

# **EFTF 96**

**Tenth European Frequency and Time Forum**

**CONFERENCE  
PUBLICATION 418**



ORGANISED BY THE INSTITUTION OF ELECTRICAL ENGINEERS

# **10th European Frequency and Time Forum**

**5 - 7 March 1996**

**Organised by**

**The Institution of Electrical Engineers**

**Venue**

**Brighton Centre, Brighton, UK**

## EXECUTIVE COMMITTEE

R J Besson (France)  
J J Gagnepain (France)  
C Audouin (France)

M Ecabert (Switzerland)  
P Kartaschoff (Switzerland)  
B Schlueter (Switzerland)

## SCIENTIFIC COMMITTEE

M J Underhill (UK) - Chairman

J P Aubry (France)  
A Ballato (USA)  
R J Besson (France)  
M Brunet (France)  
L Cutler (USA)  
A De Marchi (Italy)  
F Deyzac (France)  
J Dowsett (UK)  
P Eskelinen (Finland)  
A Godone (Italy)  
M Granveaud (France)  
D Hauden (France)  
K Hilty (Switzerland)  
K Kalliomaki (Finland)  
M Kihara (Japan)  
P W Krempf (Austria)  
K Nakagiri (Japan)  
B W Neubig (Germany)  
L Prost (Switzerland)  
A Reid (UK)  
J Steele (UK)  
A Uljanov (Russia)  
J Vanier (Canada)  
G Winkler (USA)  
G Zampetti (USA)

C Audoin (France)  
A Bauch (Germany)  
R Boloix (Spain)  
G Busca (Switzerland)  
G De Jong (The Netherlands)  
P Defranould (France)  
K Dorenwendt (Germany)  
H Duchaussoy (France)  
S Feltham (The Netherlands)  
E Graf (Switzerland)  
P Hartl (Germany)  
J Heighway (Switzerland)  
C K Hruska (Canada)  
P Kartaschoff (Switzerland)  
D Kirchner (Austria)  
S Leschiutta (Italy)  
H Nau (Germany)  
P W Nieuwejaar (Norway)  
T J Quinn (France)  
W Schluter (Germany)  
P Thomann (Switzerland)  
R Unverdross (Germany)  
F L Walls (USA)  
N Yannonin (USA)  
W Zingg (Switzerland)

## 1996 LOCAL ORGANISING COMMITTEE

E Kentley (Chairman)  
J K A Everard  
A Reid

J Dowsett  
D Henderson  
M J Underhill

## **Author Disclaimer**

"While the author and the publisher believe that the information and guidance given in this work are correct, all parties must rely upon their own skill and judgement when making use of it. Neither the author nor the publisher assume any liability to anyone for any loss or damage caused by any error or omission in the work, whether such error or omission is the result of negligence or any other cause. Any and all such liability is disclaimed."

## **Copyright and Copying**

"All Rights Reserved. No part of this publication may be produced, stored in a retrieval system, or transmitted in any form or by any means - electronic, mechanical, photocopying, recording or otherwise - without the prior written permission of the publisher."

## **Conditions of Acceptance of Advertisements**

The Institution reserves the right to refuse to insert any advertisements (even if ordered and paid for) and/or to make alterations necessary to maintain its standards.

It is not guaranteed that any advertisement will be placed in any specified position or on any specified page unless agreement has been entered into and the agreed surcharge paid.

Every effort will be made to avoid errors but no responsibility will be accepted for any mistakes that may arise in the course of publication of any advertisement. These mistakes may include non-insertion, insertions other than those ordered and errors and omissions within the advertisement.

Notice to cancel any advertisement must be received in writing ten days prior to its next scheduled appearance, otherwise a charge will be made.

No responsibility will be accepted for repetitive errors unless the advertiser's correction has been accepted in respect of that error.

No responsibility will be accepted for loss or damage alleged to arise from errors within advertisement copy, non-appearance of an advertisement or delay in forwarding box number replies.

Advertisers are required to ensure that the content of advertisements conforms with all legislation currently in force affecting such matters. They shall further indemnify the publisher in respect of any claims, costs and expenses that may arise from anything contained within the advertisement and published on their behalf by the Institution of Electrical Engineers.

The placing of an order or contract for insertion of an advertisement in any newspaper or journal published by the Institution of Electrical Engineers whether in writing or by verbal or telephone instructions will be deemed an acceptance of each and all the above conditions.

Published by the Institution of Electrical Engineers, London, ISBN 085296 661 X ISSN 0537-9989.

This publication is copyright under the Berne Convention and the Universal Copyright Convention. All rights reserved. Apart from any copying under the U.K. Copyright, Designs and Patents Act 1988, Part 1, Section 38, whereby a single copy of an article may be supplied, under certain conditions, for the purposes of research or private study, by a library of a class prescribed by The Copyright (Librarians and Archivists) (Copying of Copyright Material) regulations 1989: SI 1989/1212, no part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means without the prior permission of the copyright owners. Permission is, however, not required to copy abstracts of papers or articles on condition that a full reference to the source is shown.

**Multiple copying of the contents of the publication without permission is always illegal.**

© 1996 The Institution of Electrical Engineers.

# Contents

The Institution of Electrical Engineers is not, as a body, responsible for the opinions expressed by individual authors or speakers

## Page No

- Invited Paper**
- 1 'Pulsars as Clocks'  
Professor A G Lyne  
University of Manchester, UK
- SESSION 1: RESONATORS I**
- 8 'High frequency fundamental mode resonators equivalent electrical data dependence on electrode properties and blank surface'  
E Seydel and L Weitzel  
KVG Kristall-Verarbeitung Neckarbischofsheim GmbH, Germany
- 14 'Characteristics of high frequency fundamental rectangular quartz crystal resonators'  
S Dalla Piazza  
Asulab, Switzerland  
B Studer and W Zingg  
Micro Crystal, Switzerland
- 21 'Energy trapping in bulk acoustic wave composite resonator'  
G D Mansfeld  
Russian Academy of Sciences, Russia
- SESSION 2: CLASSICAL FREQUENCY STANDARDS**
- 25 'PTB's primary clock CS3: Uncertainty evaluation and operational experiences'  
A Bauch and T Heindorff  
Physikalisch-Technische Bundesanstalt, Germany
- 26 'State selection in a bipolar gradient magnet for a high C-field cesium beam resonator'  
G Costanzo, A de Marchi and M Repetto  
Politecnico di Torino, Italy
- 30 'Optimization of passive hydrogen standard frequency instability'  
S A Kozlov and V A Logachev  
Institute of Electronic Measurements 'KVARZ', Russia

# Contents

## Page No

### SESSION 3: RESONATORS II

- 31 **'Evaluation of mini-BVA resonators for space applications: Tests program and results'**  
J J Boy, R J Besson, N Franquet and D Thiebaud  
ENSMM / LCEP France  
R Petit  
DGA-CELAR, France  
Ph Baviere and M Brunet  
CNES, France
- 37 **'Comparision of passive and active aging of SC-cut and AT- cut crystals'**  
B W Neubig  
Tele Quartz Group, Germany
- 44 **'Relationship between amplitude and resonant frequency in quartz crystal resonators'**  
E S Ferre-Pikal and F L Walls  
National Institute of Standards and Technology, USA  
J F Garcia Nava  
Centro Nacional de Metrologia, Mexico

### SESSION 4: POSTER SESSION

- 50 **'Durability of spectral lamps with vapors  $^{87}\text{Rb}$  for optically pumped frequency standards'**  
A Gevorkyan, G Smirnova and V Khutorshchikov  
Russian Institue of Radionavigation and Time, Russia
- 55 **'Improvement of a digital automatic cavity tuning system in the active hydrogen maser CH1-75'**  
N A Demidov, V A Logachev and G M Chernov  
Institute of Electronic Measurements 'Kvarz', Russia
- 56 **'Collisional effect with optically pumped atomic clock'**  
M Kajita  
Communications Research Laboratory, Japan
- 62 **'A regularization approach for the determination of the time of flight distributions in the Cesium Beam Standards'**  
A Makdissi and E de Clercq  
Observatoire de Paris, France

# Contents

## Page No

- 66 **'Theoretical description and experimental evaluation of the effect of the interrogation oscillator frequency noise on the stability of a pulsed atomic frequency standard'**  
G Santarelli, Ph Laurent and A Clairon  
Observatoire de Paris, France  
G J Dick and C A Greenhall  
Jet Propulsion Laboratory, USA  
C Audoin,  
CNRS, France
- 72 **'Stabilized laser diode at 729NM for a CA<sup>+</sup> optical frequency standard'**  
B Fermigier, M Houssin and F Plumelle and M Desaintfuscien  
Université Paris-Sud, France
- 75 **'Frequency response of a Fabry-Perot cavity excited by a frequency modulated stochastic laser field'**  
E Bava and F Massari  
Politecnico di Milano, Italy
- 76 **'On the estimating of power spectral density of fractional frequency fluctuations from the sampled phase deviations'**  
M P Dinca  
University of Bucharest, Romania
- 81 **'Clock prediction and cross-sigma'**  
A Lepek  
Tech Projects, Israel
- 86 **'Preliminary result of time scale comparison uncertainty by means of the computer controlled A-724M-02 GLONASS receiver'**  
A Al'shina, N Koshelyaevsky and S Pushkin  
Institute of Metrology for Time and Space, Russia  
P Bogdanov, A Gevorkyan and V Ovchinnikov  
Russian Institute of Radionavigation and Time, Russia  
V Koucherov  
State Time and Standard Frequency Commission  
Yu Medvedev  
ATZUS, Russia
- 87 **'RIRT activity in clock synchronization area by means of meteor bursts radio link'**  
Y G Gouzhva, A G Gevorkyan, I M Bolotov and S N Pluzhnikov  
Russian Institute of Radionavigation and Time, Russia
- 88 **'Transfer of the phase process through the border between two time scales'**  
A Dobrogowski  
Poznan University of Technology, Poland

# Contents

## Page No

- 92 **'A computer controlled multi-reference clock system'**  
P Eskelinen, J Engelberg, A Salpala and M Matola  
Kotka Institute of Technology, Finland
- 96 **'A method for quartz oscillator synchronisation by GPS signal'**  
A Lisowiec and A Czarnecki  
Tele and Radio Research Institute, Poland  
Z Rau  
Telecommunication Institute, Poland
- 100 **'Influence of the ionospheric delays on the subnanosecond accuracy of KU-band two-way satellite time and frequency transfer'**  
P Urich  
Observatoire de Paris, France
- 105 **'GPS disciplined rubidium oscillator'**  
C Little and C Green  
Quartzlock (UK), UK

## SESSION 5: POSTER SESSION

- 111 **'Frequency jumps on BVA and other precision quartz crystal resonators and burst-noise on overtone mode high-frequency quartz crystal resonators'**  
K K Tuladhar and G Jenni  
Oscilloquartz SA, Switzerland
- 116 **'Quartz crystal units with bar piezoelements vibrating in double-shear mode'**  
M E Bogoush and P G Pozdnyakov  
Scientific Research Institute 'Phonon', Russia
- 117 **'Theoretical and experimental investigations of frequency transient process caused in crystal plates under local heating'**  
A N Lepetaev and A V Kosykh  
Omsk State Engineering University, Russia
- 118 **'Influence of loss on the critical frequencies of ceramic resonators'**  
J Kelly and A Safari  
Rutgers University, USA  
A Ballato  
US Army Research Laboratory, USA

# Contents

## Page No

- 123** **'Electromagnetic and mechanical behaviours of whispering gallery mode resonators'**  
O Di Monaco, B Piranda, J S Armand, Y Gruson, I Lajoie and V Giordano  
Laboratoire de Physique et de Métrologie des Oscillateurs du CNRS,  
France  
M Chaubet  
Centre National d'Etudes Spatiales, France  
J Piranda  
Laboratoire de Mécanique Appliquée, R Chaleat, France
- 127** **'Poisson's ratio for piezoceramics'**  
A Ballato  
US Army Research Laboratory, USA
- 132** **'Contribution to the design of the wave filters composed from the elastically coupled two resonators structure'**  
I Mateescu  
Institute of Physics and Technology of Materials, Romania  
A Serbanescu  
Technical Military Academy, Romania  
J Zelenka  
Technical University of Liberec, Czech Republic
- 138** **'Thermoluminescence of swept and doped irradiated quartz crystals'**  
M Smaali and P Zecchini  
LCCM, UFR Sciences, France  
J J Boy  
LCEP/ENSMM, France
- 144** **'Temperature dependence of SH-APM delay-lines on quartz: Design rules for (bio) chemical sensors'**  
C Déjous, D Rebière and J Pistré  
Université Bordeaux I, France  
R Planade  
Centre d'Études du Bouchet, France
- 150** **'Influence of quartz crystal origin on ferroelastic twinning observed by x-ray topography'**  
P L Guzzo and J J Boy  
ENSMM / LCEP, France  
B Capelle  
LMCP, France
- 155** **'SAW resonators frequency trimming by plasma etching'**  
K Weiss and M Tupaj  
Tele and Radio Research Institute, Poland

# Contents

## Page No

- 159 **'Parasitic bulk waves in wide band saw filters designed on 128° rotated Y-cut of LiNbO<sub>3</sub>'**  
P J Naglowski and H Majewska  
Institute of Electronic Materials Technology, Poland
- 165 **'An analysis of BAW to STW mutual transformation arising from periodic corrugation on the surface of a reciprocal coupler'**  
B Dulmet and H Watcheung  
Laboratoire de Chronométrie Electronique et Piézoélectricité, France
- 170 **'Sensitivity of inductances to magnetic fields'**  
R Brendel and G Claudel  
Laboratoire de Physique et Métrologie des Oscillateurs du CNRS, France  
P Bavière  
Centre National d'Etudes Spatiales, France
- 176 **'High-precision IC quartz oscillator'**  
S Kovita and A Myasnikov  
Russian Institute of Radionavigation and Time, Russia  
A Isakov  
St Petersburg Elektrotechnical, Russia  
A Borisenko and O Sokolov  
St Petersburg Technical University, Russia
- 180 **'Effects of crystal mount resonance, under random vibration, on the close to carrier phase noise'**  
N D Hardy  
C-MAC Quartz Crystals Ltd, UK
- 184 **'High performance, crystal oscillators for space applications'**  
S E Morris, F S McClellmont, A J Byrne A L Haston and M G McGovern  
GEC - Marconi Materials Technology Ltd, UK
- 190 **'Development status of the microcomputer compensated crystal oscillator'**  
B E Rose and H Phillips  
Q-Tech Corp, USA
- 196 **'A specific LSIC intended for synthesizer-type DTCXO'**  
A Kosykh and B Ionov  
Omsk State Engineering University, Russia
- 200 **'PRARE one resp. Two years in orbit - performance analysis of the instruments' oscillators onboard ERS-2 and Meteor-3/7'**  
S Bedrich and F Flechtner  
DLR, Germany

# Contents

## Page No

- 206 **'Delay stability measurements made within a two way time transfer system using satellite ranging from several locations'**  
J A Davis  
National Physical Laboratory, UK
- SESSION 6: NEW ATOMIC FREQUENCY STANDARDS**
- 212 **'A microwave frequency standard based on trapped, buffer gas-cooled  $^{171}\text{Yb}^+$  ions-some accuracy issues'**  
P T H Fisk, M J Sellars, M A Lawn and C Coles  
CSIRO National Measurement Laboratory, Australia
- 218 **'The LPTF preliminary accuracy evaluation of cesium fountain frequency standard'**  
A Clairon, S Ghezali, G Santarelli, Ph Laurent, E Simon, M Bahoura, S Weyers and K Szymaniec  
Observatoire de Paris, France  
S Lea  
National Physical Laboratory, UK
- 224 **'Continuous extraction mechanisms from a Cs magneto-optical trap'**  
N Sagna, G Dudle, P Berthoud and P Thomann  
Observatoire Cantonal, Switzerland
- 230 **'The L.H.A 2D magneto-optical trap: On the way to a continuous beam of cold atoms'**  
E Aucouturier, N Dimarcq, P Petit, C Valentin and S Weyers  
Université Paris Sud, France
- SESSION 7: RESONATORS III**
- 234 **'Influence of electrode mass-loading on the electrical equivalent circuit of the trapped-energy at-cut quartz resonators'**  
J Zelenka  
Technical University of Liberec, Czech Republic
- 238 **'Double-mode excitation of quartz crystals without dips of activity on the C and B modes'**  
I V Abramson  
Omsk Research Instrumentation Institute, Russia
- 239 **'SC-cut resonator operating in anharmonic modes with B-mode reduction'**  
R Bourquin B Dulmet and J J Boy  
Ecole Nationale Supérieure de Mécanique et des Microtechniques,  
France

# Contents

## Page No

### SESSION 8: LASER REFERENCES AND STABILISATION

- 244 **'Phase coherent frequency measurement of visible radiation'**  
H Schnatz, B Lipphardt, J Helmcke, F Riehle and G Zinner  
Physikalisch - Technische Bundesanstalt (PTB), Germany
- 245 **'FM noise reduction limits of a laser diode frequency stabilized on an atomic vapor'**  
M Bahoura, Ph Laurent, G Santarelli, S Ghezali and A Clairon  
Observatoire du Paris, France  
N Dimarcq  
Université Paris Sud, France
- 250 **'Frequency measurements on H I and the Kr line at 1.548  $\mu\text{m}$ '**  
D J E Knight  
D K Research, UK  
M Zucco, K I Pharaoh and D A Humphreys  
National Physical Laboratory, UK
- 256 **'High stability laser frequency references at 1.5 $\mu\text{m}$  using molecular lines: State of the art prospects'**  
M de Labachellerie  
Laboratoire de Physique et de Métrologie des Oscillateurs du CNRS, France  
K Nakagawa  
Tokyo Institute of Polytechnics, Japan  
Y Awaji, M Kouroggi and M Ohtsu  
Tokyo Institute of Technology, Japan

### SESSION 9: DIELECTRIC RESONATORS

- 262 **'Frequency stability of  $1 \times 10^{-13}$  in a compensated sapphire oscillator operating above 77K' (Invited)**  
G J Dick, D G Santiago and R T Wang  
California Institute of Technology, USA
- 266 **'Performance of a cryogenic sapphire dielectric resonator'**  
C D Langham and J C Gallop  
National Physical Laboratory, UK
- 270 **'Dielectric loaded HTS resonators as frequency standards and low phase noise oscillators'**  
J C Gallop and C D Langham  
National Physical Laboratory, UK  
F Abbas  
UMIST, UK
- 275 **'Enhanced transposed gain microwave oscillators'**  
M A Page-Jones and J K A Everard  
University of York, UK

# Contents

## Page No

### SESSION 10: SPACE TIME COMPARISON

- 279 **'GPS time transfer using carrier-phase and P-code measurements'**  
G Petit, P Moussay and C Thomas  
Bureau International des Poids et Mesures, France
- 286 **'Timing options and frequency control concepts for meeting user requirements with the second-generation global navigation satellite system (GNSS2)'**  
C C Hodge and J A Davis  
National Physical Laboratory, UK
- 296 **'A GPS parallel common-view experiment'**  
J Cermák  
Academy of Sciences of the Czech Republic, Czech Republic
- 302 **'Use of direct broadcasting satellite signals for time transfer and orbitography'**  
P Guillemot, J Folliard and M Brunet  
CNES, France  
F Meyer and M Vincent  
Observatoire de Besançon, France  
M Granveaud  
LPTF, France  
P Grudler  
OCA, France

### SESSION 11: MATERIALS

- 308 **'Studies of GaPO<sub>4</sub> crystals and resonators' (Invited)**  
A Zarka and B Capelle  
Université P et M Curie, France  
J Detaint  
CNET, France  
D Palmier and E Philippot  
Université de Montpellier II, France  
O V Zvereva  
Institute of Crystallography, Russia
- 313 **'Tolerable orientation errors in determination of the material constants of crystals'**  
C K Hruska  
York University, Canada
- 321 **'Defects in quartz related to interstitial Li and Na ions as revealed by luminescence'**  
A Halperin  
The Hebrew University, Israel

# Contents

## Page No

- 328 'Composite resonator spectroscopy of quartz ferroelastic twinning under uniaxial stress'**  
P L Guzzo and R J Besson  
Ecole Nationale Supérieure de Mécanique et des Microtechniques, France  
G D Mansfeld  
Russian Academy of Sciences, Russia

## **SESSION 12: FUNDAMENTALS AND MEASUREMENTS**

- 334 'On the foundations of the continuum physics including electromagnetic interactions and interfacial properties'**  
N Daher  
LPMO-CNRS, France
- 339 'An advanced ultra-sensitive detector for the measurement of AM and PM noise in microwave components'**  
R A Woode, E N Ivanov and M E Tobar  
The University of Western Australia, Australia
- 344 'Spectral analysis of irregularly spaced timing data: Comparison of several methods'**  
F Vernotte, G Zalamansky and M McHugh  
Observatoire de Besançon, France  
E Lantz  
CNRS, France
- 350 'A quantum Chaologist's view of quartz resonators'**  
M Oxborrow  
University of Oxford, UK

## **SESSION 13: SENSORS**

- 357 'Low dose proton radiation sensitivity of quartz resonators'**  
P E Cash and D A Emmons  
Frequency and Time Systems, USA  
W Stapor  
US Naval Research Laboratories, USA
- 364 'Contribution of a 3D numerical simulation to the design of micro-machined quartz microstructures'**  
C R Tellier, T Messaoudi and T Leblois  
Ecole Nationale Supérieure de Mécanique et des Microtechniques, France
- 370 'The influence of PcCu layer crystalline structure on the parameters of SAW gas sensors'**  
J Hechner and T Wrobel  
Institute of Electronic Materials Technology, Poland

