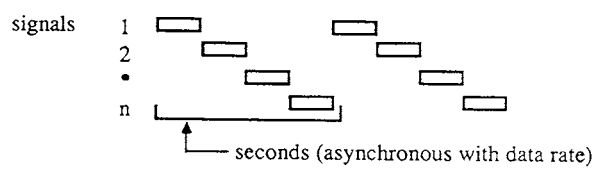


Figure 3: Sequential measurement output

The tracking time constant of the ASN-16 on one satellite is 2 seconds and the reacquisition time for the next satellite is only a fraction of a second. In the baseline version the ASN-16 is cycling through the optimal four satellite set, which is selected from the visible GLONASS constellation by aid of a minimum GDOP (Geometric Dilution of Precision) criterion. The software upgrade mentioned above enables the user to command the satellite frequency numbers to be tracked from an external PC. This means that an arbitrary set of satellites may be selected for tracking by the user.

Data transfer between ASN-16 and PC is possible via an ARINC 429 interface. The PC is equipped with an ARINC 429 IF board (50 KHz) including two receive channels and one transmit channel. The board contains a hardware control logic, memory mapped dual port RAMs and a FIFO for the transmit channel. A data decoding and logging software written in Turbo PASCAL decodes, scales and displays the incoming data. The display pages shown in the following figures can be selected by the user. Data can also be stored on disk for later analysis.

All satellites can be successively selected with function key F6. The almanac can be stored on disk by selecting function key F3.

Satellite position and velocity data consist of two ARINC 429 32 bit words, acceleration and ephemeris epoch data consist of one word. Every update of ephemeris data is stored in the logging file. Data logging can be enabled with function key F2. Function keys also allow to switch to other display pages.

The most important data is shown in the upper right hand side of the navigation page. The tracked GLONASS satellite or frequency number is displayed in the first place and the instantaneous pseudorange in nanoseconds as well as the pseudodoppler in kHz are listed below. Also very important are the time corrections listed on the following lines. The first two corrections are to be used for reconstruction of the pseudorange time of measurement relative to the 1 pps reference output of the ASN-16. With the third correction the time of measurement for the pseudodoppler can be reconstructed relative to the pps.

```
                G L O N A S S  Almanac Page (F5)
 Satellite number      :  4                 Frequency number  :  4

 Almanac age           : 1418    days
 USSR time correction  : -0.0000130236148834 23   sec.
 Time scale correction : -0.000300587084148728   sec.

 First equatorial crossing   :   3243.4687500000  sec.
 Longitude of ascending node : 115.97064971923828 10 °

 Draconitic period      : 40543.7949218750  sec.
 Change of drac. period : 0.00189208984375  sec./orbit

 Orbit eccentricity     :  0.0061273574829101 56
 Inclination            :  64.8137740418314934 00  °
 Argument of perigee    : 182.1533203125000000 00  °

 SOFTWARE WORKING INFORMATION
 Store Almanac : No
 Almanac data of satellite :     6

Store F3 | Nav. Page F4 | First Page F5 | Next Page F6 | Quit F10
```

Figure 4: Almanac display page of GLONASS satellite #4

```
                G L O N A S S  Ephemeris page (F7)   (draft)
                    Tracked satellite number                Protocol: No
label          1          2          3          4
4          AE2A0804 56861004 5ADC1E04 8FD82604
20         907A0E14 B7601614 96B21E14 CF282614
36         083E0824 22B21024 0F2A1824 31742024
52         00000834 E0C01634 E0C01E34 E0C02634
68         0DA60844 86481044 3D401844 9EA62044
84         D1200E54 E0C01654 1F401854 0FA02054
100        04AA0864 43501064 E09E1E64 B6082664
120        55A40878 6DD61078 524A1878 F48A2678
132        A9280E84 BB041684 9FC41E84 908E2084
148        6AA40894 E7B01094 2CBC1E94 87302694
164        3E840EA4 937410A4 31A81EA4 FCC026A4
180        724208B4 724210B4 724218B4 724220B4
196        B4240EC4 2D5E16C4 A75018C4 F4B620C4
204        FF700ECC 011C10CC FCA61ECC FEE226CC
212        0FA008D4 000010D4 D1201ED4 000020D4
248        166808F8 1B3616F8 051E18F8 748E26F8

  Time of Ephemeris set :  16:12:35

Protocol toggle F2 | Almanac Page F5 | Next Page F8 (test) | Quit F10
```

Figure 5: Ephemeris display page (HEX) of all tracked satellites

```
                G L O N A S S  Navigation Page (F4)
 Moscow Time :  16:14:16 not confident              Protocol: No

 Latitude  :  48.098917 °    ok     Tracked Frequency Number : 20
 Longitude :  11.542597 °    ok     Range   :  439698517.0010 ns   ok
 Altitude  :    541.188 m    ok     Doppler :    3301.8442057 kHz  ok

 N/S  Vel. :   0.00723 m/s  ok     Time correction DT1 : 938920.000 ns
 E/W  Vel. :   0.00080 m/s  ok     Time correction DT2 : 21.0000000 us
 Vert. Vel. :  0.01366 m/s  ok     Time correction DT3 : 0.69000 s ok

 No IRS data                  A S N  Status   ASN-37-4  Status: healthy
 Nav with 4 satellites                        ASN-16    Status: healthy
 Geometric factor sufficient                  ASN-16-8A Status: healthy
 IRS data not received                        ASN-16-46 Status: healthy
 Test end sign: not ended                     ASN-16-45 Status: healthy
 Operation Mode : navigation                  ASN-16-44 Status: healthy
 Initial data,almanac included: yes           ASN-16-43 Status: healthy
 Navigation field sufficient                  ASN-16-42 Status: healthy
 ASN-16 output data confident                 ASN-16-41 Status: healthy
 Time of Ephemeris set :  16:12:35            ASN-16-2  Status: healthy

Protocol toggle F2 | Almanac Page F5 | Ephemeris Page F7 | Quit F10
```

Figure 6 : Navigation and raw data display page

4. POSITION DETERMINATION ACCURACY

In order to get data about the position determination accuracy of the GLONASS ASN-16 receiver several measurement sessions have been performed since September last year.

As a basis for this paper it was tried in February to perform actual GLONASS measurements. Unfortunately the coverage of the GLONASS satellites was degraded for operational purpose since the beginning of this year. The reason for this is obviously the positioning of new GLONASS spacecrafts, because a lot of orbital manoeuvering and shifting of the satellites was noticed. In addition new frequency numbers are now seen in the almanac.

Therefore the time intervals where a complete 3-d fix under good geometric conditions is possible are currently rare and only position data over a short period of time can be shown in the following:

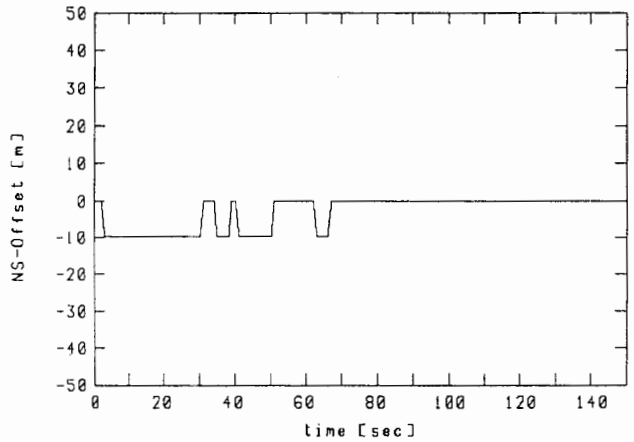


Figure 7: Variation in latitude

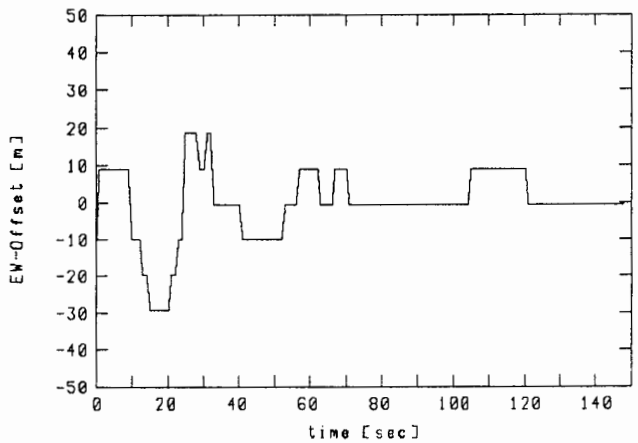


Figure 8: Variation in longitude

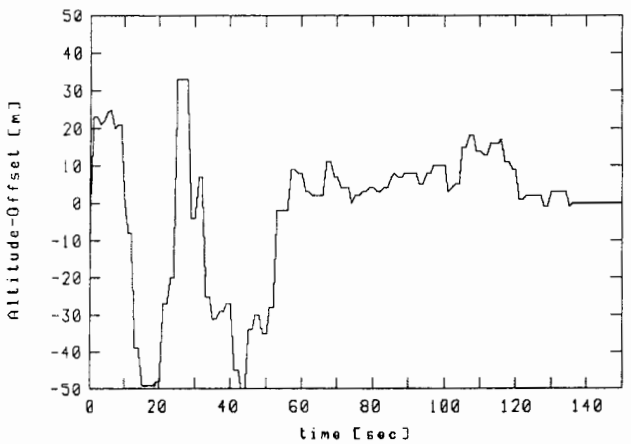


Figure 9: Variation in height

The patterns in Figure 7 and Figure 8 are due to the quantization level of the ARINC 429 interface. It is tried to get a higher resolution by additional modifications in the future.
It is shown that the variation in the horizontal coordinates is between ± 30m and the variation in height between ± 50m.

5. GLONASS CONSTELLATION DATA

At the beginning of March twelve GLONASS satellites are flagged healthy in the almanac. These are the satellite numbers 1..5, 7, 8, (orbital plane 1) 17, 19, 20, 22 and 24 (orbital plane 3). Satellite number 8 was probably part of the satellites launched in December and included in the almanac at the end of February. No satellites of orbital plane 2 are yet available (Sat no. 9..16).

The following table shows the corrspondence between satellite and frequency numbers of the GLONASS system as valid at the end of February 92:

Sat. :	01	02	03	04	05	07	08	17	19	20	22	24
Freq. :	17	09	22	04	19	13	02	21	03	15	11	14

The following plots show the number of satellites and PDOP curve for the GLONASS system covering a 24 hour period (19th Febr. 92).

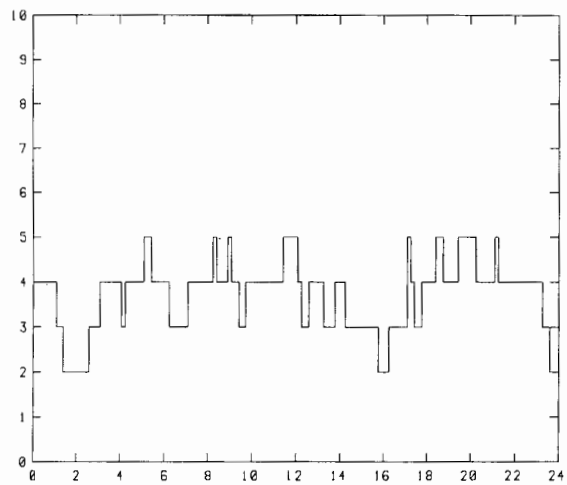


Figure 10: Number of visible GLONASS satellites at Munich

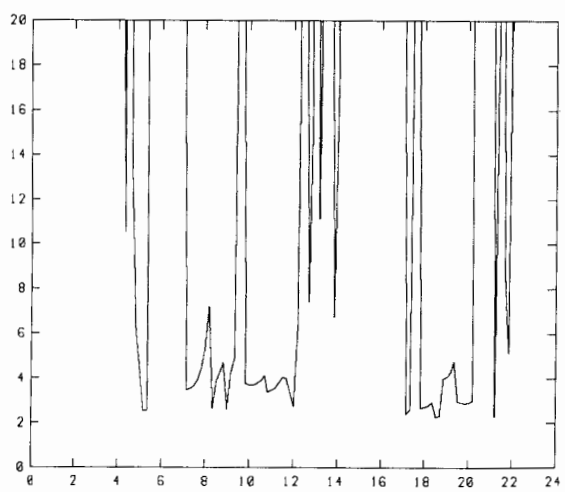


Figure 11: GLONASS PDOP over time at Munich

6. CONCLUSIONS

We have the impression that the GLONASS constellation built-up is still going on and that the three GLONASS satellites foreseen to be launched in December last year are now in orbit.

In summary:

- in Europe the first GLONASS receiver was delivered to Kayser-Threde in June 1991
- the ASN-16 provides position, velocity and time as well as pseudorange, pseudodoppler, ephemeris and almanac
- the measured horizontal position errors are about ± 30m
- the measured height errors are about ± 50m
- at the moment twelve GLONASS satellites are declared healthy in the almanac
- also new satellite frequency numbers are observed
- a lot of satellite constellation shifting is done at the moment.

STABILITY OF NONLINEAR PHASE MOTION FROM THE STUDY OF OSCILLATOR MAPPINGS

M. Planat

Laboratoire de Physique et Métrologie des Oscillateurs du C.N.R.S.
associé à l'Université de Franche-Comté-Besançon
32 avenue de l'Observatoire - 25000 Besançon - France

ABSTRACT

We have found a fundamental approach to study the intrinsic stability of phase motion at resonances of oscillator differential equations. A new averaging technique has been developed to replace the continuous problem by a discrete phase mapping (eq. (3)). Second a criterium of structural stability of resonances has been found to plot numerically the diagram of frequency-locked states (refs. 2, 3). Typical diagrams have several layers corresponding to separate power laws in the frequency fluctuation. By studying Hölder continuity of frequency in parameter space (Ω, c) the phenomenon of Hölder exponent resonance associated to fractal sets has been discovered. In the case $\varepsilon = 0$, previous power laws associated to mode lock-in or irrational windings are recovered (ref. 5). Scaling laws associated to the decay in time of frequency fluctuations have also been studied. In the case of a small $\varepsilon \neq 0$, a wide region of $1/f$ frequency noise is observed at the bottom of Arnold tongues.

THE NONLINEAR MODEL

The model is based on the study of resonances in oscillator differential equations. Let us consider the electronic signal $v(t)$ in an oscillator loop as given by solving Van der Pol's equation

$$\ddot{v} + v = \eta \dot{v}(1 - \mu v^2) + a \cos(\Omega^{-1}\tau + \Phi_0) \qquad (1)$$

where $\Omega^{-1} = \omega/\omega_0$ is the order of resonance, $\tau = \omega_0 t$ is the dimensionless time and η and μ are positive constants depending on the physical device. Equation (1) is first rewritten in polar coordinates $v = r\cos\Phi$ and $\dot{v} = r\sin\Phi$. In the unperturbed motion ($a \sim 0$) the amplitude is $r = r_0$ (constant) and the phase varies linearly in time : $\Phi = \Omega^{-1}\tau$. In the slightly perturbed case ($a \neq 0$) this is still true over a coherence time τ_c of order $1/a$ (Ref. 1, p. 143). Averaging during this time intervall and neglecting amplitude fluctuations we get the phase mapping (Refs. 2 and 3)

$$\Phi_{n+1} = \Phi_n - 2\pi\Omega + c\cos\Phi_n / (1 - c\sin\Phi_n) \qquad (2)$$

where $c = \pi a \Omega / r_0$, $r_0 = 2/\sqrt{\mu}$ and $\tau_c = 1$ for convenience's sake.

In the small amplitude limit $r_0 \sim 0$ ($c \gg 1$) no stable oscillation region can be seen since the perturbating term -cotg ϕ_n leads to Cauchy phase jump density. This probability density was observed recently in the instantaneous phase of a quartz oscillator measured by means of a phase bridge optical setup with ps resolution (Ref. 3).

In this paper I focus upon **the high amplitude region** $c < 1$. Changing Φ_n to $\pi/2 - \Phi_n$, eq. (2) is rewritten in the form of the so-called Arnold mapping (Ref. 1, p. 110).

$$\Phi_{n+1} = (1 - \varepsilon)\,\Phi_n + 2\pi\Omega - c\sin\Phi_n \qquad (3)$$

where the small parameter $\varepsilon \sim 0$ has been introduced to account for a small additive frequency shift. I believe that most phase and frequency unstabilities observed in high quality oscillator devices can be modeled with eq. (3) or its variants.

Let us begin with a typical plot of the Allan variance $\sigma_y^2(\tau)$ in the high amplitude region ($c = 0.5$; $\Omega = 1/1$; $\varepsilon = 10^{-4}$) (see Fig. 1). Here $y = \Delta\nu/\nu_0$, and ν is the winding number (perturbed frequency) which is defined as

$$\nu = \frac{\Phi_n - \Phi_o}{2\pi n}, \quad n \to \infty \qquad (4)$$

The parameter $\tau = n$ is the length of the time sample.

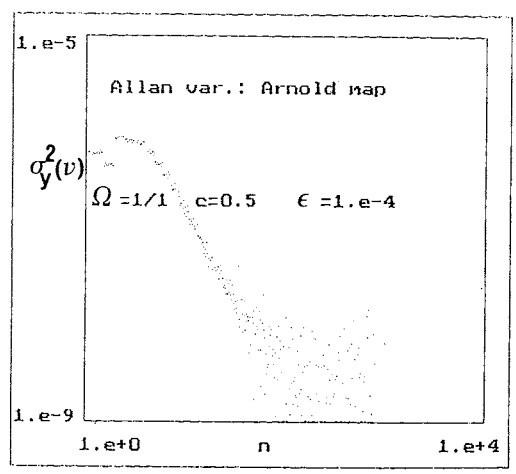


Fig. 1 : *Allan variance in perturbed Arnold circle map (3)* ($c = 0.5$; $\Omega = 1/1$; $\varepsilon = 10^{-4}$)

The best stability is seen in a window at the center of the plot ($n = 500$, $\sigma_y{}^2 = 10^{-8}$) which is typical of oscillator measurements. To get more insight into the structure of phase motion, I will now develop the concept of Arnold tongues and the associated power (or scaling) laws as it merges from nonlinear analysis.

STABILITY AND POWER LAWS IN ARNOLD MAPPING $\varepsilon = 0$

Concepts described in this subsection were introduced by Arnold (Ref. 1) and developed in a recent paper (Ref. 5). The most important one is that of **structural stability**. If $c = 0$, the curve $\nu = f(\Omega)$ is the straight line $\nu = \Omega$: no structurally stable region is seen since any slight variation of Ω leads to a new resonance. If $c \neq 0$, the curve $\nu = f(\Omega)$ has small flat steps around each resonance $\Omega = p/q \in Q$ and is called a **devil's staircase** ; they are also flat regions in the $\nu = g(c)$ curve. Each region with a constant winding number ν is a so-called **Arnold tongue** : the stable neighbourhood associated to a given resonance p/q. At $c = 1$, the devil's staircase is complete (the length of all intervalls corresponding to unlocked regions has zero measure) (see Fig. 2, curve (a)). At $c > 1$, mapping (3) is no longer invertible (it changes from monotonous increasing to cubic) and locked tongues may overlap eventually leading to chaotic motion.

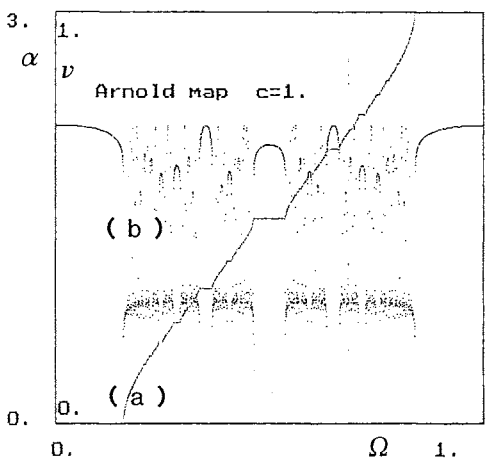


Fig. 2 : *The devil's staircase (curve a) and its associated Hölder exponent (curve b) in the critical case c = 1 in Arnold mapping*

Arnold tongues defined as above describe period 1 cycles. At $c > 1$, structurally stable cycles of period 2, 4 or more may be observed (Refs. 2 and 3). To discriminate these cycles I introduced **generalized Arnold tongues** as follows. Using the real vector parameter x (Ω, c), the perturbed rotation number is defined by the function ν : $x \rightarrow \nu(x)$ as above in (4).

Generalized Arnold tongues are defined as **Hölder continuous regions of the winding number** : two points y and z belongs to the same Arnold tongue $A_{p/q}$ if for a given real value of α :

$$\| \nu(y) - \nu(z) \| \leq M \| y - z \|^{\alpha} \qquad (5)$$

where $\| \ \|$ is the euclidien distance and M a finite real number.

In previous plots of Arnold tongues, the Lipschitz condition $\alpha = 1$ and $M = 0.9$ was used (Refs. 2 and 3).

In the subsequent paper (Ref. 4), the Hölder exponent α has been defined as $\alpha = \ell n \, \|\Delta\nu\| \, / \, \ell n \, \|\Delta x\|$, i.e. by assuming the power law $\|\Delta\nu\| = \|\Delta x\|^{\alpha}$ in some region.

Curve (b) in Fig. 2 represents Hölder exponent α versus Ω for the critical case $c = 1$. Windows of stability around each resonance p/q are clearly discriminated (by $\alpha > 1$) from irrational windings (when $\alpha < 1$). Hölder continuous regions of Arnold mapping (3) are shown in Fig. 3 (with $n = 250$ and $\Delta\alpha < 5 \times 10^{-2}$). The color code ℓ was choosen according to $\ell = [1 \, / \, |1 - \alpha|]$. Three main regions are seen :

a) the low c's region where $\alpha \sim 1$ which is unlocked but nonchaotic
b) the locked regions where α is continuous larger than 1
c) the large c's region with chaotic windings.

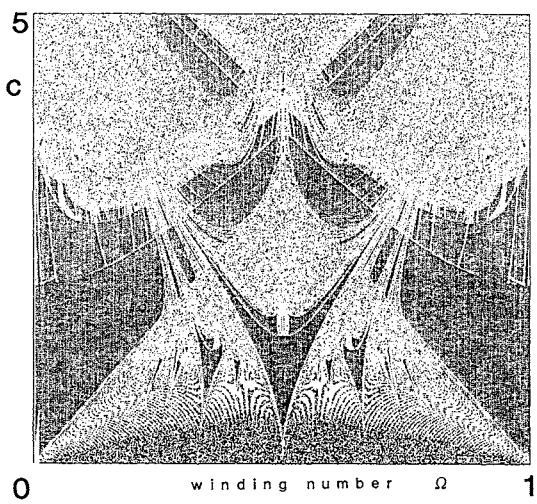


Fig. 3 : *Arnold tongues in Arnold map according to the Hölder continuity criterium*

As shown in Fig. 2 (curve (b)), the Hölder exponent is very irregular through the parameter space ; it is useful to introduce concepts of probability theory to study the set of α's. The random exponent α may be studied through its probability density $n(\alpha)$ defined as the number of occurences $N(\alpha)$ of α in a small interval $\Delta\alpha$, i.e. $n(\alpha) = N(\alpha) \, / \, \Delta\alpha$. Another useful measure is the multifractal distribution function defined as

$$f(\alpha) = log\,(N(\alpha) \, / \, \Delta\alpha) \, / \, log\,(1 \, / \, \|\Delta x\|) \qquad (6)$$

It has the advantage to be a multiplicative normalized density and a **generalized** definition of **Hausdorff dimension** which allows to characterize scaling sets with constant α.

A plot of the α-multifractal signature corresponding to the critical case of Fig. 2 is given in Fig. 4. It allows to introduce the new concept of **Hölder exponent resonance** and characterize the existence of fractal sets of the form $\|\Delta\nu\| = \|\Delta x\|^{\alpha_0}$. According to Fig. 4 irrational windings scales with the exponent $\alpha = 0.87$ in agreement with previous works (Ref. 8). The main peak at the right is found at $\alpha \sim 2.2$ and correspond to the scaling of mode lockings at the golden mean $\Omega = (\sqrt{5}-1)/2$ (Ref. 9).

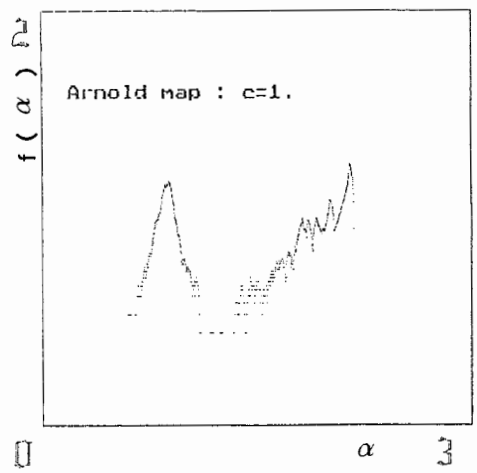


Fig. 4 : *The f(α) curve at the critical value c = 1*

The α-signature of the transition to chaos is shown in Fig. 5 with Ω = 1/1. Four main regions are observed. The one with α ≥ 3.0 is the most stable period 1 region of small c's ; the second one with 2 ≤ α ≤ 3 is the period 2 region ; the third one 1 < α ≤ 2 is an unlocked transition region where the phase jumps turn from $1/f^2$ spectrum to 1/f spectrum (Ref. 3) ; finally the region α < 1 characterize chaotic motion.

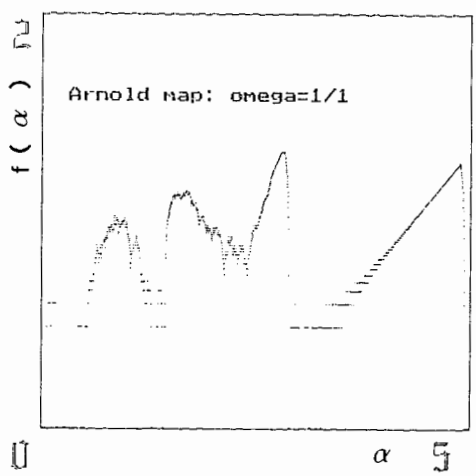


Fig. 5 : *The f(α) curve at fixed winding number 1/1*

STABILITY AND 1/f FREQUENCY NOISE IN ARNOLD MAPPING WITH ε ≠ 0

In § 2 I described frequency stability of Arnold mapping (3) in parameter space (Ω, *c*). I now introduced power laws versus discrete time *n*. Assuming a power law of the form

$$\Delta\nu = \nu_{n+1} - \nu_n = (\Delta n)^{-\beta} \qquad (7)$$

the **decay exponent β** may be computed at each point in parameter space as was α in the previous paragraph. The β-continuous regions of winding number ν in Arnold mapping (3) for the two cases ε = 0 and ε = 10^{-3} are shown in Fig. 6a and Fig. 6b respectively.

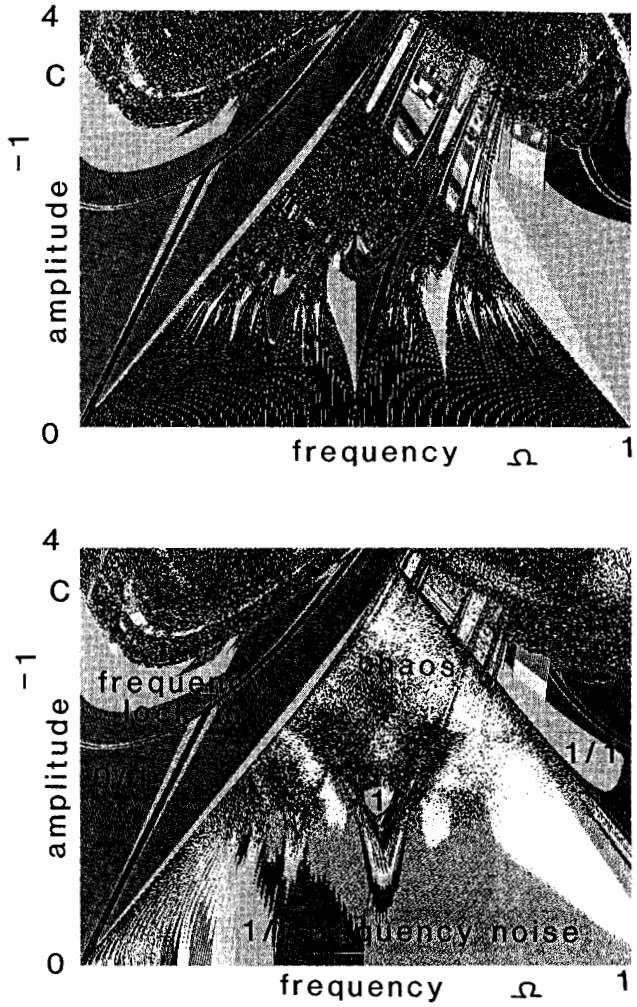


Fig. 6 : β-*continuous regions in Arnold mapping a) unperturbed mapping b) perturbed mapping* ε = 10^{-3}

Most stable regions are reminiscent of α-continuous Arnold tongues plotted in Fig. 3. However decay exponent β is not so sensitive to the lack of convergence ; this can be appreciated by looking at the $g(\beta)$ multifractal distribution defined as

$$g(\beta) = log\,(N(\beta) / \Delta\beta) / log\,(1/\Delta n) \qquad (8)$$

which mimics eq. (6). A plot of $g(\beta)$ at the critical value $c = 1$ in Arnold mapping is given in Fig. 7 (see Fig. 4 for comparison).

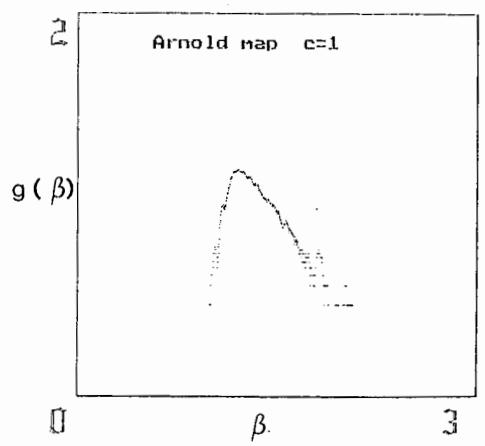


Fig. 7 : *The f(β) curve at the critical value c = 1*

Nevertheless it turns out that $g(\beta)$ is very useful to characterize the perturbed case $\varepsilon \neq 0$. In Fig. 6a ($\varepsilon = 0$) the high amplitude irrational region (α and $\beta < 1$) is very discontinuous in β. At the opposite in Fig. 6b ($\varepsilon \neq 0$) this region is very uniform.

A closer inspection reveals that in this region $\beta \sim 0.5$ quite uniformly. It means that the power spectrum of frequency scales approximately as $1/f$; I believe that this is a relevant model to explain the widely spread low frequency noise observed in solid state physics. Fig. 8a shows that time stability is reached by steps through a kind of slippery staircase. The corresponding $g(\beta)$ distribution is given in Fig. 8b : the higher peak at $\beta \sim 0.5$ corresponds to the high amplitude region of $1/f$ frequency noise while secondary resonances correspond to the steps of staircase. It should be noticed that in the 1/f frequency noise region α is very irregular with values spreading around $\alpha = 1$. A plot of Allan variance at $c = 0.5$ and $\Omega = 1/1$ was given in Fig. 1.

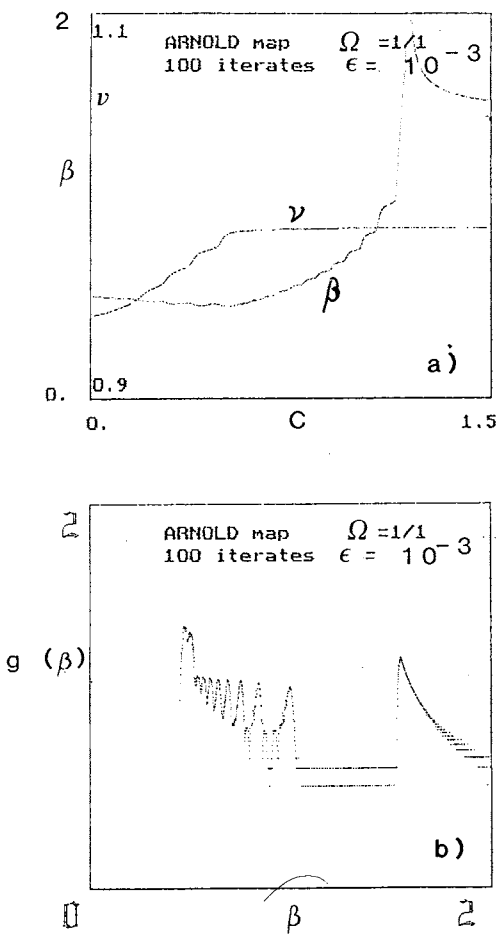


Fig. 8 :
The transition to chaos from the decay exponent β

As a conclusion, in Arnold mapping (3), when $\varepsilon \neq 0$, we have discovered a wide region of small c's where the frequency Hölder exponent remains at the boundary $\alpha \sim 1$ between order and chaos. In this region the windings are not locked in parameter space but are locked in time with $\beta \sim 0.5$. The $1/f$ frequency noise region is separated from the Arnold tongues region by time locked states forming a slippery staircase.

REFERENCES

1. V. Arnold 1980, Chapitres supplémentaires de la théorie des équations différentielles ordinaires, Editions Mir, Moscou.

2. M. Planat 1991, Nonlinear structure of phase motion from the study of differential equations near resonant tori, Lecture Notes in Physics, Springer Verlag, p. 292.

3. M. Planat, J.A. Miehe 1991, Structural stability of resonant tori modeling enhanced phase diffusion observed in quartz oscillators, Proc. of the 1991 International Conference on Noise in Physical Systems and 1/f fluctuations, edited by T. Musha, Ohmsha Ltd, p. 489.

4. Chronos 1991, La mesure de la fréquence des oscillateurs, Masson, Paris.

5. M. Planat, Hölder exponent resonances and scaling laws in the mode lock-in structure of oscillator mappings, Chaos, an interdisciplinary journal of nonlinear science (submitted).

6. M.C. Gutzwiller and B.B. Mandelbrot 1988, Phys. Rev. Lett., 60, 673.

7. B. Mandelbrot 1982, The fractal geometry of nature, New York : Freeman.

8. M.H. Jensen, P. Bak and T. Bohr 1983, Phys. Rev. Lett., 50, 1637.

9. P. Cvitanovic, B. Shraiman and B. Södelberg 1985, Physica Scripta, 32, 263.

EXPERIMENTAL RESULTS OF THE PERFORMANCE OF THE NEW PHASE AND FREQUENCY CONTROLLER

Jari Mannermaa, Kalevi Kalliomäki

University of Oulu, Department of Electrical Engineering, Finland

Abstract

The Maximum Entropy Method (MEM) has been applied to control phaselocked crystals of Loran -C receivers at the Electronic Measurements Lab of the University of Oulu and at Technical Research Centre of Finland (VTT) in Espoo. In Oulu the performance optimization of the MEM controller has achieved the result of 9 ns and $7 \cdot 10^{-12}$ for phase and frequency residuals, respectively. These values are comparable with 100 ns and $9 \cdot 10^{-12}$ of the older conventional PI controller, respectively. At VTT the performance of the MEM controlled Loran -C receiver is at least 50 % better when compared with the identical receiver equipped with the PI controller. The results verify the presumed superiority of the MEM controller over the conventional PI.

Introduction

The maximum entropy method has been studied about ten years at the Electronic Measurements Lab of the University of Oulu. The developed MEM controller has been utilized over three years to control Loran -C clocks. The performance of the maximum entropy method (MEM) controller in Loran -C receiver (with the HP 10544A oscillator) was found to be better than that of the conventional PI /1/.

A goal of this study has been to optimize the performance of the MEM controller. Therefore a series of tests with different combinations (about 100) of the parameters N, M, TC and G were carried out. Here N is the sample size, M the order of the predictor error filter, TC the control time and G the gain of the controller.

The other goal is to compare the MEM controller with conventional ones. Because the MEM controller is governed with many parameters it is not easy to set real limits for the performance of it. The attempt to find optimal combination of parameters in the case of a Loran -C receiver is not made easier by the sky noise disturbing the measurements every now and then by producing about 50-500 ns peaks lasting from 10 to 120 minutes.

However, it is possible at VTT to compare two identical Loran -C receivers to each other in the same conditions when the other one is equipped with the MEM controller and the other with a PI one.

The principal properties of the whole systems, especially the sources of noise, were first mapped. The system itself is described in the form of a block diagram in Fig. 1.

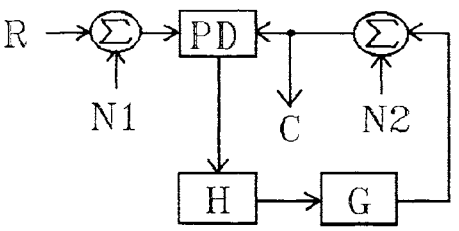


Figure 1. Block diagram of the process, in which C is the phase, G the crystal, H the controller, PD the phase difference detector, R the phase setting value, N1 the Loran -C noise (sky noise) and N2 the noise of the oscillator.

Using the notations of Fig. 1, we can derive simplified equations of the transfer function of the noises as follow:

If R = 0 and N2 = 0 then

$$\frac{C}{N1} = \frac{HG}{1 + HG}$$

if R = 0 and N1 = 0 then

$$\frac{C}{N2} = \frac{1}{1 + HG}$$

where the variables can be either Laplace or z -transformed.

It is purposed to optimize the performance of the MEM controller with respect to the phase and frequency noises, respectively. The block diagram of the MEM controller used is presented in Figure 2.

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

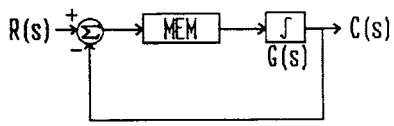


Figure 2. Block diagrams of the MEM control systems used.

The use of MEM in control systems is based on the autoregressive process having N recorded and known values:

$$x(n) = a(1)x(n-1)+a(2)x(n-2)+...+a(M)x(n-M)$$

where

x(n) is a predicted value of the known time series,
x(i)s are measured values of the time series,
a(j)s are coefficients calculated by the MEM,
M is named as the degree of the auto-regressive model or the length of the prediction error filter.

The predicted value or values are used to control the system.

The ability of the MEM controller to exploit the predictions which are calculated two or more control time interval ahead. At first the parameters of N=5, M=2, TC=30 s and G=2 were used.

The phase difference between the Loran -C signal and the crystal output of the Loran-C receiver is used as a control variable. The square of the phase residuals between the crystal and an independent rubidium clock is utilized as a measure of the performance of the controller.

Results

An interesting problem was the significance of each parameter N, M, TC and G to the final results. There are not samples enough to study the MEM parameters N and M statistically, but in Table 1 are some comparable results.

Table 1.
Test results for the MEM parameters N and M. The measuring interval is 10 minutes, TC=60 s and G=40. The measurements are done in Oulu during the period 15.08.1991-30.08.1991. The values of "Phase" and "Frequency" are calculated by using the formula of the standard deviation.

Test No.	N	M	Phase ns	Frequency 10^{-12}
1	5	2	49	70.0
2	5	4	40	66.5
3	7	2	49	82.8
4	7	4	53	77.7
5	7	6	40	84.0
6	15	2	42	68.9
7	15	8	41	74.2
8	15	14	50	120.5

Phase residuals seem not to be as sensitive to increases in N or/and M as frequency residuals.

The phase and frequency residuals as a function of gain G measured in Oulu are presented in Figures 3 a) and b). If a straight line is fitted by means of the well-known method of least squares then the phase and frequency residuals are related to gain G with the slopes of about 1 ns / gain unit and $3.8 \cdot 10^{-12}$ / gain unit, respectively.

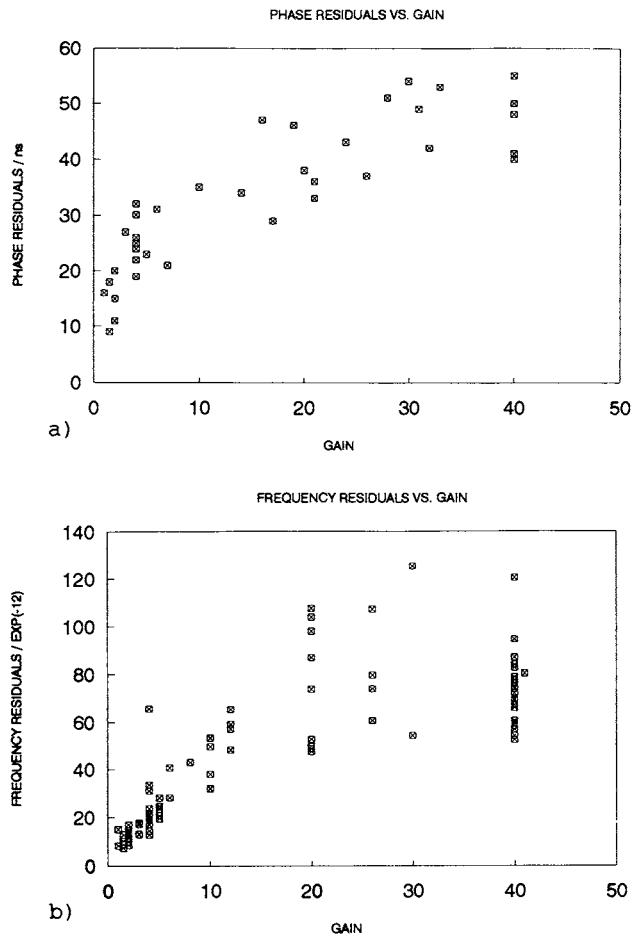
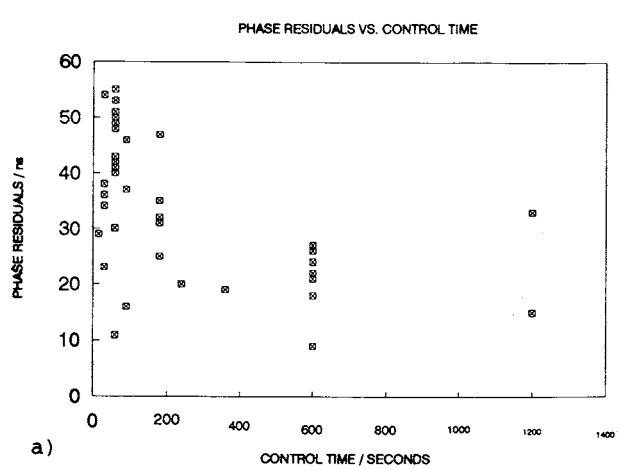


Figure 3. Phase (a) and frequency (b) residuals as a function of gain G measured in Oulu during the period 25.06.1991-11.02.1992.

Figure 4 shows the influence of the control time interval of the MEM controller on the phase and frequency residuals.



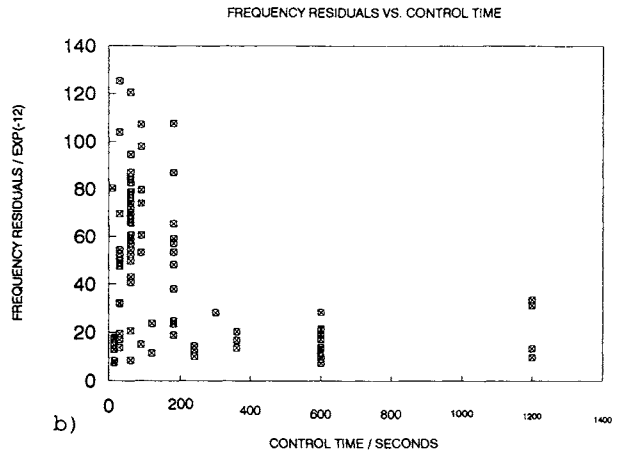


Figure 4. Phase and frequency residuals as a function of the control time interval. Measurements were done in Oulu during the period 25.061991-11.02.1992.

The effect of the sky noise on the phase residuals is studied by calculating the distribution of it during the time period 25.10.1991-13.03.1992 with the parameters N=5, M=2, TC=60 and G=40. Figure 5 shows the results.

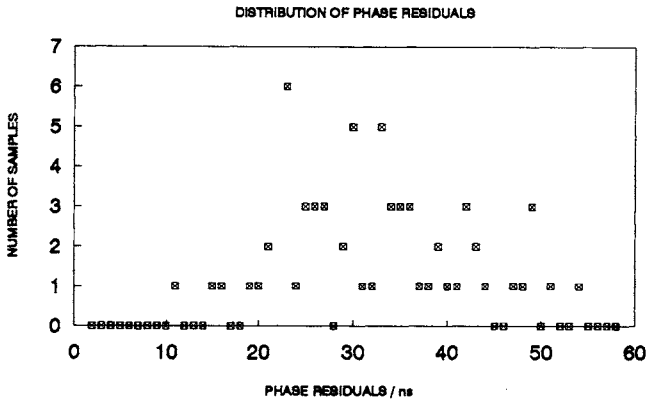


Figure 5. Distribution of the phase residuals with the parameters values N=5, M=2, TC=60 and G=40. Measurements are collected in Oulu during the period 25.06.1991-11.03.1992.

The phase and frequency residuals of different measuring time intervals are compared in Table 2.

Table 2.
Test results of the phase and frequency residuals of different measuring time intervals for the MEM controller employed in Oulu during the period 20.08.- 29.08.1992. The measuring interval is 2 hours in the case A, 10 minutes in the B one and 30 seconds in C. The values of Phase "Ph" and Frequency "Freq" are calculated by using the formula of the standard deviation.

Test No.	Ph A ns	Freq A 10^{-12}	Ph B ns	Freq B 10^{-12}	Ph C ns	Freq C 10^{-12}
9	28	3.21	40	59.3	32	16.1
10	28	3.21	40	59.3	39	21.6
11	28	3.21	40	59.3	27	15.0
12	28	3.21	33	57.0	31	17.7
13	28	3.21	49	82.8	20	16.7
14	28	3.21	51	58.5	47	21.6
15	28	3.21	53	77.7	54	15.2

The values of Table 3 are achieved by exploiting the MEM controller to calculate the predictions two control time intervals ahead.

Table 3.
Test results for the MEM controller used in Oulu during the period of 09.10.-15.10.1991. The measuring interval is 10 minutes. The values of "Phase" and "Frequency" are calculated by using the formula of the standard deviation. The parameter values N=5, M=2, TC=30 s and G=2 were used.

Test No.	Phase ns	Frequency 10^{-12}
16	33	13.3
17	19	15.8
18	38	9.9

The preliminary results with two advanced prediction steps show the average values of the phase and frequency residuals to be about 30 ns and $13 \cdot 10^{-12}$, respectively.

Table 4.
Test results for the MEM controller employed in Oulu during the period of 11.10.1991 -01.03.1992. The measuring interval is 2 hours. The values of "Phase" and "Frequency" are calculated by using the formula of the standard deviation.

Test No.	Phase ns	Frequency 10^{-12}
19	23	3.03
20	46	3.56
21	48	4.55
22	33	3.78
23	39	4.31
24	30	3.94
25	43	4.12
26	34	4.22

A typical power density spectrum of the data of a Loran -C receiver controlled by the MEM is presented in Figure 6. The number of initial measurements is 1008 (every ten minutes during a week).

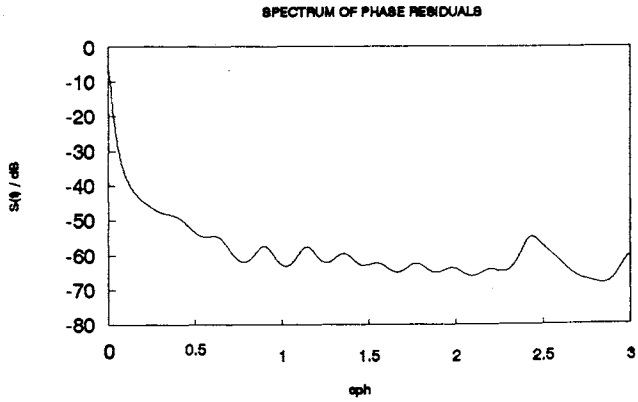


Figure 6. The power density spectrum of a typical phase data of a Loran -C receiver controlled by the MEM.

The Allan variance of a typical data after a Loran -C receiver is presented in Figure 7.

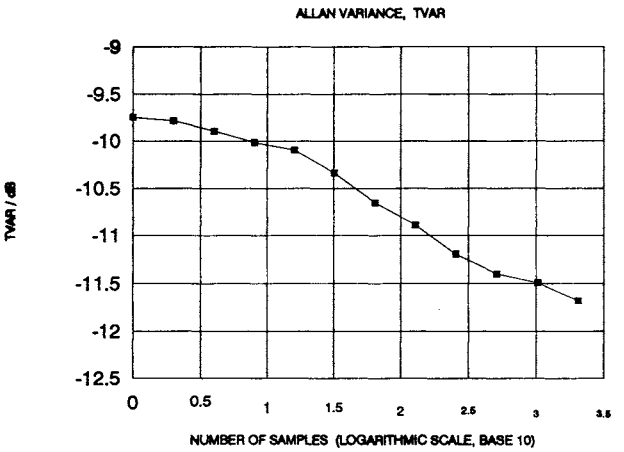


Figure 7. The Allan variance of a typical phase data of a Loran -C receiver controlled by the MEM.

Figures 8 and 9 show the influence of the gain and control time interval on the phase and frequency residuals of a Loran -C receiver at VTT.

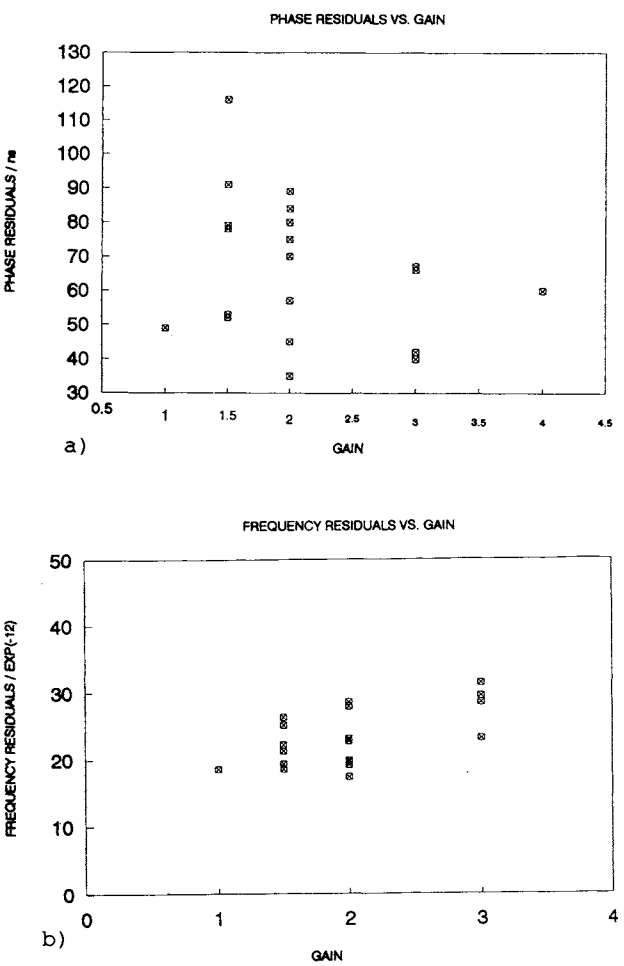


Figure 8. Phase (a) and frequency (b) residuals as a function of gain G measured at VTT during the period 25.101991-11.02.1992.

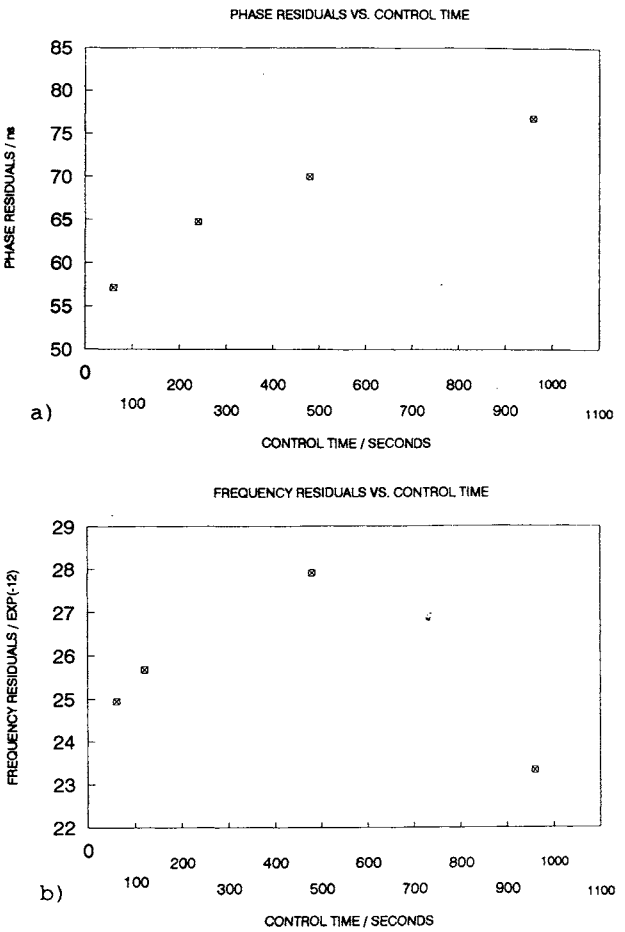


Figure 9. Phase (a) and frequency (b) residuals as a function of the control time interval measured at VTT during the period 25.061991-11.02.1992.

Table 5.
Test results of two identical Loran -C receivers for the MEM (B) and PI (A) controllers used at VTT during the period of 25.10.1991-04.02.1992. The measuring interval is 10 minutes. The values of Phase "Ph" and Frequency "Freq" are calculated by using the formula of the standard deviation.

Test No.	Ph A ns	Freq A 10^{-12}	Ph B ns	Freq B 10^{-12}
27	47	9.55	42	28.5
28	47	9.81	40	31.4
29	65	9.39	48	42.0
30	57	10.9	45	19.2
31	44	13.7	35	17.4
32	82	10.6	49	18.6
33	114	11.1	91	25.1
34	66	8.10	53	22.2
35	56	7.42	52	18.6
36	98	9.50	80	28.6
37	107	11.6	70	27.9
38	88	10.4	66	29.5
39	127	12.9	84	23.1
40	107	12.7	79	21.3
41	129	12.0	78	19.4
42	106	11.3	75	19.7
43	109	11.3	67	23.2
44	100	11.0	60	21.7

Table 6.
Test results of two identical Loran -C receivers for the MEM (B) and PI (A) controllers used at VTT during the period of 25.10.1991-04.02.1992. The measuring interval is 2 hours. The values of Phase "Ph" and Frequency "Freq" are calculated by using the formula of the standard deviation.

Test No.	Ph A ns	Freq A 10^{-12}	Ph B ns	Freq B 10^{-12}
45	129	7.22	122	6.26
46	110	7.62	104	6.42
47	124	9.65	111	6.85

The results of the short term control time interval show that the MEM controller is better than the PI one in time keeping but worse in frequency keeping. In the long term ones the MEM controller is better in both cases.

Discussion

The phase difference between the noisy Loran -C signal and the crystal output of the Loran -C receiver is a real control variable. This is the reason to use the square of the phase residuals between the crystal and an independent rubidium clock as a measure of the performance of the controller. Frequency is only a computational result obtained by differentiating phase.

At this moment there are several successful parameter combinations but the best achieved results is the phase residuals of 9 ns and frequency standard deviation of $6.71 \cdot 10^{-12}$ with the parameters of N=10, M=5, TC=600 s and G=1.5. The recent values of the digital PI controller are about 100 ns and $9 \cdot 10^{-12}$. The studies have pointed out that the use of the values of G<1 is not reasonable. However, the control time and gain have a remarkable influence on final results whereas the MEM parameters of N and M are mysteries yet.

When the values of the control time and gain of the MEM controller increase then the phase and frequency residuals have a tendency to increase too.

The typical spectrum of the Loran -C receiver phase reveals the white noise process in the signal and the same results is verified by the Allan variance having the slope of about -0.6.

At VTT in Espoo two identical Loran -C receivers are compared to each other under the same conditions execpt the other one is equipped with the MEM controller and the other with PI one. The preliminary results indicate that the performance of the MEM controller is at least 1.5 times as effective as that of PI one in keeping time.

In general the phase and frequency residuals seem to improve when the gain decreases and the control time interval increases. This tendency in the results is quite clear both in Oulu and Espoo.

The MEM controller is able to exploit the predictions which are calculated two or more control time intervals ahead. At first the parameters of N=5, M=2, TC=30 s and G=2 were used. The preliminary results with two advanced prediction steps show the average values of the phase and frequency residuals to be about 30 ns and $13 \cdot 10^{-12}$, respectively.

Reference /1/:

Jari Mannermaa, Kalevi Kalliomäki; "A New Way To Control Phaselocked Frequency Standards", EFTF91, Besancon, France.

TIME STABILITY : AN IMPROVEMENT OF THE MULTI-VARIANCE METHOD FOR THE OSCILLATOR NOISE ANALYSIS

F. Vernotte*, E. Lantz**

*Observatoire de Besançon
41 bis avenue de l'Observatoire, BP 1615, 25010 BESANÇON CEDEX - FRANCE
**Laboratoire d'Optique P. M. Duffieux, U. A. CNRS 214
Faculté des Sciences et des Techniques, 25030 BESANÇON CEDEX- FRANCE

ABSTRACT

Generally, oscillator noise analysis is performed with a single variance (i. e. Allan variance, modified Allan variance, ...). Since the noises altering the output signal of oscillators may be modeled as power laws in the spectral density of frequency deviation, the oscillator noise analysis consists in measuring the quantity of each power law noise. For each noise type, a variance yields a particular slope versus the integration time. Thus, it is possible to measure the different noise type contributions by interpolating a variance result curve.
However, the sensitivity of the single variance method may be insufficient. In order to increase the measurement accuracy, the multi-variance method was developed. This method uses as many variances as there are noise types to measure. The principle of this method consists in obtaining the N noise type contributions with the N variance results for each integration time value. Despite the convergence of the average values of these noise coefficient measurements, the different measures exhibit a large dispersion versus the integration time. It is then necessary to perform this measurement for several integration time values in order to average out the dispersion.
An improvement of this method consists in operating simultaneously the whole data obtained for different variances and for different integration times. Thus, an only measure is obtained for each type of noise, and this measure is the most probable in the sense of the least squares. Moreover, the previous multi-variance and single variance methods are included as particular cases in this new method. The performances of this method will be compared to the ones of the previous method.

Keywords : frequency stability, noise, spectral density, variance, transfer function, least squares.

1. INTRODUCTION

For several years, the quantity of systems using high-quality oscillators such as quartz oscillators, frequency synthesizer or atomic clocks, has been widely increasing. Thus, many researchers and engineers working in various fields have been interested in the oscillator stability characterization.
The oscillator instabilities can be deterministic or random. The random fluctuations are more difficult to study. Different types of noises are generally distinguished, each of them corresponding to particular properties. The origins of these noises may be linked to the oscillator environment (temperature changes, electromagnetic fields, shocks, vibrations) but also they may be internal (thermal noise...) (Refs. 1, 2). However, the links between these causes and the different types of noises are not still well-known. The reason is the lack of accuracy of the measurement techniques of the contributions characterizing each type of noise.
The aim of the job presented in this paper consists in attempting to improve the measurement accuracy in order to better distinguish the different types of noises. Following the presentation of the different notion, tools and mathematical models used in the frequency and time field, the classical measurement method using one single variance will be recalled. The multi-variance method, based on the use of several variances, will be described. An improvement of this multi-variance method will be proposed and the measurements accuracy of these methods compared.

2. BASIC QUANTITIES AND RELATIONSHIPS. MATHEMATICAL MODELS.

The output signal V(t) delivered by an oscillator of nominal amplitude V_0 and nominal frequency ν_0 may be modelled by the following relationship :

$$V(t) = V_0 \sin(2\pi\nu_0 t) \tag{1}$$

In order to take the random phase fluctuations into account, the relationship (1) may be rewritten as :

$$V(t) = V_0 \sin\left(2\pi\nu_0 t + \varphi(t)\right) \tag{2}$$

The phase noise may be expressed as frequency noise by using the relationship :

$$\Delta\nu(t) = \frac{1}{2\pi}\frac{d\varphi(t)}{dt} \tag{3}$$

However, the most often used quantity in the Frequency and Time domain, is the instantaneous normalized frequency deviation y(t) defined as :

$$y(t) = \frac{\Delta\nu(t)}{\nu_0} \tag{4}$$

The instantaneous normalized frequency deviation is linked to the phase noise by the following relationship :

$$y(t) = \frac{1}{2\pi\nu_0}\frac{d\varphi(t)}{dt} \tag{5}$$

However, only the average of this quantity over a time interval is observable. Generally, y_k samples are defined as :

$$\bar{y}_k = \frac{1}{\tau}\int_{t_k}^{t_k+\tau} y(t)dt \tag{6}$$

These samples may be calculated from the measures provided by a phase comparator :

	Variance			
	Allan	Modified Allan	high-pass	band-pass
Measurement sequence	$\sigma_y^2(\tau) = \frac{1}{2}\left\langle (\bar{y}_2 - \bar{y}_1)^2 \right\rangle$	$\sigma_M^2(\tau) = \frac{1}{2n^2}\left\langle \sum_{i=1}^{n} (\bar{y}_{i+n} - \bar{y}_i)^2 \right\rangle$	—	—
Transfer function (square of the modulus)	$\frac{2\sin^4(\pi\tau f)}{(\pi\tau f)^2}$	$\frac{2\sin^6(\pi\tau f)}{(\pi\tau f)^2 n^2 \sin^2(\pi\tau_0 f)}$	$\frac{2}{(\pi\tau f)^2}\frac{f^4}{f_h^4 + f^4}$	$\frac{8}{\pi^2}\frac{(2\tau f)^2}{(2\tau f)^4 + \left[(2\tau f)^2 - 1\right]^4}$

Table 1 : Transfer functions and measurement sequences of 4 different variances

$$\bar{y}_k = \frac{\varphi(t_k + \tau) - \varphi(t_k)}{2\pi\nu_0\tau} \qquad (7)$$

Thus, in order to study the frequency (or phase) fluctuations of an oscillator, a sequence of N consecutive y_k samples is commonly calculated (N must be large enough to allow statistical processing : we generally used 4096 or 8192 samples). These samples may be contiguous ($t_{k+1} = t_k + \tau$) or not (dead time sampling).

Generally, in order to express the previously seen quantities in the frequency domain, spectral densities are used. The spectral density $S_y(f)$ of a function y(t) is defined as the Fourier Transform of the autocorrelation of this function :

$$S_y(f) = F.T.\left[\int_{-\infty}^{+\infty} y(\theta)y(\theta - t)d\theta\right] = \left|F.T.[y(t)]\right|^2 \qquad (8)$$

The main advantage of the spectral densities is their determinstic nature even in the case of random processes. Different behaviors of the spectral density of instantaneous normalized frequency deviation $S_y(f)$ have been experimentally observed. It is commonly assumed that 5 different spectral density patterns may be encountered in oscillator output signals according to 5 different power laws (Refs. 1, 2) :

$$S_y(f) = \sum_{\alpha=-2}^{+2} h_\alpha . f^\alpha \qquad (9)$$

Each of these types of noises corresponds to a physical cause (vibrations, changes in the temperature, electronic agitation). However, the links between these causes and these effects are not well-known. The reason for this lack of knowledge is the mediocre accuracy of the noise coefficient h_α.

3. SINGLE VARIANCE METHOD

3.1. PRINCIPLE

In order to study random processes, statistical methods are generally used. Thus the IEEE Subcommitee on Frequency Stability has proposed to choose the Allan Variance (two sample variance) in order to stantardize the oscillator measurement methods (Ref. 3). This variance is defined by the following measurement sequence :

$$\sigma_y^2(\tau) = \frac{1}{2}\left\langle (\bar{y}_2 - \bar{y}_1)^2 \right\rangle \qquad (10)$$

In the frequency domain, the Allan variance may be calculated with its transfer function $H_y(f)$ which is the Fourier Transform of its measurement sequence (Refs. 1, 2) :

$$\sigma_y^2(\tau) = \int_{-\infty}^{+\infty} S_y(f).\left|H_y(f)\right|^2 df \qquad (11)$$

where

$$\left|H_y(f)\right|^2 = \frac{2\sin^4(\pi\tau f)}{(\pi\tau f)^2} \qquad (12)$$

	Variance			
	Allan	Modified Allan	high-pass	band-pass
$S_y(f)$	$\sigma_y^2(\tau)$	$\sigma_M^2(\tau)$	$\sigma_{HP}^2(\tau)$	$\sigma_{BP}^2(\tau)$
$h_2.f^2$	$\frac{3h_2 f_h}{4\pi^2\tau^2}$	$\frac{3h_2}{8\pi^2\tau^3}$	$\frac{2h_2 f_h}{\pi^2\tau^2}$	$\frac{h_2}{\sqrt{2}\,\pi\,\tau^3}$
$h_1.f^1$	$\frac{h_1[1.038 + \mathrm{Log}2\pi f_h\tau]}{4\pi^2\tau^2}$	$\frac{h_1[24\mathrm{Log}2 - 9\mathrm{Log}3]}{8\pi^2\tau^2}$	$\frac{h_1 \mathrm{Log}[1 + (\pi\tau f_h)^4]}{2\pi^2\tau^2}$	$\frac{h_1}{\sqrt{2}\,\pi\,\tau^2}$
$h_0.f^0$	$\frac{h_0}{2\tau}$	$\frac{h_0}{4\tau}$	$\frac{h_0}{\sqrt{2}\,\tau}$	$\frac{\sqrt{2}h_0}{\pi\tau}$
$h_{-1}.f^{-1}$	$2h_{-1}$	$\frac{h_{-1}[27\mathrm{Log}3 - 32\mathrm{Log}2]}{8}$	$\frac{\pi h_{-1}}{2}$	$\frac{5h_{-1}}{\pi\sqrt{2}}$
$h_{-2}.f^{-2}$	$\frac{2\pi^2 h_{-2}\tau}{3}$	$\frac{11\pi^2 h_{-2}\tau}{20}$	$\frac{\pi^2 h_{-2}\tau}{\sqrt{2}}$	$\frac{8\sqrt{2}h_{-2}\tau}{\pi}$

Table 2 : Responses of variances for the different types of noises

Several other variances have been defined by choosing new transfer functions : modified Allan variance, filtered Allan variance, high-pass variance, band-pass variance, etc... (Refs. 1, 2, 4). It can be noticed that these last two variances may be defined only in the frequency domain since their measurement sequences are impossible to calculate in the time domain. Table 1 shows the transfer functions of these variances and their measurement sequences when it is possible to calculate them. Since the variances exhibit different slopes for the different types of noises versus the integration time τ (see table 2), the principle of a variance measurement consists in calculating the variance with different integration time τ and then, in plotting those measurements on a log-log graph versus τ. The identification of each slope on this graph implies the identification of each type of noises in the spectral density $S_y(f)$ of the oscillator output signal. For a f^{α} noise, the variance measurement graph exhibits a τ^{μ} slope and generally $\mu = -1-\alpha$ (Refs. 1, 2) :

$$S_y(f) = h_\alpha . f^{\alpha} \Leftrightarrow \sigma_y^2(\tau) = C_\mu . \tau^{\mu} \qquad (13)$$

3.2. Results

In order to estimate the sensitivity of a variance for the measurement of the different noise coefficients h_α, the Allan variance was applied to signals of which noise coefficients were perfectly known. The simplest way to know these coefficients consists in generating signals by computer.

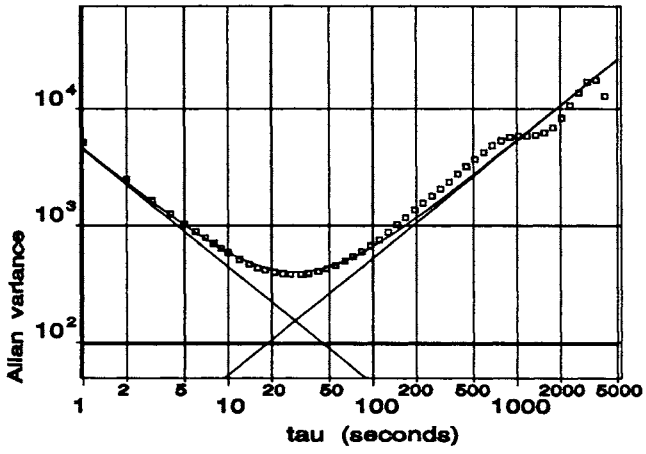


Figure 1 : Allan variance applied to a signal composed of f^0, f^{-1}, f^{-2} noises (h_0=10000, h_{-1}=10, h_{-2}=1).

Figure 1 shows the Allan variance measurement graph for a signal composed of 3 types of noises : white frequency noise (f^0 noise), flicker frequency noise (f^{-1} noise) and random walk frequency noise (f^{-2} noise). This simulated signal is constituted of 8192 samples figuring in a sequence of 8192 real y_k samples. Table 3 collects the results obtained for signals composed of different proportions of the 3 types of noises. This table shows that the accuracy of this method for these types of signals is about 20% if each of the 3 types of noises is successively prevailing in the frequency range (h_0=10000, h_{-1}=1000, h_{-2}=1).

However, if one type of noise is swamped by the others, the measurement of the h_α of this type of noise decreases drastically (see figure 1 and table 3 for h_0=10000, h_{-1}=10, h_{-2}=1 and h_0=10000, h_{-1}=10000, h_{-2}=1). Thus we were induced to define the dynamics of this method :

> ***the dynamics is the range in which a h_α coefficient can vary without swamping the other ones nor being swamped by the other ones.***

The limits of this range has been chosen to keep an accuracy of 25% for all coefficients. Under these conditions, the dynamics of the single variance method (Allan variance applied to a 8192 sample sequence composed of 3 types of noises) has been estimated at 15 dB (h_{-1} may vary from 100 to 3000 while preserving a measurement accuracy of the three h_α coefficients better than 25%).

Simulation values			Calculated values					
h_0	h_{-1}	h_{-2}	h_0	$\frac{\Delta h_0}{h_0}$	h_{-1}	$\frac{\Delta h_{-1}}{h_{-1}}$	h_{-2}	$\frac{\Delta h_{-2}}{h_{-2}}$
10000	10	1	8950	10.5 %	70	600 %	0.803	20 %
10000	1000	1	9253	7.5 %	1099	10 %	0.742	25 %
10000	10000	1	10660	6.6 %	10520	5.2 %	0.354	65 %

Table 3 : Results obtained with the Allan variance applied to simulated signals.

4. MULTI-VARIANCE METHOD

4.1. Principle

The aim of this method consists in using N different variances for calculating N h_α coefficients (Ref. 2). Let us consider the system constituted of 3 variances : the Allan variance, the high-pass variance ant the bandpass variance. Table 2 provides the responses of these 3 variances for the 3 previously considered types of noises :

$$\begin{cases} \sigma_y^2(\tau) = h_0 \cdot \frac{1}{2\tau} + h_{-1} \cdot 2\text{Log}2 + h_{-2} \cdot \frac{2\pi^2\tau}{3} \\ \sigma_{HP}^2(\tau) = h_0 \cdot \frac{1}{\sqrt{2}\tau} + h_{-1} \cdot \frac{\pi}{2} + h_{-2} \cdot \frac{\pi^2\tau}{\sqrt{2}} \\ \sigma_{BP}^2(\tau) = h_0 \cdot \frac{\sqrt{2}}{\pi\tau} + h_{-1} \cdot \frac{5}{\pi\sqrt{2}} + h_{-2} \cdot \frac{8\sqrt{2}\tau}{\pi} \end{cases} \qquad (14)$$

The resolution of this system by the Cramer method allows us to rewrite it in such a way that each h_α coefficient is equal to a linear combination of the 3 variance results :

$$\begin{cases} h_0 = \frac{4\sqrt{2} - \frac{5\pi}{2}}{\Delta}\tau.\sigma_y^2(\tau) + \frac{\frac{5\sqrt{2}\pi}{3} - \frac{16\sqrt{2}\,\text{Log}2}{\pi}}{\Delta}\tau.\sigma_{HP}^2(\tau) + \frac{\sqrt{2}\pi^2\,\text{Log}2 - \frac{\pi^2}{3}}{\Delta}\tau.\sigma_{BP}^2(\tau) \\ h_{-1} = \frac{\pi - \frac{8}{\pi}}{\Delta}.\sigma_y^2(\tau) + \frac{\frac{4\sqrt{2}}{\pi} - \frac{2\sqrt{2}\pi}{3}}{\Delta}.\sigma_{HP}^2(\tau) + \frac{\frac{\sqrt{2}\pi^2}{3} - \frac{\pi^2}{2\sqrt{2}}}{\Delta}.\sigma_{BP}^2(\tau) \\ h_{-2} = \frac{\frac{5}{2\pi} - \frac{1}{\sqrt{2}}}{\Delta}\frac{1}{\tau}.\sigma_y^2(\tau) + \frac{\frac{5\sqrt{2}\,\text{Log}2}{\pi} - \frac{2}{2\sqrt{2}\pi}}{\Delta}\frac{1}{\tau}.\sigma_{HP}^2(\tau) + \frac{\frac{\pi}{4} - \sqrt{2}\,\text{Log}2}{\Delta}\frac{1}{\tau}.\sigma_{BP}^2(\tau) \end{cases} \qquad (15)$$

where Δ is the determinant of the system (14) :

$$\Delta = 2\sqrt{2} - \frac{5\pi}{4} + 2\pi\,\text{Log}2 - \frac{16\,\text{Log}2}{\pi} + \frac{5\pi}{3} - \frac{\sqrt{2}\pi^2}{3} \qquad (16)$$

Thus, it is possible to obtain a measure of each h_α coefficient for each integration time value τ whereas, with the single variance method, we must perform an interpolation over many variance measurements obtained with different τ values.

4.2. Results

In order to efficiently compare these two methods, we applied the multi-variance method to the same simulated signals as those used with the single variance method. Figure 2 shows the h_α measurements obtained by this method for different τ values. This figure shows that these h_α measurements are not constant versus τ at all. However, the

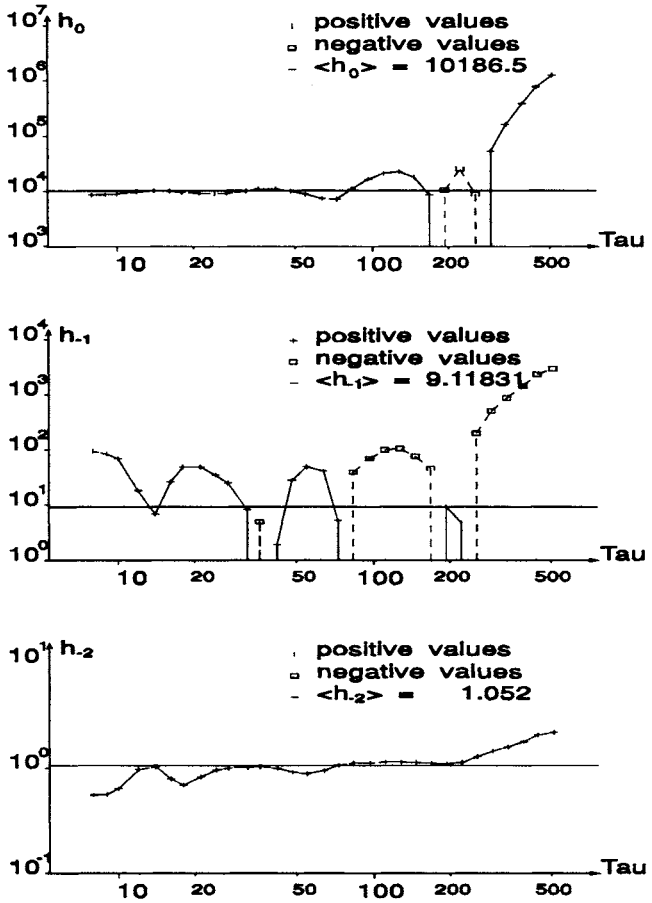


Figure 2 : Multi-variance method applied to a signal composed of f^0, f^{-1}, f^{-2} noises (h_0=10000, h_{-1}=10, h_{-2}=1).

mean value of these h_α measurements (plotted as a solid line) remains allways very close to the real h_α values entered for the simulation.

Table 4 shows the results obtained by the multi-variance measurement for the signals previously studied by the single variance method. The accuracy is better than 5% and the dynamics of this method is about 30 dB.

5. IMPROVEMENT OF THE MULTI-VARIANCE METHOD

5.1. PRINCIPLE

Because of the large dispersion of the h_α measurement around the mean value, it is necessary to perform several variance measurements for different τ values. Instead of calculating each h_α coefficient for each τ value and then calculating the mean value, it is better to calculate the h_α coefficients which fit at best all the variance measurements. This calculation is performed by minimizing the relative difference between the variance measurements and the theoretical variance curve. The result is then the most probable h_α coefficients in the sense of the least squares. Moreover, this new multi-variance method can run with any number of variances. The single variance method and the previous multi-variance method are then particular cases of this new method (one single variance or one single τ value). Let us consider N different variances and M different τ values. Denoting $\sigma^2_i(\tau_k)$ the result of the variance i for the integration time τ_k and $\phi_{i\alpha}(\tau_k)$ the response of the variance i for a f^α noise, it comes :

Simulation values			Calculated values					
h_0	h_{-1}	h_{-2}	h_0	$\frac{\Delta h_0}{h_0}$	h_{-1}	$\frac{\Delta h_{-1}}{h_{-1}}$	h_{-2}	$\frac{\Delta h_{-2}}{h_{-2}}$
10000	10	1	10186	1.9 %	9.1	8.8 %	1.052	5.2 %
10000	1000	1	10010	0.1 %	1030	3.0 %	1.021	2.1 %
10000	10000	1	8477	15.2 %	10453	4.5 %	0.701	30 %

Table 4 : Results obtained with the multi-variance method applied to simulated signals.

$$\left\{\begin{array}{l}\left[\sigma_1^2(\tau_1) = \sum_{\alpha=-2}^{+2} h_\alpha \Phi_{1\alpha}(\tau_1)\right] \\ \quad \cdots \\ \left[\sigma_i^2(\tau_1) = \sum_{\alpha=-2}^{+2} h_\alpha \Phi_{i\alpha}(\tau_1)\right] \\ \quad \cdots \\ \left[\sigma_N^2(\tau_1) = \sum_{\alpha=-2}^{+2} h_\alpha \Phi_{N\alpha}(\tau_1)\right] \\ \quad \cdots \\ \left[\sigma_i^2(\tau_k) = \sum_{\alpha=-2}^{+2} h_\alpha \Phi_{i\alpha}(\tau_k)\right] \\ \quad \cdots \\ \left[\sigma_i^2(\tau_M) = \sum_{\alpha=-2}^{+2} h_\alpha \Phi_{i\alpha}(\tau_M)\right] \\ \quad \cdots \end{array}\right. \qquad (17)$$

The condition for the minimization of the relative difference between the variance measurements and the theoretical variance response is then :

$$\sum_{k=1}^{M}\sum_{i=1}^{N}\left[\frac{\sigma_i^2(\tau_k) - \sum_{\alpha=-2}^{+2} h_\alpha \Phi_{i\alpha}(\tau_k)}{\sigma_i^2(\tau_k)}\right]^2 \text{ is minimum} \qquad (18)$$

It can be demonstrated that this condition is verified by the h_α coefficients obtained by the resolution of the following 5 equations with 5 unknowns system :

$$\left\{\begin{array}{l}\sum_{\alpha=-2}^{+2} A_{\alpha-2}.h_\alpha = B_{-2} \\ \sum_{\alpha=-2}^{+2} A_{\alpha-1}.h_\alpha = B_{-1} \\ \sum_{\alpha=-2}^{+2} A_{\alpha 0}.h_\alpha = B_0 \\ \sum_{\alpha=-2}^{+2} A_{\alpha+1}.h_\alpha = B_{+1} \\ \sum_{\alpha=-2}^{+2} A_{\alpha+2}.h_\alpha = B_{+2}\end{array}\right. \qquad (19)$$

where :

$$A_{\alpha\beta} = \sum_{k-1}^{M}\sum_{i-1}^{N} \frac{\Phi_{i\alpha}(\tau_k)\Phi_{i\beta}(\tau_k)}{\left[\sigma_i^2(\tau_k)\right]^2} \qquad (20)$$

and :

$$B_\beta = \sum_{k-1}^{M}\sum_{i-1}^{N} \frac{\Phi_{i\beta}(\tau_k)}{\sigma_i^2(\tau_k)} \qquad (21)$$

5.2. RESULTS

This new multi-variance method has been applied to the previously studied signals. Table 5 shows the results obtained with 3 variances (Allan variance, high-pass variance, band-pass variance) and the table 6 the results obtained with 4 variances (the same ones plus the modified Allan variance). Figure 3 presents the graph of the Allan variance measurements and of the theoretical curve fitting at best the measurements (fit obtained with 4 variances and 19 τ values). This figure may be compared to Figure 1 since the simulated signals are the same. It is possible to see that the single variance method do not allow to measure the h_{-1} coefficient in this case, because it is too small with regard to the other ones, whereas the new multi-variance method measure it with a 10% accuracy. The dynamics and the accuracy of this new method with 3 variances is almost the same as the previous multi-variance method (accuracy 5%, dynamics 30 dB). But, with 4 variances, the dynamics of the new multi-variance method

Simulation values			Calculated values					
h_0	h_{-1}	h_{-2}	h_0	$\frac{\Delta h_0}{h_0}$	h_{-1}	$\frac{\Delta h_{-1}}{h_{-1}}$	h_{-2}	$\frac{\Delta h_{-2}}{h_{-2}}$
10000	10	1	9974	0.3 %	6.4	36%	1.039	4%
10000	20	1	9941	0.6.%	23	17 %	1.004	0.4%
10000	10000	1	9445	5.6 %	9999	0.0 1%	0.801	20 %
10000	20000	1	12864	28 %	19283	3.6 %	2.457	146%

Table 5 : Results obtained with the new multi-variance method (3 variances) applied to simulated signals.

is about 44 dB. This new multi-variance method seems to be at least as good as the previous multi-variance method.

6. CONCLUSION

Besides its good dynamics and accuracy, the new multi-variance method presents several advantages :

- at first, it is a modular method. In order to get a quick estimation of the noise levels in a signal, it is possible to use one single variance. But if it is necessary to obtain a better measurement accuracy, 3, 4 or more variances may be used.
- secondly, this method may present the results as an Allan variance curve. This can be particularly useful since the Allan variance is a well-known tool, universally used in the research field as in the engineering field.
- at last, this method lends itself well to a dispersion analysis of the variance measurements with regard to the theoretical curve. This dispersion could give some information about the accuracy of the h_α measurements.

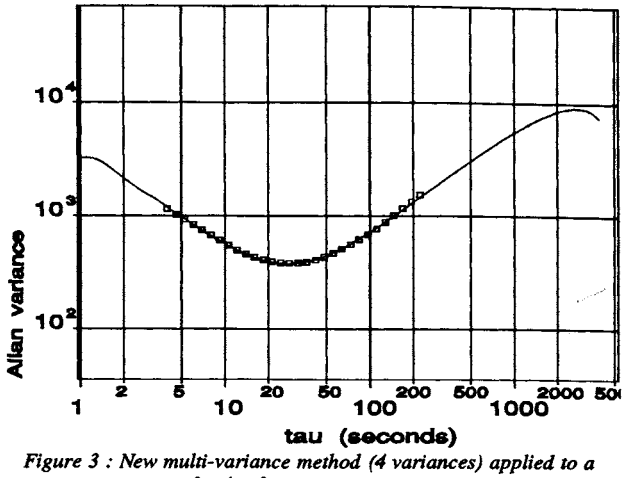


Figure 3 : New multi-variance method (4 variances) applied to a signal composed of f^0, f^{-1}, f^{-2} noises (h_0=10000, h_{-1}=10, h_{-2}=1). Allan variance result curve

Simulation values			Calculated values					
h_0	h_{-1}	h_{-2}	h_0	$\frac{\Delta h_0}{h_0}$	h_{-1}	$\frac{\Delta h_{-1}}{h_{-1}}$	h_{-2}	$\frac{\Delta h_{-2}}{h_{-2}}$
10000	2	1	10011	0.1 %	2.76	38 %	1.015	1.5%
10000	5	1	10010	0.1 %	5.82	16.3 %	1.015	1.5%
10000	50000	1	8932	10 %	49679	0.6 %	0.842	16 %
10000	100000	1	8030	20 %	99364	0.6 %	0.595	40 %

Table 6 : Results obtained with the new multi-variance method (4 variances) applied to simulated signal

7. ACKNOWLEDGEMENT

The authors express their thanks to F. Meyer for his help during the redaction of this paper.

8. REFERENCES

1. Rutman J 1978, Characterization of phase and frequency instabilities in precision frequency sources : fifteen years of progress, Proceedings of the IEEE, 66, 9, 1048-1075.

2. Vernotte F 1991, Stabilité temporelle des oscillateurs : nouvelles variances, leurs propriétés, leurs applications, Ph. D. Thesis, Université de Franche-Comté, N# 199.

3. Barnes JA, Chi AR, Cutler LS, Healey DJ, Leeson DB, McCunigal TE, Mullen JA, Smith WL, Sydnor RL, Vessot R, Winkler GMR 1971 : Characterization of frequency stability, IEEE Trans. Instrum. Meas.,IM-20, 105-120.

4. Groslambert J, Gagnepain JJ, Vernotte F, Walls F 1989 : A new filtered Allan variance and its application to the identification of phase and frequency noise sources : 43rd annual Frequency Control Symposium, Denver.

ANISOTROPIC ETCHING OF SILICON IN KOH SOLUTIONS

A. Brahim-Bounab, C.R. Tellier

L.C.E.P., E.N.S.M.M., Besançon, France

ABSTRACT

New procedures can be used to predict etched shapes encountered in anisotropic micromaching processes which start with the dissolution slowness surface. Changes in the etch rate and in dissolution shapes of differently-oriented silicon plates are analysed to obtain the slowness surface describing the anisotropic etching of silicon in KOH solutions. Theoretical etched shapes are then derived and compared with experimental results. The adequation of the proposed three-dimensional representation of the slowness surface is discussed.

Keywords : Anisotropic etching, dissolution slowness surface, silicon, micro-machining.

1. THEORY

1.1 Expressing the dissolution slowness surface

The surface of a crystal can be decomposed in a succession of planar surface elements whose orientations (φ, θ) correspond to those related to crystallographic planes {hkl}. To construct numerically the etched surface we have to track the displacements of the surface elements within the crystal. Several crystals (Si, Ge, Ga As and quartz) suffer highly anisotropic chemical attacks (Ref. 1), hence it is usual to assume that the dissolution slowness, $\vec{L}$, depends only on the orientation (φ, θ) of a surface element (Refs 1, 2). Under this assumption the trajectories of surface element are found to be linear and Tellier et al showed (Ref. 2) that the displacement of a surface element represented by a propagation vector, $\vec{P}$, can be completely determined starting from the equation L (φ, θ) expressing the representative surface of the dissolution slowness.

Since the slowness vector, $\vec{L}$, lies parallel to the unit inward normal, $\vec{n}$, to a surface element we can write its magnitude, L, in terms of dissolution tensors D_i, D_{ij}, D_{ijk},... :

$$L = D_o + D_i n_i + D_{ij} n_i n_j + D_{ijk} n_i n_j n_k \qquad (1)$$

1.2 Dissolution constants for the class m3m

To reduce the number of dissolution constants we use the symmetry properties of point group m3m :

* Class m3m is centro-symmetric
* The four rotation axes 3 along the cube diagonals permute cyclically the subscripts
* Class m3m possesses three rotation axes 4 parallel to the cartesian reference axes

Applying the symmetry transformations related to point group m3m to tensors of maximum rank, N_{max}, we obtain a number n_N of independent dissolution constants which satisfies the relation (Ref. 2)

$$n_N = 2 N_{max}$$

For $N_{max} \leq 10$. The results for $N_{max} = 12$ are summarized in table 1.

i=1	i=1, j=2	i=1, j=2, k=3
$D_{12(i)}$	$D_{2(i)\ 10(j)}$ $D_{4(i)\ 8(j)}$ $D_{6(i)\ 6(j)}$	$D_{2(i)\ 4(j)\ 6(k)}$ $D_{2(i)\ 2(j)\ 8(k)}$ $D_{4(i)\ 4(j)\ 4(k)}$

Table 1 : Identification of independent constants for $N_{max} = 12$

1.3 The generalized equation for the slowness surface

A doubly rotated surface element of orientation (φ, θ) has a unit inward normal $\vec{n}$ whose components are given by :

$$n_1 = \sin\varphi \cos\theta,\ n_2 = -\cos\varphi\cos\theta,\ n_3 = -\sin\theta \qquad (2)$$

Expanding the compact Eq.(1) in terms of non vanishing dissolution constants and using Eq. (2) leads to the final expression for L (φ, θ) :

$$L(\varphi, \theta) = S_1^* + S_2^* \qquad (3)$$

with

$$S_1^* = \sum_{k=0} a_k \left[\cos\theta \sin\theta\right]^{2k} \qquad (4)$$

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

$$S_2^* = \sum_{l=1} \sum_{m=0} b_{l,m} \left(\sin\varphi \cos\varphi \cos^2\theta\right)^{21} \left(\cos\theta\right)^{2m} \quad (5)$$

Symmetry considerations give additional relations between the coefficients a_l and $b_{l,m}$:

$$\sum b_{l,m} = a_l \quad (6)$$

which are finally expressed in terms of the dissolution constants (Ref. 2).

2. EXPERIMENTAL RESULTS

2.1 Experimental procedure

The silicon slices used were initially circular singly-rotated ($\varphi = \varphi_0, \theta = 0°$)polished plates (1 mm thick, 13.2 mm in diameter) cut in a silicon ingot. The etchant employed was KOH with a concentration of 35%. The plates were etched at a constant temperature of 80°C for various periods of time. The etch rates were evaluated from decrements in thickness, Δd, measurements. The geometrical features of etched surfaces were characterized by Scanning Electron Microscopy (SEM) and by using a microprocessor-based surface profilometer.In addition the various shapes of the etched contour of silicon plates were studied by using a Talyrond analyser.

2.2 Dissolution rate of silicon plates

The etch rate $R = \frac{1}{2} \Delta d/dt$ is found to depend markedly on the angle of cut φ_0. Table 1 gives some typical results ; it should be noted that the measured ratio R{100} takes a value close to that evaluated by others investigators (Ref.3)

Table 2 : Variations of R with the angle of cut φ_0

φ_0	0°	14°	26°	34°	45°
R(µm/min)	≈ 1.1	≈ 1.8	≈ 2.4	≈ 2.4	≈ 2.1

2.3 Surface profilometry and SEM observations

To study the changes in the surface texture with the orientation φ_0 of etched plates, profilometry traces were made along two specified directions : the cristallographic axis Z and the rotated direction X'. Profilometry traces made on four differently oriented plates are shown in Figures 1 and 2. The X' and Z surface profiles exhibit marked changes on the angle φ_0. In particular a prolonged etching causes the formation of X' profiles with an alternate shape for plates with φ_0=14° and φ_0=34° and the development of a ridged Z background on {101} plates. SEM provides additional informations (Figure 3) and reveals that the geometrical features of dissolution figures are essentially determined by the orientation of the plate on which they are formed. For example a concave background develops on {100} plates whereas etchings produce a linearly texture structure on {110} plates.

2.4 Out-of-roundness profiles

The out-of-roundness analyser results displayed in Figure 4 indicate that the KOH etchant gives rapidly rise to out-of-roundness profiles which present accentuated maxima and minima. These extrema occur for angles Ψ_ε which remain quasi-unchanged with repeated etchings. Moreover the resulting shape of out-of-roundness profiles is strictly

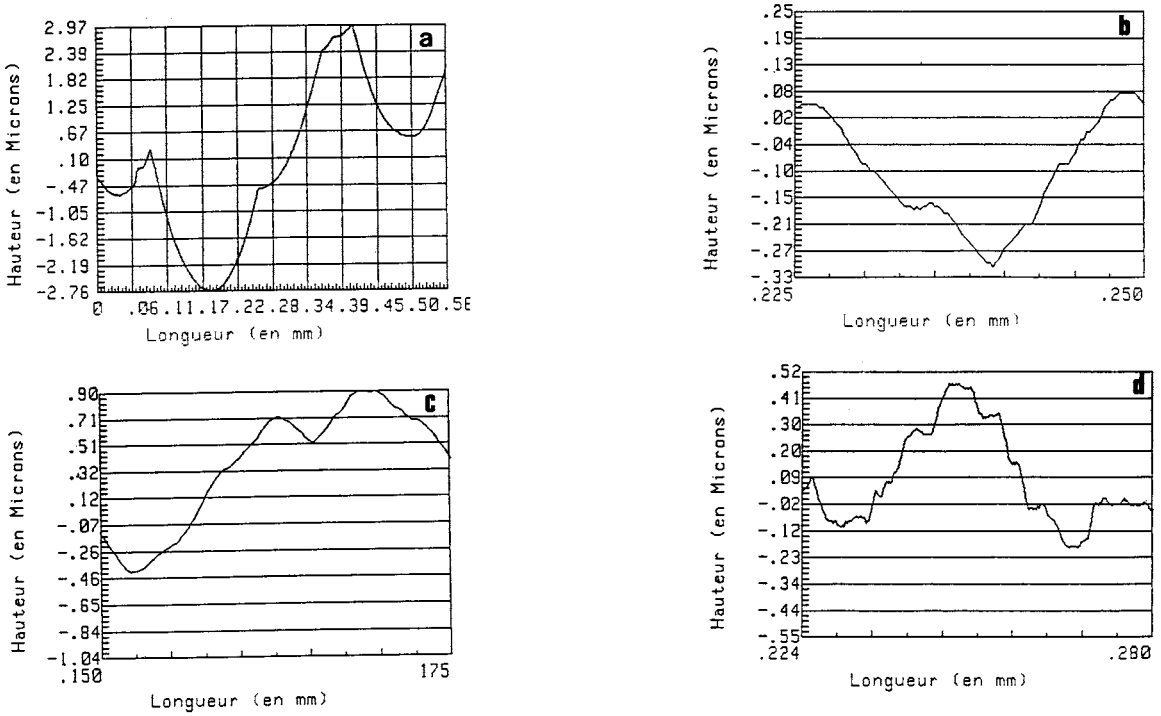


Figure 1 : X' profilometry traces for singly rotated plates : $(\varphi_0, \theta^\circ = 0)$, a) $\varphi_0 = 0^\circ$, b) $\varphi_0 = 14°$, c) $\varphi_0 = 34°$ d) $\varphi_0 = 45°$

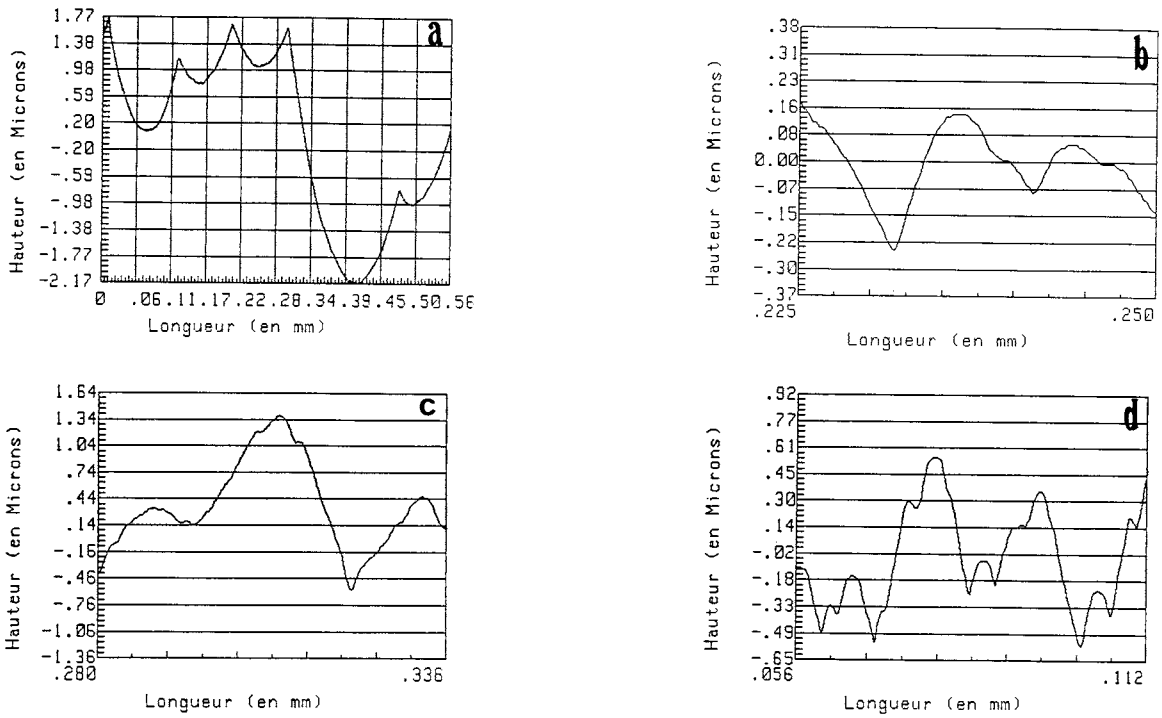


Figure 2 : Z profilometry traces for singly rotated plates : a) $\varphi_0 = 0^\circ$, b) $\varphi_0 = 14°$, c) $\varphi_0 = 34°$, d) $\varphi_0 = 45°$

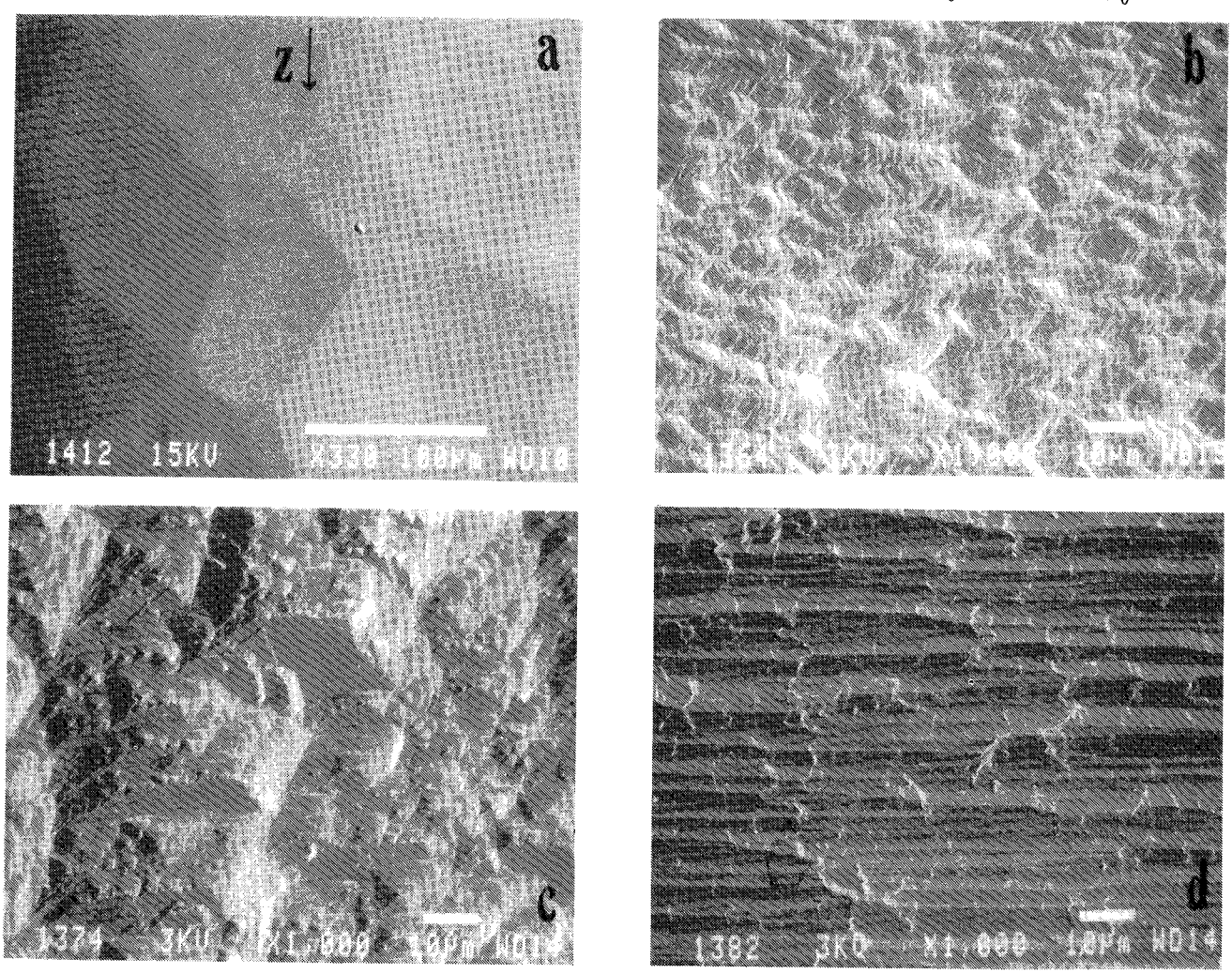


Figure 3 : S.E.M. observations : a) $\varphi_0 = 0^\circ$, b) $\varphi_0 = 14°$, c) $\varphi_0 = 34°$, d) $\varphi_0 = 45°$

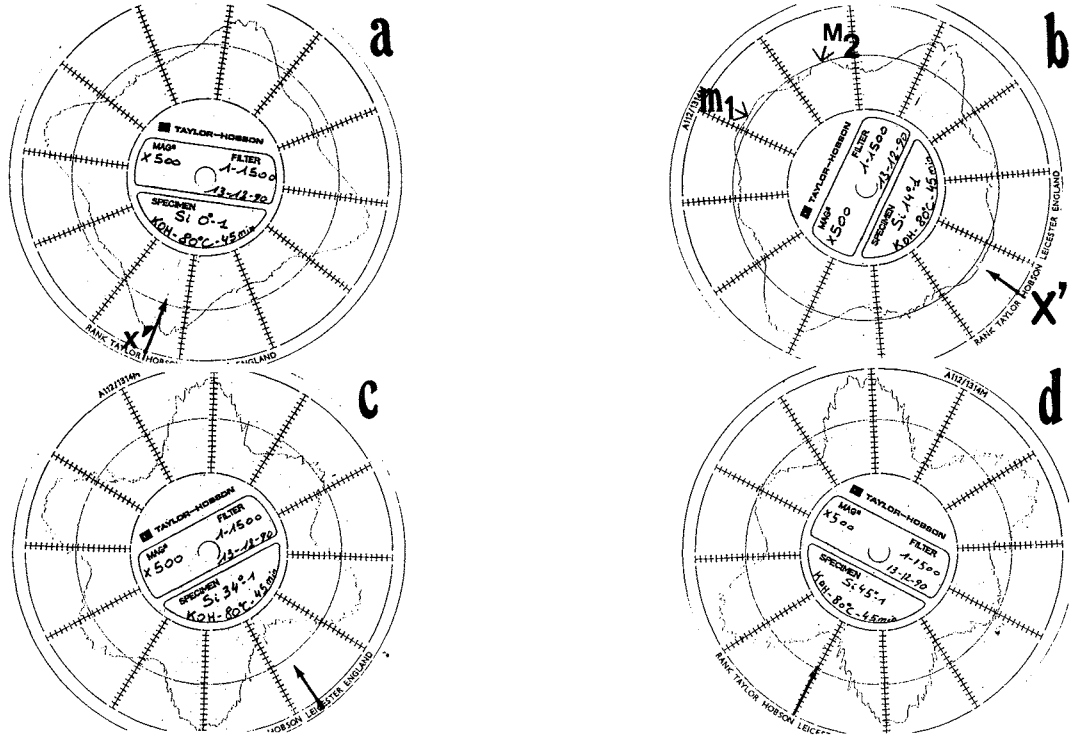


Figure 4 : Out of roundness profiles of etched plates a) $\varphi_0 = 0^\circ$, b) $\varphi_0 = 14°$, c) $\varphi_0 = 34°$, d) $\varphi_0 = 45°$

connected to the orientation of the starting circular sections. A digonal symmetry is clearly depicted on all out-of-roundness profiles because the sections contain a <100> axis. Moreover the more accentuated maxima in the {110} out-of-roundness profile are found to correspond to {111} planes.

3. SIMULATION OF DISSOLUTION SHAPES

3.1 The dissolution slowness surface of silicon crystal

The anisotropic dissolution of crystals is governed by the dissolution criteria [Refs. 1, 2, 4] which state that a concave intersection formed by two initial planes is stable provided there is no plane between them with a higher dissolution slowness and that conversely a convexe intersection remains stable when there is no plane with a lower dissolution slowness. Then when we start with a cylindrical crystal the final etched cross-sectional shape is composed of limiting planes associated with minima in L, i.e. with the valleys of the out-of-roundness profiles. It is the reason why a out-of-roundness profile constitutes a crude representation of the corresponding cross-section of the slowness surface (Ref. 5). Making use of this property and of the experimental L vs φ_0 plot (section {100}) we have derived, by successive iterations, an approximate three-dimensional representation of the slowness surface (Figure 5). To obtain a slowness ratio greater than 30 for the {111} : {110} planes together with a relative agreement between experimental and theoretical cross-sectional shapes it has been necessary to include tensor of rank 12 in the iteration procedure.

3.2 Theoretical cross-sectional shapes

A numerical simulation of the etched shape of an initially circular section of orientation φ_0 can be accomplished starting with the corresponding polar diagram of $\vec{L}$ (Refs 1,6). The final out-of-roundness profiles related to sections with φ_0 = 0°, 14°, 34° and 45° are reproduced in Figure 6. We observe a very good agreement between theoretical and experimental profile graphs of {100} sections. The agreement remains rather satisfactory for φ_0=34° and φ_0 = 45° but several deviations are depicted for φ_0=14°. In particular the secondary minimum m_1 (Figure 4) does not appear in the theoretical graph (Figure 6). Moreover as compared with the experimental graph the maximum M_2 in Figure 6 is too accentuated.

3.3 Theoretical surface profiles

To perform the numerical simulation of etched profiles we start with a real digitalized surface profile as worked by a microprocessor-based profilometer.

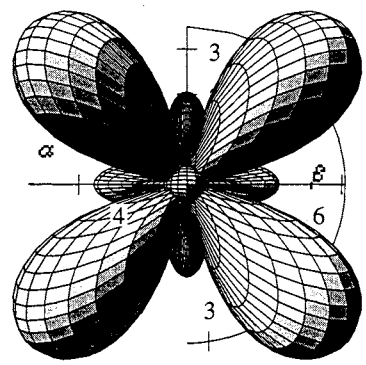


Figure 5 : The dissolution slowness surface as plotted in the form lnL versus (φ,θ) : the direction of observation is {100}

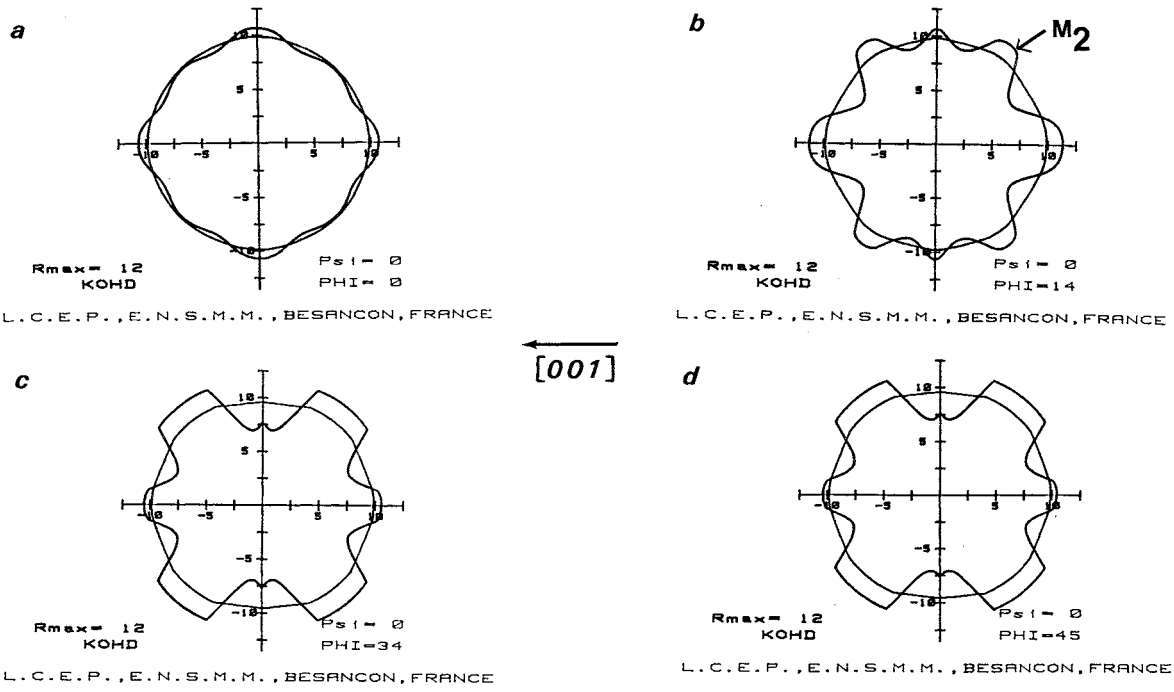


Figure 6: Theoretical cross-sectional shapes a) $\varphi_0 = 0^\circ$, b) $\varphi_0 = 14^\circ$, c) $\varphi_0 = 34^\circ$, d) $\varphi_0 = 45^\circ$

The changes in shapes with angle φ_0 are illustrated in Figures 7 and 8 for X' and Z traces. Turning our attention to theoretical X' profiles we observe that as the angle of cut φ_0 increases from 0° to 14° the etched concave background is transformed into an alternate concave-convex profile. For φ_0=45° the numerical simulation predicts the early formation of a concave background whereas for φ_0=34° we obtain without ambiguity an alternate convex-concave etched X' profile. All these theoretical features fit remarkably well the experimental observations (Figure 1). The experimental shapes (Figure 2) related to traces made along the Z axis do not yield so clearly a close agreement with the theory (Figure 8) except evidently for φ_0=0°, a cut for which the Z trace doesn't differ from the X' trace. For plates with φ_0=14°, 34° and 45° the theoretical etched shape is found to depend on the geometrical features of the starting profile. All portions of the initial profile composed of elements with small slopes give rise to etched concave portions while the rougher regions favour the development of etched profiles with persistent slopes. This theoretical scheme is specially verified for φ_0=34° and φ_0=45°.

4. DISCUSSION AND CONCLUSION

The changes in shape of etched profiles with orientation are usually discussed in terms of the dissolution criteria presented in section 3.1. Thus it is clear that any departures between the experimental and the theoretical etched shapes of starting circular cross-sections result from a poor agreement between the geometrical features of exact and predicted slowness surfaces. Comparing the out-of-roundness profiles with the theoretical cross-sectional shapes it seems that the predicted slowness surface presents undoubtedly geometrical imperfections for φ_0 in the vicinity of 14°. In particular on the theoretical polar diagram $L(\varphi_0 = 14°, \theta)$ we distinguish less extrema than on the corresponding out-of-roundness profile (Table 3). For φ_0 =14°, 34° and 45° the anisotropy of the dissolution appears to be more considerable than for φ_0=0°. The prediction of Z surface profiles is thus performed with etching times shorter than those used to derive X' profiles. In this condition the etched shapes remain sensitive to the initial surface damage and do not really give useful additional informations on the φ_0 range for which deviations are expected.

However the linearly textured structure which develops on {110} plates can be understood in terms of a polar diagram $L(\varphi_0 = 45°, \theta)$ whose shape around θ =0° approaches that depicted in Figure 4d. Effectively the persistence of slopes in experimental Z profiles is certainly correlated to the existence of two minima in L lying symmetrically with respect to the maximum occuring at θ =0°. Moreover the complete agreement between theory and experiments ({100} cross-sectional shape, X' surface profiles) indicates that for φ in the vicinity of 0° the predicted surface represents accurately the real slowness surface.

Theoretical values		Experimental values	
Minima, θ_m	Maxima, θ_M	Minima, θ_m	Maxima, θ_M
15°, 59°	0°, 35°,90°	0°, 29°, 54°	10°, 40°, 90°

Table 3: Values of the extrema in L as computed and as estimated from out-of-roundness profiles

From the preceding discussion it appears that for $\varphi_0 \approx 14°$ we have to modify the theoretical polar graph to support experiments. At this point it should be remarked that the procedure of generation of the slowness surface consists in a first step to derive a {100} polar graph which fits very well the experimental polar plot (the discrepancy is less than 2%

for θ varying from 0° to 45°). Suppose now that we are concerned with the dissolution tensor of rank 12 (seven independent dissolution constants). To obtain small departure between predicted and experimental {100} polar graphs we necessary fix the values of four dissolution constants. Thus to generate other singly rotated polar graphs we work only with three constants. The adjusment becomes a difficult task and it is the reason why the procedure inevitablably introduces geometrical deviations. The only possibility to overcome this difficulty consists of combining a greater number of dissolution constants. This implies to work with tensors of higher rank which enable us to generate more complicated slowness surfaces [Ref. 2]. Actually tentative are made to use tensors of rank 16 and 18 to derive more detailed dissolution slowness surfaces.

REFERENCES

1 - Sangwal K 1987, **Etching of Crystals**, North-Holland, Amsterdam.

2 - Brahim-Bounab et al 1991, **J. Mater. Scl**, 26, 5585-5594.

3 - Seidel H 1990, **J. Electrochem. Soc.**, 137,11,3612-3626.

4 - Irving B 1962, in **The Electrochemistry of Semiconductors**, (Holmes P J, ed) Academic Press, London.

5 - Tellier C 1991, **J. Mater. Scl.**, 26, 5595-5607.

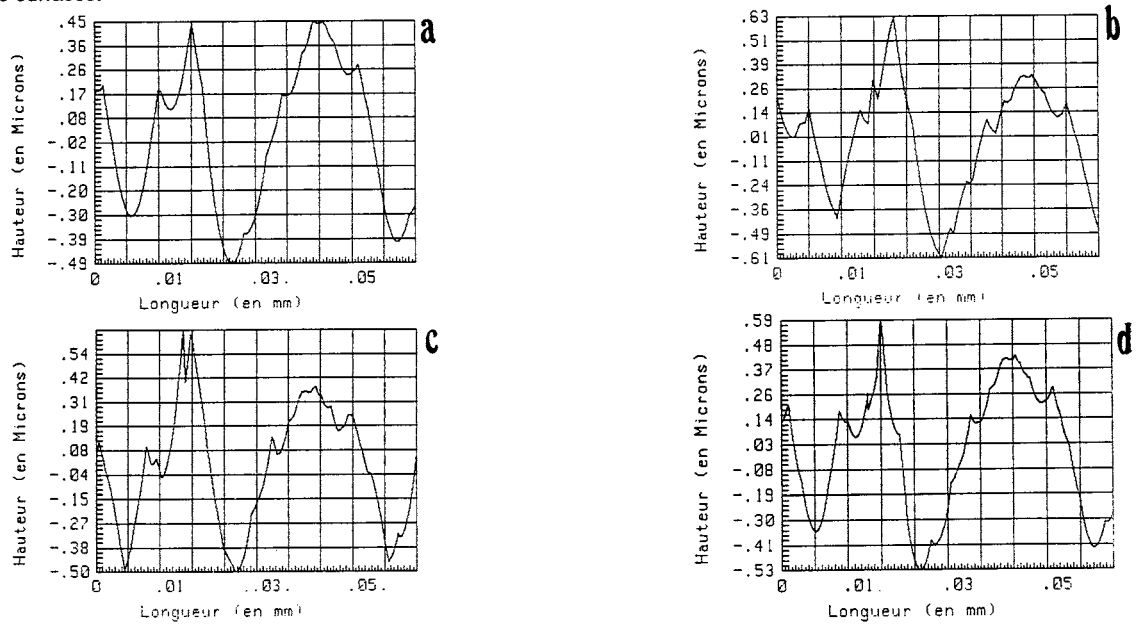


Figure 7 : Theoretical X' surface profiles a) $\varphi_0 = 0^\circ$, b) $\varphi_0 = 14^\circ$, c) $\varphi_0 = 34^\circ$, d) $\varphi_0 = 45^\circ$

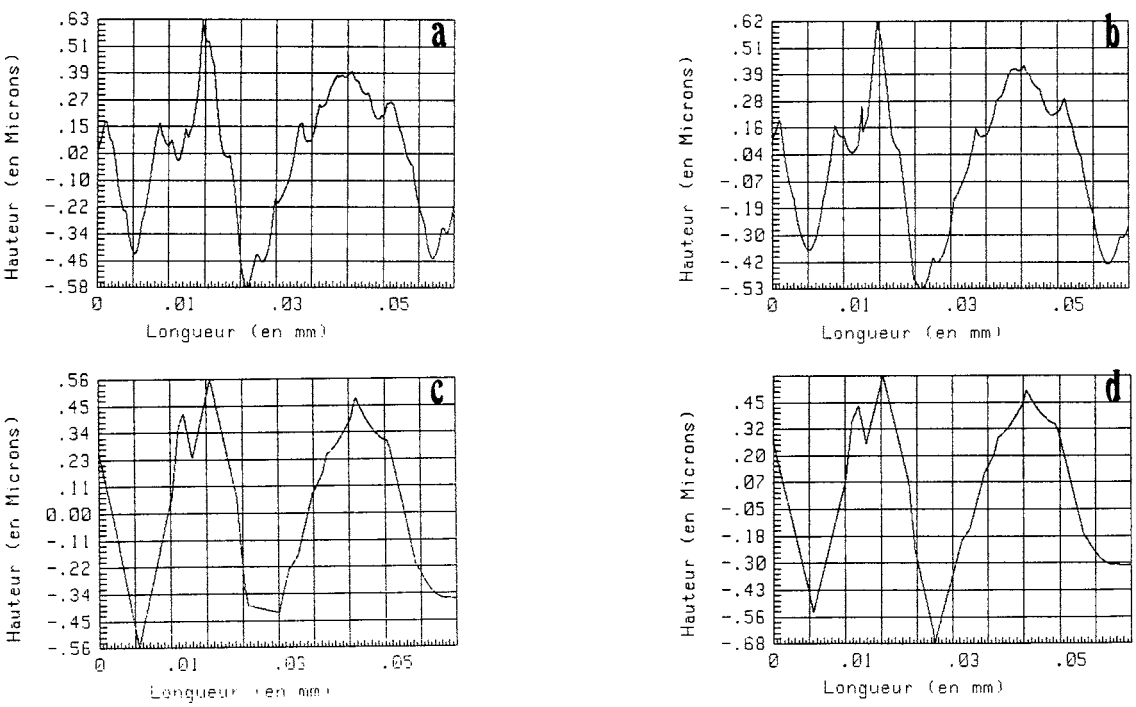


Figure 8 : Theoretical Z surface profiles a) $\varphi_0 = 0^\circ$, b) $\varphi_0 = 14^\circ$, c) $\varphi_0 = 34^\circ$, d) $\varphi_0 = 45^\circ$

MICROWAVE SYNTHESIZERS IN SPACE

A. Vulcan and M. Bloch

Frequency Electronics, Inc.
Mitchel Field, NY 11553 , USA

ABSTRACT

Synthesizers are used in satellites as local oscillators, reference signal generators for phase locked loops, sources for calibration systems, and many other purposes. The design drivers are small size, low weight, high reliability, and low power consumption. The frequency sources are used in redundant configurations to ensure that there are no single point failures. Space synthesizers have many self test telemetry outputs which provide indications of the unit health and operating mode.