# Contents

## Page No

- 376** **'Measurement of mechanical quantities using quartz sensors'**  
R Grossmann, J Michel, T Sachs and E Schrüfer  
Technical University of Munich, Germany
- SESSION 14: SPACE I**
- 382** **'Comparison of two-way satellite time and frequency transfer and GPS common-view time transfer during the intelsat field trial'**  
J A Davis and K A Bartle  
National Physical Laboratory, UK  
W Lewandowski  
Bureau International des Poids et Mesures, France  
D Kirchner  
Technical University of Graz, Austria  
J A De Young, W Klepczynski and A McKinley  
United States Naval Observatory, USA  
P Hetzel  
Physikalisch-Technische Bundesanstalt, Germany  
T Parker  
National Institute of Standards and Technology, USA  
G De Jong  
Nmi, Van Swinden Laboratorium, The Netherlands  
A Söring  
Deutsche Telekom, Germany  
F Baumont, P Claudon and P Grudler  
Ovservatoire de la Cote d'Azur, France  
H Ressler and R Robnik  
Space Research Institute, Austria  
L Veenstra  
RSI, Comsat World Systems, USA
- 388** **'GLONASS common-view time transfer between North America and Europe and its comparison with GPS'**  
W Lewandowski  
Bureau International des Poids & Mesures, France  
J Danaher  
3S Navigation, USA  
W J Klepczynski  
United States Naval Observatory, USA
- 393** **'Common view' clock synchronization of remote atomic clocks using GPS and PRARE onboard ERS-2'**  
J Hahn  
DLR Institut für Hochfrequenztechnik, Germany  
S Bedrich  
Geo ForshungsZentrum, Germany

# Contents

## Page No

- 399 **'Improving the NIST ionospheric measurement system'**  
M A Weiss, L Gaudron and S Jefferts  
NIST Time and Frequency Division, USA

## SPECIAL SESSION

- 405 **'Specification of long term reliability of frequency control products'**  
D E Brown  
C-MAC Quartz Crystals, UK  
J Clarke  
Attwater & Liell, UK

## SESSION 15: SAW

- 412 **'SAW two resonator filter using synchronous coupling interdigital transducers'**  
W Soluch  
Institute of Electronic Materials Technology, Poland
- 416 **'Frequency stability control of carrier-envelope resonance in a surface acoustic wave delay line oscillator'**  
F Lardet-Vieudrin, G Martin, S Dos Santos, G Marianneau and M Planat  
CNRS, France
- 420 **'Hybrid wide-band VCOs using low-loss leaky SAW ring filters'**  
S Doberstein, A Martynov and V A Malyukhov  
ONIIP, Russia

## SESSION 16: SPACE II - F/T TRANSFER

- 421 **'Frequency and time dissemination using spread-spectrum coded signals transmitted via direct-TV broadcasting satellites'**  
G Bethscheider and J Giannandrea  
SES, Luxembourg  
W Schäfer  
TimeTech GmbH, Germany
- 422 **'Time transfer using TV satellite TC2A'**  
F Meyer and M Vincent  
Observatoire de Besançon, France
- 423 **'One-way frequency and time transfer in the frame of the intelsat two-way experiment'**  
G Brida, F Cordara, V Pettiti and A Godone  
Istituto Elettrotecnico Nazionale Galileo Ferraris, Italy
- 430 **'Time transfer with geodetic GPS receivers using code and phase observations'**  
P Baeriswyl, Th Schildknecht, T Springer and G Beutler  
University of Berne, Switzerland

# Contents

## Page No

### SESSION 17: OSCILLATORS I

- 436 'Low noise oscillators - Theory and design' (*Invited*)  
J K A Everard  
University of York, UK
- 442 'Quartz oscillators: Deriving oscillation condition by symbolic calculus'  
N Ratier and R Brendel  
CNRS, France  
P Guillemot  
Centre National d'Etudes Spatiales, France
- 447 'Phase noise figures comparison in transistor amplifiers of different types'  
R J Besson, M Mourey and S Galliou  
Laboratoire de Chronométrie, Electronique et Piézoélectricité, France  
Ph Guillemot  
CNES, France
- 452 'Spectral improvement of direct digital frequency synthesisers and other frequency sources'  
M J Underhill and M J Blewett  
University of Surrey, UK

### SESSION 18: TERRESTRIAL F/T TRANSFER

- 461 'Two-way time transfer through SDH and SONET systems'  
S Jefferts, M A Weiss, J Levine, S Dilla and T E Parker  
NIST Time and Frequency Division, USA
- 465 'System configuration for standardizing SDH-based time and frequency transfer'  
M Kihara and A Imaoka  
NTT Optical Network Systems Laboratories, Japan
- 471 'Temperature sensitivity of a fibre optic timing link'  
O Buzek  
Academy of Sciences of the Czech Republic, Czech Republic
- 475 'Instrumentation for CS-time transfer to calibration laboratory'  
K Kalliomaki  
University of Oulu, Finland  
T Mansten  
VTT Automation, Finland

# Contents

## Page No

### SESSION 19: OSCILLATORS II

- 476 **'Reduction of phase noise in single transistor oscillators' (Invited)**  
M J Underhill  
University of Surrey, UK
- 491 **'Analysis of medium-term frequency instability of the high stability quartz oscillators'**  
B Kalinowska and B Gniewinska  
Tele and Radio Research Institute, Poland
- 497 **'An Insitu technique for the resolution of aging contributions between quartz resonators and oscillator circuits'**  
G Weaver, W Hanson and T Wickard  
Piezo Crystal Company, USA
- 515 **'Precision, low power, analogue TCXO using a single integrated circuit'**  
H Rokos  
C-MAC Quartz Crystals Ltd, UK

### SESSION 20: TIME SCALES AND DISSEMINATION

- 520 **'Stability and accuracy of international atomic time TAI'**  
C Thomas  
Bureau International des Poids & Mesures, France
- 528 **'Toward  $10^{-15}$  and sub-nanosecond international frequency and time metrology'**  
R Giffard, L Cutler and J A Kusters  
Hewlett Packard Labs, USA  
M Miranian  
USNO, USA  
D W Allan  
Allan's TIME, USA
- 534 **'A novel slave-clock implementation approach for telecommunications network synchronisation'**  
R Urbansky and W Sturm  
AT & T Network Systems, Germany
- 540 **'Time from European time keeping laboratories via telephone modem'**  
P Hetzel and T Polewka  
Physikalisch-Technische Bundesanstalt, Germany  
A Söring and R Metzger  
Forschungs und Technologiezentrum der Telekom AG (FTZ), Germany

# List of Authors

Page No.	Page No.
Abbas, F ..... 270	Daher, N .....334
Abramson, I V ..... 238	Dalla Piazza, S ..... 14
Allan, D W ..... 528	Danaher, J .....388
Al'shina, A ..... 86	Davis, J A .....206, 286, 382
Armand, J S ..... 123	De Clercq, E .....62
Aucouturier, E ..... 230	De Jong, G .....382
Audoin, C ..... 66	de Labacherlerie, M .....256
Awaji, Y ..... 256	de Marchi, A .....26
Baeriswyl, P ..... 430	Demidov, N A .....55
Bahoura, M ..... 218, 245	De Young, J A .....382
Ballato, A ..... 118, 127	Déjous, C .....144
Bartle, K A ..... 382	Desaintfuscien, M ..... 72
Bauch, A ..... 25	Detaint, J .....308
Baumont, F ..... 382	Dilla, S .....461
Bava, E ..... 75	Di Monaco, O .....123
Baviere, P H ..... 31, 170	Dick, G J ..... 66, 262
Bedrich, S ..... 200, 393	Dimarcq, N ..... 230, 245
Berthoud, P ..... 224	Dinca, M P .....76
Bethscheider, G ..... 421	Doberstein, S A .....420
Besson, R J ..... 31, 328, 447	Dobrogowski, A .....88
Beutler, G ..... 430	Dos Santos, S .....416
Blewett, M J ..... 452	Dudle, G .....224
Bogdanov, P ..... 86	Dulmet, B ..... 165, 239
Bogoush, M E ..... 116	Emmons, D A .....357
Bolotov, I M ..... 87	Engelberg, J .....92
Borisenko, A ..... 176	Eskelinen, P .....92
Bourquin, R ..... 239	Everard, J K A ..... 275, 436
Boy, J J ..... 31, 138, 150, 239	Ferminger, B .....72
Brendel, R ..... 170, 442	Ferre-Pikal, E S .....44
Brida, G ..... 423	Fisk, P T H .....212
Brown, D E ..... 405	Fletcher, F .....200
Brunet, M ..... 31, 302	Foillard, J .....302
Buzek, O ..... 471	Franquet, N .....31
Byrne, A J ..... 184	
Capelle, B ..... 150, 308	Galliou, S .....447
Cash, P E ..... 357	Gallop, J C ..... 266, 270
Cermak, J ..... 296	Garcia Nava, J F .....44
Chaubet, M ..... 123	Gaudron, L .....399
Chernov, G M ..... 55	Gevorkyan, A ..... 50, 86, 87
Clarke, J ..... 405	Ghezali, S ..... 218, 245
Clairon, A ..... 66, 218, 245	Giannandrea J .....421
Claudel, G ..... 170	Giffard, R .....528
Claudon, P ..... 382	Giordano, V .....123
Coles, C ..... 212	Gniewinska, B .....491
Costanzo, G ..... 26	Godone, G .....423
Cordara, F ..... 423	Gouzhva, Y G .....87
Cutler, L ..... 528	Granveaud, M .....302
Czarnecki, A ..... 96	Green, C .....105

Greenhall, G A	66, 86	Lepek, A	81
Grossmann, R	376	Lepetaev, A N	117
Grudler, P	302, 382	Levine, J	461
Gruson, Y	123	Lewandowski, W	382, 388
Guillemot, P	302, 442, 447	Little, C	105
Guzzo, P L	150, 328	Lipphardt, B	244
Hahn, J	393	Lisowiec, A	96
Halperin, A	321	Logachev, V A	30, 55
Hanson, W	497	Lyne, A G	1
Hardy, N D	180	Majewska, H	159
Haston, A L	184	Makdissi, A	62
Hechner, J	370	Malyukhov, V A	420
Heindorff, T	25	Mansfeld, G D	21, 328
Helmcke, J	244	Mansten, T	475
Hetzel, P	382, 541	Marianneau, G	416
Hodge, C C	286	Martin, G	416
Houssin, M	72	Martynov, A	420
Hruska, C K	313	Massari, F	75
Humphreys, D A	250	Mateescue, I	132
Imaoka, A	465	Matola, M	92
Ionov, B P	196	McClemont, F S	184
Isakov, A	176	McGovern, M G	184
Ivanov, E N	339	McHugh, M	344
Jefferts, S	399, 461	McKinley, A	382
Jenkins, P		Medvedev, Y	86
Jenni, G	111	Messaoudi, T	364
Kajita, M	56	Metzger, R	540
Kalinowska, B	491	Meyer, F	302, 422
Kalliomaki, K	475	Michel, J	376
Kelly, J	118	Miranian, M	528
Khutorshchikov, V I	50	Morris, S E	184
Kihara, M	465	Mourey, M	447
Kirchner, D	382	Moussay, P	279
Klepczynski, W J	382, 388	Myasnikov, A	176
Knight, D J E	250	Naglowski, P J	159
Koshelyaevsky, N	86	Nakagawa, K	256
Kosykh, A V	117, 196	Neubig, B W	37
Koucherov, V	86	Ohtsu, M	256
Kourogi, M	256	Ovchinnikov, V	86
Kovita, S	176	Oxborrow, M	350
Kozlov, S A	30	Page-Jones, M A	275
Kusters, J	528	Palmier, D	308
Lajoie, I	123	Parker, T	382, 461
Langham, C D	266, 270	Petit, G	279
Lantz, E	344	Petit, P	230
Lardet-Vieudrin, F	416	Petit, R	31
Laurent, Ph	66, 218, 245	Pettiti, V	423
Lawn, M A	212	Pharaoh, K I	250
Lea, S	218	Philippott, E	308
Leblois, T	364	Phillips, H	190
		Pistré, J	144

Piranda, B .....	123	Urbansky, R .....	534
Piranda, J .....	123	Valentin, C .....	230
Planade, R .....	144	Veenstra, L .....	382
Planat, M .....	416	Vernotte, F .....	344
Plumelle, F .....	72	Vincent, M .....	302, 422
Pluzhnikov, S N .....	87	Vernotte, F .....	344
Pushkin, S .....	86	Walls, F L .....	44
Polewka, T .....	540	Wang, R T .....	262
Pozdnyakov, P G .....	116	Watcheung, H .....	165
Ratier, N .....	442	Weaver, G .....	497
Rau, Z .....	96	Weiss, K .....	155
Rebière, D .....	144	Weiss, M A .....	399, 461
Repetto, M .....	26	Weitzel, L .....	8
Ressler, H .....	382	Weyers, S .....	218, 230
Riehle, F .....	244	Wickard, T .....	497
Robnik, R .....	382	Woode, R A .....	339
Rokos, H .....	515	Wrobel, T .....	370
Rose, B E .....	190	Zalamansky, G .....	344
Sachs, T .....	376	Zarka, A .....	308
Safari, A .....	118	Zecchini, P .....	138
Sagna, N .....	224	Zelenka, J .....	132, 234
Salpala, A .....	92	Zingg, W .....	14
Santarelli, G .....	66, 218, 245	Zinner, G .....	244
Santiago, D G .....	262	Zucco, M .....	250
Schäfer, W .....	421	Zvereva, O V .....	308
Schildknecht, T .....	430		
Schnatz, H .....	244		
Schrufer, E .....	376		
Sellars, M J .....	212		
Serbanescu, A .....	132		
Seydel, E .....	8		
Simon, E .....	218		
Smaali, M .....	138		
Smirnova, G M .....	50		
Sokolov, O .....	176		
Soluch, W .....	412		
Söring, A .....	382, 540		
Springer, T .....	430		
Stapor, W .....	357		
Studer, B .....	14		
Sturm, W .....	534		
Szymaniec, K .....	218		
Tellier, C R .....	364		
Thiebaud, D .....	31		
Thomann, P .....	224		
Thomas, C .....	279, 520		
Tobar, M E .....	339		
Tuladhar, K K .....	111		
Tupaj, M .....	155		
Uhrich, P .....	100		
Underhill, M J .....	452, 476		

Dr Paul Vigoureux attended the conference dinner as guest of the Local Organising Committee, in recognition of a long and distinguished career in frequency metrology. Although Dr Vigoureux retired officially in 1974 following some five decades of work in metrology, he continues to contribute to various research projects at NPL. His expression of thanks is reproduced here.

“Mr Chairman, Members of the Organising Committee, Ladies and Gentlemen, I thank the Organising Committee for the great honour done to me by asking me to be here with you this evening; my thanks go also to the Professor Underhill for his kind remarks about my scientific activities. When I joined the National Physical Laboratory in 1925, my activity became precisely your own, for I worked under D W Dye, a brilliant experimenter who, like Cade on the other side of the Atlantic, was a pioneer in the application of quartz crystals to the measurement of frequency and time. After Dye’s untimely death in 1932, his work on time and frequency was continued by Essen with great success that I think you all know.

Your conference is being held, I gather, under the auspices of the Institution of Electrical Engineers. I have appreciated being associated with the Institution ever since I joined it as a student in 1923. I have enjoyed attending lectures at Savoy Place, enjoyed serving on committees at the Institution, enjoyed membership of the Dining Club, which I still hold. Now that I have resigned from the committees, I still enjoy an occasion like this evening, and when the IEE News arrives, I enjoy opening it to read Dipole’s article. My sincere thanks to the Institution, Mr Chairman, as well as to you all.

Paul Vigoureux

Faint, illegible text, possibly bleed-through from the reverse side of the page.

# PULSARS AS CLOCKS

A. G. Lyne

University of Manchester, Nuffield Radio Astronomy Laboratories,  
Jodrell Bank, Macclesfield, Cheshire SK11 9DL, UK.

## 1 INTRODUCTION

Radio pulsars are neutron stars, highly condensed bodies formed in the gravitational collapse of the cores of massive stars, following exhaustion of fuels for nuclear burning. Enormous amounts of gravitational energy are released in the collapse, resulting in the bulk of the star being blasted into space with velocities of perhaps 10,000 km/sec. These violent events are witnessed as supernova explosions, such as that which formed the Crab Nebula and its central pulsar in 1054 AD. The collapsed body is a highly compact object, a neutron star, predicted by Baade & Zwicky (1934), and first observed at Cambridge in 1967 (Hewish et al. 1968). Consisting mostly of neutrons, they have masses of about 1.4 times the mass of the Sun, but confined in a star of only 10–15 km in radius. With mean densities of  $10^{12}$  gm cm<sup>-3</sup>, their central densities are perhaps 1000 greater, far exceeding the density of matter in the atomic nucleus. During the collapse, angular momentum is conserved and the neutron star ends up rotating perhaps a hundred times a second, the enormous gravitational field keeping the star from flying apart through centrifugal forces.

We would not normally expect to be able to detect such small, dead bodies, except for the fact that they are mostly endowed with intense, dipolar magnetic fields, which were probably formed in the squeezing of the magnetic field of the progenitor star during the collapse and maintained by the superconducting properties of the few electrons and protons inside the star. The rotation of the star and electromagnetic induction result in large electric fields along the magnetic field lines above the magnetic poles. These fields accelerate particles to highly relativistic velocities, causing them to emit beams

of electromagnetic radiation, roughly out of the magnetic poles. As the star rotates, these beams of radiation cross the line-of-sight to the Earth and we receive a pulse of energy, much in the way that a lighthouse beam is observed to produce pulses of light.

Such an object is known as a pulsar, of which over 700 are now known, having rotation periods between 1.5 milliseconds and 5 seconds. Being massive, inert flywheels, spinning in essentially empty space, and whose rotation can be observed, it is perhaps not surprising that they are potentially superb clocks. In this review, we describe the technical means by which this promise can be realised in practice and the limitations to their precision as clocks. This is followed by examples of the use of these remarkable astrophysical clocks for studying general relativity and other areas of fundamental physics. For a more general review of pulsars and their properties, the reader should refer to the books by Manchester & Taylor (1977) or Lyne & Smith (1990) on these remarkable objects, the most condensed state of matter outside a black hole.

## 2 TECHNICAL ASPECTS

There are a number of technical problems to be surmounted before accurate arrival times of pulses can be obtained from pulsars.

Firstly the pulses have passed through the interstellar medium, which causes them to be dispersed, so that low radio frequencies are delayed relative to high frequencies, resulting in a broadening of the pulses when observed with a wide receiver bandwidth. Observing in many narrow bandwidth frequency channels permits the effects of such dispersion to be removed. Secondly, since the pulses are so weak after travelling for

many hundreds or thousands of years, they have to be observed with large radiotelescopes and integrated, typically for intervals of a few minutes, in order to achieve a satisfactory signal-to-noise ratio. To obtain an arrival time, the observed pulse profile is then convolved with a template, a noiseless “ideal” profile. The time measured for the pulse is relative to the observatory clock, usually based on a rubidium frequency standard. This is calibrated continuously against UTC, which is transferred by time-averaged measurements using the GPS navigation system to an accuracy of about 100ns.

These pulse arrival times are measured on the surface of the Earth which is, of course, by no means in an inertial frame. The nearest point approximating to an inertial frame is the barycentre of the Solar System, which is moving uniformly through space, apart from the small acceleration due to its motion around the Milky Way. The barycentric arrival times,  $t_b$ , are then obtained by correcting the observed times,  $t_o$ , for 3 main effects:

$$t_b = t_o + \frac{\mathbf{r}_o \cdot \hat{\mathbf{s}}}{c} + t_r - t_s, \quad (1)$$

where  $\mathbf{r}_o$  is the position of the observatory relative to the Solar System barycentre, and  $\hat{\mathbf{s}}$  is a unit vector in the direction of the pulsar. The first correction term is just the light travel time across the Solar System, amounting to up to 500 seconds. The last two terms arise from general relativistic effects:  $t_r$  is a clock correction due to the fact that a clock on the Earth runs at a variable rate compared with one in an inertial frame because it is moving up and down the gravitational potential well of the Sun, since the orbit is slightly eccentric.  $t_s$  is the Shapiro delay which arises from the passage of the radiation through the gravitational potential well of the Sun. These latter terms amount to a few milliseconds in total.

The final process is to compare these barycentric measured times of arrival (TOAs) with a model for the pulsar position and its rotation. The differences between the TOAs and the model times are called timing residuals, and the model is then adjusted to minimise these residuals. For example, the top panel of Fig. 1 displays the timing residuals for a millisecond pulsar, PSR J0437–4715, showing a sine wave of one year period and increasing in amplitude due to

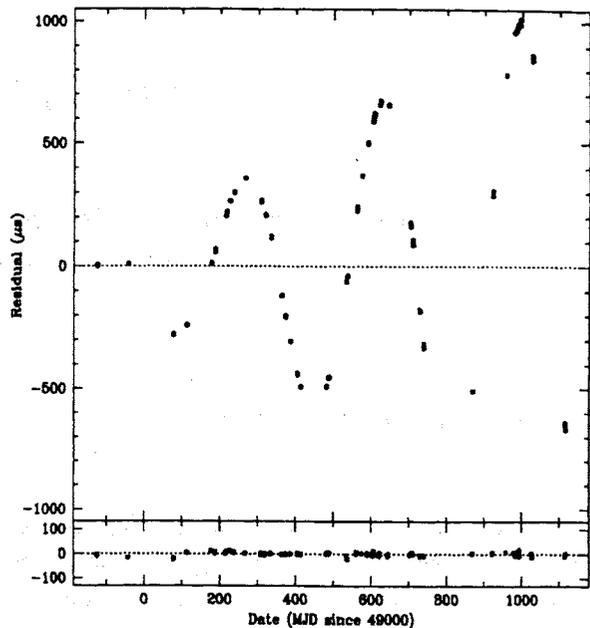


Figure 1: Timing residuals for PSR J0437–4715 relative to a simple model. The top display shows an annual sinusoid of increasing amplitude, the effect of the motion of the pulsar across the sky at a rate of 0.14 arcsec/year. The lower diagram shows the residuals after adjusting the model to include the effects of the pulsar’s motion.

a movement of the pulsar across the sky at a rate of about 0.14"/year. After fitting a new model which includes this motion to these data, the residuals shown in the lower panel of Fig. 1 are very small. Such measurements provide extremely precise determinations of position, frequently with errors of less than a milliarcsecond, and often far superior to optical astrometric methods.