They operate either directly from the spacecraft primary power bus or from a centralized DC to DC converter whose outputs are semi-regulated voltages. The long term frequency stability depends upon the reference source which may or may not be internal to the synthesizer. Atomic frequency standards or quartz crystal oscillators are used depending on mission stability requirements.

1. SYSTEM DESCRIPTION

Figure 1 shows the architecture of a typical redundant frequency generator used in a satellite. This configuration accepts N reference sources from crystal oscillators or atomic frequency standards, and delivers the reference signals to M synthesizers. The N X M summer/splitter is a passive device consisting of conventional transformer/balun type power combiners and splitters. Only one of the N sources is energized to provide the reference signals to all of the synthesizers. If crystal oscillators are used, the ovens of the standby sources are often powered, resulting in approximately 1.5 watts per oven of additional DC power. One of the synthesizers is chosen to provide the outputs signals with the others in cold standby. The synthesizers feed 1 X P splitters and M X 1 switches or power combiners. In some instances the output combiners are configured as quadrature hybrids which have two isolated inputs, either of which feeds two outputs. In some applications the M X P splitter/summer circuits are configured as a non-blocking switch matrix. This approach, which uses latching electromechanical switches for microwave distribution, has lower insertion loss since the passive power summer losses are eliminated. In the case of high power outputs this is certainly a desirable situation to minimize system DC power. Since most satellite frequency sources have multiple outputs at different frequencies, the output network is duplicated for each frequency. Cross-strapping between primary and redundant units is either performed externally using semi-rigid coaxial cable or two synthesizers are mounted in a single mechanical assembly and the cross-strapping is implemented with coaxial interconnections within the housing. This approach enhances reliability since several pairs of mating connectors are eliminated.

The frequency generator can be powered from the primary spacecraft power bus which may vary from 20 to 36 volts or from a secondary source which is semi-regulated with a typical variability of ±5%. Most modern generation systems operate from the primary spacecraft bus and have dedicated DC/DC converters with efficiencies of 80%. These converters, also redundant, use hybrid pulse width modulators and low drop linear regulators to maximize the efficiency. The operating voltages are routed to the reference sources and synthesizers by commandable latching relays which consume no power once set. Since noise and ripple generated in the power supplies can compromise the system performance, ripple rejection of 100 dB is required to minimize spurious generation. At lower frequencies this is accomplished in the regulator circuits and at ripple frequencies above 100 kHz discrete LC and RFI feedthru filters are employed.

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

2. SYNTHESIS TECHNIQUES

Many approaches are utilized to synthesize frequencies for spacecraft subsystems. These include direct and indirect methods using analog and digital techniques. For a given application the best suited architecture becomes apparent from a review of the noise, spurious levels and switching speed requirements as well as the number of frequencies and their spacing. Many applications require direct analog synthesis when phase noise must reach theoretically minimum levels. Direct synthesis has the noise advantage when the residual phase noise of the individual components is below that of the multiplied reference noise at various points in the system. Another advantage of direct synthesis is high speed switching operation although this requirement is usually not important in satellite applications. The major disadvantage is that a moderate increase in system complexity is required to keep spurious signals at low levels. Direct synthesis consists of frequency multiplication, mixing and division with bulk crystal, SAW, helical, cavity and waveguide filters used to provide high spurious signal rejection. The filters generally pass only a single frequency and their bandwidth is sufficiently narrow to ensure a high degree of unwanted signal rejection. Filter losses are compensated for with additional signal amplification. By appropriate selection of internal synthesizer frequencies and system configuration, sufficient separation can be maintained between the main signals and multiplier subharmonics, mixer harmonic distortion products, and LO leakage so that spurious signals remain below -80 dBc.

An example of a low noise synthesizer providing a 10.15 GHz output is shown in Figure 2. Two low noise Voltage Controlled Crystal Oscillators (VCXO) are phase locked to a common 10 MHz input reference. The 100 MHz VCXO is locked directly to the tenth harmonic of 10 MHz without the use of digital devices since divider noise would be enhanced by 60 dB. The 100 MHz phase locked oscillator also filters 10 MHz reference signal sidebands which would be enhanced 40 dB by the two ×10 multipliers. The fixed LO signal at 10 GHz feeds a double balanced mixer along with a 156.25 MHz signal derived from the frequency doubled output from a phase locked 78.125 MHz VCXO. This oscillator is digitally locked to a 125 kHz reference signal. This approach is used since the phase noise of the doubled 78.125 MHz phase locked loop output is not enhanced by subsequent multiplication. A 156.25 MHz crystal filter reduces the noise floor to -165 dBc/Hz. Due to the small separation between the two mixer inputs, an X-Band cavity filter is used to reduce LO leakage and the lower sideband to -75dBc. The phase noise of the 10.15 GHz source is shown in Figure 3. The RFTU consumes 5.5 watts of power and weighs 3.5 kilograms.

The synthesizer configuration used in the High Temperature Super-conducting Satellite Experiment (HTSSE) is shown in Figure 4. This application required a 4.25 to 5.25 GHz output commandable in 5 kHz increments. Long term stability requirement for a 10 year mission life and temperature excursion of -20 to +60°C was 1×10^{-7}. Consequently, a high stability 10 MHz ovenized crystal oscillator (OCXO) was used as the reference. The indirect synthesizer loop consisting of a 500 MHz to 750 MHz VCO locked to a 5 kHz reference signal provides one input to an upconverter mixer. The switching speed requirement for this application is 10 milliseconds so a single loop configuration was selected. The wideband nature of the synthesizer requires the use of switchable filter banks at the LO input and mixer output ports. Four LOs are generated from a 250 MHz comb spectrum. The 15th, 16th, 17th and 18th harmonic are selected with 5 pole cavity filters which reject adjacent subharmonics by 65 dB. The mixer output is filtered with four similar filters each covering a 250 MHz portion of the band. The HTSSE synthesizer consumes 15 watts of DC power and weighs 3.5 kilograms. Dual synthesizers are used in the application with cross-strapping external to the units. A DC to DC converter accepts primary spacecraft power covering a voltage range of 24V to 36V. The outputs are ±15V, +5 Vdc and +8 Vdc. Unregulated power is used for the oscillator ovens to maximize efficiency.

3. SYNTHESIZER PERFORMANCE CHARACTERISTICS

Table 1 tabulates the performances characteristics of ten satellite synthesizers having various mission requirements. The sources used as local oscillators on the ACTS, EUTELSAT, FSU and RFTU programs have very low phase noise at 10 Hz to 100 kHz offsets with spurious signals typically below -70 dBc. The 5 MHz and 10 MHz references, having daily stabilities of 2×10^{-12}, results in maximum drift of 1×10^{-8} for a 10 year mission. This maximum frequency excursion includes the effects of radiation, temperature, and component drift. Missions which require only 5×10^{-6} lifetime drift enable higher frequency references such as 50 MHz or 100 MHz to be used. Higher input frequencies reduce multiplier complexity, reduce spurious signal levels, and generally result in less complex source.

The LFA and INMARSAT requirements for selectable low spurious output signals require the generation of multiple microwave frequencies using direct synthesis techniques with switching matrixes selecting the appropriate output frequencies. For these applications, the switching speed is typically 10 milliseconds and high isolation in the switches is required to reduce spurious signals.

The RGA has a phase noise floor of -172 dBc/Hz which is maintained at offset frequencies above 1 MHz. The low spectral density of the 500 MHz output requires low residual noise multipliers and the inclusion of SAW resonator filters to reduce noise and spurious signals.

4. TELEMETRY

The health of the various subsystems which comprise the frequency generator are monitored and fed to the ground via telemetry subcarriers. Parameters which are often monitored are as follows.

MEASURED	TYPE	LEVEL
RF Output Level	Analog	0 to 2 Volts
Phase Locked Loop Status	Bistate	TTL
System Configuration	Relay Contact	Open or Closed
Critical DC Voltages	Analog	0 to 5 Volts
DC Power	Analog	0 to 5 Volts
Frequency Selection	Digital	TTL
Temperature	Analog	Resistance

Telemetry Outputs from the frequency generator are buffered and protected to ensure that certain maximum levels will not be exceeded and output perturbations are not reflected back to the system to cause incorrect operation.

5. CONCLUSION

A brief overview of synthesizers used in the space environment has been presented. Included are examples of both direct and indirect synthesizers whose designs are driven by specific requirements such as low phase noise, good long term frequency stability or multiple frequency stepping capability. In most cases direct synthesis techniques are employed since low phase noise is the most important feature.

TABLE 1. SYNTHESIZER CHARACTERISTICS

MISSION	TYPE	OUTPUT FREQUENCY	REFERENCE FREQUENCY	LONG TERM FREQUENCY STABILITY	SPURIOUS SIGNALS dBc	PHASE NOISE dBc/Hz	DC POWER WATTS	WEIGHT KG
ACTS	DIRECT	3.24 GHz, 4.13 GHz 5.4 GHz, 6.48 GHz	5 MHz EXTERNAL	1×10^{-8}	-85	-65 at 10 Hz -135 at 100 kHz	20.0	4.7
HTSSE	INDIRECT 5 kHz STEPS	4.25 - 5.25 GHz	10 MHz INTERNAL	1×10^{-7}	-60	-40 at 100 Hz -95 at 100 kHz -125 at 1 MHz	15.0	3.5
EUTELSAT	DIRECT	1.50 GHz, 2.55 GHz, 3.30 GHz	50 MHz INTERNAL	1×10^{-6}	-80	-60 at 10 Hz -135 at 100 kHz	8.0	3.0
RDU	DIRECT	50 MHz, 52.5 MHz, 400 MHz	52.5 MHz EXTERNAL	5×10^{-6}	-85	-107 at 10 Hz -145 at 100 kHz	4.7	1.8
RGA	DIRECT	100 MHz, 400 MHz, 500 MHz	25 MHz EXTERNAL 100 MHz INTERNAL	5×10^{-6}	-100	-106 at 10 Hz -155 at 100 kHz -172 at 1 MHz	4.0	2.5
LFA	DIRECT SELECTABLE OUTPUTS	5.5 GHz - 16 GHz	500 MHz EXTERNAL	5×10^{-6}	-90	-70 at 10 Hz -140 at 100 kHz	19	9.5
INMARSAT	DIRECT SELECTABLE OUTPUTS	1.95 GHz, 2.0 GHz	100 MHz INTERNAL	1×10^{-5}	-70	1^{o} RMS, 10 Hz to 100 kHz	2.5	0.2
REFCAL	DIRECT AND INDIRECT	90 MHz - 2.1 GHz	100 MHz EXTERNAL	5×10^{-6}	-60	-65 at 10 Hz -140 at 100 kHz	8.0	5.3
FSU	DIRECT	20.5 GHz	50 MHz INTERNAL	5×10^{-6}	-70	-60 at 10 Hz -130 at 100 kHz	4.0	1.5
RFTU	DIRECT AND INDIRECT	10.15 GHz	10 MHz EXTERNAL	1×10^{-8}	-75	-60 at 10 Hz -120 at 100 kHz	5.5	3.5

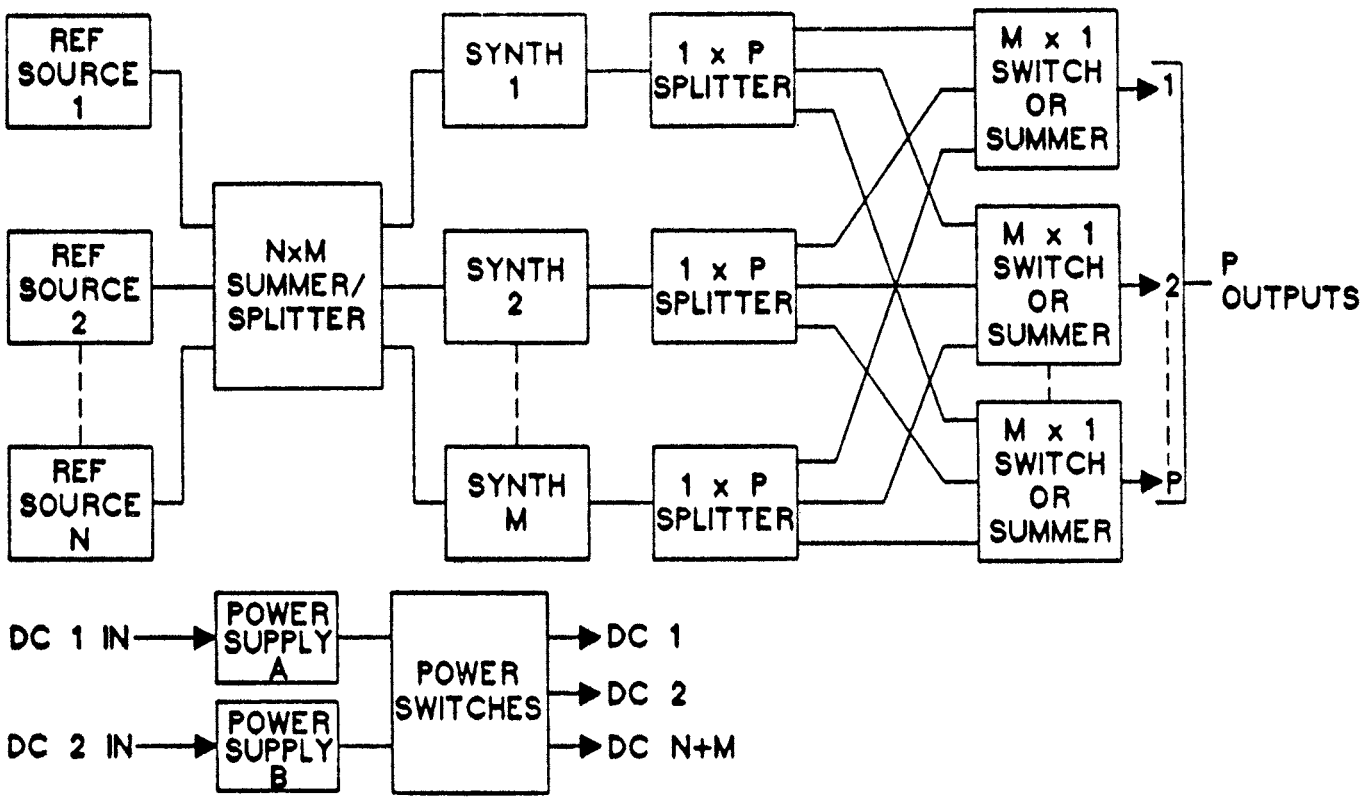


Figure 1. NXMXP Redundant Frequency Generator

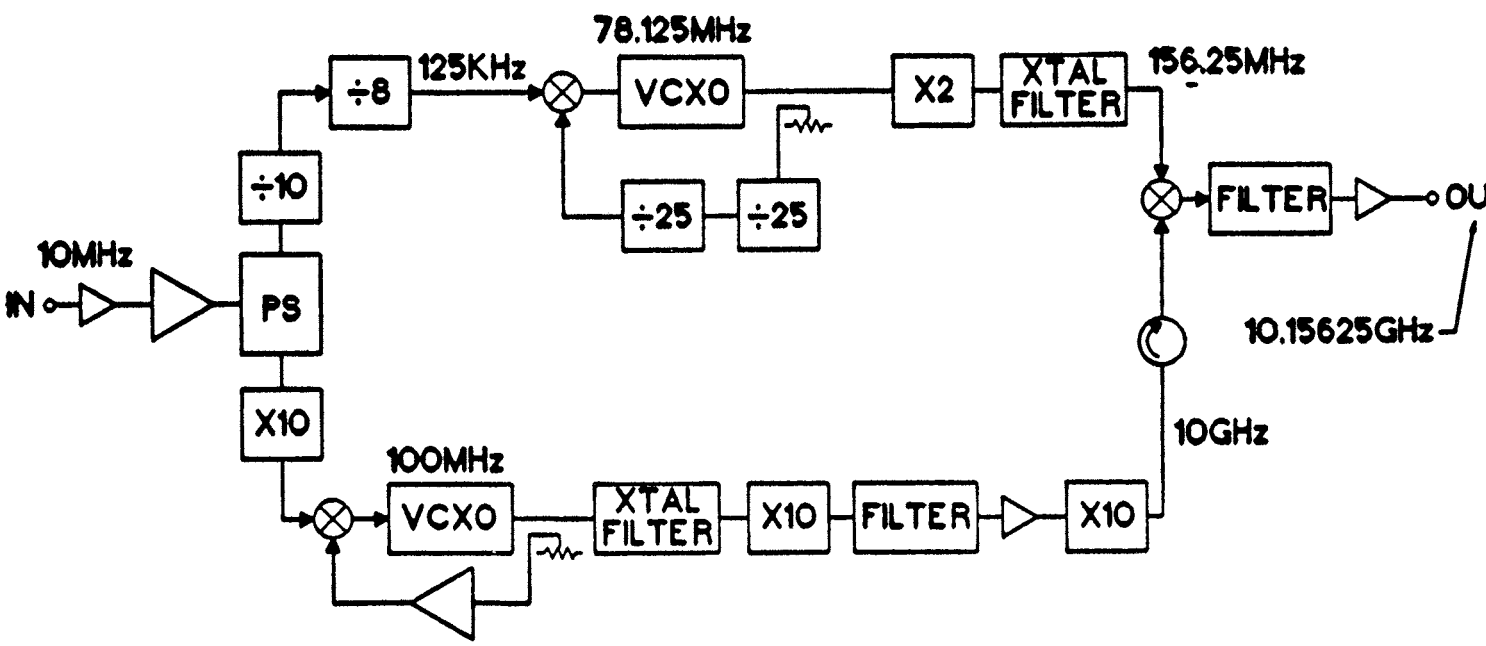


Figure 2. RFTU Local Oscillator

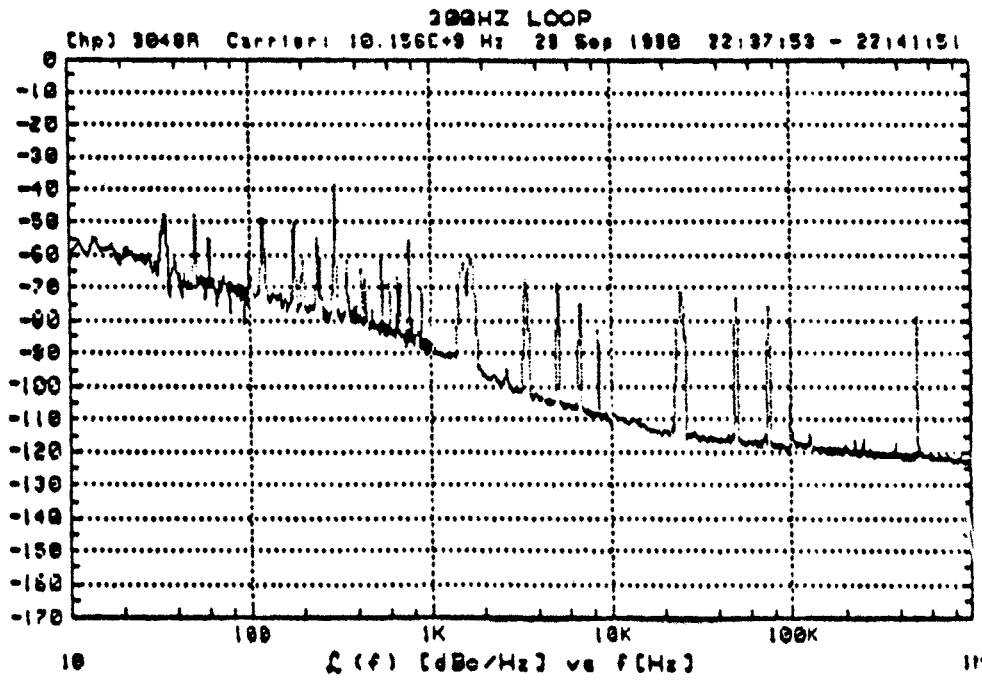


Figure 3. RTFU Phase Noise

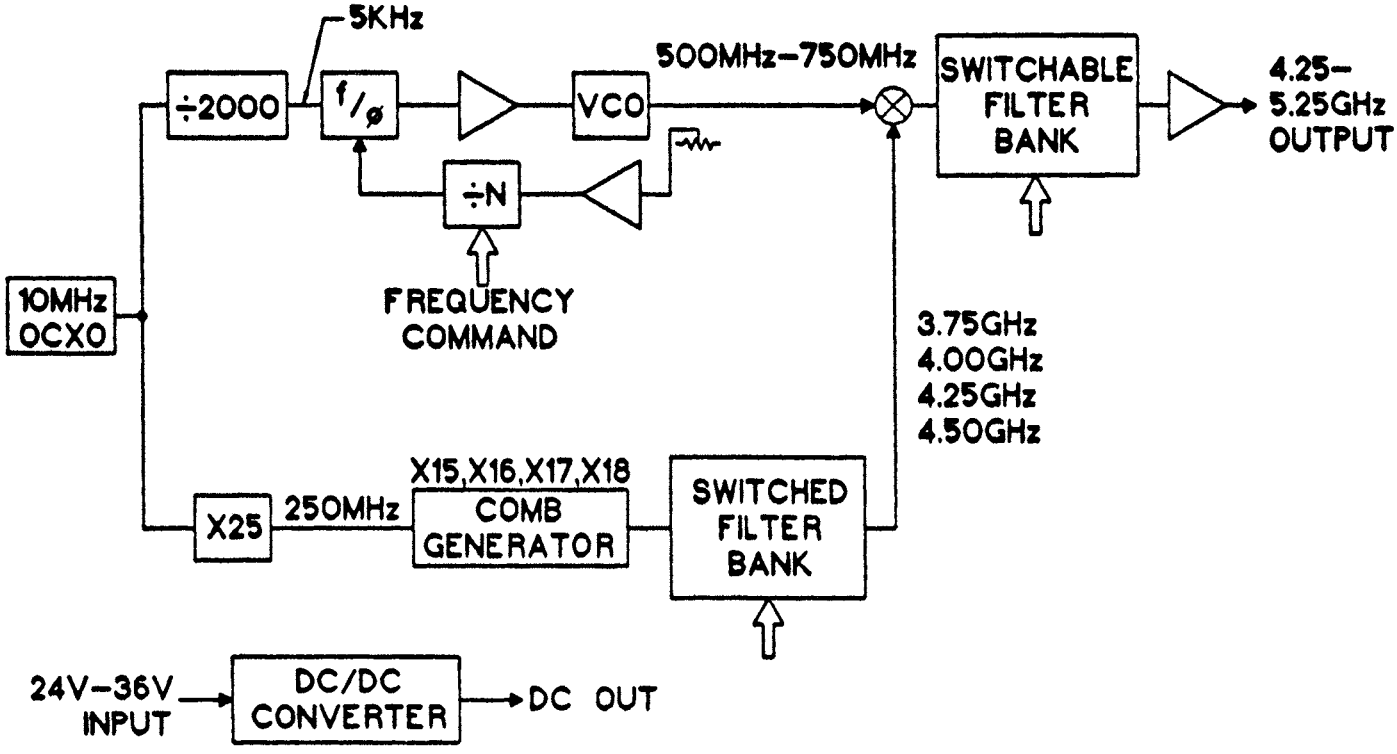


Figure 4. HTSSE SYNTHESIZER

VERY WIDE BANDWIDTH HF CRYSTAL FILTERS

P. K. Wall.

Peter Wall ElectroAcoustics, Penrith, U.K.

ABSTRACT

Recently, specialized communication equipments have demanded highly selective filters of 5% BW in the 20 MHz region. Inharmonic modes of fundamental AT crystals appear above, but relatively close to, the main mode. These unwanted modes are troublesome in the filter passband because energy once lost cannot be replaced without amplification, which may be unacceptable for noise reasons.

This paper presents measurements on 1 MHz wide crystal filters at 20 MHz with less than 6 dB insertion loss, 1 dB unwanted response heights and with high selectivity and stopband attenuation.

Calculations show that unwanted responses in the 100,00 ohm region are required. Measurements of crystals made for these filters are presented. They show that these levels of unwanted response are achievable. Comparisons with a high accuracy reference method show acceptable correlation.

Keywords: Wideband crystal filters, unwanted response measurement.

1. CRYSTAL FILTER DESIGN

1.1 Constraints

It has been known for a long time (Ref. 1) that the maximum bandwidth of crystal filters is limited by the capacitance ratio of the piezoelectric resonators. Without the addition of inductors, the fractional bandwidth of crystal filters constructed of sections each with 2 crystals is limited to 1/r approximately where r is the capacitance ratio of the resonators.

If inductors are used, the ratio increases to $1/\sqrt{r}$.

It might be thought that increasing the number of crystal resonators in each section would increase the available bandwidth. This is not true however.

Low frequency resonators are available with capacitance ratios of 125 which means bandwidths of 9% are possible theoretically.(Ref. 2).

At 20 MHz, fundamental AT crystals with a plate capacitance ratio of 200 must be used. The practical capacitance ratio is nearer 250, due to the metal enclosure. This means that the maximum bandwidth available is approximately 6.3%.

This simple criterion for maximum bandwidth was derived on the basis of the cut-offs defined from image parameter filter design theory. Use of Insertion Loss techniques allows for a little extra, the amount depending on the insertion loss allowed at the band limits and the permissible passband ripple. (Ref. 3)

1.2 Filter Requirements

The selectivity and passband requirements for the filters are shown in Fig. 1.

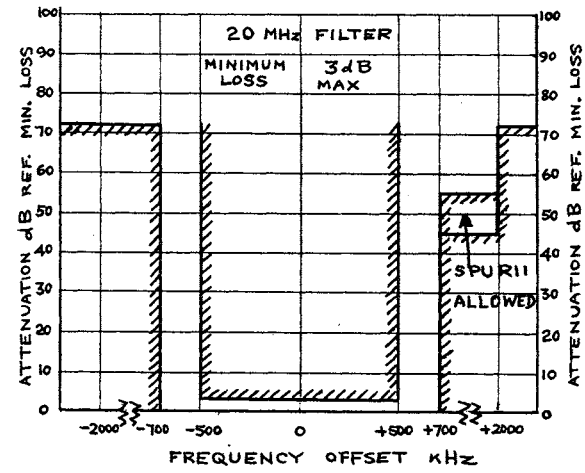


Fig. 1 Filter Requirements

1.3 Crystal Filter Design

The requirements demanded a circuit with 8 crystals if adequate margin was to be provided. There were two possibilities. The first was to use 4 sections each with 2 resonators. The second was to use 2 sections each with 4 resonators. Ref. 4 describes a method of designing widebandwidth filters with multiple resonators. The filters constructed have used the second method. The basic circuit used is shown in Fig. 2.

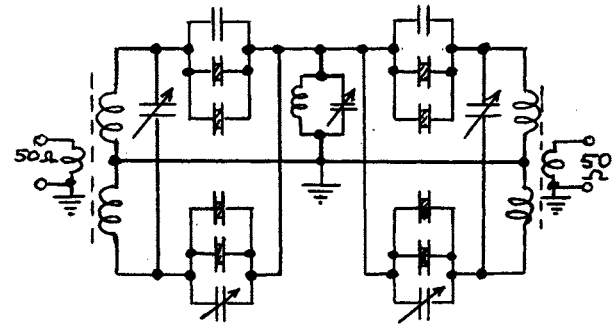


Fig. 2. Basic Filter Circuit

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

Care is necessary to minimize the capacitance at the input and output of each section of the filter and at the junction between the two sections.

This requirement affects the transformers in particular. They must be designed for low self-capacitance as well as tight coupling between the two halves of the high impedance winding.

2. CRYSTAL RESONATOR DESIGN

2.1 Representation of Unwanted Responses

These can be represented by the circuit of Fig. 3.

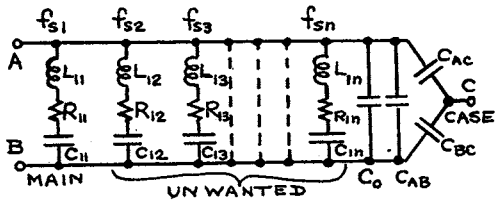


Fig. 3 Equivalent Circuit for Resonator with Unwanted Responses

Ref. 4 describes unwanted responses and their effect on filters in detail.

It may be argued, rightly, that there may be some coupling between the resonances, including the main resonance. Irrespective of how many of the resonances are coupled together and to what degree, the topology of the circuit of Fig. 4. remains valid. However the values of the elements in each arm of this circuit may vary as the coupling changes from one crystal sample to another and one temperature to another.

2.2 Unwanted Responses in the Passband

The wide bandwidth of the filter means that it is impracticable to avoid some of the crystal unwanted responses falling in the filter passband. If an unwanted response from a crystal in the first filter section coincides with one in the second section, the result is to increase the height of the passband spike. One possibility of minimizing passband spikes is therefore to select or design crystals to avoid coincidences.

2.3 Unwanted Responses in the Stopband

The requirements become less onerous as the frequency becomes more remote from the passband. In cases of difficulty, there are two strategies:

a) Design or select crystals for the two sections which have non-coincident unwanted responses.

b) Hide an unacceptable unwanted response in the transition band between passband and stopband.

2.4 Dynamic Parameters

The very wide bandwidth means that the Q of the crystal main mode need only be 10,000 or a little less, which presents no problem.

The dynamic capacitance C1 can be chosen to suit the crystal designer providing it is not too small which would give difficulties due to the resultant high impedance of the filter. In practice there is quite a sharp upper limit, beyond which it is impracticable to achieve the very high unwanted response resistances required. Values around 5 fF provide acceptable results.

2.5 Crystal Enclosure

HC-45/U holders were used. Three pin bases with centre pin connected to case are available to provide an earthed connection if desired, instead of soldering a wire to the case.

3. FILTER ENCLOSURE

Dimensions of the case, excluding pins are:

Width x Length x Height: 25 x 86 x 16.5 mm nominal

4. MEASURED FILTER PERFORMANCE

Measurements were produced on a point-by-point basis to ensure that the correct excursions of unwanted responses were recorded.

4.1 Passband Performance

The passband characteristics of two filters are shown in Fig. 4 and Fig. 5.

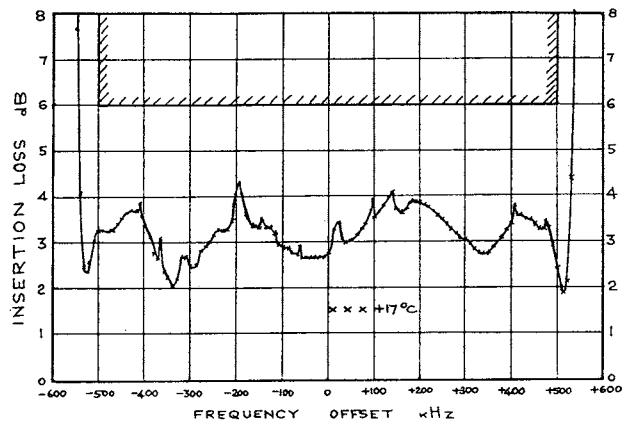


Fig. 4 Passband Characteristic of 20 MHz Filter Sample A

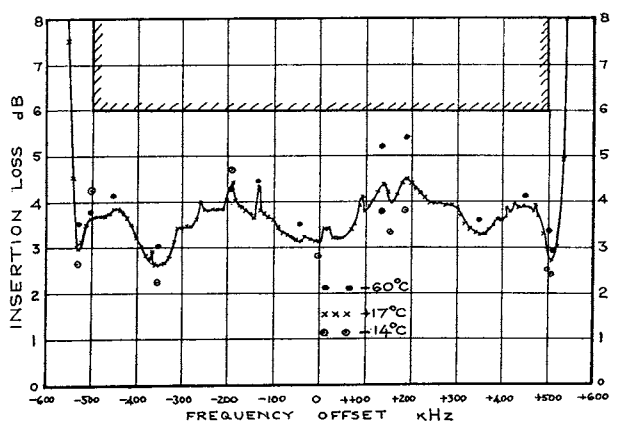


Fig. 5 Passband Characteristic of 20 MHz Filter Sample B

It will be seen that the minimum insertion loss is less than 3 dB in both cases. The maximum height of spike relative to its immediate surroundings is only 1·0 dB. The maximum excursion over the 1 MHz bandwidth is only 2·5 dB.

4.2 Stopband Performance

The stopband characteristics of two filters are shown in Fig. 6 and Fig. 7.

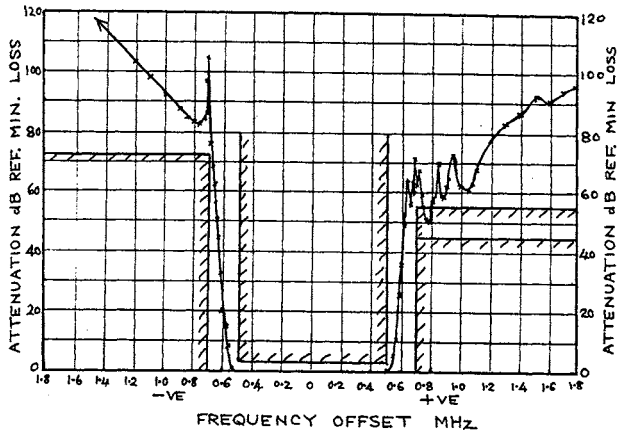


Fig. 6 Stopband Characteristic of 20 MHz Filter Sample A

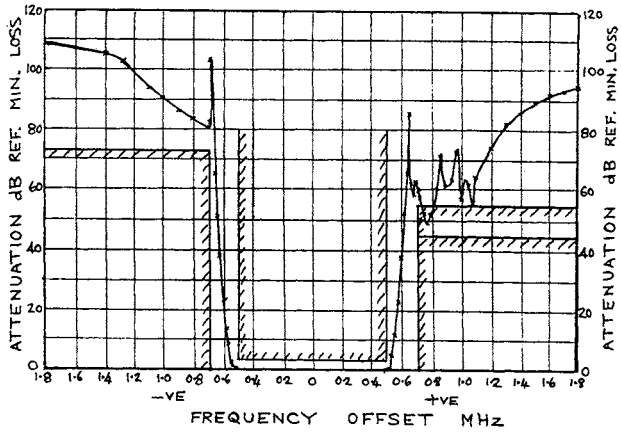


Fig. 7 Stopband Characteristic of 20 MHz Filter Sample B

The lower stopband is limited more by earth currents in the enclosure than by the inherent attenuation of the filter sections.

5. MEASUREMENT of CRYSTAL UNWANTED RESPONSES

5.1 Hybrid Measuring Method

Basically a transmission method is used with means to tune out the shunt reactance of the crystal under test in the neighbourhood of each resonance. The circuit used is shown in Fig. 8.

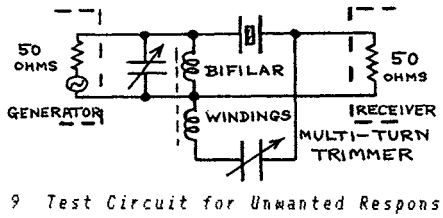


Fig. 9 Test Circuit for Unwanted Responses

This circuit is described in more detail in Ref. 4

The loss expected from an unwanted response resistance Rn in this circuit is:

$$20 \log_{10} (1 + Rn/100) \text{ dB} \qquad \text{(Eqn. 1)}$$

For an unwanted response of 100,000 ohms, the loss is 60 dB and for 200,000 ohms is 66 dB.

The measured performance of crystals from various crystal manufacturers is shown in Table 1.

Table 1: Measured Unwanted Response Performances

CRYSTAL	UNWANTEDS + kHz ref. Main Resp. (dB) in 50/50 ohm circuit					
Ser. No.	1	2	3	4	5	6
1	+129 (55.8)	+170 (64.3)	+181 (64)	+223 (59.8)	+298 (71.5)	+390 (73.5)
2	+109 (54.5)	+188 (68)	+252 (68)	+371 (64)	+558 (62)	
3	+170 (46.3)	+191 (72)	+260 (59.4)	+283 (57)	+384 (+63.7)	+604 (67)
4	+170 (70)	+278 (65)	+372 (64)	+556 (64)	+703 (65)	
5	+175 (73)	+214 (72)	+264 (65)	+386 (66)	+713 (63)	
6	+205 (69)	+315 (62)	+529 (65)	+753 (65)	+890 (71)	+982 (71)
7	+135 (70)	+203 (72)	+284 +72)	+397 (75)	+549 (66)	+754 (67)

NOTE: Crystals 1, 2 and 3 represent "poor" crystals and 4, 5 6 and 7 represent "good" crystals in the context of the performance required by these very wide filters.

5.2 Inaccuracies of the Hybrid Measuring Method

The main problem is that to measure the resistance of a particular unwanted response, the combined reactance of all other resonances including the main one must be tuned out in its neighbourhood. When the unwanted responses are close to the main resonance, this raectance changes rapidly with frequency. This means that repeated adjustment of the balancing capacitor is necessary as various unwanted resonances are explored. The correct adjustment is when the minimum loss at the unwanted response is maximized with the final adjustment of the balancing capacitor.

Errors in the method are therefore:-

a) Errors in the balancing capacitor tuning. Errors always tend to give a lower than actual resistance value.

b) Errors in the 50 ohm terminations and the shunt impedance of the hybrid transformer.

c) Resistance unbalance in the hybrid. A resistance balance of 80 db is necessary. This effect can be checked by reversing the crystal and balaancing capacitor.

d) Any dissipation across the crystal due to seals etc is automatically attributed to the unwanted response. As in 1), this results in a lower than actual resistance measurement.

5.3 Use of Reference Method for Unwanted Responses

The means for accurate measurement of crystal resonators is described in Ref.5. This method forms the basis for the UK Standard Measuring Method for Quartz Crystal Resonators.

Unwanted responses from 3 crystals were measured by this reference method. The results in comparison with the hybrid method are shown in Table 2.

Table 2: Comparison of Unwanted Response Test Methods

METHOD	UNWANTED POSITION	UNWANTED FREQ. kHz ref. main	Rn kohms	Ln H	Q
CRYSTAL 1					
Hybrid	1	+129	61.6		
UK Ref.	1	+130	61.4	1.23	2,496
CRYSTAL 2					
Hybrid	1	+109	53.1		
UK Ref.	1	+108	59.27	5.527	11,882
CRYSTAL 3					
Hybrid	1	+170	20.7		
UK Ref.	1	+170	24.16	15.7	83,000
Hybrid	3	+260	93.3		
UK Ref.	3	+261	109.2	1.518	1,784
Hybrid	4	+283	70.8		
UK Ref.	4	+283	90.96	2.275	3,214

Correlation is quite good in most cases. Where there is a significant difference between the two methods, the hybrid method gives the lower unwanted resistance value, probably because of the errors mentioned.

5.4 Use of Reference Method for Main Responses

The opportunity was taken to measure the main response parameters using this reference method and compare these with measurements carried out by the original suppliers. These comparisons are shown in Table 3.

Table 3: Comparison of Main Resonance Parameters

XTAL No.	MEASUREMENT SOURCE	PARAMETER C1 (fF)	L1 (mH)	R1 (ohms)	Q
1	UK Ref Std	(4.4636)	14.73	17.79	102,109
	SUPPLIER	not available			
2	UK Ref Std	(5.6023)	11.12	75.20	18,743
	SUPPLIER	5.62	11.09	75.31	18,656
3	UK Ref Std	(5.6696)	10.98	6.813	204,311
	SUPPLIER	5.68	10.95	6.15	225,701
4	UK Ref Std	(5.6286)	11.06	28.73	48,801
	SUPPLIER	5.64	11.03	28.50	49.060
5	UK Ref Std	(5.5782)	11.16	50.86	27,810
	SUPPLIER	5.59	11.13	50.53	27,921
6	UK Ref Std	(5.1236)	12.15	52.12	29,547
	SUPPLIER	5.12	(12.158)	50.3	30,635
7	UK Ref Std	(4.9028)	13.41	15.48	106,802
	SUPPLIER	4.89	(13.445)	15.10	109,811

() Means that the value was calculated from the other parameters supplied.

6. SUMMARY

Quartz Crystal Filters of approximately 5% bandwidth in the 20 MHz region have been shown to be feasible. Crystal resonators for these filters have been made successfully despite the rather high values of unwanted responses which are required.

7. REFERENCES

1. Mason W.P., Electrical Wave Filters Employing Quartz Crystals as Elements, *BSTJ, Vol. XIII July 1934* 405-452.

2. O'Meara T.R., The Synthesis of a Class of Wide-band Lattice Crystal Filter with symmetrical Insertion Loss Characteristics, *IRE Trans. on Circuit Theory, June 1960 121-127.*

3. Bown G.C.S., Formulae for the Wide-band Crystal Filter, *Electronic Engineering July 1963* 449-454

4. Horton W.H., Quartz Crystal Unwanted Mode Response and Measurement, *IEEE Trans. Sonics & Ultrasonics, Vol.SU-13, No.4, Oct 1966 111-115.*

5. Peach R.C. & Morris S.E., A System for Precision and Bipoles, *Proc. 39th Annual Symposium on Frequency Control 1985 IEEE 85CH2186-5* 527-534

8. ACKNOWLEDGMENT

The author is indebted to Mr. Angus Haston and Mr. Simon Morris of GEC-Marconi Ltd. Hirst Research Centre for their kind co-operation over measurement of crystal parameters using the Hirst Research Centre installation of the U.K. Standard Measuring Method for Quartz Crystal Resonators.

MEASUREMENT RESULTS OF THE CRYSTALLOGRAPHIC ORIENTATION OF SC CUT ROUND PLATES

K. Weiss

Tele- and Radio Research Institute, Poland

ABSTRACT

The temperature dependence of quartz resonator is one of very important parameters. This parameter for choosen mode of vibrations depends mainly of a quartz plate crystallographic orientation In the case of doubly rotated cuts as SC or LC the accurate crystal orientation measurements are very difficult and represent a considerable technical problem. For this measurements are used special goniometers with a very complicated instrumentation, but even with this equipment the measurements are difficult and not accurate. Very comfortable in use is the measurement method with a tilting crystal holder used with Secasi goniometers. But this method can be used only for plates with well definied edges, and its inaccuracy can result in measurement errors over 30 sec. At the Tele- and Radio Research Institute there we have been using over ten years the double reflection Secasi goniometer type PDR-1 This goniometer is equiped with a tilting holder for measurements of doubly rotated cuts. We have carried out investigations and analysis of the main measurement errors of crystallographic orientation of doubly rotated cuts. On the base of this results we developed a new procedure for this measurements.This procedure makes possible the crystal orientation measurements of rectangular and round plates, using the standard double reflection goniometer with simple additional equipment. Using this procedure we carried out the crystallographic orientation measurements of SC cut plates with reference to selected atomic planes. This paper presents a short review of factors involving the measurements inaccuracy measurement procedure, as well as measurement and calculation results for the SC round plates. The achieved measurement accuracy of crystallographic orientation of this plates was better than 5 sec. (for selected atomic planes). The time consumption of one plate orientation measurement was not longer than 5 minutes.

1. INTRODUCTION

The most important problem of crystallographic orientation measurements in the case of doubly rotated cuts is such, that usualy no one of atomic planes in quartz plate is paralel to plate edges and is not perpendicular to the plane of the goniometer. In consequence the beam reflected from the atomic plane can not enter the ionisation chamber (Fig.1). To bring this beam to ionisation chamber we must turn the quartz plate around the ω angle (Fig 2) or tilt it around the γ angle (Fig.3). In Secasi PDR type goniometers there is used the second method.

The investigations carried out at the Tele- and Radio Research Institute showed that using proposed by Secasi measurement method we make many mistakes. The main of them are the inaccuracy of the base edges of quartz plates, and the plate holder tilting error This two factors can result in crystallographic orientation errors over 1 minute. Additional, the quartz plates lapping process after crystallographic orientation measurements, can change their orientationin in the range of 4 minutes. In the case of SC cut we must keep crystallographic orientation of quartz plates in very close tolerances of $\Delta\phi \pm 5$ min. and $\Delta\Theta \pm 30$ sec. Comparing this two informations we can see that we have only one way to achieve good thermal characteristics of quartz resonators - the accurate measurements of crystallographic orientation of round plates after lapping process, and selection the plates with the proper crystallographic orientation.

2. MEASUREMENT PRINCIPLES

The diffraction intensity dependence of the ω angle rotation of the quartz plate, and the γ angle rotation of crystal holder ispresented in Fig.4. We can see that if γ angle is lover than the angle γ_0(between an atomic plane and the quartz plate surface) rotating quartz plate on crystal holder we obtain two symmetrical peeks of reflected beam intensity. If γ angle is higher than γ_0 we obtain one reflection peek lover than for γ_0 but in the same position of angle ω as in the case of $\gamma = \gamma_0$ (Fig.5).
Using this dependence I developed the following measurement procedure:

1. Positioning the crystal holder at γ angle lover than expected value.
2. Rotating slowly the quartz plate on the holder to obtain two peeks and positioning the plate between this peaks.
3. Turning the crystal holder over the maximum of diffraction to position when reflected beam in ionisation chamber yet appears
4. Rotating the quartz plate on the holder to obtain a peek extremum
5. Turning the crystal holder to obtain a maximum of diffraction peek and the γ angle reading. This angle corresponds to the angle γ_0 between choosen atomic plane and the quartz plate surface).

The calculated positioning (ω angle) of quartz plate accuracy in this case is better than 1 deg.

2. THE CHOICE OF ATOMIC PLANES

For the crystallographic orientation calculation we must measure diffracttion angles respectively to two or more atomic planes. Generaly, it is assumed thet the best for crystallographic orientation measurements are atomic planes close to the cut orientation. The position of atomic planes near the SC cut are showed in Fig.6. For the atomic planes 23.3 and 23.4 diffraction measurements are not possible using the Secasi goniometer because a big Bragg angle (maximum value of Bragg angle on this goniometer is 52 deg.).

Diffraction angle measurements accuracy depends on diffraction intensity and repetibility of maximum diffraction angle adjustment. One crystallografic orientation measurement is composed of two or more diffraction angles measurements. The accuracy of this measurement depends, in a significant degree, on the angle between diffraction angles masurement directions. We can estimate the minimal error of crystallographic orientation measurement when we know the repetibility of maximum diffraction angle adjustment and the angle between diffraction measurement directions. This information helps us to choose the best pair of atomic planes for given crystallographic orientation measurement. The main parameters of selected atomic planes for SC cut crystallographic orientation measurements are compiled in Tab.1

Tab.1 Main measurement parameters for SC cut (ϕ=22°10', Θ=34°05')

Atomic plane	Bragg angle	Orientation ϕ	Orientation Θ	γ	ω	$\Delta\gamma$ for $\Delta\omega$=1°	extremum error
11.1	20°09'	30°00'	24°37'	11°48'	-37°20'	-7"	+/-1"
11.2	25°05'	30°00'	42°16'	10°14'	34°35'	-6"	+/-1"
12.2	33°53'	19°06'	30°46'	4°13'	38°40'	-2"	<1"
12.3	39°58'	19°06'	41°45'	8°02'	-16°33'	-4"	<1"
13.2	45°35'	13°54'	23°35'	12°44'	36°42'	-7"	+/-3"
13.3	51°15'	13°54'	33°14'	6°56'	85°13'	-4"	+/-2"
22.3	47°54'	30°00'	34°17'	6°29'	86°01'	-4"	+/-2"

We can combine this atomic planes in pairs for the best measurement parameters. This pairs are presented in Tab.2 with calculated errors of ϕ and Θ angles measurement.
Tab.2. Calculated crystallographic orientation measurement errors

Paires of planes	ω_1-ω_2	$\Delta\phi$	$\Delta\Theta$
11.1 - 11.2	71°55'	+15/-2	+5/-5
11.1 - 12.2	76°00'	+8/-1	+0/-5
11.1 - 13.2	74°02'	+11/-11	+3/-11
11.1 - 13.3	42°32'	+3/-6	+3/-14
11.1 - 22.3	43°21'	+7/-3	+5/-11
12.2 - 13.3	46°32'	+3/-7	+5/-4
12.2 - 22.3	47°21'	+8/-3	+2/-8
12.3 - 11.2	51°05'	+10/-7	+5/-0
12.3 - 12.2	55°13'	+0/-10	+3/-1
12.3 - 13.2	53°15'	+5/-19	+4/-4
12.3 - 13.3	78°15'	+2/-7	+5/-2
12.3 - 22.3	77°27'	+7/-3	+6/-0

From Tab.2 we see that for crystallographic orientation measurements of SC cut the best are atomic planes pairs 12.3 - 12.2, 11.1 - 12.2 and 12.3 - 11.2. We can also use another pairs of planes with a good accuracy.

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

3. EXPERIMENTAL RESULTS

I prepared 10 plates of SC cut of diameter 14 mm and thickness 1.400 mm. This plates were lapped on Hofman planetary lap machine PR-2-32 type with 5 μm corundum abrasive. Next this plates were etched 30 min. with saturated $(NH_4)HF_2$ solution in room temperature. The flatness of this plates and the flatness of lapping plates of the lap machine were checked by means of cyclic interferometer. In both cases the flatness was better than 0,5 μm. Next I measured for all plates diffraction angles against atomic planes presented in Tab.1. For this measurements I was using above described procedure. Using presented in Tab.2 pairs of atomic planes for each pair of planes were calculated crystallographic orientations of each plate. Then for each plate mean values of ϕ and Θ angles and differences to this values were calculated. This results (rounded to whole seconds) are presented in Tab.3.

Tab.3 Measurements and calculations results

No	Pairs of planes																							
	11.1-11.2		11.1-12.2		11.1-13.2		11.1-13.3		11.1-22.3		12.2-13.3		12.2-22.3		12.3-11.2		12.3-12.2		12.3-13.2		12.3-13.3		12.3-22.3	
	Δφ	Δϑ	Δφ	Δϑ	Δφ	Δϑ	Δφ	Δϑ	Δφ	Δϑ	Δφ	Δϑ	Δφ	Δϑ	Δφ	Δϑ	Δφ	Δϑ	Δφ	Δϑ	Δφ	Δϑ	Δφ	Δϑ
1	-8	+3	-6	+5	-3	+7	+14	+18	-19	-4	+17	-10	-22	+16	+3	-3	+5	-3	+11	-1	+16	0	-20	-9
2	+6	+2	+5	+1	+9	+4	+17	-2	-20	-15	+19	-8	-23	+20	+14	-3	+12	-3	+18	-2	+18	-2	-21	-11
3	0	-3	-2	-4	+4	0	+11	-9	-15	-44	+13	-14	-72	+42	+9	-8	+6	-9	+14	-7	+13	-7	-66	-27
4	+4	+3	+2	+2	+7	+5	+12	-27	-18	-11	+13	-6	-20	+16	+16	-3	+11	-5	+19	-3	+13	-4	-18	-12
5	+8	-1	+8	-3	+10	0	+13	+1	-20	-19	+14	-6	-23	+19	0	+2	+1	+3	+4	+3	+12	+5	-21	-3
6	+5	+2	+5	+1	+8	+4	+12	+6	-26	-18	+13	-4	-29	+24	+13	-4	+12	-4	+18	-3	+13	-4	-26	-13
7	-5	-3	+1	+1	+2	+1	+6	+4	-29	-18	+7	-3	-32	+23	-5	-3	+3	-1	+5	-1	+7	0	-30	-9
8	-4	+3	-2	+5	+6	+9	-3	+4	-25	-10	-3	+6	-28	+22	+9	-4	+20	-2	+22	-1	-1	-7	-25	-12
9	-1	+2	-1	+1	+4	+5	+10	+8	-12	-8	+10	-7	-18	+12	+8	-4	+7	-4	+5	-2	+11	+5	-15	-12
10	-1	0	0	+1	+2	+2	+10	+7	-34	-8	+12	-7	-16	+12	-3	+1	-1	+2	+1	+2	+11	+5	-15	-12
$\sum \frac{\Delta}{n}$	+0.4	+0,8	+1,0	+1.0	+4,9	+3.7	10,2	+1,0	21,8	15,5	11,5	-5,9	28,3	20,6	+6,4	-2,9	+7,6	-2,2	11,7	-1,5	11,3	-1.5	25,7	11,8
$\sum \frac{\Delta i}{n}$	4,2	2,2	3,2	2,4	5,5	3,7	10,8	8,6	21,8	15,5	12,1	7,1	28,3	20,6	8.0	3.5	7.8	3,2	11,7	2,5	11.5	3,5	25,7	11,8

Mean values of orientation		
No		
1	22°09'08"	34°03'21"
2	22°09'29"	34°04'10"
3	22°09'42"	34°04'27"
4	22°08'55"	34°04'05"
5	22°09'43"	34°04'05"
6	22°09'20"	34°04'14"
7	22°09'41"	34°04'17"
8	22°08'42"	34°04'28"
9	22°09'06"	34°04'23"
10	22°09'30"	34°04'22"

We see that the experimental results are in a good agreement with prevously calculated values. Only for 22.3 atomic plane we obtain significant worse results.

4. CONCLUSIONS

Experimental results without doubt showed that proposed crystallographic orientation measurement method presented above gives a good measurement accuracy for round SC plates. For any given cut we can calculate the expected measurement errors and to chose the best atomic planes for accurate measurements.
Proposed measurement method have one inconvenience: during quartz plate rotation on the crystal holder can appear a little scratches on a plate surface. Practicaly we use this plates for plano-convex crystal elements. In this case we can make measurements on the plate side which will be designed for convex shape. In the case of plano-paralel plates we can make measurements before last lapping operation, though we must remember that measurement accuracy depends upon crystal plate thickness. Practicaly we can measure with a good accuracy crystal orientation only for plates of thickness higher than 0.2mm.

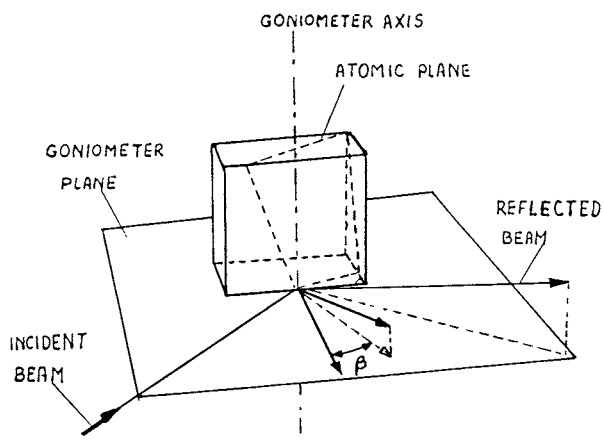


Fig.1 X-ray diffraction on doubly rotated cut

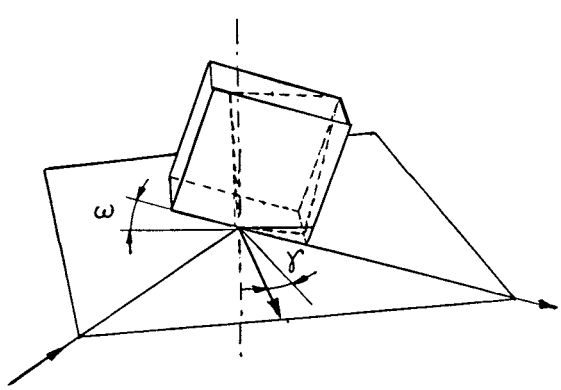


Fig.2 Quartz plate rotation

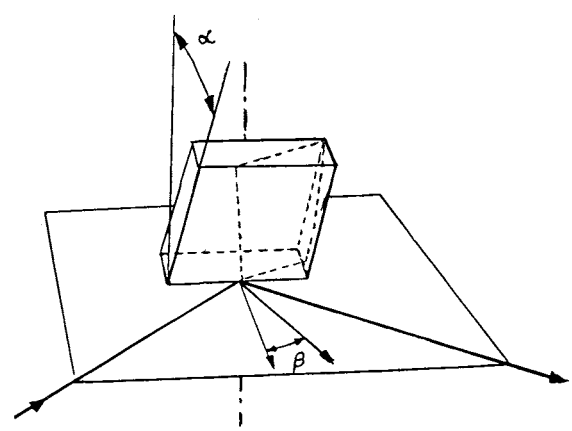


Fig.3 Quartz plate tilting

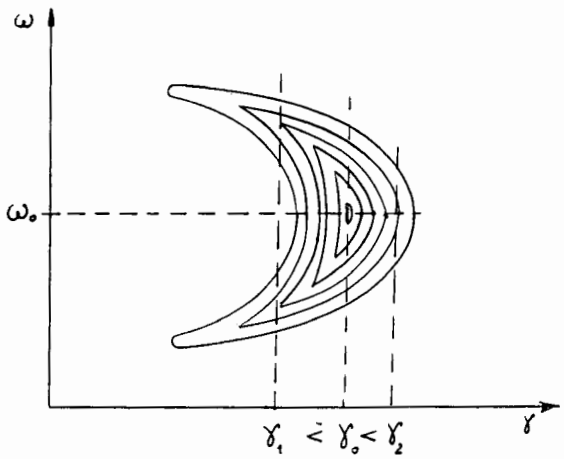


Fig.4 Diffraction intensity dependence of ω and γ angles

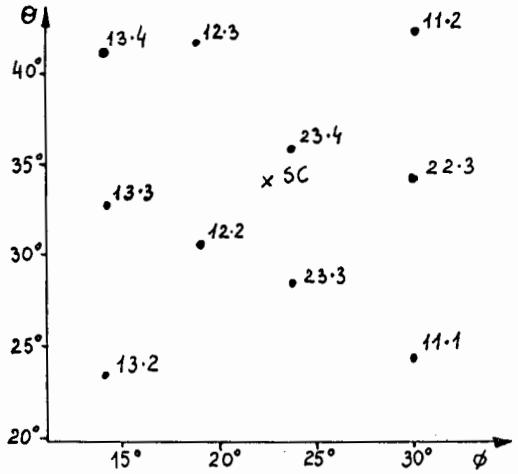


Fig.6 Atomic planes near the SC cut

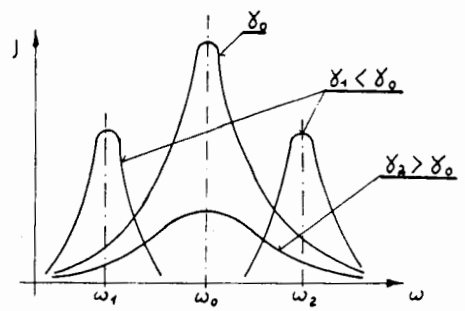


Fig.5 Diffraction intensity dependence of ω angle

A FREQUENCY GENERATING SOURCE : FREQUENCY SYNTHESIZERS

N.YAHYABEY* and S.HASSANI+
+ Laboratoire de physique nucleaire, C.D.T.N
* Laborotoire de fusion, C.D.T.A
2 Bd Frantz Fanon, Algiers, Algeria.

Abstract.
In the domain of harmonic approach we consider a novel mode of fractional frequency harmonic generation obtained by the modulation of a square wave reference signal. The produced spectrum exhibits a particular structure. The longer the fractional part of the frequency of the modulating wave is, the more figures the value of the frequency of the harmonic has. A comparison is done in terms of design purposes of component frequency generator blocks of frequency synthesizers.
Key words: Frequency synthesis, Fourier analysis, Noise

I. Introduction.

When the simultaneous generation of small number of frequencies is required, it is customary to use the brute-force or the harmonic approach |2|. These approaches use as building blocks a reference frequency source with a set of multipliers, dividers and mixers. The basic problems are then the spurious outputs generated in multiplication, division and mixing with an associated degradation in terms of phase noise performances. In this context, we propose the theory of a novel mode of fractional frequency harmonic generation.

II. Theory.

Let us define the half-cycle width modulation of a square waveform signal $\alpha(t,\infty)$ as follows,

$$\alpha(t,\infty)=\sum_{n=0}^{\infty}\left[\theta(t-T_n)\theta(T_{n+1/2}-t)-\theta(t-T_{n+1/2})\theta(T_{n+1}-t)\right] \qquad (1)$$

$T_q = qT - A\sin(\omega.q)$, q is an integer or half-integer

A , T are constants

$\theta(t)$ is the Heaviside function. qT are nominal times of frequency switch. A, defines the depth of the modulation and must be chosen $\leq T/4$ in order to avoid overlapping of adjacent T_q. We note that if the $\omega/2\pi = 1/p$ where p is an integer, the signal $\alpha(t,\infty)$ is no longer modulated. The angular frequency of the modulating wave is then of the form,

$$\omega = \frac{q}{l}2\pi \ ; \ l\neq 1 \quad \text{and } (l,q) \text{ are mutual prime numbers.} \qquad (2)$$

The Fourier coefficients calculated in |2| are given by,

$$C(b)=\frac{1}{i\pi b}\sum_{n=-\infty}^{n=\infty} J_n\left(\frac{2\pi A}{T}b\right).\psi(b+2n\varepsilon).(1-Exp\,i\pi(b+2n\varepsilon)) \qquad (3)$$

$$\psi(x)=\begin{cases}1 & , x \text{ is an integer}\\ 0 & , \text{otherwise}\end{cases} \quad \text{and} \quad C(0)=0 \ , \ \varepsilon = q/l$$

$J_n(x)$ are Bessel functions of integral order n and argument x. The frequencies b, satisfy the following fundamental relation,

$$b = \frac{\omega}{2\pi}p + p' \quad , \qquad p, p' \text{ integers} \qquad (4)$$

As depicted in figure 1 , the longer the fractional part of $\omega/2\pi = q/l$ is, the more figures the value of b has. This shows that a large range of harmonics with their desired frequency figures are made available by a adequate choice of the angular frequency ω.

III-Phase noise performances.

The square waveform signal $\alpha(t,\infty)$ may be obtained from a sinusoidal frequency source by means of a limiter and an amplifier. If we assume phase noise in the source signal, the period T_n of $\alpha(t,\infty)$ will be modified as follows,

$$T_n = nT - A\sin(\omega n) + \tau_n \quad .\tau_n \in (-\Delta T/2, \Delta T/2) \qquad (5)$$

τ_n measure fluctuations uniformly distributed in the interval $(-\Delta T/2, \Delta T/2)$ around nominal times of frequency switch where the additional perturbations introduced by the amplifier are included.

III.1-Synthesizer component frequency generator block.

According to the definition given in |1|, the coefficients Q_m and Q_d for a frequency multiplier and a frequency divider are as follows,

$$Q_m(n) = \Omega(n\delta T, b) \qquad Q_d(m) = \Omega(\upsilon.\delta T, b) \qquad (6)$$

$\Omega(\Delta X, b)$ is the Fourier transform of the distribution function of the perturbations affecting the periodicity of the output signal (limited and amplified), ΔX is the extent of the perturbations. b is the frequency order of the harmonic. n is the muliplication factor, υ is a factor dependent on the noise figure of the divider circuits. δT is the extent of the phase noise perturbations measured at the zero-crossings of the reference source signal applied to the inputs of the multiplier and the divider.

The synthesizer depicted in figure 2 generates each output frequency as a programmed, rational multiple of a reference frequency . To each input correspond the factor $Q_m(n)$. We have the Set of 4 frequencies with 4 figures each. $R=\{44.2368, 42.5984, 40.960, 39.3216\}$, with an input frequency f_o=8.

III.2 Frequency generating source

The sinusoidal modulation signal with its appropriate chosen frequency figures may be derived from the reference frequency source by means of a divider. The modulation process according to the definition given by relation (1) may be obtained by various means. The coefficient Q_s calculated in |2| at the modulator output is given for imtegral values of b,

$$Q_s = \Omega(\Delta T, b)\times\frac{J_o(x.\Omega(\delta\phi))}{J_o(x)} \qquad (7)$$

ΔT is the total extent of perturbations at the zero-crossings of the the signal $\alpha(t,\infty)$ around its nominal times of frequency switch $\omega_c = 2\pi/T$ produced by the processing chain: Limiter, amplifier and the modulator. $\delta\phi$ is the perturbation extent produced by the divider and the modulator and is dependent on their respective noise figures. Comparatively to the component frequency generator block, we obtain the following set close to R, $S=\{44.2426, 42.5978, 40.9094, 39.3116\}$ when we choose ε =.4221, with an input source of much lower frequency f_i=1 (=f_o/8)

The proposed configuration based on the above developed principle combines a small number of electronic elements and does not include <u>any frequency multiplier</u>. Its coefficient Q_s is <u>far higher than</u> $Q_m(n)$ produced at each output of the discussed synthesizer component frequency generator block. Comparatively to the <u>PLL</u> approach, it consists of no <u>voltage-controlled oscillator</u> nor a phase detector.

IV-Conclusion.

The theory of a novel mode of fractional frequency harmonic generation is presented. Results show that the produced spectrum possesses a particular structure. . We show that this property can be carried out and can open the way to a new generation of frequency synthesizers. It provides significant advantages in terms of phase noise performances of component frequency generator blocks of frequency synthesizers.

REFERENCES:

1- S.HASSANI, N.YAHYABEY "A frequency generating source"
To be submitted to IEEE Trans. Instrum. Meas.
2- Vadim MANASSEWITCH " Frequency synthesizers theory and design"
John Wiley & sons Inc, 1976.

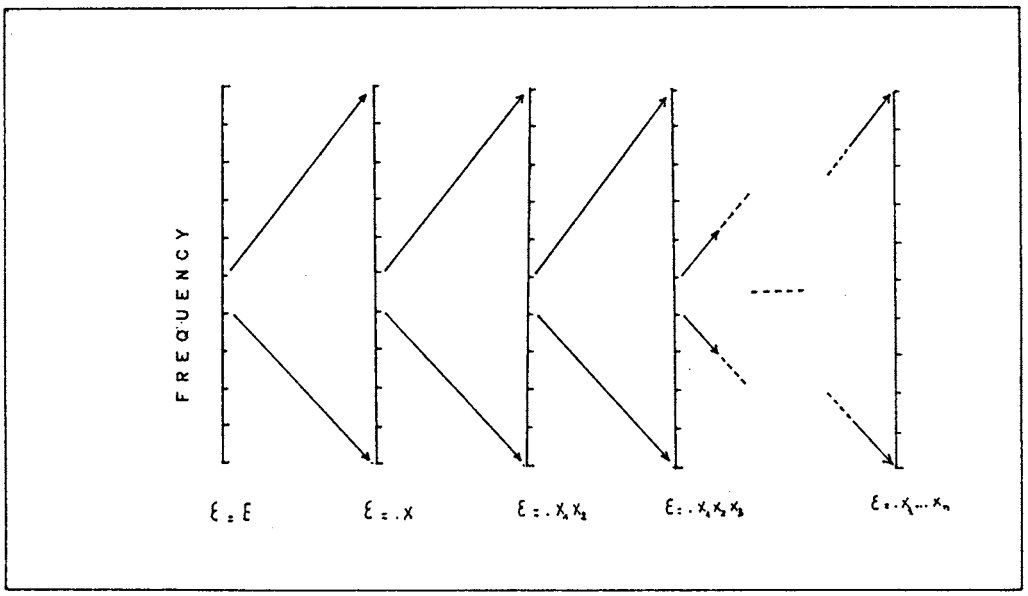
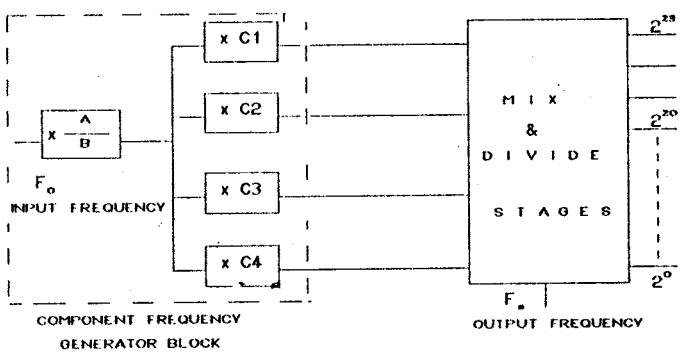


Figure 1. The structure of the produced spectrum for different figures of the modulation frequency ε, E is an integral value.



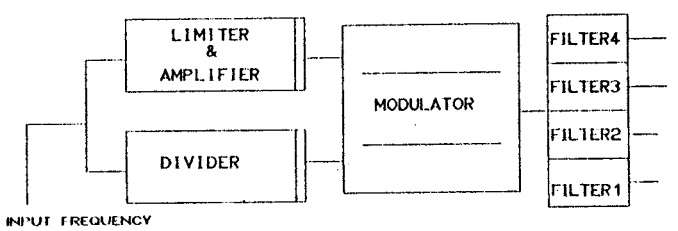


FIGURE 2.Component frequency generator block of a "cohrent" synthesizer and the proposed design approach.

PHASE DIFFERENCES OF THE FIELDS IN A RAMSEY RESONATOR FOR SUBMILLIMETRIC FREQUENCY STANDARDS*.

E. Bava[1], P. Tavella[2]

[1] Dip. Elettronica, Politecnico di Milano, Milano, Italy
[2] Istituto Elettrotecnico Nazionale G.Ferraris, Torino, Italy

* Work partially supported by the CNR (National Council of Research) of Italy.

ABSTRACT

In atomic beam frequency standards the Ramsey technique, used for a high definition of the atomic resonance, can be impaired by the resonator losses, which yield phase-shifts between the two separated fields or residual first-order Doppler effect. These errors are increased if asymmetries are present in the electromagnetic structure. Therefore low losses and suitable resonator geometries may combine to reduce these uncertainty sources. This work is a first attempt to study an open resonator suitable for standards operating in the submillimeter region, such as Mg and Ca beams, and hopefully in the infrared region. The structure is obtained by weakly coupling two resonators by means of a very thin splitter. One of the resonators is doubly folded, to allow the Ramsey technique, and it is fed in a nearly symmetric way by the other one. The resonances of the overall structure are studied, as they appear at the input, to distinguish the high excitation from the low excitation ones in the folded resonator.
This analysis is firstly performed supposing a lossless structure, then the phase relationships between the two separated fields are evaluated in a few special cases for a real structure.

Keywords: optical resonator, frequency standard.

1 INTRODUCTION

A high resolution of the resonance frequency in atomic beam devices is usually obtained by the Ramsey interrogation method, however to be fully exploited this technique requires that many unwanted effects are kept to a very low level. Among these uncertainty sources, phase-shift error and residual 1st-order Doppler shift are usually ascribed to losses and asymmetries with respect to the atomic beam in the radiation guiding structures.

To reduce these effects in primary Cs standards, experimental techniques have been devised and detailed analyses have been performed as far as the U-type waveguide cavity is concerned, as the large number of relevant references reported in [1] shows. Moreover, a different cavity design with ring cavity end parts has been studied and experimented [2,3]. Among the other frequency standards operating at very different wavelengths and requiring different cavities, a detailed analysis of the phase-shift error has been performed on the open resonator used in a Mg beam [4], a structure with very low losses (Q factor higher than 10^5), but rather asymmetric.

The object of this work is to undertake an examination of the properties of an open resonator for submillimetric standards with a high level of symmetry. Although simple in principle, it is composed by five reflectors, a semi-transparent mirror and a thin beam splitter (Fig. 1).
The resonance conditions of the real structure are difficult to be found in an explicit form due to the large number of parameters involved. To reduce complexity the resonator is firstly studied as a structure without losses and some preliminary special solutions are identified for such a case. Afterwards the behavior of the resonator is evaluated, for the same special solutions, taking losses into account.

2 DESCRIPTION OF THE RESONATOR

The resonator considered is schematically shown in Fig. 1. Radiation is supplied, through the semitransparent mirror M_1, into the section delimited by M_1 and the reflector M_2.

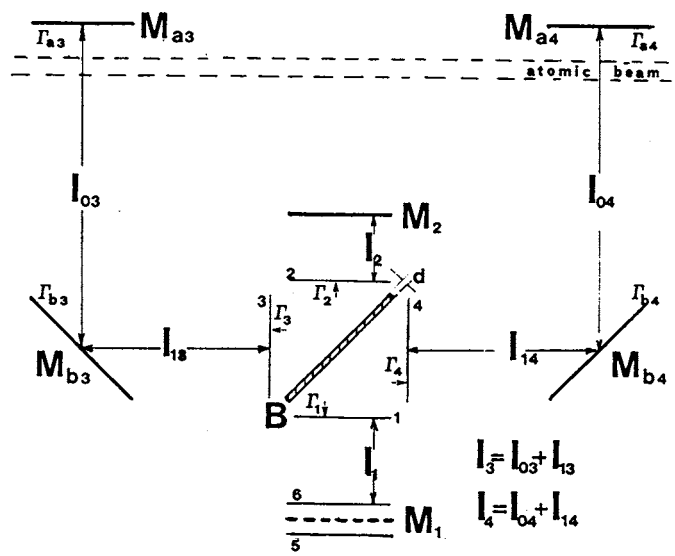


Figure 1: the resonator in a simplified drawing with flat mirrors.
M = mirrors,
B = beam splitter (thickness d),
l = distance,
Γ_i = reflection coefficient,
1, 2, ..., 6 = reference planes.

A very thin dielectric sheet (or a metal mesh) with low losses (B) acts as a beam splitter exciting the main section, defined by mirrors M_{a4}, M_{b4}, M_{b3}, and M_{a3}. This main section is twice folded in order to allow the crossing of the atomic beam in two separated regions. In the drawing the mirrors are flat, but actually they should be properly curved, when necessary, to obtain a confined electromagnetic field (e.m.f.) with plane wavefronts in the interaction regions and in the centre of the beam splitter.
The optical lengths βl_k of the four arms are determined by the reference planes (Fig. 1) and by the mirrors M_{a4}, M_{b4}, M_{a3}, M_{b3}, M_2, and M_1. Mirror and dielectric losses, distances between mirrors and

splitter and splitter thickness introduce in this problem a rather large number of parameters. Some of these are unavoidable (losses), the others (distances) can be suitably chosen to obtain a resonant structure operating as close as possible to an ideal one.

The mirror electric field reflection coefficients Γ_j, transferred to the reference planes near the beam splitter (Fig. 1), can be written as:

(1) $$\Gamma_j = \gamma_j e^{-i\psi_j} = \Gamma_{aj} e^{-i2\beta l_j} = \gamma_j e^{-i(\psi_j+\pi)} e^{-i2\beta l_j} \qquad \text{for } j = 1,2$$

where γ_j are moduli $\neq 1$ due to ohmic losses (and coupling as well, for $j = 1$), l_j are geometrical distances, ψ_j are additional phases due to losses or coupling and $\beta = 2\pi/\lambda$;

(2) $$\Gamma_j = \gamma_j e^{-i\psi_j} = \Gamma_{aj}\Gamma_{bj}^2 e^{-i2\beta l_j} = \gamma_{aj}\gamma_{bj}^2 e^{-i(\psi_{aj}+2\psi_{bj}+\pi)} e^{-i2\beta l_j} \qquad \text{for } j = 3,4$$

where Γ_a and Γ_b are reflection coefficient at normal and 45° incidence, respectively. The other symbols have the same meaning as in (1).

The reflection coefficients can be expressed by means of the normalized surface impedance $\zeta(\omega) = Z(\omega)/Z_0$ as:

(3) $$\Gamma_a = -\frac{1-\zeta(\omega)}{1+\zeta(\omega)}$$

(4) $$\Gamma_b = -\frac{1-\zeta(\omega)/\sqrt{2}}{1+\zeta(\omega)/\sqrt{2}}$$

where $\zeta(\omega)$ is a function of the plasma frequency ω_p and the damping frequency ω_τ through the relationship [5]:

(5) $$\zeta(\omega) \approx \frac{\omega_\tau}{\omega_p}\left(\frac{\omega}{2\omega_\tau}\right)^{1/2}\left[1-\frac{\omega}{2\omega_\tau}+i\left(1+\frac{\omega}{2\omega_\tau}\right)\right]$$

From [5] a conservative estimate at 600 GHz for Al is $Re[\zeta(\omega)] \approx Im[\zeta(\omega)] \leq 10^{-3}$.

The mirror transverse dimensions are large enough to yield diffraction losses much lower than ohmic and coupling losses. The mirror geometry is supposed to sustain only one paraxial propagation mode at resonance, the dielectric sheet to be thin enough to render negligible any mode conversion and, moreover, the e.m.f. exciting the resonator is shaped to match the Gaussian mode of the open cavity. Therefore the e.m.f. amplitudes in the four arms and their dependences on the excitation field can be studied by means of the scattering parameters.

Under these hypotheses, the magnetic fields in the two interaction regions are related to the exciting fields b_3 and b_4, at the planes 3 and 4 (Fig. 1 and Fig. 2), according to the relationship:

(6) $$\frac{H(z_4)}{H(z_3)} = \frac{b_4}{b_3}\frac{\Gamma_{b4}}{\Gamma_{b3}}\frac{e^{-i\beta l_4}}{e^{-i\beta l_3}}\frac{e^{-i\beta(z_4-l_4)} - \Gamma_{a4}e^{i\beta(z_4-l_4)}}{e^{-i\beta(z_3-l_3)} - \Gamma_{a3}e^{i\beta(z_3-l_3)}}$$

where $l_{1j} < z_j < l_j$. To obtain $H(z_4)$ and $H(z_3)$ in phase, the conditions

(7a) $$\beta(z_4-l_4) = \beta(z_3-l_3) + k\pi, \qquad \Gamma_{a4} = \Gamma_{a3}$$

have to be fulfilled and it must be imposed that:

(7b) $$\frac{b_4}{b_3}\frac{\Gamma_{b4}}{\Gamma_{b3}}\frac{e^{-i\beta l_4}}{e^{-i\beta l_3}} \approx \pm 1 \qquad \text{(real number)}$$

where the sign + or - is to be chosen according to k in (7a) is even or odd, respectively.

The choice of equal length for the two arms concerned with the Ramsey interrogation ($l_3 = l_4$) appears useful, for sake of symmetry, but it raises problems, from the experimental point of view, because it is quite difficult to guarantee this condition at a high accuracy level, being the operation wavelength in the submillimetric region (in the Mg beam, for example, the wavelength is about 0.5 mm).

Moreover, in order to adjust the mirror positions, only two observations are available in general without changing the operation conditions: the shape of the Ramsey pattern at the output of the device and the reflection coefficient at the resonator input.

In both cases information can be gained either by varying ω in a fixed geometry or by changing the lengths l_j, one at a time, at a fixed exciting frequency.

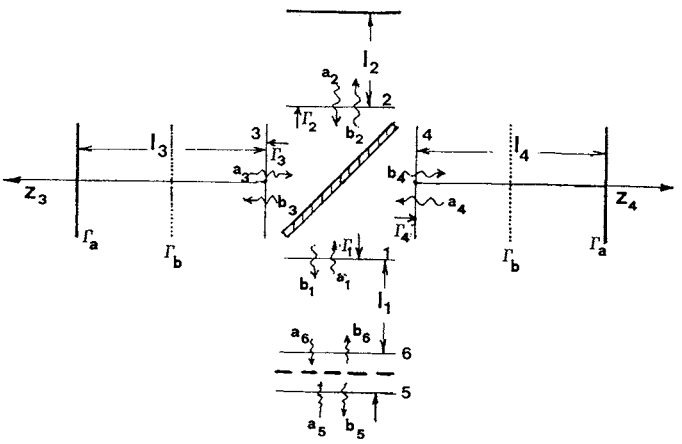


Figure 2: sketch of the resonator with reference planes (1, 2, ..., 6), backward and forward travelling waves (b_j, a_j), and reflection coefficients $\Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4$.

Γ_a and Γ_b are reflectors of the main section, with normal and 45° incidence, respectively.

One more consideration to be taken into account in the analysis of the resonator is that a good efficiency in the excitation of the main section of the resonator, that is the arms involved in the interaction with the beam, has to be reached.

3 STATEMENT OF THE PROBLEM

With reference to the sketch of Fig. 2, the resonator operation is studied by means of the scattering matrix and, with respect to the four reference planes, including the splitter, the following general relationships hold:

(8) $$b_j = \sum_{k=1}^{4} s_{jk} a_k \qquad (j = 1,\ldots 4)$$

where b_j and a_k are amplitudes of electric field waves, backward and forward respectively.

From symmetry properties, for reference planes at equal distances from the centre of the beam splitter, only two parameters are required to characterize the matrix S in (8), the transmission (t) and the reflection (r) coefficient of the beam splitter respectively.

Accordingly, the s_{ik} elements can be written in the following way:

(9) $$|s_{jk}| = \begin{vmatrix} 0 & t & 0 & r \\ t & 0 & r & 0 \\ 0 & r & 0 & t \\ r & 0 & t & 0 \end{vmatrix}$$

and, without loss of generality, the reference plane positions can be chosen in such a way that t is real and positive.

At each reference plane, a_k and b_k are linked by the relations:

(10) $$a_k = \Gamma_k b_k$$

where, according to (1) and (2), the reflection coefficients Γ_k are determined by mirror properties and positions.

Equation (8), with (9) and (10), refers to a resonator without any external source of radiation. When a source feeds the resonator through M_1, the parameters b_1 and a_1 are eliminated in (8) by means of:

(11) $$b_5 = Ra_5 + Tb_1 e^{-i\beta l_1}$$

$$a_1 = Te^{-i\beta l_1} a_5 + Re^{-i2\beta l_1} b_1$$

where R (reflection) and T (transmission) are scattering parameters of M_1 at planes 5 and 6, chosen in such a way to get R real and negative, l_1 is the distance between the reference planes 1 and 6. Therefore, from (10) it follows:

(12) $$\Gamma_1 = Re^{-i2\beta l_1}$$

The determinant Δ_H of the homogeneous system obtained from (8), (9), and (10) yields the resonator proper angular frequencies ω_H (or $\beta_H = \omega_H/c$) through the condition:

(13) $\Delta_H = 1 - t^2(\Gamma_1\Gamma_2 + \Gamma_3\Gamma_4) - r^2(\Gamma_1\Gamma_4 + \Gamma_2\Gamma_3) + (t^2 - r^2)^2\,\Gamma_1\Gamma_2\Gamma_3\Gamma_4 = 0$

When ohmic or coupling losses are taken into account, $|\Gamma_k| < 1$, $|t^2 - r^2| < 1$, $R^2 + |T|^2 < 1$, and ω_H is complex.

For a lossless resonator,

$|\Gamma_k| = 1$, $R = -1$, $t^2 - r^2 = 1$, $R^2 + |T|^2 = 1$, T and r are imaginary and ω'_H is real.

However, the interest in this case is not to find ω_H or ω'_H, but to define the phases of reflection coefficients Γ_j in order to have an open structure resonant at a frequency ω_0 determined by an atomic transition.

When, in the lossless case, a set of relations among reflection coefficients Γ_i satisfying (13) is identified, the quantities $|b_3/b_4|$ and $|b_3/b_1|$ are evaluated.

The first ratio $|b_3/b_4|$ is related to the Ramsey pattern and should be almost 1; the second one, $|b_3/b_1|$, should be larger than 1 for an efficient excitation of the interaction regions. If these conditions are satisfied, the input reflection coefficient b_5/a_5 is evaluated, with these same phase relations, for the real structure fed by a source with angular frequency ω_s and amplitude a_5. The steady state operation is given in this case by the system:

(14)

$$\begin{array}{cccccc} tT\Gamma_2 b_2 & & +rT\Gamma_4 b_4 & -e^{+i\beta_s l_1} b_5 & = & -Re^{i\beta_s l_1} a_5 \\ Tb_2 & -rT\Gamma_3 b_3 & & -tRe^{-i\beta_s l_1} b_5 & = & t(T^2 - R^2)e^{-i\beta_s l_1} a_5 \\ r\Gamma_2 b_2 & -b_3 & +t\Gamma_4 b_4 & & = & 0 \\ & tT\Gamma_3 b_3 & -Tb_4 & +rRe^{-i\beta_s l_1} b_5 & = & -r(T^2 - R^2)e^{-i\beta_s l_1} a_5 \end{array}$$

System (14) has the determinant

(15) $$\Delta = T^2 e^{i\beta_s l_1} \Delta_H(\omega_s) \neq 0$$

which is different from zero because of losses in $|\Gamma_k|$, t, and r, even in the case $\omega_s = \omega'_H = \omega_0$.

The interesting expressions to be evaluated are:

(16) $$\frac{b_5}{a_5} = R + T^2 e^{-i2\beta_s l_1} \frac{r^2\Gamma_4 + t^2\Gamma_2 - \Gamma_2\Gamma_3\Gamma_4(t^2 - r^2)^2}{\Delta_H(\omega_s)}$$

$$= R + T^2 e^{-i2\beta_s l_1} \frac{b_1}{a_1 - Re^{-i2\beta_s l_1} b_1}$$

(17) $$\frac{b_3}{b_4} = \frac{t(\Gamma_2 + \Gamma_4)}{1 + \Gamma_2\Gamma_3(t^2 - r^2)}$$

(18) $$\frac{b_1}{b_3} = \frac{r^2\Gamma_4 + t^2\Gamma_2 - \Gamma_2\Gamma_3\Gamma_4(t^2 - r^2)^2}{rt(\Gamma_2 + \Gamma_4)}$$

(19) $$\frac{b_2}{b_4} = \frac{t}{r}\frac{1 - (t^2 - r^2)\Gamma_3\Gamma_4}{1 + (t^2 - r^2)\Gamma_2\Gamma_3}$$

It is important to stress that the expression (17), (18), and (19) are formally the same as the corresponding ones obtained from the homogeneous system, which however share the condition $\Delta_H = 0$ at ω'_H (lossless case) or at ω_H (real case).

For each set of ϕ_k values defined in (1) and (2) and satisfying (13), in the lossless case, the evaluations of (17), (18), and (19) allow a first choice of solutions satisfying the desired conditions on $|b_3/b_4|$ and $|b_3/b_1|$.

Then keeping these same ϕ_k values, but introducing losses in $|\Gamma_k|$, t, and r, equation (16) is used to estimate the squared modulus $|b_5/a_5|^2_{\omega_s = \omega_0}$, whose value is made small by fixing other phase relationships and the reflection parameter $|R|$ of M_1.

The knowledge of the ratios (16), (17), (18), and (19) is important for recognizing the desired resonances obtainable by experimental mirror adjustments. With suitable losses introduced in (17), an estimate of b_3/b_4 is available and permits to evaluate the phase-shift between $H(z_3)$ and $H(z_4)$ by means of (6).

4 STUDY OF SOME SPECIFIC SOLUTIONS

In general only numerical solutions of (13) can provide the relationships among the reflection coefficients Γ_k, at the proper frequencies of the resonator ω'_H. In this first attempt to obtain some specific set of ϕ_k values giving the proper frequencies of the resonator, in the lossless case, the resonance conditions independent from the values of t (and r) are considered.

With the following substitutions:

(20) $$\Gamma_1\Gamma_2 = x = e^{-i\phi_x}$$

(21) $$\Gamma_3\Gamma_4 = y = e^{-i\phi_y}$$

(22) $$\Gamma_1\Gamma_4 = z = e^{-i\phi_z}$$

(23) $$\Gamma_2\Gamma_3 = \frac{xy}{z} = e^{-i(\phi_x + \phi_y - \phi_z)}$$

into (13), the equation $\Delta_H(\omega'_H) = 0$ becomes:

(24) $$(1 - x)(1 - y) - r^2(x + y + z + xy/z) = 0$$

symmetric in x and y and with $(-r^2 > 0)$.

These specific solutions are immediately given by:

(25) $x = 1 \quad (1 + z)(y + z) = 0$

(26) $y = 1 \quad (1 + z)(x + z) = 0$

By omitting the additive $2m\pi$ in the expressions of the phases, these solutions are:

a) $z = -1$, $y = 1 \Rightarrow \phi_z = \pi$, $\phi_y = 0$

$\Gamma_1 = -\Gamma_3$, $\Gamma_4 = \Gamma_3^{-1} \Rightarrow \phi_1 = \phi_3 + \pi$, $\phi_4 = -\phi_3$

a') $z = -1$, $x = 1 \Rightarrow \phi_z = \pi$, $\phi_x = 0$

$\Gamma_2 = -\Gamma_4$, $\Gamma_2 = \Gamma_1^{-1} \Rightarrow \phi_2 = \phi_4 + \pi$, $\phi_1 = -\phi_2$

b) $y = 1$, $x = -z \Rightarrow \phi_y = 0$, $\phi_x = \phi_z + \pi$

$\Gamma_2 = -\Gamma_4$, $\Gamma_3 = \Gamma_4^{-1} \Rightarrow \phi_2 = \phi_4 + \pi$, $\phi_3 = -\phi_4$

b') $x = 1$, $y = -z \Rightarrow \phi_x = 0$, $\phi_y = -\phi_z$

$\Gamma_3 = -\Gamma_1$, $\Gamma_1 = \Gamma_2^{-1} \Rightarrow \phi_3 = \phi_1 + \pi$, $\phi_1 = -\phi_2$

The evaluations of expressions (17), (18), and (19) for these cases are reported in the table below. The solutions a) and a') hold for $\Gamma_2\Gamma_3 \neq -1$, conversely b) and b') hold for $\Gamma_1\Gamma_4 \neq -1$, where these limitations are to be considered only for lossless problems.

	a)	a')	b)	b')
b_3/b_4	$t\Gamma_4$	0	Γ_4/t	∞
b_1/b_3	r/t	∞	0	t/r
b_2/b_4	0	t/r	r/t	∞

All the solutions found have one arm with amplitude b_j=0 (negative interference). By inspection of the table above, a') and b') must be avoided, whereas a) and b) can be accepted.

Going on in this analysis, neglecting dielectric losses, further conditions can be found forcing b_5/a_5 to be a small value, namely:

a_1) $\phi_2 + \phi_3 = \pi$

b) $\phi_1 + \phi_2 = 0$

which collapse in the same solution yielding:

(27) $$\frac{b_5}{a_5} = \frac{\gamma_2(1 - A) - |R|}{1 - |R|\gamma_2}$$

for $|r^2| \ll 1$ and with $|T^2| = 1 - R^2 - A$, where A is related to the relative ohmic losses of M_1. In this case the resonator is well matched with the choice $|R| \approx \gamma_2(1 - A)$.

Another solution is also found from the case a), adding the condition:

$a_2)\ \phi_2+\phi_3=0 \quad (\text{or } \phi_1+\phi_2=\pi)$

for $|r^2|<<1$. This solution could be sustained even in the absence of the mirror M_1. In the case a_2), b_3/a_5 is small but strongly dependent on mirror losses and on the value of r^2. An estimate for a good matching is $|R| \approx \gamma_2(1-A)(|r^2|-5\zeta)/(|r^2|+5\zeta)$, where ζ is defined in (5).

For these specific solutions, the amplitude ratios (17) and the phase shifts of $H(z_4)$ and $H(z_3)$ (6) can be easily evaluated, introducing dielectric losses as:

$$t^2-r^2=qe^{-i\phi_d}$$

where t is real by definition and ϕ_d is a very small quantity of the order of 10^{-5} (see Table 1).

For the two equivalent cases a_1) and b), (17) gives:

(28) $$\frac{b_3}{b_4}\approx\frac{t(\gamma_2-\gamma_4)e^{-i(\phi_4+\pi)}}{1-q\gamma_2\gamma_3e^{-i\phi_d}}\approx\frac{t(\gamma_2-\gamma_4)e^{-i\chi}}{\sqrt{2}(1-q\gamma_2\gamma_3\cos\phi_d)}e^{-i(\phi_4+\pi)}$$

with

$$\chi=artg\frac{q\gamma_2\gamma_3\sin\phi_d}{1-q\gamma_2\gamma_3\cos\phi_d}$$

The ratio $|b_3/a_4|$ can be rather different from 1 because of the differences between nearly equal quantities in the numerator and the denominator of (28). Moreover the contribution to the phase shift from dielectric losses is produced by $|\chi|$, which is much larger than $|\phi_d|$.

On the contrary, for the case a_2),

(29) $$\frac{b_3}{b_4}\approx\frac{t(\gamma_2+\gamma_4)e^{-i\phi_4}}{1+(t^2-r^2)\gamma_2\gamma_3}\approx te^{-i\phi_4}e^{i\phi_d/2}$$

and $|b_3/a_4|\approx 1$. The phases in the two interaction arms are related by:

(30) $$\frac{H(z_4)}{H(z_3)}=\frac{\gamma_{b4}}{\gamma_{b3}t}e^{\frac{-i\phi_d}{2}}e^{im\pi}\frac{e^{-i[\beta(z_4-l_4)-\psi_{a4}/2]}-\gamma_{a4}e^{i[\beta(z_4-l_4)-\psi_{a4}/2]}}{e^{-i[\beta(z_3-l_3)-\psi_{a3}/2]}-\gamma_{a3}e^{i[\beta(z_3-l_3)-\psi_{a3}/2]}}$$

where m is even (odd) if k in (7a) is even (odd).

For an ideal tuning the main contributions to the phase shift between $H(z_4)$ and $H(z_3)$ are given by dielectric losses (ϕ_d) and loss asymmetries in M_{a4} and M_{a3} ($\gamma_{a3}\neq\gamma_{a4}$, $\psi_{a3}\neq\psi_{a4}$). Atomic beam divergence and misalignments between standing wave antinodes should produce the same effects as discussed in [4]. However, the relationships between the phases ϕ_j cannot be obtained exactly in an experimental device and an evaluation of these residual uncertainties should be performed to determine the overall limits of this type of resonator.

5 CONCLUSION

An analysis has been performed for some special solutions of a open resonator which could be used in a submillimetric frequency standard (e. g. the Mg beam).
The phase shift of this structure is estimated when the resonator is perfectly tuned and its amount turns out to be very small if a thin dielectric sheet, made of teflon, TPX, or polyethylene is used.

REFERENCES

1 J. Vanier, C. Audoin, *The quantum physics of atomic frequency standards*, Adam Hilger, Bristol, 1989, sec. 5, pp. 699-706, 811-814.

2 A. De Marchi," A novel cavity design for minimization of distributed phase shift in atomic beam frequency standards," in Proc. 40th Annual Frequency Control Symp, Piscataway, NY, 1986, pp.441-446.

3 H. De Boer, et al.," Experimental data on a new Ramsey cavity for cesium atomic clocks," in Proc. 4th European Frequency and Time Forum, Neuchatel, CH, 1990, pp. 523-526.

4 E. Bava, A. Godone. C. Novero," Cavity phase-shift error evaluation in a magnesium atomic beam frequency standard," Appl. Phys. B 48, pp. 495-505, 1989.

5 M. A. Ordal, et al.," Optical properties of the metals Al, Co, Cu, Au, Fe, Pb, Ni, Pd, Pt, Ag, Ti, and W in the infrared and far infrared," Applied Optics, vol. 22, n. 7, pp. 1099-1119, 1983.

6 J. R. Birch, J. D. Dromey, J. Lesurf," The optical constants of some common low-loss polymer between 4 and 40 cm^{-1}," Infrared Physics, 21, pp. 225-228, 1981.

7 R. E. Collin, *Field theory of guided waves*, Mc Graw Hill, New York, 1960, pp. 87-96.

	teflon	l. d. polyethylene	TPX	teflon	l. d. polyethylene	TPX
		$\left(\frac{d}{\lambda}=\frac{1}{50}\right)$	$\lambda=0.5mm$		$\left(\frac{d}{\lambda}=\frac{2}{50}\right)$	$\lambda=0.5mm$
$\sqrt{\varepsilon}$	1.43	1.52	1.46			
ε'	2.0449	2.3104	n2.1316			
$\alpha(Np/cm)$	0.25	0.12	0.25			
ε''	4.07E-03	2.21E-03	4.24E-03			
$tg\delta=\varepsilon''/\varepsilon'$	0.0020	0.0010	0.0020			
θ_0	0.156	0.169	0.161	0.312	0.338	0.321
without losses						
t_0	0.9957	0.9934	0.9950	0.9837	0.9749	0.9810
r_0	-i 0.0921	-i 0.1151	-i 0.0996	-i 0.1797	-i 0.2227	-i 0.1939
with losses						
Re r	-0.00011	-0.00005	-0.00011	-0.00021	-0.00009	-0.00020
Im r_2	-0.0920	-0.1151	-0.0995	-0.1795	-0.2226	-0.1936
Re r_2^2	-0.0085	-0.0132	-0.0099	-0.0322	-0.0495	-0.0375
Im r_2^2	0.00002	0.00001	0.00002	0.00007	0.00004	0.00007
$Re(t^2-r^2)$	0.9986	0.9992	0.9985	0.9973	0.9986	0.9972
$Im(t^2-r^2)=Im(-r^2)$						
$\lvert t^2-r^2\rvert$	0.9986	0.9992	0.9985	0.9973	0.9986	0.9972

Table 1: *relative dielectric constants $\varepsilon'-i\varepsilon''$ of teflon, polyethylene and TPX at 600 GHz from [6] and an evaluation of the transmission (t_0, t) and reflection (r_0, r) coefficients without and with losses respectively for two different cases of dielectric thickness d. The model for these evaluations has been taken from [7].*

Session 8A

Quartz: Materials

Chairman: H. Duchaussoy

DISSOLUTION CONTROLEE D'UN MATERIAU PIEZOELECTRIQUE, LA BERLINITE ($AlPO_4$)

O. Cambon[1], A. Goiffon[2], A. Ibanez[2], E. Philippot[2]

[1] CEPE, B.P.165 Argenteuil F-95100, France
[2] URA D0407 CNRS, UM II, case 003, F-34095 Montpellier cédex 5, France

RESUME:

Les résultats présentés concernent la dissolution contrôlée de la berlinite avec pour objectif la réalisation de résonateurs AT de haute fréquence (75Mz). La dissolution a lieu dans des solutions acides (H_3PO_4 ou H_2SO_4) de phosphate d'aluminium. La cinétique de dissolution et l'évolution de l'état de surface sont contrôlés au cours du processus d'amincissement chimique par variation de la fréquence de résonance et rugosimétrie. Après une étude systématique, le bain de dissolution contrôlée conduisant aux meilleurs résultats est utilisé pour préparer des résonateurs AT de berlinite. Ceux-ci sont caractérisés par leur performances piezoélectriques et comparés à des résonateurs obtenus par bombardement ionique ou à des résonateurs quartz de même orientation.

Keywords: Contolled dissolution, chemical etching, berlinite, filter.

1 - INTRODUCTION

Depuis quelques années, l'utilisation de résonateurs piézoélectriques s'est largement développée, en particulier dans le domaine des télécommunications. Après la mise au point de l'élaboration des matériaux massifs, la mise en forme des dispositifs eux mêmes est une étape importante qui conditionne les performances obtenues. Dans le cas des résonateurs, pour une même orientation cristalline, l'épaisseur de la lame vibrante et l'état de surface de celle-ci sont les facteurs essentiels qui vont définir respectivement la fréquence de résonance et la qualité du signal obtenu. L'usinage mécanique ne permettant pas d'obtenir des épaisseurs inférieures à 40 µm, pour atteindre des fréquences élevées, il est nécessaire de faire appel à des méthodes non conventionnelles telles que le bombardement ionique (IBE) ou l'amincissement chimique. La première méthode, bien que coûteuse, est utilisée industriellement pour traiter des matériaux à faible coût comme le quartz, par exemple. Par contre, l'usinage chimique est une méthode de choix dans le cas où l'étape de mise en forme des dispositifs doit être peu onéreuse tout en conduisant à un état de surface satisfaisant. Après avoir mis au point la cristallogénèse de la berlinite $AlPO_4$, isotype du quartz α, (Réf.1) qui présente des caractéristiques piézoélectriques intéressantes (Réf.2,3), il nous a paru important et utile d'étudier son usinage chimique afin de pouvoir réaliser des dispositifs de qualité satisfaisante à des coûts raisonnables.

Les premiers essais, réalisés dans le solvant qui avait donné les meilleurs résultats en cristallogénèse, c'est à dire l'acide sulfurique, ont montré que l'usinage chimique de la berlinite était possible (Réf.4). En effet, les cinétiques de dissolution dans ce milieu, sur la coupe AT, permettent d'utiliser des temps d'usinage tout à fait compatibles avec une utilisation industrielle, tout en conservant une précision d'épaisseur correcte. Cependant, l'état de surface obtenu avec cet acide n'était pas satisfaisant : rugosité élevée, présence de piqûres d'attaque (etch pits). Si l'on se reporte à la théorie de la croissance cristalline élaborée tout d'abord dans le cas des cristaux parfaits (Réf.5) puis pour les cristaux réels à partir de phase vapeur (Réf.6,7) ou en solution (Réf.8) et étendue enfin au cas du processus inverse de la dissolution (Réf.9,10), deux points essentiels peuvent être retenus :

- la réciprocité des phénomènes de croissance et dissolution cristalline.
- l'importance du facteur de sursaturation qui conditionne la qualité de ces 2 processus.

L'ensemble des résultats peut être schématisé par la figure 1 où ΔG^*_d représente la variation d'énergie libre pour la formation d'un nucleus de croissance (a > o) ou de dissolution (a < o) sur une dislocation. C mesure la concentration du matériau dissout dans la solution alors que Co est la solubilité de ce même matériau dans le solvant utilisé, à la température considérée. Pour avoir de bonnes conditions de croissance ou de dissolution, c'est à dire pour éviter la formation de piqûres d'attaque au niveau des dislocations, il faut que la barrière d'énergie soit élevée ce qui correspond en fait à se rapprocher des conditions de l'équilibre (partie hachurée de la fig. 1) en dissolvant du matériau à usiner dans la solution d'usinage chimique. Il est toutefois nécessaire de noter que lorsque l'on se rapproche trop de l'équilibre, les vitesses de croissance ou de dissolution sont très faibles, ce qui est défavorable pour une utilisation pratique. Si l'on veut réaliser la dissolution d'un matériau dans des conditions satisfaisantes, vitesse suffisante et bon état de surface final, il est nécessaire de déterminer la concentration en matériau à dissoudre conduisant aux conditions ci-dessus imposées. On parlera alors de dissolution contrôlée.

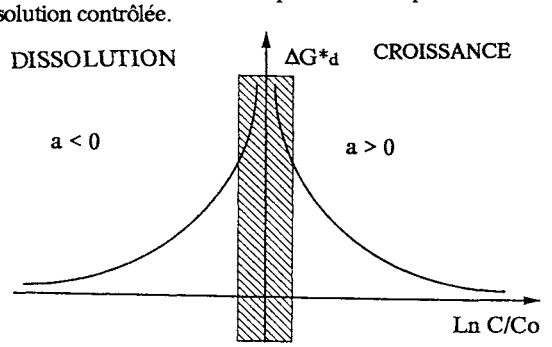


Figure 1 : Représentation de la variation de l'énergie d'activation ΔG^*_d pour la "nucléation" au niveau d'une dislocation. ($\Delta G^*_d < \Delta G^*_p$). $\Delta G^*_p = \Pi\gamma^2\Omega$ a/kTLnC/Co est la barrière d'énergie dans le cas d'un cristal parfait (Réf.9).

Au vu de ces considérations, nous avons donc poursuivi l'étude de la dissolution de la berlinite initiée en milieu acide seul par celle de sa dissolution contrôlée dans des solutions acides de phosphate d'aluminium. Nous présentons ici les résultats obtenus ainsi que leurs répercussions sur les performances des dispositifs piézoélectriques.

2 - DISSOLUTION CONTROLEE DE LA BERLINITE

Les résultats présentés ici concernent uniquement l'étude de la dissolution de lames AT de berlinite. Après un bref rappel des résultats obtenus antérieurement en milieu acide seul, c'est à dire pour C = O (Réf.4,11) nous exposerons ceux obtenus en utilisant des solutions acides, (H_3PO_4 ou H_2SO_4) de phosphate d'aluminium, soit pour C ≠ O.

2.1 - Partie expérimentale

2.1.1 - Dispositif expérimental et solutions. Celui-ci ayant été décrit par ailleurs (Réf.4) nous rappellerons simplement que la dissolution doit être menée à concentration et température du bain rigoureusement constantes ce qui nécessite un montage expérimental parfaitement isolé thermiquement et évitant tout départ de solvant.
Les solutions acides sont préparées par dilution des acides commerciaux et leurs concentrations sont contrôlées par dosage acido-basique classique et par mesure de densité.
Pour la préparation des solutions acides de phosphate d'aluminium, plusieurs voies ont été utilisées :

- en milieu phosphorique, H_3PO_4 : dissolution de berlinite, d'hydroxyde d'aluminium ou de phosphate d'aluminium hydraté dans de l'acide orthophosphorique, chaque quantité de composé étant connue.
- en milieu sulfurique, H_2SO_4 : dissolution de berlinite dans l'acide sulfurique ou dissolution de sulfate d'aluminium hydraté dans l'acide phosphorique.

2.1.2 - Echantillons et méthodes de caractérisations. Les lames AT étudiées sont des disques de 5 mm de diamètre rodées (abrasif de 3 µm) ou polies (abrasif de 0,3 µm). Elles sont issues de cristaux de berlinite élaborés selon les méthodes mises au point au laboratoire (Réf.1).
La vitesse de dissolution est mesurée à partir de la diminution de l'épaisseur d (µm). Celle-ci est déterminée par mesure de la fréquence de résonnance F (mesure air-gap) d'après la formule :

$$F\,(MHz) = K/d \qquad *$$

Dans le cas de lames AT de berlinite, la valeur calculée de K est de 1480 MHz.µm.
L'épaisseur initiale des lames est de 150 ou 70 µm ce qui correspond à des fréquences de 10 ou 21 MHz respectivement.
Parallèlement à la dimution d'épaisseur, l'état de surface est contrôlé par mesure du paramètre de rugosité Ra et observation par microscopie optique et électronique.

2.2 - Résultats et discussions

2.2.1 - Dissolution en solution acide (C = O) . Nous rappelons brièvement pour comparaison, les résultats obtenus en milieu acide chlorhydrique, phosphorique et sulfurique qui ont été détaillés par ailleurs (Réf.4,11).

* Au niveau cinétique, la vitesse de dissolution augmente avec la température et la concentration en protons selon une loi unique pour tous les acides :

$$V = A \exp(-Ea/RT)\,[H^+]^{1/2} \qquad [1]$$

* L'évolution de l'état de surface est similaire pour les 3 acides étudiés mais sa qualité finale atteinte, caractérisée par la rugosité Ra dépend de l'acide utilisé (Figure 2).
Après une augmentation régulière de la rugosité jusqu'à un maximum atteint pour la même dimution d'épaisseur quel que soit l'acide (environ 6 µm) on observe une décroissance plus ou moins rapide du paramètre mesuré.
La première partie de la courbe correspond à l'élimination de la couche superficielle perturbée par le sciage et le polissage. Après le maximum, on assiste au phénomène de "polissage chimique".

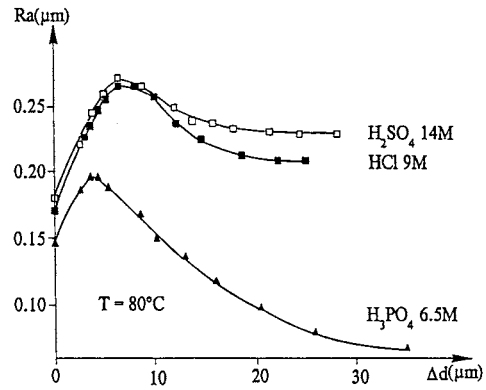


Figure 2: Variation de la rugosité Ra de lames AT de berlinite en fonction de la diminution totale d'épaisseur Δd, au cours de leur dissolution dans différents acides.

. Alors que les acides chlorhydrique et sulfurique ont un comportement voisin (valeur finale de Ra, 0,2, supérieure à la valeur initiale, 0,15), l'acide orthosphosphorique améliore l'état de surface puisque la rugosité atteinte est de l'ordre de 0,05µm. D'autre part, l'observation au microscope révèle une nette diminution des piqûres d'attaque dans ce même milieu sans toutefois les éliminer complètement. On peut expliquer ce phénomène par la présence en solution des ions phosphate PO^{3-}_4 constitutifs de la berlinite. En effet, bien qu'il n'y ait pas de phosphate d'aluminium en solution (C=O), la présence des PO^{3-}_4 peut participer à la variation du potentiel chimique $\Delta\mu_o$ qui intervient dans l'expression de l'énergie d'activation ΔG^*_p et donc créer une barrière d'énergie ΔG^*_d qui reste néanmoins faible. Pour augmenter celle-ci, la théorie prévoit qu'il suffit d'augmenter C, donc de mettre en solution le matériau considéré. C'est ce qui a été fait dans les acides sulfurique et phosphorique en éliminant l'acide chlorhydrique, du fait de ses difficultés d'emploi.

2.2.2 - Dissolution en solution acide sulfurique de phosphate d'aluminium (0 < C < Co)

* La vitesse de dissolution a été mesurée pour différentes concentrations en acide sulfurique et phosphate d'aluminium. Les valeurs reportées dans le tableau I montrent que, comme prévu, la présence de phosphate d'aluminium en solution abaisse la vitesse de dissolution. Les valeurs sont alors trop faibles pour envisager une application industrielle satisfaisante.

Tableau 1 : Vitesses de dissolution de lames AT de berlinite en milieu acide sulfurique seul (valeurs calculées d'après l'équation 1) ou en présence de phosphate d'aluminium (valeurs mesurées) dans le même acide

T (°C)	$[H_2SO_4]$ ($M.l^{-1}$)	Vcalc. ($\mu m.h^{-1}$)	$[H_2SO_4]$ + $[AlPO_4]$ ($M.l^{-1}$)	Vobs. ($\mu m.h^{-1}$)
90	8	15.7	8 / 3.6	3.4
80	5.25	6.1	5.25 / 3.5	2.3
90	5.25	11.0	5.25 / 3.5	4.4

* L'évolution de l'état de surface (Figure 3) montre une nette diminution de Ra lorsque le rapport $AlPO_4/H_2SO_4$ augmente. L'idéal aurait été de calculer le rapport C/Co. Malheureusement, les valeurs de Co, (solubilité de $AlPO_4$ dans H_2SO_4), ne sont pas connues aux températures utilisées et sont très difficilement mesurables. Cependant, même si les valeurs de rugosité atteintes (de l'ordre de 0,05 µm) semblent satisfaisantes, l'examen microscopique révèle la présence de quelques piqûres d'attaque, même dans les meilleurs cas. Il faut noter que le rapport molaire $AlPO_4/H_2SO_4$ = 0,66 est la valeur maximale que l'on peut atteindre expérimentalement pour des raisons purement chimiques (précipitation de sulfate d'aluminium en particulier).

Donc, dans le cas le plus favorable, les solutions sulfuriques ont le même comportement que l'acide phosphorique seul. Cette similitude peut s'expliquer par le fait que dans les 2 cas une barrière d'énergie ΔG^*_d est créé mais reste trop faible pour éviter toute formation de piqûre d'attaque : dans le premier cas (H_3PO_4) seuls les ions phosphate participent à cette barrière, dans le second cas, C reste trop éloigné de Co pour conduire à une valeur suffisante de ΔG^*_d.

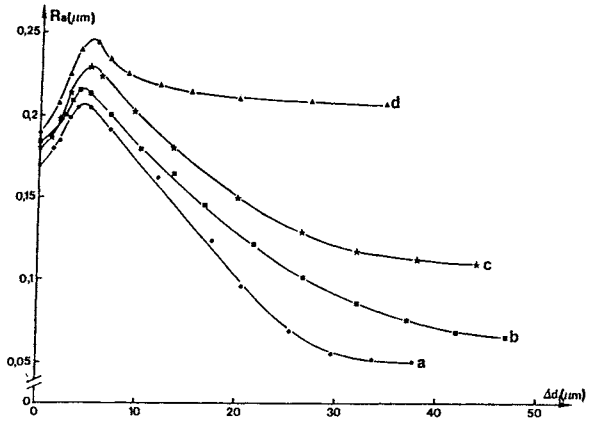


Figure 3 : Variation de la rugosité Ra de lames AT de berlinite en fonction de la dimution d'épaisseur Δd, au cours de leur dissolution dans différentes solutions de phosphate d'aluminium dans l'acide sulfurique. Valeurs du rapport $AlPO_4/H_2SO_4$: a) 0,66 - b) 0,45 - c) 0,25 - d) 0,04.

Au vu de ces constatations, nous avons donc testé les solutions de phosphate d'aluminium dans l'acide phosphorique.

2.2.3 - Dissolution en solution acide phosphorique de phosphate d'aluminium ($0 < C < Co$)

* Les vitesses de dissolution mesurées dans quelques solutions de phosphate d'aluminium dans l'acide phosphorique sont consignées dans le tableau II, en comparaison de celles calculées pour le même acide seul.

Tableau 2 : Vitesses de dissolution de lames AT de berlinite en milieu acide phosphorique seul (valeurs calculées d'après l'équation 1) ou en présence de phosphate d'aluminium dans le même acide.

T (°C)	$[H_3PO_4]$ (M.l^{-1})	Vcalc. (µm.h^{-1})	$[H_3PO_4]$ + $[AlPO_4]$ (M.l^{-1})	Vobs. (µm.h^{-1})
90	9	83.7	9 / 2.7	36.3
70	9	44.6	9 / 2.7	22.5
70	10.4	51.0	10.4 / 3.5	11.2

Comme précédemment et conformément aux prévisions, la mise en solution de phosphate d'aluminium diminue la cinétique de dissolution tout en conservant des valeurs relativement élevées (V > 10 µm/h), convenables pour leur utilisation pratique : rapidité et précision suffisantes. Dans ce type de solution, si la réaction de solubilisation peut se formuler, comme nous l'avons proposé selon (Réf.11) :

$$AlPO_4 + 3\,H^+ \qquad Al^{3+} + H_3PO_4$$

l'équation de vitesse doit s'exprimer suivant :

$$V = k_1\,[H^+]^\alpha - k_2\,[Al^{3+}]^\beta \,.\, [H_3PO_4]^\gamma$$

Cette équation reflète d'une part que la présence de phosphate d'aluminium dissout diminue la vitesse et que d'autre part, une détermination précise de la loi de vitesse nécessite la connaissance de la concentration des différentes espèces présentes. Ce problème complexe, et ceci d'autant plus que les solutions sont très concentrées, est en cours d'étude de façon à pouvoir choisir les concentrations en H_3PO_4 et $AlPO_4$ pour obtenir une vitesse préalablement fixée.

* Si l'on considère l'évolution de la rugosité (Figure 4) lorsqu'on utilise ces solutions (H_3PO_4 + $AlPO_4$), on constate une diminution rapide de celle-ci, dès le départ, pour atteindre une valeur moyenne de 0,05 µm après élimination d'environ 15 µm (7,5 µm sur chaque face) alors qu'il fallait dissoudre environ 30 µm pour obtenir un résultat comparable dans le cas de l'acide phosphorique seul (Fig. 2) ou des solutions de phosphate d'aluminium dans l'acide sulfurique (Fig. 3).

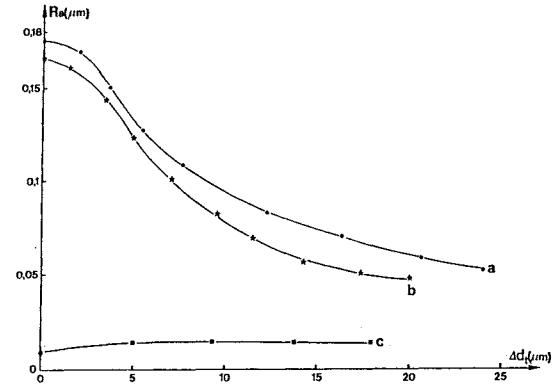


Figure 4 : Variation de la rugosité Ra de lames AT de berlinite en fonction de la diminution totale d'épaisseur Δd, au cours de leur dissolution dans différentes solutions de phosphate d'aluminium dans l'acide phosphorique à 80°C.
a) H_3PO_4 9M + $AlPO_4$ 2,7M; b) H_3PO_4 6,5M + $AlPO_4$ 3,2M;
c) H_3PO_4 9M + $AlPO_4$ 2,7M.

D'autre part, l'observation microscopique de ces lames amincies montre l'absence quasi totale de piqûres d'attaque, confirmant que, dans ce cas, la barrière d'énergie doit être suffisante pour éviter leur formation. Comme dans le cas de l'acide sulfurique, la solubilité du phosphate d'aluminium dans l'acide phosphorique ne pouvant être mesurée aux températures utilisées, une étude systématique en fonction de C/Co est malheureusement impossible.
En outre, si l'on étudie des lames polies (rugosité initiale voisine de 0,01 µm), l'état de surface est conservé tout au long de l'amincissement chimique (Fig. 4c) prouvant encore l'intérêt de ces solutions.

2.2.4 - Comparaison des différentes solutions. Les différents milieux étudiés peuvent être comparés au point de vue cinétique et au niveau de leur influence sur l'état de surface.

* Comparaison cinétique :

Le tableau III regroupe les principaux résultats à retenir en ce qui concerne les vitesses de dissolution des lames AT de berlinite.

Tableau 3 : Comparaison des vitesses de dissolution de lames AT de berlinite.

Solution	H_2SO_4	H_3PO_4	H_2SO_4+$AlPO_4$		H_3PO_4+$AlPO_4$
Equation de vitesse	$V = A \exp(-Ea/RT)\,[H^+]^{1/2}$		$V=k_1[H^+]^\alpha - k_2[Al^{3+}]^\beta[H_3PO_4]^\gamma$		
Constantes Ea	$A=2.5.10^9$ 63.4 kJ.Mol^{-1}	$A=6.25.10^{10}$ 63 kJ.Mol^{-1}	$k_1, k_2, \alpha, \beta, \gamma$ à déterminer		
T(°C)	90	90	90		90
[acide] (Mole.l^{-1})	8	9	8	9	10.4
$[AlPO_4]$ (Mole.l^{-1})			3.6	2.7	3.5
V(µm.h^{-1})	15.7	83.7	3.4	36.3	11.2

La vitesse de dissolution est toujours plus importante en milieu phosphorique que sulfurique et peut être facilement modulée en jouant sur les différents paramètres l'influençant : température et concentrations.

* Influence sur l'état de surface :

L'état de surface est caractérisé par rugosimétrie et microscopie optique.
Les résultats marquants des mesures de rugosité sont représentés sur la figure 5.

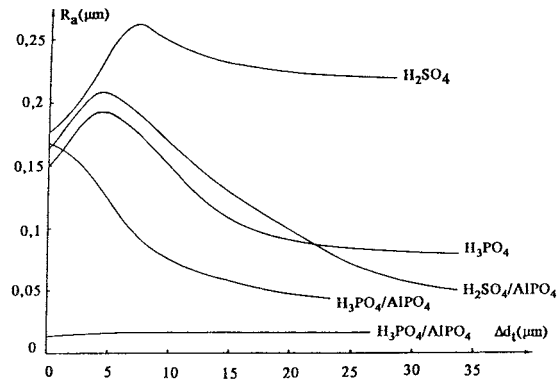


Figure 5 : Comparaison des différents milieux étudiés vis à vis de leur influence sur l'évolution de la rugosité au cours de l'amincissement chimique.

Ces résultats permettent de tirer les conclusions suivantes :

- l'acide sulfurique seul est à proscrire comme milieu d'amincissement chimique de la berlinite, il génère un état de surface trop perturbé.
- l'acide phosphorique seul ou les solutions de phosphate d'aluminium dans l'acide sulfurique conduisent à des états voisins, mais plus rapidement en milieu phosphorique. Dans les 2 cas, la présence de piqures d'attaque n'est pas favorable à la réalisation de dispositifs piézoélectriques même si la rugosité atteinte peut sembler convenable.
- les solutions de phosphate d'aluminium dans l'acide orthophosphorique H_3PO_4 + $AlPO_4$ conduisent rapidement à une surface lisse, exempte de piqûre d'attaque. Conformément aux prévisions théoriques il y a bien alors dissolution contrôllée.

3 - CARACTERISATIONS PIEZOELECTRIQUES

Cette partie concerne la réalisation et la caractérisation de résonateurs AT de berlinite à hautes fréquences.

3.1 - Préparation des échantillons.

Il s'agit d'obtenir des lames dont la fréquence est de l'ordre de 73 MHz, ce qui correspond à une épaisseur de 20 µm environ. Trop fragiles, de tels échantillons ne peuvent pas être utilisés s'ils ne possèdent pas une couronne extérieure plus épaisse pour les consolider. Le processus de fabrication comprend donc plusieurs étapes d'usinage.

- usinage mécanique : après sciage et rodage, les lames issues des blocs de berlinite sont amincies par polissage jusqu'à 15 MHz environ (100 µm).
- usinage ionique : des caches ("Iris") sont ensuite centrés sur les lames afin de protéger la couronne extérieure des lames (Fig. 6a et 6b). Introduits dans un usineur ionique, les échantillons soumis au bombardement sont amincis et amenés sur la partie centrale des pièces à 24,5 MHz (60µm).On obtient alors des lames de 100µm sur la couronne extérieure et de 60µm au centre (Fig. 7a).
- usinage chimique : les conditions retenues sont : H_3PO_4 6M/$AlPO_4$ 2M à T = 85°C. Cette étape va permettre d'atteindre 73 MHz par attaque simultanée des deux faces. La vitesse d'usinage est de 13 µm/h. La figure 7b précise les épaisseurs obtenues après usinage.

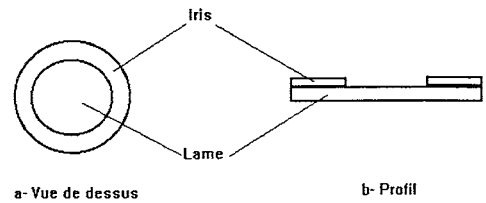


Figure 6. Procédé de masquage pour l'obtention de lames en puits.

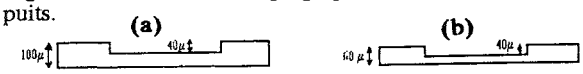


Figure 7. Profil des lames après bombardsement ionique (a) et usinage chimique (b).

3.2 - Caractérisation piezoélectrique.

Des électrodes métalliques sont déposées par évaporation et centrées sur les deux faces au milieu du puits. Le résonateur ainsi réalisé est connecté à un analyseur de réseau permettant de mesurer les différents paramètres électriques. Les résultats présentés ici concernent les performances atteintes par un résonateur en puits, usiné par voie chimique de 24,5 à 74 MHz, comparées à celles d'un résonateur usiné par bombardement ionique jusqu'à 73 MHz (Tableau 4).

Les valeurs mesurées sur les résonateurs usinés par voie chimique sont comparables à celles couramment obtenues sur des résonateurs usinés ioniquement jusqu'à la fréquence finale. L'état de surface des deux résonateurs sont donc similaires. L'usinage chimique n'a pas détérioré la lame. En comparant ces valeurs à celles du quartz pour une fréquence similaire (Tableau IV), on constate que la résistance des résonateurs berlinite est nettement supérieure à celle observée pour le quartz. Ceci a pour effet de faire chuter la valeur du coefficient de surtension Q des résonateurs. Cependant, dans le cas d'un montage en filtre, les valeurs plus faibles de la self de la berlinite constituent un avantage par rapport au quartz. De plus, le coefficient de couplage plus élevé de la berlinite (k = 11,4 % contre 8,7 % pour le quartz AT) permet d'obtenir des filtres avec une bande passante 30 % plus large que pour le quartz. La figure 9 montre le résultat d'un calcul de simulation qui donne l'évolution de la largeur de la bande passante en fonction de la fréquence pour le quartz et la berlinite et ceci pour une même configuration de filtre.

Tableau 4: Comparaison des performances de résonateurs de berlinite obtenus par usinage ionique ou chimique à ceux du quartz.

MATERIAU	BERLINITE		QUARTZ
k	11.4		8.7
	Usinage ionique (Fig. 8a)	Usinage chimique (Fig. 8b)	
Fr (MHz)	73.3175	74.315	73
R (Ohms)	49.1	52	25
L (µH)	2089	2041	3000
Q	20000	19000	55000
Q.Fr	1.4 E + 12	1.4 E + 12	4 E + 12

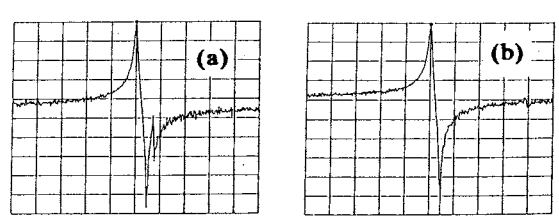


Figure 8: Réponse piezoélectrique de résonateurs usinés ioniquement (a) et chimiquement (b).

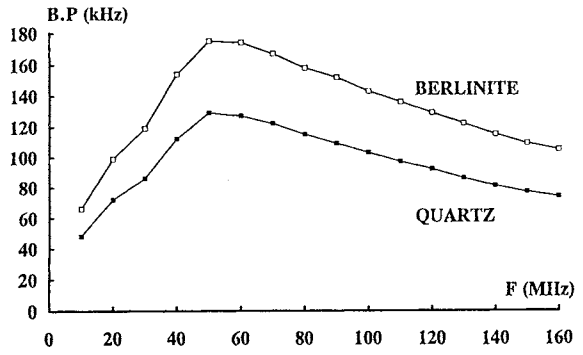


Figure 9. Evolution de la largeur de bande passante en fonction de la fréquence pour le quartz et la berlinite (calcul de simulation).

Malgré une qualité cristalline encore inférieure au quartz, tous ces résultats montrent que la berlinite présente un avantage certain puisqu'elle permet de fabriquer des filtres à ondes de volume à hautes fréquences avec de larges bandes passantes.

4 - CONCLUSION

Les premiers essais d'amincissement chimique de lames AT de berlinite menés en milieu acide sulfurique puis phosphorique ont montré que la mise en fréquence de résonateurs était possible mais l'état de surface restait défectueux. Nous nous sommes alors reportés à la théorie de la croissance et de la dissolution des cristaux en solution: elle prévoit que ces deux processus réversibles se déroulent d'autant mieux qu'ils ont lieu dans des conditions proches de l'équilibre, en particulier avec une faible sur ou sous-saturation. Nous avons alors repris nos études de dissolution dans des solutions acides de phosphate d'aluminium: les cinétiques de dissolution ont alors diminué et, conformément à nos attentes, l'état de surface final s'est nettement amélioré, conduisant à des résultats tout à fait satisfaisants dans le cas des solutions de phosphate d'aluminium dans l'acide phosphorique (rugosité $\leq 0{,}05\mu$m.et absence de piqûres d'attaque). Avec de telles solutions, qui permettent donc de contrôler la dissolution du matériau, la réalisation de résonateurs de haute fréquence de bonne qualité est tout à fait possible. La comparaison de leurs caractéristiques avec celles de résonateurs obtenus par usinage ionique montre qu'elles sont similaires et donc que cette méthode est tout à fait intéressante au niveau industriel, surtout si on garde en mémoire l'avantage de la berlinite sur le quartz en ce qui concerne la bande passante des filtres (30% plus large).

Une étude plus complète visant à établir l'équation de vitesse de dissolution quelle que soit la composition (concentration en acide et en phosphate d'aluminium) est en cours au laboratoire. D'autre part, nous envisageons d'étendre notre étude à d'autres types de matériaux.

5 - REFERENCES

1. Philippot E, Goiffon A, Maurin M, Detaint J, Schwartzel J, Toudic Y,Capelle B, Zarka A 1990, J Cryst Growth,104,713.
2. Chang Y P, Barsch G H 1976, IEEE Trans Sonics Ultrasonics,SU-23,127.
3. Wang H, Xy B, Liu X,Han J, Shan S, Li H 1986,J Cryst Growth, 79,227.
4. Cambon O, Goiffon A,Philippot E 1991, J Mat Sci,26,846.
5. Frenkel J 1945, J Phys USSR,9,392.
6. Burton W K, Cabrera N,Frank F C 1949,Nature,163,398.
7. Burton W K, Cabrera N,Frank F C 1951,Phil Trans Soc London,243,299.
8. Bennena P, Gilmer G H 1972, J Cryst Growth,13,148.
9. Johnston W G 1962, Prog Ceram Sci, 2, 1.
10. Cabrera N, Levine N 1955,Phil Mag, 1, 450.
11. Cambon O, Goiffon A, Ibanez A, Philippot E 1992, Eur J Solid Inorg Chem, à paraître.

CRYSTAL GROWTH AND PHYSICAL CHARACTERIZATIONS OF $GaPO_4$

E. Philippot[1], A. Ibanez[1], A. Goiffon[1], B. Capelle[2], A. Zarka[2], J. Schwartzel[3] and J. Détaint[3]

1 UM II, CNRS-URA D0407, case 003, F-34095 Montpellier cédex 5, France
2 Lab. Minér. Cristallog., Univ. Paris VI, F-75252 Paris cédex, France
3 CNET, PAB/BAG/MCT, F-92220 Bagneux, France

ABSTRACT

This investigation concerns $GaPO_4$ crystal growth in sulphuric and phosphoric acid media through the slow heating and vertical reverse temperature gradient methods. Systematic study of growth parameters shows a Vx growth rate always much greater than the two other ones: Vz and Vy. The most interesting result is the ability to do $GaPO_4$ epitaxy on large berlinite seeds in sulphuric acid and, after that, to use them for $GaPO_4$ crystal growth in phosphoric acid. The good epitaxy process and the crystalline quality have been checked by X-ray topography. The [OH] content, followed by infrared spectrometry, seems to prove a lower OH concentration than in the berlinite case for approximatively the same crystal growth conditions. First piezoelectric characterizations of resonators near the AT-cut show $GaPO_4$ to be a very promizing piezoelectric material with a large coupling coefficient near 16% , a quartz like thermal stability and a Q factor already nearly sufficient for the applications.

Keywords: piezoelectric material, infrared spectrometry, X-ray topography, piezoelectric characterizations.

INTRODUCTION

Since the work of Chang and Barsch (Ref. 1), many works have been devoted to growing and characterizing, first berlinite and recently $GaPO_4$ crystals, and to modelling and experiment of bulk and surface wave devices made with these materials. It was successively shown that these materials were close piezoelectric analogs of quartz, and possess similar cuts with zero frequency temperature coefficients, either for bulk or for surface wave applications (Ref.2-8).

For the same AT cut, the first results obtained for $GaPO_4$ crystals (Ref.9,10) seem to indicate a value of coupling coefficient (k>0.14) higher than those of quartz (0.085) and berlinite (0.11).This induces the possibility to obtain larger band-pass of filters. On the other hand, there is no $\alpha \leftrightarrow \beta$ quartz transition for this material and, then, it can be used in a large range of temperature, T>1000°C. We present here new crystal growth approachs for $GaPO_4$ in connection with some results of scanning electron microscopy, infrared spectrometry, X-ray topography investigations and resonator measurements.

1. CRYSTAL GROWTH STUDIES

1.1 Powder and nutrient synthesis.

Gallium phosphate being not commercially available, it has been synthesized from gallium hydroxyde with modest purity or gallium 6N metal. In the first case, the reaction :

$$H_3PO_4 + GaO(OH) \rightleftharpoons GaPO_4 + 2\,H_2O$$

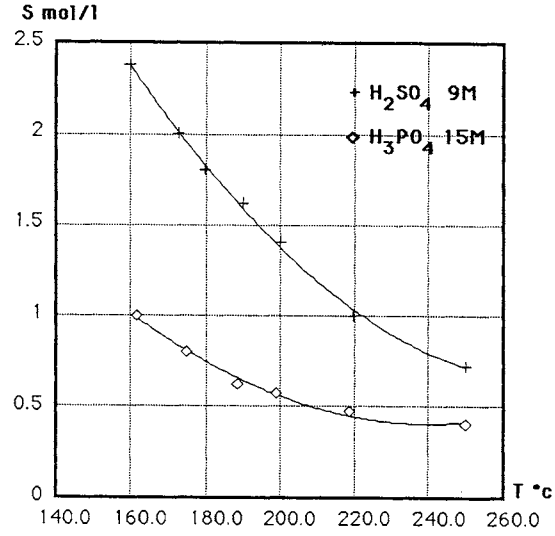


Figure 1. Berlinite solubility in sulphuric and phosphoric acids.

is carried out with an excess of phosphoric acid. As berlinite, gallium phosphate shows a retrograde solubility in all acid media, (Figure 1). The crystallization is obtained in tantalum lined autoclave. To eliminate the major part of the initial impurities, it is needful to use a re-crystallization experiment. Operating conditions are : solution 15M in H_3PO_4 saturated with gallium hydroxide, filling 80%, final temperature near 240°C. The heating rate is chosen following the wanted granularity : the slower one (# 20°C/h) for nutrient synthesis and the faster one (#100°C/h) to obtain finely powder, useful for growth solution preparations.

The second reaction needs two steps :

$$Ga + 4\,HNO_3 \rightleftharpoons Ga(NO_3)_3 + \overset{\nearrow}{NO} + 2\,H_2O$$

$$Ga(NO_3)_3 + H_3PO_4 \rightleftharpoons GaPO_4 + 3\,HNO_3.$$

The nitric solution of $GaPO_4$ so obtained is then evaporated by moderate heating and, in this case, finely powdered gallium phosphate can be obtained with an yield higher than 90% and a good purity.

1.2 $GaPO_4$ crystal growth

The $GaPO_4$ crystal growth has been carried out through the three methods already used for berlinite (Ref.) :

- the slow heating method (SHT);

- the horizontal composite gradient method (HTG);
- the vertical reverse temperature gradient method (VTG).

Only new advances on crystal growth technique will be described here :

- VTG method in H_2SO_4 acid;
- new composite gradient method in H_3PO_4 acid.

1.2.1 GaPO4 crystal growth by VTG method in H_2SO_4 acid. To obtain large seeds of gallium phosphate, it would be interesting to use directly large berlinite seeds five to eight centimeters long now available. Indeed, it must be kept in mind the continuous solid solution and the cell volume similarity for the whole pseudobinary system $Al_{(1-x)}Ga_xPO_4$, V=(230±1)Å^3 (Ref.11). This being so, the growth study of $GaPO_4$ crystals from berlinite seeds was undertaken.

In a first step, some experiments have been carried out in phosphoric acid medium . Unfortunately, the dissolution kinetics of $AlPO_4$ in this solvant is faster than the growth restart one of $GaPO_4$, and so berlinite seeds are dissolved.

Therefore, we have tried this crystal growth in sulphuric acid medium. Indeed, although all MXO_4 quartz isomorphs present their highest solubility in this solvent (#3 times higher than that one in phosphoric acid), much lower dissolution kinetics has been demonstrated (Ref.12). On the other hand, as for berlinite crystals, growth temperature must be chosen greater than 200°C to avoid major part of -OH impurity. In this case, because of too high pressure conditions, we cannot use the SHT method, but the VTG one.

In order to determine the crystal growth conditions of $GaPO_4$ in sulphuric acid solvent, we have undertaken a systematic study in terms of :

- crystal growth temperature, Tc (°C);
- temperature difference between cold and hot zone,ΔT (°C);
- acid concentration, M (mol l^{-1}).

We present here the first results in sulphuric acid solutions taking into account the following conditions :

- **berlinite seeds;**
- autoclave filling #80%;
- acid sulphuric solution, 6M, saturated in GaPO4 at the chosen growth temperature, Tc;
- growth temperature ranging from 200 to 240°C;
- vertical temperature difference (or temperature gradient): 1<ΔT<6°C;
- boring ratio of the baffle between cold and hot zones : 7,5%;
- nutrient mesh : 2-5 mm.

Although the experiments are still in progress, some interesting results can be emphasized. The most important features are :

- the corroboration of the crystal growth restart in $GaPO_4$ from berlinite seeds;
- the confirmation of kinetic results for berlinite controlled dissolution in both phosphoric and sulphuric acid media.

The fig. 2 shows the growth rates results as a function of ΔT at Tc= 210 and 240 °C. As for berlinite crystal growth, Vy growth rate is always close to 0.0 mm/d.face. On the other hand, another growth rate, Vz, is very weak, 0.01<Vz<0.02 mm/d.face and independant of Tc value.

This Vx and Vz growth rate disparity explains the strongly different morphology of crystals obtained from X or Z berlinite seeds. From a general point of view, from X berlinite seeds, all crystals are cracked along a preferential direction parallel to $(10\bar{1}2)$ plans, during the air cooling of autoclaves at the end of the crystal growth. The photo 1 compares such a crystal with its crystal structure: the plans $(10\bar{1}2)$ correspond to the maximum atom density of MXO_4 packing. These cracks are due to strong strains between $AlPO_4$ and $GaPO_4$ structures. Indeed the GaO_4 tetrahedra bigger than the AlO_4 ones for the same cell volume and explain these strong strains.

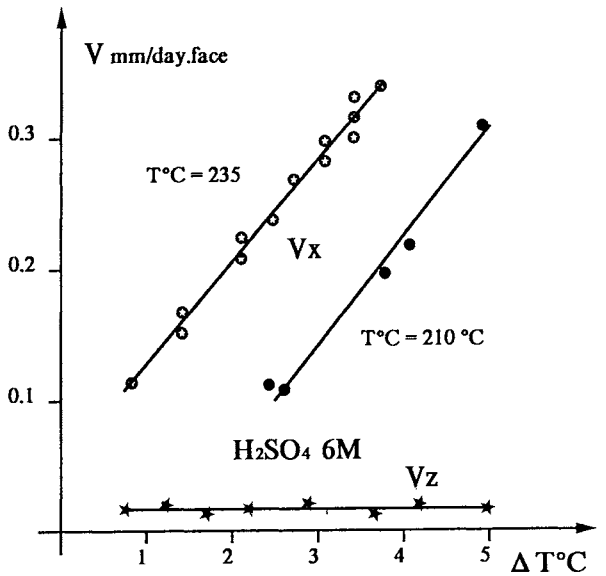


Figure 2. Vx and Vz growth rates for Tc = 210and 235°C in sulphuric acid medium, 6M

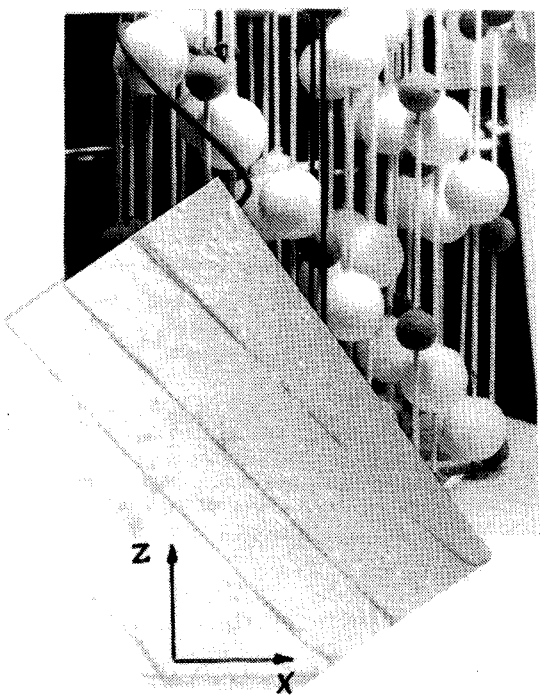


Photo 1. Cracked $GaPO_4$ crystal from X-seed berlinite with its schematic crystal structure.

The most interesting results are obtained from Z seeds. In this case, Vx being the only significant growth rate, flat crystals are obtained with a very thin epitaxy of $GaPO_4$ in the Z direction. Photo 2 and 3 give an idea of such crystal at Tc=210 and 235 °C. At lower temperature, berlinite seeds are less dissolved than at the higher one. Both kinds of seeds can be used for further growth experiments in phosphoric acid medium, where Vz ≠ 0, through the SHT method.

1.2.2 GaPO4 crystal growth by SHT method in phosphoric acid. This method has been also used for crystal growth from nucleation, but only the experiments using seeds obtained from the previous (§1.2.1) large Z crystals will be described. It is a modified version of SHT method (2), where the slow heating rate

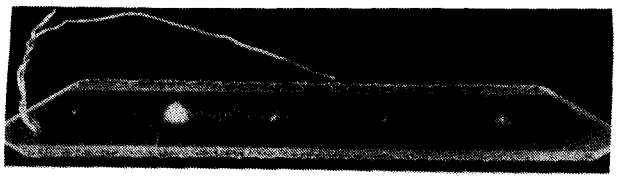


Photo 2. Flat crystal obtained from berlinite Z-seed at Tc=210°C in sulphuric acid medium, (VTG method).

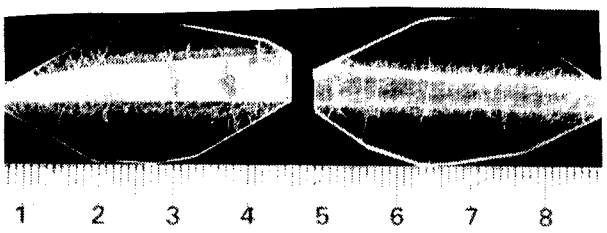


Photo 3. Flat crystal obtained from berlinite Z-seed at Tc=235°C in sulphuric acid medium, (VTG method).

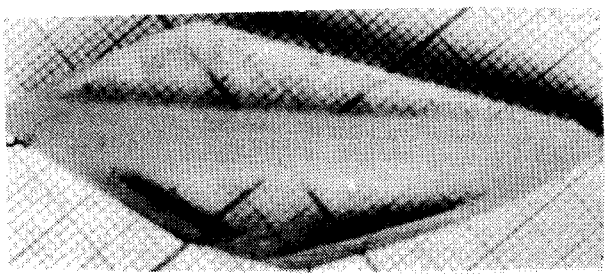


Photo 4. GaPO4 crystal obtained in phosphoric acid medium through the SHT method between 150 and 180°.

is coupled to an horizontal symetric gradient temperature as is shown in Fig.3. In these conditions, solution circulation is favoured and avoids gradient concentration near the growing crystal interface. In this modification, 150 cm^3 horizontal glass vessels are used. Experimental conditions are :

- **berlinite seeds shielded against dissolution** by thin epitaxy of GaPO 4;
- filling # 80%;
- acid phosphoric solution, 15M, saturated with GaPO4 at the initial temperature;
- growth temperature increasing from 150 to 180°C with 0.5 to 2.0 °C/day growth rate;
- horizontal temperature gradient # 5°C.

The photo 4 gives the growth result from flat crystal shown in photo 2. Whereas the Vz growth rate becomes significant, its value remains weaker than the Vx one. We observe that the Vx/Vz growth rate ratio is approximatively reverse to that of berlinite, 1/3 for $AlPO_4$ and 3/1 for $GaPO_4$.

Nevertheless, the Z-thickness is adequate to cut, in these crystals, Z-seeds which can be use in sulphuric acid, through the RTG method at higher temperature, to give crystals with low -OH content. Several alternatives to this crystal growth process, using successively the VTG method in H_2SO_4 and the SHT method in H_3PO_4, are in progress.

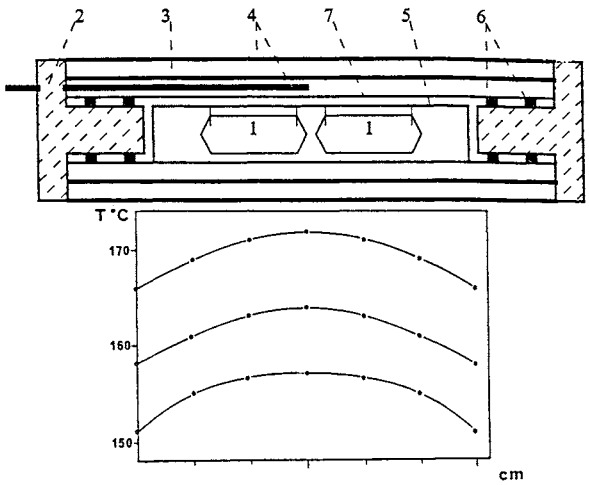


Figure 3. Schematic representation of the silica tube autoclaves, and their temperature profile, used for SHT method :
1 = seeds, 2 = PTFE plugs, 3 = regulation and control thermocouples, 4 = resistance-wound furnace between two transparent silica tubes, 5 = platinum rod, 6 = quadrangular elastomer O-rings, 7= glass tube.

At this step, some preliminary characterizations have been undertaken to check :

- the -OH content by infrared spectrometry;
- the crystalline quality by X-ray topography;
- the piezoelectric characteristics by resonator device measurements.

2. CHARACTERIZATIONS OF $GAPO_4$ CRYSTALS

2.2 -OH content determination by infrared spectrometry.

This method, as for quartz and berlinite, is the most appropriate one for the -OH content determination and can be related to the surtension coefficient. Figures 4 give an idea of transmittance and absorption infrared spectra obtained from our crystals, in the range 4000-2600 cm^{-1} characteristic of -OH absorption. The spectra of Fig. 4a are representative of low temperature crystal growth, SHT from 150 to 180°C, whereas in Fig. 4b are shown those of higher temperature crystal growth, VTG above 200°C.

In both cases, we have calculated the absorption coefficient, α, from the expression (Ref.13) :

$$\alpha = 1/d_{cm} \log(T_{3800}/T_{3167})$$

where the stronger absorption band at 3167cm^{-1} is probably due to the Ga-OH one (for berlinite, it was used the Al-OH band at 3291 cm^{-1}). For the low temperature sample, Fig. 4a, the α value is 1.78 whereas it becomes 0.04 for the higher temperature one, Fig. 4b. Therefore, it seems to exist a strong difference between berlinite and gallium phosphate behaviour. Indeed, if we compare these α values to those ones for berlinite, for approximatively the same crystal growth conditions, it can be observe a diminution of a factor 10 for gallium phosphate (Ref.8). This result presents, of course, a major interest, but must be confirmed.

2.3 Characterization of growth defects by X-ray topography.

Several plates of $GaPO_4$ were studied by X-ray transmission topography using the white beam delivered by the DCI storage ring at LURE (Orsay, France) or by conventional X-ray apparatus.

In photos 5 are given typical examples of the observation made using etched plates from crystals grown through the SHT method. The cut of photo 5a contains one isolated electrical twin (Dauphiné law) probably induced by the fixation of the seed. The right part of this plate is of quite good quality with some dislocations at the growth restart. The plate represented in photo

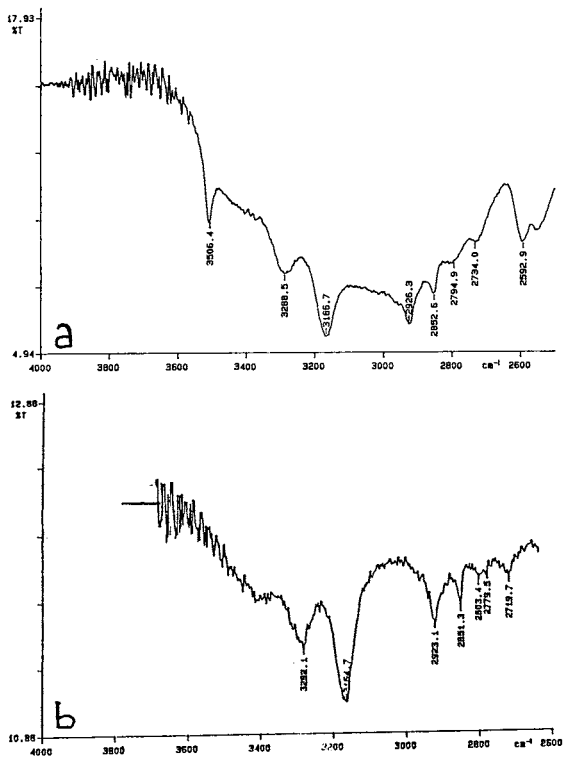


Figure 4. Infrared transmission of GaPO4 samples:
a): crystal from SHT method, 150<T<180°C
b): crystal from VTG method, Tc<200°C.

5b contains one small optical twin (Brazil law). Twinning was much more frequently found in the first crystals obtained by successive growth of small nucleations than in recent crystals obtained by the new growth method starting with high quality berlinite seeds, where only small isolated twins of the Brazil type can be observed in some rare instances.

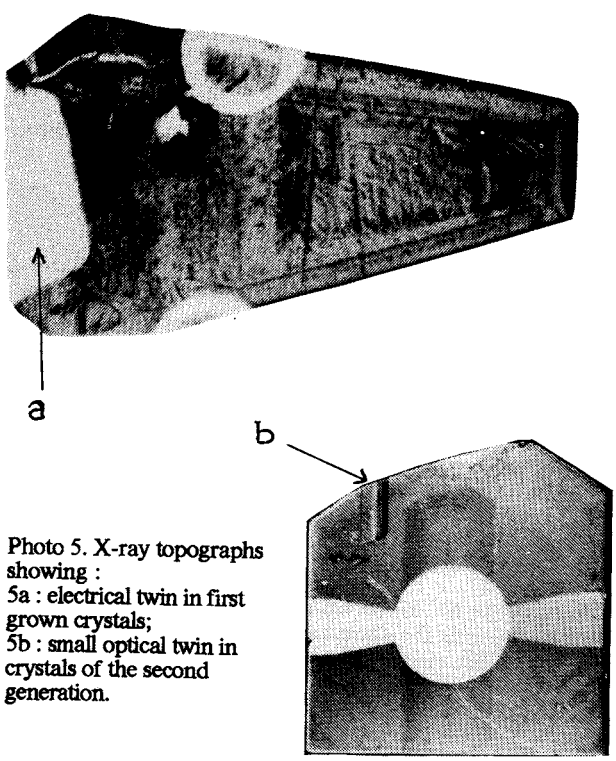


Photo 5. X-ray topographs showing :
5a : electrical twin in first grown crystals;
5b : small optical twin in crystals of the second generation.

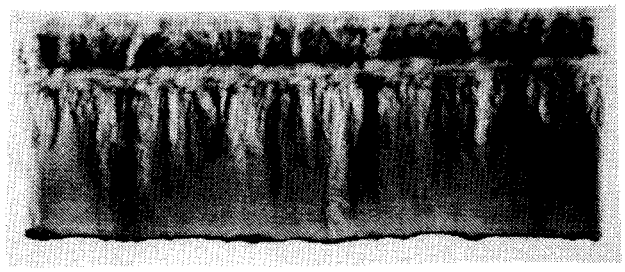


Photo 6. Topograph of a Z plate obtained by VTG method in sulphuric acid medium from berlinite seed.

Z plates obtained by VTG method in sulphuric acid medium from berlinite seeds were most often observed to be twin free. In this topograph, a large contrast, corresponding to a highly distorted region existing between the berlinite seed and theGaPO4 lateral overgrowth, can be observed. The GaPO4 growth appears to have a very good crystalline quality, photo 6.

2.3 Piezoelectric Characterizations.

2.3.1 Bulk piezoelectric wave propagation in the Y rotated cuts of GaPO4. As for quartz and berlinite, the piezoelectrically excited pure shear mode of the Y rotated cut of GaPO4 appears to be very attractive since quite simple orientations giving large coupling coefficients and a good thermal behaviour exist among these cuts. Two Y rotated compensated orientations which are of the same nature as the AT and BT cuts of quartz were reported by Krempl et al (Ref. 9,14,15). Several authors have characterized GaPO4 crystal using dielectric constant measurements (Ref. 16). Huart (Ref.17) and Krempl et al. (Ref.9,14,15) have determined several sets of elastic and piezoelectric constants of GaPO4.

Using these constants we have determined the main properties of Y rotated thickness shear resonators. For these computations we use the one dimensional model of an infinite plate with massless electrodes (Ref.18,19). On figures 5 is represented computed properties of the slow shear mode as a function of the rotation angle theta. On these figures, V is the phase velocity of the mode ($V=\sqrt{C/\rho}$), k is the electromechanical coupling coefficient; Tfr, Tfr3, Tfr5, Tfa are respectively the first order temperature coefficients of the fundamental resonance frequency, of the 3rd overtone resonance frequency, of the 5th overtone resonance frequency and of the antiresonance frequencies. The results given in figure 5 are mostly based on the constants given in (Ref.17); the calculation of the temperature coefficients is very tentative since the available data (Ref.9) constitute an incomplete set which was completed with values measured for berlinite (Ref.20).

The obtained results for V and k indicate that gallium phosphate is a material characterized by a quite small phase velocity of the shear wave and by rather large values of the coupling coefficient. Despite the some how arbitrary hypothesis made in choosing the

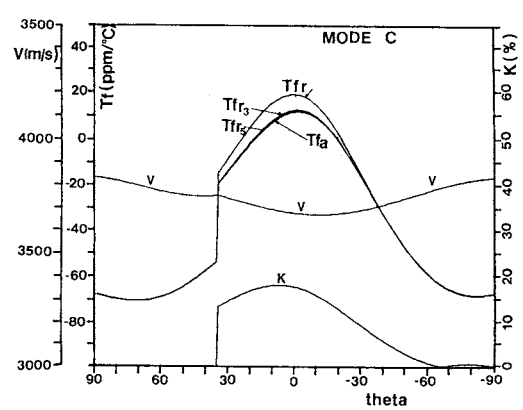


Figure 5. Computed properties of the slow shear mode.

missing temperature coefficients (TC) of the constants, the computed temperature coefficients are found to be in good agreement with all the known experimental values. Particularly the AT and BT cuts are found at angles very close to those given in Ref 9 and 15. Among the computed temperature coefficients, those relative to the antiresonance frequencies are the closest from the experimental values. This is probably a consequence of the formula used in Ref. 9 to extract the TC of the constants.

Other calculations were made using the constants given in Ref. 15 and other hypothesis concerning the unknown temperature coefficients. One example is given in figure 6 for the slow shear mode of cuts in the vicinity of the AT and of the BT cuts. The results are very similar to those given in figure 5. In this figure we can observe that in $GaPO_4$, the BT cut occurs at an angle for which the slow shear mode is the piezoelectrically excited mode whereas for quartz the mode of the BT cut is the fast shear one. The cut angle for the AT and the BT cuts are much smaller than those existing in quartz and in berlinite. This fact is an important advantage since the value of the coupling coefficient is then a

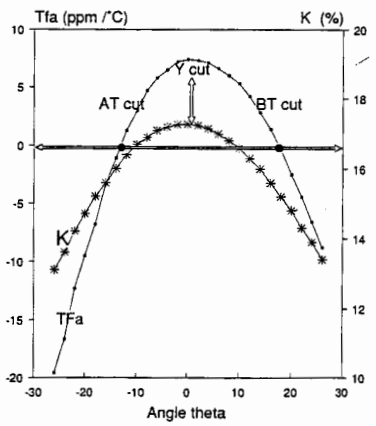


Figure 6. Slow shear mode in the vicinity of the AT and of the BT cuts.

larger fraction of the maximum possible value which, as in other quartz like materials, corresponds to the Y cut. As in other quartz like materials, the AT cut is situated very near a minimum of the phase velocity as a function of the angle.

2.3.2 Experimental results: Resonators were made from plates cut in crystals grown by the SHT method:

- by successive growth of spontaneous nucleations;
- from berlinite seeds with epitaxy of $GaPO_4$.

For these measurements, plates free of twins or having large regions free of twins were selected using X-ray topography and the optical examination of etch figures obtained in an acid solution (5% HF and 10% HCl), photos 5. The plates were electroded using Cr/Au evaporated thin films with a mass loading ranging from 0.6% to 0.9% . Electrical measurements were performed using network analysers and a high precision oven monitored by a quartz thermometer.

2.3.2a Resonators cut with the minor rhomboedral orientation. Plates parallel to the rhomboedral minor face (Y-21°) were obtained from one of the first crystals grown over a seed obtained from nucleation. About ten plates of sufficient size were obtained, 6 were determined as being free of twins or having sufficiently large area free of twins by etch figure examination. A typical case of this kind of plate is given in photo 5a . The electrical response of this kind of resonator is very clean and similar to the previously published results, nevertheless, the Q factor is still quite modest (5000 at 5 MHz). The thermal variations of the resonance frequencies of some of these resonators are displayed in figure 7. Important variations of the Q factors as a function of temperature were observed during this study (see figure 10 below). This fact is to be related to the still noticeable OH concentration existing in this type of crystal (Ref.8).

2.3.2b Properties of resonators near the AT cut. About twenty plates having two orientations around AT (Y-15° and Y-12°30') were obtained from a crystal more recently grown from a seed

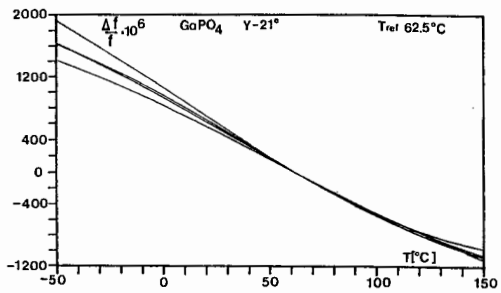


Figure 7. Thermal variations of the resonance frequencies of the first generation $GaPO_4$ crystals.

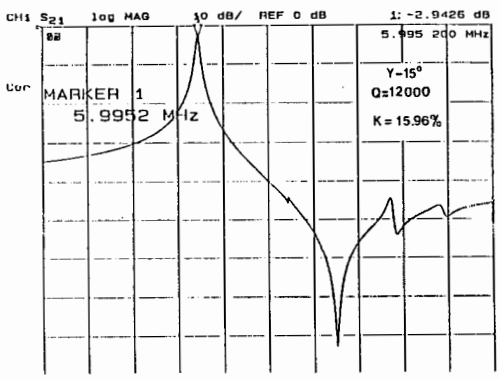


Figure 8. Typical electrical response observed for resonators from crystals grown from berlinite seeds.

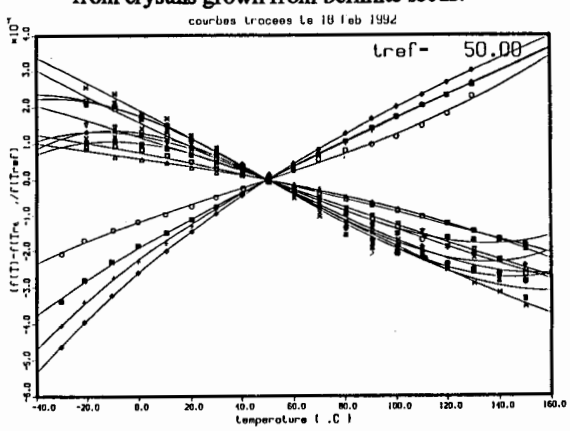


Figure 9. Thermal variations of the resonance frequencies of the second generation $GaPO_4$ crystals.

obtained by lateral epitaxy over berlinite using improved growth conditions. In figure 8 is represented a typical electrical response observed for these resonators. The Q factor of this resonator is 10000 at room temperature, which corresponds to about 12000 at the usual reference frequency of 5 MHz. A much lower dispersion than for the Y-21° cut was observed on this quantity between resonators. The large value of the relative difference between the resonance and antiresonance frequencies can be directly related to the value of the coupling coefficient which was found very near 16%. On figure 9 are represented the thermal variations of the resonance frequencies observed for resonators of the two cuts near AT.

On figure 10 we have compared the thermal variations of the insertion losses measured for the crystal used to obtain the near AT cuts and the crystal considered in 2.3.2a. In this figure, we can observe a large difference of behaviour between these two crystals. For the first crystal, an important loss peak appears in the vicinity of the room temperature whereas the second crystal displays a much smaller and more constant resistance as a function of temperature ($R=100(10^{A/20}-1)$ where A is the insertion loss). This situation was already encountered for berlinite crystals for which the existence of a loss peak at room temperature was related to the OH concentration (Ref.8,21).

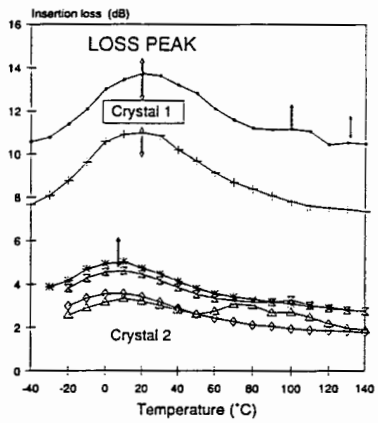


Figure 10. Compared thermal variations of the insertion losses measured for the two types of crystals.

2.3.3 Comparison of experimental and computed results, discussion. Measured and computed results are compared in Table I. In this table, a generally good agreement is found between experimental and computed values. It must be emphasized that the computed values of the coupling coefficient are attained in practically all experimental cases without any corrections for the existing parasitic capacitance of the resonators. The experimental Q factors, which are already nearly sufficient for the applications for the second crystal, underestimate probably by a factor 2 to 3 the material intrinsic Q factor that can only be measured using planoconvex overtone resonators. Somehow, higher values were measured for the overtones modes of the present plane resonators. An important observation can be made relatively to the experimental frequency constants (Nfa=V/2): they are always higher than the computed values (the accuracy of this measurement is better than 0.1%); this fact is probably significative of a much lower OH concentration in the present samples than in those used for the constant measurements. It can be observed that this difference is greater for the most recent crystal for which a greater Q factor was measured. A similar increase of the elastic constant was already observed with berlinite when the OH concentration was reduced by modification of the growth conditions. The Q factors measured for the second crystal are, to our knowledge, the best results known for this material. It is expected that higher values will be soon obtained. The performances observed are very promising for applications to filters, voltage controlled oscillators and transducers.

cut	$NFr^{\times}$ kHz.mm	$NFa^{\times}$ kHz.mm	$K^{\times\times}$ %	TFr ppm/° C	TFa ppm/° C	Q at 5MHz
E Y-21	1258.3	1266.4	14.2	- 14.1	-15.4	5000
C Y-21	1246.9	1257.7	14.5	-	-11.7	
E Y-15	1261.0	1274.1	15.97	-2.39	-3.66	12800
C Y-15	1242.5	1255.4	15.86	+1.90	-2.13	
E Y-12.5	1255.1	1269.7	15.98	+3.37	-	10100
C Y-12.5	1243.6	1257.3	16.31	+4.48	+2.60	

(C) computed, (E) experimental
× Values corrected of the mass loading
×× max. value observed for this cut

CONCLUSION

In this investigation, it has been demonstrated the feasability of $GaPO_4$ epitaxy on large berlinite seeds in sulphuric acid medium. From these new seeds, the crystal growth of large $GaPO_4$ crystals can be directly undertaken in phosphoric acid medium. At this time, the use of both solvents is needful because only Vx growth rate has a significative value in sulphuric acid medium. This process has been proved to be efficient and reproducible at the laboratory scale. In view to obtain a method more easily transposable to a large industrial scale, new investigations are currently in progress.

From first characterizations, it can be asserted the good quality of $GaPO_4$ epitaxy on berlinite seeds and the good crystalline quality of crystals obtained from these seeds. On the other hand, the [OH] content has been followed by infrared spectrometry. If we compare to berlinite crystals, it seems that the [OH] content is approximatively lowered by a factor 10 for the same growth conditions : growth temperature, acid concentration, ...

The piezoelectric measurements made on resonators cut in two types of crystal have shown that a large improvement of the properties of resonators were obtained with the refinement of the growth conditions and that a decisive advance has been made with the use of berlinite seeds. The reduction of the OH concentration and the improvement of the crystalline quality has permitted a significative gain for the Q factor and a great reduction of the thermal variations of the equivalent resistance. On the whole these results are a confirmation of the large interest of a material having simultaneously a large coupling coefficient, a quartz like thermal stability and the possibility to lead to large Q factors.

REFERENCES

1. Chang Y.P. & Barsch G.H. 1976, IEEE Trans. Sonics Ultrasonics SU-23, 127.
2. Jumas J.C., Goiffon A., Capelle B., Zarka A., Doukhan J.C., Schwartzej J., Détaint J. & Philippot E. 1987, J. Crystal Growth, 80, 133.
3. Kolb E.D. & Laudise R.A. 1978, J. Crystal Growth, 43, 313.
4. Croxall D.F., Christie T.R.A., Isherwood B.J., Todd A.J., & Birch 1979, 2nd Conf. on Crystal Growth, Lancaster.
5. Chai B.H.T., Shand M.L., Buehler E. & Gilleo M.L. 1979, IEEE Ultrasonics Symp., 577.
6. Nagai K., Ogawa T., Okuda J., Asamara J. & Taki S. 1982, 1st Inter. Symp. on Hydrothermal Reactions, 496.
7. Philippot E., Jumas J.C., Goiffon A. & Astier R. 1984, French Patent 2,559,166.
8. Philippot E., Goiffon A., Maurin M., Détaint J., Schwartzel J., Toudic Y,. Capelle B. & Zarka A. 1990, J. Crystal Growth, 104, 713.
9. Engel G., Krempl P. & Stadler J. 1989, 3rd Eur. Time and Freq. Forum, 50.
10. Engel G. & Krempl P. 1988, 2nd Eur. Time and Freq. Forum, 827.
11. Bennazha J., Cachau D., Ibanez A., Goiffon A. & Philippot E, 1992, soumis au Europ. J. Solid State and Inorg. Chem..
12. Cambon O., Goiffon A., Ibanez A. & Philippot E. 1992, Europ. J. Solid State and Inorg. Chem., sous presse.
13. Steinberg R.F., Roy M.K., Estes A.K., Chai B.H.T. & Morris R.C. 1984, IEEE Ultrasonics Symp., 279.
14. Stadler J., Krempl P. 1 Engel G. 1990, Proc. 4th E.F.T.F., Neuchatel, 567.
15. Krempl P., Stadler J., Wallnöfer W., Ellmeyer W. & Selic R. 1991, Proc. 5th E.F.T.F., Besançon ,143-147.
16. Hirano S., Kim P.C., Orihara H., Umeda H. & Ishibashi H. 1990, J. Mat.Sci., 25, 2800-2804.
17. Huard F. 1985, Thèse Université de Montpellier 27 Sept 1985.
18. Glowinski A. 1972, Ann.Telecom. 27, 3-4, 147-157.
19. Detaint J., Feldmann M., Henaff J.,Poignant Y. & Toudic Y. 1979, Proc. 33rd Annual Frequency Control Symp. 70-79.
20. Bailey D.S., Andle J.C.,Lee D.L., Soluch W., Vetelino J.F.&Chai B.H.T. 1983, Proc. Ultrasonics Symposium p.335,340.
21. Philippot E., Goiffon A., Jumas J.C., Avinens C., Détaint J., Schwartzel J. & Zarka A. 1988, Proc. 42nd Annual Frequency Control Symposium , 138-145.

MESURE DE COURANTS THERMOIONIQUES DANS LE QUARTZ

G. Jeandel[1], A. Warin [1], C. Poignon[1], J.P. Aubry[2], G. Morlot[1]

[1] Laboratoire Infrarouge (URA 809) Nancy, France ; [2] CEPE Argenteuil, France

RESUME

Nous avons caractérisé des échantillons de quartz par une mesure de courant thermoionique dite I.T.C. : Ionic Thermal Current. Une meilleure connaissance de la migration des impuretés, sous champ électrique appliqué, devrait permettre un contrôle plus efficace du processus d'électromigration utilisé pour la purification du quartz : le "sweeping". Plusieurs blocs provenant d'une même zone ont été prélevés dans un même cristal. Les lames sont taillées dans chacun des blocs perpendiculairement à l'axe Z ou l'axe Y. Nous avons étudié deux cristaux provenant de la Société S.I.C.N. : l'un est de qualité I (n° 272) et l'autre de qualité standard (n° 511). Nous présentons les courants I.T.C. mesurés pour ces deux types de cristaux. Nous avons comparé les courants thermoioniques de lames prélevées plus ou moins loin du germe. Nous montrons également l'effet du "sweeping" sur le courant I.T.C.

Mots clés : quartz, impuretés, "sweeping", courant thermoionique, électromigration.

1. INTRODUCTION

De nombreuses études ont déjà été réalisées sur des monocristaux de quartz (1,2) : elles montrent que le défaut majeur, pour des cristaux synthétiques, est la présence d'aluminium Al^{3+} en substitution de Si^{4+}. La différence de valence entre ces deux ions implique la présence d'un autre défaut permettant la neutralité électrique : les ions alcalins M^+ (Li^+, Na^+ et K^+), les protons et les trous (h^+) jouent ce rôle. La mobilité de ces ions de compensation de charge et l'attraction coulombienne favorisent la formation de paires Al-M^+, Al-OH^- ou Al-h^+ (2). La présence des OH^- est liée à l'existence des protons ; le défaut Al-OH^- peut en particulier être mis en évidence par des mesures infrarouges (3). L'étude de la relaxation diélectrique et de la résonance paramagnétique électronique a permis, en particulier, de déterminer des énergies d'activation très faibles (4). La topographie X permet d'observer des zones de forte déformation probablement dues à des accumulations de charges(5). L'évolution de la perte acoustique en fonction de la température donne aussi des informations sur les défauts (6). Dans ce qui suit, nous allons nous intéresser aux mesures de courants thermoioniques encore appelés I.T.C. Les premières mesures de ce type ont été réalisées par Bucci (7), cette méthode a déjà été appliquée au cas du quartz (8,9). Nous avons cherché à mettre en évidence l'effet de certains paramètres sur le courant I.T.C. afin de mieux connaître la nature des impuretés responsables des pics de courant.

2. MESURE DES COURANTS I.T.C.

2.1. Principe du procédé

Si un échantillon est placé sous un champ électrique, deux phénomènes peuvent apparaître : rotation de dipôles et déplacement de charges d'espace (10). Pour mesurer les charges de polarisation, nous utilisons deux sortes d'électrodes : les électrodes isolantes sont réalisées en plaçant un isolant électrique entre l'échantillon étudié et la métallisation ; les électrodes bloquantes sont des dépôts métalliques non diffusants, obtenus par évaporation sous vide. Les mesures I.T.C. sont de deux types : T.S.D. (Thermally Stimuled Depolarization) et T.S.P. (Thermally Stimuled Polarization) (11). La mesure T.S.D. s'effectue en trois étapes (figure 1) :

- l'échantillon est placé sous un champ électrique constant Ep à une certaine température T_i. Ce champ, de quelques centaines de volts par mm, peut déplacer l'ion de sa position d'équilibre initiale ,
- le diélectrique est alors refroidi lentement jusqu'à la température T_o qui doit être suffisamment basse pour que les temps de relaxation des mouvements, apparus avec le champ Ep, soient très grands devant le temps de mesure à T_o. La structure est donc "gelée",
- on amorce ensuite la remontée en température après avoir supprimé Ep, tout en court-circuitant l'échantillon par l'intermédiaire d'un électromètre (Keithley 642).

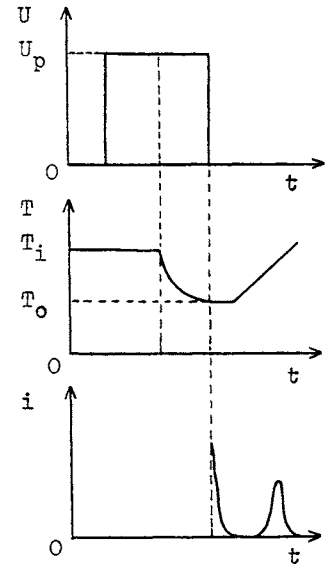


Figure 1 : Evolution de la tension appliquée U , de la température T et du courant mesuré i en fonction du temps pour une mesure T.S.D.

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

La mesure dite T.S.P. consiste à appliquer le champ électrique une fois la température To atteinte. Les mesures de courant se font donc sous champ électrique.
Nous avons préféré la première méthode (T.S.D.) car elle nous a donné des résultats plus exploitables.

2.2. Etude théorique
Comme nous l'avons indiqué précédemment, nous allons plus particulièrement nous intéresser aux relaxations dipolaires et au déplacement des charges d'espace, car ce sont ces mouvements qui sont à l'origine du courant mesuré.

Relaxation dipolaire
Nous supposerons que les relaxations ne sont pas interactives et que l'on peut considérer un seul temps de relaxation pour un phénomène donné. Le temps de relaxation τ s'écrit alors sous la forme :

$$\tau(T) = \tau_0 \exp(E/kT)$$

τ_0 est le facteur pré-exponentiel
E est l'énergie d'activation

La polarisation maximale P_0 (T) atteinte sous l'effet d'un champ statique Ep faible ($\mu Ep \ll kT$; μ est le moment dipolaire) est donnée par la loi de Langevin :

$$P_0(T) = \frac{\alpha N \mu^2 Ep}{kT}$$

N est la concentration dipolaire
α est un facteur géométrique

Pour étudier le courant I.T.C., nous supposons que seule l'orientation des dipôles crée un champ dépolarisant lorsque le milieu est soumis au champ électrique. Nous négligeons la polarisation ionique ainsi que les courants ioniques et électroniques. Le calcul de la densité de courant dépend des conditions initiales :

• si la polarisation est nulle à T_0, nous observons lors de la remontée en température une densité de courant j(t) qui vérifie la formule Debye :

$$j(t) = -\frac{dP}{dt} = \frac{P(t) - P_0(T)}{\tau(T)}$$

Si nous effectuons une remontée linéaire en température $T = bt + T_0$, nous obtenons l'expression de la densité de courant :

$$j(T) = \left[\frac{\tau^{-1}(T)}{b} \exp - \int_{T_0}^{T} \frac{dT'}{b\,\tau(T')}\right] * \int_{T_0}^{T} \frac{P_0(T')}{\tau(T')} * \left(\exp \int_{T_0}^{T} \frac{dT''}{b\tau(T'')}\right) dT' - \frac{P_0(T)}{\tau(T)}$$

• si la polarisation est égale à P_0 à T_0 :

$$j(T) = -\frac{dP}{dt} = \frac{P(t)}{\tau(T)}$$

Avec $T = bt + T_0$

$$j(T) = \frac{P_0}{\tau_0} \exp\left[-\frac{E}{kT} - \frac{1}{b\tau_0} \int_{T_0}^{T} \exp\left(-\frac{E}{kT'}\right) dT'\right]$$

Mouvement des charges d'espace
La densité de courant mesurée lors du déplacement des charges d'espace peut encore s'écrire :

$$j(t) = -\frac{dP}{dt}$$

Comme dans le cas des relaxations dipolaires, la température du maximum de courant est directement liée à l'énergie d'activation E (11).

3. RESULTATS EXPERIMENTAUX ET DISCUSSION

Pour mesurer les courants I.T.C., nous avons utilisé des électrodes bloquantes. Nous avons appliqué à chaque échantillon une tension de 1 600 V pendant 14 heures à 350 K et nous avons effectué , aprés refroidissement, une remontée linéaire en température à raison de 0,05 K/s. Tous les échantillons taillés perpendiculairement à l'axe Z ont la même épaisseur: 0,75 mm. Comme les surfaces des électrodes varient sensiblement d'un échantillon à l'autre, nous avons préféré représenter l'évolution des densités de courant en fonction de la température en considérant que l'intensité mesurée est proportionnelle à la surface de l'électrode. Les quatre figures que nous allons décrire montre l'influence d'un certain nombre de paramètres .
La figure 2 permet de comparer les densités de courant mesurées pour des échantillons du cycle 511 prélevés près ou loin du germe . Les lames ont été taillées perpendiculairement à l'axe Z . Nous constatons que les deux courbes ont la même allure dans la gamme de température explorée mais que la densité de courant est plus importante pour la lame située prés du germe . Les ions mobiles responsables de ces courants sont donc plus nombreux dans la zone localisée prés du germe . Il faut signaler que le courant mesuré à basse température est de l'ordre de 10^{-13} A.
L'enregistrement de l'intensité en fonction du temps fait apparaitre , à des températures différentes suivant les mesures, des pics de courant qui ne sont sans doute pas dûs au bruit des appareils de mesure . Il est fort probable que nous observons des effets "d'avalanche" provoqués par des états métastables .

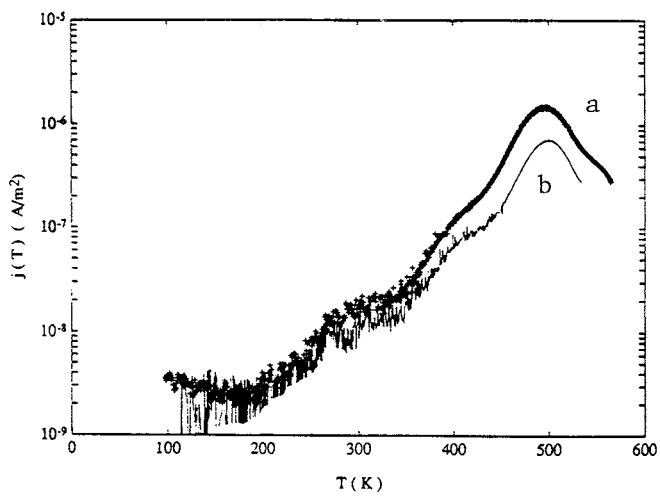


Figure 2 : Densité de courant T.S.D. pour un échantillon du cycle 511 prélevé près du germe (a) et loin du germe (b) .

La figure 3 montre l'évolution de la densité de courant pour une lame du cristal de qualité I .Nous constatons que cette courbe est trés differente de celle obtenue avec l'échantillon de qualité standard (figure 2) . Nous observons une structure beaucoup plus riche en pics (102 K - 144,5 K - 213 K -

327 K et 475 K), des inflexions sont aussi présentes aux environs de 260 K et 380 K . Jusqu'à 400 K , le courant mesuré pour l'échantillon du cycle 272 est plus important que celui provenant de la lame du cycle 511 . Cependant , la quantité de charges surfaciques recueillies jusqu'à 550 K est bien plus importante pour l'échantillon de qualité standard .

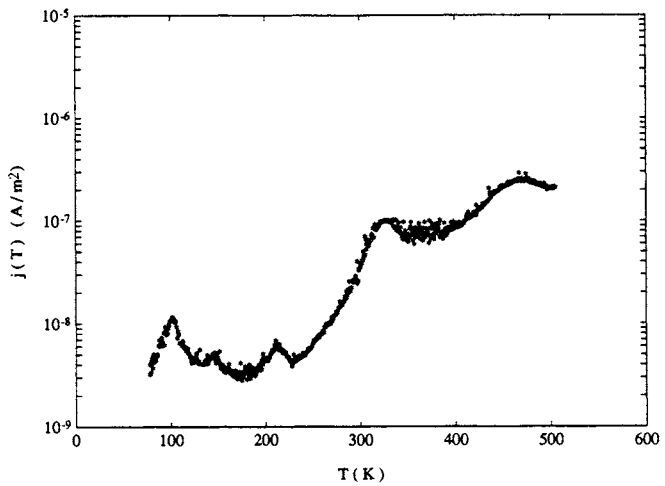


Figure 3 : Densité de courant T.S.D. pour une lame d'un cristal du cycle 272 située loin du germe.

La figure 4 presente les densités de courant de deux lames taillées perpendiculairement à Z dans un bloc ayant subi un "sweeping" dans le vide. Une des lames a été prélevée près de l'anode et l'autre près de la cathode . La différence la plus importante entre les deux courbes s'observe au delà de 400 K: le pic majeur est plus intense pour l'échantillon prélevé près de la cathode . Nous pouvons en déduire que ce pic est dû à un déplacement de cations . Le fait que nous ne puissions pas observer de grandes différences de comportement entre les deux lames aux basses températures peut s'expliquer par le choix des conditions de purification . En effet , la tension de polarisation du "sweeping" a été coupée à température ambiante : la mobilité des ions provoquant les courants de basses températures est alors suffisante pour qu'ils puissent se redéplacer avant la découpe des échantillons étudiés .

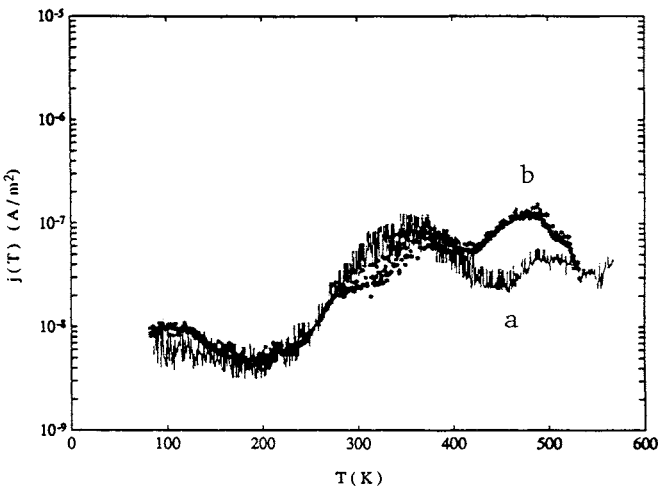


Figure 4: Densité de courant T.S.D. pour une lame d'un cristal du cycle 511 ayant subi un "sweeping" ,coté anode (a) et coté cathode (b).

Nous avons aussi mesuré la densité de courant pour une lame située près de l'anode lors d'un "sweeping" sous air (figure 5) : j(T) est plus important que celui mesuré dans les mêmes conditions pour une purification sous vide . (figure 4).

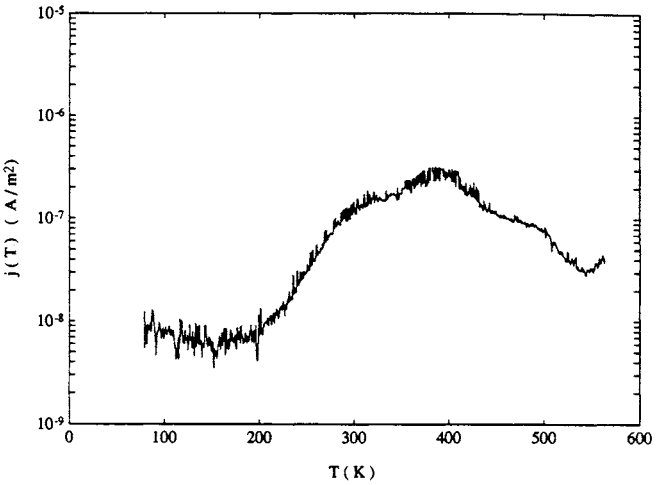


Figure 5 : Densité de courant T.S.D. pour une lame du cycle 511 ayant subi un "sweeping" sous air.

4 . CONCLUSION

Les mesures de courants thermoioniques que nous avons réalisées sur des lames de quartz montrent un comportement différent de l'échantillon suivant son origine et son histoire.
Ainsi le courant I.T.C. est plus important pour une lame prélevée près du germe que pour celle située loin du germe . La courbe de densité de courant dépend bien sûr de la qualité du cristal et nous avons aussi montré l'effet du "sweeping" sous air et sous vide . Nous cherchons actuellement à déterminer la nature des défauts qui sont à l'origine des pics observés : le calcul des énergies d'activation est un des moyens simples que nous allons utiliser .

REMERCIEMENTS : Nous remercions la société S.I.C.N qui a réalisé la synthèse des cristaux de quartz ainsi que Mrs Dolino et Bacheimer qui ont procédé au "sweeping" . Ce travail a été realisé avec le soutien financier de la DRET .

5. REFERENCES

1. KRONENBERG A.K. & KIRBY S.H. 1987, Ionic conductivity of quartz : DC time dependence and transition in charge carriers, American Mineralogist, 72, 739-747.

2. PLATA J., BRETON J. & GIRARDET C. 1988, Theoritical model for the electrodiffusion of M^+ (M = Li, Na, K) ions in a quartz crystal, Physical Review B, 38, 5, 3482-3493.

3. BACHEIMER J.P & DOLINO G .1989,Study of quartz sweeping in α and β phases,Proc. 43th Frequency Control Symposium.

4. TOULOUSE J., LING S. & NOWICK A.S. 1988, Dielectric relaxation of the aluminium hole center in α-quartz : an example of phonon-assisted tunneling, Physical Review B, 37, 12, 7070-7077.

5. DETAINT J., PHILIPPOT E., JUMAS J.C., SCHWARTZEL J., ZARKA A., CAPELLE B. & DOUKHAN J.C. 1975, Crystal growth, physical caracterization and baw devices applications of Berlinite, Proc. 39th Frequency Control Symposium, 234-246.

6. MARTIN J.J., HWANG HoB., BAHADUR H. & BERMAN G.A. 1989, Room temperature acoustic loss peaks in quartz, J. Appl. Phys., 65, 12, 4666-4671.

7. BUCCI C., FIESCHI R. & GUIDI G. 1966, Ionic thermocurrents in dielectrics, Phys. Rev., 148, 2, 816-823.

8. CALAMIOTOU M., PSICHARIS V., FILIPPAKIS S.E. & ANASTASSAKIS E. 1987, Space-charge polarisation and extinction reduction in α-quartz, J. Phys. C : Solid State Phys., 20, 5641-5653.

9. HANSON W.P. 1984, Probable ion signature in quartz electrodiffusion data, Proc. 38th Frequency Control Symposium, 38-41.

10. JEANDEL G., WARIN A. & MORLOT G. 1988, Dielectric constant and I.T.C. measurements of $B_aF_2:Dy^{3+}$, Phys. Stat. Sol. (a), 108, 185-187.

11. MULLER P. 1981, Electric relaxation currents in high-resistivity solids, Phys. Stat. Sol., 67, 1, 11-65.

THE THIRD-ORDER PERMITTIVITY OF QUARTZ OBTAINED BY DIFFERENT METHODS

Carl K. Hruska

Piezoelectricity Research Laboratory
Department of Mathematics and Statistics, York University
North York, Ontario M3J 1P3 Canada

ABSTRACT

This paper presents a comparison of three values of the third-order permittivity constant of quartz which have been obtained independently by three different methods. Each method produces a value of a different order of magnitude and there no uniformity in sign among them. The difference between the value obtained by the resonator method and the older, original value obtained by the capacitor method (Besson and Gagnepain, Ref.1) is not statistically significant. On the other hand, the value obtained by the transit-time method differs significantly from the previous two results.

Keywords: Piezoelectricity, quartz, nonlinear properties, third-order permittivity.

1. INTRODUCTION

Third-order permittivity is one of the four existing third-order electromechanical nonlinearities known in piezoelectric crystals. The material constants describing this phenomenon in this paper are the third-order nonlinear permittivity constants κ_{ijk}, i,j,k = 1,2,3. They are defined by their position in the following constitutive equation

$$D_i = \ldots + (1/2)\, \kappa_{ijk}\, E_j\, E_k + \ldots ,$$

where D_i and E_j are the components of the vectors of electric displacement and electric field intensity, respectively. The only independent nonzero third-order permittivity constant in quartz is κ_{111}.

Until recently, the only available value of κ_{111} has been the one published by Besson and Gagnepain (Ref.1) in 1972. It was obtained using the capacitor method they developed. Two indepedent least-squares estimates of quantities related to κ_{111} have been obtained recently by Hruska (Ref.2). They were calculated from experimenal data on the interactions between quartz and a dc electric field and observed using the transit-time and the resonator method.

This paper brings the three values together aiming at a comparison which would verify the original result (Ref.1) or stimulate further research into it.

The calculations done in preparation of this paper used the second-order (linear) elastic constants taken from McSkimin, Andreatch and Thurston (Ref.3), the piezoelectric and dielectric constants taken from Bechmann (Ref.4).

The numerical values of all material constants in this paper are stated for right-hand quartz, room temperature and the frame of reference according to the IEEE 1978 Standard (Ref.5). Using the notation according to Adam, Tichy and Kittinger (Ref.6), the material constants defined for the mechanical stress and electric field intensity as the independent state variables are marked with the superscript 3; those defined for the independent state variable of mechanical strain and electric field intensity are marked with the superscript 7. Differences between the isothermic and adiabatic quantities are unimportant at this stage and are ignored.

2. AVAILABLE RESULTS

The capacitor method consists in observations of frequency changes in a LC oscillator whose capacitor C contains a quartz specimen as its dielectric. The frequency changes are caused by an imposed dc potential across the capacitor which changes its capacitance. Using this method, Besson and Gagnepain (Ref.1) reported the third-order permittivity of quartz to be

$$(1) \qquad {}^3\kappa^C_{111} = (-2.4 \pm 0.4) \cdot 10^{-22} \text{ F/V},$$

where the error is understood to be the standard error. The sign at ${}^3\kappa^C_{111}$ in (Ref.1) has been adjusted to agree with the current definiton of the basic frame of reference in quartz (Ref.5). The superscript C indicates that the value has been obtained by the capacitor method.

The resonator method consists in observations of the series resonance frequency of piezoelectric resonators which changes when the resonators are exposed to a dc field bias. The method was introduced by Hruska (Ref.7).

The resonator method yields a least-squares estimate of nonlinear parameter p_1 containing third-order permittivity ${}^7\kappa_{111}$

$$(2) \quad p_1 = {}^7\kappa_{111} - 2.465 \cdot 10^{-35}\, {}^7c_{111} + 1.233 \cdot 10^{-35}\, {}^7c_{112}$$

$$+ 0.776 \cdot 10^{-35}\ {}^7c_{114} - 2.329 \cdot 10^{-35}\ {}^7c_{124}$$
$$- 0.367 \cdot 10^{-35}\ {}^7c_{144} + 0.244 \cdot 10^{-35}\ {}^7c_{155}$$
$$+ 1.234 \cdot 10^{-35}\ {}^7c_{222} + 0.038 \cdot 10^{-35}\ {}^7c_{444}$$
$$- 2.045 \cdot 10^{-23}\ ({}^7l_{12}^{rel} + 1)\ .$$

In Eq.(2) the third-order permittivity ${}^7\kappa_{111}$ is combined with several other third-order material constants of quartz. They are the third-order elastic constants ${}^7c_{ijk}$ and relative electrostrictive constant ${}^7l_{12}^{rel}$. The coefficients at ${}^7c_{ijk}$ are in Fm^2/NV and the numerical coefficient at $({}^7l_{12}^{rel} + 1)$ is in F/V. The value of parameter p_1 was found to be (Ref.2)

$$(3) \qquad p_1^R = (-29 \pm 24) \cdot 10^{-22}\ F/V\ ,$$

where the error is the standard error. The superscript R indicates that the value has been obtained using the resonator method.

The transit-time method consists in observations of the travel time of acoustic pulses through bulk piezoelectric material. Changes in the travel time are measured under the effect of a dc biasing field. The first study of the third-order nonlinearities in quartz introducing this method has been made by Reider, Kittinger and Tichy (Ref.8). More recent observations obtained by Kittinger, Tichy and Friedel (Ref.9) and used by Hruska resulted in a least-squares estimate of nonlinear parameter p_2 which includes third-order permittivity ${}^7\kappa_{111}$

$$(4)\ p_2 = {}^7\kappa_{111}$$
$$- 4.931 \cdot 10^{-35}\ {}^7c_{111} - 4.655 \cdot 10^{-35}\ {}^7c_{124}$$
$$- 0.366 \cdot 10^{-35}\ {}^7c_{144} + 0.366 \cdot 10^{-35}\ {}^7c_{155}$$
$$+ 4.931 \cdot 10^{-35}\ {}^7c_{222} + 0.038 \cdot 10^{-35}\ {}^7c_{444}.$$

The same explanatory remarks apply to Eq.(4) as made for Eq.(2). The least-squares estimate of parameter p_2 was found to be (Ref.2)

$$(5) \qquad p_2^T = (371 \pm 239) \cdot 10^{-22}\ F/V\ ,$$

where the error is the standard error. The superscript T indicates that the value has been obtained using the transit-time method.

3. COMPARISON OF RESULTS

Adam, Tichy and Kittinger (Ref.6) have derived a general relationship between the third-order permittivity constants ${}^3\kappa_{ijk}$ and ${}^7\kappa_{ijk}$. In application to quartz one can write

$$(6)\ {}^3\kappa_{111} = {}^7\kappa_{111}$$
$$- 4.931 \cdot 10^{-35}\ {}^7c_{111} - 4.655 \cdot 10^{-35}\ {}^7c_{124}$$
$$- 0.366 \cdot 10^{-35}\ {}^7c_{144} + 0.366 \cdot 10^{-35}\ {}^7c_{155}$$
$$+ 4.931 \cdot 10^{-35}\ {}^7c_{222} + 0.038 \cdot 10^{-35}\ {}^7c_{444}$$
$$+ 6.136 \cdot 10^{-23}\ {}^7l_{11}^{rel} - 6.136 \cdot 10^{-23}\ {}^7l_{12}^{rel}$$
$$- 1.931 \cdot 10^{-23}\ {}^7l_{14}^{rel}$$
$$- 1.601 \cdot 10^{-23}\ {}^7f_{111} + 1.008 \cdot 10^{-23}\ {}^7f_{114}$$
$$- 1.601 \cdot 10^{-23}\ {}^7f_{122} - 1.008 \cdot 10^{-23}\ {}^7f_{124}$$
$$- 0.159 \cdot 10^{-23}\ {}^7f_{144}\ ,$$

where ${}^7f_{ijk}$ are the electroelastic constants of quartz. The numerical coefficients at ${}^7f_{ijk}$ are in F.m/N. Additional explanatory remarks needed here are the same as made for Eq.(2).

Combining Eqs.(6) and (2) one can write

$$(7)\ {}^3\kappa_{111} = p_1$$
$$- 2.466 \cdot 10^{-35}\ {}^7c_{111} - 1.233 \cdot 10^{-35}\ {}^7c_{112}$$
$$- 0.776 \cdot 10^{-35}\ {}^7c_{114} - 2.328 \cdot 10^{-35}\ {}^7c_{124}$$
$$- 0.122 \cdot 10^{-35}\ {}^7c_{155} + 3.698 \cdot 10^{-35}\ {}^7c_{222}$$
$$+ 6.136 \cdot 10^{-23}\ {}^7l_{11}^{rel} - 4.091 \cdot 10^{-23}\ {}^7l_{12}^{rel}$$
$$- 1.931 \cdot 10^{-23}\ {}^7l_{14}^{rel}$$
$$- 1.601 \cdot 10^{-23}\ {}^7f_{111} + 1.008 \cdot 10^{-23}\ {}^7f_{114}$$
$$- 1.601 \cdot 10^{-23}\ {}^7f_{122} - 1.008 \cdot 10^{-23}\ {}^7f_{124}$$
$$- 0.159 \cdot 10^{-23}\ {}^7f_{144}$$
$$+ 2.045 \cdot 10^{-23}\ .$$

Using the value of p_1 in Eq.(3), Eq.(7) can be related to the result obtained by the resonator method. Given that for quartz the order of magnitude of third-order elastic constants ${}^7c_{ijk}$, relative electrostrictive constants ${}^7l_{ij}^{rel}$ and electroelastic constants ${}^7f_{ijk}$ is generally known not to exceed 10^{11} N/m^2, 10, and 10^0 $N/(V.m)$, respectively, it is possible to conclude that

$$(8) \qquad {}^3\kappa_{111}^R \cong p_1^R\ .$$

Combining Eqs.(6) and (4) one can write

$$(9)\ {}^3\kappa_{111} = p_2$$
$$+ 6.136 \cdot 10^{-23}\ {}^7l_{11}^{rel} - 6.136 \cdot 10^{-23}\ {}^7l_{12}^{rel}$$
$$- 1.931 \cdot 10^{-23}\ {}^7l_{14}^{rel}$$
$$- 1.601 \cdot 10^{-23}\ {}^7f_{111} + 1.008 \cdot 10^{-23}\ {}^7f_{114}$$
$$- 1.601 \cdot 10^{-23}\ {}^7f_{122} - 1.008 \cdot 10^{-23}\ {}^7f_{124}$$
$$- 0.159 \cdot 10^{-23}\ {}^7f_{144}$$

Taking into account the aforementioned order of magnitude of the third-order nonlinear constants of quartz and using Eqs.(5) and (9) it is possible to conclude that

$$(10) \qquad {}^3\kappa_{111}^T \cong p_2^T\ .$$

Eqs.(1),(8) and (10) provide three independent estimates of the third-order permittivity ${}^3\kappa_{111}$. Their comparison is made in Table I.

TABLE I. Values of the third-order permittivity ${}^3\kappa_{111}$ of quartz obtained by different methods.

${}^3\kappa_{111}^C$	$= (-2.4 \pm 0.4) \cdot 10^{-22}$ F/V
${}^3\kappa_{111}^R$	$\cong (-29 \pm 24) \cdot 10^{-22}$ F/V
${}^3\kappa_{111}^T$	$\cong (371 \pm 239) \cdot 10^{-22}$ F/V

The superscripts C, R, and T at κ_{111} indicate that the corresponding values have been obtained by the capacitor, resonance and transit-time method, respectively. The errors are standard errors.

In spite of the disagreement of one order of magnitude, the 95% confidence interval for the difference $^{3}\kappa^{C}_{111} - {}^{3}\kappa^{R}_{111}$ contains zero. Consequently, in this sense, the difference between $^{3}\kappa^{C}_{111}$ and $^{3}\kappa^{R}_{111}$ cannot be regarded as significant. Using the same criterion, the difference between the values of $^{3}\kappa^{C}_{111}$ or $^{3}\kappa^{R}_{111}$ and $^{3}\kappa^{T}_{111}$ must be regarded as statistically significant.

4. CONCLUSION

The original value of the third-order permittivity of quartz defined at constant stress has been published by Besson and Gagnepain two decades ago. This paper presents two new estimates of the same quantity obtained independently and by different experimental methods. The aim of their acquisition was to verify the original value.

The differences among the three values of the third-order permittivity constant are the largest discrepancies currently known in the entire family of the 31 third-order nonlinear electroelastic constants of quartz. They span two orders of magnitude; the three values also fail to agree in their sign. As such, the verification attempt has not produced the desired result. Further investigation into the matter will be necessary.

The current estimates of the third-order permittivity provided by the resonator and the transit-time method have excessively large standard errors. This is understandable as these estimates are a byproduct of experiments whose aim was other than the determination of the third-order permittivity of quartz. The main goal here must be to reduce the standard errors. The algorithm used to compute the standard errors suggests that this may be possible using quartz specimens suitably designed for the purpose.

The capacitor method which has produced the original value of the third-order permittivity of quartz should be used on a broader scale i.e. for a larger number of quartz specimens of different crystalographic orientation. The aim here should be to provide enough data for a full scale regression analysis uniformly covering all third-order phenomena involved in the process and exploiting the full potential of this method.

5. ACKNOWLEDGEMENTS

The author wishes to acknowledge the assistance of Pauline Hruska with the final stages of preparation of this paper. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada and by the Faculty of Arts, York University.

6. REFERENCES

1. Besson R. and Gagnepain J.-J., C.R.Acad.Sci. Paris B274(1972), pp.835-838.

2.Hruska C. K., Proceedings of the 5th European Frequency and Time Forum, Communautes Scientifiques de Besancon et de Neuchatel, Besancon, March 1991, pp. 137-142. Available from E.N.S.S.M., Universite de Besancon, France.

3. McSkimin H. J., Andreatch P. Jr, Thurston R. N., J.Appl.Phys. 36(1965), pp.1624-1632.

4. Bechmann R., Phys.Rev. 110(1958), pp. 1060-1061.

5. IEEE Standard on Piezoelectricity, IEEE Std 176-1978. Available from The Institute of Electrical and Electronics Engineers, Inc., 345 East 47th Street, New York, NY 10017.

6. Adam W., .Tichy J., and Kittinger E., J.Appl.Phys. 64(1988), pp. 2556-2562.

7. Hruska K., Czech.J.Phys. 12(1962), pp.338-353.

8. Reider, Kittinger E., Tichy J., J.Appl.Phys. 53(1982), pp.8716-8721.

9. Kittinger E., Tichy J., and Friedel W., J.Appl.Phys. 60(1986), pp.1465-1471.

THERMOELASTICITY IN QUARTZ

J. A. Kosinski, J. G. Gualtieri, and A. Ballato

US Army Electronics Technology and Devices Laboratory,
Fort Monmouth, NJ 07703-5601, USA

ABSTRACT

Experimental measurements of alpha quartz thermal expansion reported in the literature have been critically analyzed. A recommended set of best measured values over the temperature range -50° C to +150° C have been determined, as have values for the coefficient of thermal linear expansion (CTE) and the thermoelastic coefficients $\alpha_{ij}^{(n)}$. The new $\alpha_{ij}^{(n)}$ values are:

$\alpha_{11}^{(1)} = 13.65 \times 10^{-6}$ /K $\alpha_{33}^{(1)} = 7.50 \times 10^{-6}$ /K

$\alpha_{11}^{(2)} = 11.02 \times 10^{-9}$ /K^2 $\alpha_{33}^{(2)} = 8.00 \times 10^{-9}$ /K^2

$\alpha_{11}^{(3)} = -19.32 \times 10^{-12}$/K^3 $\alpha_{33}^{(3)} = -10.44 \times 10^{-12}$/K^3

The impact of using the new $\alpha_{ij}^{(n)}$ on determinations of quartz material temperature coefficients and on the calculation of temperature coefficients of frequency for the case of the AT-cut are discussed.

1. INTRODUCTION

There is an internationally recognized interest in obtaining more reliable alpha-quartz material constants [1,2]. This interest also is recognized, e.g., by ongoing projects coordinated by the Electronic Industries Association (EIA) [3], which have highlighted the need to examine all of the quartz material constants.

The thermoelastic coefficients $\alpha_{ij}^{(n)}$, also known as the thermal expansion coefficients, are critical parameters in piezoelectric crystal resonator and filter calculations. For example, the first-order temperature coefficient of the antiresonance frequency of any simple thickness mode of a crystal plate with zero mass loading can be shown to be of the form [4]:

$$2T_f^{(1)} = T_{\bar{c}}^{(1)} - T_\rho^{(1)} - 2T_h^{(1)} \qquad (1)$$

where f = frequency, $\bar{c}$ = piezoelectrically stiffened elastic constant, ρ = mass density, and h = thickness.

The quantities $T_\rho^{(1)}$ and $T_h^{(1)}$ depend directly on thermoelastic coefficients and are given by:

$$T_\rho^{(1)} = -(\alpha_{11}^{(1)} + \alpha_{22}^{(1)} + \alpha_{33}^{(1)}) \qquad (2)$$

and

$$T_h^{(1)} = +\alpha_{22}^{/(1)} \qquad (3)$$

where the $\alpha_{ij}^{(1)}$ are thermoelastic coefficients and $\alpha_{22}^{/(1)}$ is the thermoelastic coefficient rotated to the crystallographic direction of the plate thickness, which is taken as the x_2-direction. The $\alpha_{ij}^{(n)}$ are components of tensors of the second rank.

At present there are three sets of thermoelastic coefficients in general use: the Bechmann, Ballato, and Lukaszek (BBL) set [4]; the American Institute of Physics (AIP) set [5] ; and the Mizan and Ballato (MB) set [6]. The MB set is based on a digitization of a graph published by Sosman [7]. Brice [8] and James [9] have also contributed alternative sets of alpha quartz thermoelastic coefficients.

2. EXPERIMENTAL DATA

Funding and time constraints did not allow the gathering of new experimental data. Instead, a literature search and critical analysis of published data were performed. This phase was accomplished by the Purdue University - Center for Information and Numerical Data Analysis and Synthesis (CINDAS) under a program funded by ETDL. CINDAS operates the High Temperature Materials - Mechanical, Electronic and Thermophysical Properties Information Analysis Center (HTMIAC) for the Department of Defense [10].

The HTMIAC analysis involved searching, compiling, and critically analyzing experimental measurements of alpha quartz thermal expansion data as reported in the literature. The HTMIAC analysis yielded a recommended set of best measured values of the coefficient of thermal linear* expansion (CTE) over the temperature range of -50° C to +150° C, as listed in Table 1. The recommended values along the a-axis were obtained by fitting the data of White [12] and Barron et al. [13], in conjunction with the data of Jay

* The term "linear" is of historical usage, but may lead to confusion, especially inasmuch as our curves and formulas describe nonlinear variations. According to Touloukian et al. [11], it denotes unidirectional dimensional change, as opposed to volumetric change. We retain the traditional nomenclature, while recognizing that a better descriptor might be "lineal".

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17-19 March 1992 (ESA SP-340, June 1992).

Table 1: The HTMIAC recommended set of best measured values of the coefficient of thermal linear expansion (CTE) over the temperature range of −50° C to +150° C.

Temperature (° C)	a-axis CTE (X 10^{-6}/K)	c-axis CTE (X 10^{-6}/K)
−50	11.69	6.07
−40	11.97	6.31
−30	12.25	6.53
−20	12.53	6.73
−10	12.80	6.92
0	13.06	7.10
10	13.31	7.27
20	13.55	7.43
30	13.79	7.59
40	13.99	7.73
50	14.19	7.87
60	14.37	8.01
70	14.53	8.14
80	14.69	8.26
90	14.83	8.38
100	14.96	8.50
110	15.08	8.61
120	15.20	8.72
130	15.31	8.83
140	15.43	8.94
150	15.55	9.04

[14], Johnson and Parsons [15], and Lager et al. [16]. Data of Dorsey [17] and Buffington and Latimer [18] also were included. The recommended values for CTE along the c-axis were obtained from Refs. [12–16] and also included the data of Nix and McNair [19]. The experimental samples in all these studies were natural quartz. As shown in Fig. 1, none of the thermoelastic coefficient sets in general use agree well with the HTMIAC recommended data. In addition to providing a recommended set of best measured values, HTMIAC provided 3rd–, 4th–, and 5th-order power series of temperature for CTE referenced to 0° C and valid for the temperature range −50° C to +150° C. The CTE curve-fit coefficients are listed in Table 2. See Fig.2 for a plot of the fit to a 5th-order power series for CTE.

3. DATA ANALYSIS

The coefficient of thermal linear expansion is defined as [11]

$$\text{CTE} \equiv 1/h(\Theta_o)\ \partial h/\partial\Theta \qquad (4)$$

with reference temperature Θ_o = 0° C. Unfortunately, the CTE as used by HTMIAC are not equivalent to the thermoelastic coefficients, $\alpha_{ij}^{(n)}$, that are normally used by the frequency control community. The more usual thermoelastic coefficients, $\alpha_{ij}^{(n)}$, are obtained from a Taylor-series expansion of length or thickness (h) about a reference temperature Θ_o = 25° C :

$$h(\Theta)=h(\Theta_o)\ \sum_{n=o}^{m}\ 1/n!\ 1/h(\Theta_o)\left[\ \partial^n h/\partial\Theta^n\ \right]_{\Theta_o}\ (\Theta-\Theta_o)^n \qquad (5)$$

whence

$$\alpha_{ij}^{(n)} \equiv 1/n!\ 1/h(\Theta_o)\ [\partial^n h/\partial\Theta^n]_{\Theta_o} \ . \qquad (6)$$

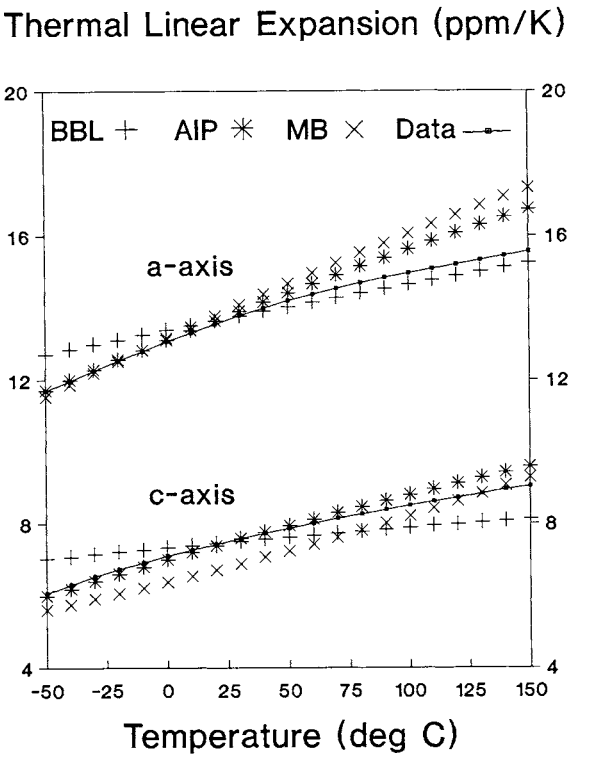


Fig. 1: Thermal linear expansion calculated using the thermoelastic coefficients of BBL [4], AIP [5], and MB [6]. There is little agreement between the HTMIAC recommended best measured values and any of the thermoelastic coefficient sets for the temperature range −50° C to 150° C.

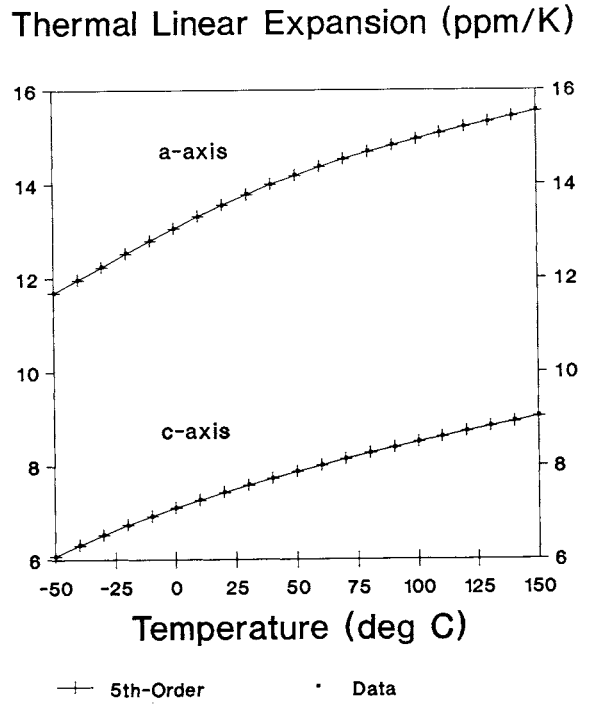


Fig. 2: Least-squares fit to the HTMIAC recommended best measured values using the CTE formalism of equations (4) and (7). The curves shown represent a 5th-order power series expansion.

Table 2: The coefficient of thermal linear expansion (CTE) curve-fit values.

	a-axis			
	Order of Fit			
	5	4	3	Units
Coeff.				
a_0	13.1	13.1	13.1	X 10^{-6}/K
a_1	26.1	25.8	25.0	X 10^{-9}/K^2
a_2	-60.6	-54.3	-64.7	X 10^{-12}/K^3
a_3	-466.2	-320.3	45.1	X 10^{-15}/K^4
a_4	4645.	1827.		X 10^{-18}/K^5
a_5	-11270.			X 10^{-21}/K^6
	c-axis			
a_0	7.1	7.1	7.1	X 10^{-6}/K
a_1	17.2	17.5	17.7	X 10^{-9}/K^2
a_2	-45.5	-50.9	-48.4	X 10^{-12}/K^3
a_3	323.0	-200.0	111.7	X 10^{-15}/K^4
a_4	-2817.	-441.6		X 10^{-18}/K^5
a_5	9503.			X 10^{-21}/K^6

In the past, most analyses have been limited to order m = 3.

Mathematical relationships between the CTE and the $\alpha_{ij}^{(n)}$ were derived, taking into account the difference in reference temperatures. The HTMIAC power series expansion of (4) can be written as

$$\text{CTE} = 1/h(0)\ \partial h/\partial\Theta = \sum_{n=0}^{m} a_n\Theta^n. \qquad (7)$$

We can express h(0), the thickness at temperature Θ = 0° C, in terms of the $\alpha_{ij}^{(n)}$ and the thickness at temperature Θ_o = 25° C as

$$h(0) = h(\Theta_o)\ \left[\ \sum_{n=0}^{m} \alpha_{ij}^{(n)}\,(-\Theta_o)^n\ \right]. \qquad (8)$$

In practice, the thickness at 0° C is very nearly the same as the thickness at 25° C, and the summation in (8) is approximately 1. Differentiating equation (5) and substituting equation (6), we find

$$\partial h/\partial\Theta = h(\Theta_o)\left[\ \sum_{n=0}^{m} n\alpha_{ij}^{(n)}(\Theta-\Theta_o)^{n-1}\ \right] \qquad (9)$$

thus we can write the relationship between CTE and $\alpha_{ij}^{(n)}$ as

$$\text{CTE} \approx \sum_{n=0}^{m} n\alpha_{ij}^{(n)}(\Theta-\Theta_o)^{n-1}. \qquad (10)$$

Equations (7) and (10) can be expanded and coefficients of equal powers of Θ equated, yielding expressions for the thermoelastic coefficients up to order six in terms of the CTE power series coefficients:

$$\begin{aligned}
\alpha_{ij}^{(1)} &= a_o + a_1\Theta_o + a_2\Theta_o^2 + a_3\Theta_o^3 + a_4\Theta_o^4 + a_5\Theta_o^5 \\
\alpha_{ij}^{(2)} &= 1/2[a_1 + 2a_2\Theta_o + 3a_3\Theta_o^2 + 4a_4\Theta_o^3 + 5a_5\Theta_o^4] \\
\alpha_{ij}^{(3)} &= 1/3[a_2 + 3a_3\Theta_o + 6a_4\Theta_o^2 + 10a_5\Theta_o^3] \\
\alpha_{ij}^{(4)} &= 1/4[a_3 + 4a_4\Theta_o + 10a_5\Theta_o^2] \\
\alpha_{ij}^{(5)} &= 1/5[a_4 + 5a_5\Theta_o] \\
\alpha_{ij}^{(6)} &= 1/6[a_5].
\end{aligned} \qquad (11)$$

Note that an analysis of CTE to order m corresponds to an analysis of $\alpha_{ij}^{(n)}$ to order m+1.

The HTMIAC results recast in thermoelastic form are listed in Table 3. Most calculations of quartz resonator temperature behavior are carried out only to third order. If any one of the HTMIAC analyses is truncated to third order, the calculated results do not reproduce the HTMIAC recommended data (see, e.g., Fig. 3). Because of this, a least-squares fit to the HTMIAC recommended data to third order in α_{ij} was developed by the authors. This fit is called KGB and is plotted in Fig. 4. The KGB thermoelastic coefficients are listed together with those of BBL, AIP, and MB in Table 4.

4. DISCUSSION

To understand the impact of using the KGB $\alpha_{ij}^{(n)}$ instead of, for example, using the BBL α_{ij}, we have examined the determination of material-constant temperature coefficients (obtained by the resonator method) as well as the calculation of frequency-temperature behavior from the material constants.

The use of KGB thermoelastic coefficients instead of BBL coefficients will result in nontrivial differences when extracting temperature coefficients of the elastic, piezoelectric, and dielectric constants from measured resonator frequency-temperature behavior. As an example, values of the temperature coefficients of selected elastic constants extracted from measurements on rotated Y-cut resonators are listed in Table 5. The previously reported [3] values obtained using BBL thermoelastic coefficients may be compared to the values obtained using KGB thermoelastic coefficients. The coefficients thus obtained are seen to

Table 3: The HTMIAC results recast in thermoelastic form.

$\alpha_{11}^{(n)}$

	Order of Analysis			
	6	5	4	Units
(n)				
1	13.673	13.668	13.659	X 10^{-6}/K
2	11.231	11.299	10.926	X 10^{-9}/K^2
3	-26.646	-23.834	-20.422	X 10^{-12}/K^3
4	-18.034	-34.400	11.278	X 10^{-15}/K^4
5	647.250	365.400		X 10^{-18}/K^5
6	-1878.333			X 10^{-21}/K^6

$\alpha_{33}^{(n)}$

1	7.508	7.513	7.515	X 10^{-6}/K
2	7.686	7.653	7.746	X 10^{-9}/K^2
3	-10.131	-12.502	-13.328	X 10^{-12}/K^3
4	25.173	38.960	27.925	X 10^{-15}/K^4
5	-325.825	-88.320		X 10^{-18}/K^5
6	1583.833			X 10^{-21}/K^6

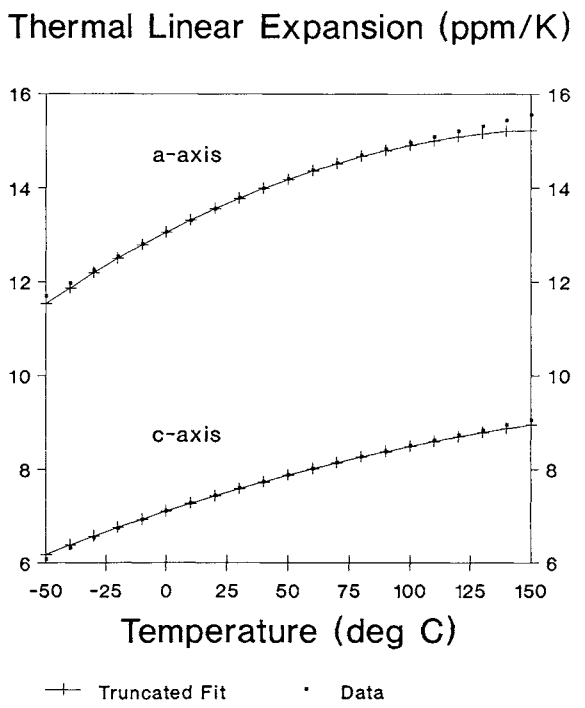


Fig. 3: Calculated thermal linear expansion using the 5th-order CTE analysis of Fig. 2 converted to a 6th-order thermoelastic analysis and subsequently truncated to 3rd-order compared to the HTMIAC recommended best measured values.

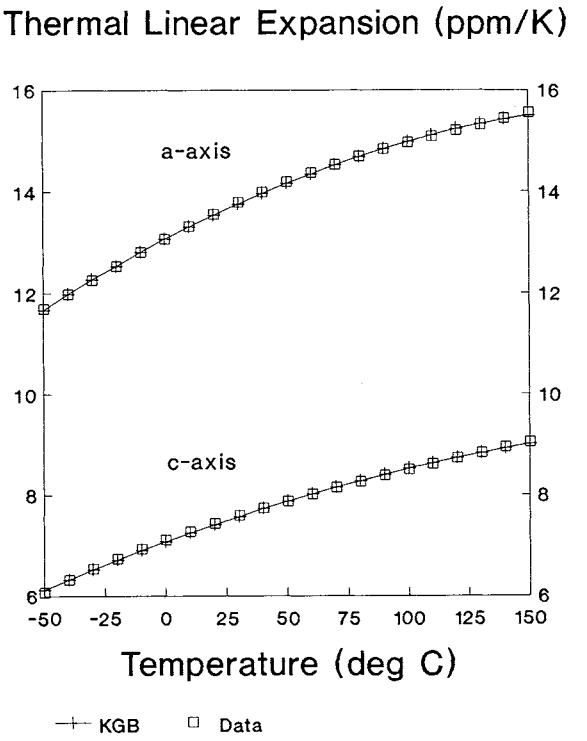


Fig. 4: Third-order thermoelastic coefficient least-squares fit (KGB) to the HTMIAC recommended best measured values.

Table 4: The KGB thermoelastic coefficients are listed together with those of BBL, AIP, and MB.

$\alpha_{11}^{(n)}$

(n)	BBL	AIP	MB	KGB	Units
1	13.71	13.77	13.92	13.65	X 10^{-6}/K
2	6.50	13.03	15.09	11.02	X 10^{-9}/K^2
3	-1.90	-6.33	-7.86	-19.32	X 10^{-12}/K^3

$\alpha_{33}^{(n)}$

1	7.48	7.48	6.79	7.50	X 10^{-6}/K
2	2.90	9.41	8.69	8.00	X 10^{-9}/K^2
3	-1.50	-5.44	6.88	-10.44	X 10^{-12}/K^3

differ from the previously reported values by up to 20%.

The influence of the $\alpha_{ij}^{(n)}$ on the calculation of frequency-temperature behavior has been examined by using (1) in conjunction with the BBL set of material constants (with the exception of the BBL $\alpha_{ij}^{(n)}$). For this purpose, it is best to rewrite (1) in the form [20]:

$$2\hat{T}_f^{(n)} = \hat{T}_c^{(n)} - \hat{T}_p^{(n)} - 2\hat{T}_h^{(n)} \qquad (12)$$

Table 5: Values of the temperature coefficients of selected elastic constants extracted from measurements on rotated Y-cut resonators.

$(T^{(n)}_{c_{ij}E})$

BBL

ij --> (n)	66	44	14	Units
1	176.9	−170.4	103.9	X 10^{-6}/K
2	163.8	−252.5	−19.4	X 10^{-9}/K^2
3	43.9	−201.1	−509.1	X 10^{-12}/K^3

KGB

1	176.9	−170.3	103.9	X 10^{-6}/K
2	158.7	−256.6	−23.9	X 10^{-9}/K^2
3	51.6	−177.3	−492.8	X 10^{-12}/K^3

where

$$\hat{T}^{(1)} = T^{(1)}$$

$$\hat{T}^{(2)} = T^{(2)} - 1/2(T^{(1)})^2 \qquad (13)$$

$$\hat{T}^{(3)} = T^{(3)} - T^{(2)}\,T^{(1)} + 1/3(T^{(1)})^3.$$

If we then assume $T^{(n)}_{\tilde{c}B} = T^{(n)}_{\tilde{c}K}$, where B denotes BBL and K denotes KGB, then

$$[\hat{T}^{(n)}_{fB} - \hat{T}^{(n)}_{fK}] = 1/2\ \{\ [\hat{T}^{(n)}_{\rho K} - \hat{T}^{(n)}_{\rho B}] + 2\,[\hat{T}^{(n)}_{hK} - \hat{T}^{(n)}_{hB}]\ \}. \qquad (14)$$

The $\alpha^{(n)}_{ij}$ values found in Table 4 together with equation (2) allow the calculation of temperature coefficients of density. Similarly, the temperature coefficients of thickness for rotated Y-cut resonators may be calculated using the equation

$$T^{(n)}_h = \alpha^{(n)}_{11}\cos^2\theta + \alpha^{(n)}_{33}\sin^2\theta\ , \qquad (15)$$

where θ is the plate rotation angle.

The differences in the temperature coefficients of frequency, density, and thickness for the AT-cut obtained from (14) are listed in Table 6. A number of observations may be made. First, the differences in the first-order temperature coefficients of frequency obtained using either set of thermoelastic coefficients are very small. For example, from Bechmann [21] we know that $\partial T^{(1)}_f/\partial\theta = -5.15$ X 10^{-6}/K°θ. Thus, the difference between the predicted AT-cut frequency-temperature behaviors is equivalent to an angular error of −14" of arc. Second, the differences in higher order temperature coefficients of frequency are not negligible. Compare, for instance, the values of the differences in temperature coefficients of frequency with the actual magnitudes of these coefficients as published by Ballato [22]:

$T^{(2)}_f = -0.45$ X 10^{-9}/K^2 and $T^{(3)}_f = 108.6$ X 10^{-12}/K^3. Such uncertainties underscore the need for new experimental measurements on currently available natural and cultured quartz materials.

Table 6: The differences in the temperature coefficients of frequency, density, and thickness for the AT-cut obtained from equation (14).

(n)	$[\hat{T}^{(n)}_{fB} - \hat{T}^{(n)}_{fK}]$	$[\hat{T}^{(n)}_{\rho K} - \hat{T}^{(n)}_{\rho B}]$	$[\hat{T}^{(n)}_{hK} - \hat{T}^{(n)}_{hB}]$	Units
1	0.02	0.10	−0.03	X 10^{-6}/K
2	−2.36	−14.14	4.71	X 10^{-9}/K^2
3	7.00	43.29	−14.65	X 10^{-12}/K^3

5. CONCLUSIONS

A critical analysis of experimental measurements of alpha quartz thermal expansion as reported in the literature has been performed, yielding an improved, average, set of higher-order thermoelastic coefficients $\alpha^{(n)}_{11}$ and $\alpha^{(n)}_{33}$. The new values are required for improved determinations of the temperature coefficients of the elastic, piezoelectric, and dielectric constants of this important crystalline material.

6. REFERENCES

[1] James B J 1988, A new measurement of the basic elastic and dielectric constants of quartz, Proc 42nd Annual Frequency Control Symposium (AFCS), 146–154, . Available from NTIS as document AD-A217275.

[2] Jin-Feng Wang, Chen Chen, Lei Zhang, and Zi-Kai Qin 1989, Determination of the dielectric, piezoelectric, and elastic constants of crystals in class 32, Phys. Rev. B, 39, 17, 12888–12890.

[3] Kosinski J, Ballato A, Lukaszek T, Mizan M, McGowan R, and Klohn K 1988, Temperature derivatives of the dynamic permittivity and permeability of the simple thickness modes of quartz plates, Proc. 42nd AFCS, 53–64.

[4] Bechmann R, Ballato A, and Lukaszek T Aug 1962, Higher-order temperature coefficients of the elastic stiffnesses and compliances of alpha quartz, Proc. IRE, 50, 8, 1812–1822.

[5] Kirby R K, Hahn T A, and Rothrock B D 1972, American Institute of Physics Handbook, ed. D. E. Gray, McGraw-Hill Book Co., New York, Ch 4, 119-140.

[6] Mizan M and Ballato A Sep 1984, Stress-frequency coefficients of doubly rotated, low temperature coefficient, quartz resonators, USAERADCOM Technical Report DELET-TR-84-10, US Army Electronics Technology and Devices Laboratory, Fort Monmouth, NJ, 76 p.

[7] Sosman R 1927, The Properties of Silica, Chemical Catalog Co., New York, 384 p.

[8] Brice J C 1985, Crystals for quartz resonators, Rev Mod Phys, 57, 1, 105-146.

[9] James B J Spring 1987, Determination of elastic and dielectric properties of quartz, Ph.D. Thesis, Royal Holloway and Bedford New College, Univ. of London, 57 p.

[10] Desai P D Jul 1990, Critical analysis of alpha-quartz thermal expansion data, HTMIAC SPECIAL REPORT 40, Defense Electronics Supply Center Contract No. DLA900-86-C-0751, 5 pp. HTMIAC is operated by CINDAS, Purdue University, 2595 Yeager Rd. West Lafayette, IN 47906.

[11] Touloukian Y S, Kirby R K, Taylor R E, and Lee T Y R, eds. 1977, Thermophysical Properties of Matter Volume 13 Thermal Expansion - Nonmetallic Solids, IFI/Plenum, New York, 34a p.

[12] White G K 1964, Thermal expansion of silica at low temperatures, Cryogenics, Vol. 4, 2-7.

[13] Barron T H K, Collins J F, Smith T W, and White G K 1982, Thermal expansion, Grüneisen functions and static lattice properties of quartz, J Phys C: Solid State Phys, 15, 4311-4326.

[14] Jay A H 1933, The thermal expansion of quartz by x-ray measurements, Proc Roy Soc (London), 142A, 237-247.

[15] Johnson W H and Parsons W H 1944, Thermal expansion of concrete aggregate materials, J Res Natl Bur Stds, 32, 101-126.

[16] Lager G A, Jorgenson J D, and Rotella F J 1982, Crystal structure and thermal expansion of α-SiO_2 at low temperatures, AIP Conf Proc, 89, 75-77. (ISSN:0094-243X/82/890075-03)

[17] Dorsey H G 1908, Further measurements of the coefficient of linear expansion at low temperatures, Phys Rev, 27, 1, 1-10.

[18] Buffington R M and Latimer W M 1926, The measurement of coefficients of expansion at low temperatures [with] some thermodynamic applications of expansion data, J Am Chem Soc, 48, 2305-2319.

[19] Nix F C and MacNair D 1941, An interferometric-dilatometer with photographic recording, Rev Sci Instr, 12, 66-70.

[20] Ballato A and Lukaszek T 1975, Higher-order temperature coefficients of frequency of mass-loaded piezoelectric crystal plates, Proc 29th AFCS, 10-25. Available from NTIS as document No. AD-A017466.

[21] Bechmann R Aug 1960, Frequency-temperature-angle characteristics of AT- and BT-type quartz oscillators in extended temperature range, Proc IRE, 48, 8, 1494 p.

[22] Ballato A 1977, Doubly rotated thickness mode plate vibrators, in Physical Acoustics, Mason W P and Thurston R N, eds, Academic Press, New York, 13, Chapter 5, 115-181.

Session 8B

Atomic Frequency Standards: Masers and Trapped ION Frequency Sources

Chairman: C. Salomon

THE DEVELOPMENT PROGRESSES OF HYDROGEN MASER AT SHANGHAI OBSERVATORY

Z.C.Zhai, H.X.Huang, J.W.He, C.F.Lin,
G.X.Jiang, W.H.Luo and J.F.Lu

SHANGHAI OBSERVATORY, ACADEMIA SINICA
Shanghai—200030

ABSTRACT

The development progresses of hydrogen maser during recent years at Shanghai Observatory are summarized as follows:

1. After the success of the first chinese hydrogen maser made at Shanghai Observatory in 1972, 5 more hydrogen masers with several improvements to the maser design, but still traditional structures, were made at Shanghai Observatory for VLBI Observations and time-keeping.
2. To equip chinese VLBI network, a new generation of hydrogen maser, an integrated, rugged and easily transportable maser, was developed successfully in 1987. And so far 5 this kind of hydrogen masers have been put to use in VLBI network.
3. For time-keeping with hydrogen masers, a kind of one-maser auto-tuning system was developed at Shanghai Observatory.
4. At present, Shanghai Observatory meanwhile is developing the third generation of hydrogen maser, a oscillating compact hydrogen maser which employs a compact cavity design, resulting in significant size and weight reductions compared to a conventional maser.This kind of maser is suitable for time-keeping because of its excellent long-term stability.

In this paper, we describe the developments and performances of hydrogen masers at Shanghai Observatory, as well as their applications.

EARLY DEVELOPMENT OF HYDROGEN MASER

Shanghai Observatory began its hydrogen maser development at the beginning of 1970 and developed 6 hydrogen masers altogether, which are of the typical constructions of traditional hydrogen maser. Their design characteristics were described in References [1-3] .Fig 1 shows two of them. These masers are being used in China for VLBI Observations and for time-keeping.

TRANSPORTABLE HYDROGEN MASER DEVELOPMENT

To equip chinese VLBI network, a new generation of hydrogen maser has been developed [4] since 1985, and so far 5 this kind of hydrogen masers have been put to use in VLBI network [5] . The new masers are compact, relatively light-weight, rugged, and easily transportable standards. Entire standard was to be self-contained in a single standard-sized rack cabinet. All electronics including the phase-lock receiver were to be mechanically and electrically integrated into the overall design. It is a relatively small instrument weighing 250kg and requiring a source of 24 Volts DC at 200 Watts to operate. There is a circulating air system in the rack cabinet. The system, which controls the temperature of the pump, the dissociator,and the upper maser electronics,is an entirely independent system with self-contained sensors and electronics. Fig 2 shows the photograph of the complete maser system with the side panel removed.

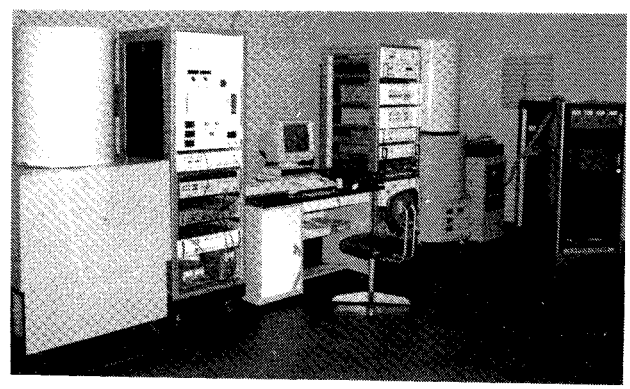


Fig 1. The early models of hydrogen maser made at Shanghai Observatory

A comprehensive series of environmental and short-term stability tests on the new masers has been evaluated [6] . Fig 3 shows the short-term frequency stability. It can be seen from the Fig 3,the stability data keep 10^{-15} levels for the periods of time beyond 100 seconds. In the environmental tests, the output frequency was carefully monitored while one of the environmental conditions was varied. The results of the environmental sensitivities of the new maser are summaried in table 1.

THE DEVELOPMENT OF ONE-MASER AUTO-TUNING SYSTEM

For time-keeping with hydrogen maser, a kind of one-maser auto-tuning system was developed successfully at Shanghai Observatory [7] . The auto-tuning systems have been installed in our hydrogen masers without much perturbing their short-term stabilities. However, the long term stabilities of the masers have greatly been improved.