### 3 REGULAR SLOW-DOWN

Pulsars slow down through the loss of kinetic energy in the form of low frequency electromagnetic waves or high energy particles. In a general slowdown, the braking torque is proportional to the frequency derivative  $\dot{\nu}$  and depends upon the rotation rate  $\nu$ , leading to the spin-down equation  $\dot{\nu} = -K\nu^n$ , where  $n$  is the braking index and depends upon the physics of the mechanism by which rotational energy is lost. If the energy is lost through dipolar electromagnetic radiation at the rotation frequency of the pulsar or by

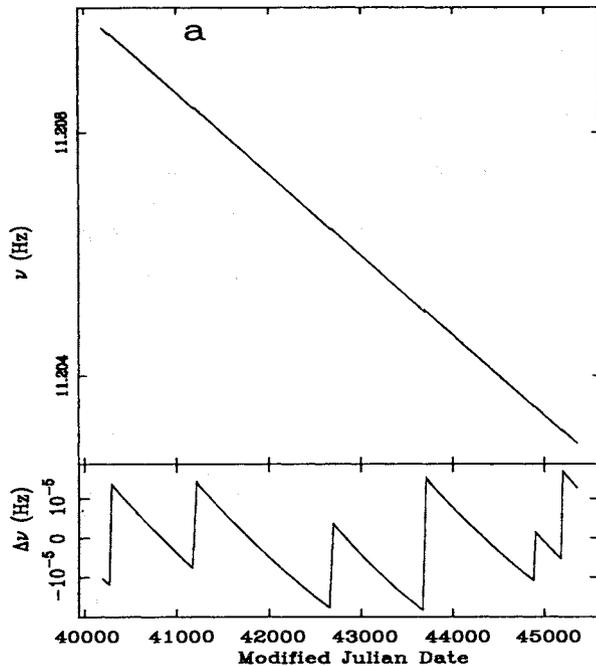


Figure 2: The slowdown of the Vela pulsar over 14 years. 6 glitches can just be seen in the top panel which is dominated by the secular slowdown. The lower panel shows the same, after subtraction of a slope, corresponding to a constant value of  $\dot{\nu}$ .

particles accelerated out of the magnetic poles, the constant  $K$  is proportional to the square of the magnetic field of the pulsar and  $n$  has an expected value of 3 (Pacini 1968; Goldreich & Julian 1969).

There are two main types of pulsar. The young, “normal” pulsars have rotation periods between about 0.1 and 2 seconds, have surface magnetic fields of around  $10^{12}$  Gauss ( $10^8$  T) and spin-down and fade on timescales of a few million years. On the other hand, the old, “millisecond” pulsars have rotation periods between about 1.5 and 20 milliseconds, surface magnetic fields of around  $10^8$  Gauss and spin-down only very slowly, on timescales of  $10^9 - 10^{10}$  years, essentially the age of the Universe! The latter pulsars are thought to be “dead” normal pulsars which have been spun-up to millisecond periods by accretion of matter from an orbiting companion star (see Bhattacharya & van den Heuvel (1991) for a review of the formation of millisecond pulsars).

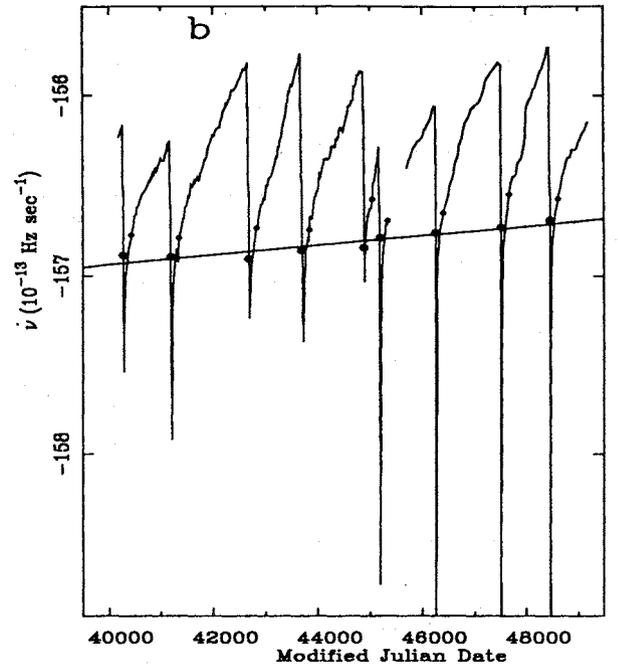


Figure 3: The slowdown of the Vela pulsar. The run of the value of  $\dot{\nu}$  over 24 years, showing the 9 glitches and quasi-exponential relaxation following each.

## 4 GLITCHES AND TIMING NOISE

The normal slowdown described above is steady and predictable. However some pulsars show erratic behaviour of two types: glitches and timing noise, which occur predominantly in young objects. Both are apparently associated with the irregular transfer of angular momentum from a fluid component in the interior (Lyne 1992) as the pulsar slows down, and we discuss each of these in turn.

Glitches are seen as sudden increases in the rotation rate,  $\nu$ , usually followed by an exponential recovery or relaxation back towards the pre-glitch frequency. For example, the slowdown of the Vela pulsar over a 14-year interval is shown in Fig. 2. During this time, 6 glitches can barely be seen in the top diagram against the normal slowdown but, after a slope is removed, they are clear, each step corresponding to a fractional increase in rotation rate of  $\Delta\nu/\nu \sim 2 \times 10^{-6}$ . The relaxations can be most clearly seen in Fig. 3 which shows the variation in  $\dot{\nu}$ , over a 25-year period. These relaxations give clear evidence for a loosely coupled neutron superfluid inside the star.

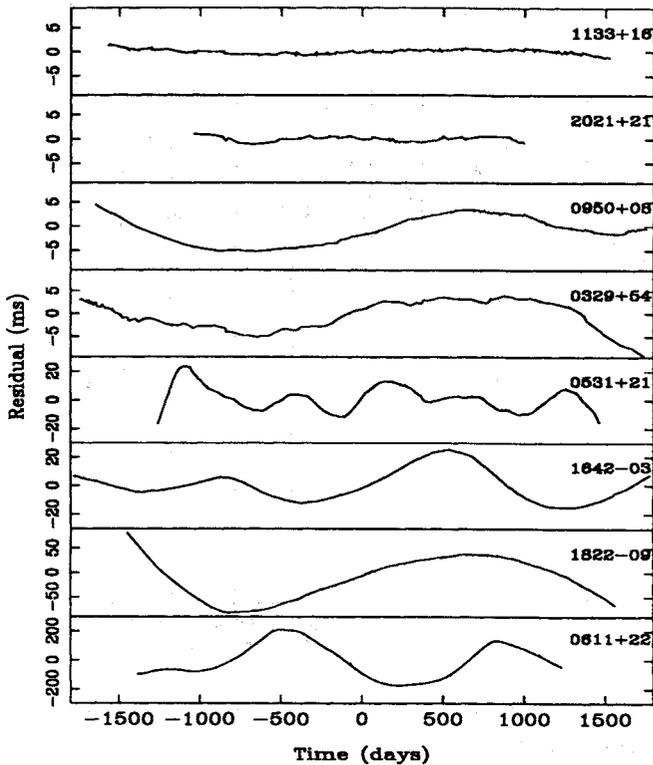


Figure 4: Examples of timing noise in 8 pulsars over about a 10 year period, showing increasing amounts of activity from the top to bottom.

Glitches are rare – only about 20 out of over 700 pulsars have suffered a total of 45 glitches and until recently the study of glitches has been limited by their small number, mainly because they occur most frequently in young pulsars which slow-down very rapidly.

The first proposed cause of these glitches was a starquake which might arise from changing ellipticity of the solid crust of the neutron star as it slows down (Ruderman 1969). The oblateness of an equilibrium spheroid will decrease as the rotation rate decreases. Stresses build up in the rigid crust as the departure from the equilibrium shape increases, until it cracks and assumes a shape closer to the equilibrium spheroid. The moment of inertia  $I$  decreases and conservation of angular momentum results in a spin-up given approximately in terms of the change in oblateness:  $\Delta\epsilon = \Delta I/I = -\Delta\nu/\nu$ .

While this provided a satisfactory explanation for the rather small and infrequent glitches in the Crab pulsar, it soon became clear that the magnitude and frequency of the glitches in the Vela pulsar were too great to be sustained by this mechanism (Pines, Shaham & Ruderman 1972) – it is too spherical. More likely,

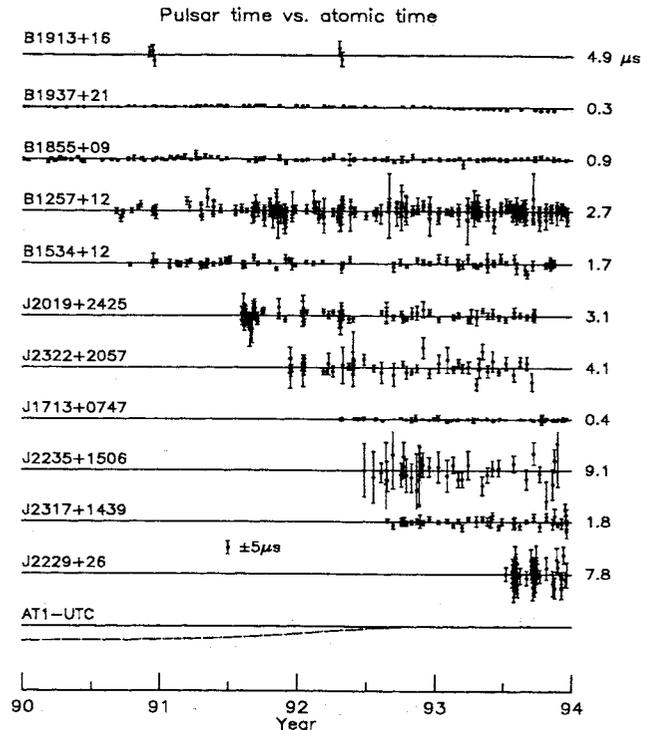


Figure 5: Examples of timing residuals from 11 millisecond pulsars observed at Arecibo Observatory over period of a few years (Camilo & Taylor, private communication). The vertical scales are the same for all plots, the separation between the vertical lines corresponding to about  $20 \mu\text{sec}$ .

it seems that in this case the glitches result from the catastrophic unpinning of superfluid neutron vortices from the nuclei of the solid crust (Baym, Pethick & Pines 1969; Anderson & Itoh 1975). The study of these glitches allows a sort of rotational seismology in order to investigate the internal structure of these neutron stars.

Timing noise is characterised by a continuous, unpredictable, phase wandering of the pulses relative to a simple slow down model. It is seen most prominently in the Crab and other pulsars with large period derivatives (Cordes & Helfand 1980). Fig. 4 shows examples of timing noise in a few young pulsars, typically amounting to a few 10s of milliseconds over a few years. On the other hand, Fig. 5 shows the timing residuals for a number of millisecond pulsars, in which the scale is measured in microseconds. In particular, the first millisecond pulsar to be discovered, PSR B1937+21, has arrival time measurement errors of only about 100 ns, although

systematic trends due to timing noise can be discerned in the diagram. This pulsar has a large period derivative compared with most millisecond pulsars and timing noise is judged to be responsible for wandering of order  $2 \mu\text{sec}$  over an interval of about 10 years. Some of the more recently discovered objects promise to be even more stable.

The amount of this timing noise can be quantified by measuring the magnitude of the second derivative of the period ( $P = 1/\nu$ ) in the timing residuals relative to a simple slow-down model as described above. Fig. 6 shows this plotted as a function of the period derivative for a number of pulsars (Martin & Lyne 1996). The millisecond pulsars with very small period derivatives are found to be very stable, as indeed might be expected from the extrapolation of the main trend in this diagram to their position on the extreme left of the figure (Arzoumanian et al. 1994).

The most natural explanation for timing noise is that it is due to an irregular flow of angular momentum from the fluid component in the interior of the star, which was revealed by the studies of glitches.

## 5 ACCURACY AS CLOCKS

While young pulsars display much glitch and timing noise activity, we have seen that the older, millisecond pulsars, which are slowing down only very gradually, are remarkably stable rotators. While the objects with the most timing noise have stabilities  $\Delta\nu/\nu \sim 10^{-10}$ , on timescales of a few years, the millisecond pulsars have  $\Delta\nu/\nu \leq 10^{-14}$ .

## 6 BINARY SYSTEMS

Most pulsars are solitary objects, having lost any companion stars in the violence of their formation supernova explosions. Nevertheless, several dozen of the 700 known pulsars are in binary systems whose Keplerian orbital parameters can be very precisely determined using the timing techniques described above. The classical Keplerian parameters provide some information about the mass of the companions of these systems through the mass function, a function of

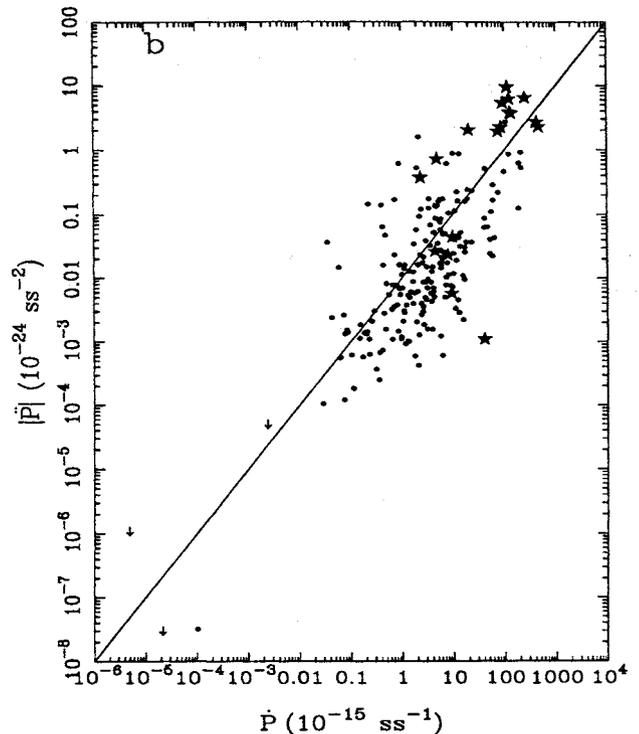


Figure 6: The magnitude of  $\ddot{P}$ , a measure of the amount of timing noise, plotted against  $\dot{P}$ , showing high level of timing noise in young, rapidly braking pulsars. Pulsars which glitch are shown by a star. The line has a slope of unity.

only two observables, the projected semi-major axis of the pulsar orbit  $a_p \sin i$ , and the orbital period  $P_{\text{orb}}$ :

$$f(m_p, m_c) = \frac{4\pi^2 (a_p \sin i)^3}{G P_{\text{orb}}^2} = \frac{(m_c \sin i)^3}{(m_p + m_c)^2}, \quad (2)$$

where  $m_p$  and  $m_c$  refer to the masses of the pulsar and companion respectively and  $i$  is the unknown angle of inclination of the orbit to the plane of the sky. This places very loose constraints upon the 3 unknowns  $m_p$ ,  $m_c$ , and  $i$  which can not be determined uniquely without other information.

### 6.1 General Relativity

However in 1975, Russ Hulse and Joe Taylor discovered the first pulsar in a binary system, PSR B1913+16, in an eccentric 7.8-hour orbit. Moving at speeds up to  $300 \text{ km s}^{-1}$ , the whole orbit was no larger than the size of the Sun and the absence of any occultation by the companion star showed that it too was compact and proba-

bly another neutron star. We thus had a superclock moving up and down the gravitational potential well of another point-like mass at mildly relativistic velocity - an excellent laboratory for testing general relativity.

The finite travel time of gravity waves across the system causes the gravitational forces between the two stars to be non-radial and the major axis of the orbit to precess, at a speed  $\dot{\Omega}$  which is easily measured and which is determined by the total mass of the system,  $m_p + m_c$ . Finally, measurement of the gravitational red-shift and time-dilation give a third observable,  $\gamma$ , which is also a function of  $m_p$  and  $m_c$ . The 3 equations can be uniquely solved (Damour & Taylor 1991) to provide values of  $m_p = 1.4411(7) M_\odot$ ,  $m_c = 1.3874(7) M_\odot$  and  $i = 43(1)$  deg.

Such experiments have been carried out on 3 such systems, providing six neutron star mass estimates which all lie in the range  $1.4 \pm 0.1 M_\odot$ . For most systems which have very small eccentricity, such measurements can not be carried out. However, in the circumstance when the inclination is such that the radiation from the pulsar passes very close to the companion star, the Shapiro delay can be measured. This has been carried out for PSR B1855+09 and its companion (Ryba & Taylor 1991) again permitting the system to be completely solved.

## 6.2 Gravitational Radiation

In all these cases in which the configuration and motion of the component stars is completely specified, general relativity theory can be used to predict the energy that will be lost through gravitational radiation, resulting in a spiralling in of the orbit and a decreasing orbital period. This has been observed in PSR B1913+16 and agrees to within 0.5% with the theoretical value, once the accelerations of the Earth and pulsar system in the Galactic gravitational potential are taken into account (Damour & Taylor 1991).

This system provides a wonderful test of general relativity and demonstrates, for the first time, the existence of gravitational radiation. For the discovery of this system Hulse and Taylor were awarded the Nobel Prize for Physics in 1993.

## 7 MASS DISTRIBUTION IN GLOBULAR CLUSTERS

In 1987, the first pulsar, PSR B1821-24 was discovered in a globular cluster, M28, a dense gravitationally bound swarm of about a million very old stars. Since then about 3 dozen pulsars have been found in such clusters, most of them millisecond pulsars, probably old neutron stars spun-up in the many binary systems formed in the frequent collisions that occur in the dense stellar cauldrons of the cores of these clusters.

These millisecond pulsars are accelerated in the gravitational potential field of their parent cluster, and the component of acceleration along the line-of-sight is registered as a change in the period derivative of the pulsar. For those pulsars near or in the cores of clusters, this dynamical contribution to the derivative dominates the intrinsic slow-down derivative and tells us about the size of the gravitational field. With several pulsars in a single cluster, it is possible to study the mass distribution and hence "weigh" the cluster. Such measurements show that there is a significant amount of "dark matter", material that can not be observed through its normal emission of radiation.

## 8 THE DETECTION OF PLANETARY SYSTEMS

The greatest use of these clocks in astronomy lies in their use for detecting very small motions. For instance, a movement of the pulsar by only 300 m will be detected as a  $1 \mu\text{s}$  change in the timing residuals. In this way, Alex Wolszczan and Dale Frail (1992) detected 3 bodies in a planetary system around the millisecond pulsar PSR B1257+12. The pulsar moves about 600 km due to the gravitational pull of the orbiting planets and results in timing residuals of a few milliseconds relative to a simple slow-down model. The planets are small, being only a few times the mass of the Earth and lying about half the Earth's orbital radius from the pulsar. Small perturbations in the orbits due to their mutual gravitational interaction have been detected, confirming their presence and making them the first known extra-solar planets.

## 9 CONCLUSION

While most pulsars are demonstrably imperfect clocks, the old, millisecond pulsars have extremely good rotational stability and provide wonderful tools for astrophysical studies of binary systems and fundamental physics, a few examples of which we have seen here. Furthermore, with stabilities of better than  $10^{-14}$  at present and with measurement precision expected to increase in the future, they are approaching the stability of the best laboratory clocks. For these reasons, intense efforts have been made by the large radio telescopes at Jodrell Bank, Arecibo and Parkes over the past 4 years to find more of these millisecond pulsars of which there are now more than 60 known. It is not clear whether the limit to their use as clocks will be placed by their intrinsic rotational stability or by the sensitivity of our instrumentation.