The principle of the auto-tuning system is described in Fig 4 [7,8] .When a frequency modulated mi-

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

Table 1. Environmental sensitivities [6]

Conditions	Sensitivities	
	slow coefficient	fast coefficient
Temperature (22℃~31℃)	$\triangle T$ up 2.9×10^{-14}/℃ $\triangle T$ down 3.0×10^{-14}/℃	1.2×10^{-14}/℃ 1.3×10^{-14}/℃
Barometric pressure (±25 ㎜)	$< 1\times10^{-14}$	
Magnetic field (±0.4G)	1.1×10^{-12}/G	

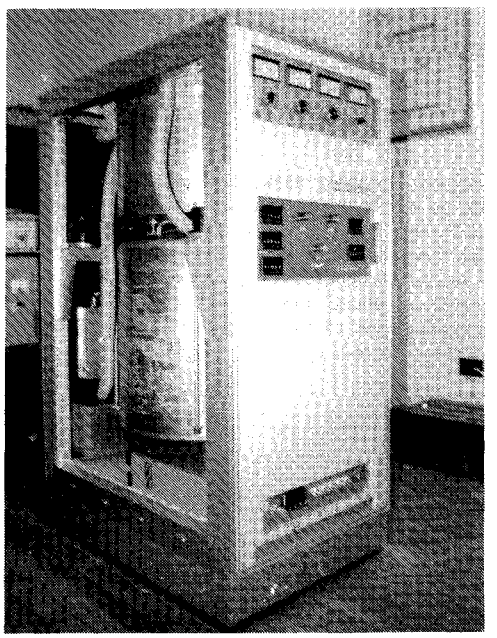


Fig 2. The photograph of the transportable hydrogen maser made at Shanghai Observatory

crowave signal f(t) is transmitted through the cavity, an amplitude response signal A(t) is obtained in the output signal,then applied to the varactor coupled to the cavity by a lock-in amplifier, so that the cavity resonant frequency response is maintained to the desired solid curve throughout. A schematic diagram of the auto-tuning system is shown in Fig 5 [7]

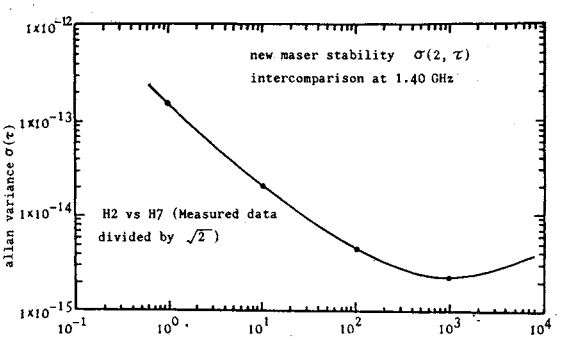


Fig 3. The stability data for the transportable hydrogen maser

A comprehensive series of tests was also made to determine the operating characteristics of the maser with and without the auto-tuning system.Allan Variance data plots, turning on and off the system, are contained in Fig 6 [7] , and clearly show that there is only minimal degration of the maser short term stability due to the introduction of the injected signal of the cavity auto-tuner.However,the long term stability of the maser has greatly been improved.

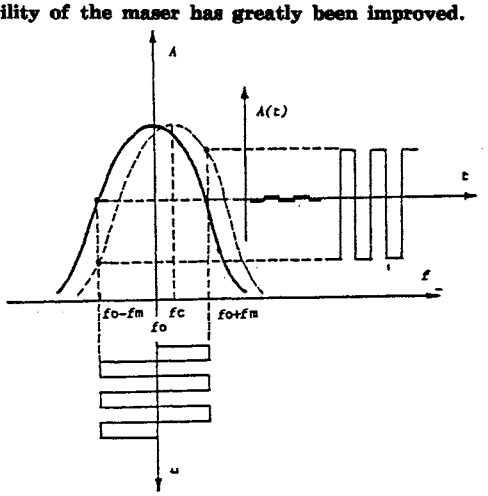


Fig 4. Response of cavity to frequency modulation signal

THE DEVELOPMENT OF SMALL HYDROGEN MASER FOR TIME-KEEPING

At present, Shanghai Observatory meanwhile is developing the third generation of hydrogen maser,a small Q-enhanced maser, which employs a compact cavity design, resulting in significant size and weight reductions compared to a conventional maser.

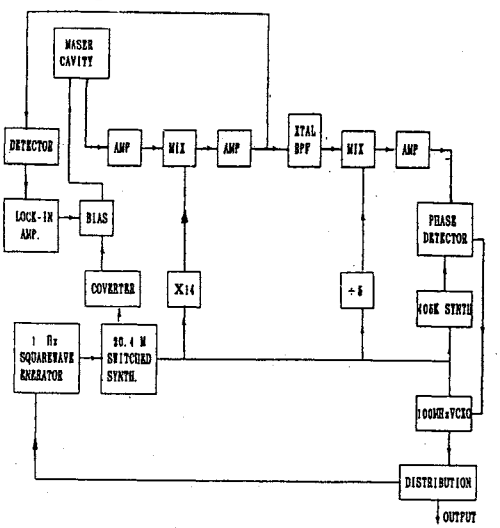


Fig 5. The schematic diagram of the auto-tuning system

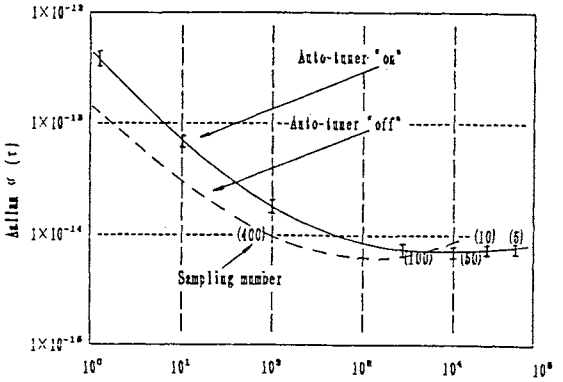


Fig 6. The stability of the hydogen maser with auto-tuner

The operation principle of a Q-enhanced maser oscillator is shown in Fig 7 [9], which is similar to a conventional maser. The major difference is that the losses in the compact cavity in a Q-enhanced maser are such that the maser oscillation condition is difficult to satisfy. This limitation is overcome by positive feedback, as shown in Fig 7. Thus, the

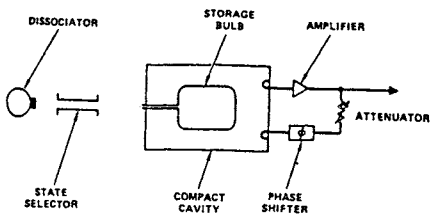


Fig 7. The operation principle of a Q-enhanced maser oscillator

cavity losses are effectively reduced or cavity Q-enhanced and sustained maser oscillation can be obtained. However, the cavity in a Q-enhanced oscillator is not an isolated component as in a conventional maser. In fact, the cavity and the feedback loop form a resonant system that is susceptible to environmental perturbations.It is therefore essential to have a cavity frequency stabilization servo system.

A functional block diagram of Q-enhanced maser is shown in Fig 8 [10]. The cavity Q enhancement and frequency stabilization servo are located, respectively, in the central portion and in the left side of the diagram.

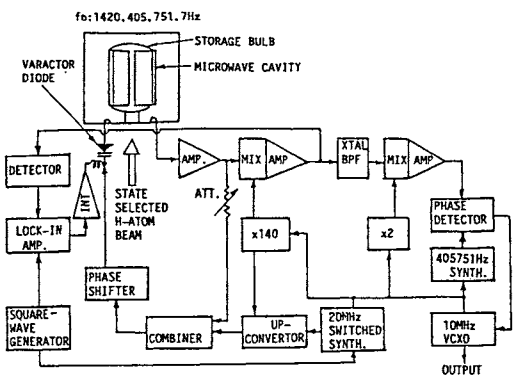


Fig 8. Functional block diagram of Q-enhanced maser oscillator with cavity stabilization servo system

Fig 9 shows the photograph of our Q-enhanced maser oscillator system including cavity stabilization servo [10] developed at Shanghai Observatory. At present,the maser is under assembling and testing.

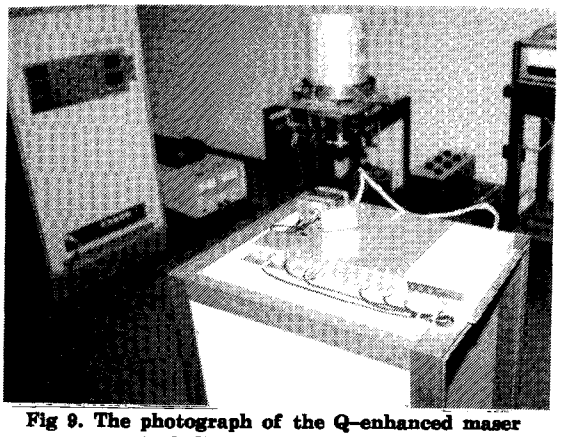


Fig 9. The photograph of the Q-enhanced maser system including the cavity stabilization servo

APPLICATIONS

The major applications of hydrogen masers at Shanghai Observatory have been VLBI Observations. During VLBI experiments, hydrogen maser performances are monitored by a microcomputer real-time measuring system. The stability data is shown not only by Allan Variance,but also in a form that might be more useful for VLBI experimenters. Fig 10 shows phase difference in seconds with only a frequency offset term removed from the data. Notice that the data has the quadratic behavior associated with uniform frequency drift. Fig 11 shows the same data with a uniform frequency drift term also removed.

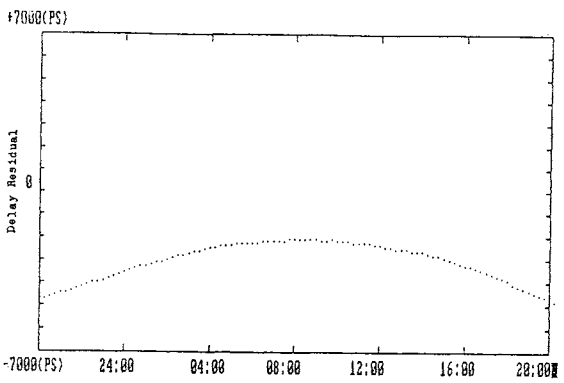


Fig 10. Phase difference with frequency offset removed

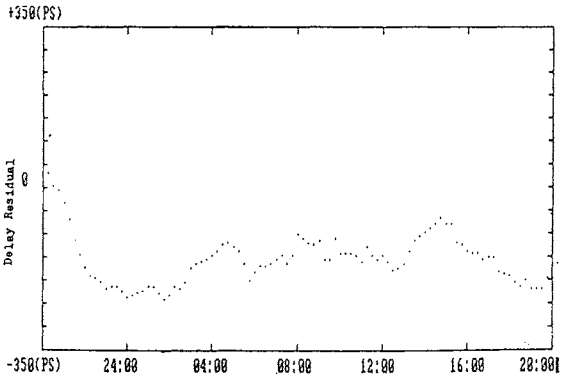


Fig 11. Phase difference with frequency drift removed

Shanghai Observatory's new transportable hydrogen maser is also being used for mobile tracking station, in which very short-term stabilities of the hydrogen maser must be emphasized. The very short-term stability data are as follows:

Sample time	Measuring system Bandwith	Stabilities
10ms	1.6KHz	8.5×10^{-11}
	400Hz	5.1×10^{-11}
100ms	100Hz	7.1×10^{-12}

Hydrogen masers, of course, can be used for time-keeping, especially, the hydrogen masers with auto-tuning.The masers with cavity auto-tuning system have been working continuously for time-keeping at Shanghai Observatory since June, 1991. Fig 12 shows the comparison data measured between the maser clock and UTC by means of GPS receiver.

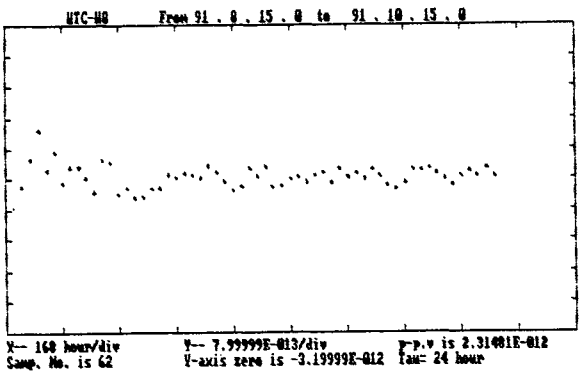


Fig 12. The comparison data between UTC and the maser with auto-tuner

CONCLUSION

Shanghai Observatory's new transportable hydrogen masers have been demonstrated to be the easily transportable and rugged enough to withstand the normal rigors of being moved from one station to another. They are suitable satisfactory frequency standards for VLBI Observations and tracking stations. Doubtlessly, the features of easy transportation and convenient to operate make the masers to be also the ideal time and frequency standards for other fields.

ACKNOWLEDGEMENTS

We are very much obliged to thank Dr.R.F.C. Vessot at SAO, U.S.A. for offering the first author the opportunity to work with him at SAO during the past years.We also greatfully wish to thank our colleagues for supporting these works.

REFERENCES

[1] . Zhai Zao-Cheng, "Shanghai Observatory's Hydrogen maser" , ETE Vol. 27, 11, 1981.

[2] . Zhai Zao-Cheng, etc, "A Brief introduction to the atomic time and frequency standards at Shanghai Observatory" , Proc. of 38th Annual Symposium on Frequency Control, 1984.

[3] . Zhai Zao-Cheng, etc, "Activities on time and frequency at Shanghai Observatory" , Proc. of 18th PTTI meeting, 1986.

[4] . Z.C.Zhai, H.X.Huang, etc, "The new generation of hydrogen maser at Shanghai Observatory" , Proc. of 41st Annual Symposium on Frequency Control, 1987.

[5] . Z.C.Zhai, H.X.Huang, etc, "A new hydrogen-maser time and frequency standard at sheshan VLBI station of Shanghai Observatory" , Proc. of 21st PTTI meeting, 1989.

[6] . Z.C.Zhai, C.F.Lin, etc, "Performance evaluation of the Shanghai Observatory's new H-maser" ,Electromagnetic metrology,International Academia publishers, 1989.

[7] . C.F.Lin, J.W.He, etc, "A hydrogen maser. with cavity auto-tuner for time-keeping" , to be published in the Proc. of 23rd PTTI meeting, 1991.

[8] . C.F.Lin, Y.Ohta, "A cavity frequency stabilization servo system for the oscillating compact hydrogen maser" ,Journal of the CRL, Vol. 36, No.149, 1989.

[9] . H.T.M. Wang, "An oscillating compact hydrogen maser" ,Proc.of 34th Annual Symposium on Frequency Control, 1980.

[10] .Z.C.Zhai, C.F.Lin, J.W.He, etc, "A kind of small hydrogen maser for time-keeping" ,to be published in Proc of 23rd PTTI meeting, 1991.

INVESTIGATIONS OF THE FREQUENCY INSTABILITY OF CH1-75 HYDROGEN MASERS

N.A. Demidov[1], E.M. Ezhov[1], B.A. Sakharov[1], B.A. Uljanov[1], A. Bauch[2], B. Fischer[2]

[1]Institute of Electronic Measurements KVARZ, N. Novgorod, Russia
[2]Physikalisch-Technische Bundesanstalt, Braunschweig, Germany

ABSTRACT

The construction and characteristics of the CH1-75 active hydrogen maser, produced by the Institute of Electronic Measurements KVARZ, Nishnij Novgorod (Russia), are described. Technical features providing low frequency instability are reported in detail. Methods and problems of microwave cavity auto-tuning are analyzed. Frequency instability caused by the auto-tuning microwave cavity system and the crystal oscillator synchronization system are also discussed.

The comparison between two CH1-75 H-masers mutually and with PTB's primary clocks CS1 and CS2 during a 1 year period are presented.

Keywords: Hydrogen maser, time-keeping

1. INTRODUCTION

Hydrogen masers nowadays have the lowest frequency instability among all commercial frequency standards. Their advantage is especially evident at averaging times up to one day - the best hydrogen masers have a frequency intability of $6 \cdot 10^{-16}$ for averaging times of about 10^4 s /1/. Therefore hydrogen masers are indispensable in VLBI (Very Long Base Line Interferometry). For longer averaging times the frequency instability is deteriorated mainly due to detuning of the cavity and possible ageing of the storage bulb coating, resulting in a drift of the wall- shift. The latter may have an essential influence for averaging times longer than one month.

Hydrogen masers of the CH1-70 and CH1-80 type are widely used in the national metrological institutions for time-keeping in the former SU. While operated in the cavity auto-tuning mode, they demonstrated their excellent long-term frequency instability, resulting in a systematic frequency drift of a few parts in 10^{16} per day /2/.

This paper describes the frequency measurement results of two CH1-75 hydrogen masers. CH1-75 is the newest model of hydrogen masers designed by the KVARZ. Two hydrogen masers CH1-75 were delivered to the Physikalisch-Technische Bundesanstalt (PTB) in Braunschweig (Germany) and put into operation in September 1990. Here, they were compared with each other and the PTB's primary clocks CS1 and CS2. From January 1991 on the readings of the CH1-75 hydrogen masers are sent to the BIPM (Paris).

2. DESIGN PRINCIPLES OF THE CH1-75 HYDROGEN MASER

The schematic representation of the physical package of the hydrogen maser is shown in Fig. 1.

- Vacuum system. The maser has two separated vacuum chambers. The first one includes storage bulb, state selector magnet and atomic hydrogen source; the second one cavity, ovens and magnetic shields. Such a construction permits to achieve excellent vacuum conditions in the storage bulb.

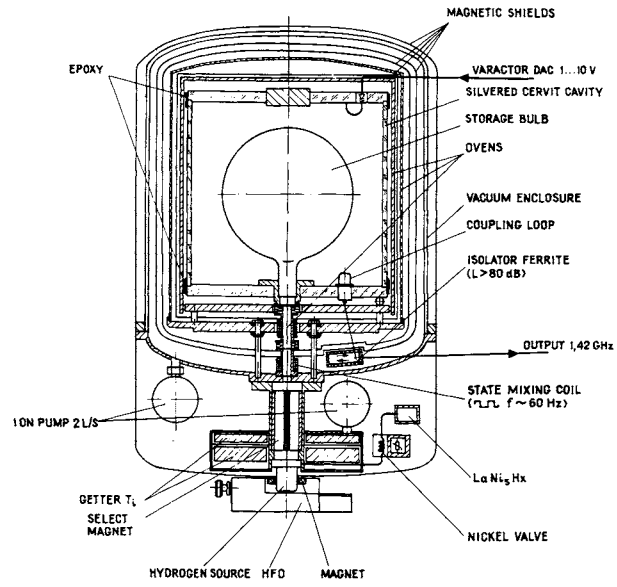


Fig. 1: CH1-75 hydrogen maser's schematic

Hydrogen pumping is performed by means of a sorption pump with titanium as getter material. The getter is activated at 800°C by an internal tungsten heater. The weight of the getter is 1 kg and is sufficient for 5 years of continuous operation. Residual gas is pumped in both vacuum chambers by a (1-2) l/s ion pump.

• Magnetic state selection. The magnetic state selection is performed by means of a quadrupole magnet. The bore diameter of the magnet is 1.6 mm, its length 75 mm and its outer diameter 30 mm. The magnetic induction at the pole tips is 1 T. A glass bead of about 0.5 mm in diameter is located at the exit of the magnet to remove the nearly undeflected high velocity hydrogen atoms from the atomic beam.

• The source of atomic hydrogen. The discharge bulb of the atomic hydrogen source is made of quartz glass. Its diameter is 30 mm. The atomic hydrogen beam is formed by a multichannel collimator, made of Pyrex glass. The channel dimensions are: inner diameter 0.01 mm, outer diameter 0.5 mm, length 0.8 mm. The discharge is excited by a 120 MHz HF-generator. Its power consumption is (10-12) W.

• Supply of molecular hydrogen. The hydride $LaNi_5H_{6,8}$ is used to supply the molecular hydrogen. Its 250 g contain 25 l of hydrogen at normal pressure. The hydrogen pressure is (2-5) atm at a temperature of (30-50)°C.

• Storage bulb. The storage bulb is made of quartz glass and has a geometric storage time of 0.7 s. The bulb is a sphere with a volume of 2 l. The inner surface of the bulb is coated with 3 layers of fluoroplastics F-10. Among other known materials this material provides the smallest wall-shift and spectral line broadening. The wall-shift for a 2 l bulb is $-3.5 \cdot 10^{-12}$ at a temperature of +50°C, the atomic line quality factor Q_l at threshold atomic beam intensity is $(2.5\text{-}3) \cdot 10^9$ /3,4/. At the operating beam intensity it decreases to $Q_l = (1.3\text{-}2.0) \cdot 10^9$.

• Microwave cavity. The cavity resonator is made of SO-110M astroglass ceramic. Its linear thermal expansion coefficient is less than $2 \cdot 10^{-7}$/°C. The temperature coefficient of the resonance frequency of the quartz storage bulb loaded microwave cavity is (250-350) Hz/°C. The dimensions of the cavity are: inner diameter 280 mm, height 250 mm. Metallization of the cavity is performed by galvanization. The Q-factor of the cavity is $(40\text{-}45) \cdot 10^3$. The special design of this cavity minimizes the influence of vibrations, shocks and atmospheric pressure changes.

• Ovens. The maser has a multizone two-stage oven, placed in vacuum. To reduce heat emission, the surfaces of the ovens and the magnetic shields are coated with aluminum deposited under vacuum by evaporation. Heat losses amount to only 0.03 W/°C. The temperature is controlled to about 10^{-4}°C. The temperature coefficient of the maser frequency is less than $3 \cdot 10^{-15}$/°C.

• Magnetic shields. The maser has a five-layer 81 HMA permalloy magnetic shield of 0.5 mm thickness. Four shields are located inside the vacuum system and the fifth one covers the whole maser. The dynamic shielding factor of the shields is $3 \cdot 10^4$ for a change of the field strength, ΔH, of 0.2 Oe and about $5 \cdot 10^4$ for ΔH = 1 Oe. To prevent undesirable mixing of Zeeman sub-levels in the drift region, a ring magnet, placed on the discharge bulb of the atomic hydrogen source assures a smooth transition between the quadrupole magnet and the apertures of the magnetic shields. Besides, this magnet improves the operation of the atomic hydrogen source.

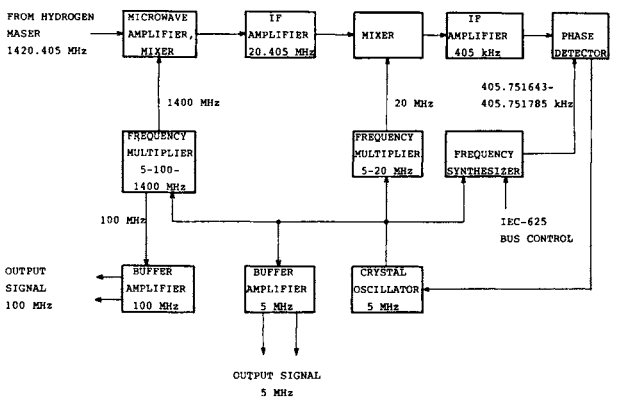


Fig. 2: Block diagram of the synchronization unit

3. ELECTRONIC SYSTEM

The electronic system of the hydrogen frequency standard CH1-75 mainly consists of three subsystems - the synchronization unit, the automatic cavity tuning (CAT) unit and the system of electrical and operational support (power supply, atomic beam intensity stabilizer, temperature control).

• The synchronization unit. A block diagram is shown in Fig. 2. It provides amplification and conversion of the hydrogen maser signal of 1420 MHz, accomplishes the synchronization of the signal phase of the 5 MHz crystal oscillator to the signal phase of the hydrogen maser and delivers output signals at 5 MHz and 100 MHz. The main requirements to the synchronization unit are low phase noise and low instability of the signal phase. This all is provided by the use of a preamplifier for the hydrogen

maser signal with a noise factor of only 2 dB, of buffer amplifiers and frequency multipliers with a phase noise floor of -(145-150)dB/Hz and a temperature sensitivity of the phase of (10-15) ps/°C as well as of a quartz oscillator with a noise spectrum below -(145-150) dB for frequency offsets of more than 1 kHz. A tunable frequency synthesizer permits to change the maser output frequency up to $1 \cdot 10^{-10}$ in steps of $1 \cdot 10^{-14}$.

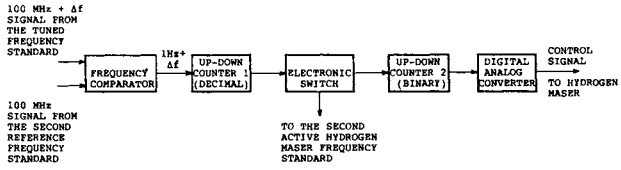


Fig. 3: Block diagram of the automatic cavity tuning system

• Automatic cavity tuning unit.
For automatic cavity tuning the method of modulation of the atomic line quality factor Q_1 is used. In case of a detuning of the frequency of the microwave cavity with reference to the atomic resonance frequency a change of Q_1 results in a maser frequency change due to cavity pulling. The cavity tuning system comprises a frequency comparator, an up-down counter and a simple processor with an output to a digital-analog converter as shown in the block diagram of Fig. 3. The cavity tuning system implemented in the CH1-75 demands a 100 MHz reference frequency of low instability. Therefore the tuning of the cavity is usually accomplished in the following way. The 100 MHz output signals of two frequency standards CH1-75 with a frequency difference Δf are the input signals of the built-in frequency comparator of one of the masers, where an output signal with a pulse repetition frequency of (1 Hz + Δf) is formed. This frequency is measured with the following up-down counter. The counter is in the up (down) mode when Q_1 is high (low) for 100 s. During cavity tuning, Q_1 is periodically changed in the maser to be tuned by applying and removing an additional line broadening inhomogeneous magnetic field. Therefore at the end of every tuning cycle the output of the up-down counter shows a result characterizing the detuning of the cavity of the corresponding device. The tuning is performed in steps. The magnitude of a tuning step depends on the counter's measuring result. If this value exceeds $1 \cdot 10^{-11}$ (more than the retuning range of the cavity by means of the varicap (Fig. 1)), the voltage applied to the varicap is not changed and an error message is displayed. Otherwise, if the absolute value of the counter's reading is in the limits ($9 \cdot 10^{-12}$ - $1 \cdot 10^{-13}$) then, after the appearance of a similar result for two times in succession, the varicap voltage is changed by a large step. Its value may be selected to produce a maser frequency change of 2 to $80 \cdot 10^{-14}$. If the counter's reading is within $9 \cdot 10^{-14}$ - $1 \cdot 10^{-15}$, the cavity frequency is changed by a small step, so that the standard's output frequency changes by $(2\text{-}4) \cdot 10^{-15}$. By continuous tuning, the cavity resonance frequency is tuned in several steps to the center of the atomic resonance line. An example is shown in Fig. 4. For tuning both masers, the role of reference maser and maser to be tuned are exchanged after each cycle.

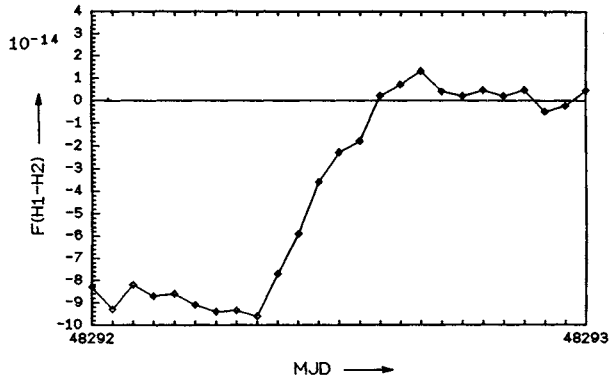


Fig. 4: Relative frequency difference of maser1 (H1) and maser2 (H2). H2 was running free for a week and afterwards CAT was switched on at MJD = 48292, 8h.

In the previous models of KVARZ hydrogen masers CH1-70 and CH1-80 the modulation of Q_1 was performed by changing the atomic beam intensity via the current of the flow controlling nickel valve (Fig. 1). The same can be done in the CH1-75, too. In this case the amount of maser frequency detuning depends on the nickel valve current $I_{\text{Nickel valve}}$, the pressure in the supply of molecular hydrogen and the operating conditions of the discharge of the atomic hydrogen source. It was noted, however, that the lasting modulation of the hydrogen pressure in the discharge bulb detoriorates the stability of the discharge during its long-term operation. That is why it was concluded to use the Q_1 modulation by means of a constant inhomogeneous magnetic field in the masers CH1-75 during routine operation. This magnetic field across the storage bulb is perpendicular to the cylindrical C-field and generated by two coils with oppositely flowing currents. For to avoid a possible frequency shifting influence it was necessary to operate the maser with an unpolarized atomic beam.

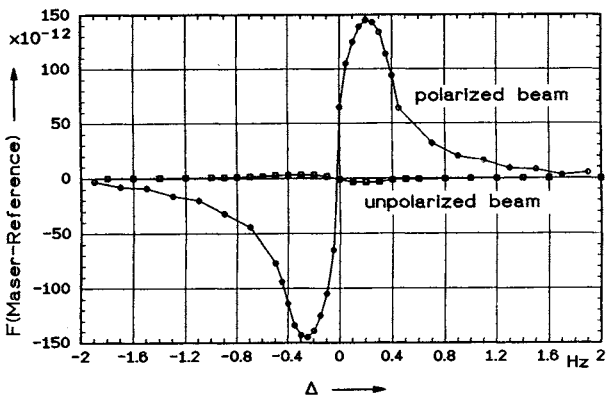


Fig. 5: Relative maser frequency shift as a function of introduced low-frequency (f) irradiation near Zeeman frequency $f_z(\Delta=f-f_z)$

Equal population of the two Zeeman sublevels m=1 and m=-1 of the state F=1 is obtained by creating a nondiabatic rapid passage for hydrogen atoms in the region between the magnetic shields periodically reversing the direction of a small magnetic field. This field is generated by a small coil placed inside the external magnetic shield. To the coil a square wave modulated 80 Hz signal is fed. The degree of state mixing between Zeeman levels m=+1 and -1 is checked by measuring the sensitivity of the maser frequency to Zeeman transition excitation when operating with a polarized compared to an unpolarized atomic beam. In Fig. 5 the results of frequency measurements during simultaneous excitation of Zeeman transitions are shown. Near the true value of the Zeeman frequency ($\Delta=0$) this sensitivity is reduced by more than a factor of 20-30. The maser power is reduced only by 5-10 % when operating with an unpolarized beam. The frequency difference of a maser tuned once by an intensity modulation of the atomic beam and another time by the method described above has been measured to be $(1-2)\cdot10^{-13}$.

In addition to the possibility of using the method described for the Q_1-modulation, maser operation using an unpolarized beam is favourable as the influence of external inhomogeneous magnetic fields is reduced.

4. MEASUREMENT OF FREQUENCY DIFFERENCE AND FREQUENCY INSTABILITY OF THE MASERS

The two hydrogen masers were located in PTB's atomic clock hall, a shielded room with temperature controlled environment where PTB's primary and commercial cesium clocks are installed.

The measurement of the frequency difference and frequency instability of both masers was performed with the help of the phase comparator CH7-48, designed by KVARZ. The phase comparator contains four independent channels each operating with two different input frequencies of 5 or 100 MHz, respectively. By error multiplication output signals of 1 Hz, 1 MHz and 5 MHz plus different error multiplication factors are obtained. In the measuring set-up the comparator operating with the highest error multiplication factor (10^4) was connected to an HP digital time interval meter and furthermore to a PC processing the measuring results. The comparison of the maser frequencies with PTB's primary clocks CS1 and CS2 was performed by measuring phase time differences.

5. RESULTS

Fig. 6 (curve CAT OFF) depicts the short-term frequency instability of a single maser with CAT switched off. The maser frequency instability is at $(2.5-3)\cdot10^{-13}$ for 1 s and reaches its minimum of $(5-7)\cdot10^{-16}$ for sample times of 10^4 s. For long sample times the frequency intability is deteriorated (dotted line) mainly due to cavity detuning. But this demonstrated short-term frequency instability of a free-running CH1-75 maser is comparable with the data from the best hydrogen masers commercially available /1,5/. For the purpose of time-keeping the most interesting point is the performance of hydrogen masers for long sample times ($\tau>10^4$ s) in the CAT mode.

Since November 1990 continuous frequency comparisons between the two masers and the primary clocks CS1 and CS2 have been made. The masers' behaviour in the CAT mode during several days is shown in Fig. 7. Every one measuring value is an one hour average value. Fig. 8 gives an overview of the whole comparison period using one day average values. Around MJD = 48289, 48436 and 48673 maser H2 was operated for one week with CAT switched off, and the intention was to measure its individual frequency drift rate.

A frequency jump attracts the attention, which took place after the output frequency of maser H2 was offset by $810\cdot10^{-15}$ by means of the internal synthesizer at MJD 48469. This constant offset is suppressed in the data shown. Both masers had been delivered to PTB carefully frequency calibrated to UTC(SU). Therefore the existence of a mutual frequency pulling effect was

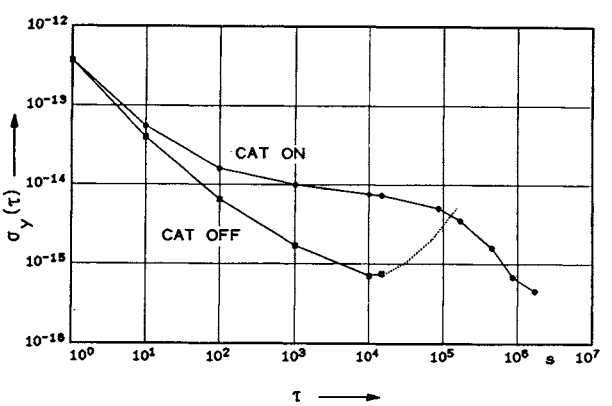


Fig. 6: Frequency instability $\sigma_y(\tau)$ of the hydrogen maser CH1-75 (MJD = 48239 to 48469)

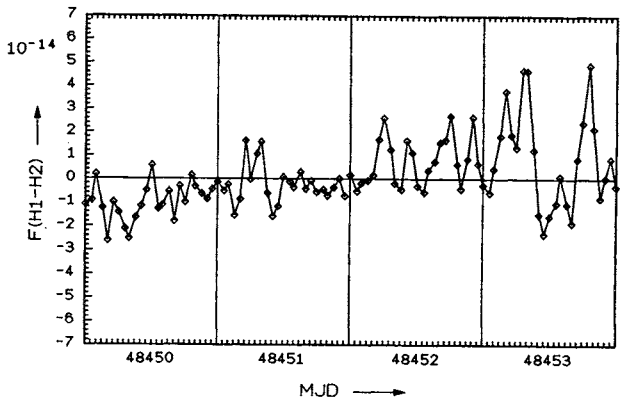


Fig. 7: Relative frequency difference F(H1-H2) of hydrogen masers H1 and H2 in the CAT mode (one hour average values)

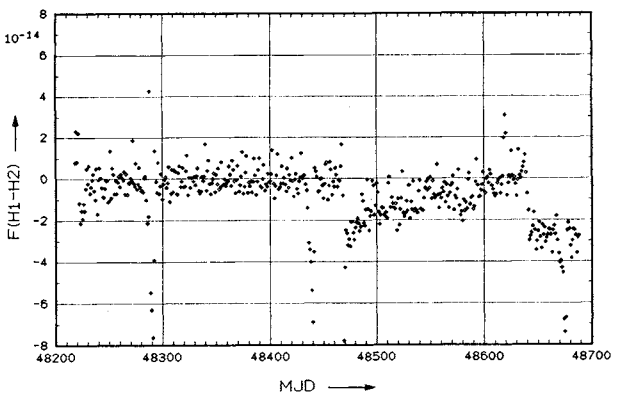


Fig. 8: Results of a long-term frequency comparison between masers H1 and H2 (1-day average values)

suspected and the previous frequency difference of a few parts in 10^{15} was enlarged by the before-mentioned amount. An explanation of such a pulling effect is not yet found. The tests performed at KVARZ with other masers didn't reproduce this effect. The same experiment, performed with a different hydrogen maser CH1-75 at the laboratory of G. Busca (ON, CH) did not show this effect, too /6/.

At MJD 48641 the supply of hydrogen was replaced in maser H1. After replacing the hydrogen supply the frequency difference of both masers was changed by about $2 \cdot 10^{-14}$. The linear frequency drift of the masers' frequency difference in the interval MJD 48239 - MJD 48639 was less than 10^{-16} per day. Fig. 9 and 10 show the relative frequency difference between each single maser and PTB's primary clocks. Obviously both masers' frequency is changing in time similarly. The drift of both masers compared to CS1 and CS2 in the interval MJD 48289 - MJD 48679 was about $-1.5 \cdot 10^{-16}$ per day.

The frequency instability $\sigma_y(\tau)$ of the masers' frequency in the CAT mode is shown in Fig. 6 for sample times τ of 1 s - 20 days. It is evident that the CAT system deteriorates the instability of the maser frequency in the sample time interval $\tau=(10^2 - 10^4)$ s whereas in each maser the tuning cycle is performed once every 460 s. The prolonged disturbance found in the maser frequency is possibly connected with the presence of flicker noise processes with typical time constants of $(10^3 - 10^4)$ s.

The CAT system for averaging times longer than 10^5 s reduces the instability to $(5-7) \cdot 10^{-16}$ for $\tau=20$ days. Avoiding overinterpretation of these results which have been obtained for two very well kept masers, nevertheless they demonstrate for the first time that the frequency instability of hydrogen masers may drop below the 10^{-15} barrier for $\tau=10^6$ s. The traditional increase in frequency instability of hydrogen masers at $\tau>10^5$ s has not been observed in our data.

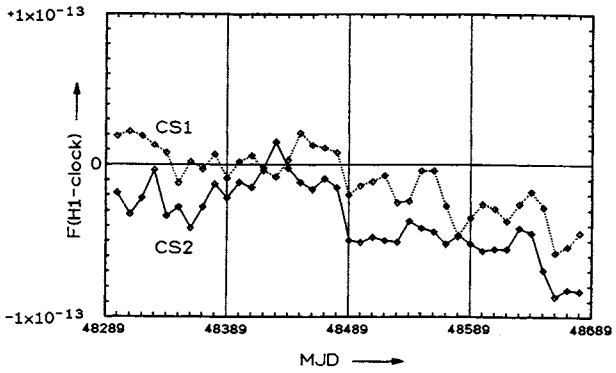


Fig. 9: Relative frequency difference between maser H1 and PTB's primary clocks CS1 and CS2

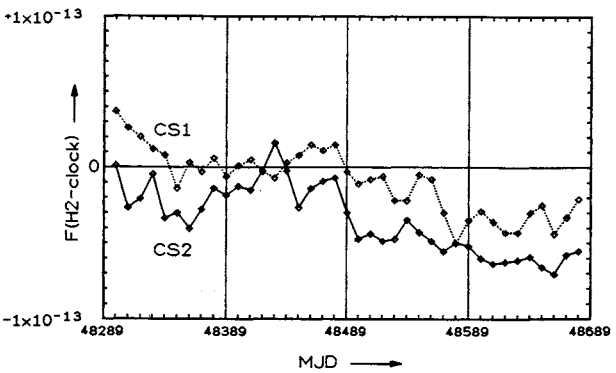


Fig. 10: Relative frequency difference between maser H2 and PTB's primary clocks CS1 and CS2

From the readings of the digital-to-analog converter (DAC) of the CAT system it is possible to evaluate the detuning of the microwave cavity in time. Simultaneously, the potential frequency drift rate of the

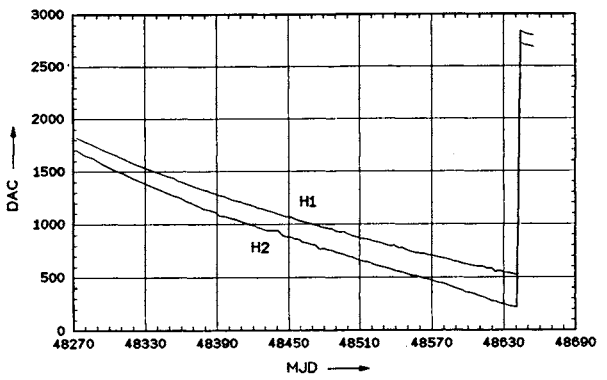


Fig. 11: Both masers digital-analog converter (DAC) indication as a function of time

masers, in the free-running mode, as would be observed, can be inferred. Fig. 11 shows the DAC reading of both masers with time. The frequency tuning by one step of DAC in maser H1 corresponds to $2.8 \cdot 10^{-15}$ and in maser H2 to $2.1 \cdot 10^{-15}$, respectively. From Fig. 11 it is estimated that in the beginning of the measurements the state of DAC in maser H1 was changing by 5 steps per day, in maser H2 by 6 steps per day, which implies a frequency drift of $-1.4 \cdot 10^{-14}$ per day for H1 and $-1.3 \cdot 10^{-14}$ per day for H2. The results for H2 have been confirmed in periods when its CAT system was switched off (Fig. 7). In January 1992, the frequency drift of the cavities is smaller by a factor of 2. The drift value is two times larger than expected from the experiences with other masers /1,4/.

The most important operating parameters of a H-maser are the line quality factor Q_l of the atomic resonance line, the ratio of its maximum to minimum value during cavity tuning, the Zeeman frequency and the variation of these parameters in time. These data are collected in Tab. 1. The first data have been taken at the date of the manufacturing of the masers.

Table 1

Date	Maser parameters							
	$I_{Nickel\ valve}$, A				mode CAT ON		f_z, Hz	
	0.45	0.72	0.48	0.68				
	H1		H2		H1	H2		
	$Q_l \cdot 10^{-9}$				Q_l^{max}/Q_l^{min}		H1	H2
6.11.1989	-	-	2.1	1.3	-	-	-	-
4.01.1990	2.0	1.0	-	-	-	-	-	-
20.11.1990	-	-	-	-	1.5	1.48	601.7	566.0
20.01.1992	1.95	1.05	1.6	1.1	1.5	1.5	605.8	569.5

It is obvious, that Q_l of maser H1 practically did not change. A small change of Q_l is seen in maser H2, possibly at the very beginning because from 20.11.90 till 20.01.92 there was no change in the ratio Q_l^{max}/Q_l^{min} in this maser. The magnetic field in the masers and therewith their Zeeman frequencies changed only by (3.5-4) Hz, which corresponds to a frequency change of maser H1 by $+4.7 \cdot 10^{-15}$ and $+3.8 \cdot 10^{-15}$ of maser H2, respectively.

6. CONCLUSIONS

The comparison of the output frequencies of two CH1-75 H-masers mutually was performed and also that of every individual maser with reference to PTB's primary clocks during more than one year.

The instability of the maser frequency in the free-running mode was $\sigma_y(\tau)=(2.5\text{-}3) \cdot 10^{-13}$ for a sample time τ=1 s and attains its minimum value of $\sigma_y(\tau)=7 \cdot 10^{-16}$ for $\tau=10^4$ s. In the automatic cavity tuning mode the instability of the maser frequency is $5 \cdot 10^{-16}$ for τ=20 days, which essentially shifts the typical increase of the frequency instability of conventional H-masers towards more prolonged sample times. During this comparison the drift rate of the frequency difference of both masers was less than 10^{-16} per day, whereas the frequency drift rate of each maser was measured to be about $-1.5 \cdot 10^{-16}$ per day when referenced to PTB's primary clocks CS1 and CS2. These measuring results demonstrate the feasibility of this type of hydrogen maser to generate a very uniform time scale.

7. REFERENCES

1. A.A. Uljanov, N.A. Demidov, E.M. Mattison, R.F.C. Vessot, D.M. Allan and G.M.R. Winkler, Performance of Soviet and U.S. Hydrogen Masers, Proc. 22nd Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, p. 509 (1990)
2. N.B. Koshelyaevsky, S.B. Pushkin, USSR National Time Unit Keeping over Long Interval Using an Ensemble of H-Masers, Proc. 22nd Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, p. 97 (1990)
3. N.A. Demidov, E.M Ezhov, V.A. Phedorov et al., Frequency Wall-Shift Investigation in the Hydrogen Maser in New Materials, Proc. VNIIFTRI issledovaniya v oblasty izmereny vremeny i chastoty, Moscow, 1978, v. 37 (67), p. 81-85
4. N.A. Demidov and A.A. Uljanov, Design and Industrial Production of Frequency Standards in the USSR, Proc. 22nd Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, p. 187 (1990).
5. Active Hydrogen Maser Frequency Standard EFOS H-Maser, ON Observatoire Cantonal, Neuchâtel (CH)
6. G. Busca, private communication.

ION TRAP FREQUENCY STANDARD RESEARCH AT PTB

D. Schnier, A. Bauch. R. Schröder, Chr. Tamm

Physikalisch-Technische Bundesanstalt,
Bundesallee 100, D-3300 Braunschweig

ABSTRACT

The transition between the hyperfine levels F=1, m_F=0 and F=0, m_F=0 in the ground state of $^{171}Yb^+$ is investigated as a reference transition for a trapped-ion frequency standard. The report covers a description of the experimental set-up and the measurement procedure. The line has been recorded with a linewidth of 16 mHz. Results of the frequency instability and the first steps of an evaluation of the uncertainty contributions are given.

Keywords:

Frequency standard, stored ions,
optical and microwave spectroscopy

1. INTRODUCTION

The high resolution spectroscopy of ions stored in traps is regarded as a promising method to realize advanced atomic frequency standards. Efforts in this direction have been undertaken in laboratories all around the world and the potential merits have already been reported /1/.

At the PTB, the objective was to set up a frequency standard suffering from a mimimum of systematic offsets. But the decision has been made to introduce the new techniques into the laboratory step by step, starting with a fairly flexible set-up of a small ion trap (inner diameter 6 mm) which first allows microwave spectroscopy to be applied to a cloud of buffer gas-cooled ytterbium ions. This is exactly the state which has recently been reached and which is documented in the following.

Next, a brief review is given of ion storage in rf traps. The energy level scheme of Yb^+ is illustrated and used to clarify the spectroscopy methods in use. Subsequently, the experimental set-up is described and the results of measurements are then presented and discussed.

2. ION STORAGE IN A RADIO-FREQUENCY TRAP

The confinement of charged particles is achieved inside a trap structure made of hyperboloids of revolution, a ring electrode and two endcaps. A voltage $U(t) = U_0 + V_0 \cos \Omega t$ is applied. The potential inside the trap is then given by

$$\phi(r,t)=U(t)(x^2+y^2-2z^2)/(r_0^2+2z_0^2)+\phi_0. \quad (1)$$

The motion of the ions, the condition of stability etc. have been extensively described in the literature /2,3/. It has components of oscillation in radial (r = x or y) and axial (z) directions with the frequencies

$$\omega_r = \beta_r \cdot \Omega/2 \qquad \omega_z = \beta_z \cdot \Omega/2 \quad (2)$$

where

$$\beta_z = \sqrt{a_z+q_z^2/2} \text{ and } \beta_r = \sqrt{-a_z/2+q_z^2/8} \quad (3)$$

and a_z, q_z are quantities which depend directly on the trapping conditions:

$$a_z = -16\, q\, U_0 / (m\Omega^2(r_0^2 + 2z_0^2))$$
$$q_z = 8\, q\, V_0 / (m\Omega^2(r_0^2 + 2z_0^2)) \quad (4)$$

The ion's charge to mass ratio is q/m. The micromotion of frequency Ω is superimposed on the motion with frequencies ω_r, ω_z and its amplitude increases with the distance from the trap centre as it originates from the trapping fields.

When optimum storage parameters are used, the average kinetic energy in ion clouds has been found to be about a tenth of the depth of the pseudopotential well /4/. The latter is defined by /2/.

$$D_r: = D_x = D_y = q_z \cdot V_0/16 + U_0/2$$
$$D_z = q_z \cdot V_0/8 - U_0/2. \quad (5)$$

More detailed investigations suggest that the pseudopotential model might be of a limited value /5/. Yet the simple model may serve to give an idea of the energies involved and to compare various experimental situations.

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

The ions' kinetic energy can be reduced by collisions with a light buffer gas such as helium. Cooling to about T_ω = 560 K has been reported for a cloud of mercury ions in an atomsphere of 10^{-5} mbar of helium. Here, T_ω is understood as the kinetic energy of the macromotion, expressed in temperature units /6/.

The presence of a buffer gas atmosphere usually increases the ions' storage time in the trap. It also modifies the number density distribution of the ion cloud and thus the charge distribution. This in turn modifies the storage parameters a_z and q_z and thus ω_z. Altogether, the spectral line shape and width of optical and microwave transitions, and the resonance frequency of the clock transition depend in a complex fashion on the storage parameters, ion number and buffer gas pressure. Several theoretical approaches have been developed which account for various experimental situations, and have recently been compiled by Blatt /7/.

3. ENERGY LEVELS OF YB^+ AND THE SPECTROSCOPICAL METHOD

The rare earth element ytterbium was chosen in the experiments as it exhibits several advantageous features. First, singly-ionized ytterbium Yb has an alkali-like $^2S_{1/2}$-$^2P_{1/2}$ transition with a wavelength of 369 nm, within reach of frequency-doubled continuous-wave dye, solid-state and diode lasers. The large oscillator strength of the $^2S_{1/2}$-$^2P_{1/2}$ transition allows sensitive optical detection of the ions, optical pumping among the hyperfine levels of the ground-state for the state-preparation, and efficient laser cooling of the trapped ions.
Figure 1 shows a section of the energy-level system of Yb^+ in which the hyperfine splitting has been neglected. The calculated ratio of the decay rates from the $^2P_{1/2}$ level to the $^2D_{3/2}$ and $^2S_{1/2}$ levels is p = 1/152 /8/. The radiative lifetime of the $^2D_{3/2}$ state was measured and found to be $\tau(D_{3/2})$ = 52 ms /9/. Without rapidly returning ions from the $^2D_{3/2}$ level to the ground state, the maximum average fluorescence rate is limited by the spontaneous decay to that rather long-living state at high excitation rates.

One characteristic of the energy-level system of Yb^+ is that, apart from the levels due to alkali-like electron configurations, additional low-lying levels arise from configurations where the $4f^{14}$ shell is excited. The lowest-lying excited state is a $4f^{13}\,6s^2\,^2F_{7/2}$ state with an extraordinarily long lifetime of more than 8 days, as has been concluded from measurements /10/. A theoretical treatment predicts a lifetime of 1533 d /8/.

In the presence of a buffer gas, ions residing in the $^2D_{3/2}$ state are transferred to the quasi-stable $^2F_{7/2}$ state. Fine-structure mixing, inelastic collisions followed by spontaneous decay to the $^2F_{7/2}$ level have been suggested as a likely decay mechanism /11/. Almost complete suppression of population trapping in the $^2F_{7/2}$ level can be achieved if the $^2D_{3/2}$ level is depleted using a second laser field at λ = 935 nm /12/.

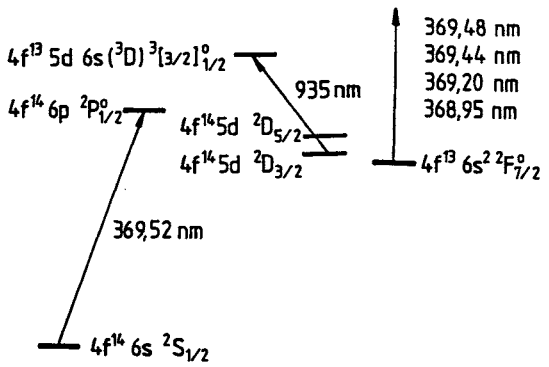


Fig. 1: Relevant energy levels in Yb^+ and used transitions for optical pumping and detection of ions.

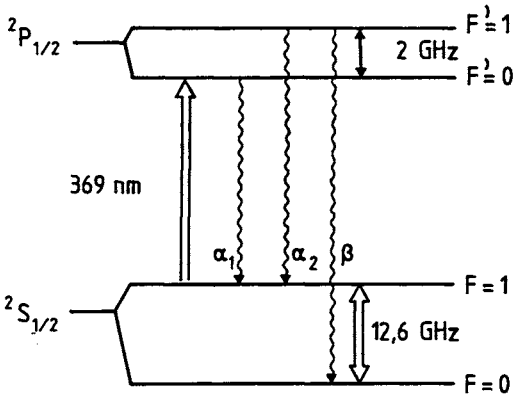


Fig. 2: Hyperfine structure of the ground and the first excited state in $^{171}Yb^+$.

To perform optical - microwave double - resonance spectroscopy, $^{171}Yb^+$ which has a nuclear spin of I = 1/2 is used. The hyperfine levels of the ground and excited states are shown in Fig. 2. The three optical excitations α_1, α_2 and β are indicated; note that the transition (F=0)-(F'=0) is forbidden by selection rules. The ground state hyperfine splitting is 12.64 GHz, and that of the $^2P_{1/2}$ state 2.15 GHz. In analogy to the case of the cesium clock, the (1,0)-(0,0) ground-state transition can be used as a "clock transition" since its resonance frequency shows only a quadratic dependence on the magnetic field:

$$f_{0,0}(B) = f_0 + \kappa_0 \cdot B^2 \qquad (6)$$

with f_0 being the unperturbed resonance frequency. The magnetic field-dependent transitions can be used to control the magnetic field, as their resonance frequency is given by

$$f_{0,\pm1}(B) = f_0 + \kappa_1 \cdot B \qquad (7)$$

κ_0 and κ_1 have the values $\kappa_0 = 3.1 \cdot 10^{10}$ Hz/T^2 and $\kappa_1 = 1.4 \cdot 10^{10}$ Hz/T /13/.

To record one of the microwave transitions, the ion population is first transferred to the F=0 state by optical pumping. The microwave field is then applied in a single pulse of duration τ_i while the laser beams are blocked. Then the laser radiation is admitted to the trap again and the fluorescence photons are counted while the ions are again transferred to the F=0 level. The laser frequency is set to the low-frequency wing of the line α_1. Although due to the Doppler effect, the optical transition frequencies are broadened to a width of ≈ 2 GHz, under the condition chosen the cycling transition F=1 ⇒ F'=0, α_1, is excited more frequently than the α_2 transition. A large number of fluorescence photons are therefore emitted before the ions are finally trapped in the F=0 state. As mentioned above, ions having undergone a spontaneous decay to the $^2D_{3/2}$ level are repumped by the laser field at 935 nm which is also admitted to the trap during the detection phase.

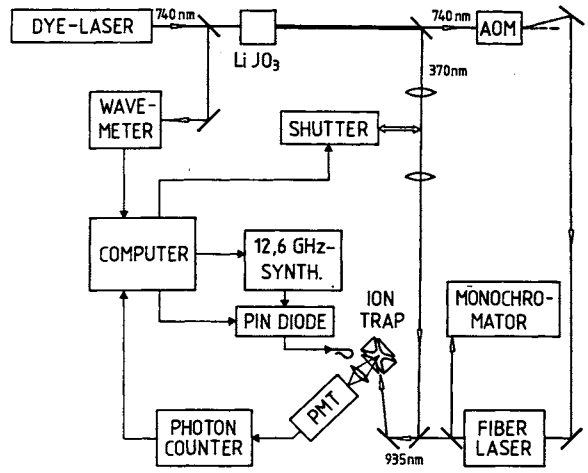


Fig. 3: Experimental set-up.

4. EXPERIMENTAL SET-UP

The experimental set-up is shown in Fig. 3. The output of a continuous-wave single-mode ring dye laser (Coherent CR 699-21) is frequency-doubled to the wavelength around 370 nm in a single pass through a lithium iodate crystal. An optical power of 3 μW is obtained with 0.5 W of fundamental power. The radiation at 935 nm is generated with a fiber laser /12/, operating as a three-level laser on the $^4F_{3/2}$-$^4I_{9/2}$ transition of Nd^{3+}. The laser is pumped by the fundamental dye laser radiation. 10 mW of output power at λ = 935 nm is obtained in an emission bandwidth of approximately 1 GHz. The radiation is sent through the trap collinearly with the UV beam. Resonance fluorescence photons at λ = 370 nm are collected by photomultiplier connected to a photon counter. The scattering ligth background is typically < 100 photons per second. The overall detection efficiency of photons is calculated at about 10^{-3}.

Fig. 4 shows a sketch of the ion trap. The trap electrodes are machined from copper. The gap between the ring and the endcaps of ≈ 1,5 mm allows access of the laser beams. The UV polarisation is perpendicular to the plane shown in the diagram. The source of atomic ytterbium is a metallic pellet, spot-welded to a Pt-Ir filament which is heated during the loading of the trap. Two such "ovens" are used, one with the natural isotope mixture of ytterbium, one with isotope-enriched material consisting of 89 % ^{171}Yb /14/. Ionizing electrons are admitted through a bore in one endcap. Fluorescence light passes through another bore in the opposite endcap and is collimated to a roughly parallel beam by a quartz lens inside the vacuum system. Outside the vacuum enclosure, which is made of quartz glass with brazed optical windows, the light passes an interference filter and is focussed through a variable aperture before reaching the photomultiplier.

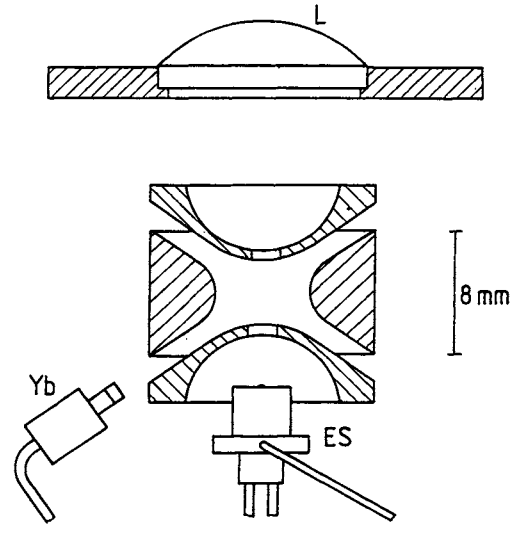


Fig. 4: Schematic drawing of the ion trap used; ES: electron gun, Yb: source of atomic Yb ("oven"), L: quartz lens.

The vacuum system consists of the trap, a variable leak valve and a quadrupole mass-analyzer. During bake-out at 250°C it is rough-pumped with a turbomolecular pump. By means of a 20 l/s getter pump, a base pressure of 10^{-7} Pa is then reached. Ion spectroscopy was usually performed using a helium buffer gas of 10^{-5} - 10^{-3} Pa and a deactivated ion pump. Occasionally, the pump remained in operation and a small helium pressure was adjusted by slightly opening the variable leak valve. Under both conditions storage times of several hours, even days, were achieved.

Some relevant data on the trap and the storage parameters are listed in table 1.

Table 1 Trap and storage data

r_0 = 3 mm, z_0 = 2.4 mm
DC voltage U_0 = 9.5 V
AC voltage V_0 = 400 V
driving frequency $\Omega/2\pi$ = 2.23 MHz
a_z = -0.03, q_z = 0.56
pseudo potential D_z = 19 eV
axial oscillation frequency $\omega_z/2\pi$ = 350 kHz.

Ion clouds are also detected electronically /15/. The maximum signal was interpreted as being due to $4 \cdot 10^4$ trapped ions but $1 \cdot 10^4$ ions is the more common figure. The theory developed by Garboriaud et al. was used to determine the number of ions /16/.

In a region of 15 cm around the ion cloud, no magnetic materials, not even stainless steel, are used. The quartz vacuum housing is surrounded by an inner magnetic shield (one layer of Mu-metal 1 mm thick). Inside the shield, three orthogonal sets of coils were mounted to compensate for residual fields. The inner shield is contained inside a lighttight housing which is surrounded by another layer of magnetic shield.

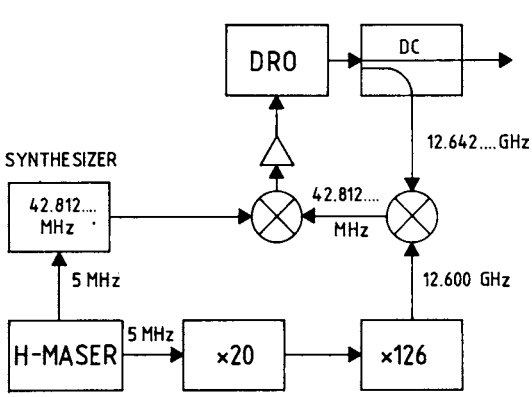


Fig. 5: Electronics for the generation of the microwave signal to probe the ions; DRO: dielectric resonator oscillator; DC: directional coupler.

The microwave-generating system is a spin-off from the development of the electronics for the PTB's primary cesium clocks. No servo-loop operation was performed; instead, the line centre of the clock transition was measured using the output frequency of an active hydrogen maser as reference. This frequency is known on the long term with respect to the PTB's CS2 to ~10^{-14} . Fig. 5 shows a diagram of the frequency synthesis chain. The maser 5 MHz output signal is multiplied to 12.6 GHz. The first three steps are low-phase noise transistor multipliers. The fourth step (100 MHz to 12.6 GHz) is made using a step recovery diode. The signal at 12.6 GHz is mixed with the output frequency of a dielectric resonator oscillator at f_p = 12.642812 GHz. The difference frequency is mixed with the 42.812 MHz output frequency of a frequency synthesizer which is also referenced to the maser signal. The beat signal is used to control the DRO. The synthesizer allows a computer-controlled stepping of the microwave frequency with a resolution of 1 mHz.

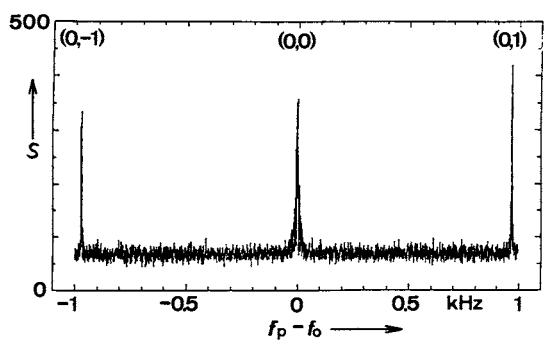


Fig. 6: $^{171}Yb^+$ ground state HFS resonance spectrum; S: fluorescence signal; f_p: microwave probing frequency; f_0: resonance frequency of the (0,0)-transition.

5. EXPERIMENTAL RESULTS

According to (6) it is advisable to operate with a reasonably small value of B with the aim to reduce the sensitivity of the clock transition's frequency to changes of the ambient magnetic field. Making use of the orthogonal coils a mean field averaged over the cloud diameter of less than 10^{-7} T was achieved. The scanning of the frequency f_p of the microwave transitions results in a fluoresence signal S as shown in Fig. 6. The (0,1)-transition has been recorded with a linewidth of 0.5 Hz and no indication of broadening due to inhomogeneities or temporal fluctuations of the magnetic field was observed. By choosing τ_i to 50 s, the linewidth of the clock transition was reduced to 16 mHz; Fig. 7 shows an example. The lineshape obeyed the theoretical relation /13/ in all cases including those where the Rabi frequency b was increased by a factor of 10 compared to the optimum excitation condition $b \cdot \tau_i = \pi$.

It can be read from Fig. 7 that the peak fluorescence signal, S, is of the order of 1000 counts while the electronic detection signal is interpreted as being due to 10000 ions in the trap at the same time. As the total detection efficiency is about 10^{-3}, 100 photons per atom must be generated in the quasi-cyclic excitation before the atoms are transferred to the F=0 level by optical pumping. It has been observed that the optimum excitation wavelength depends on the buffergas pressure in the trap as the ion's kinetic energy and thus the Doppler-width of the optical lines is affected by collisions. A systematic evaluations of this effect is underway.

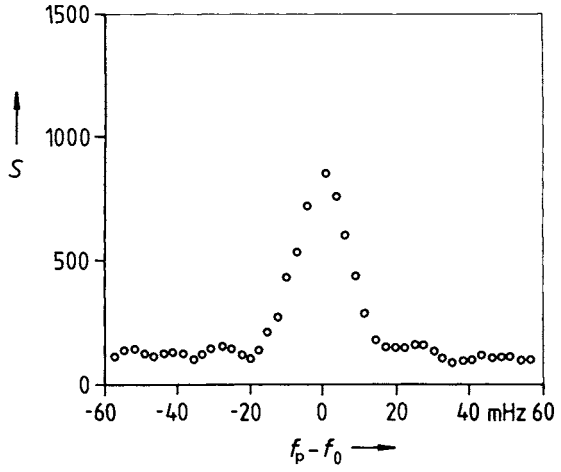


Fig. 7: $^{171}Yb^+$ clock transition line, recorded with interaction time τ_i = 50 s.

The instability of the line-centre has been analyzed by using square wave frequency modulation of the frequency f_p with the modulation width adjusted to the point of the maximum slope of the line. A sequence of counting results S_-, S_+, S_-, S_+.. is collected. Having determined the slope $\Delta S/\Delta f$, the differences $S_- - S_+$ are converted to frequency differences which can be analyzed in the conventional manner. In Fig. 8 the σ-τ diagram for three different measurement series are shown where τ_i had been chosen to 0.8 s, 2.8 s and 9.8 s. The minimum averaging time is $\tau_0 = 2\cdot(\tau_i + 0.5\ s)$, 0.5 s being the duration of laser irradiation and signal processing. In comparison, the σ-τ characteristic of the PTB's CS2 and the hydrogen maser are plotted in Fig. 8, too.

For the time being, the actual level of instability for the Yb^+-line is worse than expected from the signal count-rate and the line Q. Occasionally, the signal height changes with a long time constant, and the initial frequency data contain obvious outliers, typically 1 %. They are caused by sudden drops in the counting signal S. These findings are attributed to fluctuations in the mode spectrum of the fiber laser, occasionally leading to insufficient depletion of the $^2D_{3/2}$ level and subsequent population trapping in the $^2F_{7/2}$ level. This should be verified and eventually be improved in the future.

The determination of the absolute value of the unperturbed F=0 - F=1 transition frequency is largely affected by the determination of the frequency shift due to the

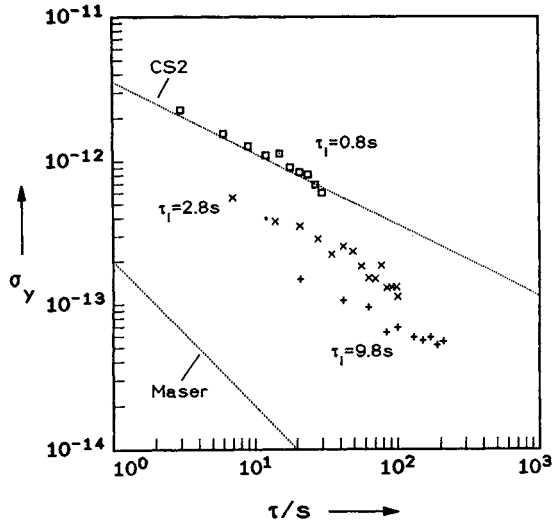


Fig. 8: Instability σ_y of the centre frequency of the $^{171}Yb^+$ clock transition at three different interaction times; for comparison: characteristic of the PTB's CS2 and the reference hydrogen maser.

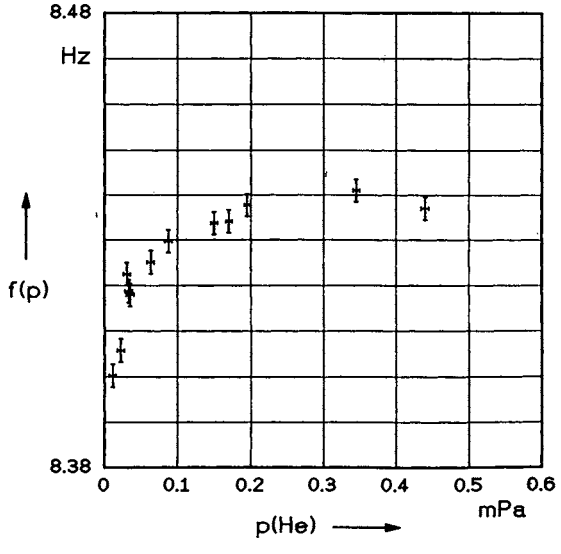


Fig. 9: Centre frequency f_0 of the $^{171}Yb^+$ clock transition as a function of the Helium buffergas pressure. It was f_0 = 12,642,812,110 Hz + f(p), the value has not been corrected for systematic offsets.

ions' kinetic energy and their interaction with the buffer gas. For the time being, no definite result can be given. It shall be analyzed whether the dependence of the fluoresence signal on the excitation wavelength can be used to determine the ions' kinetic energy. Fig. 9 gives an impression of the dependence of the centre frequency f_0 on the helium pressure in the trap. The results are similar to those reported in /6/.

6. CONCLUSION

A status report on the performance of a frequency standard based on ytterbium ions in an rf trap has been given. Promising results have been optained concerning the achievable instability. The evaluation of the uncertainty is underway, yet it can be stated that it will be difficult to surpass a level of 1 part in 10^{13}. The kinetic energy of the ions in the cloud will hamper further improvement. Logically we aim to cool the ions by laser radiation, using smaller ion samples or single ions. It is also planned to study the $^2F_{7/2}$-$^2D_{5/2}$ transition at λ = 3.43 μm as a potential frequency standard reference line. At this wavelength, an absolute frequency measurement will be possible with the PTB's frequency chain. A particular probing laser source, an atomic-iodine laser /17/ is currently under construction.

Another aspect to be addressed is the simplification of the experimental set-up frequency doubling of a λ = 740 nm diode-laser in a passive ring-cavity has already been tested, and more than 10 μW optical power at λ = 370 nm were obtained. The fiber laser could also be replaced by a diode laser so that a reasonably compact and reliable experimental set-up seems to be feasible.

7. REFERENCES

/1/ D.J. Wineland et al., Proc. 4th EFTF, Neuchâtel 1990, p. 267/272.

/2/ H.G. Dehmelt, Advances in Atomic and Molecular Physics 3, 53-72 (1969) and 5, 109-154 (1969)

/3/ E. Fischer, Z.f.Physik 156, 1-26 (1959).

/4/ H. Schaaf, U. Schmeling, G. Werth, Appl.Phys. 25, 249 (1981).

/5/ I. Siemers et al., Phys.Rev. A38, 5121 (1988)

/6/ L.S. Cutler et al., Appl.Phys. B39, 251-259 (1986).

/7/ R. Blatt, Universität Hamburg, Habilitations-Schrift (unpublished) (1988).

/8/ B.C. Fawcett, M. Wilson, Atomic Data and Nuclear Data Tables 47, 242

/9/ Ch. Gerz et al., Z.Phys. D8, 235-237 (1988)

/10/ H. Lehmitz et al., Phys.Rev.Lett. 62, 2108-2111 (1989).

/11/ A. Bauch, D. Schnier, Chr. Tamm, Journal of Mod. Optics 38, (1992)

/12/ Chr. Tamm, D. Schnier, accepted for publication in Optics Communications 1992

/13/ J. Vanier, C. Audoin, The quantum physics of atomic frequency standards. Adam Hilger, Bristol and Philadelphia, 1989

/14/ Oak Ridge National Laboratory, Isotope distribution document: isotopic analysis data

/15/ G. Rettighaus, Z.Angew. Physik 22, 321-326 (1967)

/16/ M.N. Garboriaud, M. Desaintfuscien, F.G. Major, Int. J. Mass Spectr. Ion Phys. 41, 109-123 (1981)

/17/ H.H. Kim et al., IEEE J. Quantum Electronics QE-4, 308-911 (1968)

CONTROLLING AN OPTICAL FREQUENCY BY A SINGLE ION

I. Siemers, V. Enders, R. Blatt, W. Neuhauser, and P. E. Toschek

Institut für Laser-Physik, Universität Hamburg

W-2000 Hamburg 36, Germany

ABSTRACT

The control of the oscillation frequency of a single-mode signal laser by a narrow resonance of a single trapped ion may result in frequency generation of both satisfactory precision and unprecedented accuracy. The signature of resonant excitation cannot be derived from the very weak fluorescence on the control transition. Instead, the rate of dark intervals of simultaneously excited resonance light may be used as a discriminant for the signal resonance. Alternatively, two lasers may excite an E2 transition whose resonance is indicated by a "dark line" in the resonance fluorescence. We have quantitatively studied perturbative effects on the dark lines of a single trapped Ba^+ ion the narrowest of which is 63 kHz and also demonstrated a dark line of a single Yb^+ ion. Systematic shifts of the signal frequency may be eliminated with residual inaccuracy smaller than 10^{-15}.

1. ION CLOUD OR SINGLE ION?

It is generally accepted that ions confined in a trap |1| are sustained in the most benign environment which exposes them to as small perturbation as possible. This seems indeed so up to a high level of experimental sophistication. Advanced experiments on the interaction of microwaves with ensembles of ions and the control of that signal frequency, which include $^9Be^+$ ions |2|, $^{199}Hg^+$ |3|, and $^{171}Yb^+$ ions |4|, have been exploring limits of the stability of the microwave frequency as well as systematic shifts from perturbations.

The residual instability of a frequency locked to a resonance of quality Q is defined as $\sigma_y(\tau) = (Q \cdot S/N)^{-1}$, where S/N is the signal-to-noise ratio of the resonance, which varies as $\tau^{-1/2}$ if N is dominated by shot noise. On 5...10 × 10^3 optically cooled $^9Be^+$ ions in a Penning trap, the control of 303-MHz radiation has been demonstrated with $\sigma_y = 3 \times 10^{-12}\tau^{-1/2}$ and $\sigma_{ymin} = 3 \cdot 10^{-14}$ |2|. Characteristics of the 40.5-GHz hyperfine resonance of $^{199}Hg^+$ ions stored in a linear rf trap are the width $\Delta\nu = 160$ mHz, $Q = 2.5 \times 10^{11}$, and $\sigma_y = 1.6 \times 10^{-13}\tau^{-1/2}$, with $50s < \tau < 800s$, and even $\Delta\nu = 30$ mHz, $Q = 1.3 \times 10^{12}$, and $\sigma_y = 5 \times 10^{-14}\ \tau^{-1/2}$ have been observed |5|. Recently, the 12.6-GHz hyperfine resonance of some 10^6 trapped $^{171}Yb^+$ ions with $Q = 3 \times 10^{10}$, S/N = 50 has been used for the control of the microwave generator, i.e. as a "clock", with $\sigma_y = 2 \times 10^{-11}\tau^{-1/2}$, $\tau < 600s$ |4|. Impressive as these figures are, they do not address the question of systematic shifts, or "accuracy" of the frequencies.

An ensemble of trapped ions, or "cloud" is subject to systematic perturbations of the ions' resonances, since they interact with each other and spread over a substantial volume in inhomogeneous fields: 1. Coulomb collisions result in line shifts and broadening. 2. Stark shifts of the signal line by the trapping field and residual dc fields appear. 3. Gradients of the magnetic field broaden the line. 4. The second-order Doppler effect, even with the modest optical cooling achievable with an ion cloud, may exceed 10^{-13}.

Frequency standards in the optical regime promise extreme Q values but are rather more susceptible to perturbation. Now, the linear Doppler effect still must be compensated or eliminated.

The remedy to all this systematics, it seems, is the use of a *single* ion for long-term measurement of maximum accuracy. Here, the Doppler effect of second order is $< 10^{-18}$ with ideal localization of the ion in the node of the rf trapping field. Optical cooling is so efficient that secular motion and micro-motion vanish, and both the Doppler and Stark shifts disappear.

Locking laser light to a narrow line of an individual ion is not straightforward. There is no way of detecting the very weak fluorescence from the decay of the excited state of a narrow, dipol-forbidden line (typically E2), which is to control the light frequency. However, absorption on that line may be detected by the fluorescence on a resonance line being switched off when excitation of the single ion to the upper level of that forbidden signal line has taken place ("electron shelving, |6|). The random succession of quantum jumps that terminate "on" and "off" intervals |7,8| is determined upon tuning the signal light: Then, the maximum "off" rate marks the resonance of the signal with the E2 transition.

A foremost candidate for such an *optical* standard is a single $^{199}Hg^+$ ion prepared in the lowest vibrational level of the trap potential |9|. The E2 resonance, the $S_{1/2} - D_{5/2}$ transition at 281.5 nm, was two-photon excited by 563-nm light. The narrowest line recorded so far is 172 Hz wide |10|. – The corresponding line of a single Ba^+ has been recorded by scanning 1.76-μm signal light across the E2 resonance; its width is 30kHz |11|.

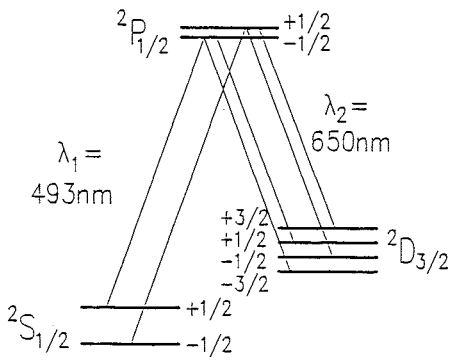


Fig. 1: Ba^+ levels and magnetic sub-levels (m_J). The shown transitions correspond to linear polarization of lasers perpendicular with magnetic field B_z.

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

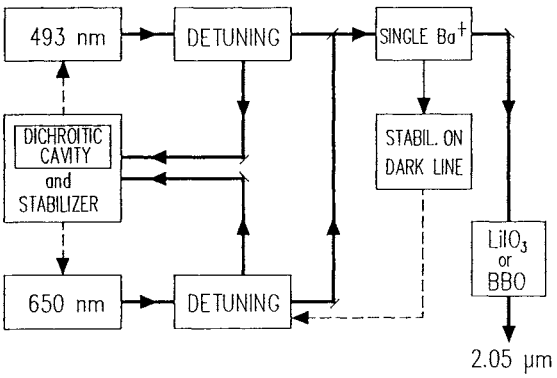


Fig. 2: Generation of 2.05-μm light whose frequency, controlled by a single Ba^+ is stable *and* highly accurate.

2. THE DARK LINE OF SINGLE BARIUM ION

An alternative approach to this scheme for the preparation of a highly perturbationless spectral discriminant includes the use of an E2 transition of very small width, which is excited by two light beams whose *frequency difference* corresponds to the signal resonance. With a Ba^+ ion (Fig. 1), relevant levels are the $6^2S_{1/2}$ ground state, the $6\ ^2P_{1/2}$ resonance level, and the $^2D_{3/2}$ metastable state. They are connected by two dipole transitions driven by the 493-nm and 650-nm light of tunable monochromatic dye lasers. The fluorescence from the ion with this Λ-shaped configuration of states vanishes at equal detunings of the light fields from their respective dipole resonances. This is a "dark line" |12,13| which results from optically pumping the ion into a "trapped state", a non-absorbing superposition of $^2S_{1/2}$ and $^2D_{3/2}$ states |14|. Such a spectral structure was observed in the fluorescence of a trapped and cooled single Ba^+ ion upon scanning the 650-nm laser light across its corresponding resonance |15| while the 493-nm light was kept down-tuned from its line centre for optically cooling the ion |16|.

The frequency of one of the light beams be controlled by the dark line. According to the frequency relation of the electronic Raman transition, the difference frequency of the light fields generated in a nonlinear medium (Fig. 2) will show high stability and extreme lack of systematic shifts. We have studied in detail various effects which might be thought to modify the spectral shape or position of the dark line of a single Ba^+ ion, i.e. which may reduce the *accuracy* of the frequency of the light locked to the dark line, or which are otherwise related. These perturbative effects are (*i*) the Zeeman effect, (*ii*) a light shift induced by the laser beams, to the neighbouring "bright" line and (*iii*) power broadening by the light beams.

The experimental procedure is as follows: One Ba^+ ion is generated inside the trap by electron impact, and its laser-induced fluorescence is detected by photon counting. The trap consists of a ring electrode with 1-mm internal diameter and pin-shaped cap electrodes. Initially, the ion does not reside in the node of the rf field, but is subject to secular motion in the pseudo-potential of the trap, and to the driven micro-motion. The secular motion will be laser-cooled. However, since the centre of the secular orbit is spatially shifted, by the mean light force and, in particular, by spurious electric charge on the surfaces of the electrodes, off the node of the ac field, the ion undergoes residual micro-motion which causes Doppler phase modulation. This modulation is detected by the variation of the counting rate of fluorescence photons – whose signal is processed by a time-to-digital converter – in synchronism with the rf drive of the trap. Four extra electrodes allow compensating the ion's spatial shift by setting their potentials such as to cancel the modulation of the counting rate. In this way the micro-motion is effectively eliminated, and laser-cooling makes the ion approach rest at $T_{fin} < 3$mK.

The light at 493 nm and at 650 nm is generated by tunable dye lasers frequency-controlled by external optical cavities. The fluctuations of the controlling circuits broaden the output spectra by less than 1 Hz and 100 Hz, respectively, when the sampling time of such a measurement is between 100 μs and 1 s. The contribution of mechanical and electrical fluctuations of these reference cavities is a few kHz, and their linear drift is less than 270 Hz/s, as measured by an iodine-stabilized He-Ne laser.

While the green light tuned some 10 MHz below the line centre excites the resonance fluorescence and cools the ion, the red light is scanned across the $P_{1/2} - D_{3/2}$ line. When the detunings of the light fields equal each other, the fluorescence vanishes, and the $D_{3/2}$ state coherently adds to the ground state. This superposition comes as an even and an odd variant, the latter of which does not interact with the light. As a result, the ion is optically pumped into that odd state and remains unexcited. The emerging dark line is Zeeman-split by a magnetic field in up to eight components; their strength and number depend on field direction and light polarization. There are two particular situations: (*i*) The coaxial light beams feature linear polarization in the same direction which extends right angles ($\beta = \pi/2$) with the magnetic field. Four components appear which correspond to combinations of the $\sigma^{\pm}$ transitions on both dipole lines (see Fig. 1). (*ii*) Linear polarizations and magnetic field are parallel: There are Zeeman sublevels $m = \pm 3/2$ in the $^2D_{3/2}$ state which are uncoupled from the light fields, and by optical pumping into these sublevels the ion quickly ceases light scattering.

The optical Bloch equations have been solved for the three-level, eight-sublevel system. The full line in Fig. 3 marks a solution, where the light intensities, the detuning of the green laser, the effective decay rate γ_{sd} of the "dark" superposition of the ionic states $^2S_{1/2}$ and $^2D_{3/2}$, and the magnetic field have been fitting parameters. Their values are derived in the way unambiguously and with best precision since the ion itself acts as a field probe. With various settings of an external magnetic field B_z spectra as those in Fig. 3 have been recorded. Their Zeeman splitting $\mu_B B/h$ is shown, *vs* B_z, in the insert of Fig. 3. Since the total field B is made up of B_z and the ambient field B_0 that consists of residual magnetization and the earth's magnetic field, the splitting never disappears unless the ambient field is precisely compensated.

A remarkable feature of the dark lines is the absence of light shifts, since on resonance the ion decouples from the light. At unbalanced Rabi frequencies of the light fields, the wings of the dark line become unsymmetric, and a "bright" line appears. This line marks the Raman resonance of the *even* superposition of the $^2S_{1/2}$ and $^2D_{3/2}$ states. Its two-photon light shift has been measured |17|; it spans a few MHz with the Rabi frequencies being on the same order.

According to the 17-s lifetime of the $^2D_{3/2}$ level |18|, the observed widths of the dark lines mainly result from power broadening (see insert in Fig. 4). Values of the line widths derived from calculated spectra that were used to fit the observed traces are plotted *vs* the effective intensitiy $(I_{green} + I_{red}) \propto \Omega_e^2$. Each dot corresponds to a spectrum; the widths and depths of whose lines determine Ω_e and

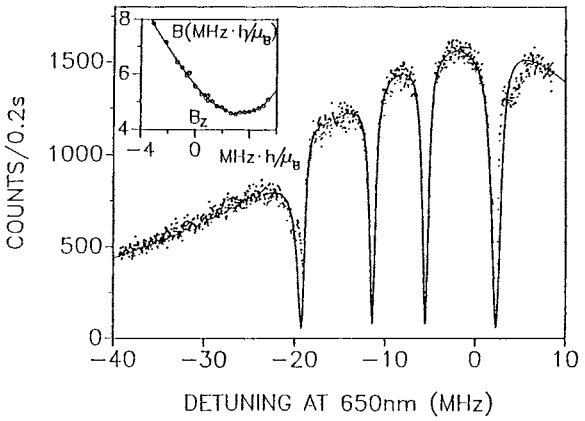


Fig. 3: 650-nm excitation spectrum of the 493nm and 650nm fluorescence of single cooled Ba^+ in a rf trap showing a Zeeman-split dark line. Components at $\pm$ 11u/5, $\pm$ 3u/5 off the green light tuning, – 8.5MHz, off the resonance line. Parameters: $u = \mu_B/hB = 4.7$MHz (B = 0.35mT), $I_g = 10$mW/cm^2, $I_r = 25$mW/cm^2, $\gamma_{sd} = 12$kHz$\times 2\pi$, $\beta = 90^o$. Insert: determined local field B (from splitting) *vs* applied magnetic field B_z.

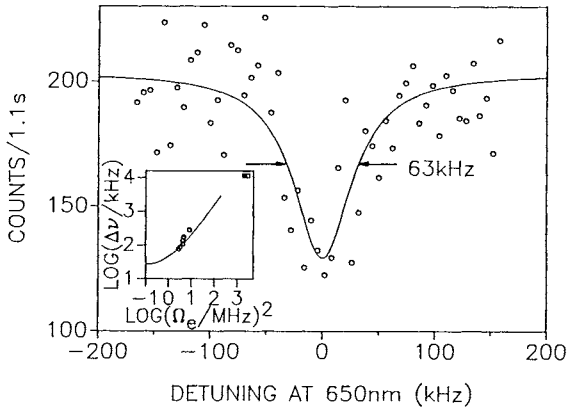


Fig. 4: Dark line at low light intensity. Line $Q \simeq 6 \times 10^9$. $I_g = 0.8\text{mW/cm}^2$, $I_r = 0.5\text{mW/cm}^2$. Insert: Full width $\Delta\nu$ of dark line *vs* $\Omega_e{}^2$, calculated (–,*) with data from fit of recorded spectra, and with Ω_e evaluated of line depth (○), and from measured I_e (□).

the decay rate of the superposition of the $S_{1/2}$ and $D_{3/2}$ states, $\gamma_{sd} = 2\pi \times 12$ kHz. A solution of the optical Bloch equation with the use of this γ_{sd} is shown as the full curve. Note that, with large Ω_e , power broadening makes the line width vary as $\Omega_e{}^2$, whereas in the small-signal limit the width approaches the decay rate γ_{sd} which includes residual noise of the external cavities for the frequency control of the lasers, and fluctuations of the ambient magnetic field.

The narrowest dark line recorded so far is shown in Fig. 4. Its width is 63 kHz, or 75 kHz after the correction for the cavity drift is included, and the quality factor is $\nu_r/\Delta\nu = 6 \times 10^9$, or $(\nu_g - \nu_r)/\Delta\nu = 2 \times 10^9$. This line, with its S/N = 3.5 and $\tau = 1$s sampling time, might control the difference frequency which corresponds to light at 2.05μm with residual instability 1.5×10^{-10}, at $\tau = 1$s. Tenfold reduction seems possible by the use of only one optical cavity with improved stability for the control of both lasers, of a low-noise photodetector, and by shielding the ambient magnetic field. The concomitant reduction of the magnetic fluctuations will benefit the contrast of the dark line. Although even this figure of frequency stability does not look overly impressive, the scheme reveals its intrinsic merits when checked for its *systematic* reliability, or accuracy, to be made use of in long-time calibration of a frequency of an optical oscillator:

1. When the ion is properly laser-cooled and placed in the node of the ac trapping field, both secular and micro-motion vanish. The dc field of the auxiliary electrodes compensates the field of spurious surface charges, the spatial shift from light pressure – which may be cancelled, in a standing light wave, e.g. – , but also the tiny displacement by gravitation. Therefore, in principle the net dc field is non-zero. However, the corresponding frequency shift by the quadratic dc Stark effect is much smaller than 10^{-16}.
2. The second-order Doppler effect of the cold ion is less than 10^{-18}.
3. Shifts from fluctuations of the magnetic field are made to cancel by successively probing *two symmetric* Zeeman components in the process of interrogating the ion for Raman resonance.
4. Earth rotation is known to shift the phase of circularly polarized light |2|. This effect is absent with linearly polarized light of the lasers.
5. The interaction of the light-induced electric quadrupole $S_{1/2} - D_{3/2}$ with an electric field of quadrupolar symmetry shifts the resonance by about 1 Hz per 1 V applied to the ring electrode of the ion trap |19|. If the dc potential of the ring is known with 10mV precision, the fractional uncertainty of the frequency is not larger than 10^{-16}.
6. A frequency shift caused by a lagging servo control is, in principle, improvable by higher gain in the circuit.
7. The shift of the dark line by the blackbody radiation |20|, a few times 10^{-16}, seems to be the largest systematic effect that afflicts a dark line of a cold single ion whose position is perfectly matched with the ac centre of a radio-frequency trap. This shift, however, is calculable with high precision.

3. SINGLE YTTERBIUM ION

Another ionic species of interest for the control of optical frequencies is Yb^+ |21| whose resonance fluorescence is excited by light at 369nm (see Fig. 5). Like the barium ion, Yb^+ shows low-lying $^2D_{3/2,\ 5/2}$ levels, although the $^2D_{5/2}$ level decays to the ground state in 50 ms, unlike Ba^+, and to the metastable $^2F_{7/2}$ state in 25 ms. The $^2D_{5/2}$ level may be excited by two photons of 822-nm signal light from a diode laser. With a single Yb^+ ion of an even isotope, resonant two-photon excitation of this E2 transition will go along with the maximum rate of dark intervals in the simultaneously excited resonance fluorescence at 369nm. Such a single Yb^+ has been prepared recently, and dark intervals of the fluorescence – or "quantised fluorescence" – although induced by collisions with background gas, have been observed. On the other hand, also the excitation, by 467-nm light, of the $^2F_{7/2}$ level whose lifetime even exceeds 7d will be monitored by quantised fluorescence. This E3 resonance would be characterized by a line quality factor $Q > 10^{21}$.

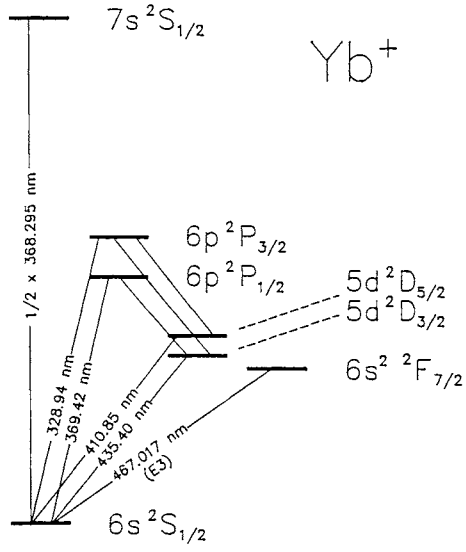


Fig. 5: Simplified level scheme of Yb^+ including dipole-forbidden lines for the control of optical frequencies.

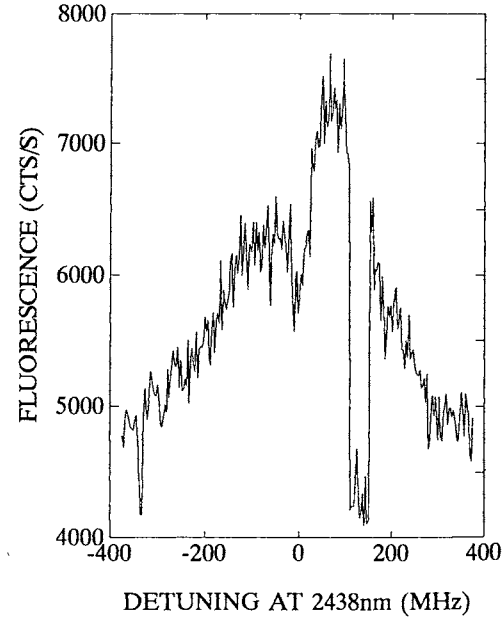


Fig. 6: Excitation spectrum of (total) resonance fluorescence, at 369.4nm, of *single* Yb^+ *vs* detuning of 2nd laser at 2438nm. Dark line is fitted with solution of Bloch equations.

Also, the $^2S_{1/2} - {}^2D_{3/2}$ dark line in the laser-excited 369-nm fluorescence of a single Yb^+ has been recorded by slightly downtuning the UV laser from the resonance line for cooling the ion, and scanning an F_2^+ colour centre laser (based on KCl:Li) at 2.44μm across the corresponding $P_{1/2} - D_{3/2}$ line. A preliminary excitation spectrum is shown in Fig. 6. The 50-MHz width of the dark line close to zero detuning is caused mainly by power broadening and the ion's location deviating from the rf field node. During the scan of the line's right wing, the ion, jumping to the $D_{5\ 2}$ level, interrupted the emission of fluorescence for a few seconds.

In summary, single trapped ions are potentially suitable for the long-time frequency control of a light field. There are two pathways, either of which shows specific merits: 1. In a microscopic version of the double resonance method, a narrow dipole-forbidden line is excited and the resonance monitored by the rate of the "off" intervals in the simultaneously excited resonance fluorescence. 2. An E2 line is excited, in a Raman-type interaction, by two light fields, and the "dark" resonance is made to control the light frequency.

Schemes for the frequency control of light by a single ion are best suited for long-time calibration of an optical frequency, and for the identification of systematic shifts of the signal resonance.

REFERENCES

1. Toschek P E , Atomic particles in traps, *New Trends in Atomic Physics*, Grynberg G and Stora R (eds.) 1984, *Les Houches Session XXXVIII*, North Holland, Amsterdam, 381 p.

2. Bollinger J J , Prestage J D, Itano W M, and Wineland D J 1985, *Phys.Rev.Lett.* 54, 1000.

3. Jardino M, Desaintfuscien M, Barillet R, Viennet J, Petit P, and Audoin C 1981, *Appl.Phys.* 24, 107. – Cutler L S, Flory C A, Giffard R P, and McGuire M D 1986, *Appl.Phys.* B 39, 251.

4. Casdorff R, Enders V, Blatt R, Neuhauser W, and Toschek P E 1991, *Ann.Phys.* 48, 41.

5. Prestage J D, Dick G J, and Maleki L 1991, *IEEE Trans. Instr. Meas.*, 40, 132.

6. Dehmelt H G 1975, *Bull.Am.Phys.Soc.* 20, 60.

7. Nagourney W, Sandberg J, and Dehmelt H G 1986, *Phys.Rev.Lett.* 56, 2797.

8. Sauter Th, Blatt R, Neuhauser W, and Toschek P E 1986, *Phys.Rev.Lett.* 57, 1696; *Opt.Communic.* 60, 287.

9. Diedrich F, Bergquist J C, Itano W M, and Wineland D J 1989, *Phys.Rev.Lett.* 62, 403.

10. Bergquist J C, Diedrich F, Itano W M, and Wineland D J, Hg^+ single ion spectroscopy, *Laser Spectroscopy IX*, Feld M S, Thomas J E, Mooradian A (eds.) 1989 Academic Press, Boston, 274 p.

11. Nagourney W, NanYu, and Dehmelt H 1990, *Opt. Communic.* 79, 176.

12. Hänsch T, Keil R, Schabert A, Schmelzer Ch, and Toschek P E 1969, *Z. Physik* 226, 293.

13. Hänsch Th and Toschek P E 1970, *Z. Physik* 236, 213.

14. Alzetta G, Gozzini A, Moi L, and Orriols G 1976, *Nuovo Cimento* 36 B, 5.

15. Toschek P E and Neuhauser W, Spectroscopy on localized cooled ions, *Atomic Physics 7*, Kleppner D, Pipkin F M, (eds.) 1981, Plenum, New York, 529 p.

16. Neuhauser W, Hohenstadt M, Toschek P E, and Dehmelt H G 1978, *Appl.Phys.* 17, 123 – *Phys.Rev.Lett.* 41, 233.

17. Siemers I, Schubert M, Blatt R, Neuhauser W, and Toschek P E 1992, *Europhysics Lett.* 18, 139.

18. Schneider R and Werth G 1979, *Z. Physik* A 293, 103.

19. Wineland D J, Itano W M, Bergquist J C, Bollinger J J, Diedrich F and Gilbert S L, High accuracy spectroscopy of stored ions, *Frequency Standards and Metrology*, A. De Marchi, (ed.) 1989, Springer-Verlag, Berlin, 71 p.

20. Hollberg M and Hall J L 1984, *Phys.Rev.Lett.* 53, 230.

21. Blatt R, Casdorff R, Enders V, Neuhauser W, and Toschek P E 1990, New Frequency Standards Based on Yb^+, *Frequency Standards and Metrology, loc.cit*, 306 p.

Session 9A

Quartz: Space Oscillators

Chairman: G. Busca

QUARTZ OSCILLATORS, FREQUENCY SYNTHESIZERS, AND TIMING SYSTEMS FOR GEODESY AND RADIO SCIENCE APPLICATIONS

J. J. Suter, J. M. Cloeren, J. R. Norton, and M. J. Reinhart

The Johns Hopkins University Applied Physics Laboratory
Laurel, Md. 20723-6099 USA

ABSTRACT

The Johns Hopkins University Applied Physics Laboratory (APL) has developed a series of high-quality frequency standards including quartz-crystal oscillators, frequency synthesizers, and hydrogen masers. The demand for the increased performance of these standards provided the impetus for improvements to these instruments. This paper reports on recent advances made in these areas and the performance of spaceflight BVA oscillators and frequency synthesizers developed for the Topex–Poseidon Mission. As a member of the Bureau International des Poids et Mesures (BIPM) network, APL's Time and Frequency Standards Laboratory contributes data for the dissemination of International Atomic Time. Data on the performance of the APL atomic clocks in the BIPM ensemble will, therefore, also be presented. The ongoing development of quartz-crystal oscillators and synthesizers and APL's participation in the BIPM network address the needs of several U.S. and European geodesy and radio science space programs.

Key words: oscillator, radiation susceptibility, hydrogen maser, environmental effects.

1. TOPEX FREQUENCY REFERENCE UNIT

Recently, APL developed a spaceflight frequency reference unit for the Topex–Poseidon Mission. This system incorporates redundant 5-MHz BVA quartz oscillators, 4.096- and 19.056391-MHz frequency synthesizers, and a frequency distribution unit. Signals from this frequency reference unit support the Topex radar altimeter, a Global Positioning System (GPS) experiment, and the Tracking and Data Relay Satellite (TDRS) Transponder. As shown in Figure 1, the BVA 5-MHz, 3rd overtone, stress-compensation (SC) quartz oscillators, developed for the Topex–Poseidon Mission, demonstrated Allan variances of less than one part in 10^{-13} for 10-, 100-, and 1000-s sampling times (Ref. 1). Twenty-four-hour aging rates for these oscillators varied between 2×10^{-12} and 4×10^{-11}. The stability of these oscillators showed reduced sensitivities to magnetic field variations and ionizing radiation compared with conventional oscillators using conventional resonators. Single-side band-phase noise performance of this type of BVA oscillator is typically –131 dBc/Hz at a 1-Hz offset, descending to a noise floor of –158 dBc/Hz at a 100-kHz offset (Fig. 2). Extensive radiation testing showed a radiation-induced frequency change of 1×10^{-11} per rad (Si) of accumulated 1.25-MeV cobalt 60 radiation (Ref. 1). These frequency changes have been shown to be smaller for BVA resonators than for conventional resonators using electrodes plated directly onto the quartz (Refs. 2 and 3). The 5 gm/cm^2 of equivalent aluminum shielding presents an optimum amount of shielding material for radiation levels encountered in the 1334-km, 63° inclination Topex orbit (Ref. 4). Typical radiation exposure intervals for the Topex satellite consist of a 42-min exposure and a 70-min annealing period, where the 42-min interval represents the accumulation of radiation in the South Atlantic Anomaly. The dose rate over this period is about 0.6 rad (Si). Extensive orbit simulation tests have been conducted, consisting of exposures at 0.6 rad (Si) over 42 min (0.014 rad (Si)/min), followed by a 70-min annealing interval. Figure 3 shows the relative frequency shift for two BVA–SC resonators. Note that the beginning and end of the 42-min exposure cycles are clearly visible. Preconditioning of the BVA resonators with 20 krad of 1.25-MeV photons reduced the frequency offset. This effect was also observed in conventional resonators after preconditioning (Refs. 2 and 3).

The two space-qualified frequency synthesizers for the Topex–Poseidon Mission generate a 4.096-MHz signal for the GPS experiment and a 19.056931-MHz signal for the TDRS Transponder (Ref. 5). As shown in Figure 4, the single-side band-phase noise of the 4.096-MHz signal is –110 dBc/Hz at a 1-Hz offset from the carrier and –147 dBc/Hz at a 1-kHz offset. The 19.056391-MHz output measured –95 dBc/Hz at a 1-Hz offset and –130 dBc/Hz at a 1-kHz offset. The frequency synthesizers were exposed to the simulated ionizing radiation environment representative of the Topex–Poseidon Mission. During two sequential exposures (42 min each) and annealing cycles (70 min each), the synthesizers maintained phase lock. The change in the 5-V phase-lock loop-control signal was of the order of a few millivolts.

Hence, the relatively severe Topex radiation environment, as depicted in Figure 5, had little effect on the required performance of the Topex Frequency Reference Units, oscillators, and frequency synthesizers. However, a precise understanding of the radiation susceptibility of quartz-crystal resonators and its relation to quartz materials properties has not yet matured. Further studies are under way to identify the mechanisms that cause electrodeless BVA resonators to be less sensitive to ionizing radiation than resonators with adhered electrodes.

2. HYDROGEN MASERS

The Applied Physics Laboratory, as part of the BIPM network, maintains a Time and Frequency Standards Laboratory consisting of four APL-designed hydrogen masers and two Hp cesium standards (Ref. 6). The performance of these clocks is reported periodically to BIPM together with that of fifty-five other collaborating laboratories. Figure 6 shows the clock comparison data between Universal Coordinate Time (UTC) and UTC (APL) between Mean Julian Date 47809 and 48619. Extensive statistical analysis at APL has established the existence of a correlation between hydrogen maser frequency offset and barometric pressure variations, which could broaden the understanding of the APL maser's long-term stability (Ref. 7).

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

Results on the barometric pressure–induced offsets in APL's masers are basically due to dimensional changes of the TE_{011} resonant cavity. We observed that these changes are independent if the masers are free running or autotuned by the spin-exchange method. The pressure-induced effects are nonlinear, and the relationship between pressure and frequency offset is of a reverse nature; that is, increasing changes in pressure introduce decreasing changes in frequency offsets.

Figure 7 presents an example of these barometric pressure–induced frequency offsets. The barometrically induced compression of the resonance cavity results in a change of the masers' output frequency. The APL maser also equally tracks both increases and decreases in barometric pressure; that is, the magnitude of the positive or negative change in frequency shift is equal for a fixed-size increase or decrease in pressure.

Recent experiments have concentrated on establishing an absolute measurement of this phenomenon. By this we mean to exclude possible barometrically induced offsets in the reference maser.

3. DISCUSSION

Environmental effects (e.g., pressure-induced changes) may indeed limit the stability of frequency standards. These effects are often difficult to describe analytically, simply because effects are often interrelated. That is, analytical expressions for environmentally induced frequency shifts (temperature, humidity, pressure, radiation) not only involve terms due to the effects separately, but coupled terms as well, leading to a combination of two or more effects in one frequency offset term. For example, high kinetic doses of flash X-ray and neutron radiation may cause resonators to heat up. Hence, two competing effects will introduce frequency shifts in the resonator. The total frequency shift is then a combination of the shift due to the accumulated radiation dose and the radiation-induced thermal changes. Separation of these effects is often not possible, for example, for high dose levels of kinetic radiation. However, for low-Earth–orbit radiation exposure, as previously described for the Topex Mission, the dose levels in question do not generate any substantial temperature increases in the crystal bank. Simply, the radiation is *not* absorbed by the quartz-crystal lattice in terms of thermal energy but is rather used for the creation of electron-hole pairs. This physical phenomenon is representative of the energy losses of proton radiation above 80 MeV as well as 1.25-MeV photons.

Further studies on the radiation susceptibility of quartz resonators and environmental influences on masers (and resonators) are warranted, given future requirements for frequency stability and time keeping as dictated by geodesy and radio science applications. These studies must address so-called coupled effects in order to understand their physical nature. For low-Earth–orbit radiation susceptibility of quartz resonators, one only needs to address radiation-induced offsets. For masers, however, the barometrically induced offsets can often involve humidity and temperature effects, resulting in offsets with up to three coupled terms (pressure, humidity, and temperature). Experiments are therefore difficult to execute in order to gain an understanding of the physical nature of those effects.

4. SUMMARY

The authors have presented information on recent advances made by the Time and Frequency Section of The Johns Hopkins University Applied Physics Laboratory in support of U.S. and European geodesy and radio science space programs. These efforts have significantly advanced the performance of various types of frequency and timing sources. In addition, data were presented on the performance of the APL clocks in the BIPM ensemble, and future studies into the causes of environmentally induced frequency changes, which presently limit the performance of frequency standards, were discussed.

The work presented in this paper demonstrates the continued dedication of APL's Time and Frequency Section to making meaningful contributions to the International Time and Frequency Community.

REFERENCES

1. Norton, J. R., "BVA-Type Quartz Oscillator for Spacecraft," in *Proc. Ann. Symp. Freq. Cntrl.*, pp. 426–403 (1991).

2. Suter, J. J., Bates, A. G., Cloeren, J. M., Norton, J. R., Schlueter, B., and Peier, U., "The Susceptibility of BVA-SC-Resonators to Proton Ionization Effects," in *Proc. Third European Frequency and Time Forum,* pp. 11–12 (1989).

3. Norton, J. R., Cloeren, J. M., and Suter, J. J., "Results from Gamma Ray and Proton Beam Radiation Testing on Quartz Resonators," *IEEE Trans. Nucl. Sci.* **NS-31**, 1230 (1984).

4. March, P. A., Horton, C. M., and Stassinopoulos, E. G., *TOPEX Radiation Environment,* JPL Report D-2116 (1985).

5. Reinhart, M. J., "A Space-Qualified Frequency Synthesizer," in *Proc. Ann. Symp. Freq. Cntrl,* pp. 442–446 (1991).

6. Lewandowski, W., Douglas, R. J., Klepczynski, W. J., Strange, W., Suter, J. J., and Weiss, M. A., "Positioning of GPS Antennas in Time-Keeping Laboratories of North America," in *Proc. Ann. Symp. Freq. Cntrl.*, pp. 218–224 (1989).

7. Dragonette, R. A., and Suter, J. J., "Barometric Pressure-Induced Frequency Offsets in Hydrogen Maser," in *Proc. Ann. Symp. Freq. Cntrl.*, pp. 586–590 (1991).

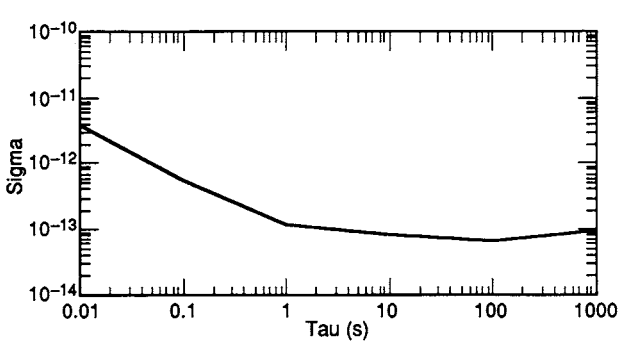


Figure 1. Allan variance of a 5-MHz BVA oscillator.

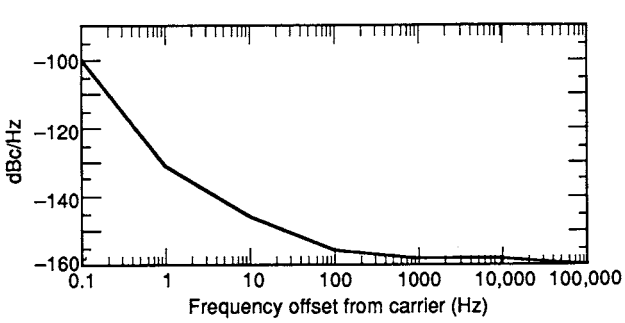


Figure 2. Single-side band-phase noise of a 5-MHz BVA oscillator.

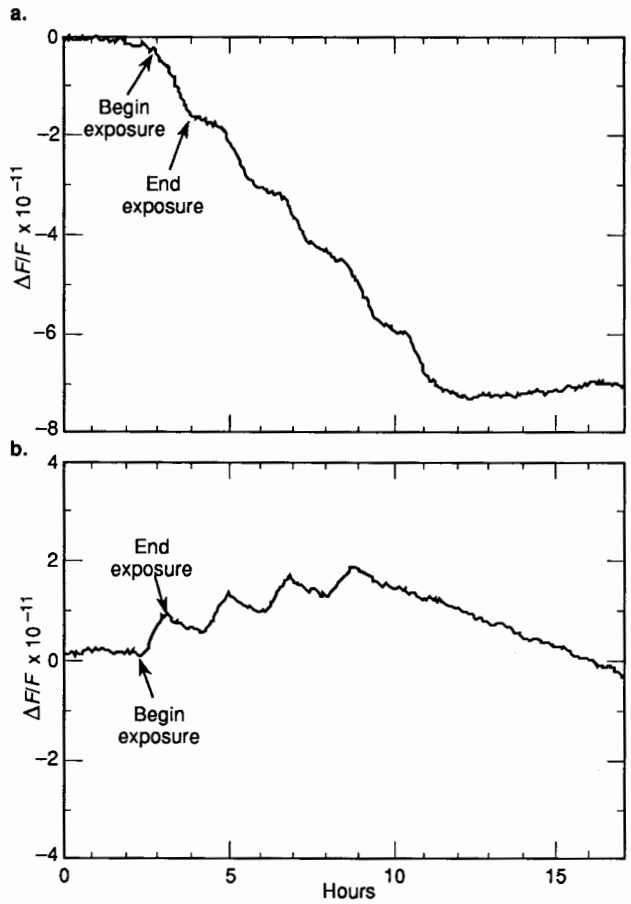


Figure 3. Frequency response of a BVA–SC quartz-crystal resonator (S/N 23) to 1.25 MeV of cobalt 60 radiation before (a) and after (b) preconditioning with 20-krad, 1.25-MeV cobalt 60 radiation. (F = frequency.)

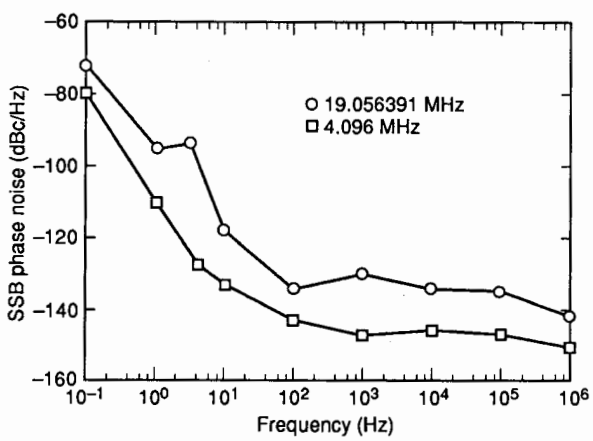


Figure 4. Single-side band-phase noise of 4.096- and 19.056391-MHz frequency synthesizers.

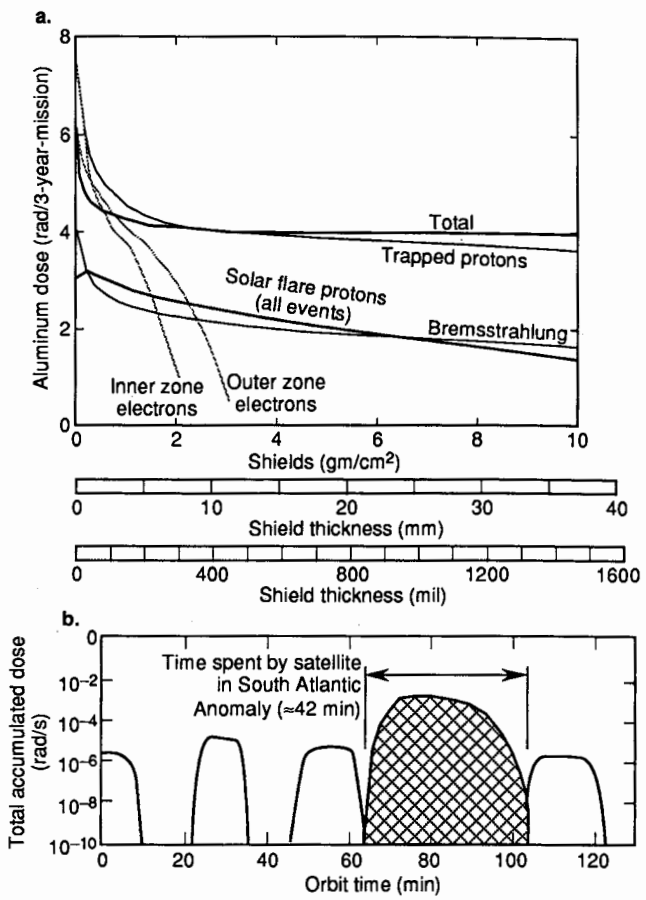


Figure 5. Topex radiation environment for a three-year mission (after March). (a) Total dose environment. (b) Environment as a function of position.

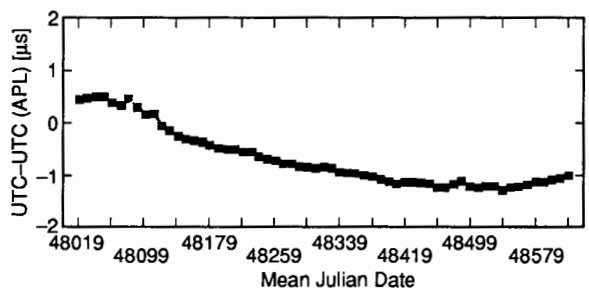


Figure 6. Clock comparison between UTC and UTC (APL) for the period (Mean Julian Date) 48019–48619.

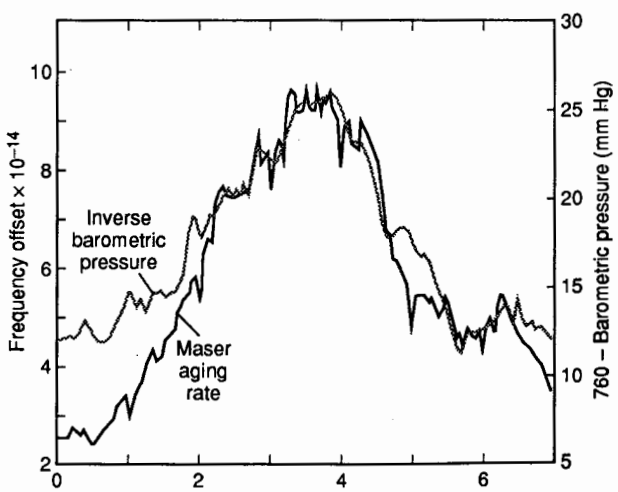


Figure 7. Barometric pressure–induced frequency offsets in an APL maser.

SPACE RADIATION SUSCEPTIBILITY OF ULTRA STABLE CRYSTAL OSCILLATORS

M. Brunet*, J.F. Dutrey*, J.B. Laporte*
J. Bourriau**
P. Calvel***

* Centre National d'Etudes Spatiales, CNES, Toulouse, France.
** ONERA - CERT - DERTS, Toulouse, France.
*** Alcatel Space, Toulouse, France.

ABSTRACT

This paper reports on the radiation susceptibility of several quartz crystal oscillators for space use (U.S.O.) with classical and BVA-type resonators, AT and SC cut. Susceptibility of each USO was measured at low dose rate, 1 and 0.1 rad per hour, using a Cobalt 60 source.

In parallel, a space radiation exposure analysis was made taking into account the radiation environment during the in flight data recording. A deposited dose calculation analysis has been conducted using an up to date 3 D Monte Carlo transport code. Results of aging rate per day are then compared to several in flight data.

1. INTRODUCTION

Piezoelectric quartz crystal resonators are widely used in satellite systems. The performances of these systems require quartz crystal resonators with superior stability on the order of 1.10^{-13} (τ = 1 to 10^3 seconds) for precise positioning systems, and 1.10^{-12} per day for telecommunication systems.

The apparent effect of radiation is frequency shifts. Previous experiments in U.S.A. (refs 1, 2, 3) have indicated that those phenomena are dependent on radiation dose rate, and particule flux and fluence. In our experiment, we have taken several USOs, from three different manufacturers, for space and ground applications (Location and Telecommunication). In the first step, ONERA - DERTS have conducted gamma ray exposure during six months; the frequency shift of each U S O was measured by CNES, using an automatic acquisition system. The second step, conducted by ALCATEL SPACE, consists of performing a space radiation exposure analysis.

The result, in form of quartz deposite dose versus time, allows prediction of the USO in flight response, in term of frequency shift per day. The last step was to compare these predictions with the drift of several USOs in orbit, obtained by three different location systems conducted by CNES.

2. GAMMA RAY RADIATION EXPERIMENT

2.1 Definition of tested USOs

We have tested ten different USOs. Table 1 gives several elements of each USO

Unit number	USO spec.	Manufacturer	Year code	Frequency in MHz	Model	Components	Quartz resonators
1	TC2	THOMSON	88	10	B.B.M.	Mil std	QAS, SC, Swept
2	and		89	10	E.M.	Mil std	QAS, SC, Swept
3	DORIS		89	10	E.M.	Mil std	QAS, SC, Swept
4	TC2	FEI	87	5	E.M.	Mil std	Classique, SC, swept
5	DORIS		88	5	E.M.	Mil std	Classique, SC, swept
6	ARGOS		74	4.7	B.B.M.	Mil std	Classique, AT, unswept
7			75	4.7	B.B.M	Mil std	Classique, AT, unswept
8			76	4.7	F.M.	Hil.Rel	Classique, AT, swept
9	ARGOS	OSA	76	5	B5400	Mil std	Classique, AT, unswept
10	DORIS		84	5	8600	Mil std	BVA, AT, unswept

B.B.M. : Bread Board Model
E.M. : Engineering Model
F.M. : Flight Model
Q.A.S. : Quartz Auto Suspendu
B.V.A. : Electrodeless quartz resonator

Table 1: Constitution of irradiated USOs

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

USOs number 1 to 5 were developed for Telecom 2 and DORIS French systems for space payloads. USO number 10 for DORIS ground station. USOs number 6 to 9 were developed for ARGOS system, number 8 for space and number 6, 7 and 9 for ground application.