## REFERENCES

- Anderson P. W., Itoh N., 1975, *Nature*, 256, 25  
 Arzoumanian Z., Nice D. J., Taylor J. H.,  
 Thorsett S. E., 1994, *ApJ*, 422, 671  
 Baade W., Zwicky F., 1934, *Proc. Nat. Acad. Sci.*,  
 20, 254  
 Baym G., Pethick C., Pines D., 1969, *Nature*, 224,  
 673  
 Bhattacharya D., van den Heuvel E. P. J., 1991,  
*Phys. Rep.*, 203, 1  
 Cordes J. M., Helfand D. J., 1980, *ApJ*, 239, 640  
 Damour T., Taylor J. H., 1991, *ApJ*, 366, 501  
 Goldreich P., Julian W. H., 1969, *ApJ*, 157, 869  
 Harrison E. R., Tademaru E., 1975, *ApJ*, 201, 447  
 Hewish A., Bell S. J., Pilkington J. D. H.,  
 Scott P. F., Collins R. A., 1968, *Nature*, 217,  
 709  
 Lyne A. G., Smith F. G., 1990, *Pulsar Astronomy*.  
 Cambridge University Press  
 Lyne A. G., 1992, *Philos. Trans. Roy. Soc. London*  
 A, 341, 29  
 Manchester R. N., Taylor J. H., 1977, *Pulsars*.  
 Freeman, San Francisco  
 Martin C. E., Lyne A. G., 1996, *MNRAS*, In  
 preparation  
 Pacini F., 1968, *Nature*, 219, 145  
 Pines D., Shaham J., Ruderman M. A., 1972,  
*Nature Phys. Sci.*, 237, 83  
 Ruderman M., 1969, *Nature*, 223, 597  
 Ryba M. F., Taylor J. H., 1991, *ApJ*, 371, 739  
 Wolszczan A., Frail D. A., 1992, *Nature*, 355, 145

# HIGH FREQUENCY FUNDAMENTAL MODE RESONATORS EQUIVALENT ELECTRICAL DATA DEPENDENCE ON ELECTRODE PROPERTIES AND BLANK SURFACE

Eberhard Seydel , Lothar Weitzel

KVG Kristall-Verarbeitung Neckarbischofsheim GmbH  
R&D quartz crystals  
D-74922 Neckarbischofsheim , P.O.Box 61  
Tel. 49/7263/648-82 FAX 49/7263/6196

## INTRODUCTION

High frequency fundamental mode crystals in the frequency range from 40 to 200 MHz show different electrical properties compared to conventional overtone quartz crystals in the same frequency range. They show lower electrical Q-values, different pullabilities and inharmonic modes as result of variations of the blank-electrode thickness ratio .

We present the behaviour of inharmonic modes and  $C_0/C_1$ -ratios under various conditions using FEM methods, but also compare our results with experimental data of planoparallel crystals and inverted MESA-structured crystals, respectively.

## METHOD OF MEASUREMENT

The equivalent electrical data of these high frequency crystals were measured according to IEC 444-5 with PI-Network and calculated  $C_0$ -compensation. In the frequency range up to 1.1GHz the electrical properties of the mounting system and enclosure have large influence on the parameters of the whole crystal unit. The capacitance and the inductance of the mounting system shifts the electrical characteristic. For the future and the frequency range over 400 MHz, a ceramic flat pack is the best solution for minimizing the influences of the mounting system and enclosure. The equivalent circuit will be shifted (fig.1). For high frequency blanks and especially for the overtone vibration the quotient of equivalent film thickness to wavelength approaches approximately one percent (for standard crystals this value is about 0.1%). This great influence of the metallic film is equivalent to an Q-value-shunt. The  $C_{0k}$  for compensation is not only imaginary, the real part has significant values (see table 1).

## THEORETICAL BACKGROUND AND CALCULATION

The calculation of vibration-modes of quartz crystal resonators with different dimensions and surface contours is based on the same method as in publications from Seydel et al [2,3,4].

In this way it is possible to calculate the vibrations of crystal resonators of various shapes, of different cut angles, with different electrode dimensions and different temperatures.

With this program variations in the amplitude distribution of each mode of vibration can be simulated as a function of the blank and electrode geometry, temperature, and cut angle. In this way predictions of crystal behaviour can be made.

The distribution of amplitude across the resonator allows us to determine the effects of the edge and the damping of vibration modes (e.g. inharmonic modes).

The frequency shift of modes , the ratio of amplitudes and the ratio of  $C_0/C_1$  show dependences on the geometrical parameters, film thickness and surface roughness..

## THEORETICAL RESULTS

We have calculated the dependences of the frequency shift of modes, the ratio of amplitudes and relative ratio  $C_0/C_1$  versus the film thickness and the surface roughness at different fundamental frequencies. The dependence of wavelength on the frequency is shown in figure 2. If the surface roughness or the equivalent thickness of the metallic film becomes a value in the order of 0.5 % or more of the wavelength then the behaviour of the quartz crystal will be disturbed.

The  $C_0/C_1$ -ratio as a function of the film thickness and of the surface roughness for different blank geometry is shown in fig. 3 and 4 for a frequency of approximately 60 MHz. The  $C_0/C_1$ -ratio increases with higher values of film thickness or surface roughness and also with higher frequency (fig.5). This is equivalent to decreasing pullability.

Figures 6,7,8 show the dependencies of amplitude, inharmonic mode frequencies versus the film thickness and surface roughness. The suppression of inharmonic modes decreases with higher frequency at the same film thickness and also the frequency offset from the main mode decreases. The dependence of the inharmonic mode amplitudes and frequency shift is shown in figure 9,10,11,12. Standard blank and inverted MESA blank exhibit a different behaviour. This is caused by the different geometry, (diameter and boundary conditions). Figure 19 and 20 show the theoretical amplitude deviations. Thick films cause a step function in the amplitude gradient at the film edge. The roughness causes another type of steps in the amplitude gradient curve at scattering points which correspond to the surface roughness.

Another critical point is the dependence of the zero cut angle on the film thickness. Figure 14 demonstrates the approximated values of zero angle variation at the same metal film thickness versus frequency. The zero cut angle differs about 10 minutes from gold to aluminium at 200 MHz fundamental frequency.

### EXPERIMENTAL RESULTS

We have measured many crystals with different frequencies in the range from 50 MHz to 1.1 GHz at fundamental mode and overtone modes respectively. The measured values are summarized in table 1. The measured values of  $C_0/C_1$  ratio are shown in figure 13. The measured values approximately correlate to the calculated values.

A new effect at high frequency fundamental mode resonators is a new type of dips on the frequency temperature curve. Dips are the deviation of the measured frequency-temperature-values from the values calculated by the method of least-square. The dip in the frequency temperature curve is accompanied by an increase in R1. The temperature range in which the effect occurs is less than 20 degrees. Active dips are caused by a mechanical coupling of the thickness-shear-mode with other vibration forms (e.g. face-shear or flexure). The magnitude depends on drive level .

High frequency dips which are caused by the reflexion or scattering of the thickness-shear mode on the surface roughness . These are not dependent on the drive level. (see fig 15, 16, 17, 18)

### CONCLUSION

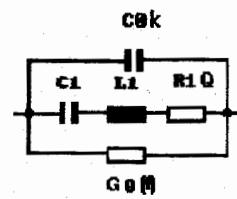
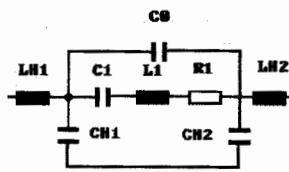
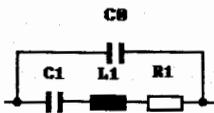
It is possible to measure all electrical data of high frequency quartz crystals with a measuring system according to IEC 444-5 and full complex compensation of  $C_0$ -effects. The stability of the measured values is better than 0.5 ppm in frequency, 0.5% in R1 and 1% in motional capacitance.

The theoretical calculations make it possible to determine the direction of variation of equivalent parameters in the high frequency range.

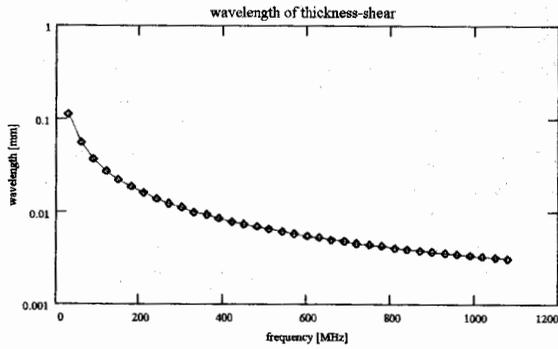
The quality of high frequency blanks is defined by the equivalent metal film thickness, the film material and the surface roughness of the blank. These values determine the equivalent electrical data and the temperature frequency curve.

### LITERATURE

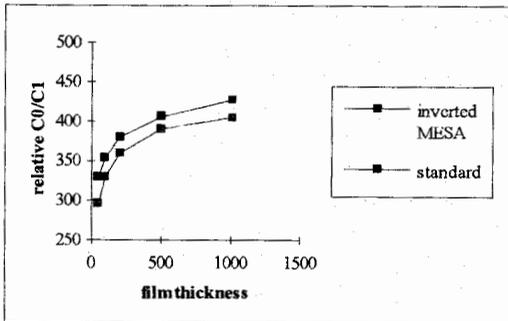
1. Beaver W.D.,1973, DESIGN EQUATIONS FOR BI- AND -PLANO-CONVEX AT-CUT RESONATORS, AFCS,27, 11
2. Seydel E., Priess U.,1994, OPTIMIZATION OF QUARTZ CRYSTAL DESIGN FOR TCXO- AND VCXO-APPLICATIONS; EFTF 1994
3. Seydel E., Priess U.,1994; QUARTZ CRYSTAL FOR MICROBALANCE; PIEZO'94
4. Seydel E., Landmann A.,1995, THE INFLUENCE OF SURFACE STRUCTURE ON AMPLITUDE DEVIATION AND ELECTRICAL PARAMETERS OF QUARTZ CRYSTALS; EFTF 1995



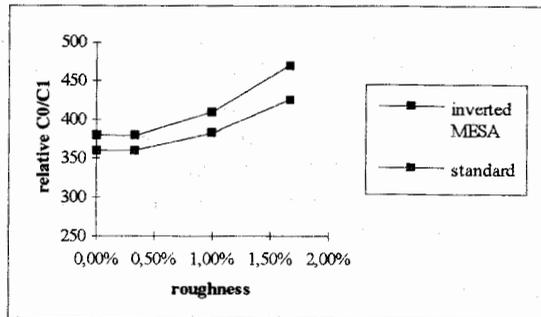
quartz-crystal      quartz-crystal with mounting system      high frequency quartz crystal  
**figure 1 different equivalent circuits**



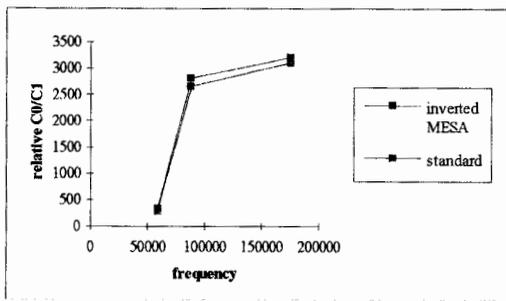
**figure 2 wavelength of thickness-shear vibration**



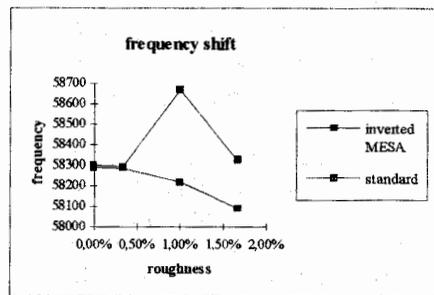
**figure 3 C<sub>0</sub>/C<sub>1</sub>-ratio versus film thickness**



**figure 4 C<sub>0</sub>/C<sub>1</sub>-ratio versus roughness**



**figure 5 C<sub>0</sub>/C<sub>1</sub>-ratio versus frequency**



**figure 6 frequency shift versus roughness**

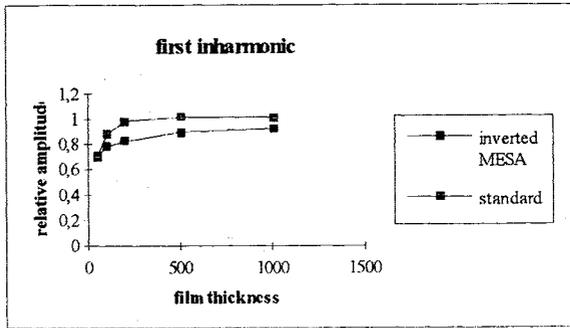


figure 7 relative amplitude of first inharmonic versus film thickness

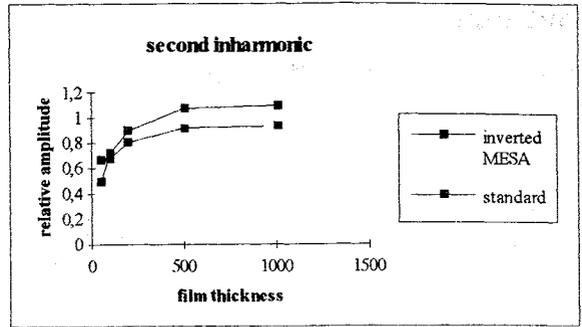


figure 8 relative amplitude of second inharmonic versus film thickness

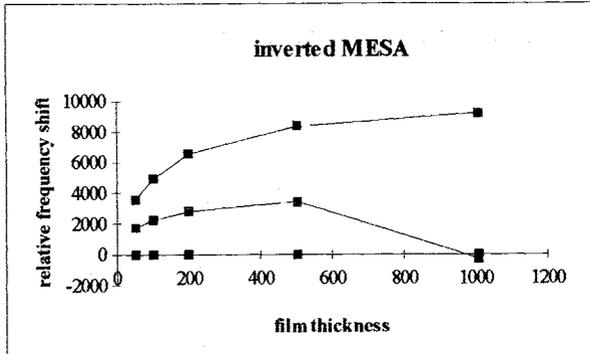


figure 9 relative frequency shift of inharmonic modes versus film thickness

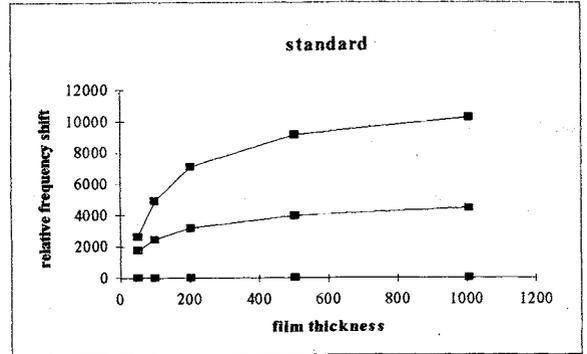


figure 10 relative frequency shift of inharmonic modes versus film thickness

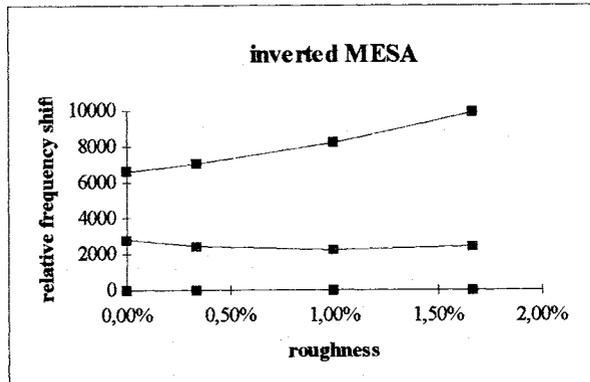


figure 10 relative frequency shift of inharmonic modes versus roughness

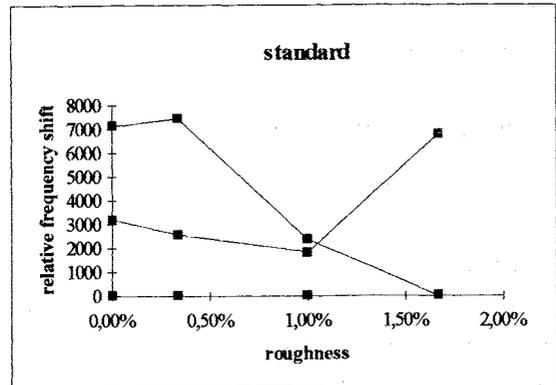


figure 11 relative frequency shift of inharmonic modes versus roughness

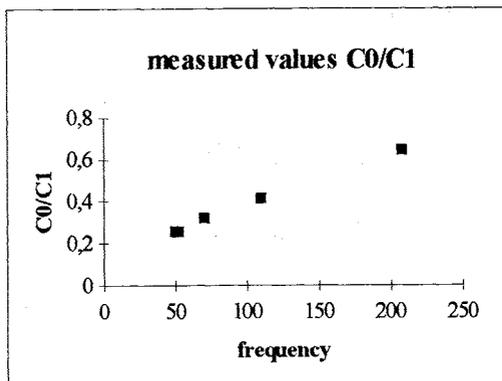


figure 13 measured C<sub>0</sub>/C<sub>1</sub> values versus frequency

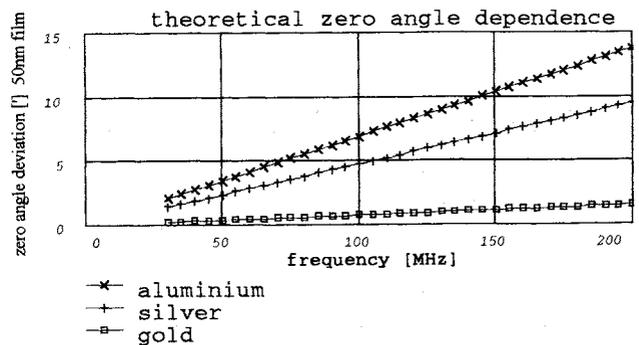
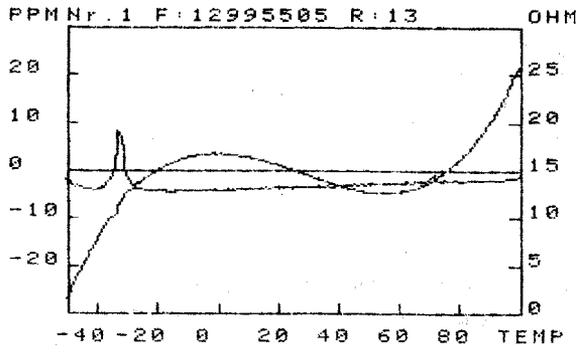
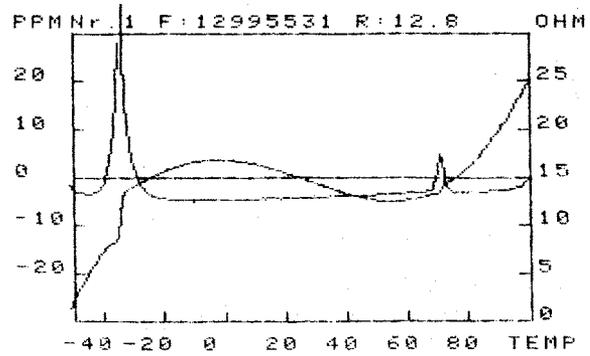


figure 14 zero cut angle

**active dips**

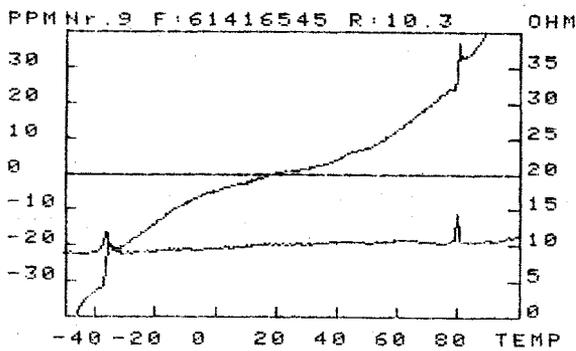


**figure 15 drive level 0,1 mW**

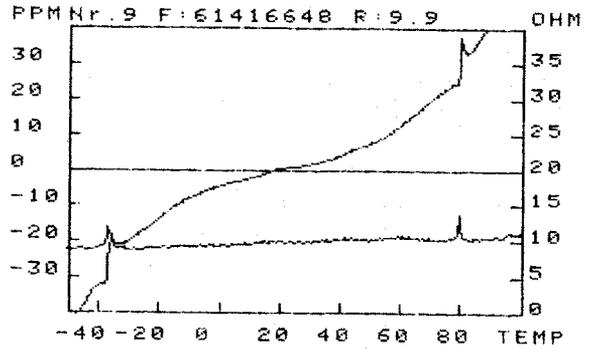


**figure 16 drive level 0,5 mW**

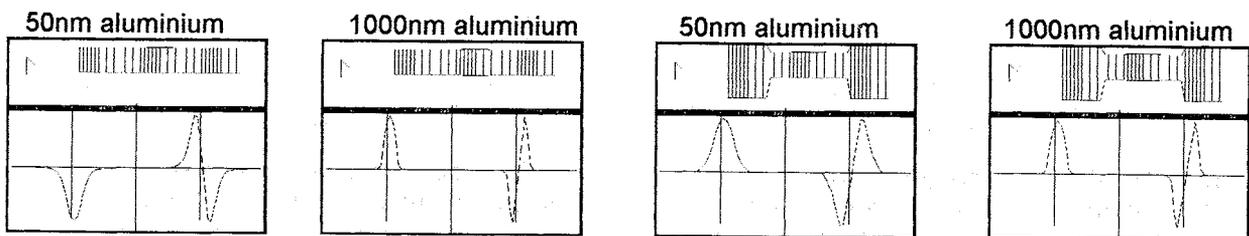
**high frequency dips**



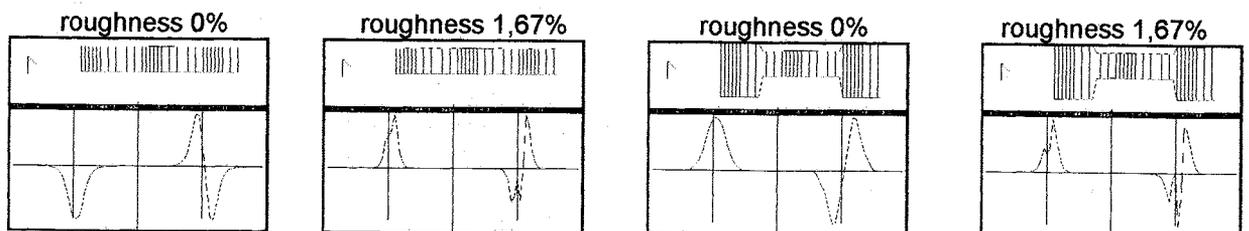
**figure 17 drive level 0,1 mW**



**figure 18 drive level 0,5 mW**



**figure 19 amplitude dependent on film thickness for standard and inverted MESA**



**figure 20 amplitude dependent on roughness for standard and inverted MESA  
50nm aluminium**

**Table 1**  
**measured equivalent data**

frequency [Mhz]	overtone	format	R1 [Ohm]	C0 [pF]	C1 [fF]	Q [k]	COk [pF]	G0k [Ohm]
49,80	1	MESA	11,0	1,80	6,950	40		
51,84	1	standard	11,9	2,26	9,260	28		
51,84	1	MESA	9,0	2,22	8,710	39		
61,44	1	standard	12,1	3,70	14,800	15		
70,00	1	MESA	9,0	2,60	8,000	33		
110,00	1	MESA	10,0	3,70	9,000	17		
207,00	1	MESA	9,3	2,04	3,150	26	2,38	11400
622,00	3	MESA	55,3	2,04	0,177	26	2,40	1230
1038,00	5	MESA	116	2,04	0,111	12	3,20	250

# CHARACTERISTICS OF HIGH FREQUENCY FUNDAMENTAL RECTANGULAR QUARTZ CRYSTAL RESONATORS

B. Studer, W. Zingg, Micro Crystal, Div. of ETA SA, 2540 Grenchen  
and S. Dalla Piazza, Asulab, Rue des Sors 3, 2074 Marin, Switzerland

## Abstract

The interest for high frequency fundamental (HFF) quartz crystals is soaring up in the fields of telecommunications and instrumentation. High speed communication systems require high frequency references and fundamental mode resonators have preference when either a large pulling range, a large bandwidth or a low spurious level in the spectrum are required.