Resonators of USO numbers 1, 2, 3, 9, and 10 were not pre-irradiated. Resonators of FEI USOs were pre-conditioned.

2.2 Radiation Exposure

2.2.1 Exposure in orbit

Radiation exposure in orbit in term of total absorbed dose and dose rate, at crystal level, depends on several factors such as orbit (LEO, GEO), inclination (70, 99°), duration in orbit (3 years, 10 years), epoch and also nature and thickness of shielding (spacecraft, equipment, USO).

As an example, Table 2 gives the cumulated total doses received, for three missions, by crystal FEI USO.

Telecom	ARGOS - NOAA DORIS - SPOT	DORIS - TOPEX
GEO 10 years 1600 rads	LEO, 800 km 3 years 600 rads	LEO, 1340 km 3 years 5200 rads

Table 2: Cumulated total doses received by crystal

Concerning dose rate, we can consider, for GEO orbit, a continuous dose rate, 1600 rads in 10 years, equivalent to 2.10^{-2} rad per hour.

In LEO orbit, satellites are periodically exposed to larger amounts of radiation doses over certain time intervals. For TOPEX, the dose rate as a function of time was determined by STASSINOPOULOS - JPL, and reported in Ref. 2, figure 4. The maximum dose rate, for 4 g/cm^2 aluminium shielding, is about 0.6 rad during 42 min every 130 min orbit period, when the spacecraft cross the South Atlantic Anomaly.

2.2.2 Gamma-ray ground test exposures

USO irradiation exposure was performed by ONERA - CERT - DERTS, using a Cobalt 60 source (activity 100 GBq), at room temperature and atmospheric pressure, according to the test program given in figure 1.

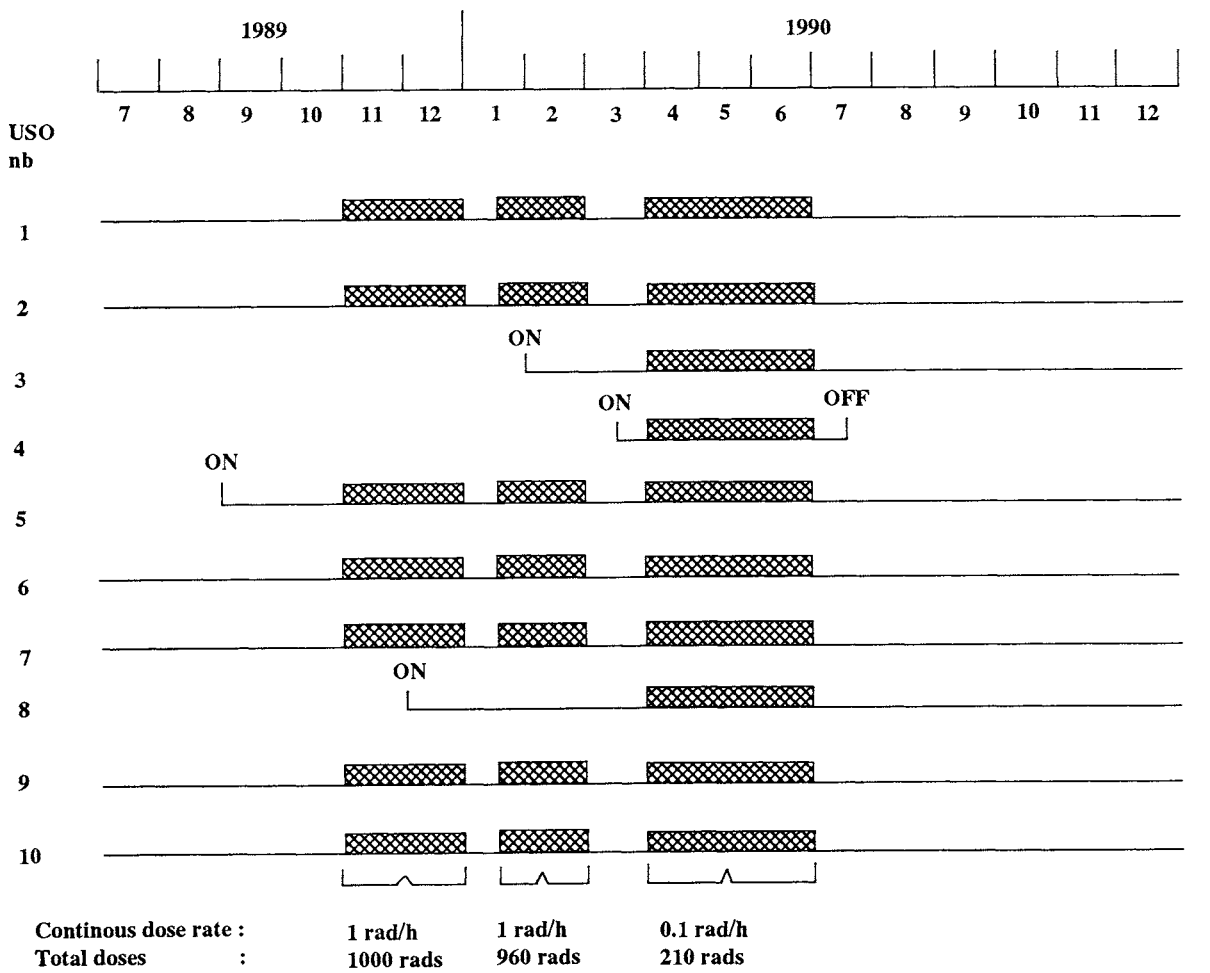


Figure 1: Gamma-ray radiation test program

The dose rate of 1 rad per hour was obtained at 2 meters far from the gamma-ray source, after deposited dose measurement, using graphite loaded LIF-7 thermoluminescent detectors. The dose rate of 0.1 rad per hour was obtained using 10 centimeters of lead shield.

Frequency of each USO was measured during several months, before, during and after exposures, using a CNES automatic data acquisition system (see Figure 2). This system is capable to measure the frequency shift ΔF/F of each USO. The beat frequency (1 MHz) is applied to a frequency error multiplier, and then to a counter that was interfaced to an HP computer system. Concurrently the beat frequency was converted by a digital to analog converter, whose output was observed by a 12 inputs recorder. At the same time, temperature and humidity of the room were measured and stored.

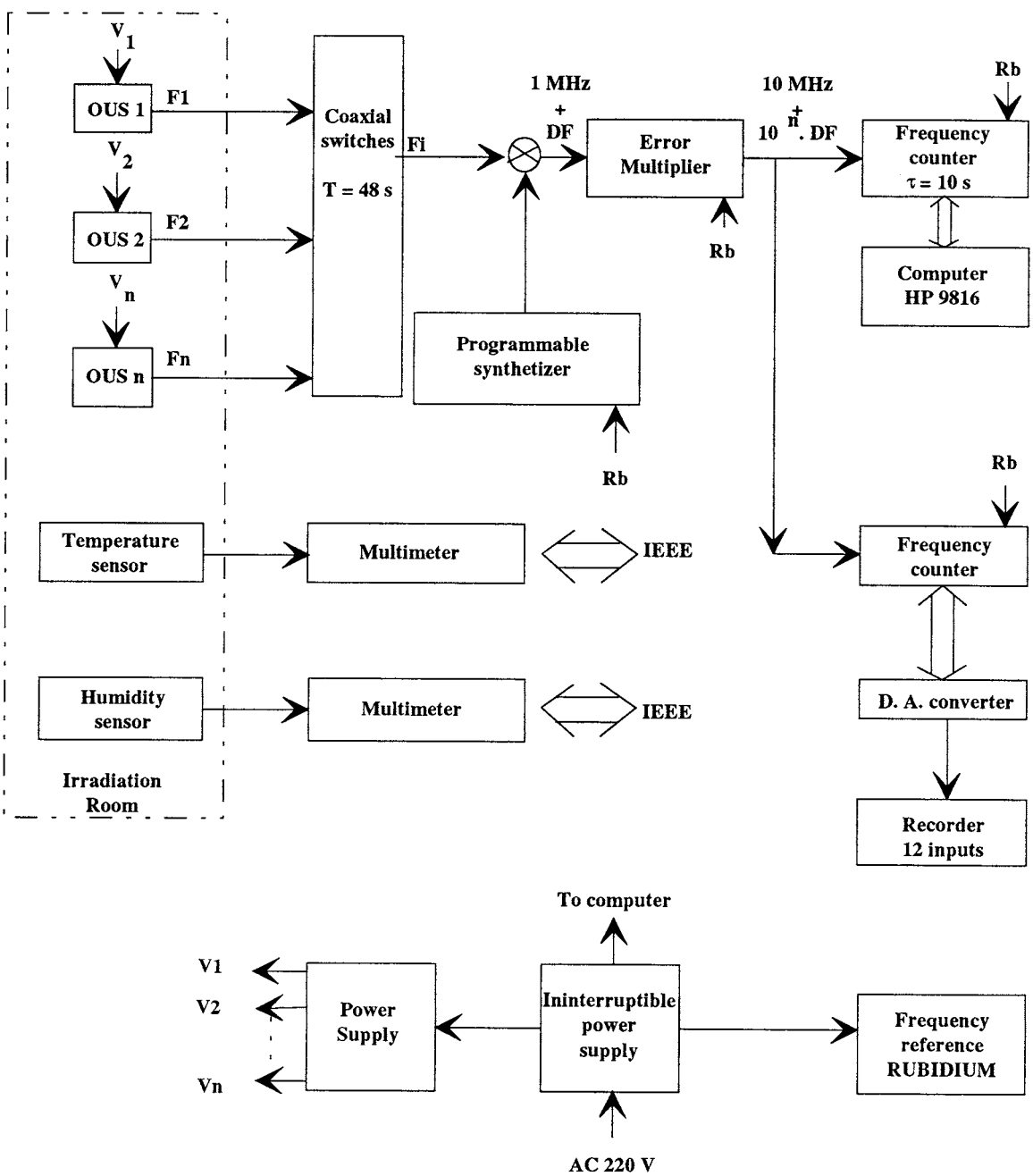


Figure 2: Data acquisition system for frequency shifts of quartz crystal oscillators

2.3 Results

2.3.1 Total frequency shift during gamma exposures

Table 3 gives total frequency shifts obtained on each USO, during the three exposures.

Manufac-turer	Unit Nb	Radiation 1 1 rad/h 1000 rad	Radiation 2 1 rad/h 940 rad	Radiation 3 0.1 rad/h 210 rad
THOMSON	1	$-1.2\ 10^{-9}$	$-1\ 10^{-9}$	$-1.3\ 10^{-9}$
	2	$-3.5\ 10^{-9}$	$-9\ 10^{-9}$	$-9\ 10^{-10}$
	3	-	-	$-9\ 10^{-10}$
FEI	4	-	-	$+2.3\ 10^{-9}$
	5	$+5.2\ 10^{-9}$	$+1.4\ 10^{-9}$	$+2.1\ 10^{-9}$
	6	$-4.6\ 10^{-8}$	$-3\ 10^{-8}$	$-3.5\ 10^{-9}$
	7	$-3.4\ 10^{-8}$	$-3\ 10^{-8}$	-
	8	-	-	$-8\ 10^{-9}$
OSA	9	$-3.1\ 10^{-8}$	$-1\ 10^{-8}$	$-5.8\ 10^{-10}$
	10	$+4\ 10^{-8}$	$+1.3\ 10^{-8}$	$-3.6\ 10^{-9}$

Table 3: Total frequency shift during continuous low dose rate gamma exposures

Discussions:

1 - USO numbers 1 to 5 can be compared, with their selected space components and SC crystals. They have a total frequency shift of several 10^{-9}. Thomson USOs with no preconditioning crystal have a negative drift, whereas FEI USOs have a positive drift.

2 - Units 6 to 8 have also a negative shift; units 6 and 7 indicate a very important drift (several 10^{-8}) due to non-swept crystal. However, unit 8 Flight Model AT cut swept crystal shows an important drift (-8.10^{-9} for 210 rad at 0.1 rad/hour).

3 - OSA USOs, with their nonswept crystals, shows an important shift, negative for B5400-type, positive for BVA-type.

Figure 3 gives, as an example, frequency drift of three USOs before, during and after tradition exposure.

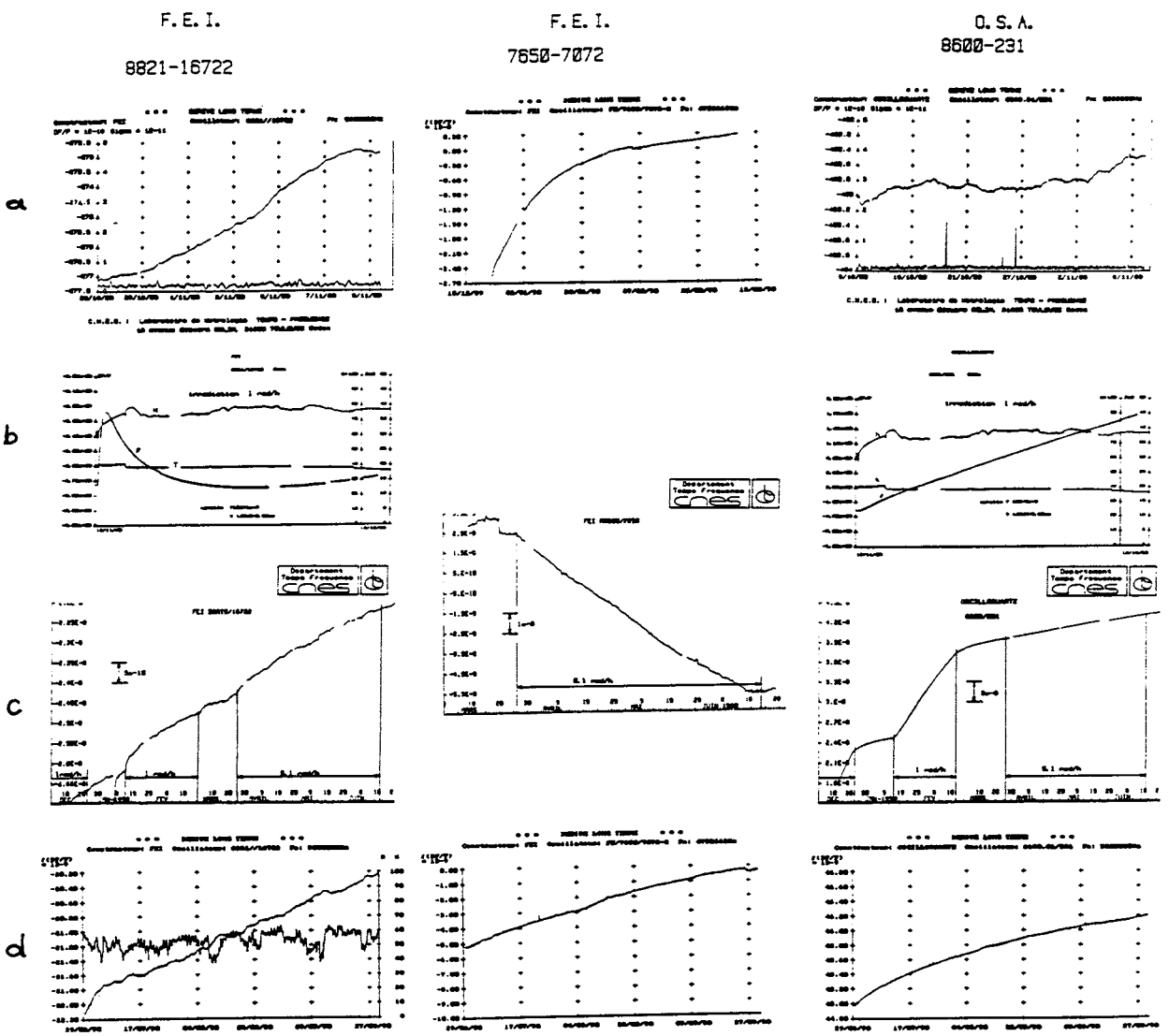
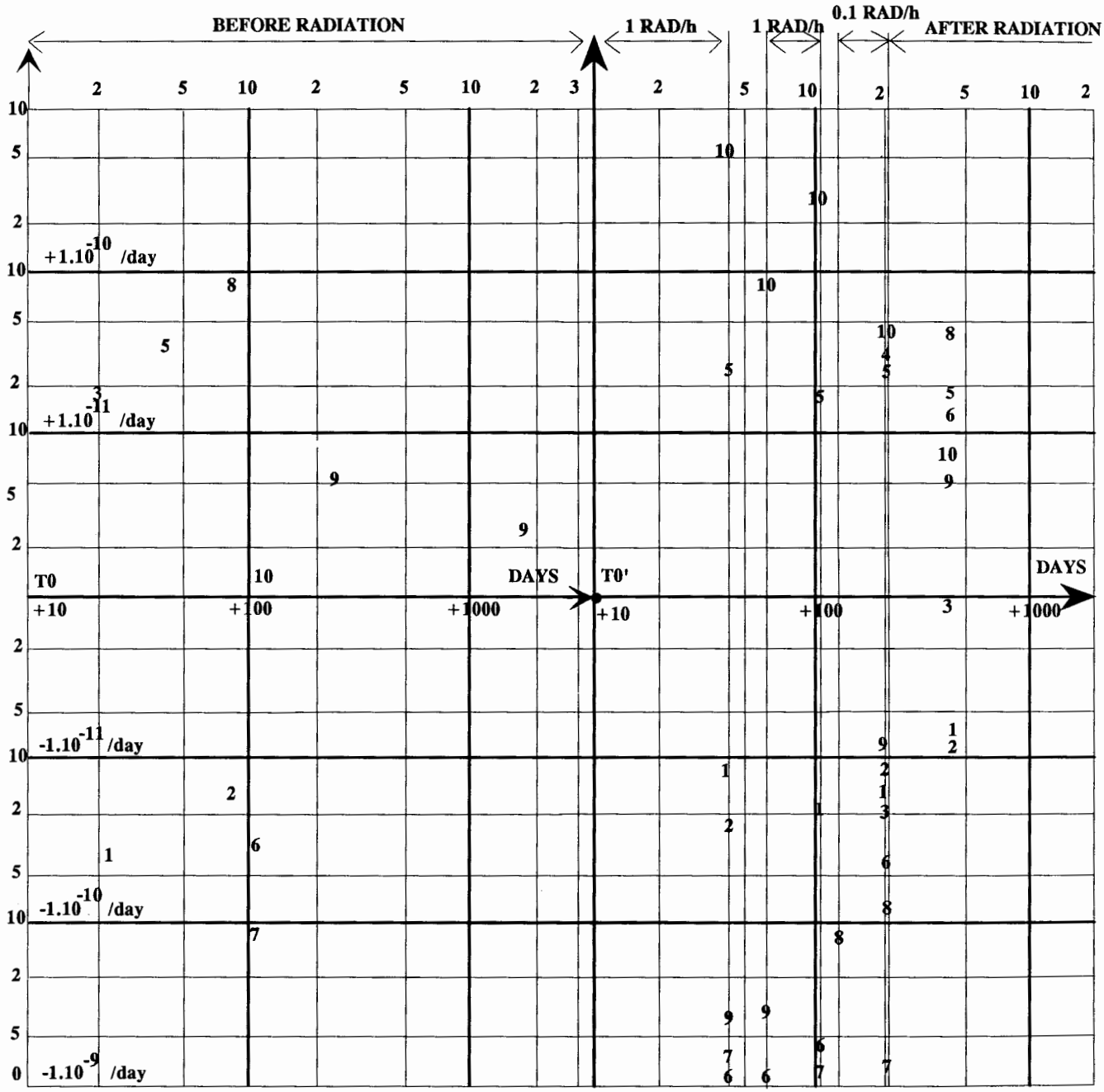


Figure 3: Example of frequency drift of three USOs, before (a), during (b, c), and after (d) gamma radiation tests.

2.3.2 Daily frequency drift during exposure

As mentioned before, we have measured the frequency of each USO, before, during and after radiation exposure. It is interesting to compare the frequency drift slope, expressed in $\Delta F/F \times day^{-1}$.

Figure 4 gives the mean frequency slope of each USO versus time.



T0 : Date of each USO power-on

T0': November 10, 1989

Figure 4: Mean frequency slope versus time

2.3.3 Radiation sensitivity determination

If we assume that ΔF/F is the sum of both aging and total dose, and that all other parameters are constant, the total slope ST is given by:

$$S T = S AG + S RAD \quad (1)$$

where S AG is the Aging Slope and

S RAD is the Radiation Slope

S RAD can be expressed by the following formula

$$S RAD = S' * DR$$

where S' is the radiation sensitivity in rad^{-1} (SiO_2) and DR is the daily deposited Dose Rate in rad/day

From these hypothesis, it is possible to calculate each term from experimental aging and radiation tests given in figure 4. Table 4 gives the radiation sensitivity S' of each USO in 10^{-11}. rad^{-1} at 1 and 0.1 rad/hour.

USO nb / DR	1	2	3	4	5	6	7	8	9	10
1 rad/h	N.M.	N.M.	-	-	N.M.	-2.3	-3.1	-	-0.8	+1.3
0.1 rad/h	N.M.	N.M.	N.M.	N.M.	N.M.	-1.1	-	-7.0	-0.2	+1.6

Table 4: Radiation Sensitivity S' in 10^{-11}/rad.

N.M. Non Measurable

Discussions:

1 - USOs number 1 to 3 have negative slopes,before, during and after the radiation tests. It is not possible, due to the sensitivity of an external parameter, to calculate accurately their radiation sensitivities. But in any case, we estimate S' less than 5.10^{-12} per rad.

2 - USOs number 4 and 5 have positive slopes, before, during and after the radiation tests. From our data ST is on the same order of SAG. Radiation Sensitivity is not measurable accurately, probably less than 1.10^{-12}/rad. USOs number 6 to 8 have a negative radiation sensitivity between -2 to -7.10^{-11} per rad.

3 - USO number 9 has a negative sensitivity (less than 1.10^{-11} per rad) whereas BVA AT USO number 10 has a positive sensitivity ($\simeq 1.10^{-11}$ per rad).

3. SPACE RADIATION ANALYSIS

In order to correlate total dose Gamma tests and in-flight data, it is necessary to get an accurate estimation of deposited doses on crystal in flight.

3.1 Predicted deposited doses

Due to the fact that we have determined the frequency drift of USOs in ARGOS/SARSAT/NOAA and DORIS/SPOT payloads, a simulation tool, the state of the art radiation shielding computer code "NOVICE", has been used.

A 3 D modeling of each spacecraft and a very detailed modeling of both FEI USO were performed including part case modelings (see reference 6).

The NOAA and SPOT 2 orbits are very similar, therefore, the same environment model was used for both evaluations. In order to get a best estimation of environment, models have been used taking into account the solar activity in the range of 1985 - 1991. Due to the lack of data concerning the solar maximum activity during 1989 - 1991, only the well known Anomalously Large Event Flare (August 1972) have been considered in these calculations. For the solar maximum period, calculations have been made in two cases: with and without 1 ALE (August 1972).

The 3 D forward and adjoint Monte Carlo simulation of electron and proton transport results are given in table 5.

	SOL min 85-88	SOL max (0 ALE) 89-91	SOL max (1 ALE) 89-91
Electrons	0	0	0
Bremmstrahlung	0.3	0.3	0.3
Trapped Protons	351	241	241
Solar Protons (1 ALE Aug. 72)	0	0	66
Total Dose	351	241	307

1 ALE = 1 Anomalously Large Event Flare (August 1972)

Table 5: Predicted deposited dose values in rad (SiO_2) per year for FEI USO

3.2 USO frequency drift in orbit

USO frequency drift in orbit has been obtained from three different space programs of CNES: ARGOS (Location and Data Collection of meteorological parameters by satellite), SARSAT (Search and Rescue by satellite), DORIS (Doppler precise positioning System). ARGOS and SARSAT are joint CNES, NASA and NOAA programs. French payloads are flown aboard TIROS spacecraft NOAA - A to N, since 1978, launched at intervals that maintain two satellites simultaneously in orbit. DORIS, a new dual frequency precise positioning system, was studied by CNES to locate, within 10 centimeters, spacecrafts and ground beacons. DORIS was on-board the French Earth Observation satellite SPOT 2, launched by ARIANE in February 1990. DORIS is on-board also on SPOT 3, SPOT 4, and on TOPEX - POSEIDON NASA spacecraft. Each system used one way Doppler technique, based on high frequency stability of on-board receiver USO, and ground transmitter automatic beacon USO. Using, at Toulouse CNES center, a cesium clock as reference in each Master Beacon, it is possible, for each pass of the satellite over Toulouse, to determine the mean frequency deviation of the on-board USO. For ARGOS and SARSAT, the frequency determination accuracy was estimated to $\pm 5.10^{-10}$ and $\pm 1.10^{-9}$ respectively.

For DORIS, due to very precise Doppler measurements and modeling technique, the accuracy was estimated to $\pm 4.10^{-12}$, and confirmed by an experiment at Observatoire de la Côte d'Azur (O.C.A.) in France, using an hydrogen maser.

Figure 5 gives the drift in orbit of ARGOS -NOAA 9 Flight Model USO, which are similar to FM USO number 8.

Figure 6 gives the drift in orbit of DORIS - SPOT 2 FM USO, which are similar to USO numbers 4 and 5.

From in orbit data, we have determined the Drift Total Slopes ST of several USOs expressed in 10^{-11} per day, function of time, given in figure 7.

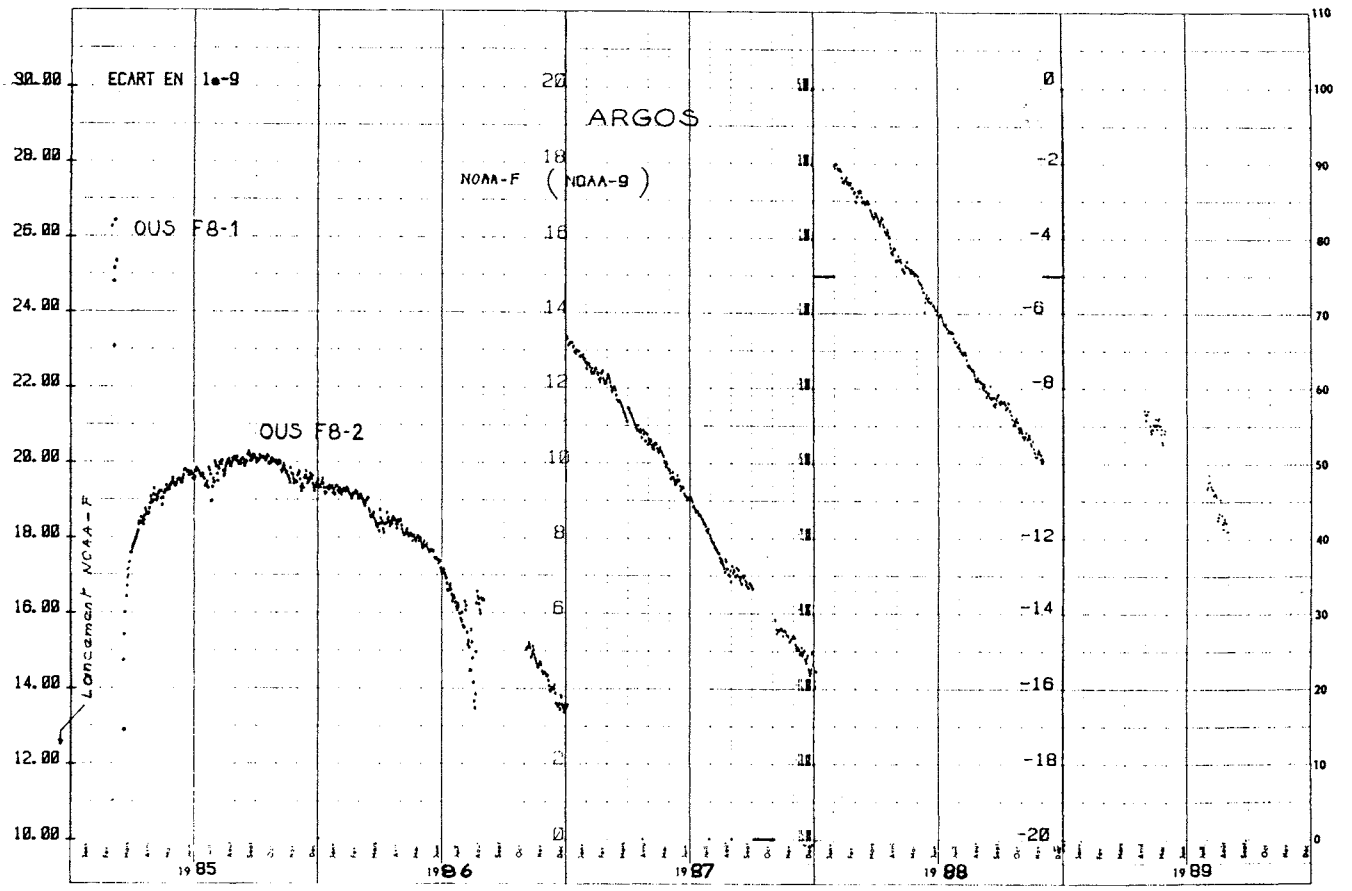


Figure 5: In orbit drift of ARGOS - NOAA 9 USO

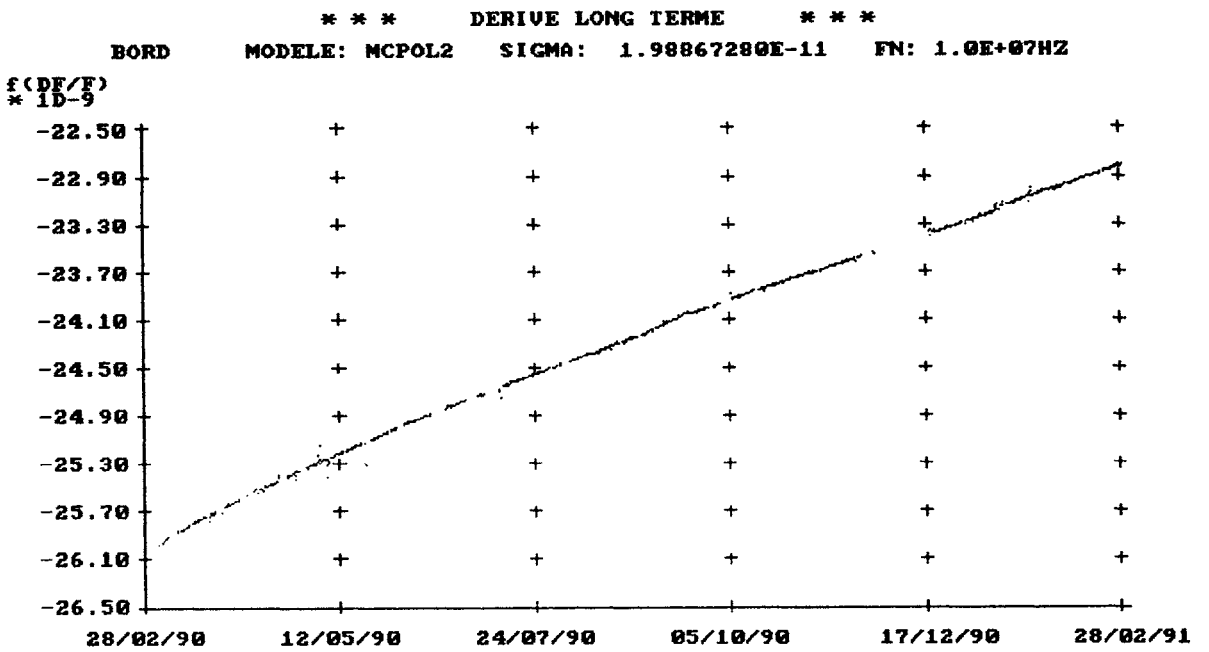


Figure 6: In orbit drift of DORIS - SPOT 2 USO

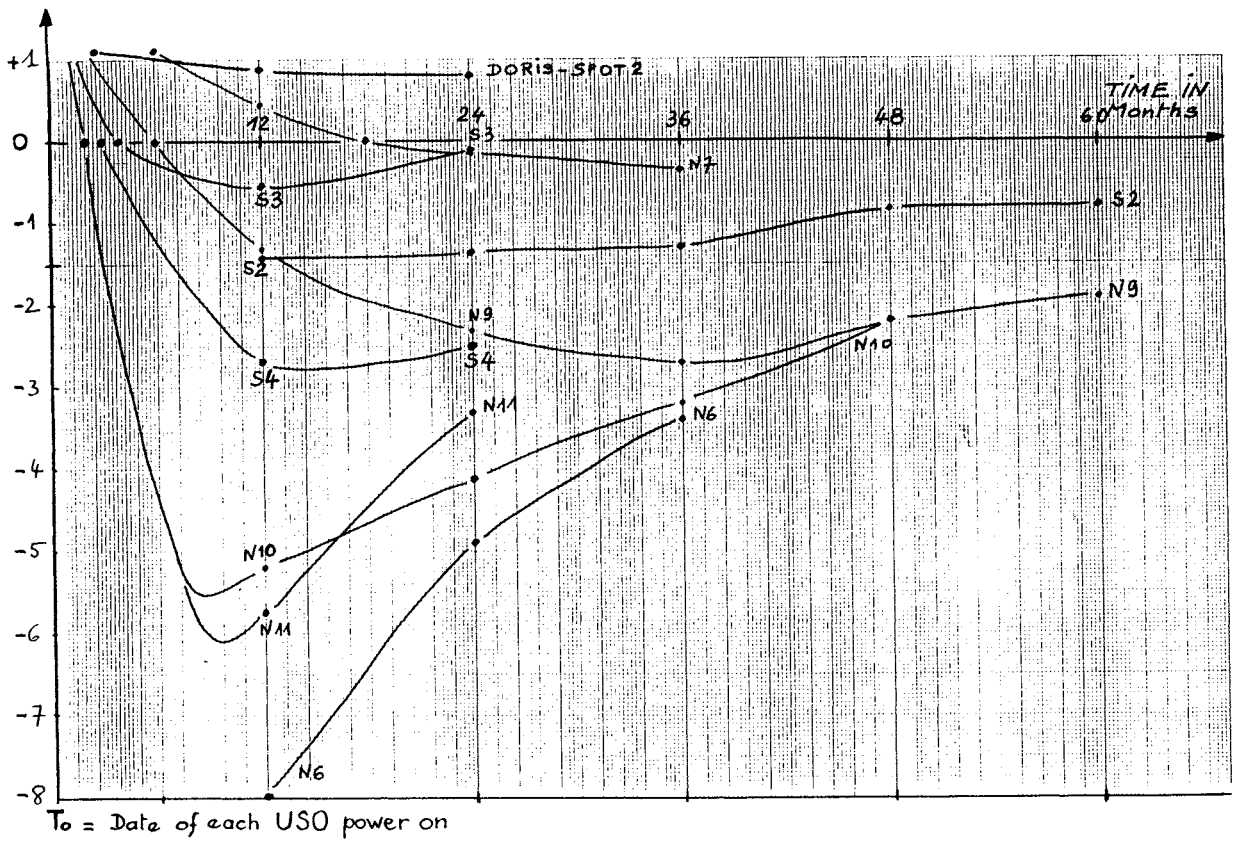


Figure 7: In orbit Total Slope ST of several USOs, in 10^{-11} per day versus time

ARGOS - SARSAT/NOAA USOs have a negative slopes in orbit, between 0 to -5.10^{-11} per day after 2 years of operation, whereas DORIS - SPOT 2 USO has a position slope, +1 to $0.8\ 10^{-11}$ per day.

3.3 Comparison with gamma radiation tests

In order to compare drift in orbit and gamma radiation tests, we have assumed:

1 - The AGING SLOPE, SAG, is positive, and a decreasing function of time. We have measured the drift of unit 8 during one year, and extrapoled until 60 months.

2 - The Radiation Slope, SRAD, is proportional to the daily deposited dose DR in rad/day, and to the Radiation Sensitivity S'.

3.3.1 ARGOS - SARSAT USO

From USO number 8 data and Table 5, we have calculated predicted Total Slope ST versus time, in Table 6.

		T0 + 12 m	T0 + 24 m	T0 + 36 m	T0 + 48 m	T0 + 60 m	Unit
USO nb 8	DR	1	1	1	0.84	0.84	rad/day
	S x DR	-7.0	-7.0	-7.0	-5.9	-5.9	10^{-11}/day
	SAG	+3	+1.5	+1.0	+0.5	+0.5	10^{-11}/day
	ST	-4	-5.5	-6.0	-5.4	-5.4	10^{-11}/day
ARGOS - NOAA 9		-1.3	-2.3	-2.7	-2.2	-1.9	10^{-11}/day
Ratio		0.3	0.4	0.45	0.4	0.35	
Year		1986	1987	1988	1989 (1 ALE)	1990 (1 ALE)	

Table 6: Predicted Total Slopes versus time and comparison with ARGOS - NOAA 9 USO

For each USO, we have determined predicted slopes, and compared with in orbit data by calculating the ratio R = (in orbit ST)/(pred ST) versus time in Table 7.

	T0 + 12m	T0 + 24 m	T0 + 36 m	T0 + 48 m	T0 + 60 m
ARGOS - N6	2.0	0.9	0.56	-	-
ARGOS - N7	-0.25	0.02	0.2	-	-
ARGOS - N9	0.3	0.4	0.45	0.4	0.35
ARGOS - N10	1.3	0.75	0.4	0.5	0.3
ARGOS - N11	2.0	0.75	-	-	-
SARSAT2 - N9	0.35	0.1	0.2	0.15	0.15
SARSAT3 - N10	0.13	0.02	-	-	-
SARSAT4 - N11	0.9	0.6	-	-	-

Table 7: Ratio R for several ARGOS-SARSAT USOs

3.3.2 DORIS - SPOT 2 USO

For DORIS - SPOT 2 USO, we have proceeded in the same manner, with positive SAG and S' = $+1.10^{-12}$/rad.

		T0 + 12 m	T0 + 24 m	Units
USO nb 5	DR	0.84	0.84	rad/day
	S.DR	0.084	0.084	10^{-11}/day
	SAG	+1.5	+1.0	10^{-11}/day
	ST pred	1.58	1.58	10^{-11}/day
DORIS SPOT 2	ST in flight	0.9	0.8	10^{-11}/day
Ratio R		0.56	0.50	

Table 8: Ratio R for DORIS-SPOT 2 USO

Comparisons between in flight data and predicted slopes show that drifts in orbit are lower than gamma radiation test. The Ratio R is between 0.1 to 1 after 24 months continuous operating, and between 0.2 to 0.5 after 48 months. Discrepancies are due to the fact that Aging Slope and Radiation Sensitivity are different between each unit.

4. CONCLUSIONS

We have tested, in irradiation, several Ultra Stable Oscillators (USOs) from three manufacturers. We have measured the frequency drift before during and after gamma radiation exposures, at very low dose rates (1 and 0.1 rad per hour).

We have confirmed that unswept crystals have a high radiation sensitivity, several parts of 10^{-11} per rad, positive or negative.

Space USOs using swept resonator and selected electronic components have a lower radiation sensitivity. Due to the fact that frequency slopes are the same with or without gamma radiation, it is not possible to have an accurate sensitivity determination. Our estimation is 1 to 5.10^{-12} per rad.

We have compared these results with drift in orbit, from several data, ARGOS - SARSAT and DORIS programs. Slopes in orbit are less important than predicted, by a factor of 2 to 3. The discrepancy observed comes from several uncertainties, such as aging slopes function of time, or absorbed dose calculations.

Flight data don't show any abrupt change at any time during 1989 period, whereas USOs have received, in October 1989, a typical solar flare of 80 rads.

Complementary investigations will be carried out in the next step of this work, using very precise DORIS - SPOT 3 and DORIS - TOPEX data.

REFERENCES

[1] J.R. NORTON, J.M. CLOEREN, J.J. SUTER, APL.
Results from Gamma Ray and Proton beam radiation testing of quartz resonators.
38th Frequency Control Symposium, 1983.

[2] J.J. SUTER, A.G. BATES, J.M. CLOEREN, APL.
The radiation susceptibility of BVA-AT resonators to simulated medium earth orbit radiation
2nd EFTF, 1988.

[3] J.J. SUTER, A.G. BATES, J.M. CLOEREN, J.R. NORTON, B. SCHLUETER, U. PEIER
The susceptibility of BVA-SC resonators to proton ionization effects.
3rd EFTF, 1989

[4] T. JORDAN: "NOVICE"
A Radiation Transport Shielding Code, Experimental and Mathematical Physics Consultants, Report # EMP.L 82.001, January 1982

[5] E.G. STASSINOPOULOS, J.M. BARTH, R.L. SMITH
METSAT CHARGED Particule Environment Study
NASA/GSFC report # x-600-87-11-August 1987.

[6] M. BRUNET, JB. LAPORTE, JF. DUTREY CNES
P. CALTEL ALCATEL ESPACE
J. BOURRIEAU ONERA-CERT-DERTS
Space Radiation susceptibility of ultra stable oscillators.
Ionizing Radiation workshop of M.S.F.C., NASA, 16-18 April 1991, Huntsville, Alabama, USA.

FREQUENCY GENERATION AND SYNTHESIS IN THE DRSS

AUTHORS : P. VAN LOOCK, J. DE VLIEGHERE

ALCATEL BELL TELEPHONE

Abstract

The Tunable Frequency Converter concept envisaged on board of the Data Relay System Satellites requires a synthesizer with state of the art phase noise performance that is locked to a common Ultrastable Oscillator.

Key words : Data Relay Satellite System, Synthesizer, Phase noise

1. DRSS Ka BAND COMMUNICATION LINKS [1]

1.1 Data Relay System

The Data Relay System (DRS) comprises three segments : ground segment, DRS satellites (DRSS) and user space terminals (UST).
The DRS supports forward communication links from ground via DRSS to user space terminals and return links from user space terminals via DRSS to ground. The forward links comprise the forward feeder links from ground to DRSS and the forward Inter-Orbit Links (IOL) from DRSS to the UST. The return links comprise the return IOL from UST's to DRSS and the return feeder links from DRSS to ground. The Ka-band communication links frequency plan is provided in fig.1.

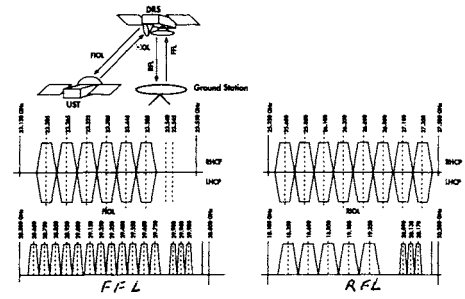


fig.1 : Ka band communication links frequency plan

1.2 Tunable Frequency Converter

The assigment of IOL channels to feeder link channels requires a high degree of flexibility and interconnection within the DRSS payload.
In fig.2, a traditional channel selection scheme using one fixed frequency conversion followed by a filter bank is compared to a scheme where a set of tunable converters is used followed by a filter at a single IF frequency.

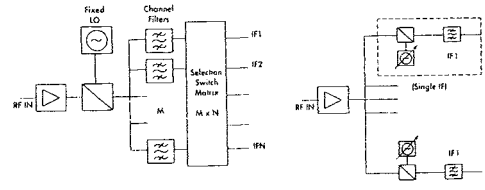


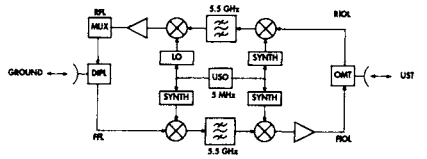
fig.2 : Channel selection concepts

In order to select this more advanced concept, the tunable synthesizers shall have an RF performance comparable to that of a fixed frequency local oscillator and they shall have a small mass, size and power consumption.

Since the DRSS shall be able to support signals for range rate measurements to and from user spacecraft, all frequency translations must be locked to one common reference. Fig.3 provides a simplified scheme of the DRSS ka band payload showing the different types of frequency converters.
In the frame of ESTEC contract Nr.8138/88/NL/PR the RIOL TFC is being designed at Alcatel Bell.

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

The aim is to provide an autonomous unit demonstrating that this advanced channel selection scheme is feasable. This functional autonomous TFC concept allows M/N redundancy schemes compared to the conventional 2/1 redundancy.



Type	RF Frequency (GHz)	LO Frequency (GHz)	Number of Frequency Steps	Frequency Multiplication Factor FLO/USO
FFL	28.5 - 30.0	23.0 - 24.5	300	4600 - 4900
RIOL	23.12 - 23.55	17.62 - 18.05	86	3524 - 3610
EIOL	25.25 - 27.5	19.75 - 22.0	450	3950 - 4400
EFL	18.1 - 20.2	12.6 - 14.7	Fixed	2520 - 2940

fig.3 : Simplified DRSS Ka band payload

2. FREQUENCY SYNTHESIZER

2.1 Concept trade-off

Electrically, the SSB phase noise density is the driving criterium in the selection of the synthesizer configuration.
The specified limit for the Ka band RIOL TFC is provided in fig. 4 .

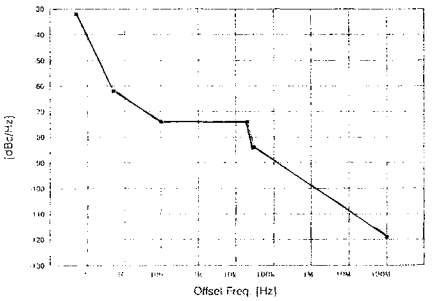
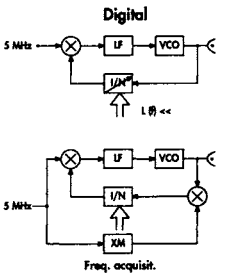


fig.4 : RIOL TFC Phase Noise Specification

In addition to this phase noise requirement,the synthesizer shall provide frequency flexibility in steps of 5 MHz over operating frequency band of 19.875 to 22 GHz.
This can be achieved in a number of ways:

. Digital concept. Frequency flexibility can be achieved in digital loops by varying the divider value (fig.5). N will be realised by a prescaler and a programmable dividerfeaturing inherent acquisition capability. The necessary use of GaAs prescalers with high division ratio lead to poor phase noise performance.
Converting the VCO frequency to a low IF frequency eliminates the need for high speed dividers and reduces the division ratio.

. Analogue concepts are based on frequency multipliers and samplers which are popular in high frequency applications (fig.6). A summing loop is introduced to limit the multiplication factor to the required application. Frequency acquisition is not inherent and a tunable discriminator for loop acquisition is required.



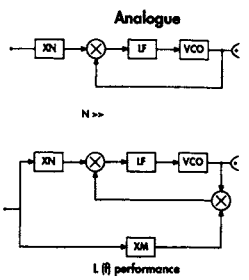


fig.5: Digital concept fig.6: Analogue concept

Superiour phase noise performance is achieved with this type of loops.
This can be seen from fig.7 where the TFC frequency sampler loop is compared with a feedback divider loop. For the feedback division ratio or sampler multiplication factor envisaged in this TFC, the sampler solution provides 10 dB improvement compared to a digital approach.

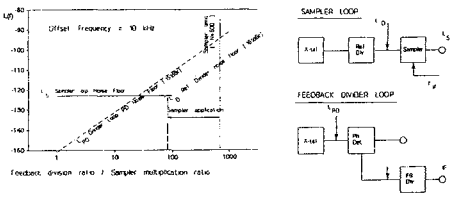


fig.7: Application Limit for a Frequency Sampler

Maximum frequency flexibility at Ka band with minimal phase noise penalty is achieved by a combination of both analogue and digital techniques (fig.8) providing respectively low phase noise performance and automatic frequency acquisition.

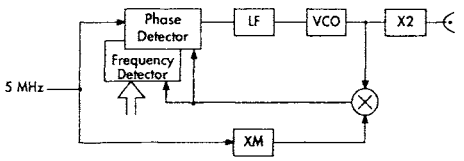


fig.8: Analogue phase/Digital frequency loop

2.2 Key Elements

. Sampled phase lock loop. The summing loop IF frequency (ranging from 275 -> 1400 MHz) is sampled with a repetion rate of 2.5 MHz thus providing a multiplication factor of 560.

. Voltage controled oscillator. Excellent phase noise performance is obtained with a bipolar VCO followed by a GaAs buffer amplifier. The measured performance over the 10% tuning range is provided in fig.9.

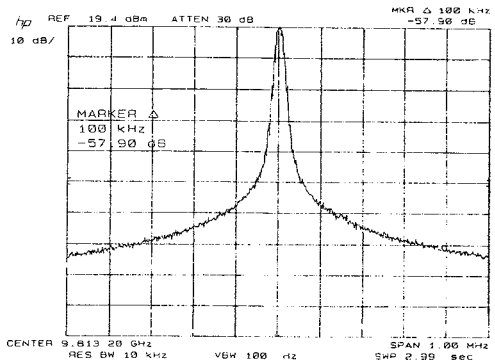


fig.9: VCO phase noise performance

3. REFERENCE GENERATOR

3.1 White noise filtering

The final reference frequency required in the summing loop for the RIOL TFC is 9.6 GHz.
The set-up of the frequency multiplier chain from 5 MHz to 9.6 GHz is driven by the phase noise requirement.
Phase noise contribution can be considered as originating from two kinds of sources. First components inherent phase noise originating from the 1/f component noise characteristics for which a selection is made at components and circuit level. The second source is multiplied up thermal noise.
The set-up followed to counteract multiplicative phase noise consists of cascading along the frequency multiplication chain a number of narrowband filters (fig.10).
A VCXO phase lock loop at 100 MHz is used as an active filter to reduce the multiplicative noise contribution of the USO in the range from 100 Hz to 100 KHz.
The phase noise performance of this VCXO is provided in fig.11.
Given the hardware complexity of this reference generator, commonality is pursued to enable its use for all TFC's, leading to reduced mass and power consumption at system level.

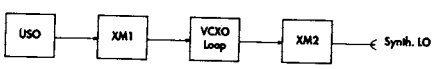


fig.10: Reference generator Implementation

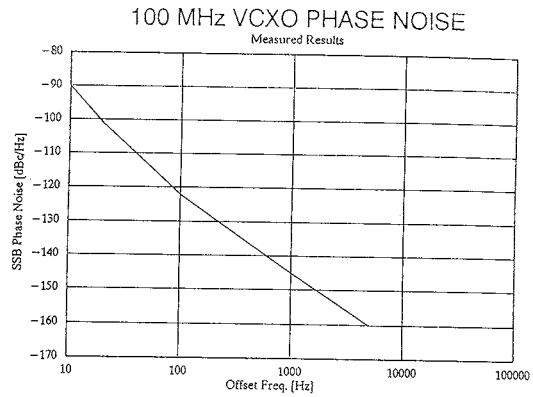


fig.11: 100 MHz VCXO measured phase noise

3.2 Commonality

Since the RIOL TFC requires the largest tuning range, the architecture for this TFC can be used for the other TFC types.
This concept is driven by the maximum operating frequency of the sampled phase lock loop of 1500 MHz and the highest possible frequency at which active filtering can be performed in the reference loop (set by the VCXO frequency of 100 MHz). The combination of these two factors lead to the choice of the 9.6 GHz (400 MHz x 24) reference frequency for the summing loop.
Table I provides the frequency plan for all TFC's when a common 400 MHz reference input is used.

TFC	RF Freq. (GHz)	LO Freq./2 (GHz)	IF Synth (MHz)	Synth LO (GHz)	M	Ref. (MHz)
RIOL	25.25 - 27.5	9.875 - 11	275 - 1400	9.6	24	400
FIOL	23.12 - 23.55	8.81 - 9.025	410 - 625	8.4	21	400
RFL	18.1 - 20.2	6.3 - 7.35	300 - 1350	6.0	15	400
FFL	28.5 - 30.0	11.5 - 12.25	300 - 1050	11.2	28	400

Table I : Reference frequency selection

4. HARDWARE REALISATION

A picture of the synthesizer/converter is provided in fig.12,while the common 400 MHz reference generator is shown in fig.13.

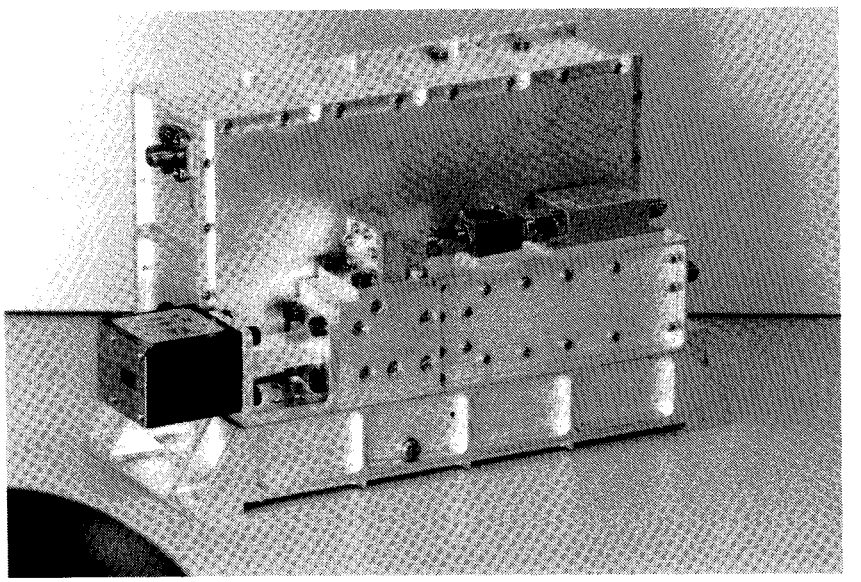


fig.12: RIOL TFC

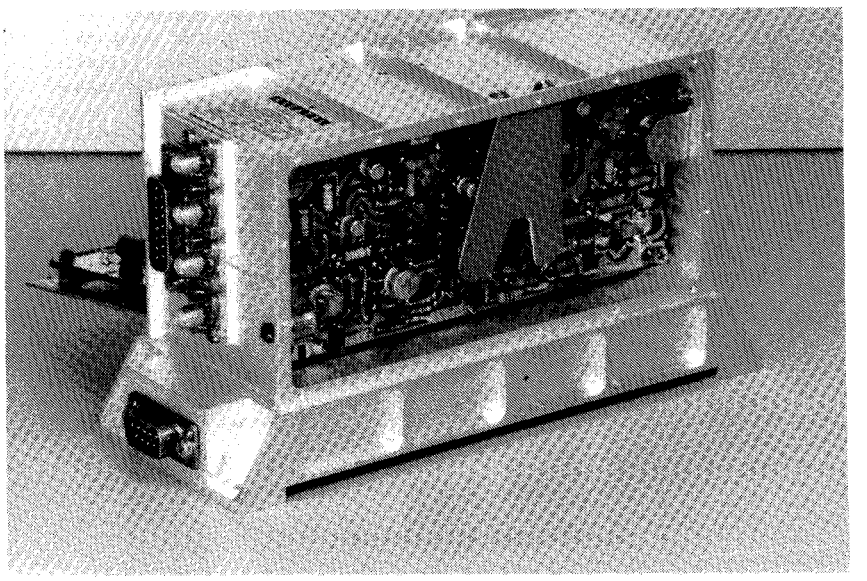


fig.13: Common 400 MHz reference generator

5. CONCLUSION

Under the ESTEC contract nr. 8138/88/NL/PR we have demonstrated the feasability of an autonomous Tunable Frequency Converter for the RIOL for the DRSS meeting the required phase noise performance.
This unit allows payload reconfigurability via frequency setting.
The concept is extendable to all TFC types on board of the DRSS.

6. REFERENCES

[1] J.Sandberg " Tunable Frequency Converter Statement of Work and Technical Specification CD/JS/525/CMM"

OSCILLATORS FOR MOBILE GPS RECEIVERS

JP AUBRY, P CANZIAN and V CANDELIER

C.E.P.E. Compagnie d'Electronique et Piézo-électricité
44 Avenue de la Glacière - 95100 ARGENTEUIL FRANCE

The GPS receivers, used either for time transfert or for localization, have to meet stringent specifications.

As a general rule, an oscillator specification is determined from :

* the overall target accuracy of the system
* the environmental conditions applied to the local receiver.

Most of the GPS oscillators we are dealing with in this paper are devoted to mobile receivers : naval or airborne.

The common specifications which make these oscillators "difficult" to design and to produce are size, G-sensitivity, power consumption and warm-up time.

The technologies used in ovenized oscillators are doubly rotated cut resonators -(SC cut)- in HC37 or HC 40 (for QAS 10 MHz design, ie for the lower S/G asked for), SMD assembly, monolithic thermal structure, PID driven loop.

A design optimization was conducted on the main concepts used in these oscillators :

* crystal resonators
* oven construction and thermal behavior
* mechanical structure
* electronics networks,

and leads to three generic OCXO types and one TCXO:

The design procedure and optimization are presented, along with typical statistical results regarding the resonator's performance.

1/ G SENSITIVITY

1/ A Induced side band modulation

Mechanical vibrations applied to an oscillator induce a sine displacement :

$$X = Xo \cos (2\pi\ fv\ t)$$

which leads to an maximum acceleration,A, given by :

$$A = (2\pi\ f_V)^2\ Xo \qquad \text{-in G units-.}$$

The instant frequency of an oscillator submitted to a sine vibration is then given by :

$$F(t) = Fo\ (1 + \Gamma\ A\ \cos(2\pi\ f_V\ t)\ ).$$

where A is the level of acceleration and Γ is the G-sensitivity of the resonator.

The output voltage of this oscillator is then given by :

$$V(t) = Vo \cos(\phi\ (t)\ )$$

where $\phi(t)$ is equal to 2π F(t).

TYPE	SIZE mm3	S/G	VIBRATIONS	DC POWER STABLE 25°C	DC POWER TURN ON	WARM-UP TIME
Naval H stable	65x47x66	$3\ 10^{-10}$	500 Hz	4 W	16 W	$5\ 10^{-8}$ 10mn/-25°C $5\ 10^{-9}$ 10 mn/+25°C
Naval Std	50x40x25	$8\ 10^{-10}$	2000 Hz	2,4 W	10 W	$2\ 10^{-7}$/10mn/ -10°C $5\ 10^{-8}$/4mn 25°C
Airborne	50x40x25	$1\ 10^{-9}$	2000 Hz	2,85 W	18 W	$2\ 10^{-7}$/5mn /-40°C $1\ 10^{-6}$/3 mn /-40°C
Airborne TCXO	40x20x15	$1.3\ 10^{-9}$	2000 Hz	0,1 W	0,1 W	1 s

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

The ouput signal becomes :

$$V(t) = Vo * [\, Jo(\beta) \cos (2\pi\ Fo\ t) \pm J1(\beta) \cos (2\pi\ (Fo \pm f_V)\ t) \pm Jn(\beta) \ldots .] ,$$

where Jo and J1 are the Bessel fonctions with argument $\beta = \dfrac{\Gamma\ A\ F_O}{f_V}$

An oscillator running at a frequency Fo submitted to a level of vibration of "A" G at a frequency f_V, using a resonator which sensitivity is Γ (in G^{-1}), exhibits an induced side band spectrum given by :

$$L^{(n)}(f_V) = 20\ Log\ (\ Jn(\beta) / Jo(\beta)\ ).$$

The first induced side band, at a frequency offset f_V from the carrier, rises a level (vs the main frequency level) given by :

$$L^{(1)}(f_V) = 20\ Log\ (\Gamma\ A\ Fo / 2\ f_V) \qquad \text{-in dBc-.}$$

1/B Resonators performances

The key factor in the relationships between resonator and oscillator performances under vibrations is the G sensitivity of the crystal resonator.

It is now widely known that G sensitivity is a complex non linear phenomenon in which the most critical items are :

* the symmetry of the mounting
* the acoustic wave symmetry

Various technological improvements and resonator designs have been suggested to provide G sensitivity better than some parts in 10^{-9} (the best of classical AT cut), down to a few part in 10^{-10}.

Some years ago BVA design and QAS design were introduced, and they can achieve (on an industrial and reproducible basis) G sensitivity better than 1 to 3 parts in 10^{-10}.

As we will see in the following lines, to-day GPS oscillators ask for dimensions no longer compatibles with HC40 crystal cans, which is the standard can for QAS and mini-BVA resonators. Most of the R'D work spent by CEPE in the past few years on the G topic was devoted to the optimization of low G-sensitivity resonators in HC 37 can.

To-day, G-sensitivity in the range of 1 to 5 10^{-10} G^{-1} is currently obtained at 10 Mhz, SC cut third overtone resonator, QAS design in a HC 40 can.

Figure 1a gives the G-sensitivity results of a batch of 60 HC 37 10MHz 3rd overtone resonators. 30% of the batch have a G sensitivity better than 5.10^{-10} G^{-1}.

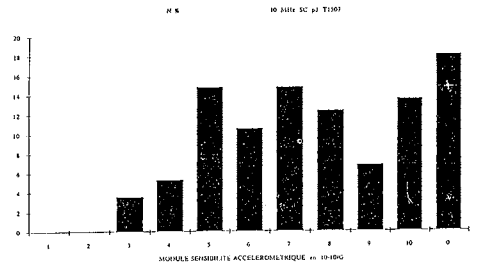
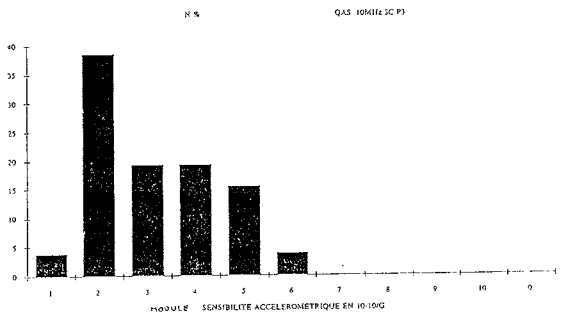


Figure 1b gives the G-sensitivity results of a batch of 60 HC 40 QAS 10MHz 3rd overtone resonators. 65% of the batch have a G sensitivity better than $3\ 10^{-10}$ G^{-1}.



2/ OVENIZED OSCILLATORS

2/A Design criteria

The basic electrical requirements of mobile GPS ovenized oscillator are :

* frequency stability in a wide temperature range
* fast warm-up
* low power consumption
* high mechanical environmental operation
* minimum size
* long term stability
* low fourrier frequency stability.

The basic design of an ovenized oscillator is given on figure 2. On this figure, OV stands for "oven", OL for "oven driving loop", OA is "output amplifier", VR is "voltage regulator", OSC is "oscillator" Xtal is the resonator.

On can built a "sensitivity matrix" between the design parameters of the oscillator and the performances required by a mobile GPS unit :

	XTAL	OSC	OA	OL	OV	VR	ME
F vs temp	***	**		**	**	***	
F vs Time		*		*		**	
DC Power	**				***		
Warm up	***			***	***		**
Vibrations	**				**		***
G Sensitivity	***						***
Size	***		**		***		***

where ME stands for "Mechanical assembly".

2/B Optimization

1) Thermomechanical structure

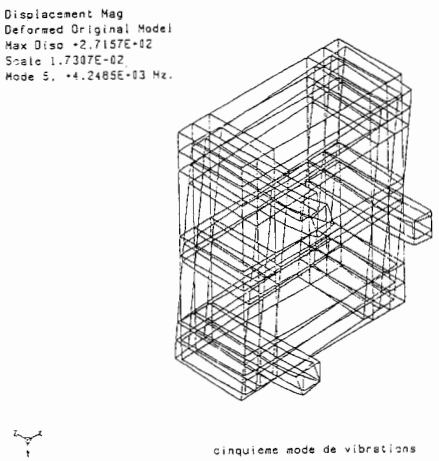
The criteria regarding size and thermomechanical behavior are so taugh that a finite element analysis and optimization is generaly required.

The basic rules are :

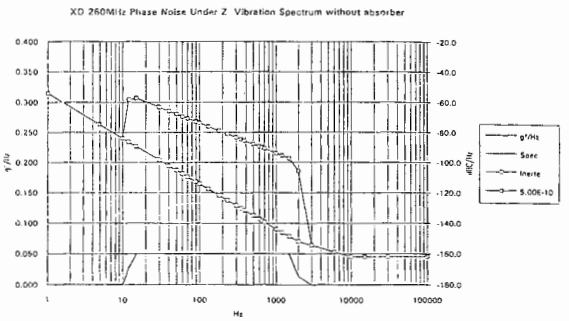
- avoid any mechanical resonance within the mechanical spectrum,
- good thermal coupling between resonator, oven and thermal sensor,

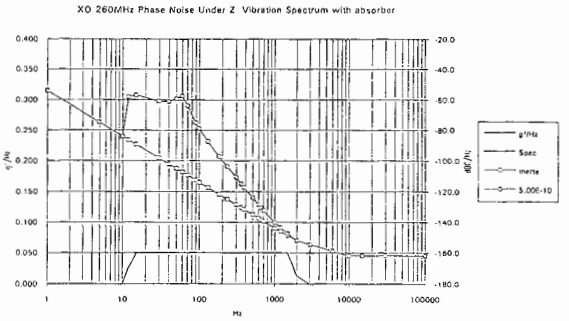
Once the mechanical behavior is determined, one can simulate what would be the phase noise spectrum under vibrations.

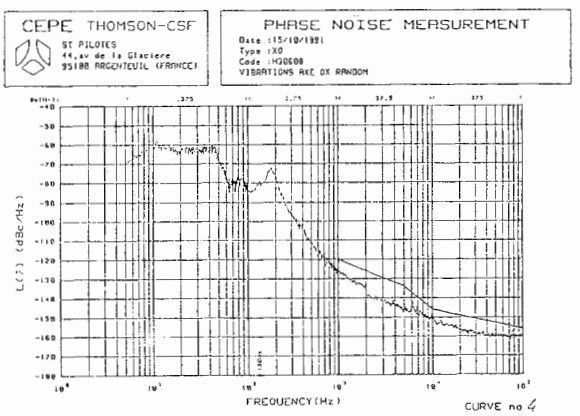
Figure 2 gives the mechanical analysis of an airborne ovenized oscillator. Figure 3 gives the theoretical phase noise spectrum and figure 3b gives the theoretical phase noise with absorbers and gives the experimental one.



cinquieme mode de vibrations







2) Crystal selection

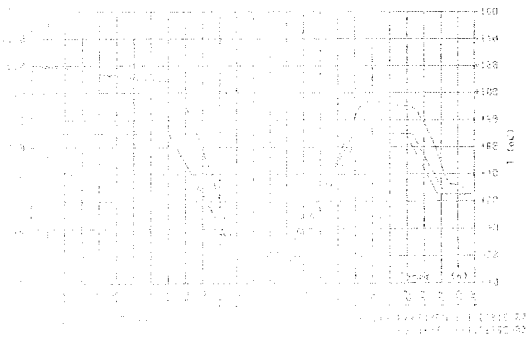
AT and SC cuts are widely used. It is well known that the thermal behavior of SC cut resonator, both static and dynamic, offers better ovenized performances , both from the static or dynamic behavior.

Due to fast warm-up requirements, we will generaly use SC cut resonators because the dynamic thermal behavior which expresses a relationship between instantaneous frequency and temperature variations through :

$$\frac{\delta f}{f}(t,T) = a_0 + a_1 T + a_2 T^2 + a_3 T^3 + @\,\frac{\delta T}{\delta t}$$

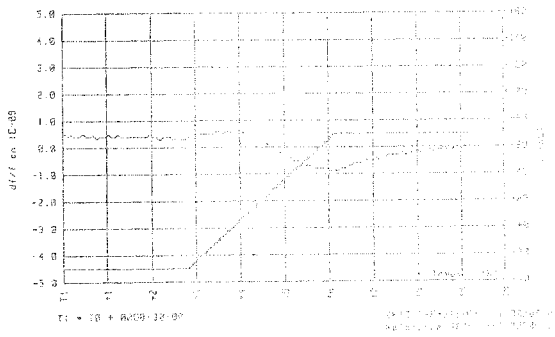
is much smaller with SC than AT cut (@ is ~ 0 in SC and @ ~3 10^{-7} in AT).

Furthermore, the static thermal behavior, seen as the radius of curvature of the f(T) curve around the Turn Over temperature point, is approximately 10 times better with an SC cut compared to an AT, which allows to improve the f(T) behavior of the oscillator.



Rayleigh Naval DTC

C.E.P.E.
ETUDES OSCILLATEURS
DATE 23 Aug 1988



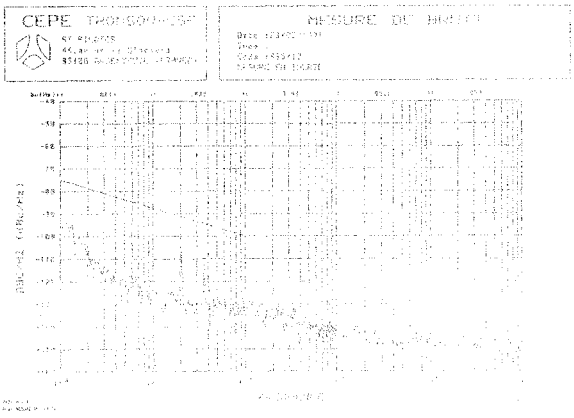
3 C/ Airborne oriented GPS

For airborne GPS applications we have developped an OCXO whose mains features are:

- 2 10^{-9} G^{-1} up to 2000 Hz (no mechanical resonances up to 2000 Hz)
- allowable vibrations spectrum : 0.2 g^2 / Hz up to 2000 Hz (30 G peak)
- 4 10^{-8} within -40, 70°C 1 10^{-8} within -55, 85°C fast warm-up : 10^{-8} - 3min / - 40°C 2 10^{-7} / 5 min / -40°C

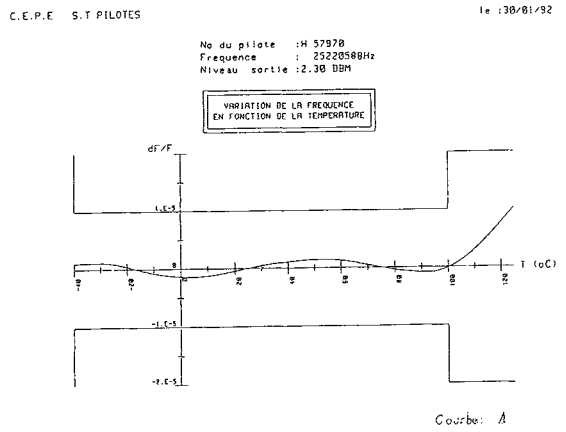
The resonator used in this oscillator is a SC cut 10 MHz 3rd overtone HC 37 can, and the oscillator dimensions are 50*40*25 mm^3

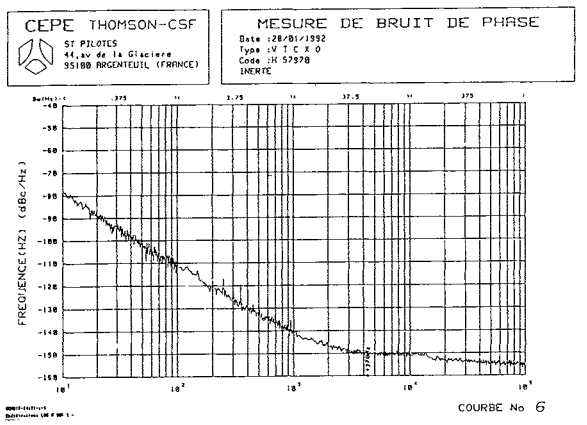
Figure 6a gives the phase without vibration, 6b is the vibrations level and to get the results of figure 6c. Figure 6d gives the warm-up behavior of the oscillator

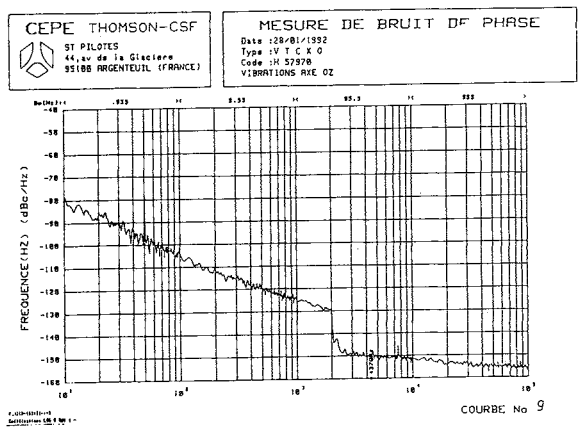


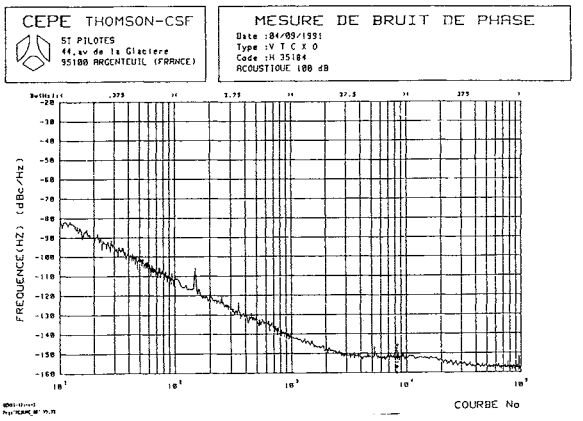
The resonator used in this oscillator is an ATcut HC 35 can, and the oscillator dimensions are 40*20*15 mm^3.

Figure 7a gives the f (T) behavior in a wide temperature range (-40, +120°C). Figure 7b gives the noise spectrum of this oscillator (without vibration), the figure 7c gives the noise spectrum with vibration. Figure 7d gives the noise behavior under a 100 dBA stress level f acoustic noise







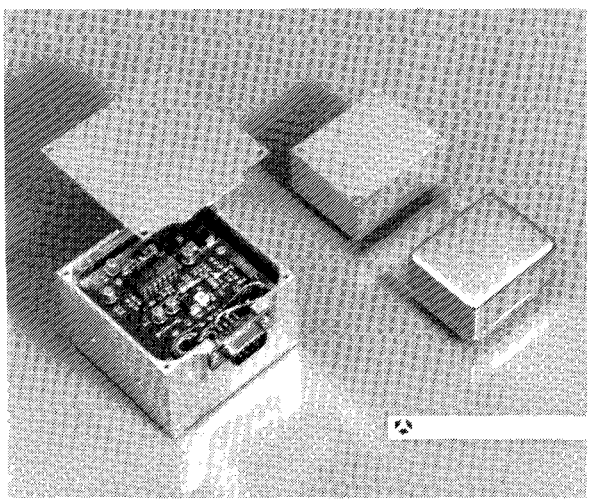


4/ CONCLUSION

The GPS oscillators we have shown are designed to meet the GPS system requirements, either for time transfer or localization, in different environment.

These oscillators focus either on the absolute accuracy, through low G-sensitivity and high stability for the naval ones, or on the warm-up, through the use of the airborne TCXO, which provides an accuracy good enough with low power and low size.

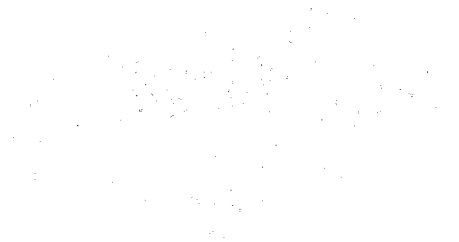
One of the key point is that most of the application require low oscillator size. That mean that the low G-sensitivity have now to be achieved with a low size resonator. Significant improvements have been carried on HC 37 can resonators which can exhibit now G-sensitivity better than 5 10^{-10} per G.



Session 9B

Atomic Frequency Standards: Electronics

Chairman: R. Vessot



MICROWAVE INTERROGATION FREQUENCY NOISE AND CLOCK FREQUENCY STABILITY: EXPERIMENTAL RESULTS

R.Barillet, V.Giordano, J.Viennet and C.Audoin

Laboratoire de l'Horloge Atomique - Unité propre du CNRS associée à l'Université Paris-Sud- Bat 221- 91405- Orsay Cedex- France

ABSTRACT

It has been shown theoretically that the frequency noise of the microwave interrogation may limit the short term frequency stability of passive atomic frequency standards. In order to investigate experimentally this effect, two identical microwave interrogation systems have been realized. The PSD of their frequency noise can be increased and measured. The validity of the theoretical model is confirmed by the experimental results. Measurement shows that the frequency stability of an experimental cesium beam tube operated at LHA is close to the limit associated with the use of a high quality 10 MHz VCXO. It will be necessary to take care of this effect in future passive frequency standards using optical pumping and/or cooled atoms.

Keywords: atomic clock, frequency stability, frequency noise

INTRODUCTION

Recent improvements in an experimental cesium beam tube operated at LHA have shown that the corresponding frequency stability cannot be deduced from the optical S/N ratio, but is limited by another phenomenon. This effect has been theoretically explained (1). It is related to the non linear behaviour of the cesium resonator, to the necessity of modulating the frequency of the interrogation signal and to the presence of frequency noise in the interrogation signal. The spectral components of this frequency noise at Fourier frequencies which are even multiples of the modulation frequency f_M are mixed with odd components of f_M which are present in the response of the resonator to the frequency modulation. This creates a disturbing signal at f_M which is synchronously detected, and added to the true error signal in the frequency servo loop. The corresponding limitation of the frequency stability, in terms of the Allan variance, is given by

$$\sigma_y^I(\tau) = \left[\sum_{n=1}^{\infty} \left(\frac{P_{2n-1} - P_{2n+1}}{P_1} \right)^2 S_y(2nf_M) \right]^{1/2} / 2\tau^{1/2}$$

where the factor P_{2n+1} is the Fourier coefficient of rank (2n+1) in the response of the resonator. It is usually a decreasing function of this rank. In order to investigate experimentally this limitation, we have realized two identical microwave interrogation systems DRO1 and DRO2. The P.S.D. of frequency fluctuations of DRO2 can be increased by means of an external noise generator, and measured by direct phase comparison with DRO1.

1- MICROWAVE GENERATORS DRO1 AND DRO2

To synthesize a microwave signal at 9192.6...MHz showing an adjustable phase noise- from the lowest possible one to 30 dB above-, we have used a dielectric resonator oscillator (DRO) which is phase locked to an ultra stable oscillator at 10MHz. The overall apparatus is shown in fig1.
After division by a factor of 27, the phase of the 270MHz VCO is compared to the phase reference at 10MHz and the operational filter H1 phaselocks the 270MHz signal to the 10MHz reference. The phase noise performance of the locked 270MHz signal is of prime importance. It is very dependent on the behaviour of the divider by 27, which is realized with ECL dividers and an ECL to TTL translator. All these circuits are sensitive to their supply noise. An ordinary supply provides an excessive level at the harmonic frequencies of the mains, whilst voltage regulators show an excessive level of wideband noise which is not compatible with our requirements. In our microwave generators, the supplies of the dividers have been filtered by passive RC cells having a 1s time constant. This ensures very low levels of noise and spurious signals. The bandwidth of the PLL is of about 100Hz to provide the phase locked 270MHz VCO with the best Sy(f).

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

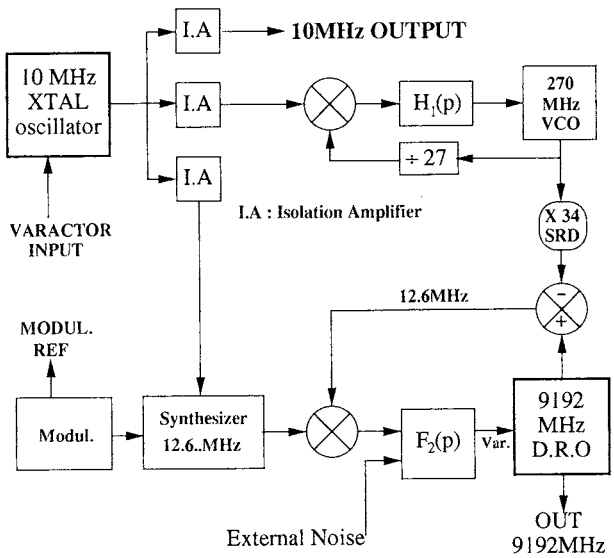


Figure 1: Microwave synthesis DRO1 or 2

A Step Recovery Diode module HP3300 generates harmonics of the 270MHz signal, which are mixed with the 9192MHz signal of the DRO. Among them, the 34th harmonics yields a beatnote at 12.6...MHz which is selected by a low-pass filter (f_c=150MHz) not mentionned in figure 1. The beatnote is mixed with an external 12.6...MHz signal generated by a commercial synthesizer whose frequency reference is the 10MHz VCXO signal, and the error signal of the PLL is obtained. The operational filter F2 phase locks the DRO signal to the 10MHz reference. An external noise can be added to the error signal to increase the phase noise (and therefore the frequency fluctuations) of the microwave signal. Typically the injected external noise is a broadband white noise filtered by a single pole low-pass filter showing a cut-off frequency of about 2kHz. The bandwidth of the PLL is larger than 10kHz so as : i) to transfer properly the frequency modulation present in the 12.6..MHz synthesizer signal and the phase modulation related to the additive noise, ii) to ensure the best Sy(f) for the locked DRO in the absence of additive noise, and iii) to preserve locking of the DRO when the additive noise has a high level.

The phase noise of the 9192.6...MHz signal can be kept very low or increased by as much as 30dB near the carrier (Δf< 1kHz). For excessive levels of the additive noise, the DRO signal cannot be locked. When the additive noise is not injected, the phase noise of the microwave signal is very close to the multiplied phase noise of the 10MHz VCXO for f<100Hz, of the 270MHz VCO for 100Hz < f < 10kHz, and equal to the phase noise of the free running DRO above 10kHz.

The main microwave output is used to interrogate the cesium atomic transition. An auxiliary 9192MHz output not mentionned on the figure makes it possible to measure the frequency of the microwave signal. The 10MHz available output is used for the frequency stability measurements.

Finally, the 10MHz VCXO can be locked to the cesium resonance by means of its varactor input and of an operational filter included in the microwave rack. The time constant of this servo is 0.4s .

2- PHASE NOISE MEASUREMENTS AT 9192MHZ

The measurement system is shown in figure 2. The 10MHz VCXO included in the microwave synthesis DRO1 is phase locked to the 9192MHz signal delivered by DRO2, with a PLL bandwidth f_c equal to 3Hz. Let $S_{D_1}(f)$ and $S_{D_2}(f)$ be the phase Power Spectral Densities (PSD) of the signals at 9192MHz, and K the gain of the mixer used as a phase detector. The PSD $S_V(f)$ of the voltage noise observed at the output of the mixer (2) may be written as

(1) $$S_V(f) = K\,[S_{D_1}(f)+S_{D_2}(f)]$$

for Fourier frequencies greater than f_c=3Hz.

When the external additive noise is not applied to DRO2, the phase noises of the two microwave signals are about statistically equivalent. Their contributions $S_{D_1}(f)$ and $S_{D_2}(f)$ are supposed to be equal and their common estimate is

(2) $$S_D(f) = S_V(f)/\,2K^2 \qquad (f >> f_c)\ .$$

When the external noise is applied to DRO2, the phase noise present in the DRO2 signal is much larger than the one of DRO1, and $S_{D2}(f)$ is given by

(3) $$S_{D2}(f) = S_V(f)\,/\,K^2 \qquad (f >> f_c)\ .$$

In our experimental work, Fourier frequencies of interest are larger than the modulation frequency f_M and the condition f>>fc is always fulfilled.

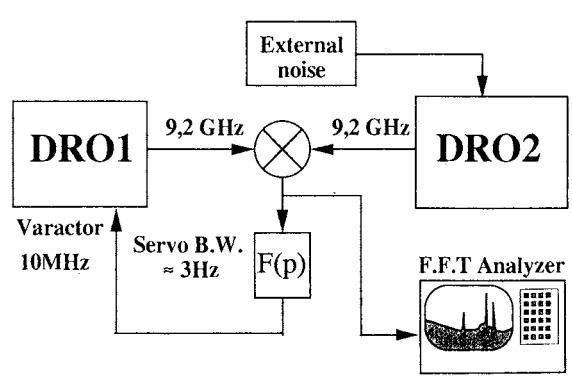


Figure 2:Phase Noise Measurement at 9.2GHz

3- EXPERIMENTAL RESULTS

The aim of this experimental work was to check quantitatively the limitation of the frequency stability related to the frequency noise of the microwave interrogation. The validity of the model is limited by the quasi-static approximation, which means that the modulation frequency f_M must be much smaller than the half-width of the cesium resonance. In our case this leads to $f_M << 250Hz$.

As indicated by the formula in the introduction, this limitation $\sigma^1_y(\tau)$ is due to the frequency noise at even multiples of the modulation frequency. Factors (P_{2n+1}/P_1) have been measured. It appears that only P_1 and P_3 must be considered and we have $P_3/P_1 = 0.15$. In those conditions, the effect of the frequency noise is mainly due to the second harmonics of the modulation frequency and its theoretical value for a white phase noise is then given by

$$(4) \qquad \sigma^1_y(\tau) = 0.45\ S_y(2f_M)^{1/2}/\tau^{1/2}$$

The overall experimental apparatus is shown in figure 3. The upper right part with DRO2 generates the useful microwave signal, with more or less phase noise according to the setting of the external noise level. The response of the cesium resonator enables us to lock the 10MHz VCXO included in DRO2, with an adjustable level of frequency fluctuations. The middle left part measures the phase noise present in the interrogation signal. The calculation of the PSD of phase is made according to (2) if there is no additive noise and according to (3) in the opposite case. After that, it is converted into PSD of fractional frequency fluctuations. The lower part measures the frequency stability, with a H-maser as a reference.

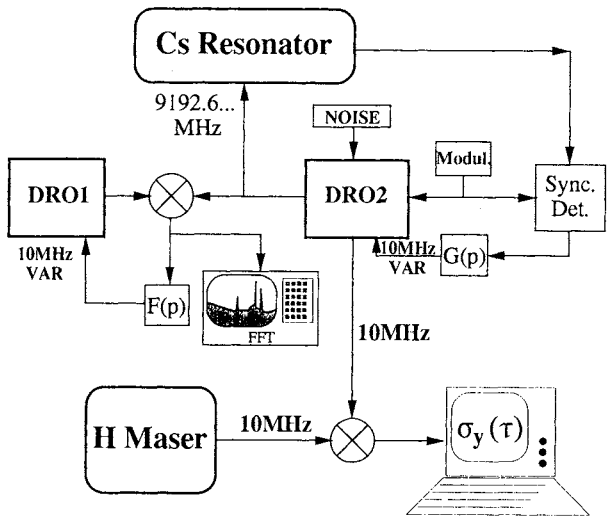
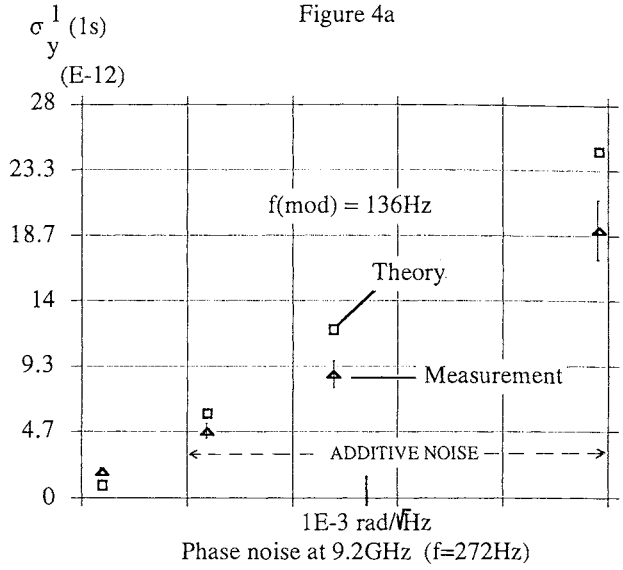


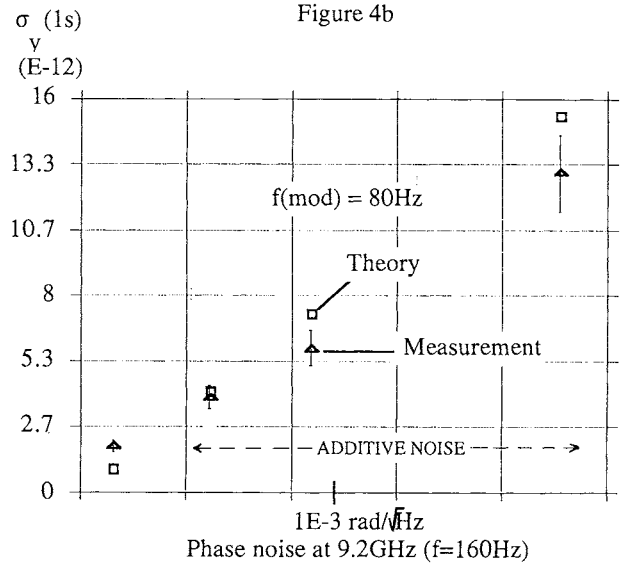
Figure 3: Experimental set-up

Figure 4a



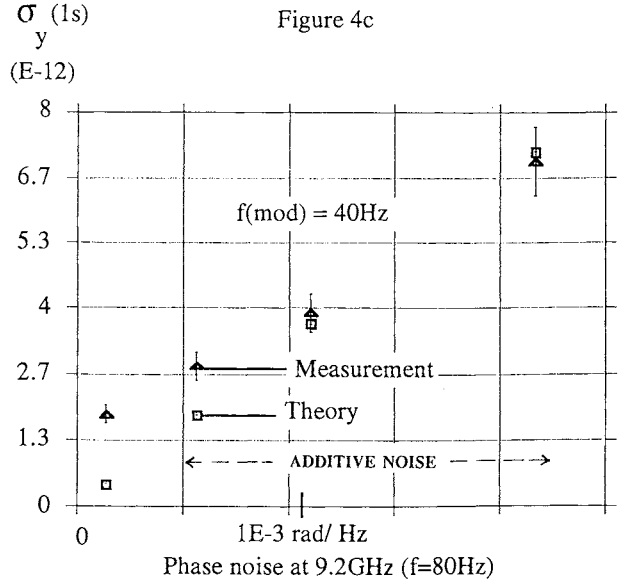
Phase noise at 9.2GHz (f=272Hz)

Figure 4b



Phase noise at 9.2GHz (f=160Hz)

Figure 4c



Phase noise at 9.2GHz (f=80Hz)

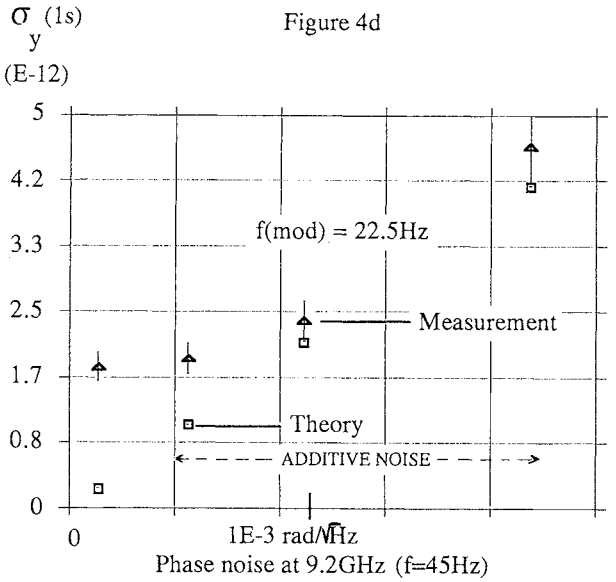


Figure 4d

Phase noise at 9.2GHz (f=45Hz)

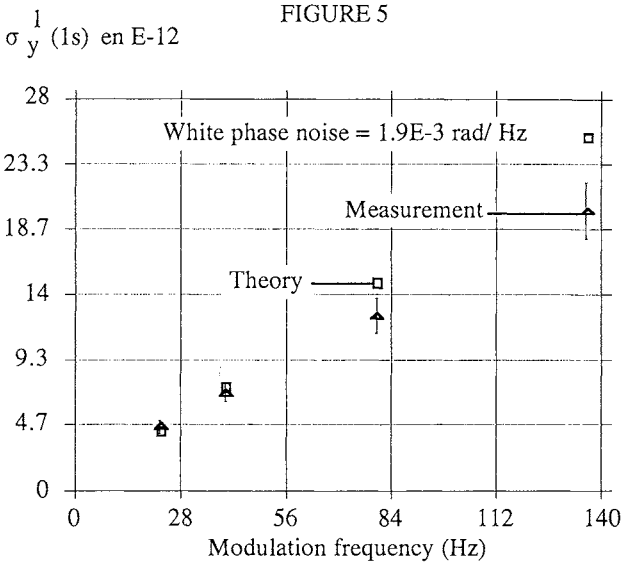


FIGURE 5

We have experienced four modulation frequencies - 136Hz, 80Hz, 40Hz and 22.5Hz- corresponding to figures 4a to 4d . For each modulation frequency, four levels of noise have been used, three with purposely added noise and one with the undisturbed noise of the microwave synthesizer. The frequency stability has been measured for a sampling time τ equal to the period of the beatnote Cs-MH which was 2.35s, then converted to τ=1s according to the $\tau^{-1/2}$ dependence .

In figure 4a, the modulation frequency is 136Hz and the measured effect is about 20% lower than the theoretical value, for large frequency noises. This difference may be due to the fact that the quasi static approximation is no longer valid for such a modulation frequency. When there is no additive noise, we reach a frequency stability of 1.92E-12. In that case, the contribution of the noise at $2f_M$ amounts to 1.E-12.

In figures 4b to 4d, one can see that the experimental points are close to the theoretical ones for high levels of additive noise. When there is no external noise and when f_M is low (22.5Hz), the measured frequency stability is limited by the atomic resonator and reaches 1.72E-12 (τ=1s). Indeed, in that case, the frequency stability calculated from the optical S/N ratio (17700) is equal to $1.55E\text{-}12\tau^{-1/2}$ and the limit related to the frequency noise is $2.6E\text{-}13\tau^{-1/2}$.

The effect of a white phase noise equal to 1.9E-3 rad/√Hz is shown in fig 5 for the previous modulation frequencies. The corresponding frequency noise is 1.9E-3f Hz/√Hz where f is the Fourier frequency. One sees that the limit associated with the effect of interest is proportional to the modulation frequency or to the frequency noise at $2f_M$. Practically, in high performance crystal oscillators, the frequency noise shows a minimum and the lowest spurious effect will be obtained when this minimum is located at $f=2f_M$: the choice of the proper VCXO depends on the modulation frequency and on the expected performance. With the present State of the Art, the lowest limits are obtained with 5 or 10MHz ultra stable oscillators for usual modulation frequencies: they are at the level of $1E\text{-}13\tau^{-1/2}$ for modulation frequencies about 3Hz, $3E\text{-}14\tau^{-1/2}$ for modulation frequencies about 20Hz, while stabilities in the range $E\text{-}14\tau^{-1/2}$ are expected from the future frequency standards...

CONCLUSION

The predictions of the model developed to explain the limitation of the frequency stability by the frequency noise of the interrogation signal are confirmed by the experimental results.

It will be very necessary to take care of this effect in the future passive frequency standards using optical pumping and/or cooled atoms.

Acknowledgements

We are very grateful to DRET (Direction des Etudes, Recherches et Techniques) who has supported this work.

References

(1) 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 Meas. 40, n°2 April 1991, pp 121-125

(2) F.L.Walls, S.R.Stein, J.E.Gray, D.J.Glaze and D.W.Allan, "Design considerations in state-of-the-art signal processing and noise measurement systems", Proc. 30th Annual Symposium on Frequency Control, p.269, 1976.

HIGH RESOLUTION DIGITAL SYNTHESIZER FOR A Rb FREQUENCY STANDARD

E. Detoma, G. Pedrotto
FIAT CIEI S.p.A.
SEPA
Corso G. Cesare, 294-300 - 10154 TORINO [Italy]

A. Stern, B. Levi, M. Bootnik
Time & Frequency Ltd.
14, Habanai Street - HOLON 58117 [Israel]

1 ABSTRACT

In the following, we present some work that has been jointly performed by SEPA and TFL in implementing a Rb frequency standard with an internal high resolution synthesizer and incorporating a novel technique of a.c. C-field modulation.

The aim of the study was to develop a device in which the output frequency can be accurately set without changing the C-field, in order to achieve a good settability in time and frequency without perturbing the conditions of the atomic interrogation.

Moreover, the modulation technique being used (square-wave FM) allowed a straightforward implementation of the TFL-proposed a.c. C-field modulation, the latter technique being specifically intended to reduce the magnetic sensitivity of the Rb Frequency Standard.

2 INTRODUCTION

Modern communication and navigation networks place increasing requirements upon frequency coordination and time synchronization. Traditional techniques to achieve frequency control in Rb frequency standards rely upon control of the current in the C-field.

This method has several disadvantages: first of all, the frequency change is proportional to an analog voltage non intrinsically related to the stability of the standard, but relying on an external voltage or current source for stability. Moreover, changes in the C-field value affect not only the frequency but the magnetic susceptibility of the standard; possible residual magnetic hysteresis can further affect, some time in unpredictable way, the frequency stability and accuracy of the standard. In addition, the method is non linear over a wide frequency range, therefore the voltage to frequency transfer constant can be different from unit to unit, or even in the same unit due to frequency aging. This prevents frequency control in a precise predetermined manner.

An internal adjustable digital synthesizer solves many of these problems, by changing directly the ratio between the interrogation frequency (which remains fixed by the parameters of the atomic resonance line) and the frequency of the local OCXO, with the result of indirectly changing the output frequency of the standard of a known quantity, digitally settable, with accuracy and stability directly comparable to the stability of the standard, since, in reality, what is changed is only a division ratio.

To obtain the maximum resolution, a fractional frequency synthesizer has been implemented in a single ASIC Ic, in such a way that the overall resolution of the synthesizer is less than the intrinsic stability of the standard for the same integration time: this means that, in controlling the output frequency of the standard, no limitation comes from the synthesis but only from the intrinsic stability of the Rb oscillator itself.

Having a digital synthesis within the loop, the natural step was to apply a square-wave frequency modulation/-demodulation, instead of the usual sine-wave modulation; in this arrangement, the capability of the synthesizer are extended to precisely set the width of the interrogation according to the shape of the Lorentzian curve for each resonance cell, in order to optimize the operation of the standard.

Moreover, we have also tried to apply to this scheme the technique, proposed by TFL, of the a.c. C-field modulation; in our case, the C-field is switched coherently with the switching of the two interrogation frequencies for the square-wave primary modulation.

The capability to precisely tune the standard with high resolution and accuracy provides means to precisely set the phase of the output waveform with subnanosecond resolution, by stepping small frequency adjustments over a definite time interval. In this way, accurate remote control of the standard is possible in time and frequency, providing simultaneous synchronization and sintonization of primary oscillators in distributed navigation and communication networks.

3 "TRADITIONAL" ELECTRONICS FOR THE Rb FREQUENCY STANDARD

The traditional Rb Frequency Standard locks the frequency of the signal generated by a reference crystal oscillator (typically 5 or 10 MHz) to the intrinsic stability of an atomic transition of the isotope 87 of the Rb atom. The Rb^{87} atoms are used as a high-Q passive frequency discriminators, whose natural resonance frequency is approximately 6834.683 MHz. The physics of the interrogation (optical pumping and microwave interrogation) is well known and will not be repeated here; the important thing to bear in

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

mind for what follows is that the Rb Standard Physics Package (fig. 1) behaves just like a frequency discriminator, with a characteristic transfer function (output light intensity, or error signal, versus the frequency of the microwave interrogation), function that has the classical Lorentzian shape.

The microwave interrogation signal at 6834.683 MHz is derived by multiplication of the 10 MHz signal in two steps; first the reference (10 MHz) is multiplied to a suitable value (90, 120 or 180 MHz), then, for the final step of multiplication, this signal is fed to a varactor (step-recovery diode - SRD) multiplier, located inside the resonance cell microwave cavity, tuned at 6.8 GHz.

The resulting signal frequency after the multiplication (6.840 GHz) differs from the interrogation frequency of an appreciable amount: this is typically 5.3125 MHz. Therefore, a signal at this latter frequency is injected on the SRD, so that a component, at the frequency:

$$6840.000 - 5.3125 = 6834.6875 MHz$$

is generated within the microwave cavity: this component is used for the atomic resonance interrogation.

In order to extract a useful error signal, the microwave signal is usually modulated by a low-frequency sine-wave; in this way, the interrogation frequency is swept through the resonance line and can be subsequently demodulated by a synchronous demodulator to recover the amplitude and sign of the error signal. This is integrated and used to control the frequency of the 10 MHz reference oscillator, thereby closing the loop, which is a basic **frequency-locked loop (FLL)**.

For a general description and the principles of operation of the Rb Frequency Standard, refer to ref. 1.

4 FREQUENCY SYNTHESIS WITHIN A Rb FREQUENCY STANDARD

The differences between a traditional and a Rb Frequency Standard with an internal adjustable synthesizer is obviously in the electronics package, the physics package being left untouched by the modifications.

In our design, a high resolution synthesizer is used to generate the offset signal at 5.3125 MHz, with the desired resolution to finely tune the interrogation frequency of the atomic resonance, and, as a consequence, to finely tune the output frequency in parts in 10^{-12}, insuring that the resolution of the tunability of the standard is not limited by anything other than the intrinsic stability of the standard itself.

4.1 *Purpose of the synthesizer*

The objectives of the digital synthesis as applied to a Rubidium atomic frequency standard are twofold:

1. to allow the user to precisely set the output frequency of the standard, in order to have an ensemble of oscillators to run at the same frequency within the intrinsic short term stability of the standard itself.
2. to adjust the microwave interrogation of the atomic transition to the natural resonance frequency of the absorption cell, which may vary from cell to cell;

As a fringe benefit, the first feature will allow the user **to apply precise time-steps**, with resolution at the subnanosecond level, to steer an external clock driven by the synthesized standard.

4.2 *Technical characteristics*

The technical specifications for the synthesizer are:

- **a.** to generate a signal, (TTL-compatible), at a frequency around 5.3125 MHz; this is the interrogation frequency for the TFL Rb Physics Package;
- **b.** the reference signal (input to the synthesizer) is a TTL-compatible signal at a frequency of 10 MHz;
- **c.** the synthesizer generates its own modulation; the best modulation technique suited for a digital syntheiszer is obviously a **square-wave frequency modulation**; the modulation frequency is around 80 Hz, non-coherent with the primary (reference) signals, in order to avoid cross-modulation effects;
- **d.** the two discrete interrogation frequencies resulting from the square-wave frequency modulation are **independently** settable with a resolution of a few tens of Hz; this defines the central frequency for the interrogation and the width of the atomic resonance;
- **e.** once the center frequency has been set, it is possible to apply a **fine-tuning offset** with a resolution [referred to the output signal (10 MHz) of the standard] of a few parts in 10·exp(-12): this allows the fine tuning of the output frequency
- **f.** the square-wave frequency modulation leads directly to a **digital synchronous demodulation**; the detector is followed by a track-and-hold (T/H) device **to block the spurious signals** generated during the switching transient from one side frequency to the other, across the resonance line; the transients are due to the natural decay of the energy stored in the atomic resonator constituted by the atoms excited by the first frequency, when this is removed and the second frequency is applied. The energy decay occurs in a time interval (in seconds) inversely proportional to the linewidth (in Hz); since the linewidth of the atomic resonance is around 1.3 kHz, a hold time of the order of 0.8 μs or greater is required.

5 SQUARE-WAVE FREQUENCY MODULATION

In order to have a usable error signal, some kind of modulation must be used to discriminate the sign of the correction resulting from the error signal; phase or frequency modulation of the interrogation signal are usually used in the process; in the current TFL design the interrogation signal is **phase-modulated** at 444 Hz, and the error signal is demodulated with the aid of a synchronous detector driven by a signal at the same frequency after a suitable phase correction has been applied to compensate for the delay accumulated during the interrogation process [the atomic resonance acts as a very narrow bandpass filter, with a high Q in the order of 10^{-7}].

In the new design, the presence of a digital frequency synthesizer leads directly to the choice of a **square-wave frequency modulation** to modulate the interrogation microwave and recover the error signal through a double-balanced synchronous demodulator; the 5.3125... MHz signal is switched between two separate discrete frequencies, differing from each other of an amount equal to the linewidth of the resonance (roughly 1 kHz) and such that the average frequency of the two is the center frequency of the atomic transition.

The **switching occurs at a 80 Hz rate**, so that the interrogation interval at each of the two frequencies is about 6.25 ms (the full interrogation of the resonance being completed in 12.5 ms, which is the period of the 80 Hz modulation frequency). Since the decay of the energy stored in the atomic resonator is a function of the linewidth only (about 1 kHz) and it is in the order of about 1 ms, this effectively leaves about 5 ms of usable error signal. The useless

portion of the error signal is effectively eliminated in the demodulation process by the use of a track-and-hold (T/H) circuit, located between the synchronous detector and the integrator (see fig. 6).

Since the "unusable time" of 1 ms is intrinsically defined by the linewidth, this effectively limits the upper frequency of switching for the interrogation to a value of less than 200 Hz to have a reasonable signal-to-noise ratio in the error signal.

For the integration of the synthesizer in the existing Rb frequency standard, the problem is then reduced to the synthesis of the appropriate 5.3125 MHz signal.

6 SYNTHESIS BLOCK DIAGRAM

The architecture of the 5.3125... MHz synthesizer is shown in fig. 4. The fractional frequency synthesizer generates a signal at a frequency around 20.8 kHz with the required settability; this output signal can be switched at a 80 Hz rate between two discrete settable frequencies, that can be set 1 kHz apart, providing the required frequency modulation for the interrogation and the recovery of the error signal.

A dedicated ASIC (Application Specific Integrated Circuit) was developed for this application, using readily available FPGA (Field Programmable Gate Array) components; two digital dividers provide signals at a frequency of 333.333... kHz and 5 MHz. The synthesizer produces the 5.3125... MHz signal with the following relationship:

$$5.3125MHz = 5MHz + 333.\overline{3}kHz - 20.8...kHz$$

where:

$$333.\overline{3}kHz = 10MHz/30$$

$$5MHz = 10MHz/2$$

and 20.8... kHz is produced with a variable ratio division (about 480) by the fractional frequency synthesizer starting from 10 MHz.
The practical implementation can be seen in fig. 4. A phase-lockable oscillator produces a 312.5 kHz TTL signal; this is part of a commercial PLL Ic package, and it is used as a filter for the 312.5 kHz signal.

The frequency stability of the signals in this part of the electronics is somewhat relaxed by the fact that they are translated, and not multiplied, to 6840 MHz; therefore, the stability for the 5.3125... MHz signal, for instance, appears to be relaxed by a factor of about 684 (ratio of the microwave interrogation frequency to the 10 MHz output signal): in order not to degrade a $1 \cdot 10^{-11}$ stability at τ = 1 s for the output signal of the Rb frequency standard, the 5.3125... MHz signal must exhibit a stability not greater than $1 \cdot 10^{-9}$ over the same sampling interval.

Following the same line of reasoning, the stability of the 20.8... kHz signal should not exceed a mere $1 \cdot 10^{-6}$!

7 FRACTIONAL FREQUENCY SYNTHESIZER

The fractional frequency synthesizer produces a signal with a frequency of about 20.8 kHz (eventually switchable between two discrete frequencies 1 kHz apart); in the following we will give a brief description of its operation (see fig. 5 for a general block diagram of the circuit).

To obtain a frequency in the vicinity of 20.8 kHz the synthesizer has to divide by 480 the 10 MHz input signal:

$$10MHz/480 = 20.8333...kHz$$

The division by 480 requires a 10 stage digital divider; in the actual design, a 12-stage programmable counter is used, the 4 Most Significant Bits (MSBs) being wired and the 8 Least Significant Bits (LSBs) being used to set externally the division ratio, using appropriate logic levels fed at the programmable inputs of the divider.

Since two frequencies must be generated at alternate times (the switching rate is the modulation frequency), two latches are available to hold the 8 LSBs to be fed at the programmable inputs of the counters under control of the modulation signal.

The resolution for this arrangement alone is rather poor, since only integer numbers can be programmed as counts for the dividers; dividing by 480 and by 481 results in two discrete frequencies:

$$10MHz/480 = 20.833...kHz$$

$$10MHz/481 = 20.790...kHz$$

differing of 43 Hz. Since the reference input frequency is fixed, **the resolution** R_f in the frequency synthesis depends upon the division ratio N loaded in the counters:

$$R_f = \frac{f(ref)}{N} - \frac{f(ref)}{N+1} = \frac{f(ref)}{(N \cdot (N+1))}$$

where, in our case:

$$f(ref) = 10MHz = 1 \cdot 10^7 Hz$$

and N = 480, resulting in R_f = 43.31... Hz. This resolution is not great, but we considered it sufficient to set the linewidth of the interrogation without appreciable loss of signal.

The **resulting resolution over the 10 MHz output signal** can be computed by considering that the 10 MHz signal is multiplied by K = 684 (K = 9 x 76) to obtain the microwave frequency needed for the interrogation of the atomic resonance; if the microwave frequency is shifted by R_f, then the 10 MHz frequency is shifted by R_f/K:

$$R_{10} = R_f/K = 63.3mHz$$

where R_{10} is the resolution of the frequency-step at 10 MHz, resulting in the offsettability of the output frequency of the standard of $6 \cdot 10^{-9}$ with respect to the nominal frequency.

To increase the resolution we can dither the division factor preset in the counters. Let f(out) be the frequency produced as the result of the division applied to the input frequency f(ref); if, for example; the counter divides by N for half the time, and by N+1 for the remaining half, then the average frequency produced at the output is:

$$f(out) = \frac{f(ref)}{N+0.5}$$

over a sampling time interval greater then the time required to perform a full cycle of divisions by N and N+1; this effectively produces a divider ratio (N+0.5) which is not an integer any longer, but exhibits a fractional part, which relates to the name given to this device as **fractional frequency divider**.

In our design, the counter uses its own output to switch between N and N+1 (or N-1 to subtract from the fractional count), i.e., it divides by N+1 for a preset number of cycles of the output frequency f(out), then reverts to a division by N for the remaining cycles.

The **cycle count** is held by another counter, called the

small-n counter; this is a special 12-bit counter, that can be programmed for any number from 0 to 4095; its operation is straightforward. If the most significant bit is set, the dividers are programmed to divide by N+1 for one cycle of the output frequency f(ref), and to divide by N for the next cycle; this was shown to be equivalent to a division by N.5, effectively halving the resolution of synthesis of the dividers alone.

Since the period P(out) of each cycle of the output frequency is:

$$P(out) = P(ref) \cdot N$$

for N=480 (N+1 = 481) and P(ref) = 100 ns (corresponding to the 10 MHz reference signal) we have:

- $P(out) = 48.0\mu s$ for N = 480 [f(out) = 20.833 kHz]
- $P(out) = 48.1\mu s$ for N = 481 [f(out) = 20.790 kHz]

The frequency average is effective on a minimum time interval of (48 + 48.1 μs) 96.1 μs, and will be 20811.6 kHz; in general, the time interval required to average the output frequency is the period of the output signal multiplied by the number of cycles entering the fractional count. In our case, the maximum resolution of the synthesizer is related to the dimension of the **"small-n" divider**, which is a 12-bit device.

The fractional frequency dividers will be able to divide by N+1 for 1 cycle over 4096, and the remaining 4095 cycles the divider will divide by N. For an output frequency around 20.8 kHz the period is about 48 μs, and the output frequency for the maximum resolution (the least significant bit set in the "small-n" counter) must be averaged over:

$$AvgTime = P(out) \cdot 4096 = 48\mu s \cdot 4096 = 196.6ms$$

to obtain his correct value as preset. The resulting resolution r_f will be $R_f/4096$, that is:

(1)

$$r_f = \frac{f(ref)}{N \cdot (N+1) \cdot n_{max}} = 10.57mHz$$

where: f(ref) = 10 MHz, N = 480 and:

$$n_{max} = 2 \cdot 10^{12} = 4096$$

The resolution r_{10} in the settability of the output frequency (10 MHz) of the synthesized Rb frequency standard can be obtained by dividing r_f by the factor K. This is the ratio between the frequency of the reference signal (10 MHz) and the frequency of the microwave interrogation (6840 MHz). Since K = 684, the resolution r_{10} in the settability of the output frequency via the fractional frequency divider is 15.5 μHz, resulting in an overall **resolution of** $1.5 \cdot 10^{-12}$ **for the 10 MHz output** over a nominal sampling interval of 196.6 ms.

This is a considerably smaller figure than the intrinsic short term stability of the frequency standard over the same sampling interval: **therefore, the fractional frequency synthesizer resolution is not the limiting factor in the offsettability of the output frequency of the standard**, but rather the main limitation originates from the intrinsic short term stability of the standard itself.

8 REDUCING THE SENSITIVITY TO MAGNETIC FIELD BY THE A.C. C-FIELD METHOD

The a.c. C-field scheme was described in (2) and (3). The idea is to alternate the C-field between $+H_C$ and $-H_C$, and to average. The alternating field can take the form of a square-wave or any periodic function such as a sine wave. The alternating frequency should be smaller than the Rubidium linewidth (1200 Hz) but higher than the servo-loop bandwidth (2 Hz).

We present here a simplified calculation that demonstrates the method. The dependence of the frequency, f, on the magnetic field, H, is given by ref. (1)

(2)

$$f = f_0 + aH^2$$

with: $a = 575Hz/Gauss^2$, and $f_0 = 6.8 \cdot 10^9 Hz$. Writing:

(3)

$$H = \pm H_C + H_D/\alpha$$

where: $\pm H_C$ is the alternating "C" field, H_D is the parallel component of the disturbance field and α is the attenuation shielding factor.

Substituting (3) in (2) and averaging:

(4)

$$\overline{f} = f_0 + aH_C^2 + \left(\overline{\pm 2aH_C H_D/\alpha}\right) + a(H_D/\alpha)^2$$

$$= f_0 + aH_C^2 + a(H_D/\alpha)^2$$

The fractional frequency change due to the disturbance is given by:

(5)

$$\Delta f/f = a(H_D/\alpha)^2/f_0$$

Substituting α = 7000 we obtain:

(6)

$$\Delta f/f = 1.7 \cdot 10^{-15}$$

for H_D = 1 Gauss, and

$$\Delta f/f = 1.7 \cdot 10^{-13}$$

for H_D = 10 Gauss. This is compared with $1 \cdot 10^{-11}$/gauss for the case of constant "C" field.

However, alternating the magnetic field induces transients signals that arise from the zero crossing of the C-field (ref. 2). In our case, the C-field modulation is being synchronized with the frequency modulation. Then, the C-field transients occur at the same time as the transients that arise from the frequency switching. The Track-and-Hold stage, as described before, is then being used to block both transients. This is shown in fig. 8.

9 PRELIMINARY MEASUREMENTS

In the following we present preliminary results demonstrating the performance of the unit, as well as measurements relating to the physics response to the square-wave frequency modulation and the C-field modulation. The measurements were performed on a breadboard that incorporated TFL's Rubidium model TF-4000B with SEPA's synthesizer circuit.

9.1 Transients

Fig. 7 presents plots of the error signal from the photodiode (after the preamplifier) in two cases: (a) in a lock condition, (b) when unlocked. When unlocked, one observes a square wave shape that results from the square-wave fre-

quency modulation. When locked, the error signal includes negative transient signals with a period of 12.5 ms, and positive transients with a period of 50 ms.

Both transients have the same width, about 1 ms. The negative transients are related to the rapid frequency passage through resonance, resulting from the square wave modulation. This was described by (4). On the other hand, the positive transients (50 ms period) are related to the zero crossing of the C-field, as it changes from $+H_C$ to $-H_C$: see references (2) and (3). The transients are being removed by the use of the Track-and-Hold, as shown in fig. 8.

9.2 Short Term Stability

Fig. 9 shows an Allan-deviation plot. It shows an Allan-deviation of $8.9 \cdot 10^{-12}$ at τ = 1 s, and $4.7 \cdot 10^{-12}$ at τ = 10 s. The results above 5 s are related to the stability of the reference rather than the Rubidium (we expect an Allan deviation of $1 \cdot 10^{-12}$ at τ = 100 s averaging time).

9.3 Frequency settability

The smallest fractional frequency step, which is related to the Least Significant Bit (LSB), is about $1.5 \cdot 10^{-12}$ at an averaging time of 196 ms, as explained in section (7) above. However, since in our measurements it was the reference that limits the Allan deviation to $5 \cdot 10^{-12}$, we have only applied steps of $2 \cdot 10^{-11}$ and higher, in order to test the settability. These are shown in fig. 10 [(a) to (d)].

9.4 Magnetic Field Insensitivity

As shown in section (8), we expect fractional frequency changes of less than $1.5 \cdot 10^{-13}$ for a magnetic field disturbance of 10 Gauss. This is smaller than the frequency noise level resulting from the short term stability (Allan deviation).

The response of the breadboard to external magnetic disturbances fields is shown in fig. 11; the disturbing fields are applied parallel and perpendicular to the C-field. No change in frequency is observed when applying the a.c. C-field, as expected.