Micro Crystal has developed a manufacturing process combining photolithographic batch technology and wet chemical local thinning on wafer. Thus, it is possible to produce HFF rectangular AT resonators with inverted mesa structure in large quantities and at reduced labor cost.

A crystal was designed at 50 MHz (membrane thickness 33 micrometers) and sufficient quantities are being produced for technological characterisation and assessment.

The AT-HFF resonator is very robust due to single ended rigid mounting and encapsulation in a hermetically sealed SMD ceramic package (8 x 3.6 x 1.8mm<sup>3</sup>). Electrical, thermal and environmental characteristics of this type of quartz crystals are presented.

Keywords: Quartz, deep etching, batch process, wafer, tuning, surface mount package

## INTRODUCTION

Since more than ten years, considerable progress in processing HFF quartz resonators and filters with inverted mesa structure has been achieved by several manufacturers (1, 2, 3). The main techniques reported in this field are ion beam milling (4), plasma etching (2,5) and wet chemical etching with either alkaline solutions (3) at relatively high temperatures or hydrofluoric acid based solutions (6). Regardless of the technique, the goal always remains the same:

a limited inner area of an originally flat crystal blank must be thinned down (from one or both sides simultaneously) to the required membrane thickness (33 micrometers for 50 MHz fundamental AT thickness shear mode).

With regard to dry milling methods, wet chemical techniques bear intrinsic advantages like natural compatibility with batch processing and potential cost reduction through high volume production. After a feasibility study in 1993, ASULAB (R&D centre of SMH) started a project with the goal to develop an industrial batch process for producing locally thinned rectangular AT-HFF resonators. In this wet chemical process, two key aspects had to be addressed: first the etching must allow deep etching without detrimental surface modification and second some procedure for chemically adjusting the membrane thickness (tuning to thickness) on the wafer had to be worked out.

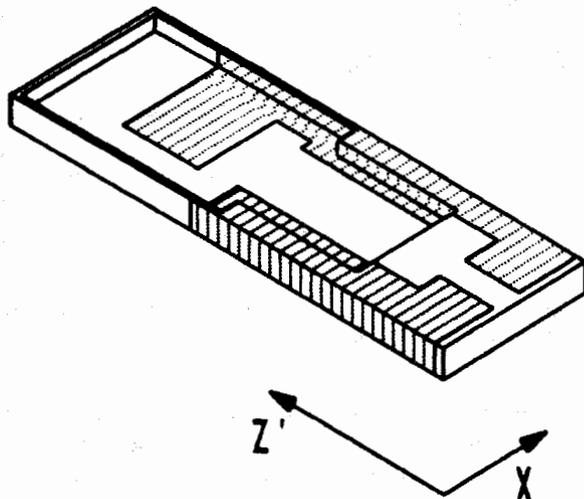
During 1995, Micro Crystal developed wet chemical thinning processes, frequency measuring techniques of membranes on wafer and automatic operating equipment. In the meantime manufacturing of 50 MHz HFF quartz crystals has reached its cruising speed. Performances of resonators encapsulated in a SMD ceramic package together with process features and capabilities are presented.

## GEOMETRY OF HFF RESONATOR

The high frequency fundamental resonator developed by Micro Crystal has the geometry shown in figure 1.

The shape results from a hybridisation between AT strips (7,8) and inverted mesa structures (1). The recessed area is thinned down to a membrane from both sides simultaneously. The length is along the Z'-axis and the width along the electrical X-axis. The membrane is surrounded by a frame of approximately 100  $\mu$ m width and thickness corresponding to the original

wafer blank thickness (125  $\mu\text{m}$ ). This frame essentially maintains a well defined geometry of the membrane.



**Figure 1:** Geometry of a 50 MHz AT-HFF crystal resonator with dimensions 5.8 mm (length along Z'), 1.9 mm (width) and 33  $\mu\text{m}$  (membrane thickness)

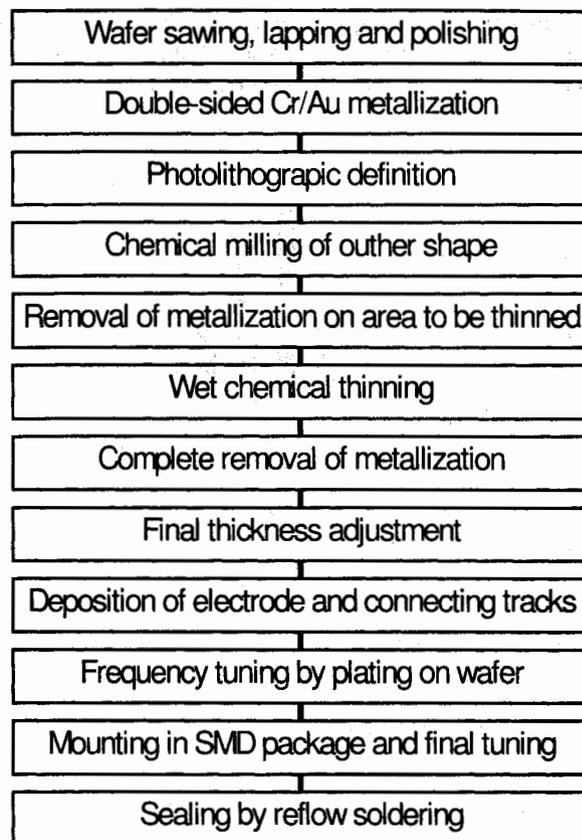
The aluminium electrodes of 1000  $\text{\AA}$  nominal thickness are deposited under vacuum through mechanical masks, the wafer being positioned between two quartz wafers used as apertures. Both electrodes are connected electrically to the mounting pads with Cr/Au tracks running along the frame. The resonator is mounted in a ceramic package with conductive epoxy as a cantilever with rigid clamping at one end. This mounting is very shock resistant and reliable for SMT applications. In addition it avoids mechanical and thermal stresses in the active area.

## TECHNOLOGY

Our goal was to implement a new wet chemical thinning process into an existing photolithographic wafer batch process (8). The technology is similar to that described by Ishii et al (6) and a simplified block diagram is given in figure 2.

### Wet Chemical Wafer Batch Processing

The first step consists in preparing the AT-cut square wafers (38.1 mm by side). As the surface roughness is very important, wafers are to be polished after sawing and lapping. Their final thickness is 125 microns. They are coated on both sides with a Cr/Au metallization performed by vacuum evaporation.



**Figure 2:** Block diagram of the wafer batch manufacturing process

Next the outer shape of the strip resonator is defined by photolithographic exposure in a double-side mask aligner followed by selective Au etching. Then the wafers are dipped into the HF/NH<sub>4</sub>F etch bath at 85 °C for shape milling. At that stage the Cr/Au metallization is removed on the area of each side where thinning has to occur.

In order to preserve the surface quality, the choice of an appropriate etching solution (9) for chemical thinning is not trivial. We performed initial tests with a commercially available etchant (10). Since this non-aqueous solution is very expensive, we finally decided to brew a customized etch bath based on concentrated hydrofluoric acid and organic surfactants. The etch rate of this solution only is 0.9 micron / h / side at a temperature of 24 °C. In addition it should be mentioned that operation with adequate exhaust ventilation is required due to stench and safety considerations.

After local chemical thinning, the membrane thickness or frequency may vary considerably from one unit to the other, on the same or different wafers. Therefore it is necessary to perform an adjustment of thickness of the individual resonators using a final etching to thickness equipment. The frequency of the thinned area is measured using the reflected resonance signal from a sweeping synthesizer. The etching to thickness of individual resonators occurs in an iterative process (usually five steps) successively with reduced etch time

i.e. finer frequency adjustments. In each step, only pre-defined crystals in any given wafer have to be exposed to the etchant, the others remaining protected.

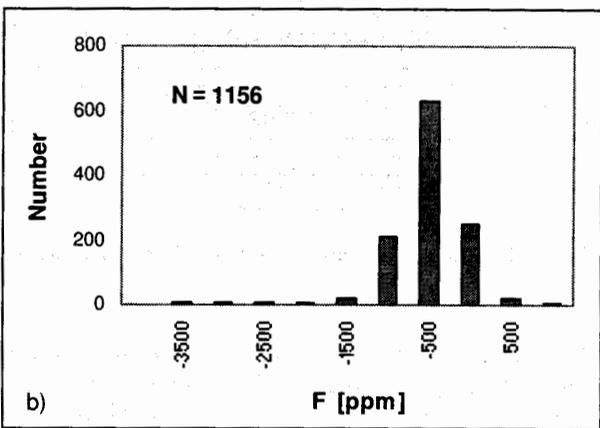
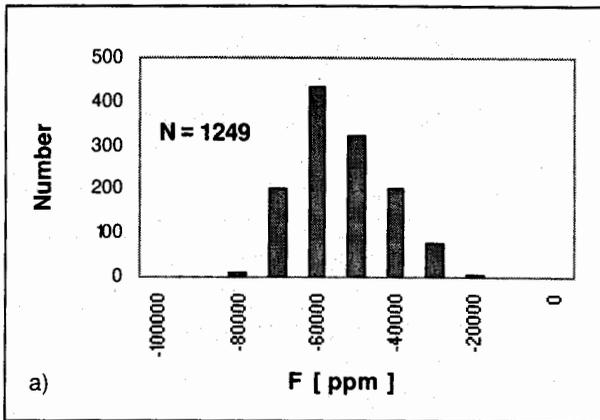


Figure 3: Histograms of frequency illustrating the capability of the etching to thickness process : a) before and b) after chemical tuning. Target frequency is 50.4 MHz

This adjustment is necessary in order to reduce the initial thickness variations. We start with 125 microns thick wafers with a tolerance of +/- .5  $\mu\text{m}$ . Assuming uniform thinning, a frequency spread of +/- 15'000 ppm would result after etching to 50 MHz. In practice, we get a distribution width approximately twice as broad. As seen in figure 3, this width is reduced typically to +/- 1500 ppm after the chemical tuning to thickness procedure.

### Electrode Metallization and Frequency Tuning

Aluminium electrodes are deposited under vacuum with an electron gun heater. The nominal thickness was chosen as 1000 Å based on experiments with variable base plating (figure 4).

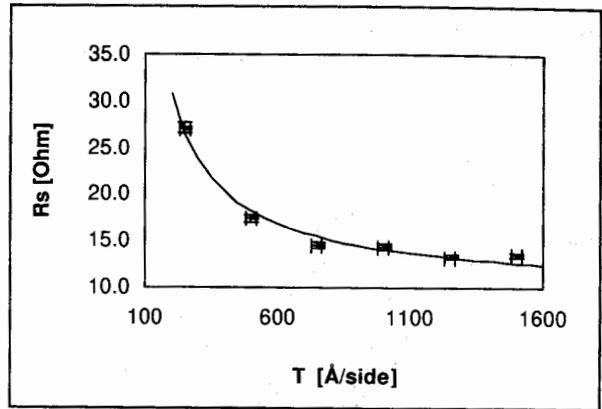


Figure 4: Series resistance  $R_s$  measured on wafer versus Al electrode thickness  $T$

The electrical connecting tracks between the electrodes and the mounting pads are obtained by vacuum evaporation of Cr/Au layers. At this stage, the wafers are ready for tuning to frequency which occurs by evaporation of gold on an automatic wafer plating equipment (11). Micro Crystal developed its own equipment for this operation but complete systems are available commercially (12).

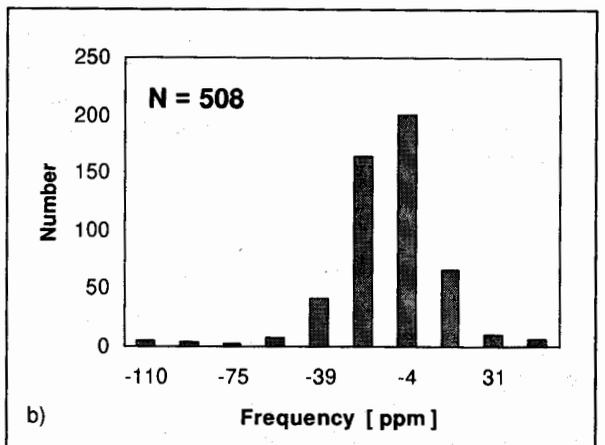
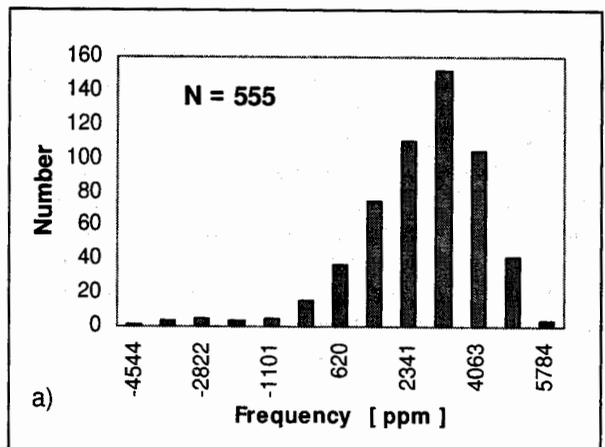


Figure 5: Histograms of resonator frequencies: a) before and b) after wafer tuning

Typical distributions before and after plating to frequency on wafer are shown in figure 5. Before tuning, the frequencies typically are 2000 to 3000 ppm above target. The width of the frequency distribution which may be +/- 4000 ppm before tuning reduces to less than +/- 100 ppm after wafer tuning. It should be emphasized that very little gold is needed for this adjustment. For a 2000 ppm frequency decrease at 50 MHz, the average thickness of gold layer only amounts to about 40 to 50 Å.

### Assembling

The wafers are electrically tested (frequency, resistance, motional capacitance) before mounting the crystals into the ceramic package (13). The resonators are attached to the metallized ceramic pad with a two component conductive epoxy resin. A final tuning to frequency may be required when tolerances tighter than +/- 50 ppm are called for. This final adjustment also has to take into account the frequency shift occurring during the sealing process ( in our case about +20 to + 30 ppm ). All SMD ceramic packages are hermetically sealed under vacuum by reflow of a Au/Sn solder preform at eutectic composition (80/20).

### CHARACTERISTICS OF RECTANGULAR AT-HFF RESONATORS

As described in the previous section, the high frequency fundamental quartz resonators are thinned locally from both sides on the wafer. The pretuned crystals are cantilever mounted in a ceramic package, tuned to final frequency and sealed hermetically under vacuum in a reflow process.

The electrical, thermal and environmental characteristics of AT-HFF crystals in SMD package are presented and discussed below.

#### Electrical Characteristics

At 50 MHz, AT-HFF crystals with the geometry shown in figure 1 and an active electrode area of  $1\text{mm}^2$  have following typical electrical parameters :

Series resistance	$R_s$ (Ohm)	15
Motional capacitance	$C_1$ (fF)	5.6
Motional inductance	$L_1$ (mH)	1.75
Shunt capacitance	$C_0$ (pF)	2.2

From above values, one obtains the quality factor  $Q = 36'000$  and the capacitance ratio  $C_0/C_1 = 380$ . It should be noted that  $C_0$  includes the capacitance of the SMD package which amounts to 1 pF. Thus the resonator itself has a shunt capacitance of 1.2 pF.

The distribution of motional capacitances of a manufacturing batch is shown in figure 6.

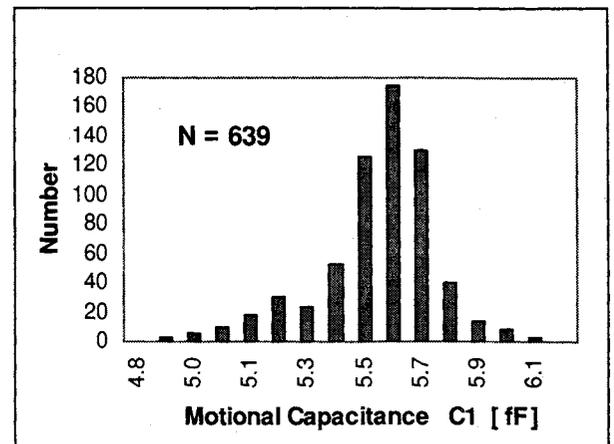


Figure 6: Distribution of C1 values of assembled crystals

The measurement of electrical parameters is performed with the help of an instrument based on a transmission method (e.g. TRANSAT CNA 300 or SAUNDERS 250 A). Although this kind of equipment is tailor-made for crystal measurement it should be emphasized that commonly available Pi-networks are not adequate for SMD packages because the clamping jaws are designed for crystals with leads (e.g. HC49). For that reason we adapted a Pi-network onto a home-made fixture for SMD packages where the crystal pads are contacted with metallic tips.

An alternative measuring method is with network analyzers. Such instruments (e.g. HP 5100) allow to display equivalent circuit parameters.

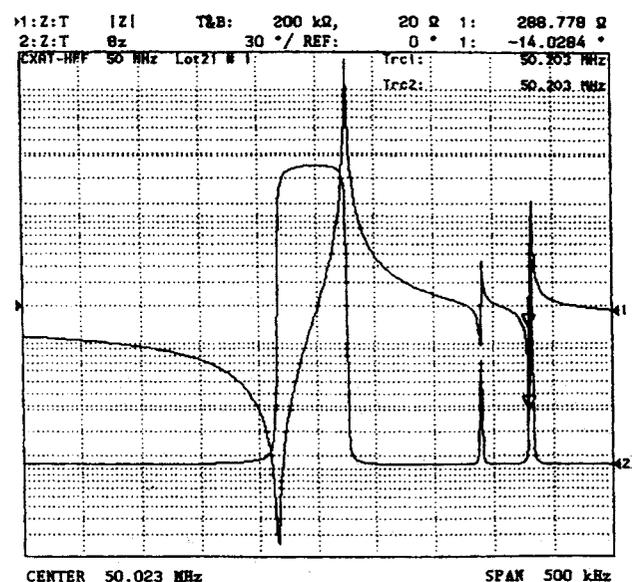


Figure 7: Spectral response with fundamental mode and anharmonic overtones

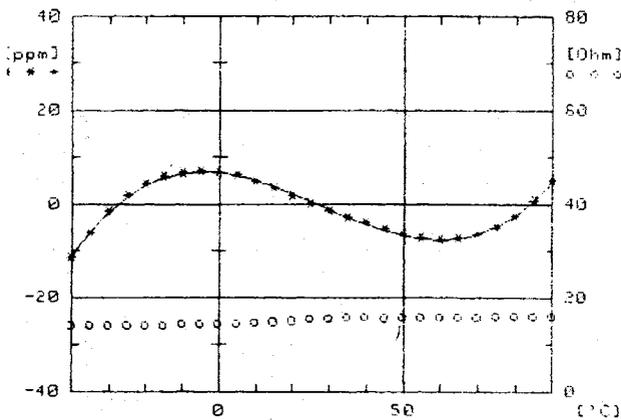
Again one should be wary about motional parameters when measuring SMD crystals in a Pi-network because erratic values may arise if electrical contacts are capacitive rather than ohmic. In addition correct crystal parameters are obtained only after calibration and optimal sweep configuration setting (frequency span, analyzer bandwidth, averaging, etc.). The impedance and phase response of a HFF crystal is shown in figure 7. The anharmonic overtone modes which appear 150 to 200 kHz above the fundamental resonance are of minor concern because they are attenuated by more than 10 dB.

### Thermal Characteristics

The dependence of frequency and resistance of a crystal on temperature is crucial in applications requiring high stability. Due attention to the temperature behaviour is at place for assessment of optimal design. The thermal response of a crystal is considered to be good if the measured frequency versus temperature curve is smooth and can be well fitted with a polynomial of third or fourth order. The response of a good crystal (figure 8) can be closely fitted with the function

$$Y = \alpha (T-T_i) + \gamma (T-T_i)^3$$

where  $Y = [F(T) - F(T_i)] / F(T_i)$  is the relative frequency variation referenced to the frequency of the individual crystal at the inflection point  $T_i$ . The linear coefficient  $\alpha$  (ppm/°C) is the slope at  $T_i$  and the third order coefficient  $\gamma$  is most conveniently expressed in  $10^{-12}$  per °C<sup>-3</sup>.

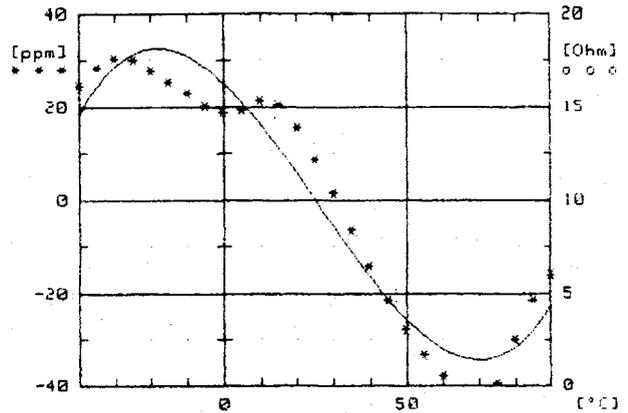


**Figure 8: Frequency versus temperature. The fitted curve closely matches measured (\*) values. Characteristic values:**

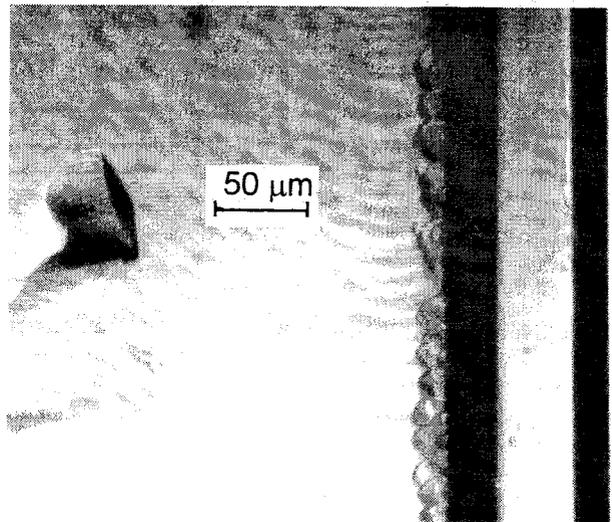
$$T_i = 28^\circ\text{C}, \alpha = -0.3 \text{ ppm}/^\circ\text{C}, \gamma = 100 \cdot 10^{-12} \text{ }^\circ\text{C}^{-3}$$

In practice a fraction of resonators (in dire plights up to 30%) show more or less dramatic deviations from a regular cubic curve. A typical distorted thermal response is illustrated in figure 9. In the present case the origin of this skew behaviour is a hillock located on the edge of the electrode (figure 10). This defect appeared

during chemical thinning because a metallic spot remained on the stripped quartz surface and subsequently acted as a mask. Geometrical flaws in the active area obviously may produce unacceptable perturbations in the vibrational mode. Most frequent defects are etch pits and hillocks. Since their lateral dimensions usually extend to tens of micrometers they can be observed easily in a binocular with magnification 20 to 80.



**Figure 9: Crystal with irregular F vs. T response**



**Figure 10: Hillock causing irregular F vs. T**

In order to get a three dimensional view of typical defects a topographical analysis with an atomic force microscope was conducted. This instrument scans the sample surface with a sharp tip fixed on a flexible cantilever whose deflection is monitored with a laser beam. Figure 11 shows the tail of an etch pit ending at the orifice of an etch channel.

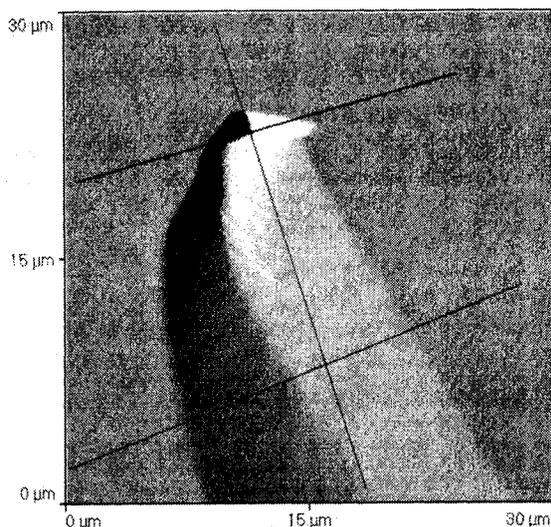


Figure 11: Topographic picture of the tail of an etch pit in the atomic force microscope

Profiles (figure 12) were taken along the straight lines. The aperture of the channel is located at the upper lines intersection. It may be seen from the corresponding profiles that the probe tip drops into the channel by more than one micrometer and that the lateral dimensions of the hole are 6 and 2 micrometers respectively. The third profile indicates that the etch pit only is approximately 200 nm deep. Measurement of surfaces without flaws yield roughness value  $R_a$  of about 4 nm.

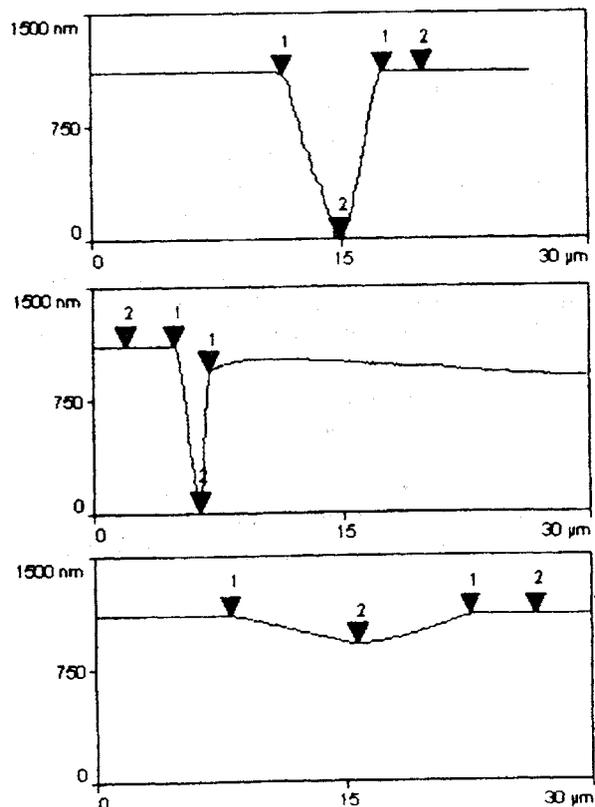


Figure 12: Atomic force microscope profiles showing the location and size of an etch channel orifice and the depth of an etch pit

## ENVIRONMENTAL CHARACTERISTICS

SMD quartz crystals must withstand a number of environmental tests for qualification and conformance to the severe conditions which they inevitably face during surface mounting (pick and place machines, belt furnace for reflow or infrared soldering, etc.). Several groups of crystals (size of samples typically 10 to 30 units) were carried through following qualification tests:

thermal shock, mechanical shock and vibration, solder dipping and accelerated aging.

### Thermal Shocks

Twelve crystals were thermally shocked between  $-55^{\circ}\text{C}$  and  $85^{\circ}\text{C}$  during 15 cycles (transition time between both levels approx. 5s, soak time approx. 30 min.). Frequency changes before and after test of all samples were less than  $\pm 2$  ppm and series resistance remained stable to within  $\pm 1$  ohm.

### Mechanical Shocks and Vibrations

Two groups of ten crystals were subjected to mechanical shocks (8000 g, 0.2 ms) and vibrations (3 runs of 20 min from 10 to 2000 Hz, 20g) respectively. The frequencies of all crystals remained within  $\pm 5$  ppm and the change of resistance was less than 1 ohm.

### Solder Dipping

This test is useful because it is very quickly done and instantaneously tells if the package has no gross leak and the glass lid withstands the thermal shock. Seven crystals were dipped into molten PbSn solder at  $200^{\circ}\text{C}$ . No one did shift by more than  $\pm 5$  ppm and resistances were stable.

### Accelerated Aging

Two groups of 16 crystals with different thickness of Al-electrode (500 and 1000 Å) were aged during two months at  $85^{\circ}\text{C}$ . The crystals were not oscillating while maintained in the oven. Measurements were made with a crystal impedance meter at room temperature. The symbols in the graph (figure 13) represent mean values of load resonant frequencies. Systematic offsets arising at different dates are attributed to adjustment errors of the CI-meter since all crystals move in chorus. In spite of this scattering an overall positive aging of some 3 ppm can be extracted.

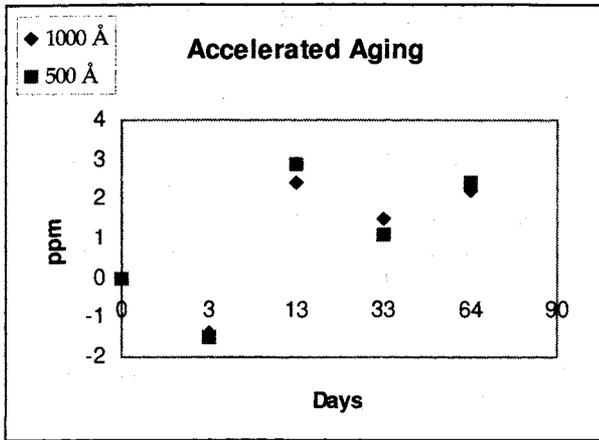


Figure 13: Aging test results (see text)

## CONCLUSIONS

Production of high frequency fundamental quartz resonators using a technique of wet chemical local thinning has been successfully implemented into an industrial photolithographic batch manufacturing process. The main technological features and capabilities have been described together with the geometry and performances of a 50 MHz crystal with inverted mesa structure (membrane etched from both sides). Careful analysis of the dependence of frequency on temperature has demonstrated the detrimental effect of etching flaws like hillocks, etch pits or groves and etch channels. This technology bears a considerable potential and outlook is promising in the context of photolithographic versatility. In order to take full benefit of its resources it is necessary to address challenging issues in crystal design, quality of raw material and control of process parameters.

## Acknowledgements

The authors wish to thank U. Hänni from the Engineering School in Bienne for his helpful topographical analysis with the atomic force microscope and they express their gratitude to S.Uebersax who had the patience to patch text and figures together on the computer.

## References

- Hunt J.R. and Smythe R.C., 1985, Chemically Milled VHF and UHF AT-cut Resonators, Proceed. 39th Freq. Control Symp., 292-300
- Stern F.M., Dowsett J., Carter D.J. and Williamson R.J., 1989, The Fabrication of High Frequency Fundamental Crystals by Plasma Etching, Proceed. 43rd Freq. Control Symp., 634-639
- Cambon O., Michel J.-P., Lamboley J., Aubry J.-P., Billon Y., Mouton D., Lozach I., 1994, An Industrial Quartz Manufacturing Process by Chemical Etching, Proceed. 8th Europ. Freq. and Time Forum, 256-262
- Berté M. 1977, Acoustic Bulk Wave Resonators and Filters Operating in the Fundamental Mode at Frequencies Greater than 100 MHz, Proceed 31st Freq. Control Symp., 122-125
- Wang J.S., Watson S.K. and Lau K.F., 1984, Reactive Ion Beam Etching for VHF Crystal Resonators, Proceed. 38th Freq. Control Symp., 101-104
- Ishii O., Morita T., Saito T. and Nakazawa Y., 1995, High Frequency Fundamental Resonators and Filters Fabricated by Batch Process using chemical Etching, Proceed. IEEE Int. Freq. Control Symp. 818-826
- Dworsky L.N., 1987, AT Quartz Strip Resonators, Proceed. 41st Freq. Control Symp., 295-302
- Studer B. and Zingg W., 1990, Technology and Characteristics of Chemically Milled Miniature Quartz Crystals, Proceed 4th Europ. Freq. and Time Forum, 653-658
- Vondeling J.K., 1983, Fluoride-Based Etchants for Quartz, J. Mater. Sci. 18, 304-314
- Lattice Etch™, Trade mark of non-aqueous isotropic etchant supplied by Xeco Inc., 1579 N. Main Street, Cedar City, Utah 84720 USA
- Varsovia J. and Chuang S.S., 1995, Wafer Backplating of Subminiature AT Quartz Resonators, 17th Piezoelectric Dev. Conf. and Exhib. 1, 40-44
- Saunders & Assoc., Inc., Scottsdale, AZ, USA, Model 5511 Wafer Plating System
- Coleman D., 1992, Ceramic Packaging Options for Surface Mount Crystals, Oscillators and SAW Devices, 14th Piezoelectric Dev. Conf. and Exhib., 1, 9-12.

# ENERGY TRAPPING IN BULK ACOUSTIC WAVE COMPOSITE RESONATOR

G.D.MANSFELD

Institute of Radioengineering and Electronics of Russian Academy of Sciences  
11, Mokhovaya str., 103907, Moscow, Russia

## Abstract

The analytical approach to the problem of energy trapping in bulk acoustic wave composite resonators is developed. It is based on a standard system of wave equations and boundary conditions for adjacent to resonator three layer structure - relatively thick layer of the substrate, thin metal electrode deposited onto its surface and the layer of piezoelectric film. Dispersion equation is solved to find small components of the wave vector in the direction in the plane of the structure outside the resonator. The obtained formula describes the attenuation of the wave in the region close to resonator. This expression is purely analytic and does not require matching procedure in calculations. The situation corresponding to microwave bulk acoustic wave resonators structures is discussed. In particular the developed approach at least qualitatively explains surprisingly perfect energy trapping which exists in almost all microwave composite resonators.

## 1. Introduction

The main purpose of this work is to obtain a simple expression describing the energy trapping phenomena in bulk acoustic wave resonators for microwaves. As a matter of fact most of such structures display surprisingly perfect energy trapping resulting in the fact that in most cases total losses in the structure are governed by only material losses - in perfect crystals by Akhiezer losses. It means that the losses connected with energy radiation from the resonator region in outer part of the sample are negligibly small. Using this fact a new composite resonator method of losses measurements in thin films and layers was proposed and successfully experimentally used [1-3].

The analysis in this work is based on the solution of the standard system of wave equations and boundary conditions for BAW composite resonator configuration which is schematically shown on Fig.1.

In our analysis we followed the usual scheme. Our structure formally consists of two regions - region I with top electrode and region II without it. Waves exited in region I can't penetrate in region II with higher eigenfrequencies owing to the appearance of pure imaginary components of wave vector in perpendicular direction. Similarly to other works we consider that these components are the same for film and substrate. The difference in eigenfrequencies in regions I and II appears due to the existence of the top electrode in region I (in resonator).

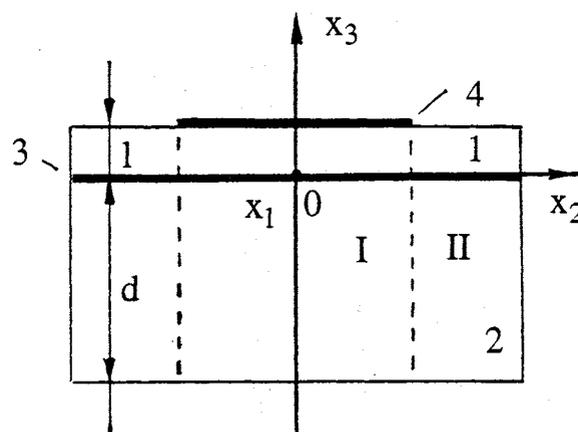


Fig.1. Schematic of BAW composite resonator structure.

We restrict our analysis by simple case of pure longitudinal mode propagating along the C-axis of hexagonal piezoelectric crystalline film 1 (ZnO) placed on the surface of cubic substrate 2 (YAG, Si, etc.). Ground metal electrode 3 is located in interface between film and substrate. Upper electrode 4 is placed on the surface of piezoelectric film. Surfaces of the structure are suggested to be ideally flat and parallel.

## 2. Basic equations

The set of wave equations and boundary conditions were the same as in [4]:

Wave equations for film are:

$$\bar{\rho} \frac{\partial^2 \bar{u}_i}{\partial t^2} = \frac{\partial \bar{T}_{ij}}{\partial x_j}, \quad i, j = 1, 2, 3. \quad (1a)$$

wave equations for substrate are:

$$\rho \frac{\partial^2 u_i}{\partial t^2} = \frac{\partial T_{ij}}{\partial x_j}, \quad i, j = 1, 2, 3. \quad (1b)$$

Mechanical stresses for film are:

$$\bar{T}_{11} = \bar{c}_{11} \frac{\partial \bar{u}_1}{\partial x_1} + \bar{c}_{12} \frac{\partial \bar{u}_2}{\partial x_2} + \bar{c}_{13} \frac{\partial \bar{u}_3}{\partial x_3} + e_{31} \frac{\partial \phi}{\partial x_3}$$

$$\begin{aligned}
\bar{T}_{22} &= \bar{c}_{12} \frac{\partial \bar{u}_1}{\partial x_1} + \bar{c}_{11} \frac{\partial \bar{u}_2}{\partial x_2} + \bar{c}_{13} \frac{\partial \bar{u}_3}{\partial x_3} + e_{31} \frac{\partial \phi}{\partial x_3} \\
\bar{T}_{33} &= \bar{c}_{13} \frac{\partial \bar{u}_1}{\partial x_1} + \bar{c}_{13} \frac{\partial \bar{u}_2}{\partial x_2} + \bar{c}_{33} \frac{\partial \bar{u}_3}{\partial x_3} + e_{31} \frac{\partial \phi}{\partial x_3} \\
\bar{T}_{12} &= \bar{c}_{44} \frac{\partial \bar{u}_1}{\partial x_2} + \bar{c}_{44} \frac{\partial \bar{u}_2}{\partial x_1}
\end{aligned}
\tag{2}$$

with electrical displacement:

$$\begin{aligned}
D_1 &= e_{15} \frac{\partial \bar{u}_3}{\partial x_1} + e_{15} \frac{\partial \bar{u}_1}{\partial x_3} - \epsilon_{11} \frac{\partial \phi}{\partial x_1} \\
D_2 &= e_{15} \frac{\partial \bar{u}_3}{\partial x_2} + e_{15} \frac{\partial \bar{u}_2}{\partial x_3} - \epsilon_{11} \frac{\partial \phi}{\partial x_1} \\
D_3 &= e_{31} \frac{\partial \bar{u}_1}{\partial x_1} + e_{31} \frac{\partial \bar{u}_2}{\partial x_2} + e_{33} \frac{\partial \bar{u}_3}{\partial x_3} - \epsilon_{33} \frac{\partial \phi}{\partial x_3}
\end{aligned}
\tag{3}$$

Mechanical stresses for substrate:

$$\begin{aligned}
T_{11} &= c_{11} \frac{\partial u_1}{\partial x_1} + c_{12} \frac{\partial u_2}{\partial x_2} + c_{12} \frac{\partial u_3}{\partial x_3} \\
T_{22} &= c_{12} \frac{\partial u_1}{\partial x_1} + c_{11} \frac{\partial u_2}{\partial x_2} + c_{12} \frac{\partial u_3}{\partial x_3} \\
T_{33} &= c_{12} \frac{\partial u_1}{\partial x_1} + c_{12} \frac{\partial u_2}{\partial x_2} + c_{11} \frac{\partial u_3}{\partial x_3} \\
T_{12} &= c_{44} \frac{\partial u_1}{\partial x_2} + c_{44} \frac{\partial u_2}{\partial x_1} \\
T_{23} &= c_{44} \frac{\partial u_3}{\partial x_2} + c_{44} \frac{\partial u_2}{\partial x_3} \\
T_{31} &= c_{44} \frac{\partial u_3}{\partial x_1} + c_{44} \frac{\partial u_1}{\partial x_3}
\end{aligned}
\tag{4}$$

Below we used following notations:

the thickness of the substrate is  $d$ , the wave vector  $\beta$ ; thickness of the film is  $l$ , wavevector -  $q$ ; thicknesses of top and ground electrodes are respectively  $b_1, b_2$ , their material densities are  $\rho_1$  and  $\rho_2$

### 3. Waves in infinite media

Here some necessary relations connected with the problem of wave propagation in the directions very close to normal to the plate plane are found. The solution in the substrate material is looked in the form:

$$u_i = A_i \cdot \exp(-i\beta x_3 - i\xi x_1 - i\nu x_2 + i\omega t) \tag{5}$$

Substitution of (5) in (1b) results in set of linear equations for  $A_i$ , with proper dispersion relation given by putting to zero of determinant:

$$\begin{vmatrix}
c_{44}\beta^2 - \rho\omega^2 & 0 & \xi\beta(c_{44} + c_{12}) \\
0 & c_{44}\beta^2 - \rho\omega^2 & \nu\beta(c_{44} + c_{12}) \\
\xi\beta(c_{44} + c_{12}) & \nu\beta(c_{44} + c_{12}) & c_{11}\beta^2 - \rho\omega^2
\end{vmatrix} = 0
\tag{6}$$

For  $\xi = \nu = 0$ ,  $\beta = \beta_0$  there is three pairs of roots corresponding to pure longitudinal and transverse modes in  $x_3$  direction. Further only the longitudinal modes with wave vectors  $\beta_0 = \pm\omega_0\sqrt{\rho/c_{11}}$  are considered. For

small  $\xi, \nu, \Delta$  ( $\beta = \beta_0 + \Delta$ ) using (6) it is possible to find relation between these quantities and the corresponding change of the frequency  $\Delta\omega$ , ( $\omega = \omega_0 + \Delta\omega$ ):

$$\frac{\Delta}{\beta_0} = -\frac{1}{2} \frac{\xi^2 + \nu^2}{\beta_0^2} \left[ \frac{(c_{44} + c_{12})^2}{c_{11}(c_{11} - c_{44})} \right] + \frac{\Delta\omega}{\omega_0} \tag{7}$$

Similar relation can be obtained and for the transducer media with  $q_0 = \pm\omega_0\sqrt{\rho/c_{33}^p}$ :

$$\begin{aligned}
\frac{\delta}{q_0} &= -\frac{1}{2} \frac{\xi^2 + \nu^2}{q_0^2} \\
&\left[ \frac{(\bar{c}_{44}^p \bar{c}_{44} + \bar{c}_{44}^p \bar{c}_{13} + \bar{c}_{13}^p \bar{c}_{44} + \bar{c}_{13}^p \bar{c}_{13})}{\bar{c}_{33}^p (\bar{c}_{33}^p - \bar{c}_{44})} \right] + \frac{\Delta\omega}{\omega_0}
\end{aligned}
\tag{7a}$$

Here  $\bar{c}_{33}^p, \bar{c}_{12}^p, \bar{c}_{44}^p$  are the piezoelectrically stiffened constants.