10 APPLICATIONS

Many applications will benefit from the capability to precisely control the output frequency of the Rubidium frequency standard.

Communications networks using reference oscillators at the primary and secondary nodes will be able to accurately tune every standard to the same frequency, with an uncertainty that is only dependent from the instrinsic stability of the standards themselves.

Since frequency steps applied for a definite amount of time produce phase (or time) steps in the output signal, the reference oscillators can be "phase-coordinated"; a failure of the oscillator at one node can be overcome by locking the failed node to another oscillator located at a remote node without phase or frequency discontinuities.

The capability to precisely adjust frequency and phase digitally will make the device ideal for **timing applications**, ranging from producing "real" outputs from **"paper clocks"** in timing labs, to implementing accurate time systems, such as in **disciplined oscillators**, or by locking the standard output frequency to an external reference, such as **in a GPS receiver or a Loran-C receiver**, both easily providing a digital output of the synchronization or sintonization errors.

Space applications of frequency standards and, in general, any application where **remote control** of the standard is desired, will benefit from the capability to apply time and frequency corrections, digitally and with the same accuracy of the frequency output of the oscillator being controlled; **spaceborne oscillators** should be the first to benefit from the new capabilities of this device.

11 CONCLUSIONS

We have described the implementation and the principles of operation of an adjustable fractional frequency synthesizer applied to a Rubidium Frequency Standard. Preliminary results of measurements that were performed on a breadboard unit are given. We are now in the process of assembling the first prototypes, where additional tests are planned. These include temperature tests as well as finer settability tests.

We expect a short term stability (Allan deviation) of parts in 10^{-13} for averaging times around few hundreds seconds. In that case one should be able to test and measure fractional frequency steps of $1.5 \cdot 10^{-12}$. One could then attempt even smaller frequency steps of, for example, $7 \cdot 10^{-13}$: this can be achieved, for instance, by stepping the frequency half the time by $1.5 \cdot 10^{-12}$ and half the time by 0 (no change). This is equivalent to increasing the length of the "small-n" counter of one (or more) stage(s).

12 ACKNOWLEDGEMENTS

One of the authors (E. Detoma) is deeply indebted to Dr. Giovanni Busca, Director of the Observatoire Cantonal de Neuchatel, for the many interesting discussions and useful insights on the theory and operation of the Rb frequency standards.

Much of the assembly and testing work during the early development phase was carried on by Mr. E. Santoro, Mr. Santa, Mr. Bertoluzzo and Mr. Tubiana of SEPA.

13 REFERENCES

1. Audoin C., Vanier J., THE QUANTUM PHYSICS OF ATOMIC FREQUENCY STANDARDS
2. Stern A. et Val, 43rd Ann. Symp. on Freq. Control, 1989, p.124
3. Stern A., Golosovsky M., 44th Ann. Symp. on Freq. Control, 1990, p.53
4. Camparo J.C., Frueholz R.P., Phys. Rev. A, 1984, p.803

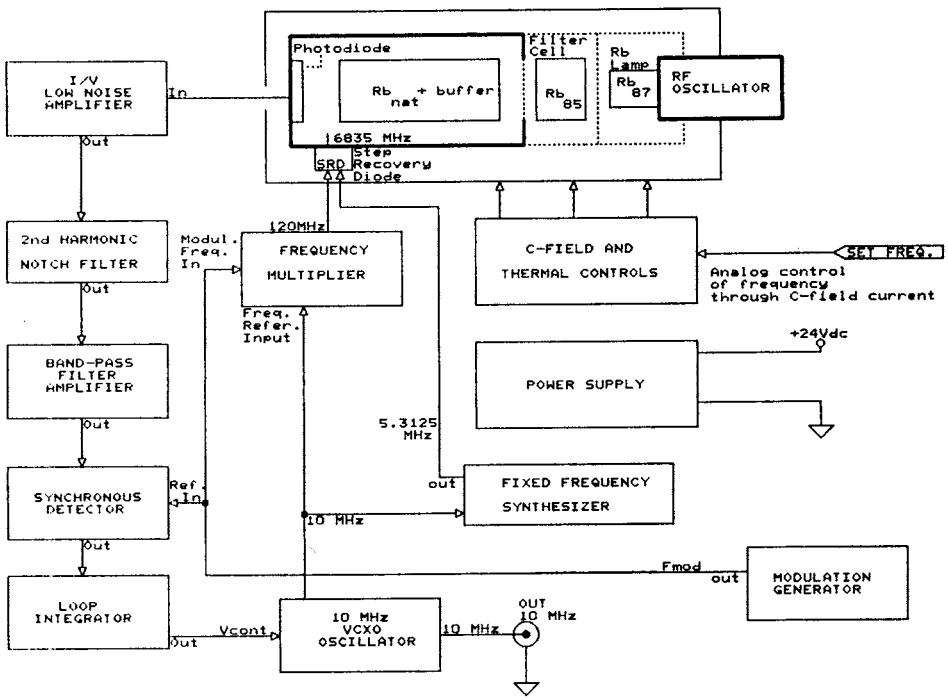


FIGURE 1:Rb Passive Atomic Frequency Standard: Traditional Design

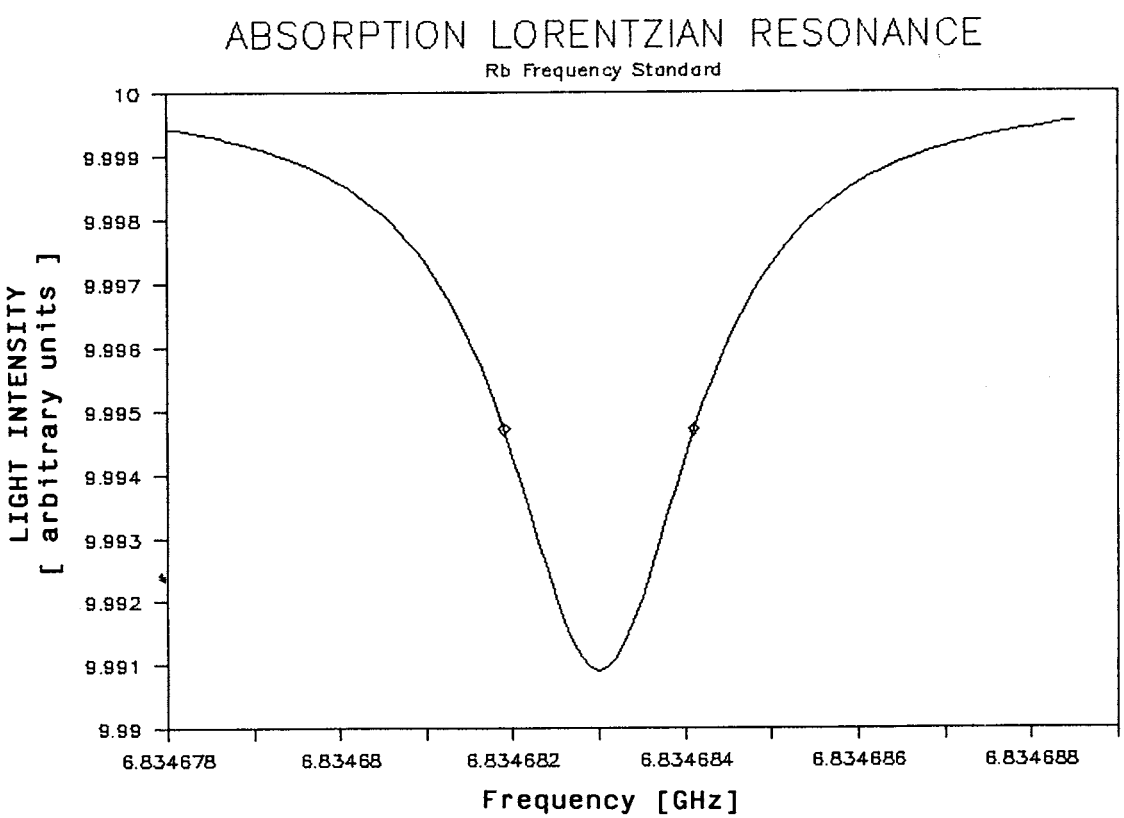


FIGURE 2:Microwave Interrogation and Light Intensity: the Lorentzian shape of the atomic frequency discriminator

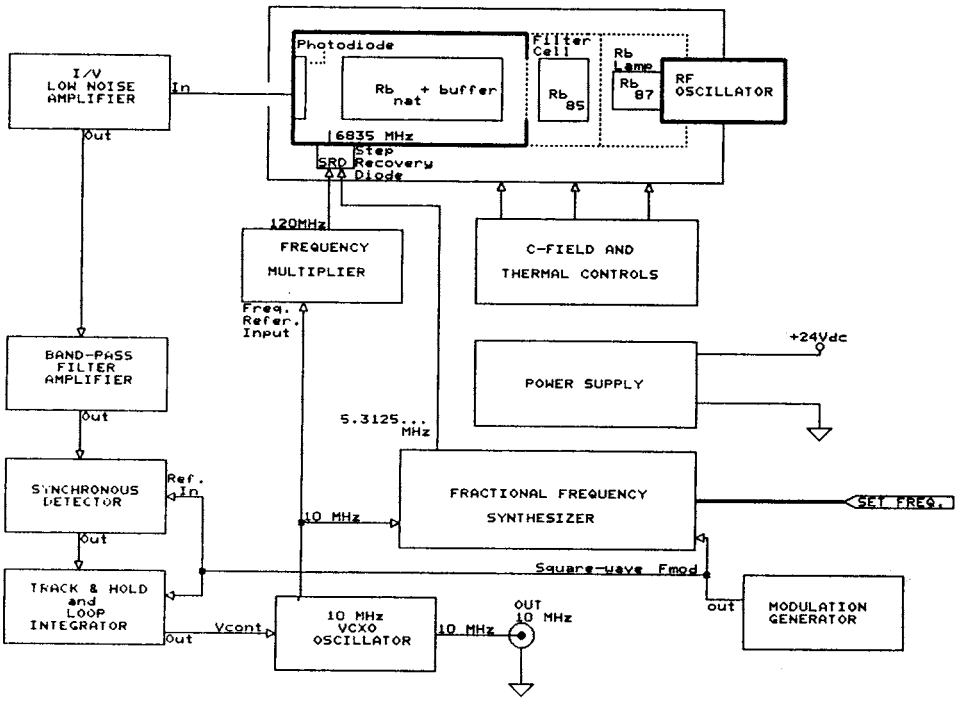


FIGURE 3:Rb Passive Atomic Frequency Standard: "Synthesized" Design

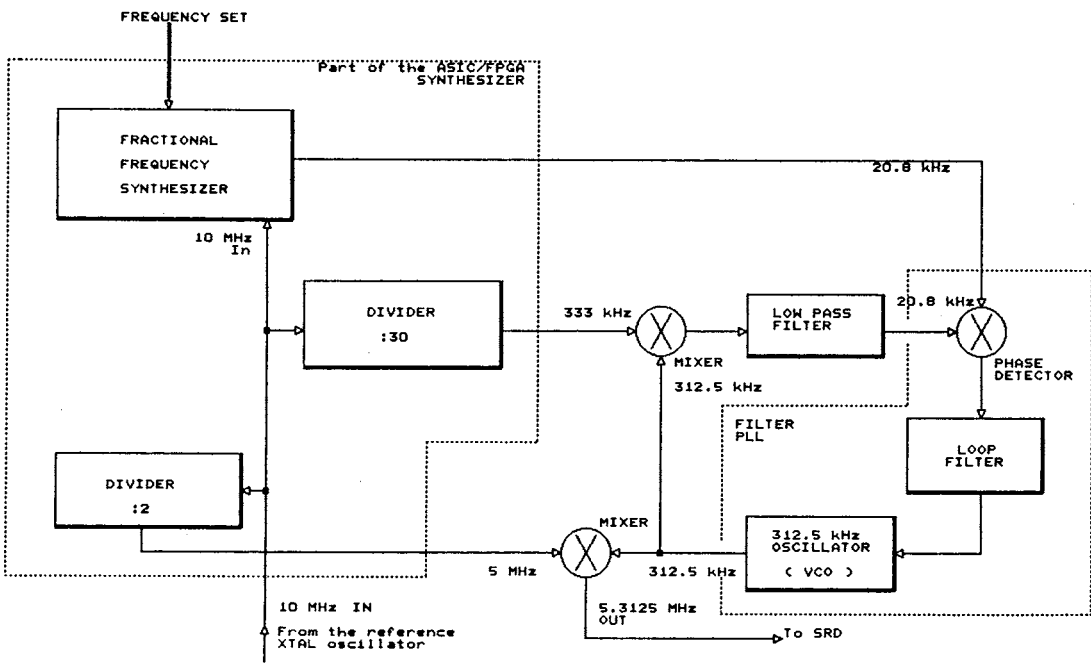


FIGURE 4:Architecture of the Synthesizer

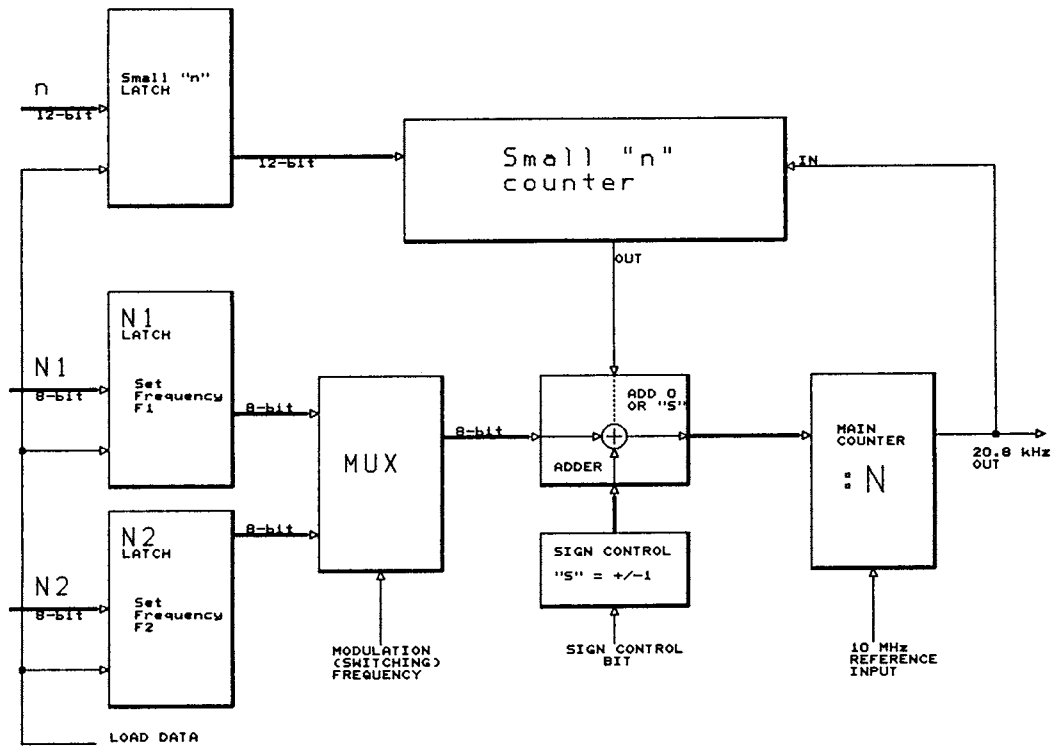


FIGURE 5:Fractional Frequency Synthesizer Block Diagram

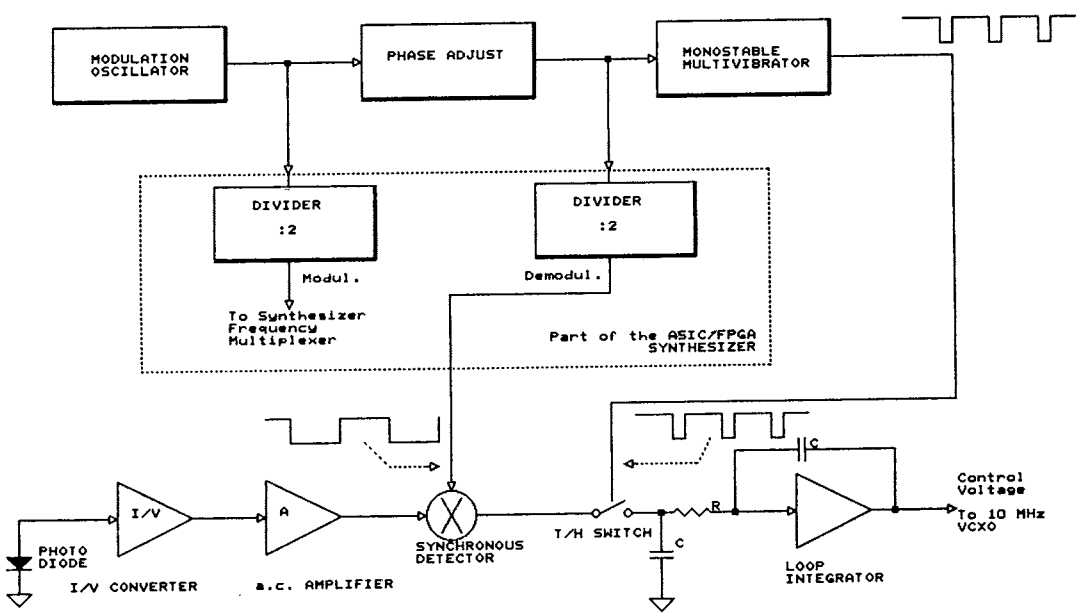


FIGURE 6:Square-Wave FM Modulation/Demodulation: Block Diagram of the Electronics

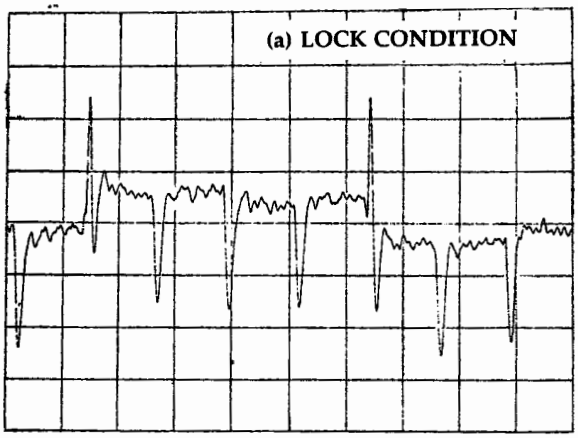
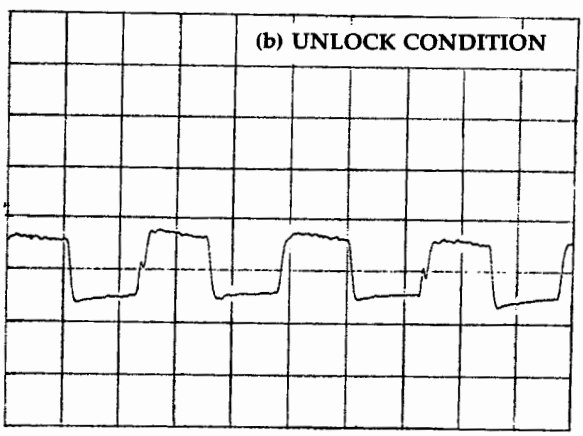


FIGURE 7:Error signal from the Photo-diode; (a) - Lock; (b) - Unlock

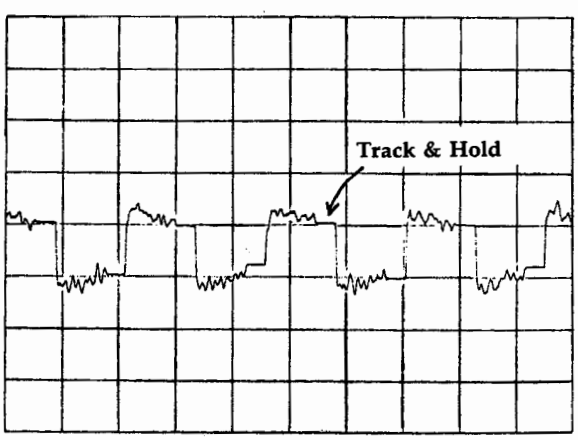


Vertical scale (a): 1.5 V/div

Vertical scale (b): 10 V/div

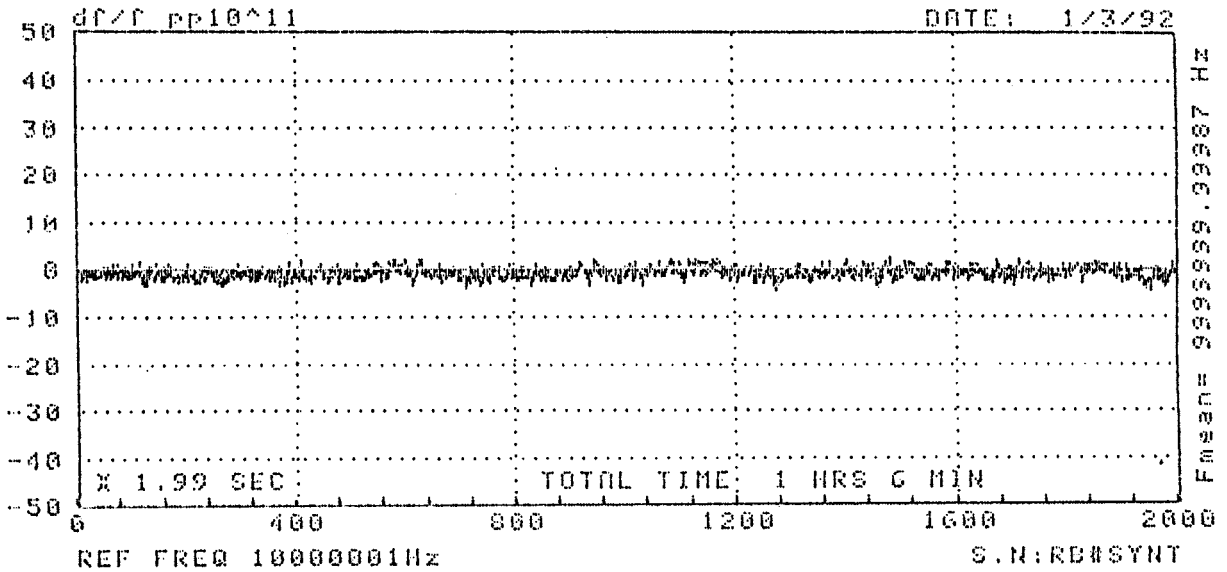
Horizontal scale: 5 ms/div

FIGURE 8:Track-and-Hold operation



Vertical scale: 0.25 V/div

Horizontal scale: 5 ms/div



SAMPLES	TAU(Sec)	Sigma
2000	1.000	8.92E-012
1000	2.000	7.02E-012
500	3.999	4.78E-012
250	7.999	4.68E-012
125	15.998	4.76E-012
63	31.996	3.48E-012
32	63.992	3.79E-012
16	127.984	3.50E-012
8	255.968	2.11E-012

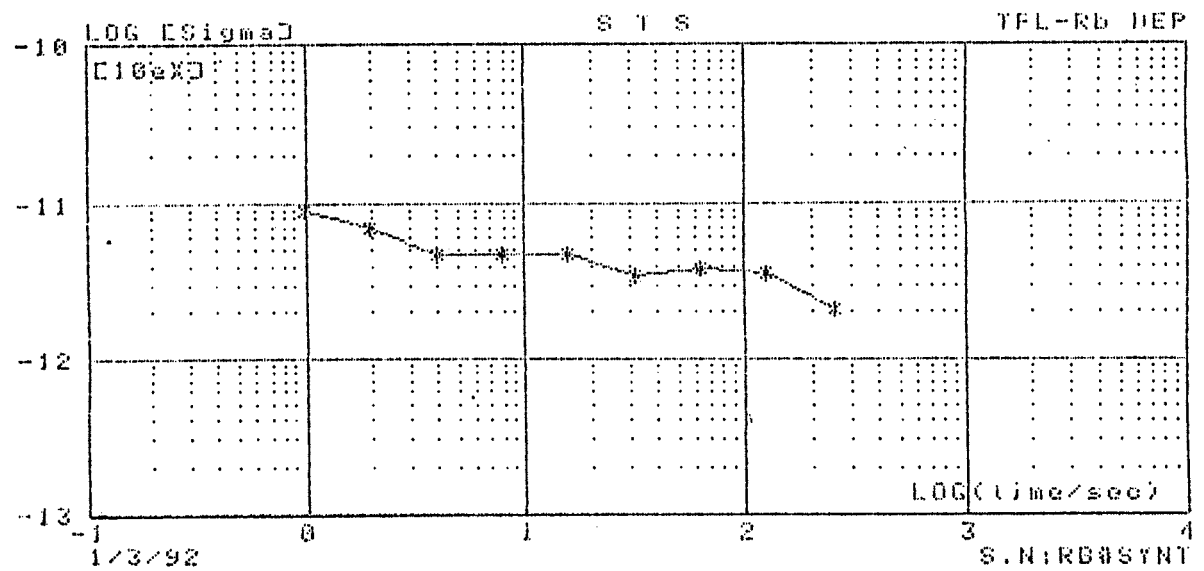


FIGURE 9:Results of Allan deviation (short term stability)

Figure 10 (a)

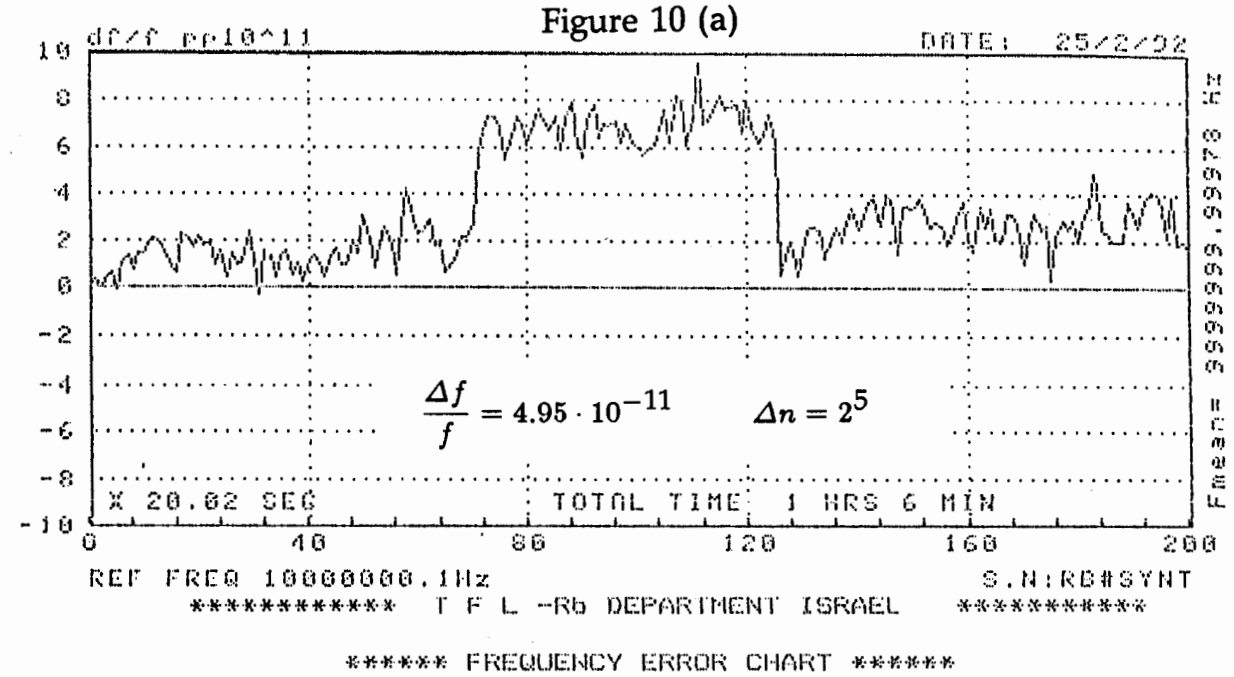


Figure 10 (b)

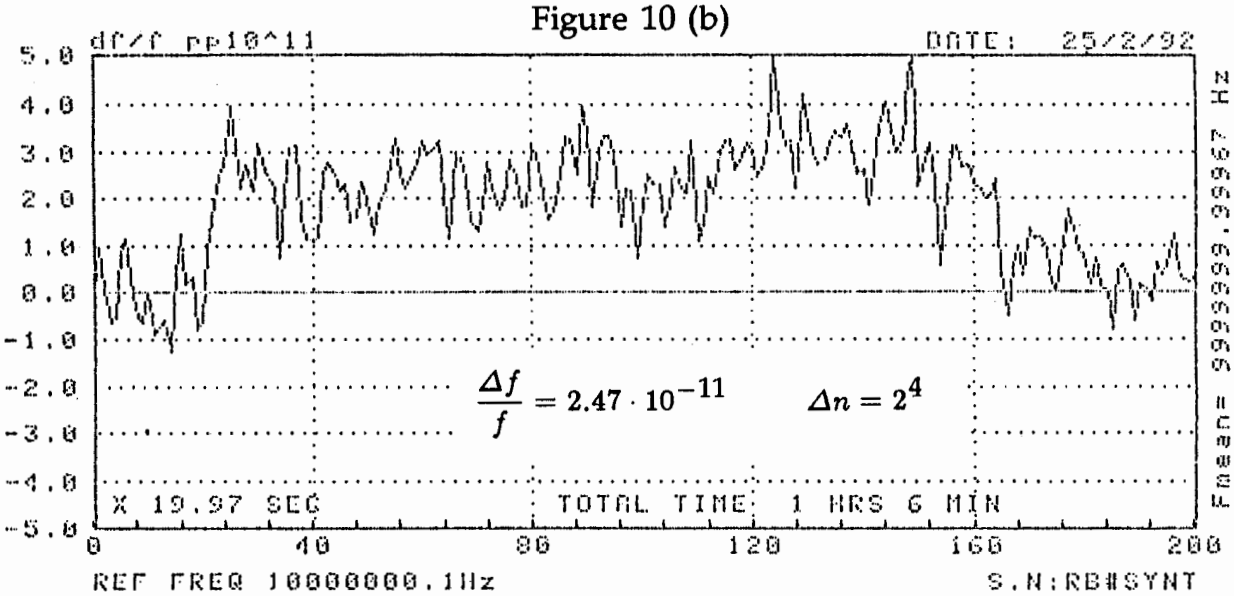


FIGURE 10:Exhibiting Frequency Settability

Figure 10 (c)

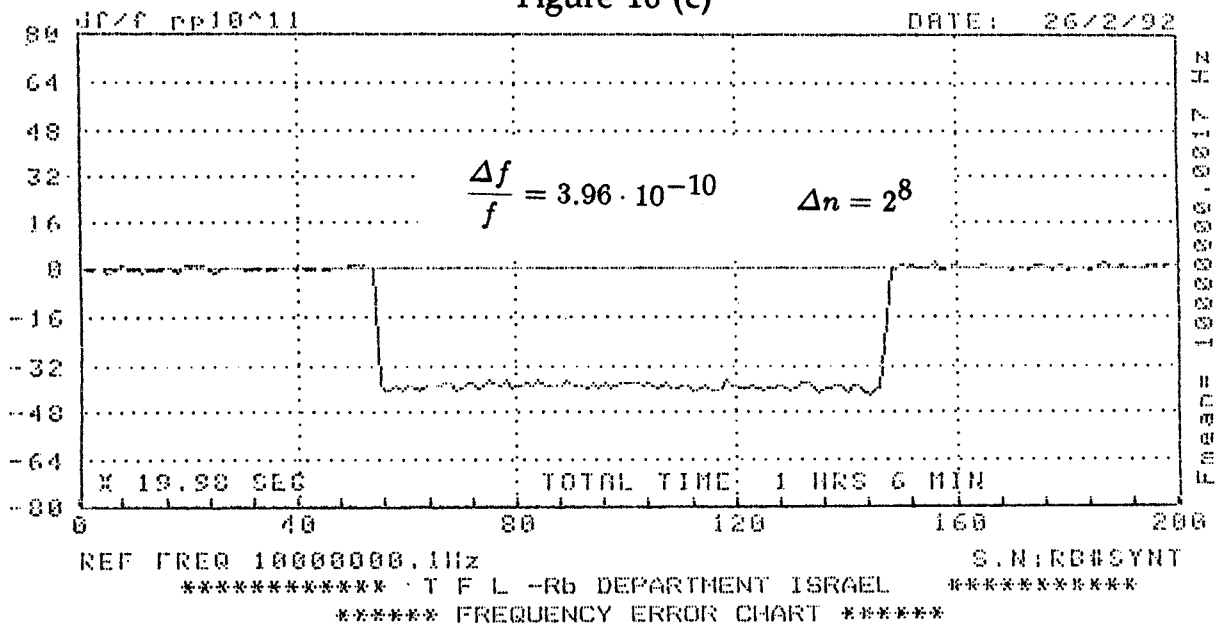


Figure 10 (d)

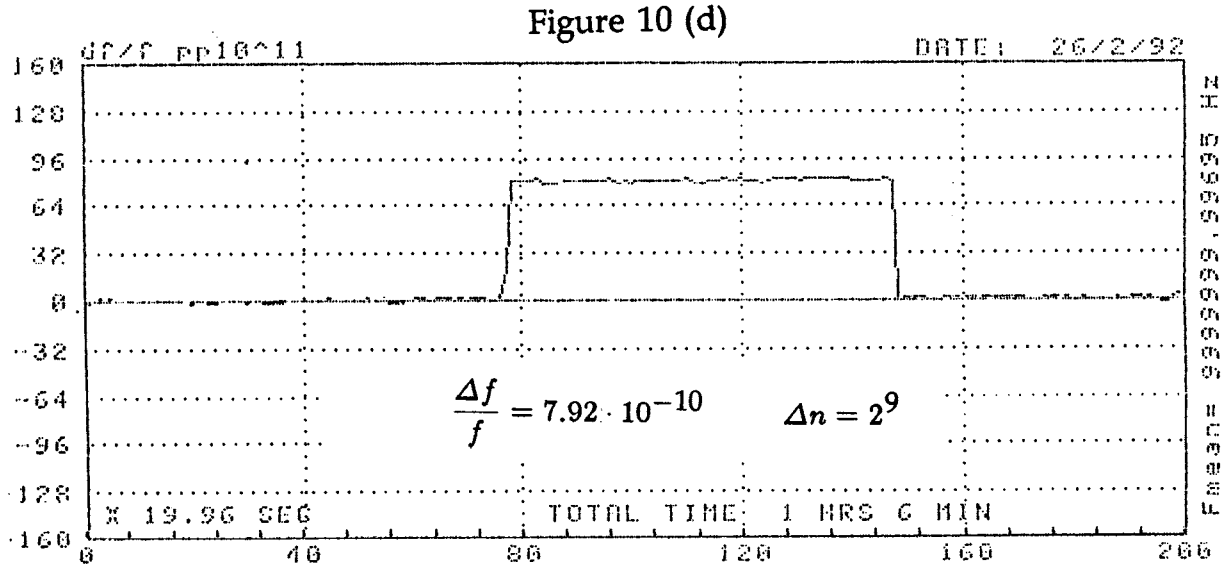


FIGURE 10:Exhibiting Frequency Settability (Cont.)

Figure 11 (a)

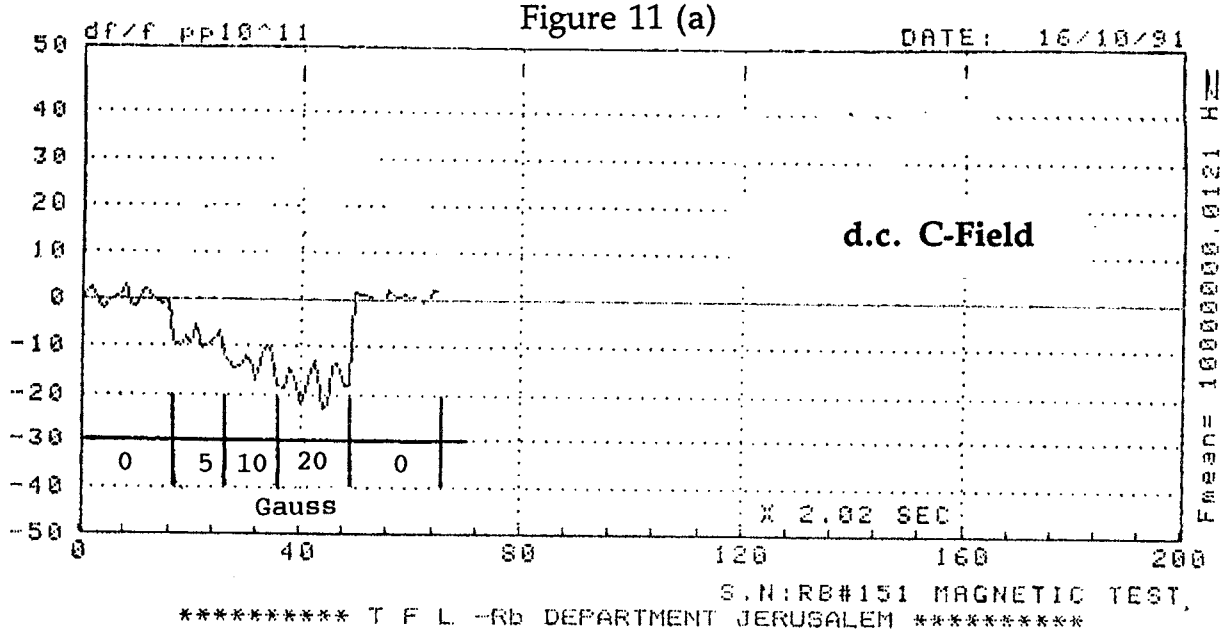


Figure 11 (b)

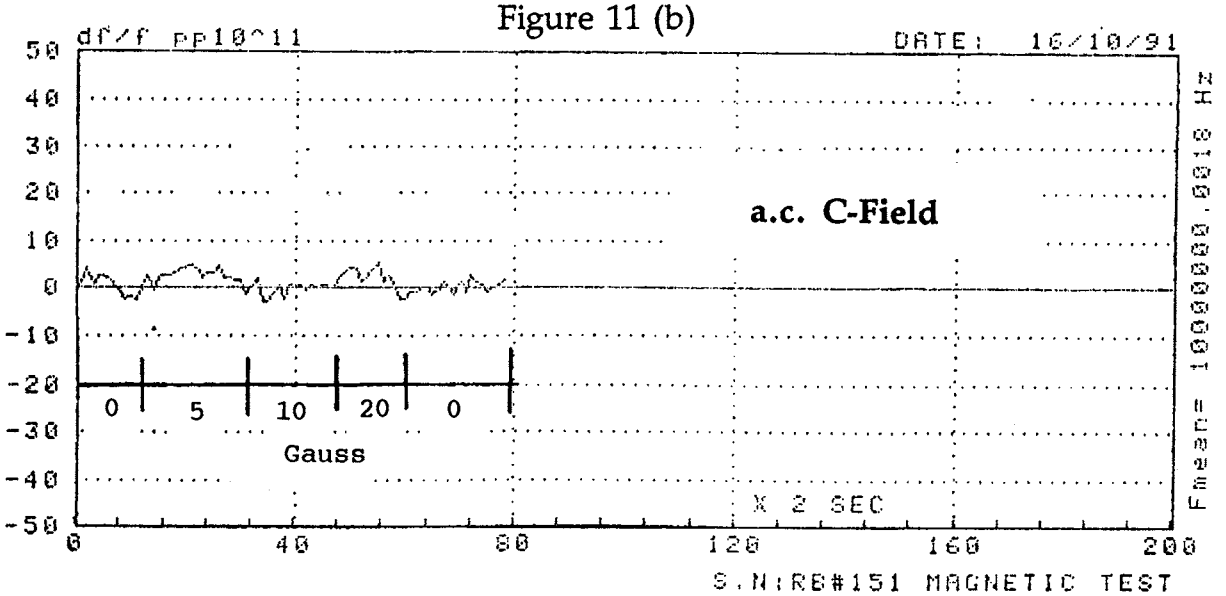


FIGURE 11:Sensitivity to external magnetic field with an a.c. and with a d.c. magnetic C-field.

SOLID STATE FREQUENCY SYNTHESIS UP TO 600 GHz: AN APPLICATION TO THE Mg FREQUENCY STANDARD

A. Godone, C. Novero

Istituto Elettrotecnico Nazionale "Galileo Ferraris"
Strada delle Cacce, 91 - 10135 Torino, Italy

Abstract

We describe in this paper an all solid-state frequency multiplication chain from 5 MHz to 601.3 GHz with spectral purity and output power suitable to excite the $^3P_1 - {}^3P_0$ $\Delta m_j = 0$ transition in the metastable triplet of Mg. By use of this chain the Mg frequency is remeasured with an uncertainty of $\pm 4 \times 10^{-12}$ with respect to the Cs standard, corresponding to an improvement of an order of magnitude with respect to the previous data.

Introduction

The extension of phase-coherent frequency synthesis up to submillimetric wavelengths is of interest in several fields such as telecommunications, radio-astronomy, sub-Doppler spectroscopy and metrology |1,2|. Two main problems have to be faced in order to synthesize a submillimetric signal suitable for the above applications: i) the high harmonic number involved in the multiplication chain imposes severe constraints on the phase-noise characteristics of the first multiplication stages and of the driving oscillator; ii) the high output frequency is a limiting factor to the available output power. The former has been considered in |3| and more recently phase-coherent frequency synthesis has been demonstrated up to 3.7 THz |4|; the latter has been overcome by use of backward-wave-oscillators (BWO) and/or optically pumped far infrared lasers |5,6,7|. In this paper we shall describe an all-solid-state frequency multiplication chain from 5 MHz up to 601.3 GHz with spectral purity and output power suitable to excite the $^3P_1 - {}^3P_0$ $\Delta m_j = 0$ ^{24}Mg transition in a metastable atomic beam. A 100 GHz Gunn oscillator, phase-locked to a 5 MHz quartz crystal oscillator through a low phase-noise frequency multiplication chain, drives a varactor doubler followed by a varactor tripler with an output power of 0.3 mW at 601 GHz. The phase-noise spectrum at different multiplied frequencies along the chain and the Gunn oscillator phase-lock loop will be discussed in detail in order to characterize the fluctuation spectrum of the output frequency at 601 GHz.
Finally we present the application of the solid-state source to the Mg beam frequency standard and, in particular, a new frequency measurement of the Mg clock transition at the 10^{-12} level, corresponding to an improvement of an order of magnitude with respect to the previous data.

Description of the chain

The development of highly efficient varactor multipliers in the millimetric and submillimetric regions |8| makes possible the design of the frequency multiplication chain reported in fig. 1.

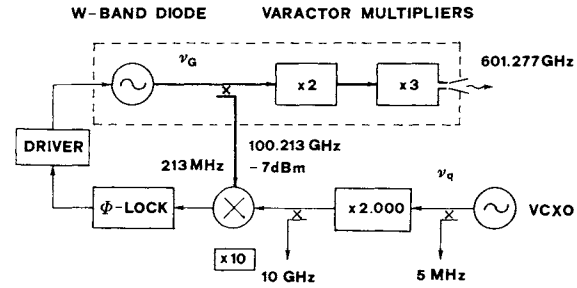


Fig. 1 - Block-diagram of the solid-state multiplication chain (5 MHz-601 GHz).

The W-band oscillator is an InP Gunn diode delivering 75 mW at ν_G = 100.2 GHz; it is electronically tunable through the bias voltage (±50 MHz) with an FM sensitivity $K_o(\omega)$ given by the following expression:

$$K_o(\omega) = \frac{K_{oo}}{1+j\omega\tau_G} \qquad (1)$$

where ω is the modulation frequency, $K_{oo}/2\pi$ = 240 MHz/V the d.c. FM sensitivity and τ_G = 1 µs the time constant defining the modulation bandwidth.
The Gunn oscillator directly drives a fixed-tuned doubler employing a two-varactor diode balanced configuration which allows a conversion efficiency as high as 25%; the power at 200.4 GHz is then about 18 mW, an optimum value to drive the fixed-tuned varactor tripler used in the last multiplication step. The output power of the tripler is shown in fig. 2 versus the Gunn bias voltage together with

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

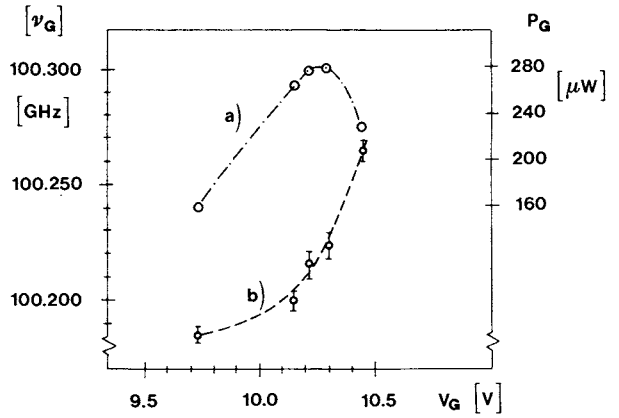


Fig. 2 - Curve a) shows the output power at 601 GHz versus the Gunn bias voltage V_G; curve b) reports the Gunn frequency ν_G versus V_G.

its tuning curve; an efficiency higher than 1% is achieved in this case. The output power is optimized at the operating frequency 601.3 GHz, what strongly reduces the FM-AM conversion factor; moreover the amplitude stability of the output power is better than 1% per hour. The submillimetric radiation is coupled out through a horn with a beam-width of 12° in the E-field direction and 15° in the H-field direction.

A fraction (-7 dBm) of the Gunn oscillator signal is sent to a GaAs harmonic mixer (harmonic number 10); the X-band output (+10 dBm) of a low phase-noise frequency multiplication chain, with input frequency ν_q = 5 MHz, is also coupled to the mixer, whose IF signal at 213 MHz is sent to the phase-lock circuit of the millimetric oscillator.

The 5 MHz - 10 GHz multiplication chain is similar to that described in |3| and its white phase-noise level has been tested at different multiplied frequencies as shown in fig. 3; the measured values

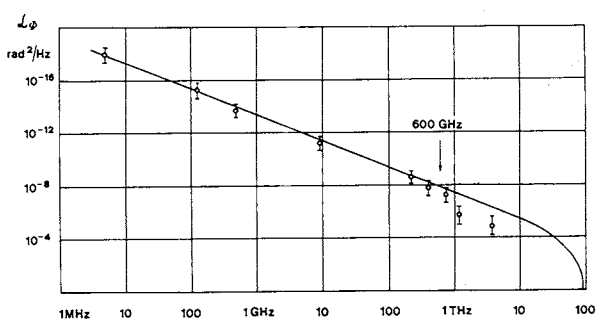


Fig. 3 - Two-sided white phase-noise level at different multiplied frequencies of the 5 MHz-10 GHz multiplication chain; $\mathcal{L}_\phi = \frac{1}{2} S_\phi$.

above 100 GHz have been obtained using as a reference optically pumped far-infrared laser emissions (see for example |4|). The flicker phase-noise of the multipliers is completely overcome by the flicker frequency-noise of the driving quartz oscillator, so that the phase-noise spectrum of the reference signal is described at any multiplied frequency by the following expressions:

$$S_{\phi_R}(f,\nu) = N^2 S_{\phi_R}(f,\nu_q)$$

$$S_{\phi_R}(f,\nu_q) = 2\times10^{-9} f^{-3} + 2\times10^{-18} \text{ rad}^2/\text{Hz} \quad (2)$$

where $N = \nu/\nu_q$ is the multiplication factor. The above spectrum is band-limited in its white phase-noise region at the Fourier frequency f = 12 kHz by a quartz filter inserted in the VHF step of the multiplication chain.

FM-spectrum of the free-running Gunn oscillator

The fractional frequency-noise spectral density $Sy_G(f)$ of the free-running oscillator has to be known for a proper design of the phase-lock loop and for the characterization of the spectrum of the locked oscillator; it may be expressed in the well-known form:

$$Sy_G(f) = \sum_{-2}^{+2} h_\alpha f^\alpha \quad (3)$$

We have performed both frequency and time-domain analysis of the frequency instabilities of the Gunn oscillator through the IF signal at 213 MHz (see fig.1) in open loop conditions, aiming to determine the h_α coefficients of expansion (3).

In fig. 4 the measured frequency spectrum $Sy_G(f)$ of

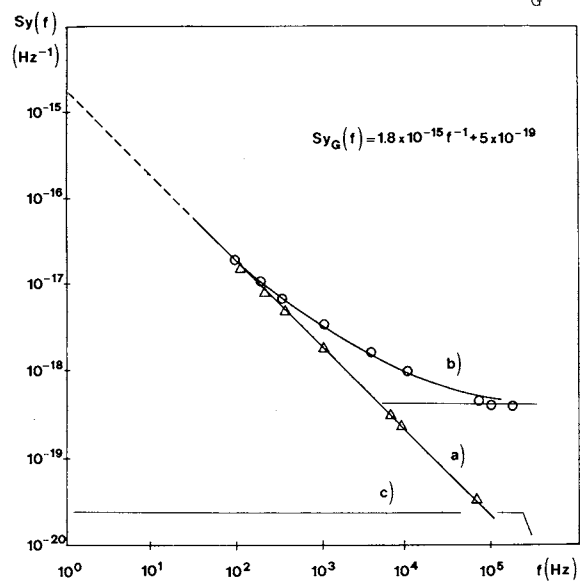


Fig. 4 - Curves a) and b) refer to the measured frequency spectrum $Sy_G(f)$ of the free-running Gunn oscillator, respectively for V_G = 8.2|V| and V_G = 10.2|V|. Curve c) reports the Gunn-driver noise spectrum, transferred to the output frequency $Sy_D(f)$.

the Gunn oscillator is reported for the Fourier frequency range 100 Hz-100 kHz: the flicker-frequency noise appears the dominant one (curve a), what is typical for high frequency low-Q sources; the

white-frequency noise is present only when the Gunn diode is driven near the maximum allowed voltage (curve b). The Gunn-driver noise spectrum, transferred to the output frequency, $Sy_D(f)$ has also been reported in fig. 4 (curve C); it is related to the noise-voltage spectral density Sv(f) by the following relation:

$$Sy_D(f) = \frac{|K_o(\omega)|^2}{\nu_G^2} \; Sv(f)$$

At the operating point (V_G = 10.2 V, I_G = 0.4 A) we have measured:

$$Sv(f) = 4\times10^{-15} \; V^2/Hz, \quad 1 \; Hz < f < 1 \; MHz$$

It is clear from fig. 4 that no contribution to the free-running frequency spectrum of the Gunn oscillator is expected from the used driver.
The Allan standard deviation $\sigma y_G(\tau)$ of the millimetric source is reported in fig. 5 for measurement time intervals τ in the range 1-100 s.

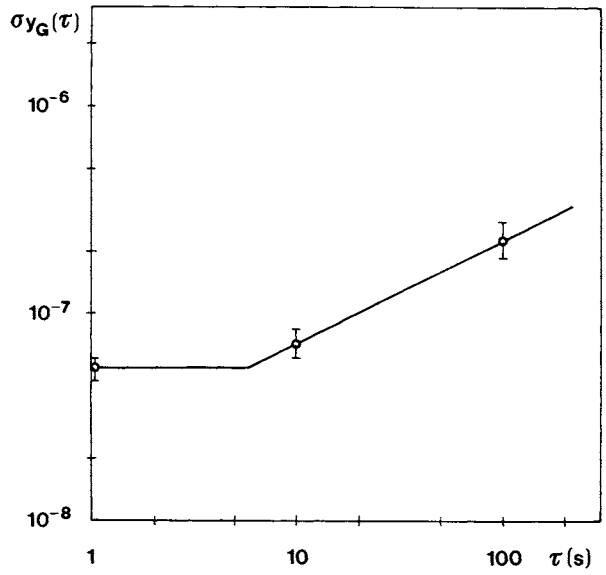


Fig. 5 - Allan standard deviation $\sigma y_G(\tau)$ of the free-running Gunn oscillator.

All the measurements reported in fig. 4 and 5 may be summarized in the following expansion:

$$Sy_G(f) = h_{-2}f^{-2} + h_{-1}f^{-1} + h_o \qquad (4a)$$

$h_{-2} = 2.5\times10^{-16}$ Hz; $h_{-1} = 1.8\times10^{-15}$; $h_o = 5\times10^{-19}$ Hz^{-1}

or by the equivalent phase-noise spectral density $S_{\phi_G}(f)$:

$$S_{\phi_G}(f) = \nu_G^2 h_{-2} f^{-4} + \nu_G^2 h_{-1} f^{-3} + \nu_G^2 h_o f^{-2} \qquad (4b)$$

In fig. 6 we report for a direct comparison the phase-noise spectral densities of the reference signal multiplied at 100 GHz, of the free-running Gunn diode oscillator and of the additive noise due to the IF section. The last noise source is given by the following relation:

$$S_{\phi_n}(f) = \frac{1}{P} \; LKFT \qquad (5)$$

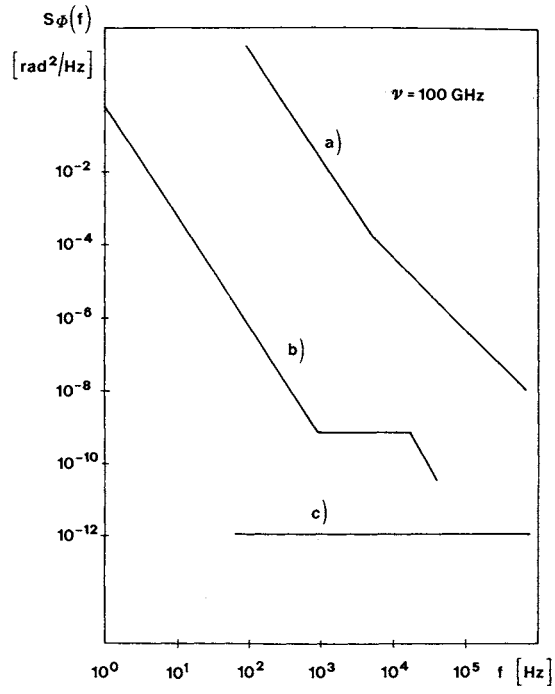


Fig. 6 - Phase-noise spectral densities of: a) free-running Gunn oscillator; b) the reference signal multiplied at 100 GHz; c) additive noise of the IF section.

where P is the local oscillator power (-7 dBm);
L is the conversion loss of the harmonic mixer (43 dB);
KT is the thermal noise in 1 Hz bandwidth;
F is the noise figure of the IF amplifier.
It turns out from fig. 6 that the following inequalities are always satisfied in the Fourier frequency range of interest:

$$S_{\phi_n}(f) << N^2 S_{\phi_R}(f) << S_{\phi_G}(f) \qquad (6)$$

Gunn oscillator phase-lock

The block diagram of the system adopted for the phase stabilization of the millimetric source is shown in fig. 7. The IF signal from the harmonic

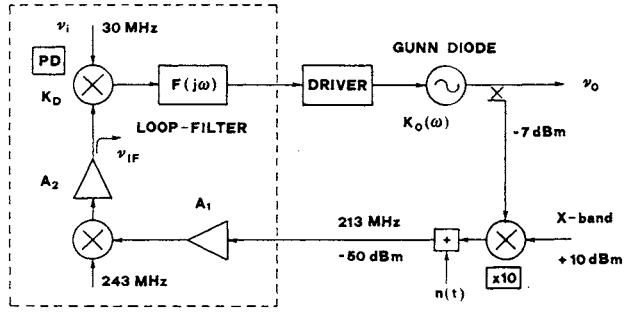


Fig. 7 - Block-diagram of the Gunn phase-lock loop; the boxed part corresponds to the Φ-LOCK block of fig. 1.

mixer is amplified by A_1 (noise figure F = 4 dB), is downconverted to the new IF frequency ν_{IF} ($\simeq$ 30 MHz) and sent to the limiting amplifier A_2. The phase-detector (PD) is driven by the limiting amplifier and by the reference signal ν_i ($\simeq$ 30 MHz); the dc output signal (error signal) is sent to the Gunn oscillator driver through a loop filter with transfer function F(jω). In the same figure we have also indicated the additive noise term n(t), whose contribution to the phase instabilities of the X-band reference is given by expression (5).

The phase spectral densities of the IF signal $S_{\phi_{IF}}(f)$ and of the Gunn signal $S_{\phi_o}(f)$ in closed loop conditions may be expressed in the following form |6|:

$$S_{\phi_{IF}}(f) = |1-H(j\omega)|^2 |S_{\phi_G}(f) + N^2 S_{\phi_R}(f) + S_{\phi_n}(f)| \quad (7)$$

$$S_{\phi_o}(f) = |H(j\omega)|^2 |S_{\phi_n}(f) + N^2 S_{\phi_R}(f)| + |1-H(j\omega)|^2 S_{\phi_G}(f) \quad (8)$$

where H(jω) is the closed loop transfer function |9|.

Introducing inequalities (6) in (7) and (8) we obtain the simplified spectra:

$$S_{\phi_{IF}}(f) = |1-H(j\omega)|^2 S_{\phi_G}(f) \quad (9)$$

$$S_{\phi_o}(f) = |H(j\omega)|^2 N^2 S_{\phi_R}(f) + S_{\phi_{IF}}(f) \quad (10)$$

In the time domain the frequency stability of the phase-locked oscillator, evaluated via the two-sample standard deviation $\sigma y_o(\tau)$ turns out from (10):

$$\sigma y_o(\tau) = \sigma y_R(\tau) + \frac{\nu_{IF}}{\nu_G} \sigma y_{IF}(\tau) \quad (11)$$

From the above expressions it can be seen that the fluctuation spectrum of the controlled oscillator reproduces the reference one only if the residual IF fluctuations are made unimportant. To evaluate the IF and output spectra as given by (9) and (10) it is necessary to specify the order of the phase-lock; in our experiment we have implemented the following loop-filter transfer function:

$$F(j\omega) = \frac{(1+j\omega\tau_2)(1+j\omega\tau_z)}{j\omega\tau_1} \quad (12)$$

with $\tau_z = \tau_G$ to compensate the pole of the Gunn modulation bandwidth (1).

We realize in such a way a second-order loop, which is the minimum required to control a flicker-noise spectrum with unconditional dynamical stability; in this case the loop may be described by the natural frequency f_n and by the damping factor ζ :

$$f_n^2 = K_{oo} K_D / 4\pi^2 \tau_1 ; \quad \zeta = \pi f_n \tau_2 \quad (13)$$

where K_D is the PD gain factor.

Moreover the main equations (7) and (8) are valid for a linear control loop: the mean-square value of the phase fluctuations at the discriminator output $\overline{\phi_e^2}$ must satisfy the following condition:

$$\overline{\phi_e^2} \ll 1 \text{ rad}^2$$

Taking into account (6) and (9), we can write for the second-order loop:

$$\overline{\phi_e^2} = \overline{\phi_{IF}^2} = \int_o^\infty S_{\phi IF}(f) df = \frac{\pi h_{-2} \nu_G^2}{4\zeta f_n^3} + \frac{f(\zeta) h_{-1} \nu_G^2}{2 f_n^2} + \frac{\pi h_o \nu_G^2}{4\zeta f_n} \quad (14)$$

being f(ζ) the flicker-variance factor |9|.

Finally the Allan variance of the IF signal $\sigma^2 y_{IF}(\tau)$ is related to the phase spectrum through the well-known expression:

$$\sigma^2 y_{IF}(\tau) = \frac{2}{\nu_{IF}^2} \int_o^\infty f^2 S_{\phi IF}(f) \frac{\sin^4 \pi f \tau}{(\pi f \tau)^2} df \quad (15)$$

Introducing in (15),(4b),(9), the definition of H(jω) |9|, (12) and (14) we obtain:

$$\frac{\nu_{IF}^2}{\nu_G^2} \sigma^2 y_{IF}(\tau) = \frac{3 \overline{\phi_e^2}}{4\pi^2 \nu_G^2} \tau^{-2} \quad (16)$$

From expressions (14) and (16) it turns out that ζ and f_n must be chosen as large as possible; in order to avoid an unpractically wide loop bandwidth and to guarantee a sufficient phase margin we have implemented the following values:

$$\zeta = 1.5, \quad f_n = 150 \text{ kHz}$$

Inserting the above values in (14) and (16) we have:

$$(\overline{\phi_e^2})^{1/2} = 0.13 \text{ rad}$$

$$\frac{\nu_{IF}}{\nu_G} \sigma y_{IF}(\tau) = 3.6 \times 10^{-13} \tau^{-1}$$

The observed experimental values are:

$$(\overline{\phi_e^2})^{1/2} = 0.2 \text{ rad}$$

$$\frac{\nu_{IF}}{\nu_G} \sigma y_{IF}(\tau) = 6 \times 10^{-13} \tau^{-1}$$

in agreement with the theoretical predictions; the linearity condition is satisfied and the residual frequency fluctuations of the Gunn oscillator are unimportant compared to the reference instabilities |see (11)|.

The observed spectrum of the IF signal, with an observation bandwidth of 10 kHz, is reported in fig. 8; this spectrum shows a carrier-to-noise ratio not

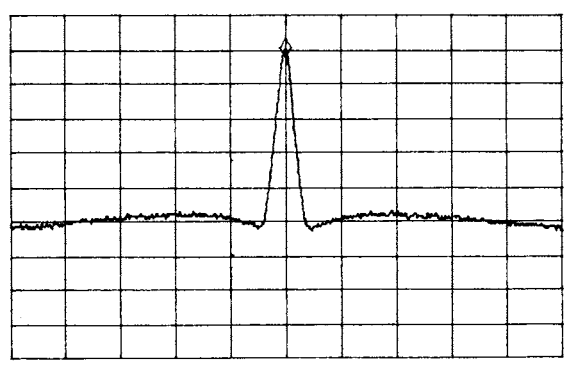


Fig. 8 - RF spectrum of the IF signal in closed-loop conditions; hor.: 10 kHz/div, vert.: 10 dB/div, BW 1 kHz; ν_{IF} = 30 MHz.

worse than the value predicted in fig. 3 for the reference signal. Taking into account equations (10),(11) and the experimental data we can conclude that the spectrum of the locked-oscillator reproduces the spectrum of the reference quartz oscillator multiplied up to 100 GHz, without any further degradation; the same conclusion is valid for the submillimetric output at 601 GHz. The results here obtained are in agreement with those reported in the literature on millimetric Gunn frequency stabilization |10,11|.

The application to the Mg frequency standard

We present in this paragraph the application of the solid-state source described so far to the Mg beam frequency standard. In all the previous experiments the $^3P_1 - {}^3P_0$ $\Delta m_j = 0$ ^{24}Mg clock transition was excited by a BWO operating at 601 GHz and phase-locked to a quartz oscillator through a frequency multiplication chain, as described in |6|. Stray magnetic fields of the BWO and phase-lock slipping cycles were recognized as the main uncertainty sources which limit the Mg standard accuracy at 10^{-11} level |2|. The above problems have been eliminated substituting the BWO with the solid-state source; the atomic beam section of the frequency standard is the same described in |2|.
The record of the Ramsey fringes obtained in this case is shown in fig. 9; the equivalent video

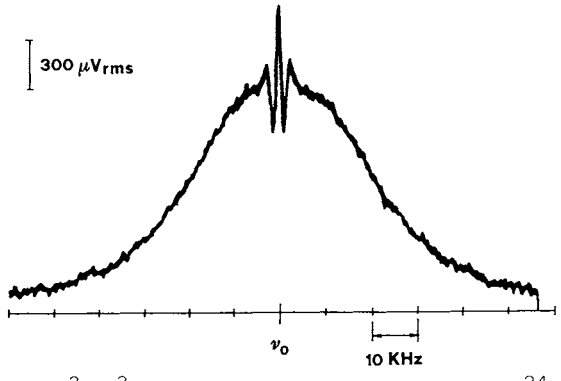


Fig.9 - $^3P_1 - {}^3P_0$ $\Delta mj = 0$ Rabi-Ramsey signal of ^{24}Mg.

bandwidth is 0.1 Hz, the sweep speed is 200 Hz/s and the RF power 1/3 of the optimum value. The 5 MHz quartz oscillator which drives the solid-state multiplication chain is locked to the peak of the central Ramsey fringe ($\Delta\nu = 1.2$ kHz, $Q = 5\times10^8$) via a 1st harmonic control-loop with a time constant T = 100 ms; the RF radiation at 601 GHz is sinusoidally frequency modulated at $f_m = 127$ Hz with peak-to-peak frequency deviation $\Delta\nu_{pp} = 1.4$ kHz. A quantization field $B_o = 65$ μT is applied in the interaction region for the observation of the π-transition ($\Delta m_j = 0$).
The measurement of the ^{24}Mg frequency ν_o has been performed in 5 different days; each day a value has been obtained by averaging a set of 30 different measurements, each of them being the mean of 10 frequency readings with a counting time $\tau = 10$ s. The zero-field extrapolation has been obtained by fitting the experimental Zeeman curve in the range $\Delta B_o = \pm 300$ μT; finally by applying the 2nd order Doppler correction we have obtained the following estimate of the $^3P_1 - {}^3P_0$ $\Delta m_j = 0$ ^{24}Mg transition, with respect to UTC(IEN):

$$\nu_o = 601277157868.4 \pm 2.6 \text{ Hz}$$

The uncertainty reported ($\pm 4\times10^{-12}$) is a conservative evaluation of the different contributions to the accuracy budget, whose estimation is now under way: 2nd order Doppler shift, 2nd order Zeeman shift, residual first order Doppler shift, Ø-shift, servo-loop electronics, cavity pulling, etc....

Conclusions

In this paper we have reported the extension of frequency synthesis up to the sub-millimeter wave range, with an all solid-state multiplication chain, with an output power suitable to excite the $^3P_1 - {}^3P_0$ transition in the metastable triplet of Mg. Moreover we have remeasured the Mg clock transition frequency with respect to UTC(IEN) with an uncertainty of $\pm 4\times10^{-12}$, an order of magnitude better than the previous measurement employing a BWO as a reference source.

References

|1| J. Lian, M. Gross, P. Goy and S. Haroche, "Circular Rydberg-state spectroscopy", Phys. Rev. A, vol. 33, pp. 437-4439, 1986.

|2| A. Godone, E. Bava and C. Novero, "High resolution frequency measurement of the $^3P_1 - {}^3P_0$ ^{24}Mg transition", Metrologia, vol. 24, pp. 133-138, 1987.

|3| A. De Marchi, A. Godone and E. Bava, "Phase-coherent synthesis and precision frequency measurements in the far-infrared", IEEE Trans. Instrum. Meas., vol. IM-30, pp. 132-138, 1981.

|4| A. Godone, M.P. Sassi, C. Caldera and E. Bava, "Frequency measurement of the NH_3 81.5 and 263.4 um lines", Opt. Communic. vol. 71, pp. 59-60, 1989.

|5| C.O. Weiss, "Frequency measurement chain to 30 THz using FIR Schottky diodes and a submillimeter backward-wave-oscillator", Appl. Phys., vol. B34, pp. 63-67, 1984.

|6| A. Godone, E. Bava and C. Novero, "Phase-lock of submillimetric backward-wave-oscillators", IEEE Trans. Instrum. Meas., vol. IM-38, pp. 794-798, 1989.

|7| A. Clairon, B. Dahmani, A. Filimon and J. Rutman, "Precise frequency measurements of CO_2/OsO_4 and $HeNe/CH_4$ - stabilized lasers", IEEE Trans. Instrum. Meas., vol. IM-34, pp. 265-268, 1985.

|8| Millitech Corp. Mod. MUS-601-6 (1990).

|9| F.M. Gardner, "Phase-Lock Techniques", 2nd ed., New York: Wiley, 1979.

|10| R.S. Arora and R. Becker, "New digital phase-lock system for millimeter-wave Gunn oscillator used in radioastronomy", Electronics Lett., vol. 24, pp. 1286-1287, 1988.

|11| B. Komiyama and T. Matsui, "Phase-lock of millimeter-wave Gunn oscillators", 409th inner CRL conference, 1990.

DESIGNING FOR FREQUENCY AND TIME METROLOGY AT THE 10^{-18} LEVEL*

F. L. Walls, L. M. Nelson, and G. R. Valdez
Time and Frequency Division
National Institute of Standards and Technology
Boulder, CO 80303

Abstract

This paper examines some of the key parameters that significantly affect the overall architecture of a system that is being designed to measure frequency or time accuracy to 3 x 10^{-18}. Specifically, we investigate the timing errors in signal transmission for distances up to 100 m, the effect of changes in temperature and rf amplitude on timing (or frequency) errors in several available phase detectors, and the influence of measurement algorithms on the allowable level of white phase noise in the system as a function of the carrier frequency.

frequency stability; phase noise; timing errors; timing stability

Introduction

Several new frequency standards have been proposed that show a potential for a short-term frequency stability of order 2 x $10^{-15}\tau^{-1/2}$ [1-4]. To measure the systematic shifts as a function of operating parameters and to measure the performance of these new clocks, it is desirable to have a measurement system that reaches the clock performance at a measurement time no longer than 10^4 s. Attaining this requires a resolution of approximately $\sigma_y(\tau) = 2 \times 10^{-17}$ at a measurement time of 10^4 s, or a timing error of 0.2 ps. Figure 1 shows the primary elements in such a measurement system.

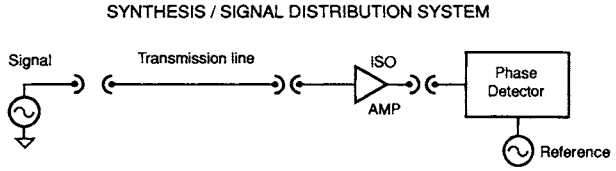


Fig. 1 Overall block diagram showing the key elements for measuring the frequency or time of a remote clock.

In this paper we investigate the timing errors in signal transmission for distances up to 100 m, the effect of changes in temperature and rf amplitude on timing (or frequency) errors in several available phase detectors, and the influence of measurement algorithms on the allowable white phase noise in the system as a function of the carrier frequency.

*Contribution of the U.S. Government, not subject to copyright.

Timing Errors in Signal Transmission

Potentially serious timing errors can be caused by the variation of transit time delay and the phase of the standing waves on the transmission line connecting the signal (clock) to the measurement system. The transit time delay of a signal is given by the length divided by the velocity of propagation. Signals transmitted to the measurement system and reflected from the load and then from the signal source lead to phase change of the signal at the load. Changes in the voltage standing wave ratio (VSWR) therefore lead to changes in the timing errors. The complete timing error from the transmission is

$$\delta t = L/(\beta c) + \rho_l \rho_s \sin\phi/(4\eta\nu_0) \qquad (1)$$

where L is the equivalent free-space length of the transmission line, c is the speed of light, ß is the propagation factor, ν_0 is the carrier frequency, η is the round trip attenuation, ρ_l is the voltage reflection ratio from the load, ρ_s is the voltage reflection ratio from the source (ρ = (VSWR - 1)/(VSWR + 1)), and ϕ is the angle of the twice-reflected wave at the load relative to the primary signal. This equation assumes that the next order of reflected waves can be ignored ($\rho_s\rho_l/\eta << 1$). The first term is independent of frequency while the second term scales as $1/\nu_0$.

The variation in the timing error due to changes in L/(ßc) is potentially quite large for long transmission cables. Typical variations for several types of semi-rigid coaxial cables versus temperature are shown in Fig. 2 [5]. Typical insertion losses are of the order 1 dB/30 m at 5 MHz and 3 dB/30 m at 100 MHz for a 0.358 cm diameter cable. Cables with solid Teflon dielectric (Curve A) exhibit large hysteresis (due to the mismatch between the thermal expansion coefficient between the dielectric and the conductors) for temperature variations from 10°C to 20°C. Various attempts have been made to produce cable with a smaller temperature sensitivity by reducing the amount of Teflon dielectric or incorporating compensating materials (See Curves B-E of Fig. 2). Curve F of Fig. 2 shows the temperature sensitivity of a cable that uses SiO_2 powder as the dielectric. The phase variation from 18.3°C to 23.8°C is less than ±0.5 ps over 30 m, while the insertion loss is 3.2 dB/30 m at 100 MHz. This generally satisfies the timing requirements for distances up to about 100 m. The primary dissadvantage of this cable is the cost. Except in cases where there are large temperature variations,

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

one of the other cables from Fig. 2 usually performs acceptably at less cost for distances up to approximately 3 m. These data are summarized in Table 1.

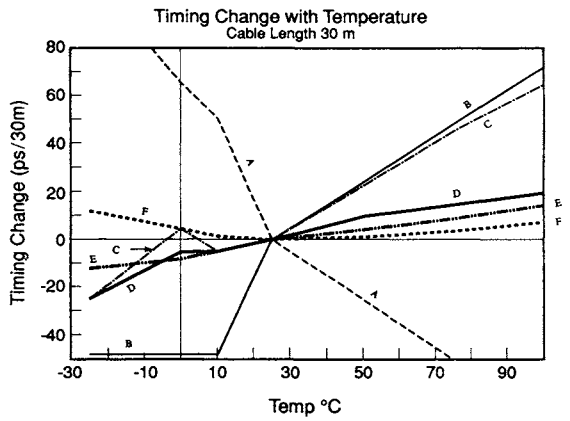


Fig. 2 Timing error versus temperature for several different semi-rigid coaxial cables. Curve A, typical 0.358 cm (0.141") solid Teflon dielectric cable. Curve B, data for 0.483 cm (0.190") low density Teflon cable from W. L. Gore. Curve C, data for 0.358 cm cable with low density Teflon cable from Precision Tube. Curve D, data for 0.358 cm air-articulated cable from Precision Tube. Curve E, phase shift data for 0.358 cm ISOCORE cable from Rogers with the temperature coefficient of the dielectric matched to the conductors. Curve F data for 0.358 cm cable with SiO_2 dielectric from Kaman Instruments [5].

At 5 MHz, $1/(4\ \nu_o)$ is 50 ns and numbers for ρ_ℓ and ρ_s are typically 1/3 (VSWR = 2). Such numbers can lead to large variations of timing error since VSWR is often dependent on temperature, rf drive level, and even supply voltage. Timing errors due to VSWR can be significantly improved by making ρ_ℓ, ρ_s, and sin ϕ small. At 5 MHz, for example, the timing error due to VSWR can be reduced below 11 ps by reducing VSWR below 1.1 at the source and load, and reducing sin ϕ below 0.1. At 100 MHz this term can be made less than 1 ps by making VSWR < 1.1 and sin ϕ < 0.1. The characteristic impedance of coaxial cables changes slightly from one batch to another, which might require a small adjustment in source and load impedances for the most critical applications.

An alternative approach to the problem is to use two-way transmission measurements to correct for variations in the transmission medium. This technique has been pursued at the Jet Propulsion Laboratory where researchers have concentrated on AM modulation of a laser transmitted over fiber optic cables to solve both the attenuation of the signal and timing errors for cable lengths up to 25 km. Their results [6], shown in Fig. 3, are very impressive. Although this system is relatively costly, it is currently the only practical approach for long transmission distances due to the problem of attenuation.

Comparison of Measurement Algorithms

The minimum signal-to-noise ratio (S/N) that is consistent with the goals is set by the measurement algorithm and the time available to average the data. High precision measurement systems are generally dominated by white phase noise in the short-term, flicker phase, and environmentally driven offsets in the medium term, and bounded processes in the long term. If $\sigma_y(\tau)$ is used as the measure of frequency stability, the observed stability depends on the level of white phase noise, the level of white frequency noise, the measurement bandwidth, and the measurement time and does not depend on the number of samples [7]. If mod$\sigma_y(\tau)$ is used as a measure of frequency stability, the observed frequency stability also depends on the number of measurements averaged together to obtain mod$\sigma_y(\tau)$ [7].

The added phase noise from single and double balanced mixers is flicker phase noise (usually $S_\phi(f) = 10^{-14}/f$ in units of rad^2/Hz) for a Fourier frequency, f, from 0.1 to 1 kHz and white phase at a level dependent on the IF amplifier at higher frequencies. Typical levels are $10^{-16.5}$ to $10^{-17.7}$ rad^2/Hz [8]. The added phase noise at Fourier frequencies below 0.01 Hz is often dominated by systematic and environmental effects which are discussed in later sections. Most practical clocks have white frequency noise far above that of the mixer.

Figure 4 shows the frequency stability as determined by $\sigma_y(\tau)$ and mod $\sigma_y(\tau)$ as a function of measurement time, and carrier frequency for $S_\phi(f) = 8 \times 10^{-30}\nu_0^2/f^2$ (white frequency) + $10^{-14}/f$ (flicker phase) + $4 \times 10^{-27}\nu_0^2$ (white phase) rad^2/Hz and a measurement bandwidth of 1 kHz [8]. The first term

TABLE 1. Timing errors due to thermal variations in cable length

BASIC TYPES OF CABLES	Loss/30 m @ 100 MHz DIAMETER	σT/K 23-30°C	APPROXIMATE COST/30 m
SOLID DIELECTRIC PTFE common semi-rigid	3.4 dB 0.385 cm	1 ps/(30mK)	$ 210
MICROPOROUS W. L. GORE	2.8 dB 0.483.cm	1 ps/(30mK)	$4000
LOW DENSITY PTFE PRECISION TUBE	2.4 dB 0.385 cm	0.9 ps/(30mK)	$ 550
AIR ARTICULATED PTFE PRECISION TUBE	2.7 dB 0.385 cm	0.4 ps/(30mK)	$ 950
COMPENSATED DIELECTRIC ISOCORE ROGERS	3.5 dB 0.386 cm	0.18 ps/(30mK)	$ 657
SiO_2 DIELECTRIC KAMAN INSTRUMENTS	3.2 dB 0.385 cm	0.06 ps/(30mK)	$2200

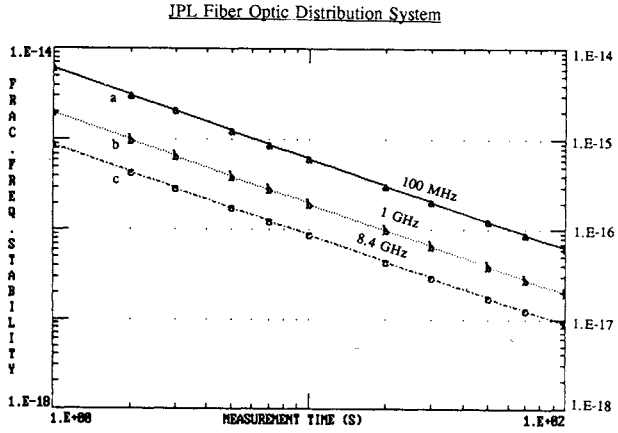


Fig. 3 Fractional timing error versus measurement time (calculated from phase noise measurements) of the JPL fiber optic system for AM frequences of 100 MHz, 1 GHz, and 8.4 GHz. Timing error is $\tau\sigma_y(\tau)$. Adapted from [6]

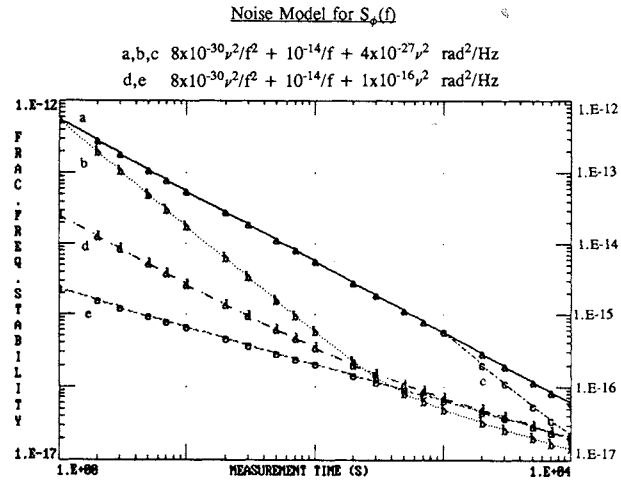


Fig. 4 Fractional frequency or timing error as a function of measurement time and measurement algorithm and input noise. Curves a, b, and c are for $S_\phi(f) = 8 \times 10^{-30}\nu^2/f^2 + 10^{-14}/f + 4 \times 10^{-27}\nu^2$ rad²/Hz and a measurement bandwidth of 1 kHz. Curve a shows the results at 5 and 100 MHz using $\sigma_y(\tau)$ to calculate fractional stability. Curve b shows mod$\sigma_y(\tau)$ at 5 and 100 MHz for a minimum measurement time of 1 s, while curve c shows mod$\sigma_y(\tau)$ at 5 and 100 MHz for a minimum measurement time of 1000 s. Curve d shows $\sigma_y(\tau)$ at 5 MHz and Curve e shows $\sigma_y(\tau)$ at 100 MHz for $S_\phi(f) = 8 \times 10^{-30}\nu^2 + 10^{-14}/f + 10^{-16}$ rad²/Hz.

originates from the assumed intrinsic frequency stability of the clock signal; the middle term originates from the mixer noise; and the last term is white phase noise assumed to originate in the clock or the measurement process. Also shown is the frequency stability at 5 MHz and 100 MHz when the white phase noise is equal to that of the mixer. Note that the frequency stability derived from mod$\sigma_y(\tau)$ is a function of the number of samples averaged to obtain mod$\sigma_y(\tau)$ when flicker phase and/or white phase noise dominates.

This example demonstrates that it is possible to achieve the desired time domain frequency stability even for relatively large levels of white phase noise, measurement times larger than 1000 s, and measurement bandwidths less than 1 kHz at either 5 or 100 MHz, assuming that the medium term noise from the measurement system is small enough. Furthermore, the use of multiple samples to compute the time of a clock reduces the timing error roughly as $n^{1/2}$ for white phase noise, ln(n) for flicker phase noise, and n for coherent signals, where n is the number of measurements averaged together to obtain the phase (time) of the clock signal [7,8]. The medium term noise is most often due to changes in temperature, rf drive levels, humidity, voltage, vibration and mechanical deformation, and pick-up of external signals.

Environmental Sensitivity of Phase Detectors

One measure of phase detector performance is the phase error $\delta\phi$ given by

$$\delta\phi = \delta V/k_d, \tag{2}$$

where δV is the dc output voltage error and k_d is the phase detection slope in V/rad [6]. For comparing different phase detector circuits as a function of carrier frequency, ν_o, it is more beneficial to use the timing error given by

$$\delta t = \delta V/(k_d 2\pi\nu_0) \quad s. \tag{3}$$

The dc voltage error originates from imbalance between the I-V curves of the diodes, imbalance in the transformers, and asymmetric stray impedances. These imbalances are a function of the temperature and drive levels and therefore lead to changes in apparent phase or time with change in these parameters. The most important parameters that characterizes these effects are the isolation between the L and R ports, and the IF port. Table 2 shows approximate dc offsets as a function of L drive and isolation [9].

Table 2. DC offset voltage in mV for various combinations of LO drive level and isolation assuming a 50 Ω resistance on all ports [9].

ISOLATION - dB \ LO DRIVE LEVEL – dBm	0	+3	+7	+10	+13	+17	+20
15	25.3	35.8	63.7	90.0	127.1	226.0	319.2
20	14.2	20.1	35.8	50.6	71.5	127.1	179.5
25	8.0	11.3	20.1	28.5	40.2	71.5	100.9
30	4.5	6.4	11.3	16.0	22.6	40.2	56.8
35	2.5	3.6	6.4	9.0	12.7	22.6	31.9
40	1.4	2.0	3.6	5.1	7.1	12.7	17.9
45	0.80	1.1	2.0	2.9	4.0	7.1	10.1
50	0.45	0.64	1.1	1.6	2.3	4.0	5.7

Figure 5 shows a typical circuit diagram for evaluating the phase detectors. The dc output (IF) port is terminated with the characteristic impedance which is typically 50 Ω for the double balanced mixers and 500 Ω for the single balanced mixer.

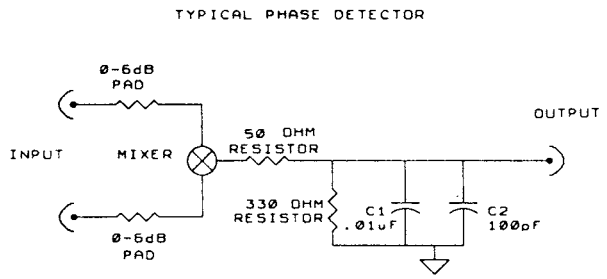


Figure 5. Typical circuit diagram for evaluating the sensitivity of various phase detctors to variations in temperature and rf level.

Figures 6 to 8 show timing errors in selected phase detectors due to variation in input power and for two different temperatures for a carrier frequency of 5 MHz. One set of curves was taken without attenuators and the others with either 3 or 6 dB attenuators on the R and L inputs. At 5 MHz, the mixer with the individually matched diodes has a temperature coefficient of approximately 6 ps/°C when used with 6 dB pads, while the mixers with the 4 diodes on the same substrate exhibit a temperature coefficient of about 25 ps/°C. The sensitivity to power change is generally improved by adding 3 or 6 dB attenuators to the input ports. Without attenuators the power sensitivity is of order 1 ms/ν_0 per dB change in the power or 250 ps/dB at 5 MHz. The improvement due to the use of attenuators on the R and L inputs is thought to originate from the reduction in the cable VSWR generated by the phase detector load.

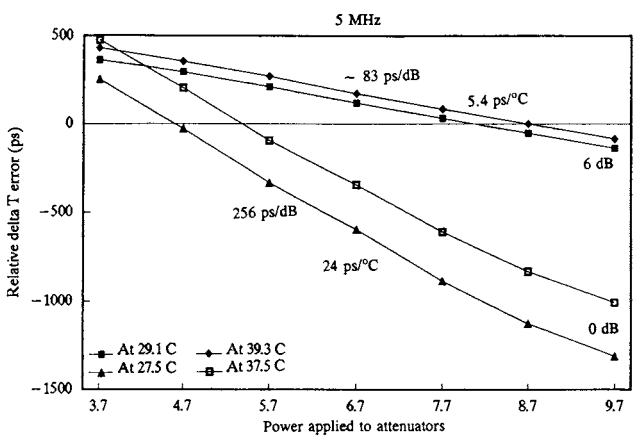


Figure 6. Timing errors in a double balanced mixer with individual diodes at a carrier frequency of 5 MHz as a function of drive power, input attenuators (0 dB and 6 dB), and temperature.

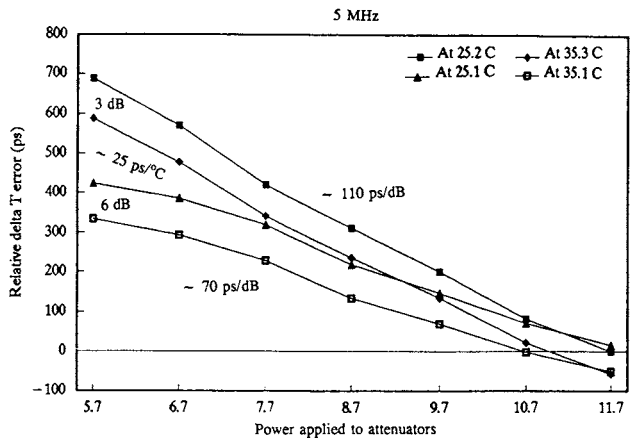


Figure 7. Timing errors in a double balanced mixer with a quad diode module, at a carrier frequency of 5 MHz as a function of drive power, input attenuators (3 dB and 6 dB), and temperature.

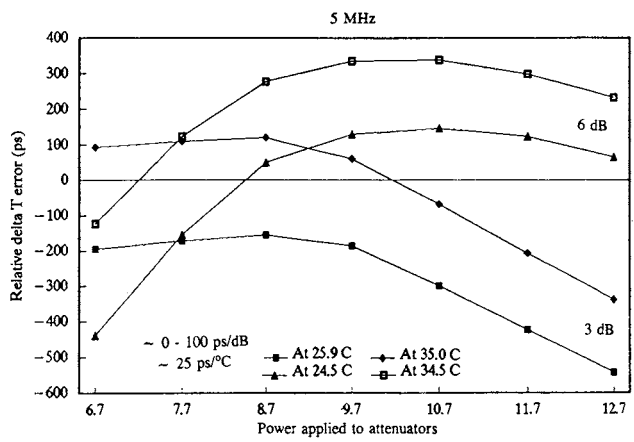


Figure 8. Timing errors in a single balanced mixer with a dual diode module, at a carrier frequency of 5 MHz as a function of drive power, input attenuators (3 dB and 6 dB), and temperature.

Figures 9 to 11 show the timing errors in selected phase detectors due to variations in input power and for two different temperatures for a carrier frequency of 100 MHz. One set of curves was taken without attenuators and the others were taken with either 3 or 6 dB attenuators on the R and L inputs. The performance at 100 MHz of all three phase detectors versus temperature and power is improved over the results at 5 MHz by roughly the ratio of carrier frequencies. Temperature sensitivity is of the order 0.2 to 1 ps/°C. The best temperature performance comes from the mixer with individually matched diodes. The power sensitivity is of the order 0.1 to 1 ms/ν_0 per dB change in power or 1 to 10 ps/dB at 100 MHz depending on mixer type, input power, and attenuation.

Discussion

It is much easier to achieve the desired time domain performance of $\delta\nu/\nu_0 = \delta t/t = 2 \times 10^{-17}$ at a measurement time of 10^4 s for a carrier frequency of 100 MHz than for 5 MHz due to the sensitivity of presently available phase

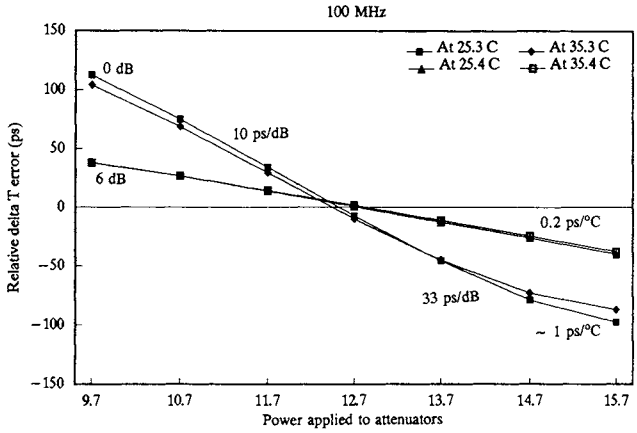


Figure 9. Timing errors in a double balanced mixer with a individual diodes, at a carrier frequency of 100 MHz as a function of drive power, input attenuators (0 dB and 6 dB), and temperature.

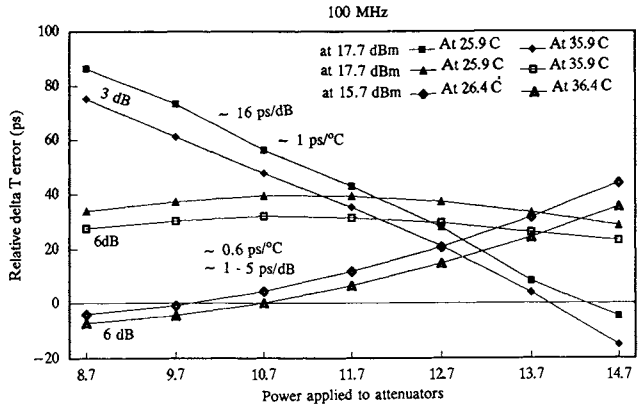


Figure 10. Timing errors in a double balanced mixer with a quad diode module, at a carrier frequency of 100 MHz as a function of drive power, input attenuators (3 dB and 6 dB), and temperature.

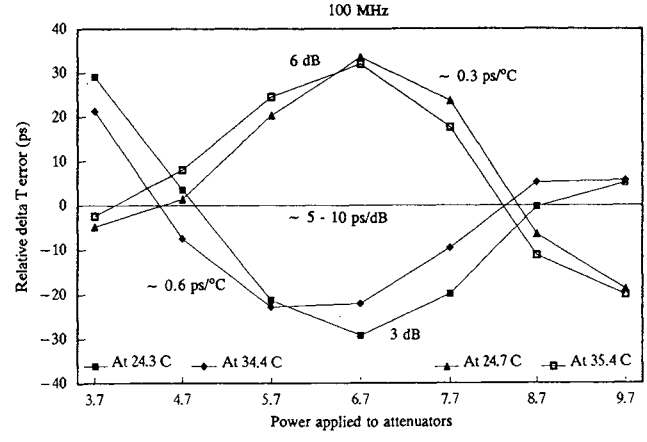


Figure 11. Timing errors in a single balanced mixer with a dual diode module, at a carrier frequency of 100 MHz as a function of drive power, input attenuators (3 dB and 6 dB), and temperature.

detectors to both temperature and rf amplitude. Even at 100 MHz it will probably be necessary to stabilize the temperature of the phase detectors to about 1°C and carefully match the characteristic impedance of the coaxial cable. For distances of up to 100 m, existing coaxial cables are probably sufficient. For longer distances we must consider either AM modulation of a laser transmitted over fiber optic cables or lower frequencies where the attenuation losses are smaller. The fiber optic system has already demonstrated excellent performance for distances up to 25 km and is probably the best choice. Going to 5 MHz would require two-way timing measurements to correct for the time dispersion in the transmission medium, much more effort to stabilize the operating parameters of the phase detector, and lowering the effects of VSWR at the source and the load.

ACKNOWLEDGEMENTS

The authors thank C. Osborn for help in construction of the test systems and the CECOM Center for Space Systems TSSRF, Ft. Monmouth, NJ, for partial support of this work.

References

1. D. J. Wineland, J. C. Bergquist, J. J. Bollinger, W. M. Itano, D. J. Heinzen, S. L. Gilbert, C. H. Manney, and M. G. Raizen, "Progress at NIST toward absolute frequency standards using stored ions," IEEE Trans. on Ultrasonics, Ferroelectrics, and Frequency Control, **37**, 515-523 (1990).

2. J. D. Prestage, G. J. Dick, and L. Maleki, "Linear ion trap based atomic frequency standard," Proc. of 44th Ann. Frequency Control Symp., 1990, pp. 82-88.

3. S. Chu, "Laser manipulation of atoms and particles," Science, **253**, 861-866 (1991).

4. J. L. Hall, M. Zhu, and P. Buch, "Prospects for using laser-prepared atomic fountains for optical frequency standards applications," J. Opt. Soc. Am. B, **6**, 2194-2205 (1989).

5. Certain commercial equipment, instruments, or materials are identified in this paper in order to adequately specify the experimental procedure. Such identification does not imply recommendation or endorsement by the National Institute of Standards & Technology, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose.

6. G. F. Lutes and R. T. Logan, "Status of frequency and timing reference signal transmission by fiber optics," Proc 45th Ann Frequency Control Symp., 1991, pp. 679-686.

7. "Characterization of clocks and oscillators," Eds. D. B. Sullivan, D. W. Allan, D. A. Howe, and F. L. Walls, NIST Tech Note 1337, 1990.

8. $\sigma_y(\tau)$ and mod$\sigma_y(\tau)$ were calculated using the SIGINT software package described in NISTIR 89-3916R. See also "Time domain frequency stability calculated from the frequency domain," Proc. of 4th European Frequency and Time Forum, 1990, pp. 197-204.

9. "Mixers as phase detectors," from 1985 Watkins-Johnson Catalog, pp. 666-673, 1985.

Session 10A

Time and Frequency: Network Synchronisation

Chairman: P. Kartaschoff

TIME GENERATION AND DISTRIBUTION SYSTEM FOR A MILITARY NATIONAL SYNCHRONIZATION

V. Mourier

TEKELEC SYSTEMES
29, avenue de la Baltique
Z.A. de Courtaboeuf
91953 Les Ulis cedex
FRANCE

ABSTRACT

The time generation and distribution system provides homogeneous dating of events throughout the country. It generates and distributes time codes and precise frequency used for high precision measurements, telecommunications, etc. Accordingly the need is to provide on each operationnal site an accurate and stable time reference based on UTC time scale. The system consistes of a common reference (GPS), independant clocks and time distribution that provides time to all the user equipment on each operational site.

Keywords : national synchronization, time distribution, operational applications.

1 - INTRODUCTION

The French Air Force has decided to equip the operational applications for the whole country with the Time Generation and Distribution System.

2 - THE REQUIREMENT

Present and forthcoming military systems require a very precise time management. The following two examples show the importance of a time and frequency reference in these systems.

To be safe from jamming, the aircraft identification systems use coded data. Since the change of codes for secure transmission depends on the precise time, all transponders and interrogators must have synchronized clocks.

The anti-jamming radiocommunication systems use pseudo-random frequency changes. Since the frequency used at a particular instance depends on the precise time, all terminals of such radiocommunication links must have synchronized clocks.

These examples put the emphasis on the need to synchronize equipment and operational systems (also navigation systems, multi-radar tracking, digital networks, etc.).

To cope with the issue of autonomy, the system need to manage accurate clocks in order to be independant of continuous synchronization.

3 - THE CONCEPT

The concept put into service by the French Air Force is to have a precise, reliable and common time reference on each operational site.

- Precise to guarantee the proper operation of systems and communication between systems.
- Reliable for operational availability.
- Common to insure the general coherence of the system.

The chosen reference is the UTC time scale in accordance with international standards and other nations' choice (NATO especially).

The concern for the reliability and integrity of the national system leads to the deployment of locally autonomous and independant systems to avoid wide spread disruption as a result of a local fault.

The need of the user systems set the uniform accuracy to about ± 1 ms between each site or system and the UTC time scale. We will see that the chosen solutions allow better capabilities.

The context evoked leads the French Air Force to build the Time Generation and Distribution System. This system is based on about 50 time reference platforms shared on operational sites and also the time distribution means. Time platforms are composed of 4 primary equipments with high autonomy and 45 secondary equipments. Figure 1 illustrates this concept.

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

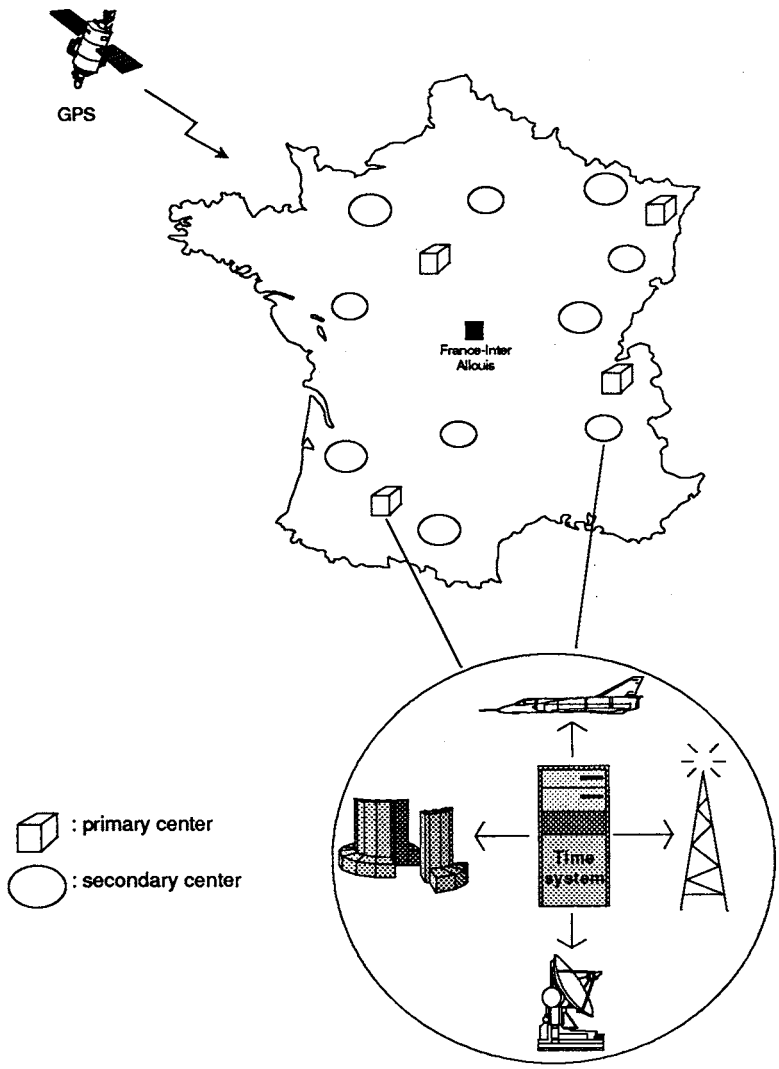


Figure 1 - French Air Force synchronization concept

4 - THE PRINCIPLE

Every time system uses three components :

- acquisition of time reference
- time keeping with high autonomy
- distribution of time and frequency to users

These three components are specially defined to meet the user requirements. A special concern is brought to reliability, evolution and use of standards. Figure 2 explains the architecture.

4.1 - Acquisition

The requirement to have in each country site a precise and UTC referenced synchronization path leads to use of the Global Positionning System (GPS) as the primary source.

Time-specialized GPS receivers used bring a 1pps signal with an accuracy better than ± 1 µs from the UTC. It is accurate enough considering the user needs and the system autonomy.

However, in order to have a back-up reference in case of GPS unavailability, two other ways of synchronization have been defined.

The first one uses an external synchronization mean composed of a time transfer equipment. This system, after synchronization using a national laboratory (Observatoire de Paris, etc.) allows the successive synchronization of the national sites.

This system guarantees a good overall integrity and a national control (time scale and equipment transfer).

The second system is constituted of a VLF receiver that works on the national broadcasting radio system (France-Inter 162 kHz). Considering its lower precision (a few milliseconds for commercial receivers), it is only used when the two other methods are unavailable.

4.2 - Time keeping

The aim is to keep a precise time according to the kind of operational site.

For primary centres where a high availability is needed, time systems have been designed with redundancy.

For secondary centres and for quantity and cost reasons, systems have been designed without redundancy.

- Primary systems :

Redundant frequency generation is assured by a Cesium standard and a slaved BVA oscillator. A monitoring unit switches on one of the two signals (Cesium or BVA) in case of loss of the other one.

The time is generated by 3 time code generators with a majority vote switch.

- Secondary systems :

Frequency is generated by a BVA oscillator slaved on GPS signals. Filtering with long time-constants minimizes the effects of eventual GPS disruption.

For both systems, a measurement of the difference between the 1pps generated and the GPS 1pps is processed. The 24 hours average difference sets under certain conditions a time resetting to compensate the system drift.

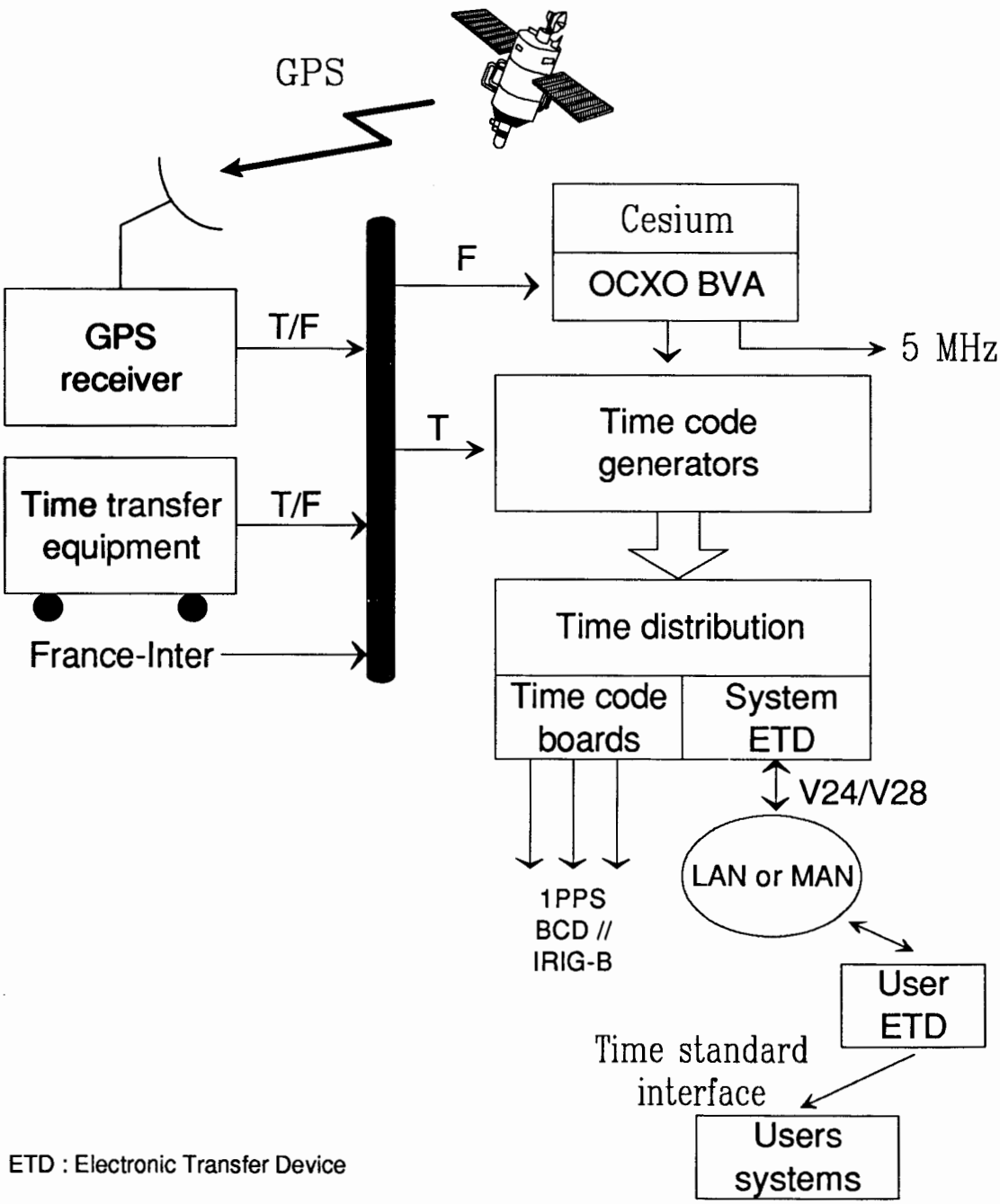


Figure 2 - Time Generation and Distribution System

4.3 - Time Distribution

Time and frequency reference signals - 1pps and 5 MHz sine from the system oscillator - are distributed to all the applications. They are the basic signals that allow the user to synchronize.

This information is completed with a time dating message. For civil or non-operational applications, time message is provided with IRIG-B code (completed with year to adapt to the French NF S 87-500 standard). This type of code does not guarantee information to users and needs a specific transmission equipment.

For operational applications, a digital transmission method for time information is defined. It uses a specific protocol between the time generation equipment and an equipment (ETD : Electronic Transfer Device) located near the users (for distant distribution). By this method, the ETD is synchronized by the generation system with an accuracy of about ± 100 µs.

This time distribution is independant of the specialized link used (wire, optical, radio, etc.). The precise time keeping in the user ETD protects systems against distribution network disruptions.

Time is generated by the user ETD as a digital message including a check-sum. It is composed of : time of day (seconds to year), management of the leap second and a time quality figure (maximum estimated difference between the offered time and UTC time). This message is in accordance with a NATO standardization (STANAG) currently under preparation.

Modular structure of the user ETD allows the distribution of time with other time codes (IRIG-B, BCD //, etc.).

Figure 3 illustrates the time distribution on an operational site such as an air base.

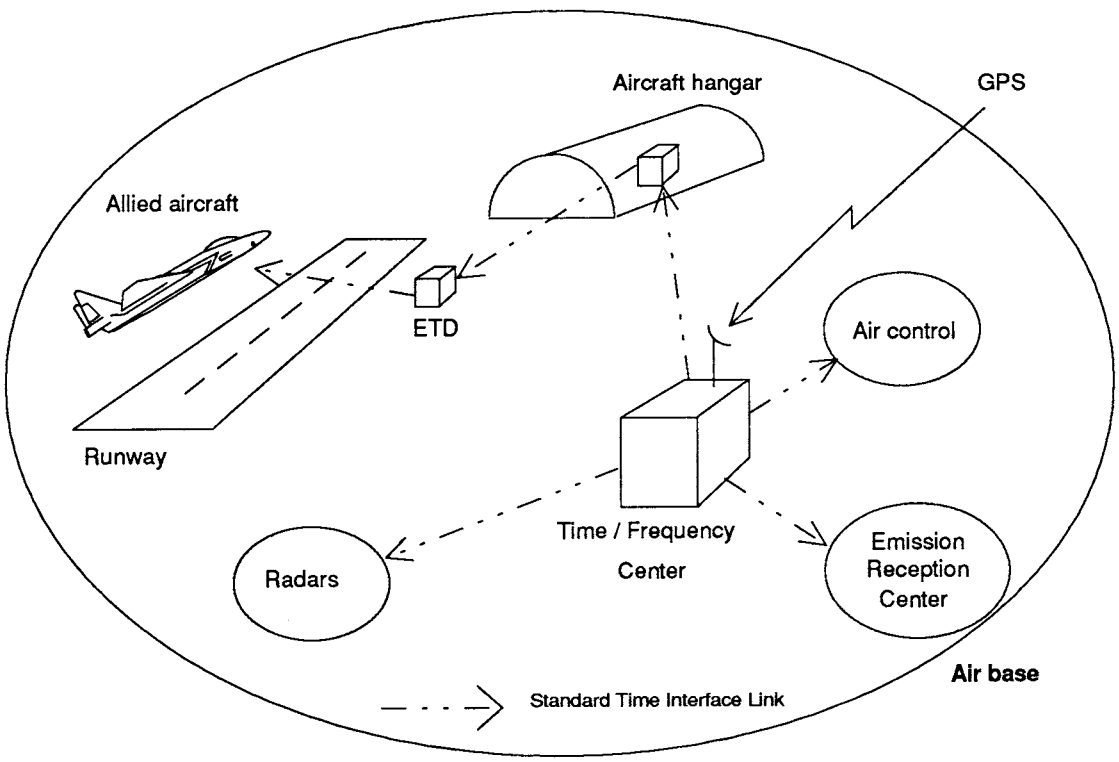


Figure 3 - Time distribution concept

The Time Generation and Distribution System defined by the French Air Force insures in the whole country a precise and homogeneous time that meets the operational applications needs.

Modular structure of the system enables continuous up-dating to take care of new applications and techniques.

The author would like to thank the French Air Force for its support during the developpement of the Time Generation and Distribution System.

TIMING AND TIME SIGNAL DISTRIBUTION IN DIGITAL COMMUNICATIONS NETWORKS

Masami Kihara and Atushi Imaoka

NTT Transmission Systems Laboratories
1-2356, Take, Yokosuka-shi, Kanagawa, 238-03 Japan
Tel:+81-468-59-3164
Fax:+81-468-59-3197
e-mail:kihara@ntttsd.ntt.jp

Abstract

The timing signal distribution characteristics of a digital communications network are evaluated to determine the Maximum Time Interval Error (MTIE) of the network; reference is made to the performance of network components such as transmission systems, slave clocks and timing distribution systems in intra-offices. The MTIE of each component is measured and used to determine the allowable MTIE of that component. The maximum number of slave node chains is shown to be 20.

Time signal distribution performance is also described in detail. It is shown that time synchronization accuracy is of the order of sub-micro seconds between nodes separated by 2400 km over a two year period. For intra-office time signal distribution, the relative time accuracy is less than 3 nano seconds using an 8 Mb/s round-trip digital interface to connect a time signal supply in an office to dispersed equipment.

Keywords: timing distribution, time distribution and network synchronization

1. INTRODUCTION

Timing signals are normally distributed in digital networks using the master-slave synchronization method to synchronize digital cross-connect and switching systems. The existing NTT network now has more 400 digital nodes that receive timing signals. This number will greatly increase with the aggressive installation of Synchronous Digital Hierarchy (SDH) systems[1], particularly if the timing signal is transferred via transmission clock rates that are multiples of 155 Mb/s as specified in the Network Node Interface (NNI). The chaining of timing distribution nodes degrades the jitter and wander performance of timing signals. The maximum jitter and wander is specified in several CCITT (International Telegraph and Telephone Consultative Committee) recommendations[2]-[4].

There are two ways to cope with this degradation of jitter and wander. First is to reduce the number of timing distribution nodes that are chained. AT&T adopts this approach; the reference timing signal is transferred from a GPS satellite instead of the master node[5]. Second is to improve the performance of network components such as transmission systems for timing distribution , slave clocks and the master clock[6]. NTT's timing distribution system is based on the second approach. NTT's system and its characteristics are described. The characteristics are evaluated by the Maximum Time Interval Error (MTIE). We also present a positive wander reduction method for the second approach which measures wander and then compensates it. This method can achieve both precise timing and time distribution. Intra-office time signal distribution characteristics are also shown.

2. TIMING DISTRIBUTION SYSTEM

2.1 Network Synchronization Stream

Timing signal distribution systems usually consist of a unique master node and slave nodes. Cesium atomic standards (the reference clock) generate reference timing signals in the master node. The slave clock in each slave node synchronizes its own timing signal to the transferred reference timing signal. This process is repeated according to the reference timing transfer hierarchical tree. Synchronous bit rate frequency components of transmission paths such as 1.5Mb/s or 2Mb/s are conventionally used as the timing distribution path to transfer the reference timing signal. Since SDH has been recommended in CCITT as a higher synchronous hierarchy, multiples of 150 Mb/s in the NNI can be used to form the synchronous bit rate frequency.

The full introduction of the SDH systems will require the synchronization of all hierarchical bit rates including transmission line frequencies to multiples of 150 Mb/s, such as 600Mb/s and 2.4Gb/s. This implies that reference timing distribution path termination will coincide with transmission line termination. Therefore, SDH installation will increase the number of links in the reference timing transfer hierarchical tree. If no remedial

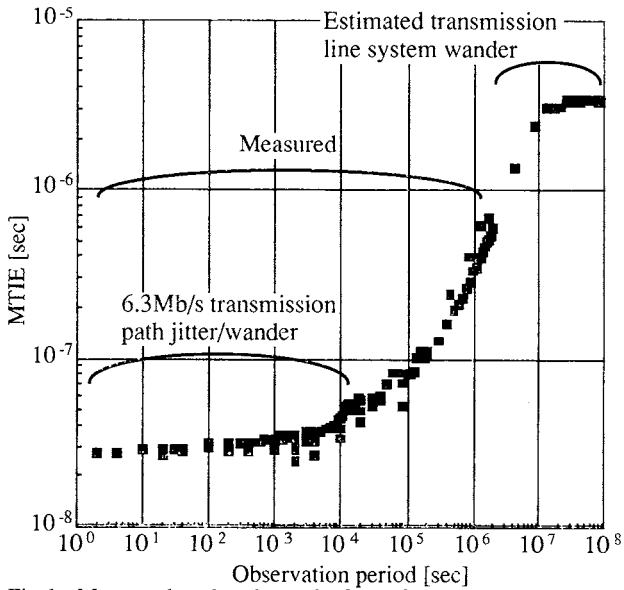


Fig.1 Measured and estimated MTIE for transmission delay variation

measures are taken, jitter and wander in the reference timing signal will certainly exceed CCITT specifications.

2.2 Time Interval Error of the Present Network

2.2.1 Transmission line systems

Wander, transmission delay variation, is induced by the transmission length variation of each timing distribution path. The current transmission delay variation rate is approximately 80ps/km/°C in lightwave line transmission systems. The Maximum Time Interval Error (MTIE) of this variation is estimated to be 4μs for NTT's network. An example of MTIE measurement for reference timing signal transfer over 2400km is shown in Fig.1. MTIE values were determined by analyzing the phase-time variation measurement results shown in Fig.4 and estimating the effects of the Japanese climate over a 10 year period. This wander can be reduced only by the proposed method which will be described in 2.3.

Jitter/wander in pulse justification multiplex/demultiplex systems appears over observation periods less than 10^4 s in Fig.1. Jitter/wander in a transmission path is reduced by selecting a higher bit rate or a higher synchronous hierarchical bit rate.

2.2.2 Slave clocks

Jitter/wander generated by the slave clocks is dominant in chained timing distribution systems, especially when more than 10 links form one chain. To suppress accumulated jitter/wander, the accuracy of tracking the transferred reference timing signal should be improved. One important point in designing a slave clock is the time constant of the phase-locked loop (PLL) used to synchronize the slave to the reference. A time constant of around 1000s can minimize jitter/wander if a Rubidium atomic standard

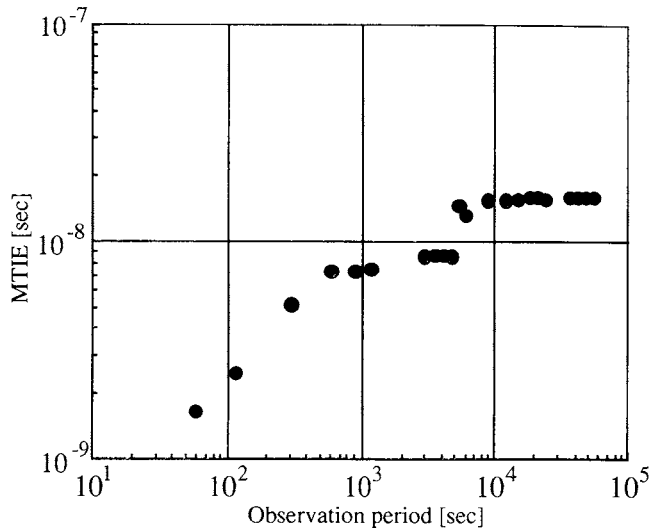


Fig. 2 Measured MTIE for a slave clock

is installed as the frequency controlled oscillator of the PLL. Much longer time constants are by no means better in terms of jitter/wander accumulation. Usage of a Rubidium atomic standard is now practical because the standard has become the same size and cost as highly stable quartz crystal oscillators although its frequency stability is much higher.