Other important result is that the ratios of amplitudes  $A_1/A_3, A_2/A_3, \bar{A}_1/\bar{A}_3, \bar{A}_2/\bar{A}_3$ , are very small and have the orders of magnitude correspondingly  $\xi/\beta_0, \nu/\beta_0, \xi/q_0, \nu/q_0$  in accordance with[4].

### 4. Dispersion equation of the composite resonator structure

We traditionally start from the consideration of the composite resonator structure without upper electrode. Free oscillations in such structure are governed by next twelve equations forming the system of boundary conditions:

$$\begin{aligned}
T_{3i}(-d) &= 0 \\
\bar{T}_{3i}(0) - T_{3i}(0) &= b_2 \rho_2 \omega^2 u_i(0) \\
\bar{u}_i(0) &= u_i(0) \\
\bar{T}_{3i}(l) &= 0 \\
i &= 1, 2, 3.
\end{aligned}
\tag{8}$$

We use this system to find the correction to the classical solution for unidimensional problem. So we shall look for the solution in the form:

$$\begin{aligned}
u_i &= A_i \cdot \exp(-i\beta x_3 - i\xi x_1 - ivx_2 + i\omega t) + \\
&+ B_i \cdot \exp(i\beta x_3 - i\xi x_1 - ivx_2 + i\omega t); \\
\bar{u}_i &= \bar{A}_i \cdot \exp(-iqx_3 - i\xi x_1 - ivx_2 + i\omega t) + \\
&+ \bar{B}_i \cdot \exp(iqx_3 - i\xi x_1 - ivx_2 + i\omega t)
\end{aligned} \quad (9)$$

To simplify the problem we take the amplitude ratios as:

$$\begin{aligned}
\frac{A_1}{A_3} &= a^{(1)} \cdot \frac{\xi}{\beta_0}, & \frac{A_2}{A_3} &= a^{(2)} \cdot \frac{v}{\beta_0}, \\
\frac{B_1}{B_3} &= b^{(1)} \cdot \frac{\xi}{\beta_0}, & \frac{B_2}{B_3} &= b^{(2)} \cdot \frac{\xi}{\beta_0}, \\
\frac{\bar{A}_1}{\bar{A}_3} &= \bar{a}^{(1)} \cdot \frac{\xi}{q_0}, & \frac{\bar{A}_2}{\bar{A}_3} &= \bar{a}^{(2)} \cdot \frac{v}{q_0}, \\
\frac{\bar{B}_1}{\bar{B}_3} &= \bar{b}^{(1)} \cdot \frac{\xi}{q_0}, & \frac{\bar{B}_2}{\bar{B}_3} &= \bar{b}^{(2)} \cdot \frac{\xi}{q_0}
\end{aligned} \quad (10)$$

Here  $a^{(1)}, b^{(1)}, \bar{a}^{(1)}, \bar{b}^{(1)}, a^{(2)}, b^{(2)}, \bar{a}^{(2)}, \bar{b}^{(2)}$  – the new unknown coefficients.

After the substitution of (9) and (10) into (8) we obtain 12 algebraic equations. Four of them for  $i=3$  are:

$$\begin{aligned}
(i\bar{c}_{13}^p \bar{b}^{(1)} \xi^2 / q_0 + i\bar{c}_{13}^p \bar{b}^{(2)} v^2 / q_0 - \\
-i\bar{c}_{33}^p q) \cdot \exp(iql) \cdot \bar{B}_3 + \\
i\bar{c}_{13}^p \bar{a}^{(1)} \xi^2 / q_0 + i\bar{c}_{13}^p \bar{a}^{(2)} v^2 / q_0 + \\
+i\bar{c}_{33}^p q) \cdot \exp(iql) \cdot \bar{A}_3 = 0
\end{aligned} \quad (11)$$

$$A_3 + B_3 - \bar{A}_3 - \bar{B}_3 = 0 \quad (12)$$

$$\begin{aligned}
(ic_{12} b^{(2)} v^2 / \beta_0 + ic_{12} b^{(1)} \xi^2 / \beta_0 - \\
-ic_{11} \beta + \rho_2 \omega^2 b_2) B_3 + \\
ic_{12} a^{(1)} \xi^2 / \beta_0 + ic_{12} a^{(2)} v^2 / \beta_0 + \\
ic_{11} \beta + \rho_2 \omega^2 b_2) A_3 + \\
+(-i\bar{c}_{13}^p \bar{b}^{(1)} \xi^2 / q_0 - i\bar{c}_{13}^p \bar{b}^{(2)} v^2 / q_0 + \\
+i\bar{c}_{13}^p q) \cdot \bar{B}_3 + \\
(-i\bar{c}_{13}^p \bar{a}^{(1)} \xi^2 / q_0 - i\bar{c}_{13}^p \bar{a}^{(2)} v^2 / q_0 + \\
-i\bar{c}_{33}^p q) \cdot \bar{A}_3 = 0
\end{aligned} \quad (13)$$

$$\begin{aligned}
(ic_{12} b^{(2)} v^2 / \beta_0 + ic_{12} b^{(1)} \xi^2 / \beta_0 - \\
+ic_{11} \beta) \cdot \exp(-i\beta d) \cdot B_3 - \\
(ic_{12} a^{(1)} \xi^2 / \beta_0 + ic_{12} a^{(2)} v^2 / \beta_0 + \\
ic_{11} \beta) \cdot \exp(i\beta d) \cdot A_3 = 0
\end{aligned} \quad (14)$$

To obtain the convenient form of the dispersion relation the used procedure of solution was a subsequent substitution of all constants in(13) to eliminate  $\bar{A}_3, \bar{B}_3, B_3$ . The result for  $\beta = \beta_0 + \Delta$ ,  $q = q_0 + \delta$  is:

$$\begin{aligned}
z \cdot \tan(\beta_0 d + \Delta d) \cdot \left( 1 + \frac{\Delta}{\beta_0} + \frac{c_{12} \mu^2 (a^{(1)} - b^{(1)})}{2c_{11} \beta_0^2} \right) + \\
+ \bar{z} \cdot \tan(q_0 l + \delta l) \cdot \left( 1 + \frac{\delta}{q_0} + \frac{\bar{c}_{13}^p \mu^2 (\bar{a}^{(1)} - \bar{b}^{(1)})}{2\bar{c}_{33}^p q_0^2} \right) + \\
+ b_2 \rho_2 (\omega_0 + \Delta \omega) = 0
\end{aligned} \quad (15)$$

Here  $\mu^2 = \xi^2 + v^2$ .

For pure longitudinal mode, when  $\xi = v = 0$  it comes to (compare with [4]) well known result:

$$z \cdot \tan(\beta_0 d) + \bar{z} \cdot \tan(q_0 l) + b_2 \rho_2 \omega_0 = 0 \quad (16)$$

here  $z = \rho \cdot v$ ,  $\bar{z} = \bar{\rho} \cdot \bar{v}$  are the material acoustic impedances of substrate and the film respectively,  $v = \sqrt{c_{11}/\rho}$ ,  $\bar{v} = \sqrt{\bar{c}_{33}/\bar{\rho}}$  – are pure longitudinal wave velocities in the plate and in the film.

To find the constants in the amplitude ratios (10) remaining equations of (8) with  $i=1,2$  were used. The result with taking into account of (16) is:

$$\begin{aligned}
a^{(1)} &= -1; \\
\bar{a}^{(1)} &= -1; \\
b^{(1)} &= -a^{(1)}, \quad \bar{b}^{(1)} = -\bar{a}^{(1)}, \quad a^{(2)} = a^{(1)}, \\
b^{(2)} &= b^{(1)}, \quad \bar{a}^{(2)} = \bar{a}^{(1)}, \quad \bar{b}^{(2)} = \bar{b}^{(1)}
\end{aligned} \quad (17)$$

In order to avoid the problem of using of variational procedures in finding of correct values of  $\xi$ ,  $v$  matched to values of  $\beta$ ,  $q$  the proper expression was simplified by subtracting (16) from (15). The approximate ultimate expression is obtained from (15-17) with the use of (7). It is:

$$\begin{aligned}
\frac{\mu^2}{q_0^2} = \frac{\xi^2 + v^2}{q_0^2} = 2 \cdot \frac{\Delta \omega}{\omega_0} \cdot \frac{\left( 1 + \frac{d \cos^2 q_0 l \rho_s}{l \cos^2 \beta_0 d \rho_t} \right)}{P + Q} \cdot \frac{d \cos^2 q_0 l c_{11}}{l \cos^2 \beta_0 d \bar{c}_{33}^p}; \\
P = \left[ r + \left( r - 2\bar{a}_1 \frac{\bar{c}_{13}^p}{\bar{c}_{33}^p} \right) \frac{\sin q_0 l}{q_0 l} \right]; \\
Q = \left[ s + \left( s - 2a_1 \frac{c_{12}}{c_{11}} \right) \frac{\sin \beta_0 d}{\beta_0 d} \right]; \quad r = \frac{(c_{44} + c_{12})^2}{c_{11}(c_{11} - c_{44})}; \\
s = \frac{\bar{c}_{44}^p \bar{c}_{44} + \bar{c}_{44}^p \bar{c}_{13} + \bar{c}_{13}^p \bar{c}_{44} + \bar{c}_{13}^p \bar{c}_{13}}{\bar{c}_{33}^p (\bar{c}_{33}^p + \bar{c}_{44})};
\end{aligned} \quad (18)$$

The relative change of the frequency due to addition of the top electrode is found from the solution of the dispersion relation for pure longitudinal mode in resonator (region I in fig.1.):

$$z \cdot \tan(\beta_1 d) + \bar{z} \cdot \tan(q_1 l + \arctan \frac{\rho_1 b_1 q_1}{\bar{\rho}}) + b_2 \rho_2 \omega_1 = 0 \quad (19)$$

where:

$$\beta_1 = \beta_0 (1 + \Delta\omega/\omega_0), \quad q_1 = q_0 (1 + \Delta\omega/\omega_0), \\ \omega_1 = \omega_0 + \Delta\omega.$$

Then subtracting (16) from (19) one obtains:

$$\frac{\Delta\omega}{\omega_0} \approx \frac{\frac{\rho_1}{\bar{\rho}} \frac{b_1}{d} (1 - x_0 \tan q_0 l)^{-1}}{\frac{1}{d} + \frac{\rho_1}{\bar{\rho}} \frac{b_1}{d} \frac{1}{1 + x_0^2} + \left[ \frac{\rho}{\bar{\rho} \cos^2 \beta_0 d} + \frac{\rho_2}{\bar{\rho}} \frac{b_2}{d} \right] \cos^2 q_0 l} \quad (20)$$

$$\text{where } x_0 = \frac{\rho_1}{\bar{\rho}} q_0 b_1.$$

The equations (18)-(20) give the solution of the problem of free oscillations in the high overtone bulk acoustic wave composite resonator structure shown in Fig.1 in the vicinity of pure longitudinal mode.

## 5. Discussion

The obtained expression (18) is a relatively simple analytic formula describing the behaviour of the transverse components of wavevector in unelectroded region. It is seen that for negative values of  $\Delta\omega$  due to the influence of the top electrode  $\mu$  may be either pure imaginary or pure real depending of the sign of the remaining multiple in (18).

From this expression it is seen that for high ratios of  $d/l$  which is typical for high overtone bulk acoustic wave resonators the wave vector component  $\mu$  is pure imaginary providing good energy trapping.

It is also seen from (18),(20) that with the increase of the frequency the absolute value of  $\mu$  increases almost linearly providing better energy trapping. It explains at least qualitatively the observed very high values of  $Q$  factors of composite resonators at microwave frequencies. In comparison with internal material energy dissipation in the region of resonator only the small part of energy is lost outside the resonator.

With the decrease of frequency the effect of trapping is much poorer and some peculiarities of this phenomenon at small ratios of  $d/l$  may take place.

More detailed analytical research will be presented later.

## Acknowledgements

Part of this work was done in LCEP ENSMM, Besançon, France.

Author is grateful to prof. R.Besson (ENSMM) and academician Yu.V.Gulyaev (IRE RAS) for support of this work and to dr. B.Dulmet (ENSMM) and dr. N.I.Polzikova (IRE RAS) for useful discussions.

## References

- [1] Mansfeld G.D., "Measurements of acoustic wave attenuation in  $\text{La}_3\text{Ga}_5\text{SiO}_{14}$  using HBAR technique". in IEEE Int. Freq. Control Symp. proc. 1994, pp.35-39.
- [2] Mansfeld G.D., Freik A.D., Moorey M., "Characterisation of losses in quartz resonator using electromagnetic waves reflection coefficient measurements". in Actes proc. EFTF-95, pp. 20-23.
- [3] Mansfeld G.D. "BAW composite resonator spectroscopy", in 1994 IEEE Ultrasonics Symp. proc., vol. 2, pp. 655-658.
- [4] Tiersten H.F., Stevens D.S. "Analysis of thickness-extensional trapped energy resonator device structure with rectangular electrodes in the piezoelectric film on silicon configuration", in J.Appl. Phys.,1983, vol.54, 5893-5910.

# PTB's Primary Clock CS3

## Uncertainty Evaluation and Operational Experiences

A. Bauch, T. Heindorff

Physikalisch-Technische Bundesanstalt

Bundesallee 100, D-38116 Braunschweig, Germany

### Summary

The primary clock CS3 is now used for routine time-keeping and for the realization of the SI second in PTB's time unit laboratory. The development of the device has lasted quite a long time and the principle as well as some details of the construction have already been published. In this contribution we present the final result of a Type B evaluation and some details about the methods involved.

In case of the following systematic effects the initial design goals have been achieved, i.e. the relevant contributions to the uncertainty budget are as small as aimed at: Second-order Doppler effect, end-to-end cavity phase difference, cavity pulling, Rabi-pulling, servo electronics, gravity, blackbody radiation. Their evaluation practically does not rely on frequency measurements.

In other cases, systematic studies have been hampered by the short-term frequency instability which is  $9 \cdot 10^{-12} / \sqrt{\tau/s}$ , because the selection of slow atoms resulted in a smaller fraction of atoms contributing to the beam signal than naively expected.

A combined standard uncertainty of  $14 \cdot 10^{-15}$  was obtained. The largest components in the list of uncertainty contributions are connected with the quadratic Zeeman effect and the end-to-end phased difference. In the long term, a mean frequency difference between CS3 and CS2 of  $15 \cdot 10^{-15}$  was found during a one year period.

The most recent results will be included in the report at the Forum.

**STATE SELECTION IN A BIPOLAR GRADIENT MAGNET  
FOR A HIGH C-FIELD CESIUM BEAM RESONATOR**

G. Costanzo, A. De Marchi, M. Repetto

Politecnico di Torino, C.so D.d.Abruzzi, 24 - 10129 - TORINO - Italy  
E-mail : cesio@polito.it

**Abstract**

In Cs beam frequency standards it is essential to achieve a large population difference between the  $F=4$  and  $F=3$  hyperfine levels. For this purpose selection systems with 2N-polar gradient magnets are often used. An advantage of this method is its velocity selectivity, which can be exploited to obtain a velocity distribution centered on reasonably slow atoms (i.e. 100 m/s). This reduces velocity-dependent biases, like second order Doppler and phase shift.

In this paper we describe a state selection system with bipolar permanent magnets, to be used in the Cesium beam tube at the Politecnico di Torino. An experimental apparatus was realized to measure the deflection of Cesium atoms. Results are compared with data obtained by numerical analysis.

**Introduction**

The Politecnico di Torino planned in 1992 to build a Cs beam frequency standard based on the high C-field concept as described in [1].

The prototype operates with an intensity of the longitudinal magnetic field of about 82 mT, such as to minimize the frequency of the  $m_F = -1$  transition, which is taken as the reference.

In this manner the closest neighboring transition falls 600 MHz away from the reference. A less careful and cheaper beam design can then be used with a wide beam for high atomic flux, without running risks usually associated with neighboring transitions, e.g. Rabi and Ramsey pulling and Majorana transitions.

On the other hand the uniformity and stability required of the magnetic field are higher than in classical clocks. In fact, to achieve an accuracy of the order of  $10^{-14}$ , relative uniformity measurements must have an uncertainty smaller than  $10^{-6}$ .

Moreover the stability of the magnetic field profile must be of the same order of magnitude, and its uniformity cannot be much worse. In the prototype, a multilambda cylindrical Rabi cavity is used, which

resonates in the  $TE_{017}$  mode. This mode has a very flat phase around the symmetry axis, and allows opening up the beam passing holes. A diameter of 8mm has been used, with a useful beam section 3 times larger than in conventional solutions. In this manner the atomic flux is maximized.

Furthermore, in order to increase line Q and decrease second order Doppler, a scheme with bipolar permanent magnets was planned, which have a wide useful area and then permit to obtain an intense atomic flux.

**The bipolar gradient magnets**

Different magnetic state selection configurations can be utilized in Cs beam tubes depending on the shape of the atomic beam. For rotationally symmetric atomic flux, multipole magnets (quadrupole, exapole) are used. However, in this kind of magnets, the useful area is small and the low signal-to-noise ratio limits the short term stability notwithstanding their focusing effect.

In other beams, different solutions have been used, like the two wire magnetic field configuration which was proposed in [2] and used in many tubes.

On the other hand, the double confocal conical state selection magnet [3] [4] has a wider useful area between the polar expansions.

The gradient magnet profile used in our prototype is ideally derived from this latter configuration, but is asymmetric and intended for use with a single beam.

Because the deflection angle of the atomic trajectories is proportional to the gradient of the magnetic flux ( $F = \mu_{\text{eff}} \nabla B$ , where  $\mu_{\text{eff}}$  is the effective magnetic moment) it is necessary to evaluate the gradient value.

A rough evaluation of the deflection is easily obtained by considering that above 1T,  $\mu_{\text{eff}}$  can be approximated by the Bohr magneton.

A monovelocity beam of 100m/s is deflected 10mrad in a magnet 50 mm long with a gradient of 50 T/m normal to the beam.

To achieve this gradient value in the gap, a magnetic induction of the order of 1T was required. For this

purpose we used a permanent magnet in the Alnico family, which has a lower temperature sensitivity than other materials in the range 0 to 100°C.

This material has a maximum operational temperature of about 500°C. Because the vacuum chamber has a cylindrical shape, a circular flux return yoke was designed. The latter is made of soft iron with a saturation flux density of about 1.8T.

After this simple magnet structure analysis, we computed with a 2-D finite element code the value of the magnetic flux density in a plane normal to the beam.

To define its thickness attention was paid to local saturation in the return yoke. The outer diameter is limited by the vacuum chamber i.d. (60mm). A thickness of the flux return yoke of 5 mm was sufficient to achieve a flux density of about 1.2T.

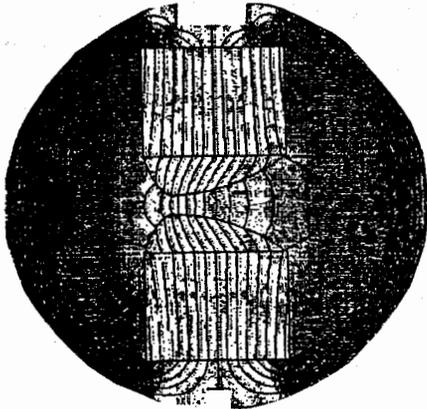


Fig.1 Pattern of the simulated flux density on the normal plane to the symmetry axis

In the same manner, with a trial and error procedure, we defined the pole piece profile. Two arcs of parabola (Fig.1) were adjusted to maximize both the useful area and the gradient value, always keeping the pole tips at the limit of saturation.

The pole piece profile designed in this way was fed to a numerical control machine with a mechanical tolerance of 0.1 mm. Magnetic induction measurements were performed on the manufactured magnets by a Hall probe with a resolution of about  $10^{-4}$ T.

In Fig.2 and Fig.3 the computed and measured magnetic flux density and gradient value along a line on the symmetry plane and normal to the beam direction are reported. No great deviations were observed. As it can be seen, the average value of about 50 T/m is obtained in a 6 mm wide area across the magnet center, which is taken as geometrical reference for the definition of the beam optics.

#### Beam optics

Using the information obtained in the measurements described above we simulated the atomic trajectories in

the Cs tube with the aim of defining the optics of the beam in the path from the oven to the hot wire detector.