Once the optimum time constant is chosen, the quantized control of the PLL is the major source of phase-time variation. MTIE measurement results of an installed commercial slave clock is shown in Fig.2. The measured slave clock can detect a phase-time difference of 8ns and has a maximum frequency control resolution of 1×10^{-12}.

There are other jitter/wander sources related to the slave clocks in intra-nodes. One is the equipment that extracts the reference timing signal component transmitted from the higher hierarchical node and passes the extracted information to the lower hierarchical node. Several PLL circuits are used to synthesize the transmission bit rate in the slave clock and other equipment. These circuits cause phase variations of several tens of ns as determined by MTIE. Another source is the interface between the reference timing signal extraction equipment and the slave clock, and between the slave clock and the reference timing signal transmission equipment. If the interface is of the order of MHz, MTIE is much less than 10 ns; however, the MTIE of a 64kHz interface is around 50ns due to the balance circuit which is composed of a pair cable and a transformer. The other source is the additional equipment used to multiply and to supply the timing signals to the many digital systems within the node. The additional timing supply uses interfaces to receive the reference timing signal from the fundamental slave clock and to transfer it to each equipment, as well as to the fundamental slave clock. The MTIE of the additional timing supply is degraded to 50ns.

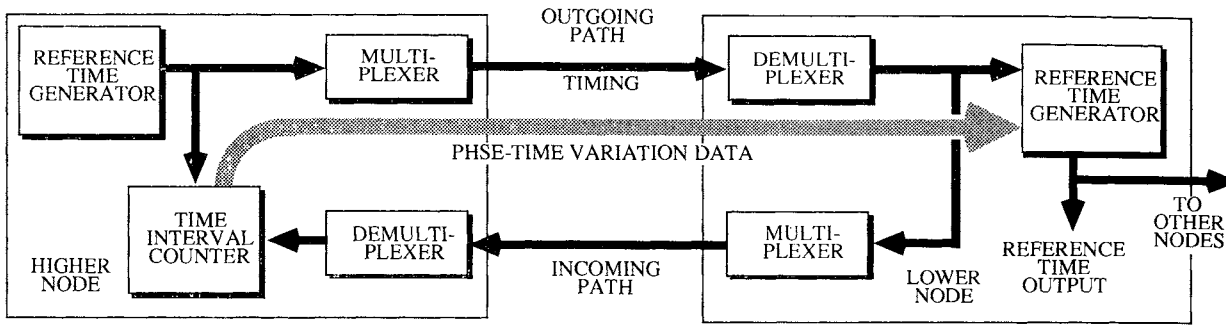


Fig.3 The basic configuration of the transmission delay compensation

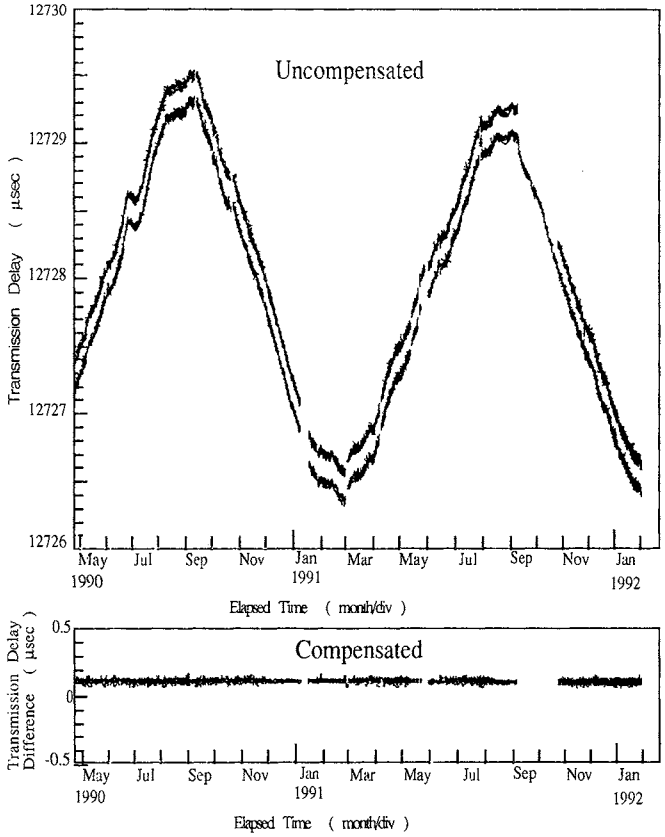


Fig.4 Measured transmission delay variation and its compensated result

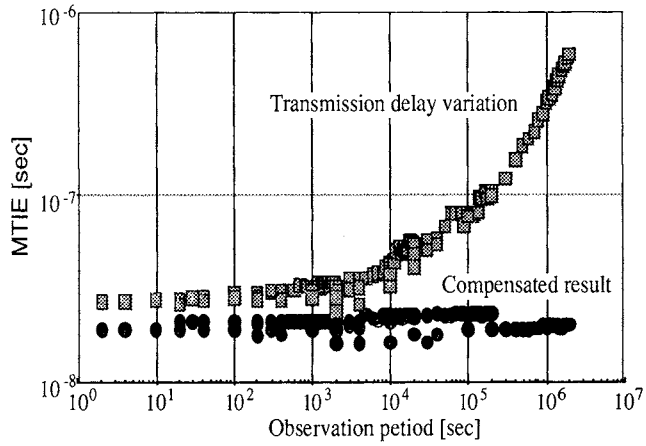


Fig.5 Measured MTIE for compensated transmission delay variation

2.2.3 Total MTIE in national networks

In designing a national synchronization scheme, the frequency deviation of the cesium atomic standard can be excluded from MTIE specifications in CCITT Recommendation. The most stringent MTIE is 10μs over a long observation period. 10μs is allocated as follows:

transmission path MTIE : 4μs

The MTIE of 10 chained slave clocks including intra-node degradation : 10x0.3μs

the reference clock MTIE based on CCITT recommendation : 3μs.

These values have some margin of safety for worst case design and security. The current maximum number of chained slave nodes is 10.

2.3 Wander Reduction with Transmission Delay Compensation System

Positive wander compensation is needed to allow 20 or more slave nodes to be chained together. A transmission delay compensation method using round-trip paths is very simple and effective in determining transmission delay and reducing wander in large terrestrial digital networks. This method, shown in Fig.3, uses both incoming and outgoing paths to transfer the timing and phase-time variation data. Figure 4 shows measured results of uncompensated and compensated phase-time variations over a 2 year period for an installed transmission path of 6.3Mb/s. Although the basic phase-time variation is about 3μs, the compensated value is within 100ns. The MTIE determined from these measured results is shown in Fig.5. The compensated MTIE is constant at about 20ns. It is possible to reduce the transmission system MTIE to 1μs by including a margin for other degradation sources such as delay measurement resolution and compensation error.

The improved MTIE allocation is:

Transmission path MTIE : 1μs

MTIE of 20 chained slave clocks including intra-node degradation: 20x0.3μs

Reference clock MTIE based on the CCITT recommendation: 3μs.

The compensation of transmission delay can allow up to 20 slave clocks to be chained; other slave clock systems such as the intra-office timing distribution and the timing extraction do not need to be improved.

3. TIME SYNCHRONIZATION

3.1 Future Network Synchronization

There is a need to distribute time in addition to timing (frequency) in telecommunication networks. One application that utilizes time is the distributed computer network used as the distributed database system for telecommunication network operation. Time is required to maintain the integrity of the database renewal process. The time stamp method is supposed to be suitable for large distributed computer networks. The time signal supply based on time synchronization enables the time stamp method to be used throughout the network. Accurate time stamps can be used to realize cell transmission delay measurement which is important in Asynchronous Transfer Mode (ATM) and other services. Time synchronization using terrestrial communication systems is needed to ensure accurate International time comparison since the capacity of satellite channels is too low to permit international time comparison over a long measurement term.

3.2 Transmission Delay Compensation Limitation in Time Synchronization

The delay compensation method described in 2.3 can achieve time synchronization as well as timing (frequency) synchronization because the transmission delay of the outgoing path is calculated from the round-trip delay measurement and can be compensated at lower slave nodes when they receive the delay data. The absolute time difference between the reference and slave nodes depends on delay measurement accuracy, that is the transmission delay difference between the outgoing and incoming paths. This accuracy is 150ns in the example shown in Fig.4. It can be estimated that the total absolute difference in time distribution is less than 500ns in NTT's digital network.

3.3 Time Signal Supply

Intra-office time signal supply systems have two important characteristics: time signal supply characteristics between the time signal supply and equipment requiring the time signal such as a work-station, and between internal circuits such as between a work-station interface and a UNIX based operation system. The time signal MTIE that corresponds to the time accuracy shown in Fig.6 is less than 3 ns with an 8 Mb/s round-trip digital interface. The time signal is supplied to individual devices with both outgoing and incoming cables as well as between nodes. The system configuration is similar to that shown in Fig.3. Absolute time

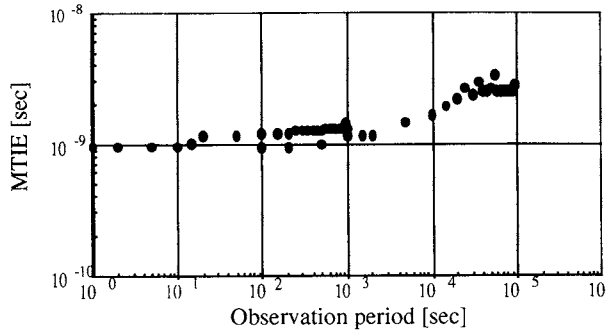


Fig. 6 Intra-office time distribution accuracy with an 8Mb/s interface

difference in a UNIX based operation system has not been measured yet; however, a stability of better than 15 micro seconds is possible if only the time recognition process is run by the work-station.

4. CONCLUSION

It is possible to synchronize a 20 node timing distribution chain such that the MTIE is less than 10μs if the MTIE specifications of the slave clocks are improved and transmission delay compensation is employed. The proposed delay compensation method achieves absolute time synchronization. Its accuracy is better than 1μs up to the equipment input interface, and overall accuracy depends on the actual processes performed within the equipment to which the time is distributed.

ACKNOWLEDGEMENTS

We are grateful to Dr. Tomonori Aoyama and Mr. Ikuo Tokizawa of NTT Transmission Systems Laboratories for their encouragement.

REFERENCES

[1]CCITT Recommendation:"Network node interface for the synchronous digital hierarchy," G.708

[2]CCITT Recommendation:"Timing requirements at the outputs of primary reference clocks suitable for plesiochronous operation of international digital links," G.811

[3]CCITT Recommendation:"The control of jitter and wander within digital networks which are based on the 2048 kb/s hierarchy," G.823

[4]CCITT Recommendation:"The control of jitter and wander within digital networks which are based on the 1544 kb/s hierarchy," G.824

[5]E. W. Butterline and J. E. Abate and G. P. Zampetti:"Use of GPS to synchronize the AT&T national telecommunications network," Proc. of the 20th PTTI, pp.1988

[6]M. Kihara and A. Imaoka:"Time concurrency / phase-time synchronization in digital communications networks," Proc. of the 22th PTTI, pp.367-374, 1990

TIMING A NATIONAL TELECOMMUNICATIONS NETWORK USING GPS

E. W. *Butterline*[1], J. E. *Abate*[2], G. P. *Zampetti*[3]

[1] AT&T, Bedminster, NJ, USA; [2] AT&T Bell Laboratories, Holmdel, NJ, USA;
[3] formerly with AT&T Bell Laboratories, Holmdel, NJ, USA

ABSTRACT

AT&T has recently converted to the use of signals from the Global Positioning System (GPS) to synchronize its national telecommunications networks. The GPS based system, termed a Primary Reference Clock or PRC, employs a GPS receiver providing long term timing accuracy along with duplicated disciplined rubidium oscillators providing short term (ie. one day) stability. The PRC system produces a digital signal used to synchronize a master clock in the node which in turn produces a signal that is used to synchronize all other clocks in the node, and distributed on a dedicated digital basis to all neighboring nodes. By use of monitoring equipment at each PRC, the performance of the master clock at every AT&T network node will be verified against the PRC signal. For the first time, a national telecommunications network is monitored and verified to be performing to a precision better than existing international standards.

Keywords: clock, digital, GPS, network, synchronization, telecommunications

INTRODUCTION

Synchronization of national telecommunication networks has become significantly more important with the deployment of digital switching and transmission facilities. The synchronization plans and equipment used for an analog network are not necessarily the optimum for use in a digital network. Many of the networks routinely used by industry and government require high performance synchronization. To provide that high quality synchronization, AT&T has developed a new verifiable synchronization network.

This paper describes the new synchronization network focusing on the new AT&T Primary Reference Clock (hereafter referred to as the PRC), its verification methodology and capabilities, and the methodology for distributing its timing reference to nodes and systems throughout its digital network. It will also touch briefly on the hardware selection process and the actual PRC performance realized.

BACKGROUND

AT&T provides synchronization signals for its national telecommunications network, to customers private networks, and to many local exchange carrier telephone companies that provide local telecommunications services. As work started to transition from an analog to a digital network in the 1980's, it became clear that the 1970's analog synchronization plan needed to be revisited. The analog network was timed using a cesium clock ensemble (Fig. 1). That approach worked well in a frequency tolerant analog network where the facilities were synchronized and the switches were not. Even the most critical frequency division multiplex equipment used in the AT&T network could tolerate a 10^{-9} frequency offset. When it became obvious that the network would grow digitally and analog facilities would eventually be retired, serious effort was begun to design a digital synchronization network.

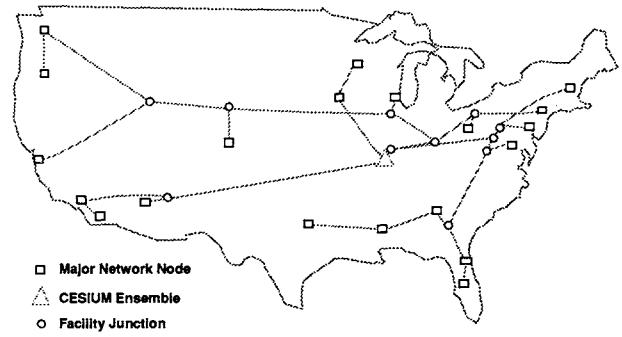


FIG. 1 PREVIOUS AT&T PRIMARY TIER ANALOG REFERENCE DISTRIBUTION SYSTEM (REPRESENTATIVE)

The new network would be composed of time division transmission and switching. Synchronization of a time division network encompasses bit, frame and network timing issues.

The bit timing level of synchronization deals with physical layer timing issues such as: clock insertion and recovery, transmission line jitter, sampling windows in eye patterns, and ones density. These issues are addressed in the fundamental design of the transport system and by adhering to engineering rules such as repeater spacing and maximum number of repeater spans.

The frame timing level of synchronization deals with identifying groupings of bits (time slots) which can be associated with a particular user. A framing signal encoded in the digital transmission system typically accomplishes this function. When time slots within a digital stream need to be processed separately, a frame alignment process is performed on an incoming stream. Once framing is determined, each time slot can

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

be identified and stored for processing. The main synchronization issues at the frame level of synchronization are reframe time, framing loss detection, and incorrect framing.

The network timing level deals with processing of time slots. Under ideal conditions, time slots are sent from the source node at a perfectly uniform rate, delivered over the transmission system with a fixed delay, and clocked into the receiving node buffer. The receiving node will then read the time slot currently in the store after a fixed delay. Assuming no random or systematic offset of the receiving node clock relative to the sending clock, this processing will continue with no variation in the buffer and complete time slot integrity.

The objective of network timing, is to approach this idealized scenario. Time slot buffers (more commonly termed frame alignment buffers) can accommodate 125μs of systematic time error and 18μs of random time error without slipping time slots. The maximum long term systematic frequency departure allowed at the output of a network clock is 1×10^{-11}. This requirement has been established for a number of years both domestically [1,2] and internationally [3,4], and is based on achieving satisfactory slip rates on an end to end basis for the most critical services (voiceband data, facsimile, video, secure voice and secure data). Provisional short term requirements allow from 1 to 10 μs of time error in a day at the output of a network clock. The short term requirements are being put in place to ensure that random or cyclical daily timing variations will not produce daily slips.

The performance objective of 1×10^{-11} for digital network timing is two orders of magnitude greater than the frequency accuracy required for analog transmission. This could be obtained using a single cesium ensemble. However, the realities of transmission and operating impairments could, at certain times, result in a node degrading to parts in 10^{-10} for a short period.

As a result, AT&T has deployed a network of distributed primary reference clocks to synchronize its digital network. These clocks will use the Global Positioning System (GPS) to provide long term timing accuracy to 16 network nodes or buildings.

This new synchronization network will have rich connectivity and unique capability of verification at several levels. Verification is the capability to directly monitor the timing performance of each office or node. This capability enables AT&T to detect and resolve timing degradation before it becomes service affecting. Verification ensures uniform compliance of domestic and international timing standards at all nodes, and is a cost effective tool to sectionalize and resolve timing problems. Two fully redundant centers are in place where trained technicians are monitoring all AT&T network clocks.

PRIMARY REFERENCE CLOCK STRATEGY

The architecture of the new verifiable timing distribution plan is illustrated in Fig. 2 . In addition to the fourteen PRC's located in the mainland United States, PRC's are located in Hawaii and Puerto Rico. This network has the capability to detect timing degradations. The PRC reference is transmitted to other nodes, called secondary nodes, which do not contain a PRC. The PRC simultaneously monitors the reference it receives from at least two other PRC nodes as well as secondary nodes. The monitoring of neighboring PRC locations provides long term stability data. By employing a n-corner hat decomposition [5] of the PRC level stability data, the long term stability of each PRC location will be tracked and verified to parts in 10^{-13}. The architecture is designed so that secondary nodes are monitored by two independent PRCs. This provides a means to sectionalize nodal timing instability from transmission path timing instability. By deploying PRC's close to secondary nodes, a uniformly stable transmission path performance is attained both for transmitting and monitoring performance. Daily transmission path stability exceed several parts in 10^{-12}.

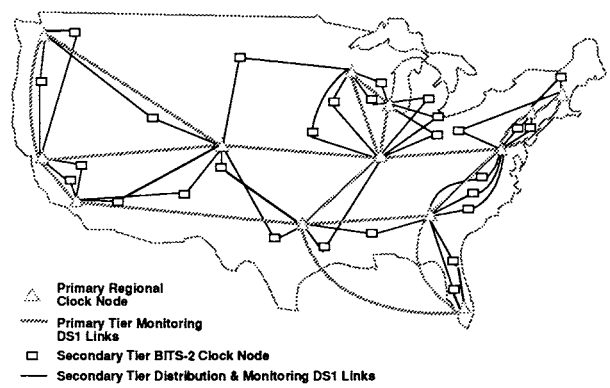


FIG. 2 VERIFIABLE DIGITAL SYNCHRONIZATION DISTRIBUTION SYSTEM (REPRESENTATIVE)

The search for a source of precise timing for a PRC led to three alternatives. The first was to continue to use a single cesium ensemble with a digital distribution system. The second and third considered regionally distributed PRC's using either LORAN-C or GPS and disciplined timing elements. The single cesium ensemble offered some economies, but did not provide for nationwide verification of secondary nodes, nor would it offer a performance improvement over the then current network. Using a nodal based system would permit a broad based verification capability. GPS was judged superior to LORAN-C in terms of time transfer accuracy, inherent operations, and future capabilities, therefore AT&T chose GPS for its long term time reference. Rubidium oscillators were selected over quartz and cesium as the disciplined timing elements.

The strategy in using the GPS system was based on using a daily observation period to obtain a time and frequency error estimate. There were many reasons for this choice. First, a daily ensemble average effectively removes the daily systematic bias errors which arise in the time transfer process of GPS from user position errors and other sources. Second, ensemble averaging of the satellite tracks reduces random error sources. Third, evaluation of ensemble tracks permits the removal of outlying data which can contaminate the steering process.

For periods less than one day, a set of duplicated rubidium oscillators provide timing stability. The stability of the disciplined rubidium over the 24 hour interval between steering updates is equivalent to the stability of the steering update data from the GPS satellite tracks. For a 24 hour steering interval, rubidium oscillators proved to provide the best cost/performance tradeoff. Quartz oscillators cannot obtain the daily stability performance achievable from GPS, and cesium oscillators provide no real stability advantage and are more costly.

A computer is used to steer the rubidium oscillators and to verify their performance relative to one another and to the GPS receiver. The block diagram given in Fig. 3 illustrates the interconnectivity of the PRC's computer, and indicates the DS1 inputs from the other nodes and clocks used for timing monitoring. A DS1 is the primary rate synchronous digital signal used in the network. It is a 1544 Kbps signal consisting of twenty four 64 kbps time slots. Each time slot typically carries

a single direction voice circuit. In the figure, the rubidium oscillators (Rb) are shown communicating with the disciplined controllers which in turn communicate with the computer. The computer performs control, performance verification, and error recovery functions.

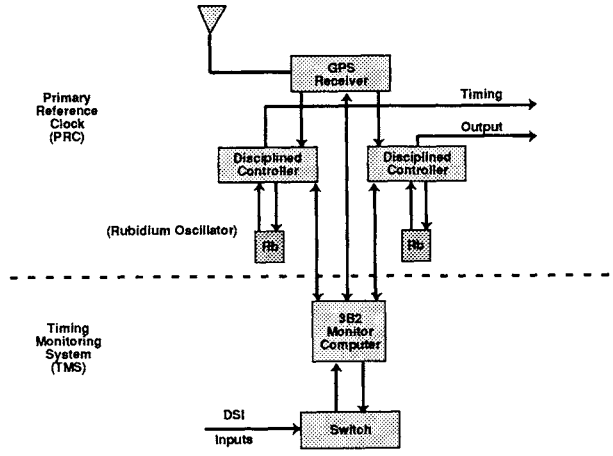


FIG. 3 PRIMARY REFERENCE CLOCK AND TIMING MONITORING SYSTEM

The timing outputs of the PRC are a pair of DS1 primary rate digital signals. The PRC outputs are used to provide timing reference to a master clock in a node containing typically many clocks. The master clock, known as the Building Integrated Timing Supply, or BITS [6], provides timing reference to all other clocks in that node as illustrated in Fig. 4. The BITS provides input to a device known as a Clock Distribution Unit whose function is to provide multiple DS1 outputs to all clocks in the node requiring digital reference. It can also provide analog synchronization signals to frequency locked clocks for any remaining analog frequency division multiplex (FDM) equipment.

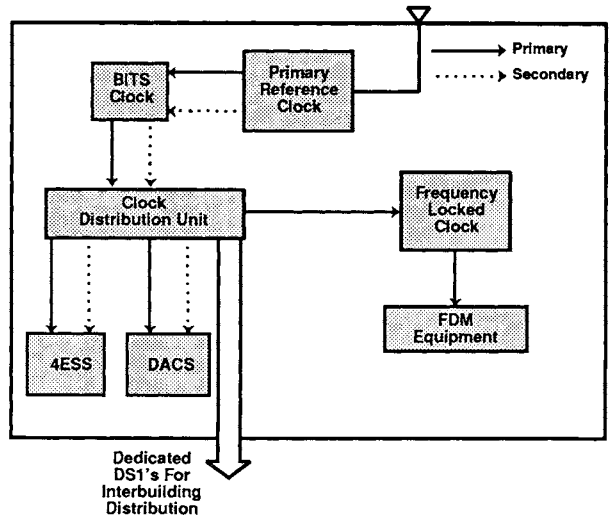


FIG. 4 PRIMARY NODE TIMING DISTRIBUTION

INTER-NODE AND INTRA-NODE REFERENCE DISTRIBUTION

Buildings containing a PRC are known as Primary Nodes. Buildings which do not contain a PRC are known as Secondary Nodes and obtain their timing reference from a Primary Node. This strategy minimizes the cascading of timing through intermediate clocks. Timing to Secondary Nodes is distributed on a digital basis using two DS1 signals which are routed over diverse facilities from different PRC's. These DS1 signals carry only synchronization reference between nodes as illustrated in Fig. 5. A dedicated synchronization network of over 300,000 kilometers of DS1 facilities is in use. The reference information is transmitted to the clock of the system selected as the BITS. Thereafter, the reference distribution is the same as in the Primary Node.

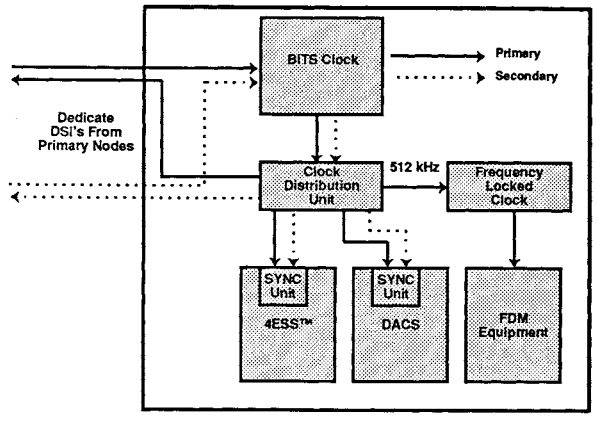


FIG. 5 SECONDARY NODE TIMING DISTRIBUTION

VERIFICATION METHODOLOGY

As illustrated in Fig. 6, there are four levels of verification in the AT&T methodology. The first is an ongoing verification of GPS performance by various laboratories in the United States such as the National Institute of Standards and Technology, the Naval Observatory, and the Network Synchronization Laboratory at AT&T Bell Laboratories. The high level verification results in long term changes in the synchronization network operations such as upgrading of steering algorithms.

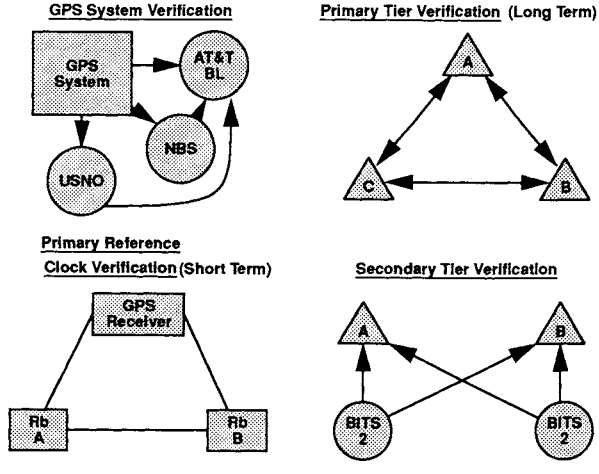


FIG. 6 LEVELS OF VERIFICATION

The second level of verification is within the PRC itself. The PRC determines the relative timing instability or uncertainty associated with the three clocking elements in the system (the GPS receiver and the two disciplined rubidium oscillators).

The third level of verification, called the Primary Tier (or long term) verification determines the long term time uncertainty or noise of each of the Primary Nodes (eg. A, B, and C in Fig. 6).

Finally, the fourth level, or Secondary Tier Verification, extends verification down to the Secondary Nodes, where the signals returning from the Secondary Node BITS clocks are monitored by at least two PRC's to determine their performance.

PERFORMANCE

Fig. 7 illustrates the performance of the AT&T Primary Reference clock as compared to the cesium ensemble and telecommunication synchronization standards. In Fig. 7, the time keeping error relative to UTC is shown in seconds on the ordinate along with the observation time on the abscissa shown in seconds.

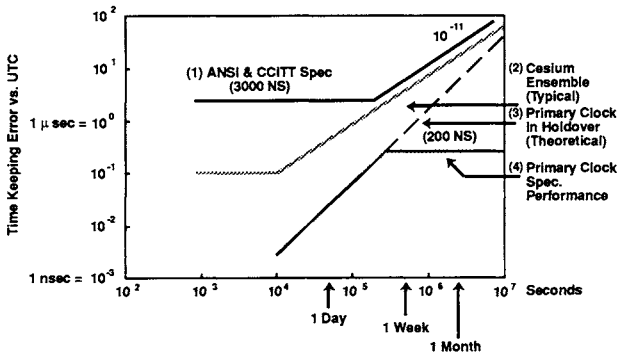


FIG. 7 PERFORMANCE COMPARISONS

In case 1, the slope of the ANSI and CCITT specification is the allowed long term frequency accuracy of a PRC of 1×10^{-11}. The ANSI and CCITT specifications allow for a maximum daily time instability of 3000 ns which is reflected by the horizontal asymptote.

Case 2 shows typical performance of the former AT&T cesium ensemble. It is monitored relative to LORAN-C and found to perform with a long term accuracy within a few parts in 10^{-12}. The daily timing stability is limited by diurnal transmission delay variation of typically 100 ns.

Case 3 shows the time keeping performance of the rubidium disciplined oscillator assuming a failure condition in which no daily steering is applied. It is based on a simplified assumption of optimal calibration of initial time, frequency and drift, and a benign environment. The thermal design maintains small ambient temperature variation (5 deg C) daily objective.

Case 4 shows the PRC performance under normal conditions, three sigma time error specification of 200 ns. The PRC three sigma performance bound is based on measurements and analysis of the PRC components and simulation of the steering algorithm. The PRC's are typically performing at a level well within their specifications [6]. A new algorithm, for use with Block II satellites, has recently been placed in service. PRC performance is still being evaluated.

CONCLUSIONS

AT&T has deployed a GPS based system, termed a Primary Reference Clock or PRC, which employs a GPS receiver providing long term timing accuracy along with duplicated disciplined rubidium oscillators providing short term (ie. one day) stability. Under computer control, these three elements are verified against each other and against identical PRC systems in other parts of the network. All these network nodes are interconnected with a dedicated network of DS1 facilities. For the first time, a national telecommunications network is being monitored and verified to be performing to a precision well in excess of existing standards.

ACKNOWLEDGEMENTS We wish to thank Sanjay Asthana, Roy Carley, Monique Cundari, Ricardo Mediavilla and Adolfo Montenegro of AT&T Bell Labs, and Robert Lopes and Doug Sisk of Network Systems Adaptive Design Engineering Center, all of whom made numerous contributions to the development of the PRC system.

[1] AT&T Technical Reference PUB 60110, "Digital Synchronization Network Plan," December, 1983.

[2] American National Standard for Telecommunications,"Synchronization Interface Standards for Digital Networks," ANSI T1.101-1987

[3] CCITT Recommendation G.811, "Timing Requirements at the Output of Primary Reference Clocks Suitable for Plesiochronous Operation of International Digital Links."

[4] CCITT Recommendation G.824, "The Control of Jitter and Wander Within Digital Networks which are Based on the 1544 kbit/s Hierarchies."

[5] C. A. Greenhall, "A Likelihood Approach to the M-Cornered Hat - Preliminary Report," Proceedings of PTTI, 1987.

[6] J. E. Abate, E. W. Butterline, et al, "AT&T's New Approach to the Synchronization of Telecommunication Networks," IEEE Communications Magazine, Vol. 27, No. 4, 1989.

STRATEGIES FOR SYNCHRONISATION IN AN EVOLVING TELECOMMUNICATIONS NETWORK

Rob Avery
Telecom Solutions Europe, Ltd
England

ABSTRACT

Achieving precise synchronisation in the telecommunications environment is not a trivial task. Transmitting the timing from node to node has been the inherent problem for all digital networks. Traditional network equipment used to transfer network synchronisation, such as digital switching systems, adds impairments to the once-traceable signal. As the synchronisation signals are passed from node-to-node, they lose stability by passing through intervening clocks.

Timing will be an integrated part of all new network and service deployments. New transmission methods, such as the synchronous digital hierarchy (SDH), survivable network topologies and the issues that arise from them, necessitate a review of current network synchronisation strategies. Challenges that face the network are itemized in this paper. It will be demonstrated why localised primary reference clocks (PRC) in key nodes and the synchronisation supply unit (SSU) clock architecture of transit and local node clocks is a technically and economically viable solution to the issues facing network planners today.

Key Words: Synchronisation, Synchronous Digital Hierarchy (SDH), Network timing, Clocks.

1. INTRODUCTION

The digital telecommunications network is becoming homogenized to the point where the traditional physical boundary has become blurred. These mid-point connections of fibre optical facilities using the synchronous digital hierarchy (SDH) carry vast amounts of valuable digital information. As the network evolves, SDH equipment will create digital synchronous islands that require connection to the existing plesiochronous digital hierarchal (PDH) network with its asynchronous equipment. The originator, intermediate carriers, and terminator of digital information must be synchronised to a common, highly-stable, clock source to ensure complete connectivity and performance objectives are met. The new technology of SDH is placing an increased demand on network synchronisation.

2. SYNCHRONISATION ISSUES

Todays strategy of mutual and hierarchal synchronisation has solved many network problems. However, timing transfer still remains an issue. Current network topologies rely on switching equipment to provide the timing distribution. With this method, the network and the network elements (NE) contribute to timing impairments. Current network synchronisation issues are divided into four categories:

- Network elements designed for other purposes are employed to distribute timing.
- Intervening clocks add inaccuracies and instabilities.
- Specifications for synchronisation are evolving and not widely understood or adopted by the industry.
- New technologies and topologies create additional concerns.

Synchronisation distribution is linked closely with the structure and topology of the current network where the routing of internode feeds are well constrained. However, timing transfer performance (jitter, wander and short-term stability) of NE has been, within certain bounds, fairly loose and not adequately specified by industry recommendations.

3. EQUIPMENT TIMING

The Synchronisation Supply Unit (SSU) concept is the preferred method of distributing phase and frequency synchronisation within an office [1]. This concept is based on the well established Bellcore BITS (Building Integrated Timing Supply) method used in N. America and Canada. In each node, only one slave clock is designated as the synchronous supply unit in that node. This clock receives timing references from other nodes and supplies timing signals to all other clocks in that node (Figure 1).

Synchronisation in the network is typically achieved by conveying the timing reference signal from one SSU to the next SSU. Hierarchies of SSUs are established with some SSUs (identified as transit node clocks) being slaved from higher order SSUs (Primary Reference Clocks; PRC) and in turn acting as master clocks for lower SSUs (local node clocks) (Figure 2). The SSU is a stand-alone device and is also known as a Digital Clock Distributor (DCD). For a "true" SSU architecture, the DCD is not a part of any network element. The advantages of a DCD is that its characteristics are designed solely for timing and are not compromised by other functions such as switching or transport.

The performance of the SSU is loosely covered by the established CCITT recommendation G.812 for slave clocks [2] and the new recommendation G.81s for SDH clocks [3]. Both these documents are currently under revision. The benefits of using the SSU concept are found in four areas:

- Increased performance.
- Reduced administration.
- Reduced maintenance.
- Maximum utilisation of resources.

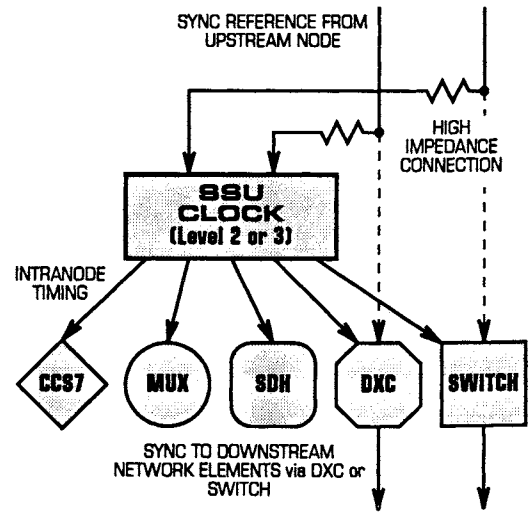


Figure 1. The Synchronisation Supply Unit (SSU) provides all timing signals within a node and filters the recovered reference.

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

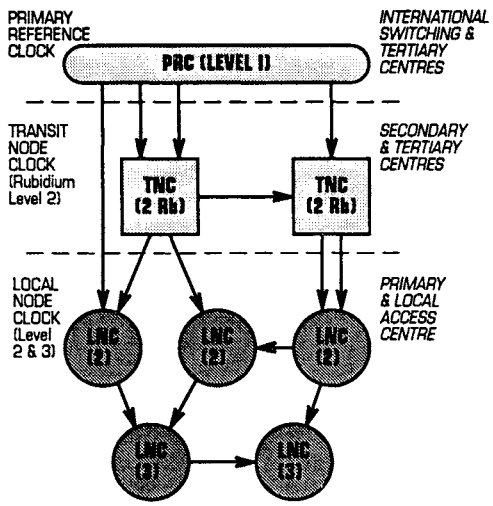


Figure 2. The synchronisation network is built from a hierarchy of clock systems from PRC (level 1) to LNC (level 3).

The fact that the SSU is location-dependent rather than service-dependent—that is, synchronisation is not a function of switching or facilities but a core support service— makes installation and maintenance considerably easier as new digital services are introduced [4]. An SSU also serves to filter perturbations from the timing signal before it is used to synchronise NEs and downstream nodes.

4. NETWORK TIMING STABILITY

As the synchronisation reference signal passes from node-to-node and on to the lowest level in the "food chain", intervening clocks and the transmission medium contribute to timing inaccuracy and instability [5,6]. Long-term variations such as wander are cumulative, resulting in poor stability below the 1,000 second range. Although the SSU clock provides stable, filtered, timing to the equipment within a node, it is rare that the SSU directly feeds the downstream network. NEs, such as digital cross-connects and digital switching systems, are employed to perform the distribution function. These NEs, as stated earlier, are not specifically designed to transfer timing and therefore add instabilities to the timing signal by the "breathing" effects of their intervening clocks and related system architectures.

The characteristics of the reference as it passes through the network and the resulting impact on network performance are still being analysed. Standards' bodies continue to provide satisfactory and detailed specifications for the understudy portion of the synchronisation signal. A draft document [7] will soon supersede the current CCITT synchronisation interface recommendations with a new, more stringent, proposal for short-term timing stability (Figure 3).

The rush to plug the holes in the synchronisation standard is being fueled by the new technology and trial deployment of the synchronous digital hierarchy (SDH). SDH, as its name implies, relies heavily on synchronisation. Therefore, the stability of its timing source directly reflects its performance. This is also true when SDH is used to transfer timing. The "timing" output of a SDH NE is a function of its input plus any inherent effects from the NE.

5. CHALLENGES IN THE EVOLVING NETWORK

SDH has, perhaps, the greatest impact on network performance. Its potential is enhanced digital service offerings and powerful networking capabilities while substantially reducing network delay. But, ironically, with this vast improvement comes a synchronisation penalty. The major challenge presented to equipment manufacturers and network planners is how to put the "S" (synchronisation) in SDH. Many of the synchronisation challenges are common to all digital networks. However, SDH presents several unique issues of its own:

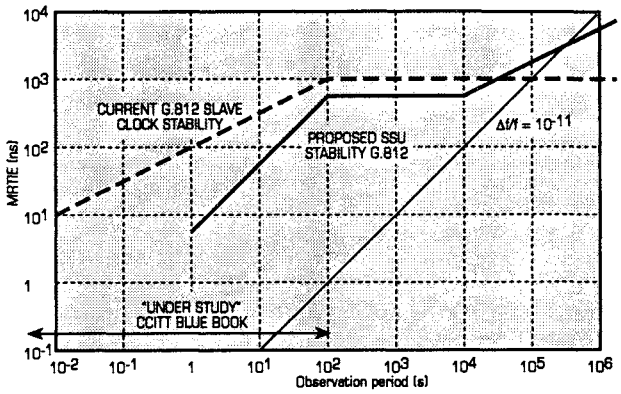


Figure 3. Proposed SSU stability requirement vs. the current G.812 slave clock short-term stability.

- SDH is a "synchronous" system vs. todays PDH and asynchronous equipment.
- Standards are "understudy" for SDH and synchronisation.
- VC (virtual container) interfaces with existing asynchronous NEs and PDH networks create timing compatibility concerns.
- Existing NEs upgrading to SDH need performance characterization.
- Intelligent networks add complexity to the network.
- Added complexity of the network creates administrative issues.
- NEs that distribute timing are intervening clocks.
- SDH adds new requirements for SSU performance and capacity.
- SSUs are not widely used where SDH will be deployed.

5.1 SDH in the PDH Environment

Synchronisation, therefore, plays an extremely important role in the performance and networking of SDH. SDH, by design, is expected to work in a purely synchronous network. However, the existing environment will be transitioning from PDH to SDH and the two must be capable of interworking [8]. One of the principle differences between PDH and SDH multiplexers is their interface buffers which absorb timing variations and contribute to —or reduce— transmission delay. The 2 Mb/s interface of a PDH mux typically has a ±125 microseconds slip buffer where the delay accumulates as a connection is built. Whereas the SDH mux has an end-to-end delay objective of no more than 50 microseconds where the buffering of the SDH is distributed (Figure 4) across the connection.

Therefore, it can be seen that relatively small timing variations will fill the SDH buffer in contrast to the larger PDH buffer. As the SDH buffering reaches the threshold due to timing variations, the 2 Mb/s signals will be readdressed in the synchronous transport module (STM-1) payload by a new pointer value. These pointer justifications appear on the output side of the SDH transport system as 8 UI phase steps. Fourth order signals (140 Mb/s) mapped into STM-*n* will undergo similar phase hits due to timing variations. The rate that the pointer justifications (8 UI phase steps) occur per day will depend on the relative clock offset of each SDH terminal and the network interfaces. Recent studies [9] have shown that the mean time between pointers could be as little as 13 seconds.

The resulting effect of these pointer-created phase steps has still to be evaluated in a functioning SDH network. But it is known that the pointer justification steps of 8 UI created by clock differences will result in jitter bursts. These jitter bursts statistically will cause buffer spills in PDH equipment resulting in slips (Figure 5). The slips will be received by an end user as "severely errorred seconds" (SES). It is therefore important to stop this error chain

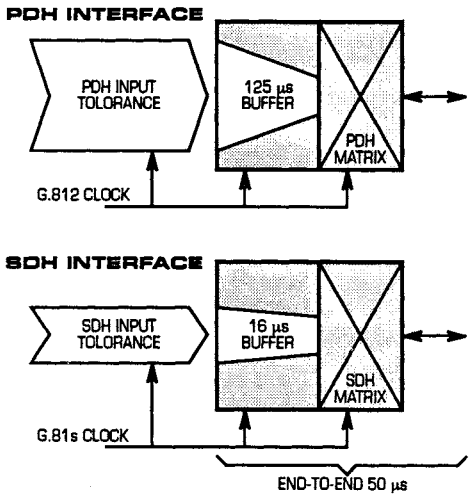


Figure 4. The interface buffers that accommodate timing variations differ in size between the PDH and SDH systems.

at the source by controlling the timing differences within very tight bounds to meet end-to-end error performance objectives [10,11].

5.2 Network Topologies

Digital NEs such as digital cross-connects (DXC) and exchanges are being offered with integrated SDH transport functions or the promise of upgrades. Most of these existing NEs are currently used to distribute timing throughout the network. With the introduction of SDH, the entire topology of today's network is being reevaluated. The operation, administration, maintenance and provisioning (OAM&P) function of SDH will allow the network to be configured for self-healing and automatic reconfiguration by the use of a flat-mesh consisting of diverse-routing, dual-homing hubs and interwoven ring architectures. While this strategy will potentially aid the long-term integrity of the network, it may actually lower the reliability of synchronisation by adding complexity and administrative problems to timing distribution by the lack of a well constrained hierarchy.

Some network providers have chosen to establish dedicated timing links from node to node via the existing PDH asynchronous NEs, which directly connect SSU clocks to SSU clocks. This is, however, a short-term work-around that allows SDH to coexist with the current network by removing the burden of timing distribution from the SDH transport.

5.3 Timing via SDH

When a 2 Mb/s signal (intended to distribute timing) is passed through SDH, it will be subject to the phase steps caused by pointer justification which may be a significant impairment to downstream clocks. 2 Mb/s signals embedded in a 140 Mb/s signal —before passing through a SDH transport system— will also have the 8 UI phase step. Since the 140 Mb/s is a much higher bit rate, an 8 UI transient at 140 Mb/s will not greatly affect the embedded 2 Mb/s' performance. However, once a 2 or 140 Mb/s signal is mapped into an SDH transport system, it is difficult to control its routing through an integrated network of DXCs and add-drop multiplexers (ADM). Because of the potential for rerouting, it is impossible to differentiate a 2 Mb/s suitable for network timing from one that is not.

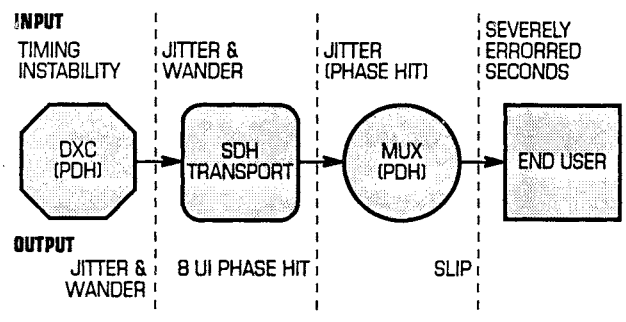


Figure 5. The timing error chain must be broken by controlling the stability of the source clock (SSU).

One potential solution recommended by CCITT [12] is to provide a method of using the optical line rate (STM-*n*) of SDH as a synchronisation transport medium. To allow the evolution of SDH based timing distribution, NEs in exchange applications are required to provide the capability to time the SSU by providing a 2 Mb/s reference derived from the STM optical carrier (Figure 6). Each SDH NE will use "timing only" ports (2 Mb/s) from the collocated SSU to time its internal clock and time the optical carrier. The receiving SDH NE will then convert the optical line rate directly to a pair of 2 Mb/s timing ports to be used by the local SSU as the reference signal. Timing signals from the SSU are then used to clock all NEs at that node.

Optical based timing is offered due to its immunity to facility rearrangements. It is not subjected to the phase steps resulting from pointer justification and has transparency from source to sink. However, this is a very simplistic view of transporting synchronisation via SDH and does not address the complex network topologies in which SDH will be used.

One solution to the network puzzle is not to rely on SDH for internode synchronisation but to continue to use the existing PDH facilities. This strategy mandates that two independent networks are maintained. This may not be practical since most telephone companies are phasing in SDH-based transport systems that will "optically isolate" exchanges by replacing existing facilities.

6. SDH NETWORK COMPLEXITIES

When timing is carried via the optical line rate in a network environment each SDH NE will regenerate the optical signal under control of the NE internal clock and pass that signal to downstream NEs. Therefore, undesirable timing characteristics will accumulate toward a maximum allowable limit as the timing is passed from node to node. The internal clocks of some NEs in service today —which will be upgraded to SDH transport systems— exhibit excessive short-term timing instabilities and/or do not meet the proposed CCITT G.81s clock specifications. Optical timing requires each NE to be evaluated on a case-by-case basis before including the NE in the timing plan. SDH ADMs will also use internal (intervening) clocks to regenerate timing and pass the timing to the next downstream ADM and, therefore, must meet the proposed stringent short-term stability requirements of G.81s.

The adoption of a survivable network by the use of SDH rings, route diversity and automatic reconfiguration places an additional requirement on the SDH NE and SSU clock. Intelligence must be provided that determines the derivation of the timing source,

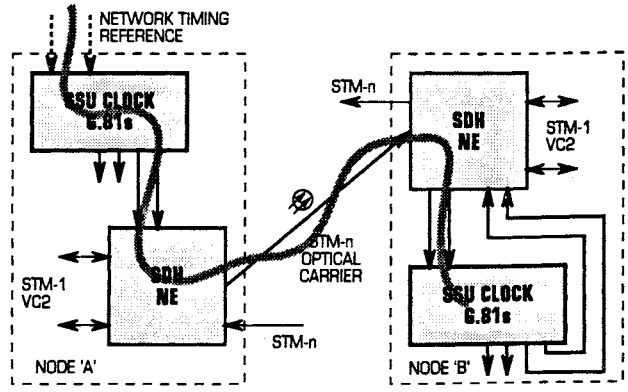


Figure 6. Timing via the STM optical carrier. The SSU provides all timing within a node.

whether the timing source is suitable as a reference, and whether or not a choice should be made to search for other possible timing sources (other optical links, diverse network feeds or holdover). Selection and administration of reconfigured paths —suitable for timing references— become very complex and need to be administered manually at each node. Each network configuration and the associated failure modes will require analysing. This network model grows geometrically as links are added.

Synchronisation status messages plays a key role in intelligent networks. As listed above, the NE must be able to identify viable timing links (Figure 7). A proposal is currently being discussed within the standards bodies [13] that will assign a synchronisation "quality indicator" (QI) message to be mapped onto the SDH overhead structure from the SSU clock. Generically, the synchronisation QI message will contain the source of timing and the status of the intermediate clock(s). This new standard is a formidable task since it requires all equipment vendors and network providers to work in concert to ratify the requirement. If "national" bits are also standardised by the ETSI committee, status messaging will be considerately easier across national boundaries.

Smart networks are not limited to SDH architectures. 4/4 and 4/1 automatic digital cross-connects can also provide the dynamic network reconfiguration, rings and route diversity that SDH offers. One problem that does arise in a total electronic office — where the manual digital distribution frame (DDF) has been replaced by an automatic DDF (ADDF) or DXC— is that there is no physical access to a 2 Mb/s signal where the SSU clock extracts timing from the upstream node. This may be the case as existing exchanges are converted to "all digital" in an effort to improve performance and reduce maintenance and operations.

7. NEW TIMING STABILITY REQUIREMENTS

Clock requirements for SDH are known to be extremely stringent. SDH vendors have expressed a need for signals with very tight short-term timing characteristics. The source for these signals will be via the new enhanced SSU filter clock specified under CCITT recommendation G.81s. Currently, the industry standards bodies are actively negotiating applicable limits but basically the new enhanced clocks' bandwidth will increase the filter characteristics from approximately 5 Hz to 0.01 Hz [7]. These limits will be based on the effects of empirical data applied to an "SDH reference model" currently known as the G.sna1 document [14]. The SSU is therefore considered an SDH timing filter that

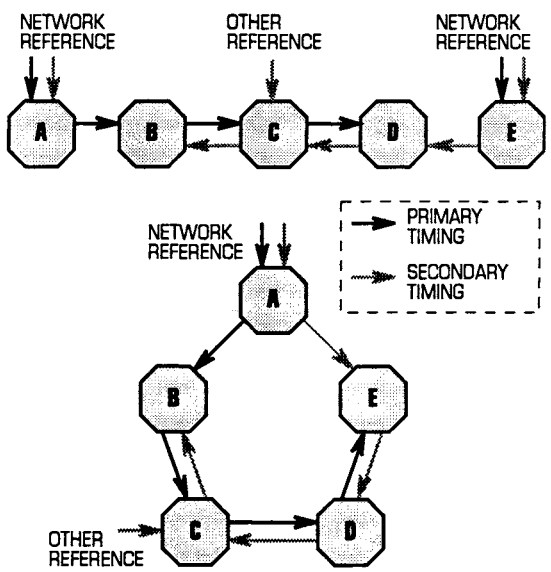


Figure 7. The SDH ring and linear network have similar concerns when identifying viable timing links. QI messages must therefore be used.

provides the very precise stability required by the NE to minimize pointer justification and assure the high-speed optical jitter performance of the SDH-based network (Figure 8). Most transit node (level II or stratum 2) SSU clocks will meet the requirements proposed however tight the stability mask. However, the specifications must be known before any judgment is made on existing local node (level III or stratum 3) SSU clock compliance.

8. DISTRIBUTED TIMING

The SSU hierarchy still relies on the transport of timing from node to node and master-slave operation with its attendant problems stated above. An alternative to transporting timing from the single primary reference clock (PRC) location to the far extremes of a network is to deploy distributed PRC clocks in key nodes. At each site the SSU can be directly coordinated with UTC (universal coordinated time), and become a PRC operating pseudo-synchronously with other PRC sites in the network as national networks do today [15].

The recent modernization of LORAN-C by the US Department of Transportation (DOT) and the trial deployment of Global Positioning System (GPS) satellites by the US Department of Defense (DOD) allows the recovery of a PRC (level I or stratum 1) signal from the LORAN or GPS transmitter's cesium atomic clock ensemble by using low-cost radio receivers [16]. LORAN type systems are available throughout most of the Northern Hemisphere, while GPS provides global coverage. Both LORAN-C and GPS PRC SSU clocks provide long-term stabilities in the range of parts in 10^{-12} (an equivalent slip rate measured in years), and short-term stabilities in nanoseconds, by locking a digital synthesiser with a rubidium oscillator standard to the received signal [17]. This stability is well within the requirement for a PRC source and proposed SDH standards.

8.1 Benefits of Distributed PRC

Localised PRCs distributed in every exchange effectively eliminates the issues and problems encountered when timing is transported via SDH or the PDH network. In this configuration, all NEs have access to the highly stable synchronisation source that is now required (Figure 9). Several significant advantages are gained by embracing a localised PRC architecture with SSU clocks:

- The localised PRC strategy is autonomous of network architecture and rearrangements and removes the requirement for timing distribution maintenance and administration.
- The network operation and performance is not compromised by slaving to a clock source from the upstream network.
- Quality timing is assured independent of the introduction and deployment of new technologies and topologies.
- Synchronisation QI messages are not required since there is no timing distribution.

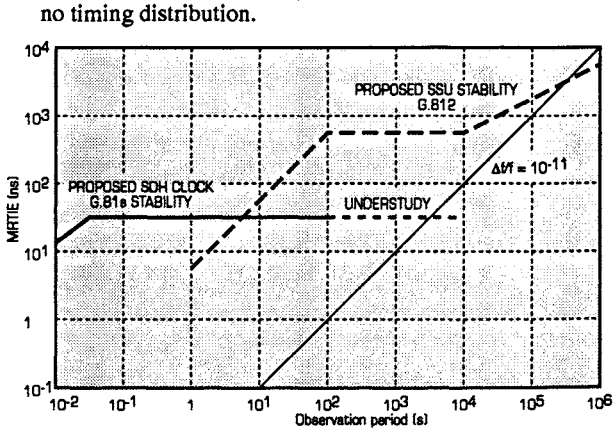


Figure 8. SDH requires very stable timing as indicated by the proposed G.81s stability which is more stringent than the G.812 slave clock.

- The SSU clock is engineered specifically for the best timing reference acquisition, generation and transfer qualities.
- Synchronisation performance is enhanced due to the removal of timing distribution which generally is the main contributor to timing impairments.
- The cost of embedded clocks in NEs can be reduced since they are not required to filter, transfer or distribute timing.

Clearly these factors cannot be ignored when analysing the localised PRC strategy with SSU clocks. Moreover, by removing the transport element from timing distribution, the instabilities introduced by the transport facility are minimised to a point where they will not be a factor for critical SDH synchronisation.

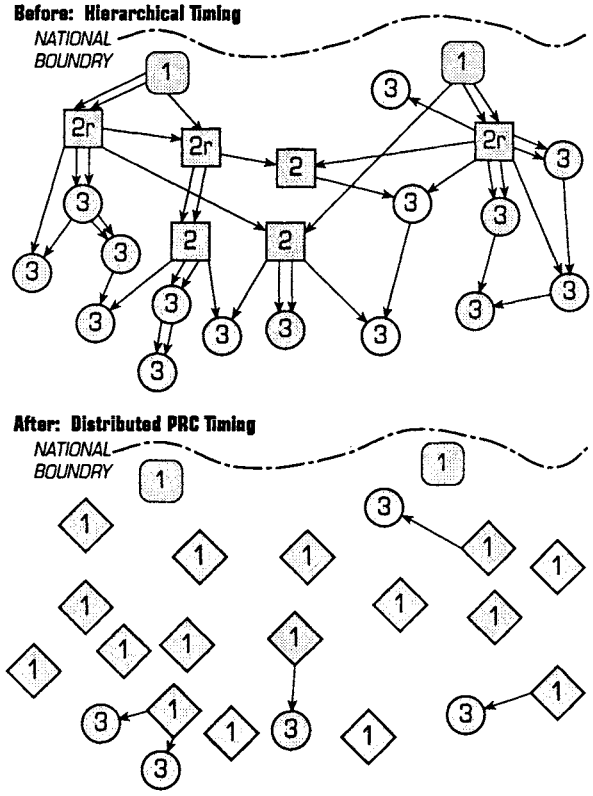


Figure 9. Distributed PRC nodes remove the timing distribution requirement. Node-to-node links now operate pseudo-synchronously.

9. SUMMARY

The network's ability to transfer timing precisely has remained the major roadblock to an ideal synchronisation plan. SDH will exacerbate the already taxed office clocks by placing increased capacity and filter performance demand upon them. The robustness that SDH can provide to synchronisation networking by using its optical carrier and self-healing/alternative routing schemes does, in fact, add complexity to timing distribution.

The challenges of today's —and tomorrow's— network can be met technically and economically by the new approach of distributing the primary reference clock source function to every node via "local" PRCs. This removes the distribution problems entirely and ensures that all network elements receive precisely the same timing – all the time.

The local PRC will provide a superior level of synchronisation that has widespread benefits for all digital services and network performance. The distributed PRC architecture with SSU filter clocks will provide the extremely stable timing required by SDH at its source without the issues of transporting timing throughout the network. This will provide the platform for future technologies and topologies to migrate into the network of tomorrow, today.

10. REFERENCES

[1] CCITT SG XVIII/WP6, *Consideration on Time Interval Error Specifications,* Melbourne, 1991.

[2] CCITT G.812, *Timing Requirements at the Output of Slave Clocks Suitable for Plesiochronous Operation of International Digital Links,* Blue Book, 1988.

[3] CCITT G.81s (draft), *Timing Characteristics of Slave Clocks Suitable for Operation in SDH Equipments,* ??, 19??

[4] Telecom Solutions, *Why Deploy BITS?,* EPL 8808-01, August 1988.

[5] CCITT G.823, *Control of Jitter and Wander within Digital Networks which are based on the 2048 Kb/s Hierarchy,* Blue Book, 1988.

[6] Telecom Solutions, *Timing Jitter and Wander Effects on Transmission,* AN 404, May 1988.

[7] CCITT SG XVIII/WP6, *Study of the Effects of Phase Discontinuity of G.812 Slave Clocks on SDH Networks,* Melbourne, 1991.

[8] K. Asatani, K. R. Harrison, R. Ballart, *CCITT Standardization of Network Node Interface of Synchronous Digital Hierarchy,* IEEE Communications, pp15-20, August 1990.

[9] T1X1.3/91-121, *Pointer Simulation Results on the Proposed TVAR Short-Term Stability Mask,* ANSI T1 Committee, October 1991.

[10] CCITT G.822, *Controlled Slip Rate Objective on an International Digital Connection,* Blue Book, 1988.

[11] CCITT G.821, *Error Performance of an International Digital Connection forming part of an Integrated Services Digital Network,* Blue Book, 1988.

[12] CCITT SG XVIII/WP7, *Proposed Amendment to G.snal,* Melbourne, December 1991.

[13] ETSI/STC TM3 WP2, *Protocol for Synchronisation Status Messages,* Dormstadt, October 1991.

[14] CCITT G.snal (draft), *Architectures of Transport Networks based on the Synchronous Digital Hierarchy (SDH),* Corfou, December 1991.

[15] CCITT G.811, *Timing Requirements at the Outputs of Primary Reference Clocks Suitable for Plesichronous Operation of International Digital Links,* Blue Book, 1988.

[16] US Department of Transportation and Defence, *Federal Radionavigation Plan,* 1990.

[17] Telecom Solutions, *Which PRC for the Telecommunications Industry, LORAN-C vs GPS,* December 1990.

SYNOPSIS OF TIMING MEASUREMENT TECHNIQUES USED IN TELECOMMUNICATIONS

G. P. Zampetti, Telecom Solutions

ABSTRACT

Historically, Maximum Time Interval Error (MTIE) and Maximum Relative Time Interval Error (MRTIE) have been the main measurement techniques used to characterize timing performance in telecommunications networks. Recently, a new measurement technique, Time Variance (TVAR) has gained acceptance in the North American (ANSI) standards body. TVAR was developed in concurrence with NIST to address certain inadequacies in the MTIE approach. This paper describes the advantages and disadvantages of each of these approaches. Real measurement examples are presented to illustrate the critical issues in actual telecommunication applications. Finally, a new MTIE measurement is proposed (ZTIE) that complements TVAR . Together, TVAR and ZTIE provide a very good characterization of network timing.

1. MEASUREMENT CRITERIA

A good starting point in evaluating timing analysis techniques is to set objectives of what constitutes a good approach. A good measurement technique is one that efficiently extracts the useful information from the timing signal in a repeatable and cost effective manner. The usefulness of the information extracted is closely tied to the use of a good practical model. Historically , the model used to described timing signals in non telecommunications applications is based on a simple decomposition of the signal into two components:

- Systematic components (Phase Offset, Frequency Offset and Frequency Drift)
- Stochastic Power law noise components, $S_x(f) = \sum_{n=0}^{4} \alpha_n f^{-n}$ (white pm, flicker pm, white fm, flicker fm, random walk fm).

This model permits the characterization of a complex timing signal with a handful of parameters. The premise that this model can be extended to telecommunication synchronization distribution applications is reasonable. Conceptually , one can treat a distributed network of clocks as a single composite super clock. The dominant sources of noise and bias may be different, but the resulting timing signal will still be characterized by this model. This approach has recently been adopted by ANSI committee T1X1 . Extensive field phase data has been analyzed using TVAR (described later) as the measurement tool to extract the key parameters of the model [ANSI T1X1.3/91-074]. The dominant noise types found in the network are white PM, flicker PM and white FM.
The fact that more divergent noise processes (such as random walk FM) are not significant indicated that the network is reasonable traceable in frequency to the primary reference sources.

1.1 Dual Role of Measurement/Modelling Approach

One role that the measurement approach fills is the need to extract the timing model parameters. The critical issues in this role are:

- Does the measurement approach extract the essential parameters ?
- Does it use the data efficiently (fast convergence of parameter estimates)?
- Is it practical to compute?
- Is it robust to practical measurement anomalies (glitches, gaps etc.)?

A key role that the modelling approach fulfills is the need for timing specifications and standards. The critical issues in this role are:

- Is the parameter set complete (does the model allow for timing signals that meet the parameters but have undesirable attributes)?
- Is the parameter set excessive (does the model over constrain the design of the network or the elements)?
- Does the model permit networking of elements and subnetworks (can one interconnect clocks and/or subnetworks and determine overall performance)?

2. MEASUREMENT TOOL DESCRIPTION

Three measurement approaches are compared in the paper:

1. M[R]TIE (Maximum [Relative] Time Interval Error) ,
2. TVAR (Time Variance),
3. ZTIE (see 2.3)

2.1 M[R]TIE

M[R]TIE is the established measurement approach for telecommunications. The bracketed [R] (Relative) is used to distinguish the reference timing signal. When the term is used with the [R] (MRTIE) the reference timing signal is the input to a timing element and the signal under test is the network element output. When no [R] is used (MTIE), the reference is considered to be ideal (practically a primary reference source). For simplicity, MTIE will be used as a general description of both measurements.

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

MTIE is fairly straightforward to describe. For a given observation interval (usually termed S) the phase error between two timing signals is observed. The peak to peak phase (delay) variation of the error signal is the MTIE for that sample interval. The observation interval can be considered as a window function. As the window slides through the phase error signal, the next sample of MTIE is calculated. The MTIE for the overall phase data set is the maximum of all the individual MTIE samples. For full overlapping processing of MTIE, there is a greater than linear growth in computation resulting from the need to search for minimums and maximums.

MTIE appears to be a useful approach when considering conventional network controlled slip rates. The maximum peak to peak delay variation impressed on a frame alignment (slip) buffer is directly related to MTIE . However, the controlled slip mechanism is a function not only of peak to peak buffer movement, but also of the type of movement. Monotonic movement will results in a slip rate governed by the frame size (typically 125 microseconds) and the rate of phase movement (frequency offset). But, cyclical movement results in slip rate governed by the hysteresis in the buffer control (which can be as small as 18 microseconds). MTIE provides little information regarding the underlying model controlling the phase motion in the buffer. The basic attributes of MTIE are:

1. Monotonically increasing with increasing observation intervals .
2. Suppresses Phase bias; asymptotically converges to frequency or drift bias for long observation intervals.
3. Captures peak information regarding delta time variation.

2.2 TVAR

TVAR (Time Variation) is a new term to some, but it is based on the well established Modified Allan Variance (MVAR). TVAR can be viewed as a filtering operation applied to the raw phase error signal. The filtering operation is diagrammed in Figure 1. The following notation is used:

1. f is normalized in units of T^{-1} where;
2. T is the observation interval (averaging int
3. $T = n \bullet T_0$ where T_0 is the sampling perio and n is number of samples in the observation in

The TVAR filter can be viewed as consisting of three sections:

1. Block Averaging
2. First Difference
3. Second Difference

The block averaging provides a T depended low pass filter characteristic. The transfer function of the block averaging function is illustrated in Figure 2.

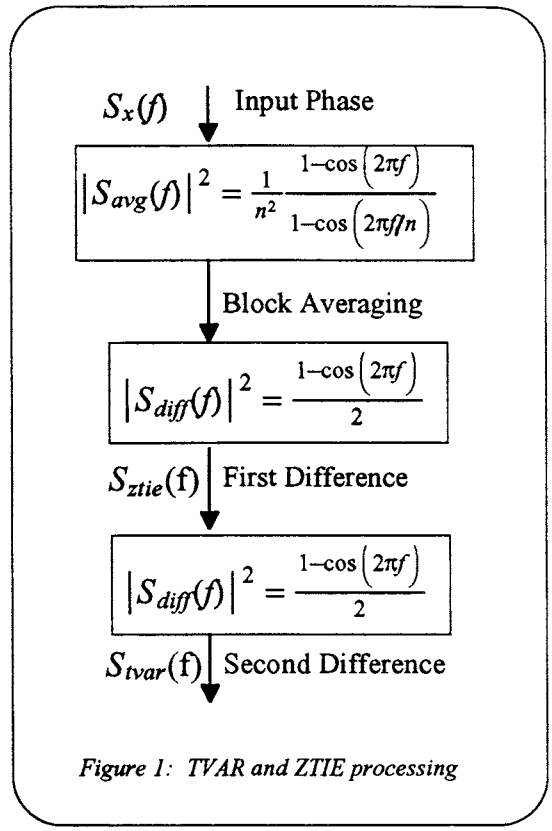


Figure 1: TVAR and ZTIE processing

The block averaging function has the characteristic main lobe with a zero at 1/T. The main function of the block averager is to provide a Tau depended low pass filter to suppress jitter and provide for discrimination between flicker PM and white PM. As will be shown in section 3 this low pass function is critical in telecommunications as timing signals can be dominated by these higher frequency noise components.

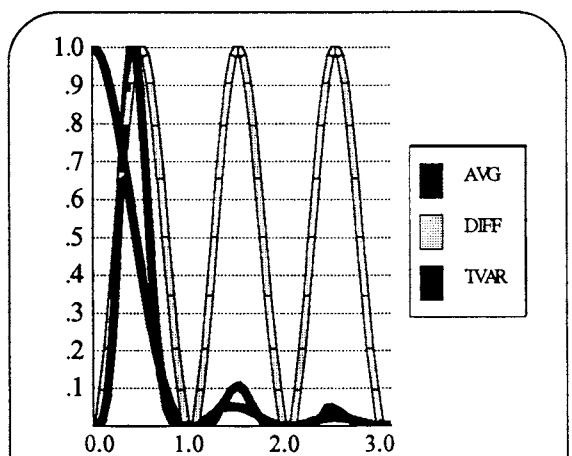


Figure 2: TVAR and ZTIE Transfer Function in units of 1/T.

The next operation is the first difference. The first difference transfer function is shown in Figure 2. The main purpose of the difference operator is to provide suppression of the f^{-n} divergent power law noise processes. The first difference provides 20 dB per decade suppression . This is adequate to

ensure a stationary noise power process for input phase noise with white noise FM as the most divergent noise (which is the expected case for traceable timing signals). The first difference operation also suppresses phase bias. In TVAR processing the difference operator is applied twice yielding an overall second difference. This ensures stationarity for noise processes up to random walk FM which is the most divergent noise expected in general oscillator applications. The second difference also suppresses frequency bias . Since frequency bias is a critical parameter, TVAR needs to be complemented with another measurement in describing telecommunication timing signals.

The composite TVAR transfer function is shown in Figure 2. It is essentially a bandpass filter with a center frequency and bandwidth controlled by the averaging time T. To complete the calculation of TVAR the variance of the filtered output signal is calculated for each observation interval (T). A close analogy can be drawn to an FFT based spectral analysis. The bandpass filter function can be viewed as a window function centered at a specific frequency (approx 1/2T). Normally TVAR is calculated in a geometric series of $2^{-n}T_0^{-1}$ frequency bins. An FFT is based on a arithmetic series of2^n frequency points. In an FFT the bandwidth of the window function is fixed. In TVAR analysis the bandwidth is proportional to the frequency bin. TVAR is tailored to being an efficient method to extract broadband power noise processed. FFTs are in general better suited for analyzing frequency over a smaller dynamic range.

2.3 ZTIE

The basic concept behind ZTIE is that a peak power measure is needed to complement TVAR. ZTIE is one possible approach to providing a bandpass measure of peak power. ZTIE measures the peak power using a bandpass filter controlled by the averaging time T. Figure 1 shows that ZTIE is calculated from the output of the first difference function used in calculating TVAR. The reason that the first difference output was selected over the second difference is that it does not suppress the frequency bias term of the model. The importance of this will be seen in the examples in the next section. ZTIE is calculated by simple measuring the peak output of the first difference operator . The peak information is critical to providing bounds to the model parameters. As the examples will show, it is a risky assumption to assume that the peak to rms ratio is guassian in telecommunication networks.

3. MEASUREMENT EXAMPLES

This section provides a selection of measurement examples to illustrate the key attributes of each measurement method. The measurement examples were generated using a simulation approach.

3.1 Power Law Noise Examples

Figures 3a,3b and 3c illustrate the measurement behavior for White PM, Flicker PM and White FM respectively. Each simulation was run for 86,400 sample points.

Figure 3a shows the results for a guassian white PM input signal. In addition to the TVAR, ZTIE and MTIE results, an

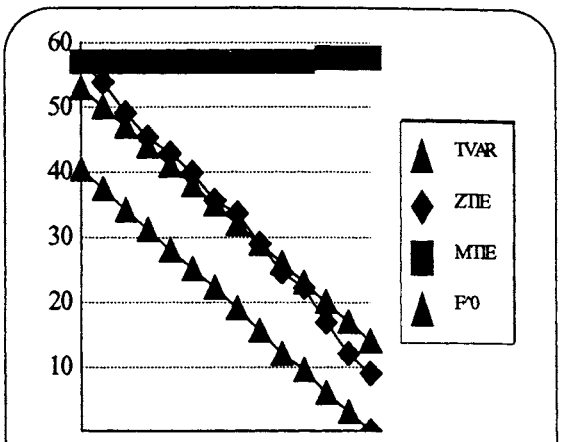


Figure 3a: Simulation results white PM case.

asymptote showing the -1/2 expected slope is plotted. The TVAR result shows the expected behavior:

1. At T_0 the TVAR (actually the root TVAR) value 40dBns (100ns) equals the rms level of the input jitter.
2. The -1/2 slope shows the root law reduction in rms error associated with averaging .

The ZTIE result shows the same trend as the TVAR data. The slight droop at longer averaging times is explained by the reduction of independent peak power measurements (less measurements reduces the likely of finding a point on the tail of the distribution). The TVAR curve is running about 14 dB above the ZTIE curve showing a well behaved peak to rms ratio.

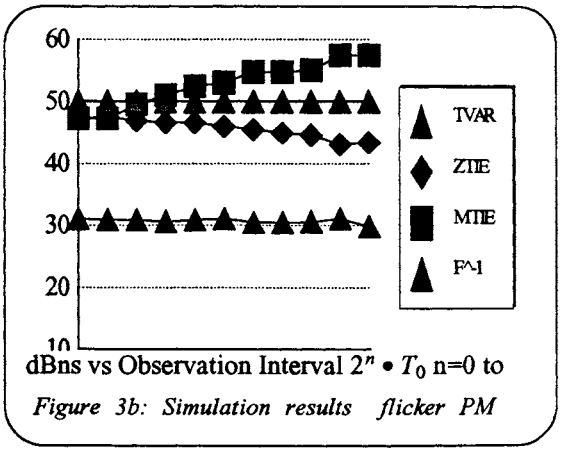


Figure 3b: Simulation results flicker PM

The MTIE result highlights the inadequacy of MTIE in extracting useful data for white PM noise. The MTIE shows a slight increase for longer observation intervals Naturally. since the bandwidth of the measurement increases with observation interval there is increasing probability of finding a higher noise peak.

Figure 3b shows the results for Flicker PM. The TVAR results shows the flicker PM noise floor as a horizontal asymptote. This is a very practical feature of TVAR. The flicker PM floor represents the best time error calibration that can be realized for a given timing signal. This is

important when consider issues like phase error during rearrangement of clock timing inputs. A clock cannot resolve phase better than the flicker PM noise floor of the reference input.

ZTIE shows a similar behavior as with white PM. The peak to rms ratio is well behaved. The peak information is very useful in this case as it bounds the peak phase error

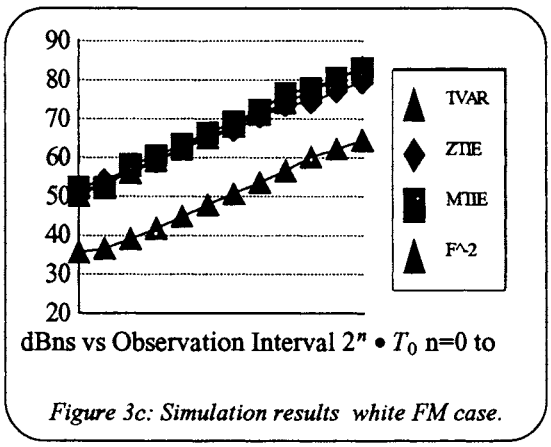


Figure 3c: Simulation results white FM case.

achievable during rearrangement (approx 220 ns in this case).

MTIE behaves much better with a divergent noise process input. However it still shows a 10 dB positive skew for longer observation intervals.

With white FM all three measures show the +1/2 slope divergence of the timing error expected. Note that ZTIE and MTIE converge for white noise FM. ZTIE has been designed to converge with MTIE for longer observation intervals where either white FM or frequency bias dominate.

3.2 Telecommunication Network Examples

The examples in this section are constructed to illustrate issue related to the telecommunications network.

3.2.1 Transient Glitch

In measuring actual network timing data there is the practical issue of managing transients. The transient may be part of the measurement system, disruption in the transport path or a clock anomaly. Regardless, a good measurement

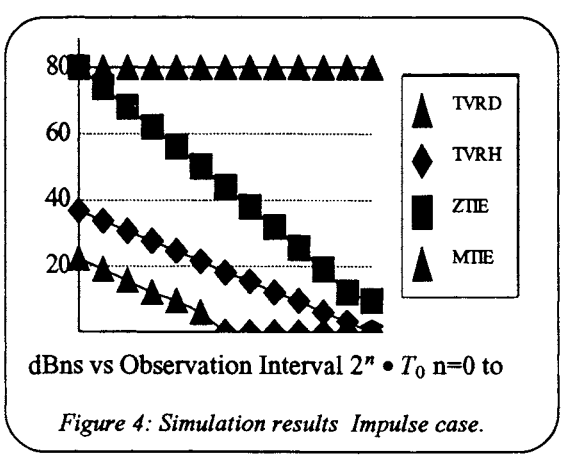


Figure 4: Simulation results Impulse case.

approach should not completely mask transients or let transients mask steady state performance. Figure 4 shows the response of the three measurement approaches to a simple phase impulse event.

TVAR is plotted for two data collection intervals (one hour and one day) . Since TVAR is an rms power estimator it tends to remove transient impulses (which are finite energy zero average power). Thus, TVAR needs to be complemented with a peak estimator to capture transient information.

ZTIE shows and interesting response to the impulse. The total peak level of the impulse is shown at T_0 (10,000 ns). The power level is then linearly suppressed for longer observation intervals.

MTIE is extremely sensitive to transients. A single transient can completely mask out the underlying noise model for the timing signal. This is a serious weakness in the application of MTIE to actual network timing analysis.

3.2.2 Jitter plus Frequency Offset

Figure 5 illustrates an example of a realistic network timing signal .

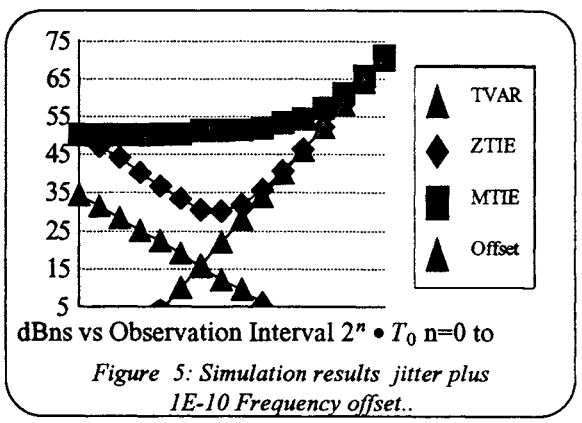


Figure 5: Simulation results jitter plus 1E-10 Frequency offset..

The timing signal is constructed with nominally 1UI of peak to peak timing jitter. This level of jitter is well within worse case network limits (5UI) , and is a reasonable bound for expected network transport jitter over the existing plesiochronous digital network. The frequency bias (1E-10) represents a timing signal that has impaired frequency traceability to the primary reference source (the next example suggests one of a number of mechanisms for this to occur).

TVAR effectively suppresses the frequency offset and correctly reports the underlying white PM transport jitter. ZTIE is very effective in showing both the short term peak jitter noise component and the long term frequency bias. MTIE again is inadequate to show the true underlying noise process.

This is a good example to consider from the point of view of timing specifications and standards. One can consider how well this measurement results would function as network timing interface standards. A key question is how well one can calibrate the frequency of slave clock to a network timing reference.? The ability to calibrate directly impacts

slip performance during reference outage conditions. For sake of illustration lets consider a clock with a 128 second averaging algorithm. At 128 seconds the MTIE value is 488 ns. Since the MTIE mask provides no insight on the underlying noise process, one can't assume white PM. In fact the only safe assumption is to assume that this 488 ns phase instability cannot be effectively filtered. This leads to frequency calibration error bound of 3.8 E-9 . In contrast, at 128 seconds ZTIE has a value of 30 ns . A mask based on ZTIE would constrain the worse case frequency calibration to 2.4 E-10 (over an order of magnitude tighter).

3.2.3 Network Timing Signal (Stress Condition)

Figure 6 shows the measurements results for a network timing signal under stress conditions.

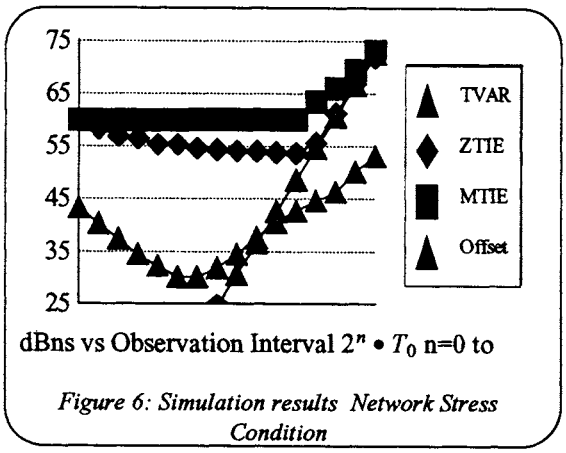


Figure 6: Simulation results Network Stress Condition

CCITT G.812 defines a clock under stress conditions. One stress condition is frequent transient disruption on the reference transport. For constructing the simulation for network stress one reference disruption per hour was used. This level of disruption is within the maintenance performance action limits for network transport. The disruption activity will lead to reference rearrangement transients. The following model for reference rearrangement was simulated:

1. The maximum total phase movement at the clock output in response to a transient is assumed to be either plus or minus 1 UI (488ns). This is within the current maximum of 1000 ns allowed for in network timing standards.
2. The direction of phase step is assumed to be bias with positive steps occurring with a 75 % probability. Bias errors can result from numerous hardware and software sources in a clock.

In addition to the hourly reference disruptions, white PM was added to timing signal to simulate network transport jitter (the jitter distribution was assumed to uniform with a 1UI peak to peak magnitude.

The TVAR result clearly shows two dominant noise processes. The the white PM transport jitter dominates for short observation intervals., and the white FM associated with the reference rearrangements dominates for longer observation interval.

The ZTIE result shows that timing signal error distribution is definitely non-guassian. The random walk in phase is quantized in 1 UI steps. These phase steps manifest themselves as impulsive noise. ZTIE complements TVAR by showing that the peak error is dominated by non-filterable transient steps rather than the underlying white FM noise process. The TVAR results suggest that the best time calibration error is 32 ns. ZTIE shows the true achievable timing calibration performance is actually limited to 500 ns.

MTIE performs fairly well for this constructed example because the phase walking process dominates. However, ZTIE still provides a correct estimate of the underlying frequency offset (which in this case is 6E-11) about ten times faster than MTIE.

4. SUMMARY

Measurement and modelling are closely related. The use of the power law model in conjunction with TVAR as a measurement technique provides a very effective platform to characterize network timing performance. However, TVAR is inadequate for establishing network timing specifications and standards as it fails to extract important information concerning peak noise power, frequency biases and transients. MTIE has been shown to be inadequate as measurement tool for a variety of practical network scenarios. ZTIE is proposed as one measure of peak power. It complements TVAR by providing information on peak noise power, frequency bias and transient activity. It is design to converge with MTIE for longer observation intervals. The use of both TVAR and a bandlimited peak power measurement like ZTIE in telecommunication applications is increasingly critical as new synchronous optical transport (SONET and SDH) require more accurate and tighter specification of network timing.

Session 10B

Atomic Frequency Standards: Miscellaneous

Chairman: F. Walls

A COMPACT AND HIGH-PERFORMANCE REFERENCE CELL FOR LASER FREQUENCY STABILIZATION

G. Dudle and P. Thomann

Observatoire Cantonal de Neuchâtel
Avenue Beauregard 3
CH - 2036 Cormondrèche, Switzerland

Abstract

Frequency stabilized lasers are required for many applications such as length metrology, telecommunications and optically pumped atomic frequency standards. Linear and saturated absorption atomic vapor cells are traditionally used as frequency discriminators to lock the laser frequency to an atomic transition (e.g. the D1 and D2 lines in Cs or Rb). The problems associated with these discriminators may be overcome by using a short and intense atomic beam as a reference. Theoretical considerations show that with a proper design of the atomic beam collimator, an atomic beam discriminator can combine a small residual Doppler width with a high signal-to-noise ratio and signal-to-background ratio. The experimental results obtained on a Cs beam show that this high-performance discriminator can be realized in practice in a very compact sealed cell, comparable in size with a simple absorption cell.

Keywords: laser stabilization, frequency discriminator, atomic beams.

1. INTRODUCTION

In usual cesium vapor cells, the Doppler width of the absorption line does not permit to resolve directly all of the six hyperfine transitions between the $6S_{1/2}$ and the $6P_{3/2}$ level (D_2 line). In many applications where stabilized lasers are required it would be very useful to have a frequency discriminator with less Doppler width. One solution to this is the method of saturated absorption [7]. The problem with this technique is the background slope due to the linear absorption signal which can be removed electronically. Since this background depends strongly on the temperature it is not easy to ensure a good signal stability. Another way of removing the Doppler broadening is the atomic beam technique, which usually implies a vacuum system of much larger size than an absorption cell and its integration into a experiment to stabilize a laser is therefore often too complicated.

It seems therefore attractive to combine the advantages of an atomic beam (no Doppler background and reduced Doppler width) and of a vapor cell (small size and of simple use) in a sealed cell which could find applications in many fields where stabilized lasers are necessary.

The design of the short beam cell (SBC) is described in section 2, then a theoretical estimate of the residual Doppler width and of the signal intensity is outlined (section 3). In section 4 the first results obtained with the SBC are presented and discussed.

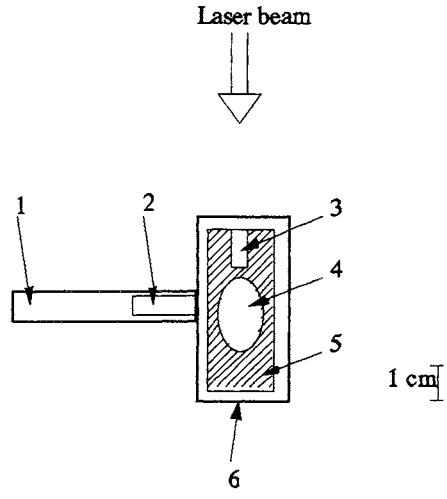


Fig. 1: Schema of the short beam cell: 1. Cs source, 2. Collimator, 3. Bulk getter, 4. Aperture for fluoresence light, 5. Graphite, 6. Optical window for laser passage

2. CONSTRUCTION OF THE CELL

The basic idea to reduce the Doppler width is to imitate a atomic beam in the cell. In a large atomic beam device three parts can be distinguished: i) the cesium source, ii) the collimator and iii) the interaction region. Figure 1 shows how these three components are realized in the sealed cell. The SBC is divided into two separate chambers. In a first one the cesium source is placed, the second encloses the interaction region. The two parts of the cell are connected by a collimator similar to the exit of a conventional cesium oven.

Source: As source we use a mesh of stainless steal on which about 1g of cesium is distilled. The stainless steal mesh acts as sponge for the Cs and prevents the metal from spilling into the collimator.

Collimator: Atoms having a velocity parallel to the laser beam induce a Doppler shift of the absorption spectrum. Velocities perpendicular to the laser propagation direction do not influence the width of the observed signal. For this reason we designed a one-directional collimator with a

directivity of 25 parallel to the light propagation and much less (2.5) perpendicular to that. This design provides a larger atomic flux then a two-directional (e.g. cylindrical) collimator would. In practice the collimator is made of several parallel plates forming 5mm long channels with a rectangular cross-section (0.2mm x 2mm) (fig. 2). Whereas a two-directional collimator has no preferred orientation angle around the axis perpendicular to the light, the one-directional collimator must be aligned in order to have the parallel plates in planes perpendicular to the laser beam. A theoretical evaluation of the angular distribution of the atomic beam produced by this collimator is made in section 3.

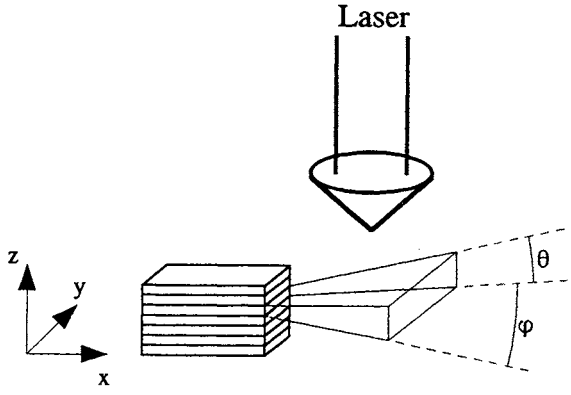


Fig. 2: Outline of the collimator

Interaction region: The fact that the three Cs-transitions can be resolved is due to a small Doppler width of the absorption (fluoresence) line. Only atoms of the beam should be able to interact with the laser beam. In other words the cesium which reaches the end of the cell should be stopped there, to avoid the residual cesium pressure to grow. In order to provide the required high pumping speed for cesium, the interaction region is machined out of a block of graphite, a material known to be a good getter for cesium. The atomic beam enters through a radial aperture in the graphite. Another hole, along the axis of the cylindrical bulk allows the laser beam through the cell. In this direction absorption measurements can be done. For the fluoresence detection a second radial aperture is made in the graphite. A bulk getter is added in the interaction region to prevent the pressure buildup due to outgassing from the walls of the cell.
The whole cell is fabricated in Pyrex glass. The two windows for the passage of the laser beam are flat and of good optical quality. No special anti-reflection-coating has been added. All parts enclosed in the cell (stainless steal mesh, collimator, graphite) are outgassed before assembly. In addition, the whole cell is carefully outgassed before introducing the Cs fill.

3. LINESHAPE CALCULATION

The calculation of the lineshape in a atomic beam is well known [2,3,4,6]. The intensity of the atomic flux is essentialy decreasing with $1/r^2$, r being the distance between the source and the interaction region. In the SBC case these calculations are no longer valid since a non negligible part of the atoms crossing the laser beam have struck the collimator wall on their trajectory, i. e. have been diffused. For this reason the calculation has to be redone. First one has to evaluate the angular distribution of the atomic flux produced by the collimator. In a second step the number of atoms at resonance for a given detuning $\delta\omega$ is calulated taking into account the Maxwellian distribution of velocity in the atomic beam.

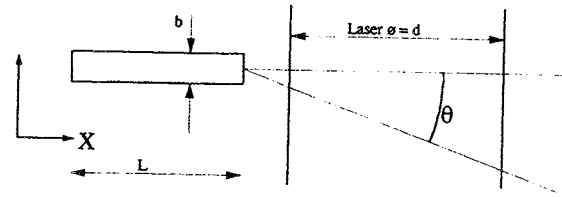


Fig. 3: Geometry of the collimator

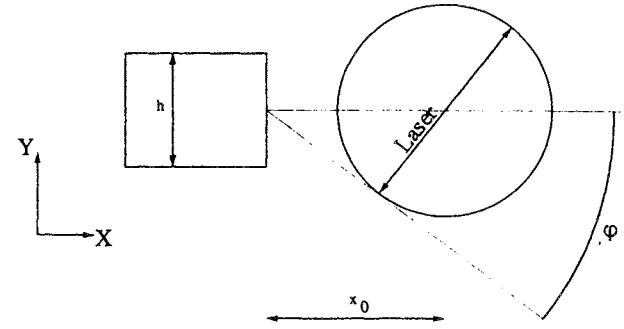


Fig. 4: Geometry of the collimator

Angular distribution: For the calculation of the angular distribution several assumptions are made. First of all the pressure in the source is low enough that the flow of atoms is molecular, i. e. the mean free path of the Cs atoms is much longer than the largest dimension of the collimator. Second, it is assumed as usual, that atoms colliding with the collimator walls are reemitted with a Lambertian distribution. Third, the density profile of atoms between the entry and the exit of the collimator is assumed to be linear, which is true to first order in b/L. The last assumption is specific to the type of collimator described here: since the section of the tube is a rectangle with one side much smaller than the other, atoms emitted from the small wall will be neglected.

The angular distribution takes a different analytical form depending on the value of the parameter a = (tg θ) L/b

$$\text{for } a < 1 \quad \frac{dN}{d\theta} = \frac{n_0 h d^2}{16} \frac{b}{L} \left\{ \frac{1-a}{1+\frac{x_0}{L}} + a\cos\theta \left[1 - \frac{x_0}{L} \ln\left(1 + \frac{L}{x_0} \right) \right] \right\} \qquad (3.1)$$

$$\text{for } a > 1 \quad \frac{dN}{d\theta} = \frac{n_0 h d^2 \cos\theta}{16} \frac{b}{L} \left\{ 1 - \frac{a x_0}{L} \ln\left(1 + \frac{L}{a x_0} \right) \right\} \qquad (3.2)$$

Equation 3.1 has two contributions: the first term of the sum between braces corresponds to atoms moving through the collimator without hitting the walls, the second term

corresponds to atoms reemitted by the collimator wall. For larger θ (a > 1, eq. 3.2) the first contribution vanishes.

dN/dθ gives the number of atoms crossing the laser beam at distance x_0 from the collimator exit, emitted at an angle θ in the xz plane. The angular density dN/dθ results from an integration over range of angle φ which is determined both by the collimator geometry and the laser beam diameter and position.

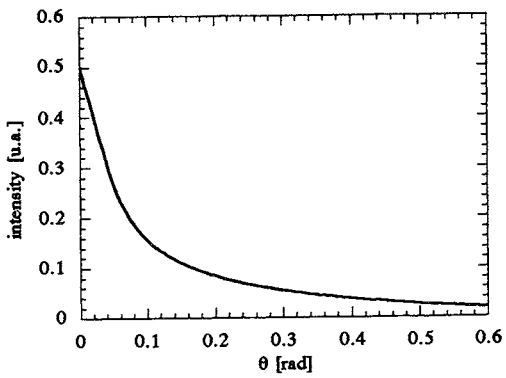


Fig 5: This figure shows the angular distribution of atoms in a typical configuration of the cell (equation 3.1 and 3.2) with the values L = 5mm, h = 2mm, b = 0.2 mm, x_0 = 5mm, d = 3mm

Detuning δω: Atoms having a velocity component parallel to the laser propagation direction need a laser-atom detuning $\delta\omega = \omega_L - \omega_{Cs}$ to be at resonance. A given detuning can be realized by different combinations of v and θ. To take into account that the atomic beam is formed by a thermal distribution of velocities, one has to integrate the Maxwellian distribution over all v. Therefore the detuning δω is expressed as δω = kv sinθ and the product of angular- and Maxwellian distribution is integrated over θ instead of v, yielding

$$\frac{dN}{d\delta\omega} = \frac{4}{\sqrt{\pi}} \int_0^{\theta_{max}} \frac{dN}{d\theta}(d\theta) \frac{\delta\omega^2}{(k\alpha \sin\theta)^3} \exp\left[-\left(\frac{\delta\omega}{k\alpha \sin\theta}\right)^2\right] d\theta \qquad (3.3)$$

This function gives the number of atoms at resonance per unit detuning at a given detuning δω. The upper integration limit is determined by the detection mode. In the case of fluoresence only a restricted area of the interaction region will be imaged on the detector. Atoms moving with an angle $\theta > \theta_{max}$ (angle determined by the detection geometry) are not seen by the detector and therefore do not contribute to the signal. In the case of absorption all atoms crossing the laser beam will contribute to the absorption spectrum, irrespective of the angle θ. The absorption and the fluorescence signals should thus show a slightly different behaviour for large values of δω. As shown in fig. 5 the number of atoms with a large Doppler detuning (mostly atoms with large angle θ) is small even for the very short atomic beams described here. As a consequence the expected discriminator curves, as detected in absorption or in fluorescence have nearly the same shape and width (see fig. 6). In the two cases the FWHM is ≈ 20MHz.

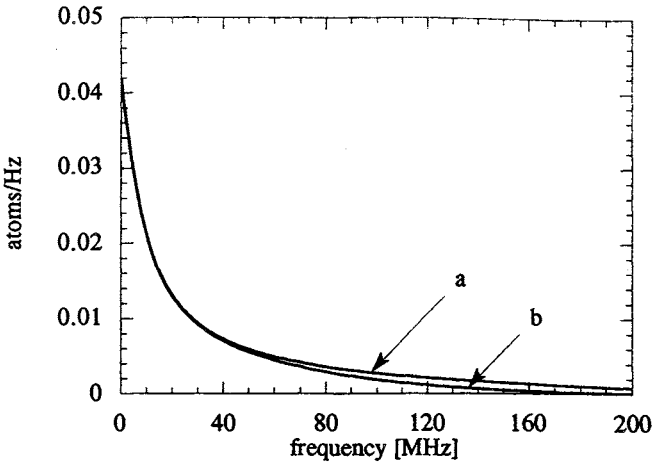


Fig. 6: Spectal density of atoms contributing to the signal. a) absorption signal, b) fluorescenc signal. Numerical parameters are the same as in fig. 5.

The sharply peaked distributions shown in fig. 6 must be convolved with the atomic absorption lineshape and the laser spectral intensity in order to obtain the actual absorption and fluorescence signals. In the case of fluoresence detection, the power received on the photodetector is given by

$$P_{fluo}(\omega) = \frac{\Omega}{4\pi} P(\omega) \int_0^{\infty} \frac{dN}{d\,\delta\omega} \frac{\sigma(\delta\omega)}{A} d\,\delta\omega \qquad (3.4)$$

where Ω is the solide angle of the collection, A is the section of the laser beam and P is the incident laser power assumed to be low enough, such that the hyperfine pumping is negligible. If the function dN/dδω is much broader than both the natural linewidth and the laser linewidth (here 5 MHz for each) Equation 3.4 can be simplified to

$$P_{fluo}(\omega) = \frac{\Omega}{4\pi} \frac{dN}{d\,\delta\omega}(\delta\omega = 0) \frac{\sigma_0}{A} P_{laser}(\omega)\ \Delta\omega_L \qquad (3.5)$$

where $\Delta\omega_L$ is the laser linewidth. With the values Ω = 1.9 sr, $\sigma_0 = 67\ 10^{-15}$ m^2 for the transition 4 -> 4', $\Delta\omega_L = 2\pi$ 5MHz, A =7 10^{-6} m^2 and P = 2.12 μW (corresponds to an intensity below the hyperfine pumping limit of 30μW/cm^2) P_{fluo} = 630 pW at room temperature and with the geometrical values of Figure 5.

4. PRELIMINARY MEASUREMENTS

To test the device we use a DBR type laser studied elsewhere [1]. A special feature of this laser is to be tuneable over a wide frequency range without changing the optical output power. The linewidth of the used laser is about 5 MHz.

Figure 7 shows the fluorescence light as a function of laser frequency obtained with the SBC. Several points may be noted on this figure. The baseline of the signal is completely flat and does not show the behaviour of a vapor cell. The absence of a Doppler background indicates that the vapor pressure in the interaction region is negligible. In other words the graphite is able to absorb all the Cs

arriving in the interaction region. A second point to be outlined is the low background light level. In the present case the fluoresced power on the 4 -> 4' transition is of the same order as the background. Finally, the linewidth of the signal is about 50MHz. This allows clearly to resolve the three transitions hidden under the Doppler broadened absorption signal of a vapor cell. However, it is higher than predicted by the theoretical angular distribution. Although the absence of a wide Doppler background precludes the possibility of Cs - Cs collisions, the effect of weak collisions from foreign gas molecules cannot be ruled out. Indeed, in a first cell the linewidth was observed to increase to 150 MHz in a few days. The subsequent addition of a bulk getter has markedly improved the long-term behaviour which is presently under test.

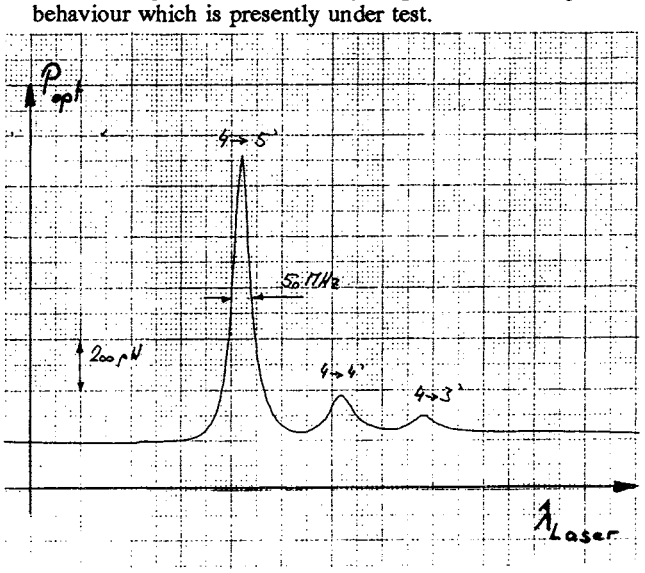


Fig. 7: Fluoresence light obtained experimentally with a SBC.

In order to compare the perfomances of the SBC and a traditional vapor cell we estimate the stability in the two cases. A laser locked on an atomic resonnance with a width $\Delta\omega = \omega/Q$ has a theoretical stability, characterized by the Allan variance $\sigma_y(\tau)$.

$$\sigma_y(\tau) \sim \frac{1}{Q\frac{S}{N}} = \frac{1}{\frac{\omega}{\Delta\omega}\frac{S}{N}} = \frac{1}{\omega} N \frac{\Delta\omega}{S} \approx \frac{1}{\omega}\frac{N}{D} \qquad (4.1)$$

where D is the maximum slope on the peak. In a vapor cell (as in a saturated absorption cell) the noise is dominated by the shot noise contribution. This is a result of the relatively high background light level in this mode. The ratio N/D is typically 400 for linear absorption and 150 - 500 for saturated absorption depending on the chosen transition [5]. In the case of the SBC, due to the fluoresence mode, the shot noise is negligible and the limitation comes from the thermal noise of the photodetector. The ratio N/D is about 1000 for the 4 - 4' transition and 300 for the 4 - 5' transition, which is of the same order as the two other discriminators. However, the center frequency and the width of the signals obtained in an atomic beam are inherently independent of various conditions that affect linear and saturated absorption signals, such as laser intensity and cell temperature; this is a significant advantage in applications where accuracy is important. Moreover, an improvement of the N/D ratio by at least one order of magnitude may be expected in the case of the SBC by increasing the operating temperature which is not possible in the linear and satureted absorption cells (the Cs consumption at room temperature in the SBC is only 5mg/year).

5. CONCLUSION

We designed a new frequency discriminator for laser stabilization, called Short Beam Cell (SBC). The SBC has the following advantages: it is simple to use, it requires a minimum number of optical components and it allows to separate all the hyperfine transitions of the D_2 line of Cs. Based on the preliminary results obtained so far, the theoretical stability of a laser locked to the SBC would be somewhat lower than using saturated absorption. However, this technique has the advantage that no Doppler background is superimposed on the signal. The laser frequency is therefore insensitive to variations of cell temperature or laser intensity. In addition further improvements are possible by increasing the atomic flux either by heating the cell or by using a larger collimator. Since the difference between the fluoresence and absorption signal seems to be insignificant, the SBC could also be used in the absorption mode. A problem to be solved is the residual total pressure in the sealed cell which is responsible of the long term degradation of the linewidth.

REFERENCES

1. Sagna N. & al. 1992, Noise Measurements in Single-Mode GaAlAs Diode Lasers, Proc 6th European Frequency and Time Forum, Noordwijk, to be published.

2. Giordmaine J. A. & Wang T. C. 1960, Molecular Beam Formation by Long Parallel Tubes, J. of App. Physics, 31, 3, 463-471.

3. Beijerinck H.C. W. & Verster N.F. 1975, Velocity distribution and angular distribution of molecular beams from multichannel arrays, J. of App. Physics, 46, 5, 2083-2090.

4. Yu. G. Abashev & al.1976, Investigation of the Angular Distribution of Cesium Atoms Diffusing From Rectangular Channels, Izmeritel'naya Tekhnika, 11, 37-39.

5. Ohshima S. I. & al 1987, Spectral Width of Saturated Absorption Spectra of Cs with a Laser Diode, IEEE J.or Quantum Electronics, 23, 5, 473-475.

6.Ramsey N. F.1956, Molecular Beams, Clarendon Press, Oxford.

7 Demtröder W. 1981, Laser Spectroscopy, Springer-Verlag, Berlin.

ACKNOWLEDGEMENTS

This work was in part supported by ESTEC under contract number 9099/90/NL/US(SC) and in part by the Swiss National Fund for Scientific Research.

LINE INHOMOGENEITY EFFECTS AND POWER SHIFT IN MINIATURIZED RUBIDIUM FREQUENCY STANDARD

G. Mileti, I. Rüedi and H. Schweda

Observatoire Cantonal de Neuchâtel
Avenue Beauregard 3
CH - 2036 Cormondrèche, Switzerland

Abstract

The role of the line inhomogeneity in determining the overall performances of a compact rubidium frequency standard is analyzed. The possibility of zero power shift operation is discussed and the effects of the main physical parameters of the standards on this particular operation mode is investigated experimentally.

Keywords: rubidium standard, microwave resonator, power shift.

1. INTRODUCTION

The Observatoire Cantonal de Neuchâtel is currently developing a compact rubidium frequency standard with variants for spatial and industrial applications. It is designed for small size, light weight, low power consumption, fast warm-up along with minimal environmental sensitivities. The main characteristics of its physics package include:

- use of the integrated filter technique (IFT).
- use of a magnetron-type microwave resonator.
- integration of the complete optical package in a sealed tube.

In this paper we will not go much into details of the technological or mechanical aspects of the standard. We are more interested here to investigate some of its physical degrees of freedom, and study how to fix their value in order to optimize the performance. Nevertheless, due to the fact that the physical behaviour of the standard is strongly connected to the structure of the device under study, we will give a description of its principal components i.e. the new magnetron-type resonator, which is the main innovation.

It is well-known that in order to have a minimum frequency drift, and/or a reduced environmental sensitivity , and maximal stability, an adjustment of some physical parameters of the standard has to be done. For example the cell temperature, the lamp temperature, the light intensity, the C-field and the micro-wave power have to be chosen properly. Although the individual influences of these parameters on frequency are pretty well understood, in their combined effect they appear as an intermixture with all the others. On the other hand, the long-term drift is still only partly understood. It is believed that it is also due to the mentioned sensitivities, and in particular to the power shift (Ref. 1).

The paper gives an illustration of the optimization of physics parameters in our particular design, and focuses on details of the problem of power shift. It is organized in the following way. Chapter 2 describes the new resonator. Chapter 3 gives a general overview on the knowledge of the power shift, and in Chapter 4 we display our experimental approach and results. The conclusion is in the Chapter 5.

2. PRACTICAL IMPLEMENTATION

2.1. General description.

Lamp and cell are placed in a cylindrical sealed tube. The C-field coils are wound around the cylindrical housing in such a way that line inhomogeneity effects and the resulting shifts are partly compensated. The tube is surrounded by a closed magnetic shield made of two half cylinders which determines the external dimensions: diameter of 32mm, length of 57 mm.

2.2. The lamp.

The RF discharge lamp contains both isotopes of rubidium with 2 mbar argon. The isotopic ratio has been chosen to minimize the light shift and the temperature coefficient at the same temperature of the absorption cell.

The bulb is surrounded by a cylidrical coil which forms the inductive load of a RF-power oscillator. It is a straight-forward Clapp oscillator which operates at a typical DC power of 1,2 W. At this power and with an operating temperature of 140°C the emission lines show already self-absorption.

In order to quantify the light level incident on the absorption cell, we use a small photocell placed near its entrance. In the following the term "incident light" refers to the voltage produced by this photocurrent at the 10kΩ resistor of the current-to-voltage converter used with this detector.

2.3. The cell.

The cylindrical cell is made of quartz. It contains natural rubidium, and a buffer gas. Its normal operating temperature is 83°C. The pressure of the buffer gas has been chosen in accordance with the isotopic ratio in the lamp in order to give a zero temperature coefficient and zero light shift at the same temperature (Ref. 2).
As we use the integrated filter technique (IFT), the entrance region of the cell acts essentially as filter. In reality, there is no sharp boundary between the two regions acting

respectively as filter and as resonance cell . We can only say that the number of pumped atoms increases from the entrance to the end of the cell, and this has to be taken into account in a modellisation.

At first approximation one can assume an exponential decay of the light along the cell. Average absorption coefficients as measured over the whole cell length are $K(80°C) = 0.37\ cm^{-1}$ and $K(83°C) = 0.42\ cm^{-1}$.

In a rough approximation, we can also assume that the number of pumped atoms in the cell increases exponentially with this factor.

2.4. The resonator.

The best choice for a microwave resonator in a rubidium frequency standard has been a cylindrical TE_{011} cavity since it has the most favorable field geometry.
However, the rubidium resonance wave length of $\lambda_0 = 4{,}4$ cm leads to relatively large dimensions for this cavity. In order to relax these constraints most rubidium standards fabricated today use a TE_{111} cavity with less favorable field geometry. If one wants to reduce the size even further without sacrifieing a good field distribution, one can partially load the cavity with a dielectric material of high permittivity ε. However, such dielectric loading inevitably introduces losses leading to a reduced cavity quality factor.

An alternative way is to load the cavity with some inductive-capacitive structures. Since such structures can be made completely metallic, no significant losses are introduced. Furthermore, the structures can be readily connected to the cavity envelope which simultanously acts as the cell oven. In such a way, efficient heat transfer to the cell and a good mechanical stability is preserved. Earlier developments which used isolated helical structures or loop-gap electrodes bounded to the cell did not show the excellent thermal homogeneity and short warm-up time reached with the present design.

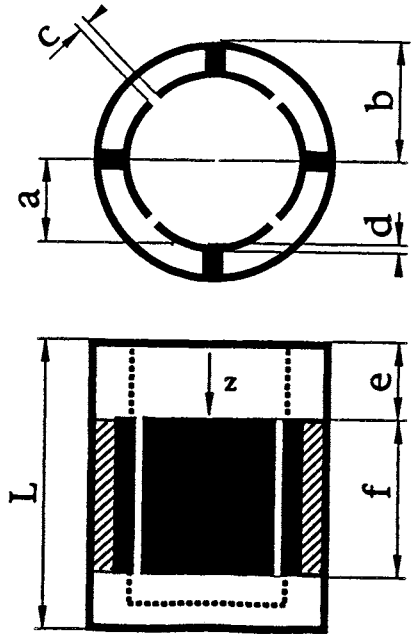


Fig 2.2:The magnetron resonator.

Fig. 2.2. shows the resonator used in the present set-up. The resonating structure consists of four quarter-cylindrical electrodes placed equally spaced around the absorption cell. The electrodes are held in place by four radial spacers. All parts are made of copper. The electrodes form a slotted inner cylinder coaxial to the cavity envelope. The inductance represented by this cylinder and the capacitance of the gaps betwen the electrodes determine in first approximation the frequency of the newly created mode.

The resonating structure used here resembles somewhat arrangements which have been tried in early magnetron designs therefore we name our cavity a magnetron-type resonator.

The measured cavity quality factor was > 500. The field distribution ressembles closely the TE_{011} geometry with the axial field concentrated in the region surrounded by the electrodes. The measured variation of the axial field amplitude along the cavity axis showed a cosine square dependence with z centered in the electrodes region.

The atomic response spectrum evidenced the absence of undesired of $\Delta mF = \pm 1$ transitions, which only occur when the oscillating magnetic field is perpendicular to the C-field. Less then 1 μW microwave power at 6,8 GHz was needed for an optimal clock signal.

3. INHOMOGENEITY EFFECTS

Due to the presence of buffer gas in the absorption cell the rubidium atoms are effectively frozen in place for times long compared to the stimulated emission period between the clock levels.In general the resonance frequencies differ slightly for atoms in different parts of the cell. In such a situation the resonance signal is an inhomogeneous mixture with contributions from all parts of the cell. An additional consequence is that the measured frequency becomes dependent on the applied RF power since individual atoms contribute to the signal by varying degrees upon saturation and the line center shifts according to RF power level. This power sensitivity depends on local values of the light intensity (lifht shift) and magnetic field strength (quadratic Zeeman shift).

Since the pump light intensity and its spectrum changes along the cell, the levels of the absorbing atoms are shifted by the light shift effects. The degree of shifting depends on the intensity of the pump light and on the mismatch between pump and and absorption profile at the site of the atoms. Thus a gradient of different resonance frequencies along the cell is created.

The pump light arrives at the cell with a definite spectrum. In the separate filter configuration pure isotopes of Rb^{87} and Rb^{85} are used in the lamp and the filter cell respectively. The lamp spectrum consists essentially of two lines corresponding to the optical transitions from the F=1 and F=2 levels of the $^2S_{1/2}$ ground state to the first excited ^2P states. To enable effective optical pumping Rb^{85} is used in the filter to attenuate the F=2 line relative to the F=1 utilizing the accidental overlap of the Rb^{87} F=2 and Rb^{85} F=3 lines. As a result the filtered light entering the absorption cell contains mainly light of the F=1 line wanted for the optical pumping process. In general the filtering process displaces the line in such a way that pump line and absorption line do not coincide exactly. In such a situation light shift occurs.

In the integrated filter approach the filtering takes place in the resonance cell itself which contains natural rubidium. As the light traverses through the cell due to the filtering by the Rb^{85} atoms the pumping efficiency steadily increases. As the pumping becomes more effective, the pump light gradually decreases until the maximum population of the F=2 levels is reached. It is obvious that in this approach the pump light change along the cell is more pronounced resulting in an even greater inhomogeneity. The description of the situation is further complicated by the fact that in IFT both the lamp and the cell contain Rb^{87} and Rb^{85} in different composition to minimize the light shift effect. It should be noted that in both approaches conditions can be founded to zero the overall light shift (by careful adjustment of the buffer gas pressure and temperature in the filter and absorption cell, respectively), but it is obvious that even then, in general the individual resonance frequencies of atoms at different sites in the cell remain quite different.

Frequency gradients along the cell can also be created when the C-field is not homogeneous. Although the frequency gradient caused by the light shift effect cannot be measured directly, one might be able to find a C-field gradient which compensates for the power sensitivity caused by the inhomogeneous light shift. As a matter of fact such empirically found compensating C-field gradients have already been implemented in industrial Rb standards. (Ref. 3 and 4)

The first report of measurements of the frequency dependence upon microwave power was published in 1978 by Risley and Busca (Ref. 5).The IFT with buffer gas was used, and a power shift around $2{,}5.10^{-11}$/dB was measured. These results were already discussed in term of line inhomogeneity in the absorbtion cell, and the effect of a magnetic gradient, as well as the light intensity, was already evidenced.

In 1980 Risley et al. (Ref. 6) showed that the power shift essentially disappears in a wall coated cell without buffer gas. This result was obtained while intense inhomogeneous line broadening occured caused by an intentionally applied strong magnetic field gradient. However, the free motion of the atoms in the cell averaged out the positional sensitivity.

A number of semi-empirical calculations to rationalize the observed effects have already been perfomed in the quoted papers using the Vanier model (Ref. 7, 8). However, a complete theoretical work which could modelise all observed effects, including C-field compensation, is still lacking. (see also Ref. 9, 10, 11) Nevertheless, we find that such a modelisation could be quite useful for the development of actual devices as well as in the efforts to optimize and control their performances. The present experiments form the basis for such theoretical work presently under progress in our laboratory.

4. EXPERIMENTS AND RESULTS

4.1. Experimental conditions.

Fig. 4.1. gives a block diagram of our experimental setup. The physics package is placed in a thermo chamber so that also external temperature coefficients can be measured.

The modulation frequency is 1kHz, which allows a fast servo-loop of the quartz oscillator. At this frequency, and with an amplitude of modulation of 1 kHz, we had typically a line width of 2,1 kHz, and a discriminator slope of 0,7 nA/Hz. The temperature of the lamp has not been varied in this study. Its value (140°C) was fixed after previous measurements.

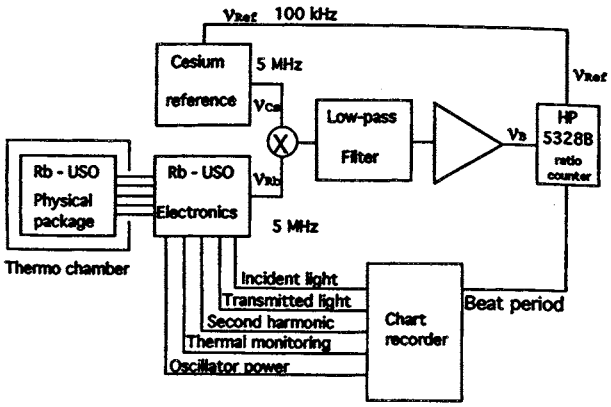


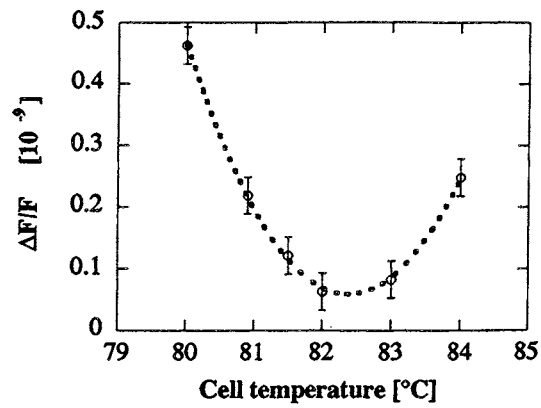
Fig 4.1:The experimental setup.

To perform our physical study, we have the following measurable parameters:

- The nominal temperatures of the lamp and the cell.
- The DC power delivered to the oscillator.
- The current delivered by a photocell placed near the entrance of the cell.
- The current delivered by the photocell for the detection after the cell.
- The microwave power delivered to the resonator.
- The signal at the modulation frequency and the Zeeman lines.

4.2. Temperature coefficient.

We first measured the variations of the frequency with the temperature of the absorption cell. The result depends of course on the values of all the other parameters given below. Figure 4.2 describes already the final result of an iterative adjustment of these parameters. The experimental points are fitted to a parabolic curve .



Incident light: 4,5 V; Transmitted light: 2,15 V;
Second harmonic: 600 μV; Microwave power: -0.15 μW

Fig 4.2:Temperature coefficient.

We want a minimal sensitivity to external temperature variations, so our working point will be around the minimum of this curve. As we will see in the next sections, this point depends on the value of the other parameters (light level, microwave power, C-field). However, by

optimizing the other parameters, the minimum does not vary by more then three degrees.

This high value of the cell temperature (82-83°C) is a result of the integrated filter technique, and allows operation at high environmental temperatures. (70°C)

4.3. Light shift

To have at the same point both minimal light shift and zero temperature coefficient, we varied the reference light at different cell temperatures. In Fig.4.3 which displays the frequency as function of the incident light, we can see that the 3 curves in the range 81.5°C-83°C parametrized by the cell temperature cross at a point (4,5 V) which corresponds to a minimal light shift at 83°C. The other parameters in this experiment were the same as those given with Fig 4.2.

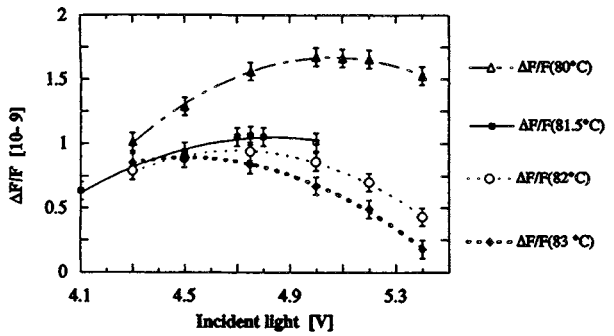


Fig 4.3:Light shift for different cell temperatures.

4.4. Power shift.

For the measurement of the frequency dependence upon microwave power, we have fixed the value of the incident light (4,5 V). Previousely, we could see that this parameter had no significant effect on the power shift. Fig.4.4. shows the frequency as a function of the microwave power, and illustrates the effect of the temperature, which is very important. The C-field setting corresponds to the Field 3 o next section.

From these results we can immediately deduce that the power shift is dominated by the filtering processes, i.e. by the inhomogeneity effects. Another argument in favour of this assertion will be seen in the next section, in which we show measurements of the power shift at different C-field configurations, with effects quite similar.

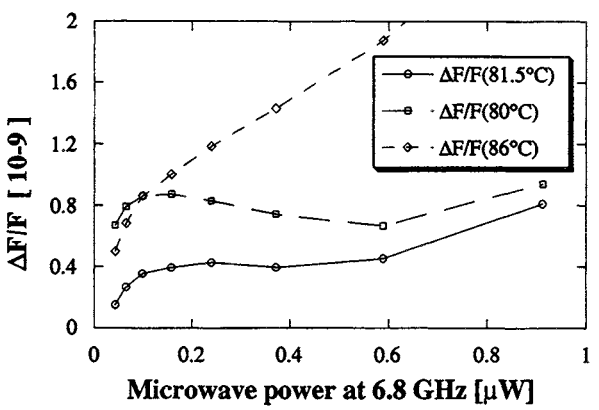


Fig. 4.4: Effect of the cell temperature.

4.5. Effect of the C-field configuration

The spatial variations of resonance frequency in the the cell which are due to variations of the light intensity and/or light spectrum, are difficult to estimate and control. The effect of the C-field H_c on the contrary, is well determined by the formula $\Delta f_c = 574\ H_c^2$, so knowing H_c everywhere in the cell, we also know Δf_c everywhere.
Figures 4.5 and 4.6 show the power shift measured in three different C-field configurations, at a cell temperature of 83°C.

The C-field is produced by two coils of different lengths surrounding the resonator. With different currents in these coils, magnetic field gradients along the cell axis could be produced. Since the coils are surrounded by magnetic shields only approximate calculations of the field geometry were possible. The average C-field as measured by the Zeeman line separation was 70 mG, whereas the calculated gradients converted into frequency variations for the last quarter of the cell were: Field 1: -0.071 Hz/mm; Field 2: +0.075 Hz/mm; Field 3: +0.33 Hz/mm. Field 3 allowed a compensation of the power shift over a wide range in the sub-μW region. The residual power shift around the typical operating level of 0.2 μW (± 0.1 μW) was less than $3x10^{-12}/\mu W$. This low operating power was a result of the improved performance of the magnetron resonator.

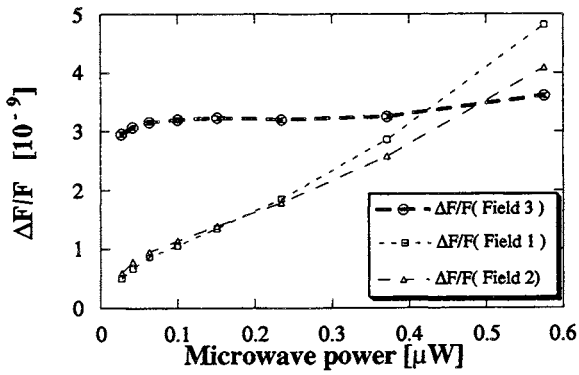


Fig. 4.5: Power shift for different C-field.

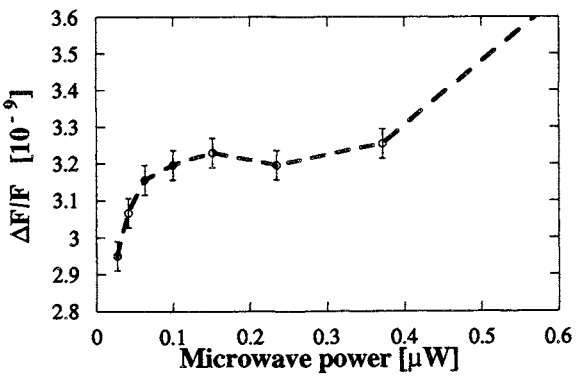


Fig. 4.6: Power shift for the Field 3 (fine).

5. CONCLUSION

The problem of power shift has been studied in a miniaturized Rb standard, in which the integrated filter technique and a novel type of microwave resonator was used.

The power shift varies as function of the cell temperature, but only very little with the light intensity. Light attenuation and filtering appears as the main cause.

Different C-field gradients shows variations in the power shift which was predominantly positive in all cases.

It was possible to find a C-field configuration, and a working point, in which, in addition to zero temperature coefficient, and zero light shift, the power shift is reduced to $< 3x\ 10^{-12}/\mu W$. This performance was possible at very low operating RF power: $< 0.5\ \mu W$ due to an improved microwave structure.

REFERENCES

1. J.C.Camparo, Proc. 18th PTTI, p.565, (1986).
2. P. Willemin, G.Busca, J.T.Liu, and P.Thomann, Proc 5th EFTF, 181 (1991).
3. B.Grover and T.M.Kwon, 15th PTTI, p.677, (1983)
4. E.B.Sarosy,W.A.Johnson, S.K.Karuza, and F.J.Voit, Proc. 23th PTTI, (1991).
5. A.Risley and G.Busca, Proc.32nd Ann. Symp. on Freq. Control, p.506, (1978).
6. A.Risley and S.Jarvis and J.Vanier, J. Appl. Phys., 51(9), 4571 (1980).
7. J.Vanier and C.Audoin, The Quantum Physics of Atomic Frequency Standards, (Adam Hilger, Bristol and Philadelphia, 1989).
8. G.Busca, M.Têtu, and J.Vanier, Can.J.Phys., 51(13), 1379 (1973).
9. P.Tremblay, N.Cyr, M.Têtu, Can.J.Phys. 63, 1563 (1985).
10. J.C.Camparo, R.P.Frueholz, and C.H.Volk, Phys.Rev.A 27(4), 1914 (1983).
11. V.V.Batygyn,M.V.Gornyi and B.G.Matisov, Opt. Spectrosc. USSR 48(1), 67, (1980).

ACKNOWLEDGEMENTS

This work was in part supported by ESTEC under contract number 9099/90/NL/US(SC) and in part by the Swiss Commission for the Encouragment of Research, project number 2159.1. We want to thank Professor G.Busca for stimulating discussions, and P. Scherler and M.Durrenberger for their experimental support.

NOISE MEASUREMENTS IN SINGLE-MODE GaAlAs DIODE LASERS

N. Sagna[1], C. Mandache[2], P. Thomann[1]

[1]Observatoire Cantonal de Neuchâtel;[2]Institute of atomic Physics,Bucarest.

ABSTRACT

Laser diodes may bring signifiant improvements to the performance of Cs and Rb frequency standards,as well as new possibilities for their miniaturization. However, the requirements on the laser are stringent, particularly as far as frequency noise and intensity noise are concerned. These requirements will be reviewed and compared with measurements of several types of single-mode lasers including Fabry-Perot CSP , quantum well and distributed Bragg reflector structures.

Keywords: laser diode,intensity noise,frequency noise,white noise ,1/f noise,spectral linewidth

1. INTRODUCTION

In metrology, as in numerous fields of atomic physics the use of laser diodes has become very important. Indeed, with its spectral purity and coherence, a laser diode provides an efficient tool for the study of interaction between atoms and light. For the past few years, they have been used extensively in frequency standards and atomic clocks to optically pump the atoms and to detect the atoms that have undergone the microwave clock transition [1-2]. Moreover there has been lately significant research and development efforts to use extremely cold atoms generated by laser cooling or trapping in frequency standards [3-4].

One of the most important requirements on the semi-conductor lasers used in frequency standards is a low intensity and frequency noise. Indeed for optically pumped frequency standards, the AM and FM stability of the lasers which provide both the selection and the detection of the atoms has an important influence on the overall frequency standard stability [5]. Besides, experiments of laser trapping have pointed out that the linewidth of the used lasers ,which is directly related to their frequency noise, must be of the order of a few MHz . Therefore, a characterization of intensity noise and linewidth broadening of lasers is of prime interest for their application to frequency standards.

In this paper, we have considered three different types of single mode GaAlAs laser diodes , namely Fabry-Perot CSP, Fabry-Perot Quantum-Well, and Distributed Bragg Reflector structures and we have measured their AM and FM noise spectra. This series of measurement reveals interesting differences between the three types of laser diodes. Besides, on the basis of an analytical development, we show the importance of the 1/f type FM noise in the laser linewidth.

2. EXPERIMENTAL METHODS AND SET-UP

2.1 General disposition

In the different noise measurements, the same current supply has been used for the three different types of laser diode. This current supply developped in our laboratory presents a low noise of 0.6 nA/√Hz between 100 Hz and 100 kHz. Such a low level noise is required when one wants to point out the specific noise features of a laser diode. Besides, knowing that interference and feedback effects as well as mode hopping may yield a considerably increased noise, an isolator and anti-reflection coated optical elements were used to eliminate the feedback, and the temperature and the current of each laser diode were adjusted so as to be far enough from a mode hop.

The measurements of the noise spectra were done in the range 0-100 kHz. Such a range was adopted for two reasons : it corresponds to the noise frequency domain that is most relevant for frequency standards and, as we will see later, that gives the most significant contribution to the laser linewidth. The transfer function of the chain : current supply-laser diode-photodetector was measured, and this calibration was taken into account in the results. The noise measurements were performed using a FFT spectrum analyser whose acquisition time, for the range 0-100 kHz, is 400 ms and that gives the single-sided power spectral density of the output signal.

2.2 Intensity noise measurements

The measurements of the intensity noise of the diode lasers were done considering the block diagram shown in fig.1.

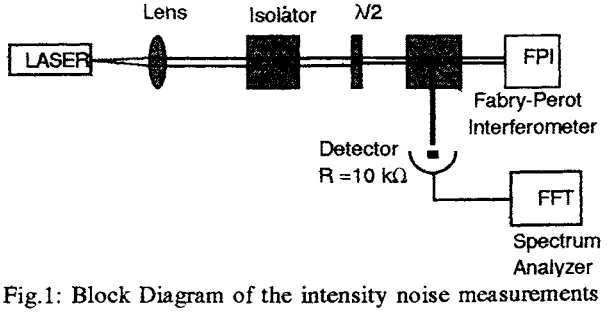


Fig.1: Block Diagram of the intensity noise measurements

Let us point out that we present in the results the relative intensity noise (RIN) related to the photocurrent I_{ph} produced by the detector , which includes other contibutions

than the RIN of the laser such as photocurrent shot noise,photodetector thermal noise ,and current supply noise.The light intensity was high enough (I_{ph} > 0.2 mA) for the contribution of the laser diode to be dominant in all measurements.

2.3 Frequency noise measurement

The measurements of the FM noise were carried out using a cesium absorption cell at room temperature and of diameter 25 mm as a discriminator (fig.2) : the laser was tuned at the point of maximum slope.

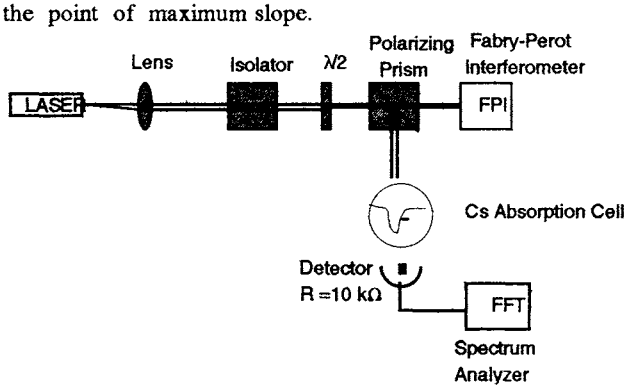


Fig 2 : Block diagram of the frequency noise measurements

The DC discrimination slope was of the order of about 10^{-13} A/Hz.In all measurements the laser power was low enough (P<15 μW/cm^2 for 3 mm laser beam) so that hyperfine optical pumping, which induces saturation effects and a significant time constant in the atoms response could be neglected. At this power level it was verified that the AM noise was negligible compared to the FM noise.

3. EXPERIMENTAL RESULTS

3.1 Intensity noise measurements

The relative intensity noise was measured at the frequency 1 kHz, as a function of the ratio I/I_{th} between the LD operating and threshold current, for the three types of lasers. The results obtained are shown in fig.3 .

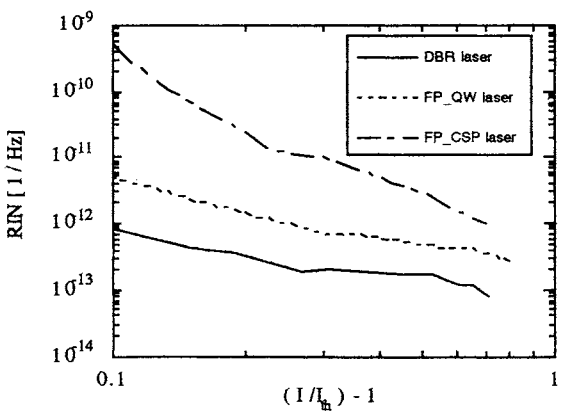


Fig 3 : Relative Intensity Noise as a function of the ratio between the LD and the threshold current for the three lasers at the frequency 1 kHz.

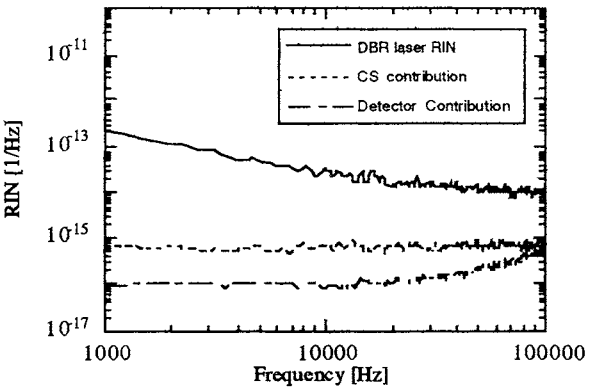


Fig 4 : Contribution of the Current Supply (CS) and the detector in the measured Relative Intensity Noise of the DBR laser

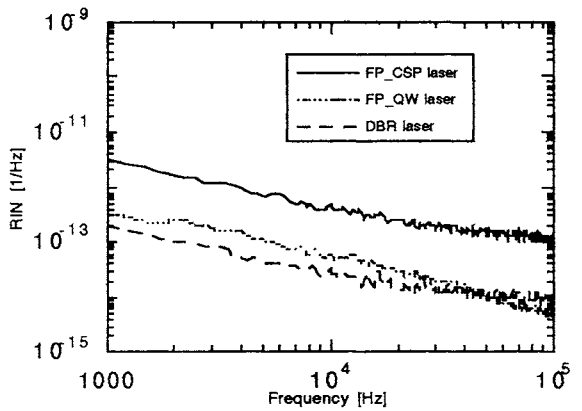


Fig 5 : Relative Intensity Noise for the ratio I / I_{th} =1.6 and for the three types of laser

The contribution of the current supply (CS) noise as well as the thermal and shot noise of the photodetector, to the total relative noise of the DBR laser were measured and are presented in fig.4. In addition, fig.5 gives the relative intensity noise spectra of the three types of laser operating at the same level above threshold (I/I_{th} =1.6).

3.2 Frequency noise measurements

In order to estimate the contribution of the current supply to the frequency noise one has to take into account the laser diode response to different frequencies of modulation of the current supply, when the laser is tuned to the point of maximum slope in the absorption curve. This contribution, and the one of the detector are compared in fig.6, to the total frequency noise of the DBR laser.

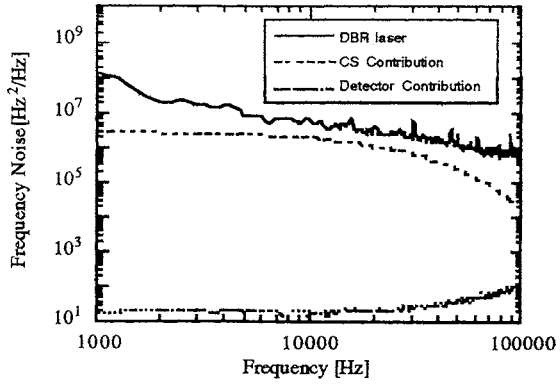


Fig 6 : Contributions of the Current Supply (CS) and the photodetector to the Power Spectral Density (PSD) of the frequency noise of the DBR laser

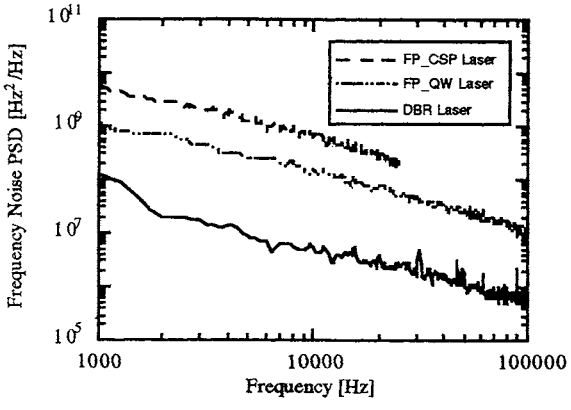


Fig 7 : Power Spectral Density (PSD) of the frequency noise for the three types of laser

Table 1: Numerical fittings of the power spectral densities of the FM noise as a function of the Fourier frequencies for the three types of laser

	DBR	FP_QW	FP_CSP
Fitting of FM PSD	$y=3.5\,10^{10}x^{-0.94}$	$y=1.9\,10^{12}x^{-1.04}$	$y=9.2\,10^{12}x^{-1.04}$
Regression Coef.	0.91	0.981	0.981

Besides, we give in fig.7, a comparaison between the frequency noise of the three types of laser, at the same level of power (P = 0.5 mW), and present in table 1, the numerical fitting of each one of the FM noise spectra.(y = Power spectral density of the frequency noise; x = Fourier frequency)

4. DISCUSSION AND CONCLUSIONS

4.1 Analysis of the intensity noise results

As shown in fig.4, the contribution of the measuring system to the laser intensity noise is negligible ; this allows to pull out from the results a few features of the laser diodes in test. Thus, for the DBR and the QW-type lasers the theoretical developments [6] that suggest that the RIN is inversely proportional to the emitted power were verified. However, the results obtained are hardly comparable with others in the litterature because most measurements of RIN for GaAlAs laser diode were done in the GHz region [7].Comparing the results obtained for the three types of laser, one can conclude that the DBR laser exhibits a better intensity stability at frequencies below 30 kHz than the two Fabry-Perot type lasers. This may eventually be explained by the fact that for the Fabry-Perot -type lasers, the gain discrimination of the dominant mode against the other modes is small; therefore a change of the threshold gain induced by the residual unwanted optical feedback may yield to increase the mode hopping and its corresponding intensity noise [6]. Let us also notice that as far as the intensity fluctuations are concerned, the QW and DBR-type lasers exhibit noise of the same order ,unlike what happens, as we will see below, for frequency noise.

4.2 Frequency noise and linewidth broadening

4.2.1 Theoretical derivation of a laser linewidth

The purpose of this section is to show in a rigorous way how the linewidth of a laser can be deduced from its experimentally measured FM noise spectrum. Let us consider a single mode laser whose intensity fluctuations are negligible compared to the phase fluctuations and whose electric field can be written:

$$E(t) = E_0 \cos(2\pi \nu_0 t + \phi(t)) \qquad (1)$$

By definition the laser spectral linewidth is given by the full width at half maximum of the power spectral density $S_E(\nu)$ of the laser electric field.In order to evaluate $S_E(\nu)$ one has to characterize the statistics of the laser signal phase change during a time τ ,defined as:

$$\Delta\phi_\tau(t) = \phi(t+\frac{\tau}{2}) - \phi(t-\frac{\tau}{2}) = \int_{t-\frac{\tau}{2}}^{t+\frac{\tau}{2}} \dot{\phi}(t)\, dt \qquad (2)$$

In the following one will assume the stochastic processes that yield to change of $\Delta\phi_\tau$ and $\dot{\phi}$, to be wide sense stationary [8], so that one can deduce from (2) that the single-sided power spectral density $S_{\Delta\phi_\tau}$ and $S_{\dot{\phi}}$ verify the relation :

$$S_{\Delta\phi_\tau}(\nu) = S_{\dot{\phi}}(\nu) \left(\frac{\sin(\pi\nu\tau)}{\pi\nu\tau}\right)^2 \tau^2 \qquad (3)$$

Using a change of units related to the relation : $\delta\nu = \dot{\varphi}/2\pi$ as well as the definition of the mean square value, one can relate the measured power spectral density $S_{\delta\nu}(\nu)$ of the frequency noise to the phase noise statistics by writing

$$< \Delta\phi_\tau^2 > = 4\pi\tau^2 \int_0^\infty S_{\delta\nu}(\nu) \left(\frac{\sin(\pi\nu\tau)}{\pi\nu\tau}\right)^2 d\nu \qquad (4)$$

A further characterization of the phase change statistics requires an assumption on the probability law followed by the phase change $\Delta\phi_\tau$. Considering the independence of the spontaneous emission processes and the central limit theorem [8], the probability distribution of the phase change will be assumed to be gaussian :

$$p(\Delta\phi) = \frac{1}{\sqrt{2\pi < \Delta\phi^2 >}} \exp\left(-\frac{1}{2}\frac{(\Delta\phi)^2}{< \Delta\phi^2 >}\right) \qquad (5)$$

Moreover, assuming the processes that yield to phase change to be ergodic, the autocorrelation function can be written:

$$G(\tau) = < E(t)\, E(t-\tau) > \qquad (6)$$

One can show that the calculation of (6) gives

$$G(\tau) = \frac{1}{2} E_0^2 \cos(2\pi\nu_0\tau) \exp\left(-\frac{1}{2} < \Delta\phi_\tau^2 >\right) \qquad (7)$$

By definition, the power spectral density of the laser electric field is given by

$$S_E(\nu) = \frac{1}{2} E_0^2 \int_0^{+\infty} [\cos(2\pi(\nu - \nu_0)\tau) + \cos(2\pi(\nu + \nu_0)] \times \exp\left(-\frac{1}{2} < \Delta\phi_\tau^2 >\right) d\tau \qquad (8)$$

where $< \Delta\phi_\tau^2 >$ is given by expression (4). Starting from the measured power spectral density $S_{\delta\nu}(\nu)$ (fig. 7), we have evaluated expressions (4) and (8) and have obtained, the spectral linewidth of each laser defined as the full width at half maximum of $S_E(\nu)$.

4.2.2 Results of the linewidth calculations

As shown in table 1, in the frequency range below 100 kHz, the 1/f noise is predominant for the three types of lasers . Due to the frequency range limitation of the spectrum analyser used we have made the approximation that the white noise level is given by the level of the power spectral density at 100 kHz; such an upperbound approximation will emphasize even better what we want to show.

Table 2 : Numerical estimates of the different contributions to the lasers linewidth taking the frequency cut-off of the order of a few Hertz and as an approximation of the FM noise PSD the expression : $S_{\delta\nu}(\nu) = C_0 + C_1/\nu$ where C_0 and C_1 are fitted to experimental data.

	DBR	FP_QW	FP_CSP
C_0 [Hz]	$6.05\ 10^5$	$1.05\ 10^7$	$1.31\ 10^7$
C_1 [Hz^2]	$3.644\ 10^{10}$	$1.94\ 10^{12}$	$9.2\ 10^{12}$
White Noise Linewidth [MHz]	1.8	30.5	39.5
1/f Noise Linewidth [MHz]	2.8	45.6	95.3
Total Linewidth [MHz]	4.2	64.6	109.5

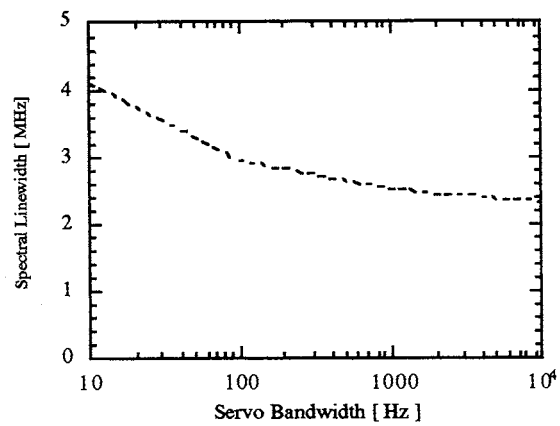


Fig.9 : Numerical detemination of the spectral linewidth of the DBR laser as a fonction of the bandwidth of a ideal electronic servo system.

For the three types of laser, the respective linewidth broadening contribution of the white and 1/f noise as well as the total spectral linewidth were numerically calculated; the results are presented in table 2. Moreover, fig.9 shows the effects of electronic servo systems of different bandwidths on the linewidth of the DBR laser.

4.2.3 Interpretation

The calculated linewidths of table 2 ,as well as the FM noise spectra of fig.7 ,point out an important difference between the three types of laser,as far as the frequency stability is concerned.Unlike the two Fabry-Perot type lasers, the DBR-types laser presents a linewidth of a few megahertz , comparable to the one of the lasers used for telecommunication coherent ligthwave applications [9-10].On the other hand,table 2 reveals a common feature to the three types of lasers: it shows ,in spite of the use of an upperbound approximation for the white noise level ,the relative importance of the 1/f noise contribution to the total linewidth.Unlike the white noise which is inversely

proportional to the laser output power,the 1/f-type noise is almost power independent [9] ; its contribution cannot be reduced by increasing the laser emitted power, therefore it limits the minimum spectral linewidth. This emphasizes as shown in fig.9 in the case of the DBR laser the effectiveness of an electronic servo system to reduce significantly such a contribution.

5. ACKNOWLEDGEMENT

This work was supported by the Swiss National Fund for Scientific Research. The authors also wish to thank Mr M. Morimura and Mr H. Hosomatsu of OMTEC Corporation for the loan of DBR laser diodes.

6. REFERENCES

1. M. Hashimoto & M. Ohtsu 1989,J. Opt. Soc.Am , 6 ,1777-1789

2. R.E.Drullinger & al 1991,Proc 5th EFTF,Besancon ,412-414

3. C.Monroe & al 1991 , Opt. Lett., 16 ,50-52

4. A. Clairon & al 1991,Proc 5th EFTF,Besancon ,228-2236

5. P. Thomann & al 1988,Proc 3rd EFTF,Besancon ,269-273

6. K. Petermann ,1991, Laser Diode Modulation and Noise, Kluwer Academic Publishers,Tokyo

7. K. Petermann & G. Arnold , 1982 ,J. Quant. Electron., 18, 543-555

8. A. Papoulis ,1984, Probability,RandomVariables,and stochastic processes , McGraw-Hill , New-York

9. K. Kikuchi , 1989 , J. Quant. Electron., 25 , 684 - 688

10. M. Okai & al 1991,IEEE Phot.Techn.Lett. , 3.,427 - 429

FUNDAMENTAL CHARACTERISTICS OF THE LASER-DIODE-PUMPED RUBIDIUM MASER WITH AN EXTERNAL FEEDBACK LOOP

OURA Nobunori, KURAMOCHI Naimu, DOMON Wataru,
KOMIYAMA Tamotsu, and SHIRATO Ryota

TOKYO INSTITUTE OF TECHNOLOGY
4259, NAGATSUTA, MIDORIKU, YOKOHAMA 227, JAPAN

ABSTRACT

The ^{87}Rb maser studied is a small one with a TE_{011} mode cavity. The feedback loop technique similar to that of a compact H maser is utilized. The continuous oscillation is confirmed from the fact that the oscillation began only when the laser diode light irradiated an ^{87}Rb maser cell in the cavity. In addition to the ^{87}Rb maser, the continuous oscillation of maser using a natural Rb cell is also obtained.

Keywords: Rb maser, TE_{011} mode cavity, optical pumping, laser diode, external feedback loop,

1. INTRODUCTION

The rubidium (Rb) maser is a kind of stable frequency sources based on the transition frequency (f=6834 MHz) between two hyperfine levels of the ground state of the ^{87}Rb. Though the conventional Rb masers show superior short-term stabilities[1,2], they have not yet appeared on the market. The problems might lie in the use of a Rb lamp and a large TE_{021} mode cavity.

For more than a decade, some attempts have been made to use the laser diode as a pumping light source for the atomic frequency standards, mainly for the cesium standard[3,4], and the passive Rb standard[5,6]. Recently, the continuous oscillation of a laser-diode-pumped ^{87}Rb maser has been reported[7]. That maser uses a conventional TE_{021} mode cavity and a frequency-modulated laser.

In this paper, an external feedback loop technique, which is used in the field of a compact H maser[8,9,10], is introduced to enhance the Q-value of a TE_{011} mode cavity. The volume of the cavity is one-eighth of the TE_{021} mode cavity.

Some experimental results on the fundamental characteristics of the Rb maser employing an ^{87}Rb cell and on the continuous oscillation of the Rb maser employing a natural Rb cell will be described.

2. CAVITY AND THE EXPERIMENTAL SETUP

2.1 TE_{011} mode cavity

The maser employs a cylindrical TE_{011} mode cavity containing a cylindrical fused quartz bulb as a maser cell filled with Rb vapor and nitrogen as a buffer gas.

The first cavity (#1) is designed so as to vary easily the resonance length for detecting the maser signal. The cavity and its surroundings are shown in Fig.1. The resonance frequency of the cavity can be changed by moving the upper end plate with three screws. Dimensions of the cavity and the maser cell are listed in Table 1.

The cavity cylinder is made of fused quartz and its inside surface is coated with conductive silver paint (DOTITE D-1240, FUJIKURA KASEI CO.). Circular copper plates with flange are mounted in the upper and lower ends of the cavity. These end plates compensate for the resonance frequency changes due to thermal expansion of the cylinder by taking advantage of the difference of thermal expansions between both materials. The cavity is tuned to 6.83468 GHz of the clock transition of the ^{87}Rb. A concave lens of 10 mm in diameter for introducing the pumping light is built into the center of the plate.

2.2 Experimental setup

A block diagram of the experimental system is shown in Fig.2. The pumping light is guided to the cavity from a laser diode (HL7806G), which emits light of a wavelength of 780.027 nm, corresponding to the wavelength of the optical transition between the 5P state and the F=1 ground state of the ^{87}Rb. The wavelength of the laser is stabilized to the spectral line of the reference ^{87}Rb cell made under the same specifications as the maser cell. The maximum optical output power delivered by the laser diode is about 4 mW.

The external feedback loop consists of a phase shifter, variable attenuator, directional coupler and rf amplifier.

2.3 Q-value of the cavity with the feedback loop

The Q-value of the cavity with the external feedback loop is measured by transmission method.

Proceedings of the 6th European Frequency and Time Forum, held at ESTEC, Noordwijk, NL, 17–19 March 1992 (ESA SP-340, June 1992).

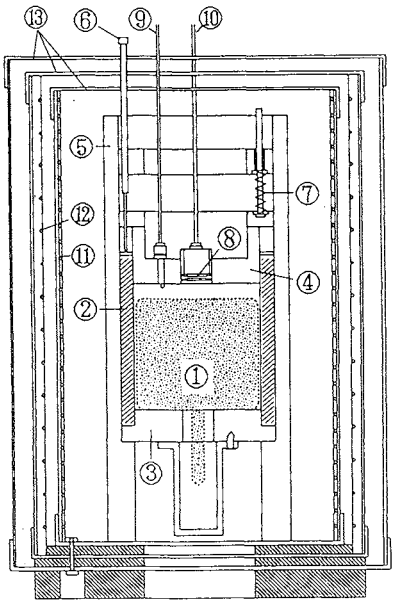


Fig.1: Cross section of the cavity #1.

1. Maser cell.
2. Fused quartz cylinder.
3. Copper lower end plate.
4. Copper upper end plate.
5. Aluminum cylinder for supporting the cavity.
6. Screws for adjusting the resonance frequency.
7. Springs.
8. Concave lens.
9. Coupling loops.
10. Stub for fine tuning.
11. Coil for C-field.
12. Heater coil.
13. Magnetic shields.

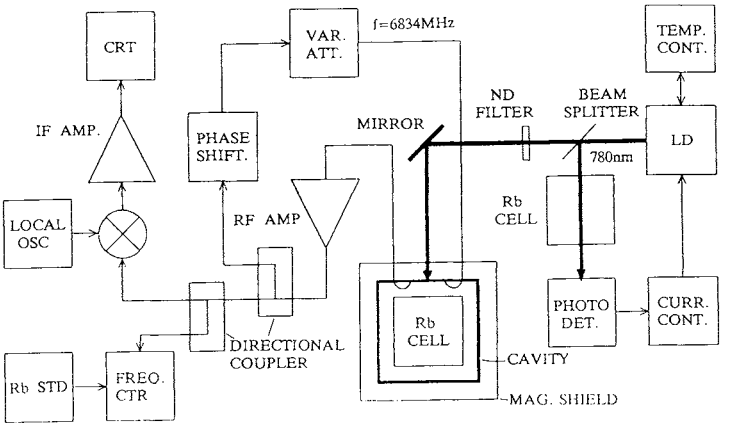


Fig.2: Block diagram of the experimental setup of the Rb maser.

Table 1: Dimensions and specifications of the cavity and the maser cell.

Cavity		
Resonance frequency	:	6834 MHz
Mode	:	TE_{011} mode
Inner diameter	:	57.0 mm
Resonance length	:	59.2 mm
Maser cell		
Outer diameter	:	56.0 mm
Length	:	54.0 mm
Buffer gas	:	11 Torr of nitrogen

The attenuation and phase shift in the feedback circuit are adjusted to amounts to cause sustaining oscillation in the circuit without pumping light. Then, the attenuation of the attenuator is increased a bit by an amount to stop the self-oscillation. Under this condition, the signal from a signal generator is fed to the circuit and the frequency is swept, then the output voltage vs. frequency plots of the circuit draw a resonance curve, from which the Q-value is calculated.

The measured resonance curves are shown in Fig.3 with the parameter values indicating the output signal levels of the signal generator. Two resonance curves in the uppermost of the figure are asymmetric. These curves are eliminated from the calculation of the Q-value. Resonance curves similar to those above are also observed with high Q crystal resonators, when they were excited to a nonlinear region[11].

The Q-values for weak excitation (under -10 dB) in Fig.3 were about 20 x 10^4 . On the contrary, the Q-value of the cavity itself is about 1.8 x 10^4 as shown in Fig.4. The use of the external feedback loop enhanced the Q-value about 10 times.

3. MASER OSCILLATION BY USING A VARIABLE RESONANCE LENGTH CAVITY

3.1 Observation of the oscillation of the Rb maser

The temperature of the maser cell is controlled at 60°C. The experimental procedure to observe the oscillation of the Rb maser is as follows.

(1) Adjust the phase and the attenuation in the feedback loop without pumping light to obtain self-oscillation of the cavity oscillator. (2) By changing the resonance length of the cavity with three screws, the oscillation frequency of the cavity oscillator can be tuned in to the clock transition of the ^{87}Rb. (3) The gain of the feedback loop is decreased to a magnitude insufficient slightly for sustaining the self-oscillation. (4) The pumping light irradiates the maser cell in the cavity. Then, the signal appears on a CRT and only when the pumping light irradiated the cell, it appears. (5) The oscillation sustained, even if the phase of the shifter was shifted by about ±5 degrees. This value is one-ninth smaller than that of the system only as a cavity oscillator. (6) The oscillating frequency was about 6.834688 GHz. This frequency coincides with the clock transition frequency of the ^{87}Rb shifted by the buffer gas of 11 Torr of nitrogen.

From these experiments, the signal observed was determined to be the continuous oscillation of the Rb maser.

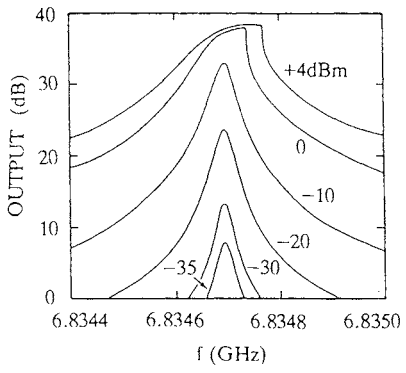


Fig.3: Obtained resonance curves of the cavity with the external feedback loop. Parameters indicate the signal levels of the signal generator.

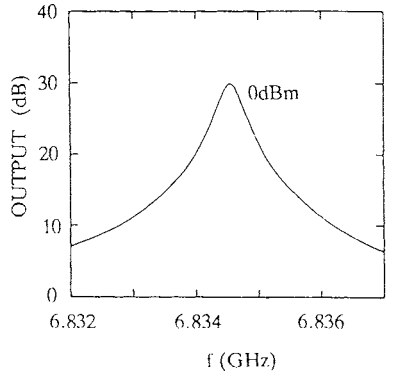


Fig.4: Resonance curve of the cavity without the external feedback loop.

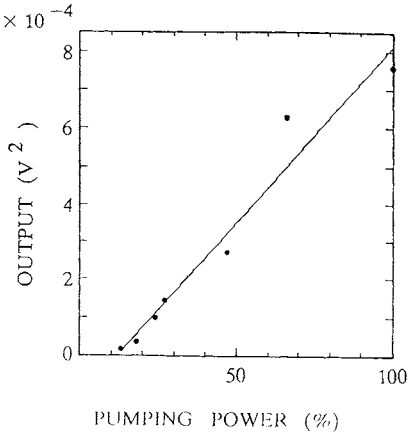


Fig.5. Oscillation output with respect to the pumping laser light power. T : Transmittance of the ND filter. T =100 % corresponds to the LD power of 3.9mW.

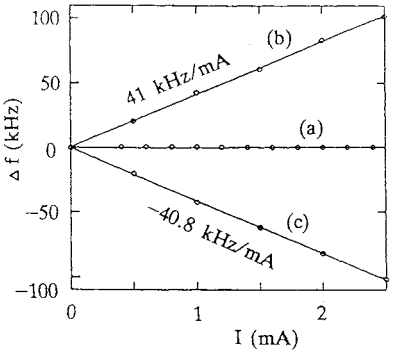


Fig.6: Magnetic field intensity vs. oscillating frequency. I : coil current.

3.2 Pumping light power vs. oscillation output

In Fig.2, the optical pumping power is adjustable by inserting a neutral density (ND) filter in the path of the laser beam. Figure 5 shows oscillation output with respect to the pumping laser power. The output on the abscissa is the squared value of the amplitude of the signal appeared on the CRT. This value must be proportional to the maser output from the directional coupler. In Fig.5, the output power is proportional to the optical pumping light intensity reduced to the transmittance(T) of the light. The maximum optical pumping power in this experiment was 3.9 mW, which was measured in the right front of the mirror in Fig.2.

3.3 Magnetic field intensity vs. maser frequency

The intensity of the static magnetic field is varied with the current in the C-field coil in Fig.1. The magnetic field intensity in the center of the cavity, where the maser cell is to be set, is 51 mOe, when the coil current (I) is 2 mA. With changing the resonance length of the cavity, while the maser was oscillating, signals of different frequencies were observed at different resonance lengths of the cavity. Figure 6 shows these characteristics as a function of the C-field current. The line (a) corresponds to the frequency of the clock transition. The lines (b) and (c) depend on the magnetic field intensity, and these frequencies are the transitions other than the clock transition.

3.4 Cavity pulling and the light shift

The oscillation frequency shift due to the optical piumping light is called the "light shift." The output frequency is altered owing to the light shift and the cavity pulling from the transition frequency. The effects of the light shift and the cavity pulling on the oscillating frequency were measured by changing the resonance length and the light intensity.

When the maser is oscillating at the transition frequency f_0 without an ND filter in the light path, the oscillating frequency shifts to a frequency f, if the resonance length of the cavity is changed.

Under these conditions, if the transmittance T of the light is changed by inserting an ND filter in the path, the frequency shifts to other frequency f' due to the light shift. The results are shown in Fig.7. From this figure, it is understandable that;

(1) The light shift depends on the maser oscillating frequency determined from the cavity pulling.
(2) The light shift is linear to the light intensity.
(3) There is a maser oscillating frequency independent of the light shift. The frequency was f_0 = 6.8346887 GHz in this experiment.

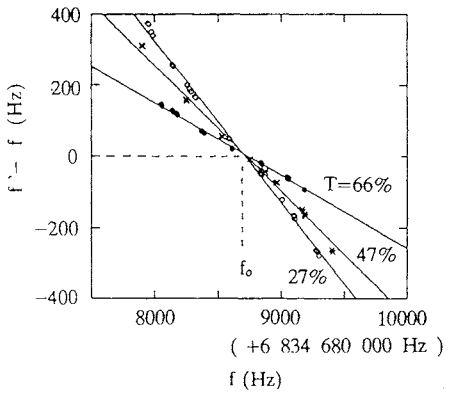


Fig.7: Cavity pulling and the light shift.
T : Transmittance of the ND filter.
f : Oscillating frequency.
f'- f : light shift.

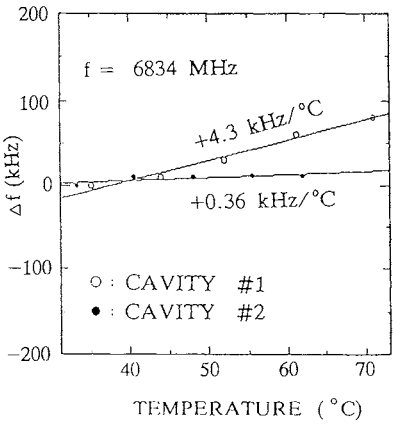


Fig.8: Resonance frequency temperature characteristics of the cavities.

4. IMPROVEMENT OF THE MECHANICAL STABILITY OF THE CAVITY

Though the resonance frequency of the cavity #1 is adjustable with screws and the cavity is appropriate for the light shift and the cavity pulling experiments, the cavity is unstable mechanically, for example uneven expansions of the screws due to temperature changes.

A new cavity (#2) without adjustable screws was made. The lengths of the upper and lower plates protruding into the cylinder are designed for aiming at more precise compensation for the resonance frequency change due to thermal expansion superior to that for the #1 cavity. Several couples of the upper and lower plate of different length by 0.1 mm steps were prepared for every maser cell. The rough frequency tuning of the cavity to the resonance frequency was performed by grinding off the cylinder length and selection of the plates commensurate with the cylinder length. The fine tuning was done by a stub. Another change on designing the #2 cavity is the use of two concave lenses to diverge optical pumping light.

The frequency temperature characteristics of the cavities are shown in Fig.8. The frequency temperature immunity of the #2 cavity was improved by an order of magnitude than that of the #1 cavity. In addition, the frequency fluctuation of the maser employing the #2 cavity is improved by an order of magnitude or more than that of the #1 maser and has an order of 10^{-10}. The frequency stability as a function of averaging time has not yet been evaluated because of the measurement system.

5. OSCILLATION OF THE RB MASER WITH A NATURAL RB CELL

The employment of the external feedback loop in the maser to enhance the Q-value made it possible to use a small size cavity and a natural Rb maser cell. The continuous oscillation of a maser with a natural Rb cell was observed. The frequency fluctuation was approximately the same as that of the maser with an isotope cell, although the output is one-third of that of the maser with an ^{87}Rb cell.

6. CONCLUSION

The continuous oscillation has been obtained for a maser pumped by a laser diode light with a TE_{011} mode cavity and with an external feedback loop. The Q-value of the cavity with the external feedback loop was measured to be about 20×10^4. The continuous oscillation of a maser using a natural Rb cell was also observed. The frequency stability by two-sample variance will be evaluated in the near future.

ACKNOWLEDGEMENT

The authors express their gratitude to Dr.Nakagiri and Mr. Ohta of Communication Research Laboratory for their useful discussions and help to this study. They also thank to Dr. Nakamura and Mr. Chinone of Hitachi Central Research Laboratory for preparing laser diodes.

REFERENCES

[1] M.Têtu, G.Busca, and J.Vanier: IEEE Trans. Instrum. Meas., Vol.IM-22, p.250, 1973.
[2] G.Busca and L.Pan: Proc. 2nd EFTF, p.409, 1988.
[3] A.Brillet: Metrologia, Vol.17, p.147, 1981.
[4] G.Avila, E.De Clercq, M.De Labachellerie, and P.Cerez: IEEE Trans. Instrum. Meas., Vol.IM-34, p.139, 1985.
[5] L.Lewis and M.Feldman: Proc. 35th AFCS, p.612, 1981.
[6] M.Hashimoto, M.Ohtsu, and H.Furuta: Proc. 41st AFCS, p.25, 1987.
[7] A.Michaud, P.Tremblay, and M.Têtu: CPEM '90 DIGEST, P.155, 1990, and IEEE Trans.Instrum. Meas.,IM-40, p.170, 1991.
[8] H.T.M.Wang: Proc.34th AFCS, p.364, 1980.
[9] P.Lesage and C.Audion: IEEE Trans. Instrum. Meas., Vol.IM-30, p.182, 1981.
[10] C.F.Lin, Y.Ohta, S.Urabe, and H.Saito: J. CRL., Vol.36, No.149, p.171, 1989.
[11] J.-J.Gagnepain, J.-C.Ponçot, and C.Pegeot: Proc. 31st AFCS, p.17, 1977.

PROGRESS TOWARDS LABORATORY FREQUENCY STANDARDS FOR THE OPTICAL COMMUNICATION BANDS

D J E Knight[1], K I Pharaoh[1], S L T Drysdale[1], J Meldau[2], G Duxbury[2] and C J Hodges[3]

[1]National Physical Laboratory, Teddington, Middlesex, TW11 0LW, UK
[2]University of Strathclyde, 107 Rottenrow, Glasgow, G4 0NG, UK
[3]Cavendish Laboratory, University of Cambridge, Cambridge, CB3 0HE, UK

ABSTRACT

This paper reviews some frequency reference points in the 1.5 μm optical communication band, and experimental results are presented on frequency references for the 1.52 μm He-Ne laser. A $^{15}NH_3$ absorption within the tuning range of the laser is found to be weak and not saturable. A diode laser was used to investigate the 1.52 μm Ne transition observed via the optogalvanic effect in a hollow-cathode discharge. Using a similar discharge tube intracavity in a He-Ne laser, optogalvanic saturated-absorption dips of about 6% were found in ^{20}Ne discharges at about 1 torr (133 Pa) pressure. The full width at half maximum was about 100 MHz.

Keywords: laser frequency standards, 1.5 μm optical communication band, He-Ne laser

1. INTRODUCTION

A recent programme in the United Kingdom is aimed primarily at providing absolute frequency standards for the optical communication bands [1,2]. The work has been a combination of searches for molecular or atomic resonances at frequencies of interest to users [3], e.g. near 1.5 and 1.3 μm wavelength [4], and attention towards improved frequency stabilisation and absolute frequency measurement of the HeNe laser near 1.523 μm, 200 THz. Improved absolute frequency measurement, following an exploration experiment [5], now waits upon new work with MIM diodes and more laser power. A grating tuned diode laser has been used to investigate optogalvanic signals from Ne at 1.523 μm.

This paper further reports experimental results towards finding frequency reference points in the 1.5 μm optical communication band. A long-known point is the centre of the ^{20}Ne gain peak at 1.523 μm generated in a He-Ne laser, and which was measured by Jennings et al. to ± 25 MHz [6]. For a frequency standard it is better to have a feature such as a Lamb dip in a laser or a saturated absorption feature in a gas cell which will eliminate the first-order Doppler effect. We report here experiments towards improvement of the ^{20}Ne reference point by observation of the saturated-absorption dip using the optogalvanic effect, and investigation of the usefulness of a $^{15}NH_3$ absorption in coincidence with the He-Ne laser line [7,8].

2. FREQUENCY STANDARDS AND FREQUENCY MEASUREMENT

There are particular difficulties for frequency standards in this spectral region. The well-known He-Ne laser at 1.52 μm does not show a strong Lamb dip, and 2-mode Zeeman stabilisation is more commonly used [9,10]. This however is not so traceable to the Ne line centre. Absorptions in general tend to be weak, and this may be connected with the use of the band for long-distance transmission of signals in silica fibres.

Table 1: Some frequency standards for the 1.5 μm optical fibre band

SYSTEM		MEASUREMENT			REF.
molecular	laser	λ/μm	f/MHz	δf/MHz	frequ.ref.used
Rbf/2 LD	diode	1.5605	192 113 990.95	0.07	0.633 I_2 via 1-m F-P [18]
HCN	diode	1.5236	196 764 158	26	1.523 HeNe [7]
C_2H_2	diode	1.5236	196 764 891	26	1.523 HeNe [7]
$^{13}C_2H_2$	diode	1.5235	196 776 026	27	1.523 HeNe [7]
^{20}Ne	HeNe	1.5235	196 780 372	25	Xe(2μm), CO_2 [6]
NH_3	diode	1.5234	196 786 923	28	1.523 HeNe [7]
$^{15}NH_3$	diode	1.5234	196 787 821	26	1.523 HeNe [4]
^{20}Ne LD	HeNe	1.1526	260 103 249.26	0.08	CH_4 via CO_2 [16]
$^{127}I_2$(i)	HeNe	0.6330	473 612 214.80	0.16	CH_4, CO_2; recommended [19]

Notes:-
Rb f/2: the absorption is at 0.780 μm; a 2nd-harmonic signal from a 1.56 μm laser is needed to interrogate it.
LD: Lamb dip, i.e. saturated-absorption dip;
Uncertainties are 1-sigma values.

Table 2: Summary of accurately measured spectral lines in the 1.5 μm band.

molecule or atom	lambda range (μm)	no.of lines	uncertainty ± parts/10^7	uncertainty ± MHz	notes	absorption /(cm.torr)	ref.
HCN	1.560 - 1.567	36	0.8	15	measured		[22]
HCN	1.514 - 1.570	259	0.8-2.4	15-45	calculated		[22]
$H^{13}CN$	1.560 - 1.568	34	0.8	15	measured		[22]
$H^{13}CN$	1.523 - 1.578	257	0.8-3.3	15-63	calculated		[22]
Ne (o.g.)	1.523491	1	6 (LD)	(0.2)			[12]
$^{13}C^{12}CH_2$}	1.522 - 1.550	(650)	5	100	measured + fitted		[24]
C_2H_2 }					const.s only given		
$^{13}ClCH_2$	1.517 - 1.544	50/63	6	120		>1.E-04	[4]
Ar (o.g.)	1.505 - 1.533	3	~10 (LD)	(0.2)			[12]
$^{15}NH_3$	1.490 - 1.530	36	6	120		>7.E-04	[4]
NH_3	1.494 - 1.520	20/54	5	100		>5.E-04	[25]
Kr (o.g.)	1.474 - 1.538	4	~10 (LD)	(0.2)			[12]

Notes:
o.g.: optogalvanic signal
LD : Lamb dip observed; bracketed uncertainty is an estimated reproducibility

Atomic transitions in general are from excited states which are weakly populated at room temperature. Thus techniques comparable to the pumping of lasers, by electrc discharges [11,12] or by optical pumping, as demonstrated by S Gilbert in Rb [13], are necessary to interrogate the transitions. Transitions in molecules in this region are typically overtone or combination-band transitions, and occur in molecules such as acetylene (HCCH) and ammonia (NH_3) and their various isotopic species.

Molecular transitions can be quite strong, e.g. in NH_3, but there are no reports of saturated absorption having been observed in the 1.5 - 1.3 μm region. The erbium-doped fibre laser is a useful, fairly powerful, tunable source for attempting this [14]. Saturation effects have been observed in atomic transitions [12], but require use of electrical discharges for excitation and of the opto-galvanic effect for observation of transitions.

A survey of requirements for frequency standards in the optical communication bands has recently been carried out in the UK [3]. The enquiry was conducted in the UK and Europe for the most part, and opinions were also taken from Japan and the USA. The findings summarised the current need as "two standards per window in the 1300 and 1500 nm windows, that the 1500 nm standards should precede those in the 1300nm window, and that the 1500 nm window standards should be available in the next 3 to 5 years. The accuracy of the standards should be developed to the level of about 1 part in 10^9 (± 0.2 MHz at 200 THz), recognising that this accuracy might have to be improved in the longer term".

For purposes of absolute frequency measurement against the caesium standard the position of 1.5 μm standards in the spectrum is as follows. The spectral region, 200 THz, lies at about the 6th harmonic of CO_2 laser frequency standards in the 33 THz region, at present just beyond the reported harmonic-mixing limit of the metal-insulator-metal (MIM) point-contact diodes [5,15]. There is reason to hope that mixing might be extended from the present limit of 5th harmonic at about 2.3 μm, 130 THz [16], and that MIM diodes might be used for e.g. difference-frequency measurements in the 1.5 μm band. A known difference frequency falling in the far-infrared region can be generated in such a diode by applying the radiations of two CO_2 lasers [17], and this might be subtracted from the 1.5 μm difference frequency to leave a measurable beat frequency. Frequency doubling, e.g. from 1.56 μm to 0.78 μm creates a signal in coincidence with strong and saturable transitions such as the D-lines in rubidium. These are already used as frequency standards at NPL [18] and have been accurately measured by wavelength techniques against visible frequency standards [19,20]. Thus the fibre band at 1.5 μm lies in a potential frequency chain [15] from the caesium standard to visible frequencies, but is a little above the demonstrated range of harmonic mixing in the wide-band MIM point-contact diodes. We shall not here discuss further the extension of MIM mixing, but shall concentrate on frequency standards in the 1.5 μm band, and possible frequency doubling to the near-visible, where sufficiently accurate references for immediate purposes are available. More is given on frequency synthesis aspects at reference [2].

3. THE PRESENT SITUATION AND OPTIONS

Frequency reference points measured in and near the 1.5 μm band are shown in Table 1. Data in the band comprises the early measurement of the ^{20}Ne laser gain centre, together with more recent measurements of unsaturated molecular peaks within ± 20 GHz of it by Sasada and colleagues [7], of which only the strongest are shown. A second point of interest is provided at 1.56 μm at the half frequency of the saturated-absorption stabilised laser at 0.78 μm referenced to the rubidium D-line. This is potentially more accurate because of the strong and narrow s.a. features at 0.78 μm, which have been measured [18] by interferometric comparison against the I_2-stabilised laser at 633 nm. Internally-generated second-harmonic radiation from a DFB diode laser has been used to interrogate the Rb transition directly [21], but it is likely in a metrology laboratory that amplification, frequency doubling in a crystal and heterodyne mixing is desirable to make the connection. A local standard within about ± 20 GHz of the half frequency is also desirable, and lines in HCN and CO are already known in this vicinity [22].

The wider range of accurate spectroscopic measurements at 1.5 μm is shown in Table 2. The band extends from about 1.48 to 1.60 μm. The lines given are mostly molecular, and a substantial number have been measured by Fourier-transform spectroscopy to accuracies better than 1 part per million. Three atomic lines are included in which saturated absorption has

been observed and which offer a potential accuracy of 1 part in 10^9. Stronger lines are noted as a fraction of the total number and refer to the absorption threshold listed at the right of the table.

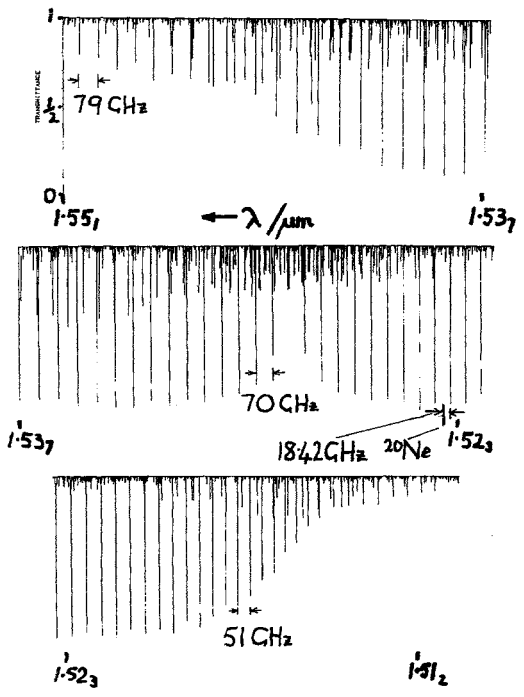


Fig. 1: $H^{13}C^{12}CH$ reference lines.

Figure 1 shows a typical microwave-interval spectral coverage available across the spectrum from, in this case $H^{13}C^{12}CH$, with stronger lines at intervals typically between 80 and 50 GHz. The strong HCCH and NH_3 lines do not extend to 1.56 μm, which, as noted above, is covered by HCN [22]. Also in this region CO has a well-calibrated spectrum, which connects with heterodyne measurements in the 2.3 μm band by Pollock et al. [23].

4. LASER APPARATUS

The lasers in use at NPL are shown in Figure 2. These comprise two 6-m long He-Ne lasers formerly used for frequency-chain work, and two grating-tuned diode lasers recently acquired. The He-Ne lasers operate on the 1.52 μm ^{20}Ne transition with 1:10 ^{20}Ne:^3He fills, to a total pressure of 4 torr (530 Pa). One has a "short" 2.5 m gain tube in the cavity and is designed to be operated with an intracavity metal-film mode selector 170 mm from the output end. It provides a tunable single-mode source with a full tuning range of about 850 MHz. The other long laser is a power oscillator, depending upon pressure broadening in the gain tube and high output coupling, 22%, to suppress modes other than the central mode, and emits about 10 mW in a single mode tunable over about ± 50 MHz. The short laser can be stepped over the mode spacing of 25 MHz by moving the metal film, and tuning without stepping is possible over shorter ranges by moving the far end reflector with a long-travel PZT at the same time. There is space in the cavity for a discharge cell or an absorption cell up to 2.5 m long. The diodes lasers emit a low-power single-mode signal in fibre over a range of about 1.46 to 1.56 μm. These can be heterodyned over more than ± 20 GHz from the HeNe lasers using a mixer diode having -3 dB conversion at 17 GHz.

5. $^{15}NH_3$ IN AN INTRACAVITY CELL

An attempt to interrogate a $^{15}NH_3$ absorption in coincidence with the 1.523 μm Ne line is demonstrated in Figure 3. At the top is shown the laser tuning range observed on an optical spectrum analyser (2 GHz free spectral range) without an absorption cell in the cavity, and below are shown distorted tuning profiles at increasing pressures of $^{15}NH_3$ in the 2.5 m cell. In the lower photographs the tall peak superimposed is the signal from the long laser which is used as a frequency reference. Its whole tuning range is shown. The effect observed was a shifted, distorted and much-reduced tuning profile with 12 torr (1600 Pa) of $^{15}NH_3$ in the cell, also reduced laser output power. Since the signal with 14% output coupler and 6 torr of $^{15}NH_3$ in the cell is similar to that at 12 torr with 1.6% output coupling, the $^{15}NH_3$ absorption coefficient appears to be of order 4 x 10^{-5} per cm.torr.

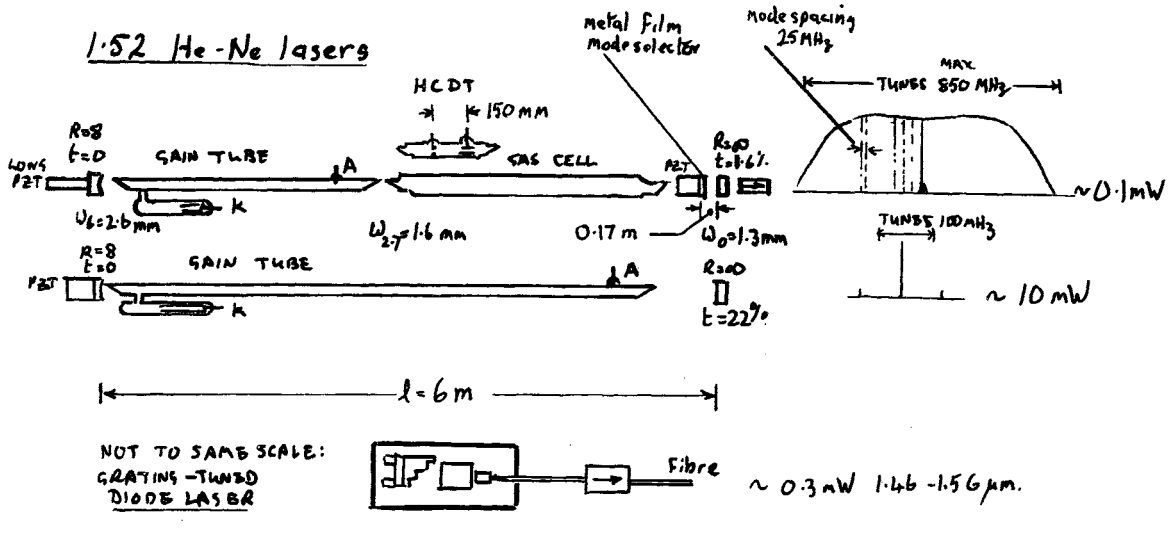


Fig. 2 1.52 μm He-Ne lasers and diode lasers.

However a useful computation of the frequency offset from the ^{20}Ne line is not possible with the information available. The difficulty is created by the pressure width of the $^{15}NH_3$ absorption, as indicated schematically in the intracavity small-signal gain/loss curves of Figure 4. The net gain would follow the dotted shifted curve, giving an output power curve similar to that observed, but this is fairly independent of the $^{15}NH_3$ offset frequency. The result is compatible with that measured by Sasada et al. [7], for 3 torr laser pressure: +248(20) MHz. Searches were also made at 1.6 and 3 torr to try to see a saturated absorption feature but none was observed. The internal counter-propagating intensity used was about 130 mW/cm^2.

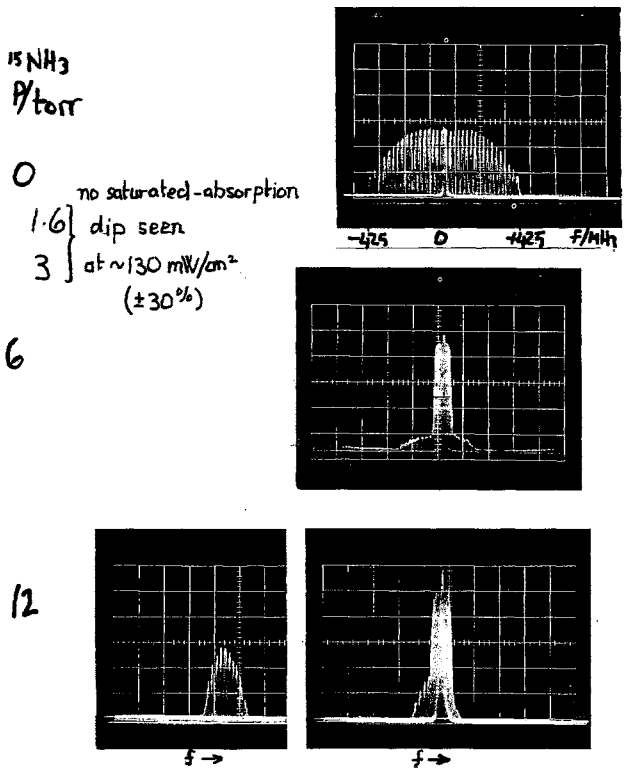


Fig. 3: 1.52 μm He-Ne laser with $^{15}NH_3$ intra cavity.

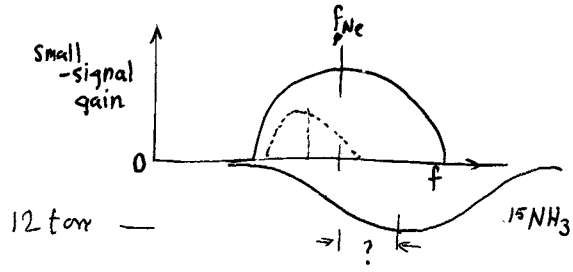


Fig. 4: Schematic effect of $^{15}NH_3$ absorption on He-Ne laser net gain & output (dotted).

6. OPTOGALVANIC SIGNALS FROM ^{20}Ne

If a pure neon discharge tube is inserted in the cavity together with an optical beam chopper, a search can be made with a synchronous detector to find effects on the discharge voltage resulting from interaction of the laser light with the discharge. This optogalvanic signal represents a photon emission or absorption process interacting with the discharge excitation process.

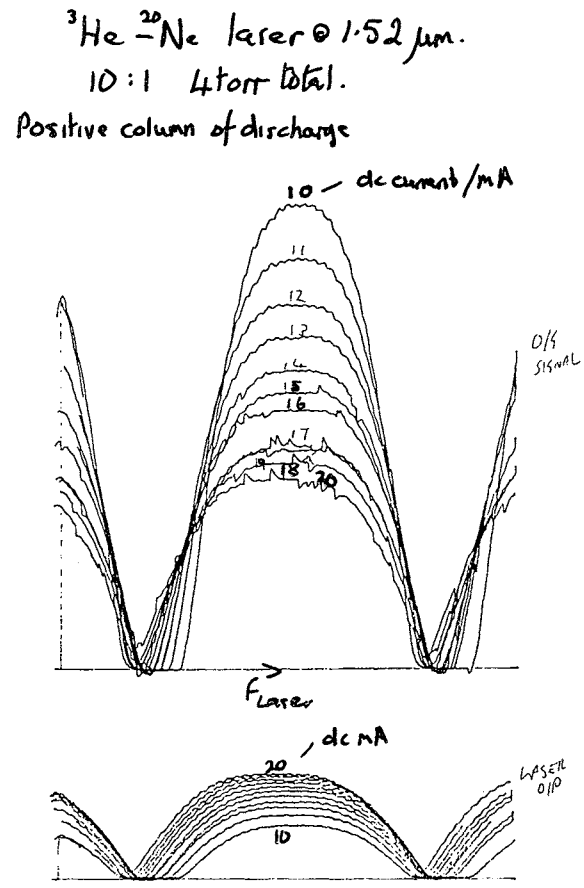


Fig. 5: Optogalvanic signals from He-Ne laser gain tube.

An initial experiment was carried out using the laser gain tube itself as the discharge, and this showed, Figure 5, that the optogalvanic signal increased as the laser discharge current was decreased, showing the effect to be stronger in a weaker discharge, in spite of the fact that the lasing signal intensity decreased at the same time. The optogalvanic signal ought to show a saturated absorption feature in the observed signal if the transition is saturated and the pressure of the gas in the discharge is low enough. In the laser gain tube the pressure is too high, and only the positive column of the dc discharge is subjected to the laser light, where one would expect a stronger intraction in the ionising cathode-fall region.

A further experiment, Figure 6, was carried out with a tunable diode laser from which the light was passed one way through a commercial hollow-cathode discharge tube (HCDT) containing neon at 6 torr. At the same time its signal was heterodyned against that of a He-Ne laser tuned near line centre. A strong optogalvanic signal was obtained showing a symmetrical profile, apart from some tuning irregularities on the left hand side, with a peak corresponding to the He-Ne gain peak within about ± 35 MHz. The width shows a combination of Doppler and pressure broadening, and the line shape corresponds to an expected Voight profile, the laser

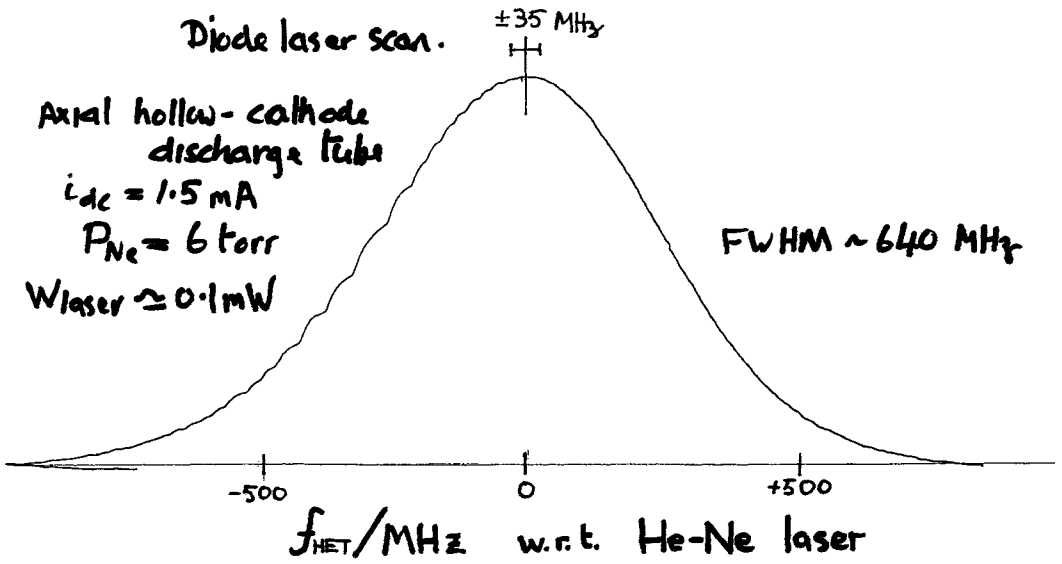


Figure 6. Ne optogalvanic signal at 1.523 μm using a diode laser.

power here remaining approximately constant across the line.

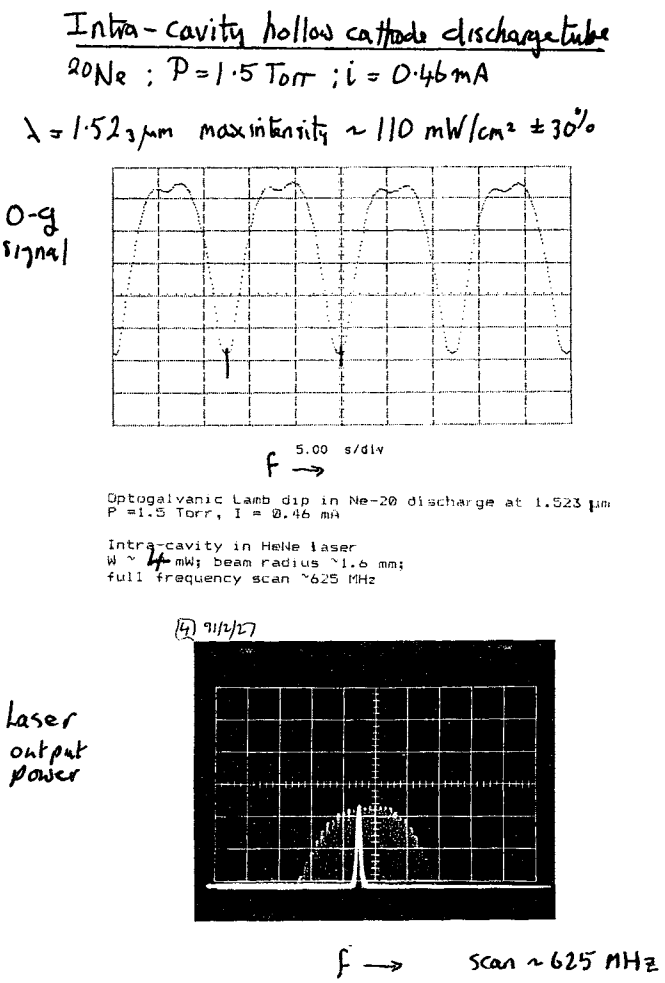


Fig. 7: Saturated absorption using hollow cathode discharge tube in the He-Ne laser cavity.

To seek saturated absorption (s.a.) signals from pure Ne a special discharge tube, with coaxial hollow cathode and ring anode of diameter 12 mm and spaced by 150 mm, was constructed and put in the laser cavity at the position indicated in Figure 2. Good optogalvanic signals were observed at pressures down to about 1 torr. At lower pressures, down to 0.5 torr (70 Pa), the discharge voltage increased sharply, together with discharge noise and cathode sputtering, and optogalvanic signals were weaker. Some s.a. dips are shown in the optogalvanic signals shown in Figure 7, taken with a pressure of 1.5 torr and a dc discharge current of 0.46 mA. The maximum intensity available for these scans was of order 110 mW/cm^2, and the laser tuning profile seen on an optical spectrum analyser is shown below. These traces were taken with a time scan of the metal film mode selector in the laser and on the oscilloscope. There is asymmetry and noise in the observed peaks. The asymmetry was reduced with good alignment of the laser.

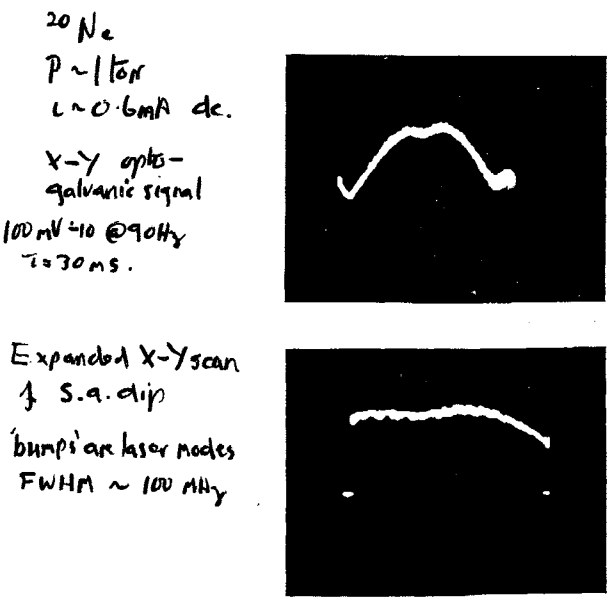


Fig.8: ^{20}Ne saturated absorption using x-y scans: optogalvanic output vs laser frequency.

These traces were taken with a slow scan taking about 10 s to cross the absorption feature, and using a chopper frequency of about 90 Hz for synchronous detection of the optogalvanic signal.

A further investigation of the s.a. dip is shown in Figure 8 at a pressure of 1 torr and a current of 0.6 mA, and for this an x-y display of the optogalvanic signal against the pzt frequency-scan signal is shown. The observed dip has a full width half maximum (FWHM) of about 100 MHz, and amounts to about 6% of the signal height. This compares with a shallow dip seen in the He-Ne laser, of order 1 or 2% when running without absorption cells in the cavity, Figure 3, top. However optogalvanic signals are affected by noise in the discharge tube, which tends to get worse at low pressures because of, e.g., discharge path flicker, and the s.a. dip seems rather broader than is expected.

7. CONCLUSIONS

Our main conclusion is that the ^{20}Ne s.a. dip is observable by the optogalvanic method, but not yet at good signal/noise ratios. Molecular lines are useful for diode lasers, and the optogalvanic atomic lines may also be observed with the latter. For better accuracy in the 1.5 μm band, and for a second main reference point one can in principle frequency double from 1.56 μm to the Rb 0.78 μm saturated absorption lines.

AKNOWLEDGEMENTS

This work was carried out under the UK National Measurement System programme for Optoelectronics.

REFERENCES

1. D. J. E. Knight "Laser frequency standards and laser frequency measurement for the optical fibre bands". Proc. 2nd European Frequency and Time Forum, 16-18 March 1988, Neuchatel, Switzerland, pp 23-33.

2. D J E Knight "Absolute laser frequency standards for the optical communication bands", Colloquium on "Wavelength standards in fibre-optics systems", IEE London, 7 February 1992, IEE Digest no. 1992/029, pp 4/1-4/8.

3. P. Hansell and H. Leeson "Absolute frequency standards for optical communications - Requirements Survey", February 1992, 23pp (UK National Measurement Programme for Opto-Electronics Doc. PRO-AFS-FR-2, National Physical Laboratory, Teddington, Middlesex TW11 0LW UK) Division of Electrical Science.

4. D. J. E. Knight "Absolute laser frequency standards and techniques of frequency measurement for the optical communication bands: research survey and experimental review. Draft DES Report: February 1992, National Physical Laboratory, Teddington Middlesex TW11 0LW UK).

5. D J E Knight, K I Pharaoh and M J Padgett "Laser frequency measurement at NPL", Proc. 4th Symposium on Frequency Standards and Metrology, Ancona, Italy, 5-9 September 1988; A De Marchi Ed. (Springer-Verlag, Berlin 1989) pp 259-260.

6. D A Jennings, F R Petersen and K M Evenson, "Frequency measurement of the 260-THz (1.15 μm) HeNe laser", Optics Lett. 4 (5) May 1979 pp 129-130, see note 10.

7. H Sasada, S Takeuchi, M Iritani and K Nakatani "Semiconductor laser heterodyne frequency measurements on the 1.52-μm molecular transitions". J. Opt. Soc. Amer. B 8 (4) April 1991 pp 713-8.

8. J W Eerkens, W Lee, R Smith and R S Vodhanel "Ammonia stabilisation of the 1.5 μm HeNe laser for fiber-optic applications", presented at the 2nd Int. Tech. Symp. on Optical and Electro-optic Science and Engineering and Instrument Display, Cannes France, 25-29 November 1985, (Melles Griot, 4pp).

9. W J Tomlinson and R L Fork "Frequency stabilisation of a gas laser", Appl. Optics 8, No. 1, January 1969 pp 121-129.

10. Wu Y, Guan K, Yu T and Tan F "Experiment on 1500 nm stabilised Zeeman laser" Electron. Lett. 23 (7) 26 March 1987 pp 318-9.

11. A J Lucero, Y C Chung and R W Tkach "Survey of atomic transitions for absolute frequency locking of lasers for lightwave systems". IEEE Photonics Technol. Lett. 3 (5) May 1991 pp 484-486.

12. A J Lucero, Y C Chung, S Reilly and R W Tkach "Saturation measurements of excited-state transitions on noble gases using the optogalvanic effect". Optics Lett. 16 (11) 1 June 1991 pp 349-351.

13. S L Gilbert "High resolution spectroscopy using fiber lasers", in Proc. 10th Int. Conf. on Laser Spectroscopy, Font-Romeu, France, June 1991 (World Scientific, M Ducloy, E Giacobino and F Camy Eds.)

14. S L Gilbert "Frequency stabilisation of a tunable erbium-doped fibre laser". Optics Lett. 16 1991 pp 150ff.

15. D A Jennings, K M Evenson and D J E Knight "Optical frequency measurements" Proc. IEEE 74 (1) January 1986 pp 168-179.

16. C R Pollock, D A Jennings, F R Petersen, J S Wells, R E Drullinger, E C Beaty and K M Evenson, "Direct frequency measurement of transitions at 520 THz (560 nm) in iodine and 260 THz (1.15 μm) in neon", Optics Lett. 8 (3) March 1983 pp 133-135.

17. K M Everson, D A Jennings and F R Petersen "Tunable far-infrared spectroscopy", Appl. Phys. Lett. 44 (6) 15 March 1984 pp 576-578.

18. G P Barwood, P Gill and W R C Rowley "Frequency measurements on optically narrowed Rb-stabilised laser diodes at 780 nm and 795 nm", Appl. Phys. B53 1991 pp 142-147.

19. BIPM "Documents concerning the new definition of the metre", Metrologia 19 1984 pp 163-177.

20. D J E Knight "A tabulation of absolute laser-frequency measurements" Metrologia 22 1986 pp 251-257.

21. M Ohtsu and E Ikegami "Novel frequency control scheme for a 1.56-μm DFB laser using an internal second harmonic and an atomic rubidium line", Proc. CLEO 1989, Baltimore Md, April 1989, paper TUD4, 2pp.

22. H Sasada and K Yamada "The calibration lines of HCN in the 1.5 μm region", Appl. Optics 29 (24) 20 August 1990 pp 3535-3547.

23. C R Pollock, F R Petersen, D A Jennings, J S Wells and A G Maki, "Absolute frequency measurements of the 2-0 band of CO at 2.3 μm; calibration standard frequencies from high resolution color center laser spectroscopy", J. Mol. Spectrosc. 99 1983 pp 357-368.

24. S Kinugawa and H Sasada "Wavenumber measurement of the 1.5 μm band of acetylene by semiconductor laser spectrometer", Japan J. Appl. Phys. 29 (3) March 1990 pp 611-612.

25. M DeLabachelerie C Latrasse, K Diomande, P Kemssu and P Cerez "A 1.5 μm absolutely stabilised extended-cavity semiconductor laser" Proc. Conf. on Precision Electrical Measurements, Ottawa, June 1990. Preprint received July 1990.

List of Exhibitors

EXHIBITORS

1. RODITI INTERN. Corp. Ltd.
   Londen, England

2. COMPOSANTS QUARTZ ELEKTRONIQUE
   Pont Ste Marie, France

3. TELE QUARTZ GmbH
   Neckarbischofsheim, Germany

4. FTS/AUSTRON Inc.
   Austin, Texas, United States

5. QUARZKERAMIK GmbH
   Stockdorf, Germany

6. AT & T
   Hilversum, The Netherlands

7. HEWLETT - PACKARD
   Santa Clara, California, United States

8. CEPE
   Argenteuil, France

9. TIME TECH GmbH
   Stuttgart, Germany

10. STEATITE INSULATIONS Ltd.
    Birmingham, England

11. TELECOM SOLUTIONS EUROPE
    Leatherhead, Surrey, England

12. TRAK MICROWAVE CORPORATION
    Tampa, Florida, United States

13. GUIDE TECHNOLOGY Inc.
    San Jose, California, United States

List of Participants

Mr. L. ADAMS
ESA/ESTEC, QCA
Keplerlaan 1
2200 AG Noordwijk
NETHERLANDS

Mr. J.P. AUBRY
CEPE
44, Avenue de la Glaciere
BP 165 Argenteuil
FRANCE

Mr. E. BAVA
Politecnico Milano
Dip. Elettronica
Piazza Leonardo da Vinci 32
20133 Milano
ITALY

Mr. F. ALFONSO
T.F.E. SA.
C. Lopez Bravo
09179 Boreos
SPAIN

Mr. C.L. AUDOIN
Lab. Horloge Atomique
Universite Paris-Sud
Batiment 221
91405 Orsay
FRANCE

Mr. A. BECK
Telequarz GmbH
Landstrasse
6924 Neckarbischofsheim 2
GERMANY

Mr. Y.A. AMAND
SAGEM
72-74 Rue de la Tour
Billy BP 72
95101 Argenteuil Cedex
FRANCE

Mr. R. AVERY
Telecom Solutions
85, West Tasman Drive
San Jose
California 951341703
UNITED STATES

Mr. J. BENEDICTO
ESA/Estec
Keplerlaan 1
2200 AG Noordwijk
NETHERLANDS

Mr. R. AMBROSINI
Radioastronomy Institute CNR
Irnerio 46
Bologna 40126
ITALY

Mr. O. BAETZ
MBB
Dept. KT155
P.O. Box 801169
8000 Munich 80
GERMANY

Mr. H.C. BERARD
CEPE
44 Avenue de la Glaciere
BP 165
95105 Argenteuil
FRANCE

Mr. E. AMSCHLER
Quarzkeramik GmbH
Gautinger Str. 23
Postfach 29
D-8035 Stockdorf
GERMANY

Mr. A. BALLATO
US Army
Electronics Techn. & Devices Lab.
Ft. Monmouth
07703-5601 New Jersey
UNITED STATES

Mr. L-G. BERNIER
OBSERVATOIRE CANTONAL
Rue de l'Observatoire 58
2000 Neuchatel
SWITZERLAND

Mr. R. ARNAUD
S.I.C.N.
4, Rue du Radar
74008 Annecy
FRANCE

Mr. W. BALZER
KVG
Postfach 61
W-6924 Neckarbischofsheim
GERMANY

Mr. P. BERTRAND
Tekelec-Systemes
29, Avenue de la Baltique
Z.A. de Courtaboeuf
91953 Les Ulis
FRANCE

Mr. J.C. AUBRIL
CEPE
44 Avenue de la Glaciere
BP 165 Argenteuil
95105
FRANCE

Mr. R. BARILLET
C.N.R.S.
Lab. de l'Horloge Atomique
UPS, Batiment 221
91405 Orsay
FRANCE

Mr. R.J. BESSON
Ecole Nat. Superieure Mecanique
La Bouloie
Route de Gray
25030 Besancon
FRANCE

Mr. J.S. BHARJ
ESA/Estec
XRM
Keplerlaan 1
2200 AG Noordwijk
NETHERLANDS

Mr. E. BIGLER
L.P.M.O.-CNRS
32, Avenue de l'Observatoire
25000 Besancon
FRANCE

Mr. G.A. BJORNSTROM
ESA/Estec, XRT
Keplerlaan 1
2200 AG Noordwijk
NETHERLANDS

Mr. L. BJORSET
ESA/Estec/YMM
Keplerlaan 1
2200 AG Noordwijk
NETHERLANDS

Mr. G. BOIME
Tekelec-Systemes
29, Avenue de la Baltique
Z.A. de Courtaboeuf
91953 Les Ulis
FRANCE

Mr. R. BOLOIX
Observatorio Armada
Cecilio Pujazon s/n
11110 San Fernando-Cadiz
SPAIN

Mr. P. BOUNIOL
LRBA
BP 914
F27207 Vernon
FRANCE

Mr. R.B. BOURQUIN
ENSMM
Route de Gray
25030 Besancon
FRANCE

Mr. J.J. BOY
ENSMM/LCEP
Route de Gray
25030 Besancon cedex
FRANCE

Mr. E. BRAENNSTROEM
Telia Research
SNT Laboratory
S-13680 Haninge
SWEDEN

Mr. B.B.A. BRAHIM BOUNAB
LCEP/ENSMM
La Bouloie
Route de Gray
25030 Besancon
FRANCE

Mr. R. BREITWEG
Alcatel SEL
Lorenzstrasse 10
W-7000 Stuttgart 40
GERMANY

Mr. R. BRENDEL
L.P.M.O.-CNRS
32, Avenue de l'Observatoire
25000 Besancon
FRANCE

Mr. K. BROD
Quarzkeramik GmbH
Gautinger Str. 23
Postfach 29
D-8035 Stockdorf
GERMANY

Mr. M. BRUNET
CNES
18, Avenue E. Belin
31055 Toulouse
FRANCE

Mr. M. BRUNET
C.Q.E.
BP 16
Rue Keller Zone Industrielle
10150 Pont Sainte Marie
FRANCE

Mr. A.C. BUCHLEITNER
Euro Quarz GmbH
Betriebstrasse 16
A-2620 Ternitz
AUSTRIA

Mr. X. BUISSON
SICN
4 Rue de Radar
74008 Annecy
FRANCE

Mr. G. BUSCA
OBSERVATOIRE CANTONAL
Rue de l'Observatoire
2000 Neuchatel
SWITZERLAND

Mr. E.W. BUTTERLINE
AT&T
RM 3c130e
Routes 202-206 North
Bedminster NJ 07921
UNITED STATES

Mr. B. CAPELLE
CNRS
B 115
4, Place Jussicu
75252 Paris cedx 05
FRANCE

Ms. A. CARLUCCI
SIRTI Spa
Via Eurico Feru 2
Cassina de Pechi,
20060 Milano
ITALY

Mr. G. CAROFALO
ESA/Estec
XRF
Keplerlaan 1
2200 AG Noordwijk
NETHERLANDS

Mr. C. CAVALLONI
KISTLER INSTRUMENTE AG
Eulachstrasse 22
CH-8408 Winterthur
SWITZERLAND

Mr. P. CEREZ
C.N.R.S.
LHA Bat. 221
Universite Paris-Sud
91405 Orsay
FRANCE

Mr. M. CHAMBON
BNM
22 Rue Monge
75005 Paris
FRANCE

Mr. D.C. CHATURVEDI
ESA/Estec
WDP
Keplerlaan 1
2200 AG Noordwijk
NETHERLANDS

Mr. A.C. CLAIRON
L.P.T.F. (BNM)
61, Avenue de l'Observatoire
75014 Paris
FRANCE

Mr. R. COIRAULT
ESA/Estec
XR
Keplerlaan 1
2200 AG Noordwijk
NETHERLANDS

Mr. P.C. COLLET
CQE
BP 16
Rue Keller Zone Industrielle
10150 Saint Marie
FRANCE

Mr. K. CONTOS
FOQ PIEZO TECHNIK
Riemenstrasse 33
D-6927 Bad Rappenau
GERMANY

Mr. F. CORDARA
Instituto Electrotecnico Nazionale

Strada Delle Cacce 91
10135 Torino
ITALY

Ms. C. COUPLET
OBSERVATOIRE NEUCHATEL
Rue de l'Observatoire 58
2000 Neuchatel
SWITZERLAND

Mr. M.P. CRACKNELL
Northern Telecom
17 Windmill Fields Harlow
Essex CM17QLQ,
UNITED KINGDOM

Mr. L.S. CUTLER
Hewlett Packard Laboratories
P.O.Box 10350
Palo Alto
Cal. 94303-0867
UNITED STATES

Mr. S. DALLA PIAZZA
ASULAB. S.A.
Passage Max-Meuron 6
CH-2001 Neuchatel
SWITZERLAND

Mr. W. DANIAU
L.P.M.O.-CNRS
32, Avenue de l'Observatoire
25000 Besancon
FRANCE

Mr. J. DAVIS
NPL
Queens Road
Teddington, Middlesex TW110LW
UNITED KINGDOM

Mr. E. DE CLERCQ
Observatoire de Paris, LPTF
61, Avenue de l'Observatoire
75014 Paris
FRANCE

Mr. G. DE JONG
Van Swinden Lab.
P.O. Box 654
2600 AR Delft
NETHERLANDS

Mr. A.D. DE MARCHI
Politecnico di Torino
Dip. Electronica
Cso. Duca Degu Abruzzi 24
10129 Torino
ITALY

Mr. R. DECKER
AT&T
Two Gateway Center
Routes 202 and 206 North
Newark, New Jersey 07102
UNITED STATES

Mr. N. DEMIDOV
KVARZ
Institute of Electronic Measurement
Nizhny Novgograd
Moscow
SOVIET UNION

Mr. G.J. DICK
Jet Propulsion Laboratory
Mailstation 298-100
4800 Oak Grove Drive Pasadena
California 91109
UNITED STATES

Dr. B.D. EISSFELLER
Kayser-Threde GmbH
Wolfratshauserstrasse 48
8000 Munchen 70
GERMANY

Mr. A. DENTON-MILLER
Telecom Solutions
17, Bridge Street
Leatherhead
Surrey KT22 8BL
UNITED KINGDOM

Mr. M. DIENERT
TIMETECH GMBH
Nobelstrasse 15
D-7000 Stuttgart 80
GERMANY

Mr. D.A, EMMONS
Frequency and Time Systems
34 Tozer Road
Beverly MA 01915
UNITED STATES

Mr. M. DESAINTFUSCIEN
Laboratoire de l Horloge Atomique
Batiment 221
Universite Paris Sud
91405 Orsay
FRANCE

Mr. R.E. DRULLINGER
NIST
325 Broadway
Boulder
UNITED STATES

Mr. T. ENGLISH
Ball Efratom
3 Parker
Irvine
California 92718-1605
UNITED STATES

Mr. J. DETAINT
CNET
Dept. MCT/CMM
196, Avenue H. Ravera
92220 Bagneux
FRANCE

Mr. H. DUCHAUSSOY
DRET
26 Bd. Victor
00460 Paris
FRANCE

Mr. S.J. FELTHAM
ESA/Estec, XRM
Keplerlaan 1
2200 AG Noordwijk
NETHERLANDS

Mr. E. DETOMA
FIAT-CIEI-SEPA
Corso Giulio Cesare 294-300
10154-Torino To
ITALY

Mr. G. DUDLE
OBSERVATOIRE CANTONAL
Rue de l Observatoire 58
2000 Neuchatel
SWITZERLAND

Mr. J.M. FERNANDEZ
TFE, S.A.
C/Lopes Bravo
Pd. Ind. Villalouquejar.
09197 Burgos
SPAIN

Mr. A.D. DEVAL
SAGEM
BP 72
95101 Argenteuil Cedex
FRANCE

Mr. B.M. DULMET
ENSMM
Route de Gray
25030 Besancon
FRANCE

Mr. A. FEU
TFE, S.A.
C/Lopes Bravo
Pd. Ind. Villalouquejar
09179 Burgos
SPAIN

Mr. F. DEYZAC
ONERA
Onera BP 72
F-92322 Chatillon-Cedex
FRANCE

Mr. M. ECABERT
F.S.R.M.
Foundation for Res. in Microtechn.
Orangerie 8
CH 2000 Meuchatel
SWITZERLAND

Mr. B. FISCHER
Phys. Techn. Bundensanstalt
Bundesallee 100
D 3300 Braunschweig
GERMANY

Mr. G. FREON
Observatoire de Paris
LPTF
61, Avenue de l'Observatoire
75014 Paris
FRANCE

Mr. M. FREY
Telequarz GmbH
Landstrasse
W-6924 Neckarbischofsheim 2
GERMANY

Mr. FRIDELANCE
OCA/CERGA
Rue Copernic
06130 Grasse
FRANCE

Mr. S. GALLIOU
LCEP/ENSMM
La Bouloie
Route de Gray
25030 Besancon
FRANCE

Mr. G. GAROFALO
ESA/Estec, XRF
Keplerlaan 1
2200 AG Noordwijk
NETHERLANDS

Mr. R.M. GARVEY
FREQUENCY & TIME SYSTEMS
34 Tozer Rd.
Beverly MA 01915
UNITED STATES

Mr. G. GATTI
ESA/Estec, XRM
Keplerlaan 1
2200 AG Noordwijk
NETHERLANDS

Mr. M.P. GAWARECKI
FREQUENCY & TIME SYSTEMS
Austron Inc.
P.O. Box 14766
TX 78761-4766 Austin
UNITED STATES

Mr. E. GERARD
CERE
44 Avenue de la Glaciere
BP 165 Argenteuil
FRANCE

Mr. V. GIORDANO
C.N.R.S.
LHA- Bat. 221
Universite Paris-Sud
91405 Orsay
FRANCE

Ms. S. GRAEFE
Quarzkeramik GmbH
Gautinger Str. 23
Postfach 29
D-8035 Stockdorf
GERMANY

Mr. E. GRAF
Oscilloquartz SA
Brevards 16
2000 Neuchatel
SWITZERLAND

Mr. M.P. GRANVEAUD
LPTF/Paris Observatory
61, Avenue Observatoire
75014 Paris
FRANCE

Mr. W. GREINER
ESA/Estec, XRF
Keplerlaan 1
2200 AG Noordwijk
NETHERLANDS

Mr. P. GRUDLER
Observatoire de la Cote Dázur
OCA/CERGA
Av. Nicolas Copernic
06130 Grasse
FRANCE

Mr. J. GUIJARRO
ESA/Estec, ORP
Keplerlaan 1
2200 AG Noordwijk
NETHERLANDS

Mr. B. GUINOT
B.I.P.M.
Rue des Soupirs
Pavillon de Breteuil
77590 Chartrettes
FRANCE

Mr. J.M. GUYON-GELLIN
DRET
26, Boulevard Victor-Paris
00460 Armees
FRANCE

Mr. A. HADDRELL
Navstar Ltd
Royal Oak Way
Daventry
Northants NN11 5PJ
UNITED KINGDOM

Mr. B. HAGERS
PTT Research Laboratorium
St. Paulusstraat 4
2264 XZ Leidschendam
NETHERLANDS

Mr. M. HAMILTON
Telecom Solutions
85, West Tasman Drive
San Jose
California 95134-1703
UNITED STATES

Mr. H. HANDREK
Quarzkeramik GmbH
Gautinger Str. 23
Postfach 29
D-8035 Stockdorf
GERMANY

Mr. J. HIEMSTRA
Klove Electronics
Industriestraat 3
1704 AA Heerhugowaard
NETHERLANDS

Mr. B.J. JAMES
GEC Marconi First Research Centre
East Lane
Wembley Middlesex
UNITED KINGDOM

Mr. H.H. HANSEN
TFL
Lyngsoe Alle 2
DK 2970 Hoersholm
DENMARK

Mr. G.B. HOLMES
Telecom Solutions
85 West Tasman Drive
San Jose
California 915341703
UNITED STATES

Mr. JANIAUD
ONERA
29, Avenue de la Div. Leclerc
92320 Chatillon
FRANCE

Mr. P.H. HARTL
University Stuttgart
Institute of navigation
Keplerstrasse 11
D-7000 Stuttgart 1
GERMANY

Mr. C.K. HRUSKA
York University
Dept. of Math, Room N520 Ross
4700 Keele Street
North York, Ont. M3J 1P3
CANADA

Mr. G. JEANDEL
LABORATOIRE INFRARONGE
FACULTE DES SCIENCES
B.P. 239
54506 Vandoeuvre les Nancy
FRANCE

Mr. D.B. HAUDEN
L.P.M.O.-CNRS
32, Avenue de l'Observatoire
25000 Besancon
FRANCE

Mr. H.X. HUANG
Shanghai Observatory
Chinese Academy of Sciences
80, Nandan Road
Shanghai
PEOPLES REP. OF CHINA

Ms. C. JIMENEZ
ESA/ESTEC, GIS
Keplerlaan 1
2200 AG Noordwijk
NETHERLANDS

Mr. J.W. HE
Shanghai Observatory
Chinese Academy of Sciences
80, Nandan Road
Shanghai
PEOPLES REP. OF CHINA

Mr. J.F HUNGERFORD
FREQUENCY & TIME SYSTEMS
Austron Inc.
34 Tozer Rd.
Beverly MA 01915
UNITED STATES

Mr. J.L. JOSSET
ESA/Estec, SC
Keplerlaan 1
2200 AG Noordwijk
NETHERLANDS

Mr. G. HERMAN
Tekelec Systemes

FRANCE

Mr. J. HUNT
ESA/Estec, IP
Keplerlaan 1
2200 AG Noordwijk
NETHERLANDS

Mr. B.J. JUOMPAN
L.P.T.F.
Observatoire de Paris
77, Bd. Denfert Rocherow
75014 Paris
FRANCE

Mr. P. HETZEL
Phys. Techn. Bundesanstalt
Bundesallee 100
D 3300 Braunschweig
GERMANY

Mr. H.I. IMAJO
Kansai Advanced Research Center
588-2 Iwaoka, Iwaoka-cho
Nishi-ku, Kobe-shi
Hyogo 651-24
JAPAN

Mr. K. KALLIOMAEKI
University of Oulu
Sf 90570 Oulu
FINLAND

Mr. P. KARTASCHOFF
Swiss PTT
Swiss PTT R en D
Seltion VD2
CH-3000 Bern 29
SWITZERLAND

Mr. R.A. KIVELA
Jutel Oy
Teknologiantie 11
SF-90570 Oulu
FINLAND

Mr. N. KURAMOCHI
Tokyo Institute of Technology
4259 Nagatsuta
Midoriku
Yokohama 227
JAPAN

Mr. S. KATTAN
Guide Technology Inc.
920 Saratoga
Ave Suite 215
San Jose CA 95129
UNITED STATES

Mr. P.K. KLEIN
STEI
18, rue de Docteur Zamenhof
92131 Issy les Moulineaux
FRANCE

Mr. E. KUSHNER
Frequency Electronics Inc.
55 Charles Lindbergh Blvd.
Mitchel Field
NY 11553
UNITED STATES

Mr. E.W. KENTLEY
RODITI Int. Corp. Ltd
130 regent Street
London W1R 6BR
UNITED KINGDOM

Mr. D.I.E. KNIGHT
National Physical Laboratory
Teddington
Middlesex TW11 OLW
UNITED KINGDOM

Mr. J. KUSTERS
Hewlett-Packard co
5301 Stevens Creek Boulevard
Santa Clara, Ca 95052-8059
UNITED STATES

Mr. A.A. KEUS
PTT Research Laboratorium
St. Paulusstraat 4
2264 XZ Leidschendam
NETHERLANDS

Mr. R. KOHL
MBB
Dept. KT155, P.O. Box 801169
8000 Munich 80
GERMANY

Mr. W.R.L. LANGE
Lange Electronic GmbH
Ganghoferstrasse 29
8031 Gernlinden
GERMANY

Mr. M.K. KIHARA
NTT Trans. Sys. Labs.
1-2356 Take Yokosuka Kanagawa
238-03
JAPAN

Mr. P.W. KREMPL
AVL LIST GmbH
Kleiststr. 48
A-8020 Graz
AUSTRIA

Mr. F.L. LAZARUS
HEWLETT-PACKARD
150 Route du Nant-d'Avril
CH-1217 Meyrin 2 Geneva
SWITZERLAND

Mr. R. KIMBERLEY
ACAL ELECTRONICS LTD
Technitron House
Redfields Park
Church Crookham Fleet GU13 ORD
UNITED KINGDOM

Mr. D. KRUYVER
Klove Electronics
Industriestraat 3
1704 AA Heerhugowaard
NETHERLANDS

Ms. T.G. LEBLOIS
ENSMM/LCEP
Route de Gray
25030 Besancon
FRANCE

Mr. D. KIRCHNER
TU-GRAZ
Inffeldgasse 12
A-8010 Graz
AUSTRIA

Mr. H. KUNIMORI
Communications Research Lab.
2-1 Nukui-Kitamachi
4-Chome, Koganei-SHI
Tokyo 184
JAPAN

Mr. CH. LEHNER
DLR
Postfach 1126
8120 Weilheim
GERMANY

Mr. S. LESCHIUTTA
Politecnico Torino
Corso Duca Abruzzi 124
10129 Torino
ITALY

Mr. R.J. LOPES
AT&T
Two Gateway Center-14th Floor
Routes 202 and 206 North
Bedminister, NJ 07921
UNITED STATES

Mr. T. MANSTEN
Tech. Res. Ctr. of Finland
VTT/MTI
P.O. Box 34
SF-02151 Espoo
FINLAND

Mr. W. LEWANDOWSKI
BIPM
Pavillon de Breteuil
92312 Sevres cedex
FRANCE

Mr. P.R. LOPEZ
TRAK Microwave Corporation
4726 Eisenhower Blvd.
Tampa, Florida 33634-6391
UNITED STATES

Mr. M.F.S. MARQUES
IPQ
R. Prof. Reinaldo dos Santos Lote 1374
1500 Lisboa
PORTUGAL

Mr. P. LIAKOPOULOS
Netcom Ltd
Metamorfoseos 4
15344 Glyka Nera-Attiki
GREECE

Mr. M. LOPRIORE
ESA/Estec, Electrical Systems Dept.
Keplerlaan 1
2200 AG Noordwijk
NETHERLANDS

Mr. S. MARRA
ESA/Estec, XEE
Keplerlaan 1
2200 AG Noordwijk
NETHERLANDS

Mr. E.E. LIJPHART
ESA/Esoc
Robert Boschstrasse 5
6100 Darmstadt
GERMANY

Mr. S. LORENZI
SIRTI Spa
Via Vida 19
Milano 20127
ITALY

Mr. D.D.E. MARTIN
ESA/Estec, SC
Keplerlaan 1
2200 AG Noordwijk
NETHERLANDS

Mr. K. LIND
Norwegian Telecom Research
Box 83
N-2007 Kjeller
NORWAY

Mr. R. LUCAS
ESA/Estec, XRT
Keplerlaan 1
2200 AG Noordwijk
NETHERLANDS

Mr. M. MARTIN-NEIRA
ESA/Estec, XRT
Keplerlaan 1
2200 AG Noordwijk
NETHERLANDS

Mr. U. LINNEKUHLE
c/o RODITI Int. Corp.. Ltd
Bergedorfer Strasse 142
2050 Hamburg 80
GERMANY

Mr. R. MANETSBERGER
DLR
Postfach 1126
8120 Weilheim
GERMANY

Mr. J.S. MAS
ESA/Estec, XRM
Keplerlaan 1
2200 AG Noordwijk
NETHERLANDS

Mr. S. LOCKE
ESA/Estec, XRM
Keplerlaan 1
2200 AG Noordwijk
NETHERLANDS

Mr. J.K. MANNERMAA
University of Oulu
Mittaustekn. Lab.
90570 Oulu
FINLAND

Mr. I. MATEESCU
IFTMIFA
Str. Armeneasca nr. 28A
et.I ap.2 Sector 2 COD 70228
Bucaresti
RUMENIA

Mr. H. MATTHEY
VIBRO-METER SA
Sagefeldweg 55
CH 2504 Biel
SWITZERLAND

Mr. J.S. MC NABB
TRAK Microwave Corporation
4726 Eisenhower Blvd.
Tampa
Florida 33634-6391
UNITED STATES

Mr. F. MEYER
Observatoire de Besancon
41 B, Avenue de l'Observatoire
BP 1615
25010 Besancon
FRANCE

Mr. A. MICHAUD
University of Oxford
Clarendon Lab.
Parks Road
Oxford OX1-3PU
UNITED KINGDOM

Mr. JP. MICHEL
CEPE
44, Avenue de la Glaciere
95100 Argenteuil
FRANCE

Mr. G. MILETI
OBSERVATOIRE CANTONAL
Rue de l'Observatoire 58
2000 Neuchatel
SWITZERLAND

Mr. Y. MINGUY
DCN Brest
29240 Brest Naval
FRANCE

Mr. M. MOUREY
ENSMM/LCEP
Route de Gray
25030 Besancon
FRANCE

Mr. V.M. MOURIER
Tekelec-Systemes
29, Avenue de la Baltique
Z.A. de Courtaboeuf
91953 Les Ulis
FRANCE

Mr. L. MUREDDU
STAZIONE ASTRONOMICA
Via Ospedale 72
09124 Gagliari
ITALY

Mr. B. NEUBIG
Telequarz GmbH
Landstrasse
6924 Neckarbischofsheim 2
GERMANY

Mr. C. NIELSEN
ESA/Esoc
Robert Boschstrasse 5
6100 Darmstadt
GERMANY

Mr. J.R. NORTON
Johns Hopkins University
Johns Hopkins Road
Laurel, MD 20703
UNITED STATES

Mr. M. NUNN
Matra Marconi Space UK
Anchorage Road
Porthsmouth
UNITED KINGDOM

Mr. M. OLIVIER
L.P.M.O.-CNRS
32, Avenue de l'Observatoire
25000 Besancon
FRANCE

Mr. E. ORGIL
Time & Frequency Ltd.

ISRAEL

Mr. N. OURA
Tokyo Institute of Technology
4259 Nagatsuta
Midoriku
Yokohama 227
JAPAN

Mr. P. PAQUET
Observatoire Royal de Belgique
Av. Circulaire 3
1180 Brussels
BELGIUM

Mr. J. PEREZ
Tekelec-Systemes
29, Avenue de la Baltique
Z.A. de Courtaboeuf
91953 Les Ulis
FRANCE

Mr. G. PETIT
BIPM
Pavillon de Breteuil
92312 Sevres cedex
FRANCE

Mr. M. PETIT
DGA/CELAR
35170 Bruz
FRANCE

Ms. L. PFAFF
Quarzkeramik GmbH
Gautinger Str. 23
Postfach 29
D-8035 Stockdorf
GERMANY

Mr. W.P. PFLAUM
Ball Efratom
3, Parker Irvine
92118-1605 California
UNITED STATES

Mr. R. PFLEIDERER
Alcatel SEL
Lorenzstrasse 10
W-7000 Stuttgart 40
GERMANY

Mr. E. PHILIPPOT
C.N.R.S.
Lab. Physicochemie & Materiaux
URA 407 Place E. Bataillin
34095 Montpellier
FRANCE

Mr. H.D. PIANOWSKI
EFG GmbH
Dueppelstr. 13
1000 Berlin 37
GERMANY

Mr. G. PILBRATT
ESA/Estec
SA
Keplerlaan 1
2200 AG Noordwijk
NETHERLANDS

Mr. I.C. PIMM
ESA/ESTEC
QCT
Keplerlaan 1
2200 AG Noordwijk
NETHERLANDS

Mr. R.L. PLACITO
Steatite Insulations Ltd
2, The Square
Broad Street
Birmingham B15 1AP
UNITED KINGDOM

Mr. M. PLANAT
L.P.M.O.-CNRS
32, Avenue de l'Observatoire
25000 Besancon
FRANCE

Mr. V.P. POINTET
CIR
Nevenburgstrasse 7
CH-2076 Gals/BE
SWITZERLAND

Mr. L. POPKEN
ESA/Estec
XRT
Keplerlaan 1
2200 AG Noordwijk
NETHERLANDS

Mr. G. PRISCO
ESA/ESTEC
WMA
Keplerlaan 1
2200 AG Noordwijk
NETHERLANDS

Mr. L. PROST
Office Federal de Metrologie
Lindenweg 50
CH 3084 Wabern/Bern
SWITZERLAND

Mr. P. PULLENS
CER bv.
Postbus 258
Roosendaal
NETHERLANDS

Mr. S. REINECK
KVG
Postfach 61
D-6924 Neckarbischofsheim
GERMANY

Mr. A. REMOND
Observatoire de Besancon
41 bis, Avenue de l'Observatoire
BP 1615
25010 Besancon Cedex
FRANCE

Mr. H. RESSLER
TU-GRAZ
Infeldgasse 12
A-8010 Graz
AUSTRIA

Mr. W.J. RILEY
EG&G FREQUENCY PRODUCTS
35 Congress Street Salem
01970 Massachusetts
UNITED STATES

Mr. E. ROBERT
CNES
Siege Social
18, Avenue Belin
Toulouse
FRANCE

Mr. Y. ROBIC
Dassault Electronique
55 Quai
Marcel Dassault 92214
St. Cloud
FRANCE

Mr. H. ROTH
Ball Efratom
8011 Hofolding
Fichtenstrasse 25
GERMANY

Mr. E. RUBIOLA
Politecnico di Torino
Dipart. Electronica
C/so Duca Degli Abruzzi 24
10123 Torino
ITALY

Mr. A.L. SATIN
AEROSPACE CORP.
2350 E. El Segundo Blvd.
Los Angeles, Ca 90009
UNITED STATES

Mr. L.E. SCHULZKE
KVG
Postfach 61
D-6924 Neckarbischofsheim
GERMANY

Ms. I. RUEDI
OBSERVATOIRE CANTONAL
Rue de l'Observatoire 58
2000 Neuchatel
SWITZERLAND

Mr. G. SATTERTHWAITE
Telecom Solutions
17, Bridge Street
Leatherhead, Surrey
UNITED KINGDOM

Mr. P. SCHUMACHER
OBSERVATOIRE CANTONAL
Rue de l'Observatoire 58
2000 Neuchatel
SWITZERLAND

Mr. J. RUEFENACHT
Swiss PTT
Technical Centre VM42
CH-3000 Berne 29
SWITZERLAND

Mr. S.L. SAUNDERS
RODITI Int. Corp. Ltd
130 regent Street
London W1R 6BR
UNITED KINGDOM

Mr. J. SCHWARTEL
CNET
Dept. MCT/CMM
196, Avenue H. Ravera
92220 Bagneux
FRANCE

Mr. C.R. RUINEAU
Tekelec Microwave
33, Avenue Faidherbe
93106 Montreuil
FRANCE

Mr. W. SCHAEFER
University of Stuttgart
Institute of Navigation
Senefelder Strasse 26
7000 Stuttgart 1
Germany

Mr. H. SCHWEDA
OBS. CANTONAL NEUCHATEL
Rue de l'Observatoire 58
2000 Neuchatel
SWITZERLAND

Mr. V. RUTSCHMANN
FSRM
Orangerie 8
2000 Neuchatel
SWITZERLAND

Mr. T. SCHILDKNECHT
Astronomical Insl. Univ. of Berne
Sidlerstrasse 5
3012 Bern
SWITZERLAND

Mr. D. SENNEDOT
DCN Brest
29240 Brest Naval
FRANCE

Mr. N. SAGNA
OBSERVATOIRE CANTONAL
Rue de l'Observatoire 58
2000 Neuchatel
SWITZERLAND

Mr. B. SCHLUETER
Keres Engineering ltd.
P.O. Box 704
CH-2001 Neuchatel
SWITZERLAND

Mr. E.J. SHANKLIN
Hewlett-Packard
van der Hooplaan 241
1180 AZ Amstelveen
NETHERLANDS

Mr. C. SALOMON
Laboratoire de Spectroscopie
Herzienne de l'Ens
24, Rue L'Homond
75005 Paris
FRANCE

Mr. D. SCHNIER
Physikalisch-Technische Bundesanstalt
Bundesallee 100
3300 Braunschweig
GERMANY

Mr. P. SILVERSTRIN
ESA/ESTEC
OPS
Keplerlaan 1
2200 AG Noordwijk
NETHERLANDS

Mr. A. SISWICK
ESA/Estec
XRF
Keplerlaan 1
2200 AG Noordwijk
NETHERLANDS

Mr. L. SVENKERUD
ESA/Estec/XRF
Keplerlaan 1
2200 AG Noordwijk
NETHERLANDS

Mr. A. ULJANOV
KVARZ
Institute of Electronic Measurements
Nizhny Novograd
Moscow
SOVIET UNION

Mr. S.C. SKYNAZY
Oscilloquartz S.A.
P.O.Box
2002 Neuchatel 2
SWITZERLAND

Mr. C.R. TELLIER
LCEP/ENSMM
Route de Gray
25030 Besancon
FRANCE

Mr. P. UYTHOVEN
AT&T Network Systems NL.
P.O.Box 18
1270 AA Huizen
NETHERLANDS

Mr. B. SOLBERG
Norwegian Mapping Authority
Kartverksveien
3500 Honefoss
NORWAY

Mr. G. THEOBALD
C.N.R.S.
LHA Bat. 221
Universite Paris-Sud
91405 Orsay Cedex
FRANCE

Mr. A. VAN DONINCK
Alcatel Bell Telephone
Francis Wellesplein 1
2018 Antwerpen
BELGIUM

Mr. A. SPALLICCI
ESA/ESTEC
MTF
Keplerlaan 1
2200 AG Noordwijk
NETHERLANDS

Mr. P. THOMANN
OBSERVATOIRE CANTONAL
Rue de l'Observatoire 58
2000 Neuchatel
SWITZERLAND

Mr. P. VAN LOOCK
Alcatel Bell Telephone
Francis Wellesplein 1
2018 Antwerpen
BELGIUM

Mr. S.K. STARKER
DLR
Muencherstrasse 1
D-8031 Wessling
GERMANY

Mr. J.M. TORRE
Observatoire de la Cote Dázur
Av. Copernic
06130 Grasse
FRANCE

Mr. F. VERNOTE
OBSERVATOIRE DE BESANCON
41 bis, Avenue de l'Observatoire
B.P. 1615
25010 Besancon Cedex
FRANCE

Mr. P. STEELE
Ball Efratom
Fichtenstrasse 25
8011 Hofolding/Munchen
GERMANY

Mr. P.E. TOSCHEK
Universitat Hamburg
Institut fur Laser-Physik
Jungiusstrasse 9
2000 Hamburg 36
GERMANY

Mr. R.F.C. VESSOT
Smithsonian Astrophysical Observatory
60, Garden street
Cambridge, Massachusetts
UNITED STATES

Mr. J.J. SUTER
The Johns Hopkins University
Applied Physics Laboratory
Johns Hopkins Road
Laurel, MD 20723
UNITED STATES

Mr. R. TOURDE
Observatoire de Paris
LPTF
61, Avenue de l'Observatoire
75014 Paris
FRANCE

Mr. M. VINCENT
Observatoire de Besancon
41 bis Av. de l'Observatoire
BP 1615
F 25010 Besancon cedex
FRANCE

Mr. A. VULCAN
Frequency Electronics Inc.

New York
UNITED STATES

Mr. A. WINTER
Rhode & Schwarz
P.O. Box 801469
D-8000 Munchen 80
GERMANY

Mr. P.K. WALL
PETER WALL ELECTROACOUSTICS
Beaconside, Beacon Edge
Penrith, Cumbria CA11 8BN
UNITED KINGDOM

Mr. N.Y. YAHYABEY
Haut Commissariat a la Recherche
60, rue Cardinal
Verdier, B.E.O., Alger
ALGERIA

Mr. F.L. WALLS
National Inst. Standards Technology
325 Broadway MS 847
Boulder co 80303
UNITED STATES

Mr. N. YANNONI
C.N.R.S.-L.P.M.O.
Rome Laboratory
32, Avenue de l'Observatoire
25000 Besancon
FRANCE

Mr. W. WEIDEMANN
Ball Efratom
3 Parker
Irvine
California 92718-1605
UNITED STATES

Mr. J.S. YERNA
Yetech Applications bv.
Postbus 674
7000 AR Doetinchem
NETHERLANDS

Mr. K. WEISS
Tele Radio Research Institute
Ratuszowa 11
03-450 Warszawa
POLAND

Mr. G. ZAMPETTI
Telecom Solutions
85, West Tasman Drive
San Jose
California 95134-1703
UNITED STATES

Mr. M.A. WHITAKER
British Railways
E601
Macmillan House
Paddington Station CP8, London
UNITED KINGDOM

Mr. A. ZARKA
L.M.C.P.
4, Place Jussieu
75252 Paris Cedex 05
FRANCE

Mr. D. WILLIAMS
ESA/Estec
CSP
Keplerlaan 1
2200 AG Noordwijk
NETHERLANDS

Mr. W. ZINGG
Micro Crystal
Muehlestr. 14, ETA2
CH-2540 Grenchen
SWITZERLAND