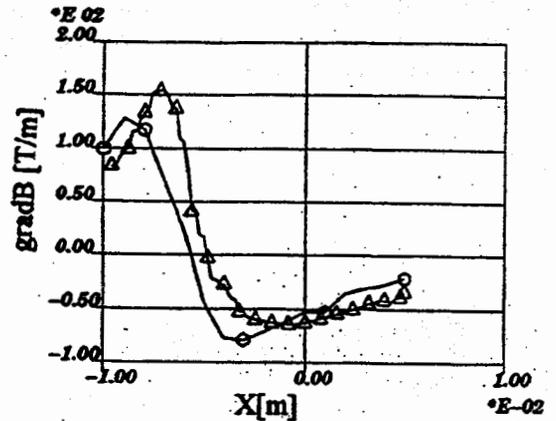


Fig.3 Computed (o) and measured ( $\Delta$ ) magnetic flux density gradient along a line normal symmetry plane of the gradient magnet.

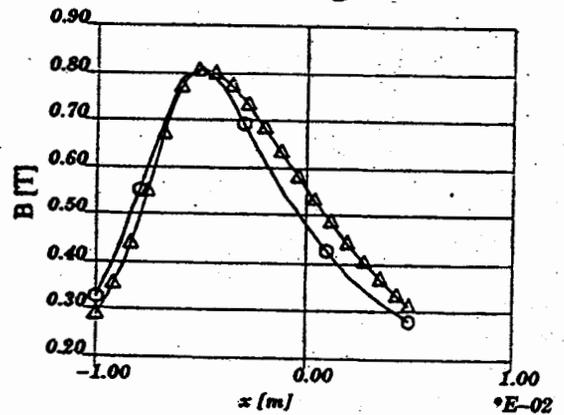


Fig.2 Computed (o) and measured ( $\Delta$ ) magnetic flux density along a line normal symmetry plane of the gradient magnet.

A FORTRAN code is realized for the computation of trajectories, and a fourth order Runge-Kutta integration method was necessary for the solution of the equation of motion in the pattern of the measured magnetic flux density on the symmetry plane of the magnet.

In the software code the mechanical iris, represented by the beam passing holes in the microwave cavity end caps, is taken into account.

In the same manner as before, a trial and error procedure was used, with the aim to find the best position of oven and detector in order to minimize the number of atoms which reach the detector in the non-desired level.

As shown in Fig.4, a full off axis scheme was used with an offset angle in the beam starting direction of about 30mrad. The divergence of the beam was fixed to 10mrad. The reason to limit the analysis to this

aperture was the fact that we were planning to use a microchannel array with 3mm long and 10 $\mu$ m i.d. channels. This has not been done as yet.

In the prototype, the oven to detector path length is about 1.2 meter and the hot wire detector used is 4 mm wide, with a useful area of the order of 24mm<sup>2</sup>.

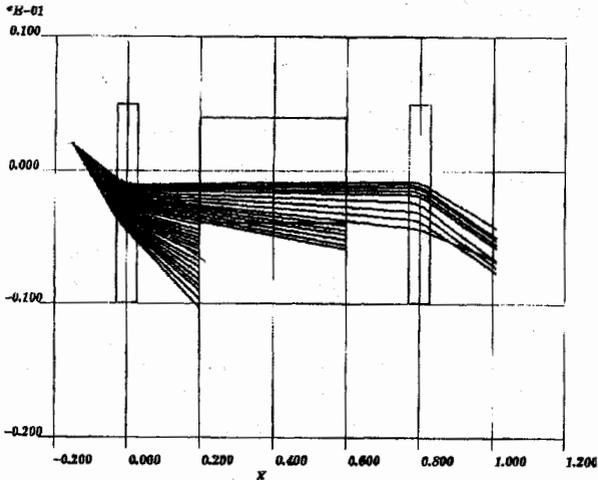


Fig.4 Atomic trajectory simulations for a monokinetic beam at 100m/s.

As shown in Fig.4, atoms in the selected level are deflected by the A-magnet toward the cylindrical cavity, where they interact with the microwaves.

The B-magnet then selects the atoms which did the transition, by deflecting them toward the hot wire detector. The latter is placed -6mm off axis.

In this manner it was possible to evaluate the transmission efficiency. This is of the order of 10% for atoms with a velocity between 75 m/s and 120 m/s. Consequently, an atomic beam of about 10<sup>9</sup> atoms/s was expected in the prototype.

#### TOF measurements with selected velocity

As illustrated in [5] a first atomic resonance was measured but the signal to noise ratio was awful. We wrote the fault to the analyzing B-magnet, because of an insufficient separation between the two beams, which were detected by the hot wire.

Therefore, we turned the attention to the gradient magnets, with the aim of evaluating the real deflection effect on the atoms. We measured the velocity distribution of the atoms with a time-of-flight (TOF) method.

For that purpose, we took out of the vacuum chamber the B-magnet and we pulsed on the atomic beam before the A-magnet a laser tuned to the  $F=3$   $F'=4$  pumping transition.

In this manner the atomic population of the selected level is increased for 1ms and we observe the TOF for 28ms, which is the chopping period. The experimental setup is shown in Fig.5: the laser diode (LD) has an

extended cavity configuration and is frequency stabilized on saturated absorption.

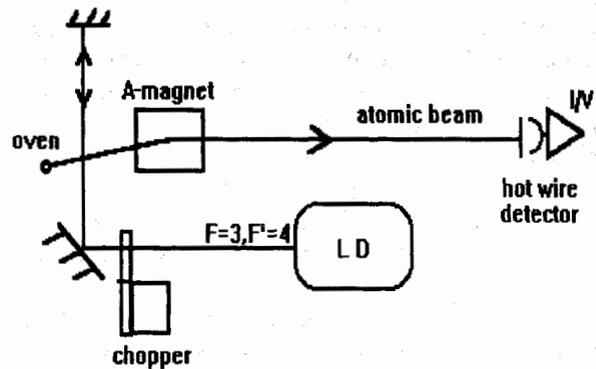


Fig.5 Experimental setup for TOF measurements

It turned out that a selection effect is achieved by the magnet over the full maxwellian velocity distribution. As shown in Fig.6, the spatial filter is centered around 130 m/s and falls gradually to zero toward the high velocity atoms.

In good accord with the resonance signal measured one year ago, the atomic intensity has its peak value centered around 200 m/s. In fact, the detected velocity distribution is the product of the spatial filter and the full maxwellian distribution.

Following that, we prepared an experiment with the purpose of measuring the deflection of the atomic trajectories.

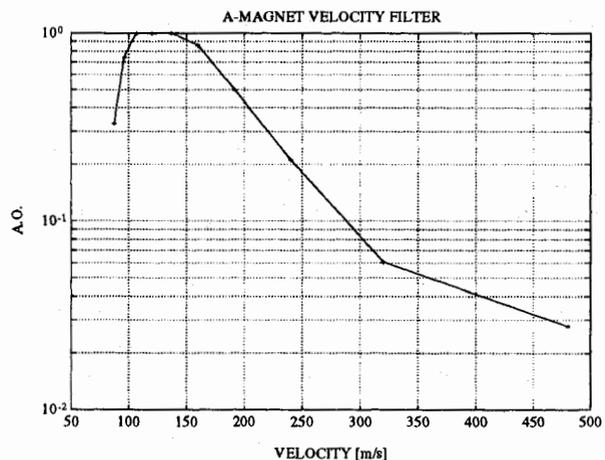


Fig.6 Spatial filter over the full maxwellian velocity distribution.

A small vacuum apparatus with a turbomolecular pump was adapted with a long flexible tube with the purpose of moving the hot wire detector sideways. The hot wire detector was mounted on a vacuum flange fixed on a translation stage.

A low noise FET transconductance amplifier converts the Cesium ion current into a voltage which is then measured with a multimeter. In this manner it is

possible to measure the beam shape over a 50mm range.

In Fig.7 the atomic beam without the gradient magnet is shown to have a trapezoidal shape. It is not clear why in the middle of the smaller base we find a lower atomic intensity.

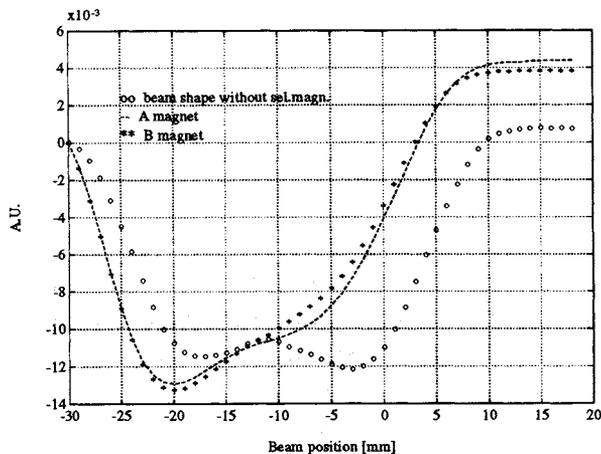


Fig.7 Beam shape with and without the A and B gradient magnets

An obvious deflection effect on the atomic trajectories was found, as it can be observed in Fig.7, by comparing the deflected beam with the undeflected one. However, several questions arise from this measurement because, in this simple Stern-Gerlach experiment, the beam of atoms in the other level was not detected.

### Conclusions and work in progress

We are presently considering the possibility of replacing the existing permanent magnets with stronger ones (i.e. rare earth as SmCo). In this manner a higher deflecting effect and, at the same time, a narrow spatial filter can be realized.

In the meantime, the existing machine will be operated in the high C-field mode in an optically pumped configuration. Solving the problem of efficient deflection of a wide beam is still considered important for the filtering effect in the velocity space that this method provides.

### Acknowledgments

The authors wish to acknowledge discussions and invaluable support by many colleagues, in particular R.E. Drullinger of NIST and G.D. Rovera of LPTF. A special thank goes to S.Thérisod, who realized the laser frequency stabilization system used for the TOF measurements.

### References

- [1] A. De Marchi, "The high C-field concept for an accurate Cesium beam resonator", Proc. 7<sup>th</sup> EFTF, 1993
- [2] N.F. Ramsey, "Molecular beams", Clarendon press, 1956
- [3] P. Kartashoff, P.E. Debely, "Cesium beam tubes using confocal conical state selectors", Proc. 8<sup>th</sup> Int. Congress Chronom., 1969
- [4] C.Xue, "Newly designed dipole beam optics system for Cesium beam frequency standards", IEEE Trans. on Instrum. & Measurements, vol. IM-29, 4, pp.301-304, Dec. 1980
- [5] G.Costanzo, E.Rubiola, A.De Marchi, "First results on the high C-field Cesium resonator at the Politecnico di Torino", Proc. 9<sup>th</sup> EFTF, 1995

## OPTIMIZATION OF PASSIVE HYDROGEN STANDARD FREQUENCY INSTABILITY

*S.A.Kozlov, V.A.Logachev*

Institute of Electronic Measurements "Kvarz",  
Prospect Gagarina, 176, Nizhny Novgorod, Russia

The frequency instability of the passive hydrogen standard due to quantum hydrogen discriminator, receiver thermal noises and crystal oscillator phase fluctuations near frequency modulation harmonics of hydrogen maser interrogation signal is analyzed.

The passive hydrogen standard block-diagram with single phase modulation interrogation signal of quantum hydrogen discriminator is considered at the condition that phase modulation frequency  $\Omega_m$  much more than hydrogen spectral line halfwidth  $\gamma$ . This standard has two automatic frequency control loops: for crystal oscillator and for microwave cavity.

In comparison to results in [1] the new dependences of frequency instability via interrogation signal parameters are defined.

Frequency instability minimum for passive hydrogen standard is at phase modulation index  $m = 2.0 - 2.2$ , saturation parameter  $S_o = 0.3 - 1.0$  and is defined generally by noise in automatic frequency control loop of crystal oscillator. Automatic frequency control loop of cavity is going to influence at  $m < 2$  and than more than less parameter  $S_o$ .

Dependence of hydrogen passive standard frequency instability via interrogation signal modulation frequency is weak for  $0,2\beta < \Omega_m < 0,5\beta$ , where  $\beta$  is halfwidthband of cavity.

It is shown that:

- crystall oscillator phase fluctuations on even harmonics of modulation frequency (but not frequency fluctuations as for another types of passive frequency standards [2]) have influence on passive hydrogen standard frequency instability;

- this influence two orders as less than for another types of passive standards and for usual phase noise level of crystal oscillators  $S_o = -150$  dBc/Hz two orders as less than influence of hydrogen maser and the receiver thermal noises.

The results are agreed with the experimental facts.

### Reference

- [1] J.Vanier and C.Audoin, The Quantum Physics of Atomic Frequency Standards, Bristol and Philadelphia, 1989.
- [2] C.Audoin, V. Candelier and N.Dimarq. A limit to the frequency stability of passive frequency standards due to an intermodulation effect. - IEEE Trans. on Instr. and Measur., vol.40, No 2, April 1991.

# EVALUATION OF MINI-BVA RESONATORS FOR SPACE APPLICATIONS : TESTS PROGRAM AND RESULTS

**J.J. BOY, R.J. BESSON, N. FRANQUET, D. THIEBAUD**  
ENSMM / LCEP - 26, chemin de l'Épitaphe - 25030 BESANÇON Cedex - FRANCE

**R. PETIT**  
DGA / CELAR - 35998 RENNES ARMEES - FRANCE

**Ph. BAVIERE, M. BRUNET**  
CNES - Av. Edouard Belin - 31400 TOULOUSE - FRANCE

## ABSTRACT

During last twenty years, there has been a growing need of ultrastable quartz oscillators operating under special environmental conditions usually required in space program applications [1]. This is the case for radiocommunications, navigation, precise positioning and scientific program especially in the case of satellite applications.

Under those conditions an evaluation program based on general rules ESA/SCC has been launched in 1992 for precise space evaluation of mini-BVA resonators. A special Process Identification Document (P.I.D.) has been established which includes fabrication processes and tests procedures under guidance of Centre National d'Études Spatiales (CNES) division of components quality. This first required to define wishable performances for Délégation Générale pour l'Armement (DGA) and CNES and to take into account space military telecommunications programs goals and the Doris program (precise orbitography and positioning system) with an accuracy of some centimeters.

21 resonators have been fabricated according to proposed P.I.D. Then a test program has been defined and resonators have been space evaluated with interesting results which will be discussed. Finally, it appeared that performances goals have been obtained. In conclusion special resonators can be fabricated by LCEP with guaranteed performances. LCEP has been qualified by CNES and is now capable to deliver resonators according to P.I.D. approved by CNES, for future DGA/CNES space programs.

## I. INTRODUCTION

In the general frame of the evolution of space program applications, BVA resonators (which means "in a box with improved ageing") represented an attempt to obtain the better performances in short term and long term stabilities and in external parameters sensitivities (acceleration barometry, magnetic field or radiations exposure) as well.

Meanwhile, though several BVA resonators have been used in space programs (in RFA or USA), BVA technology had never been evaluated in details.

So, an evaluation program based on general rules ESA/SCC has been launched in 1992 for precise space evaluation of the mini-BVA resonators under guidance of CNES and DGA for which future needs required time and frequency systems more and more stable under space programs conditions.

Indeed, the "in-demand" stabilities in the case of radiocommunications, navigation or precise positioning are of the order of  $1 \cdot 10^{-13}$  in short and medium terms and a few  $10^{-11}$  by day, including ageing and, eventually, frequency shift due to radiation exposures.

A piezoelectric resonator, which is a cavity for electroacoustic waves in a solid of given shape, configuration and dimensions, can be used for frequency and time applications. In this case, quartz resonators have been built so that their frequencies are as independent as possible from external parameters. Furthermore, though important advances have been achieved in quartz quality material, material characterisation and then selection of the blanks is absolutely necessary to avoid, as far as possible, the influence of main quartz crystal defects (chemical impurities or dislocations, for instance).

## II. BVA TECHNOLOGIES AND TACTICAL BVA RESONATOR

In this context, BVA technology includes improvements to obtain, through new technologies (for instance ultrasonics machining or chemical etching) monolithic, highly reproducible structures eventually using electrodeless techniques as in the BVA<sub>2</sub> type.

This resonator basically corresponds to the main following items :

*An electrodeless design* : all problems that relate to electrode deposition, such as damping, stresses, contamination and ion migration disappear.

*A crystal mounting made of quartz* : small "bridges" connect the vibrating part of the crystal to the dormant part to avoid discontinuities or stresses in the mounting points

Tactical BVA resonators have been progressively developed [2, 3, 4, 5, 6] using a special symmetrical mounting according to Fig. 1 :

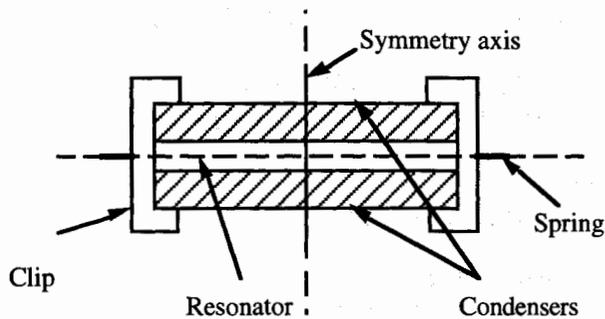


Fig. 1 : diagram of the mini-BVA (called *bva2*)

The crystal is miniaturized (*bva* resonators), and the *bva2* is a 10 MHz SC cut third overtone.

The two condensers and the vibrating part are strongly clipped together and the clips are mounted in a groove inside a steel cylinder exactly in its middle by use of four springs. Mechanical parameters of clips and springs are carefully determined and checked.

### III. THE EVALUATION PROGRAM

#### III.1. General conditions :

The general conditions (traceability, screening, delivery, ...) for qualification and acceptance of a quartz resonator flight model are defined in the specification ESA/SCC n° 3501. Furthermore, a detailed specification based on ESA/SCC 3501/003 has been defined by LCEP and DGA/CNES at the beginning of the program to be completed at the end according to obtained results during evaluation.

This evaluation includes two phases :

**Ph.1** : writing of the P.I.D., fabrication of 21 resonators, final production tests and burn-in,

**Ph.2** : Selection of 10 units and limit evaluation tests.

#### III.2. Physical specifications :

All fabrication parameters have been defined in a detailed fabrication sheet. Including also fabrication processes and tests procedures, a special Process Identification Document (P.I.D.) has been established by LCEP with the help of CNES, so that each detail of construction is documented.

For instance, this document includes :

\* detailed proceeding to obtain a great precision in the crystalline orientation of the three parts of the "sandwich" which is better than 15" for the two cut angles ( $\varphi$  and  $\theta$ )

\* details on ultrasonic machining used to make quartz bridges in the vibrating part of the resonator : our procedure can guarantee the orientation of these 4 bridges better than  $\pm 1^\circ$  in respect to the crystallographic axes.

\* precise definition of the gap depth (accuracy better than  $1 \mu\text{m}$ ) which is the distance between the vibrating part and the "condenser electrode-holder".

Finally, the good definition and reproducibility of these mechanical parameters have to insure a quasi-perfect symmetry of the entire structure around three perpendicular axes (X, Y and Z).

#### III.3. Electrical specifications :

In the general frames of this evaluation program, it has been necessary to study the evolution of the main electrical parameters after some mechanical tests as sinus and random vibrations, shocks, acceleration, rapide changes of temp. (final production tests) and burn-in.

Furthermore, in a second phase of the program, we have tested units at limit of destruction for each mechanical test (vibrations, shocks and acceleration). All these results are summarized in the paragraph.

The first table presents the applicable specifications extracted from the ESA/SCC detail specification N° 3501/003 which is named table 1. So, it indicates the parameter deviations for the entire batch of the 21 resonators to be manufactured according proposed P.I.D.

Table 1 : Detail specification

N°	Characteristic	Sym	Min	Max	Unit
1	Nom. frequency	$f_r$	9.999983		MHz
2	Reference tpre	$T_0$	74	82	°C
3	Overtone order	-	3		
4	Load capacitance	$C_L$	20	50	pF
5	Drive Level	$P_0$	0.05	0.2	mW
6	Freq. Adjust.	$\Delta f/f$	-1	+0.6	$10^{-6}$
7	Res. Resistance	$R_r$		100	$\Omega$
9	$R_r$ Var. with tpre	$\Delta R/R$	-10	+10	%
10	Oper. tpre range	$T_{00}$	-40	+85	°C
16	Q factor	Q	$10^6$		
20	Stor. tpre range	$T_{STR}$	-55	125	°C

NB : Other electrical parameters as motional inductance ( $L_1$ ), motional capacitance ( $C_1$ ) and static capacitance ( $C_0$ ) have to be defined and are presented in the next paragraph.

Table 2 presented below indicates the *guaranteed* specifications (defined at the beginning of the evaluation program) of some characteristics measured after each environmental tests. If the above electrical parameters can be measured in a conventional  $\pi$ -network, the other ones have to be checked in a appropriate oscillator whose performances largely depend on available components and technologies, particularly for the measurement of short term stability.

For each characteristic, the chosen methods are rapidly described below.

Table 2 : Electrical parameters (21 to 26)

N°	Characteristics	Symbol	Specifications	Unit	Remarks
21	Short term stability $\tau = 10$ s	S.T.	5	$10^{-13}$	Allan Variance
22	Ageing : a. per day b. per month c. per year	L.T.	3 10 5	$10^{-11}$ $10^{-10}$ $10^{-9}$	After 20 days of continuous operation
23	Accelerometric sensitivity	A.S.	10	$10^{-10}$ / g	All axes  Γ
24	Barometric sensitivity	B.S.	5	$10^{-10}$ / bar	range : 0 - 2 bar
25	Magnetic sensitivity	M.S.	5	$10^{-12}$ / gauss	range : 0, ± 4 gauss
26	Radiation sensitivity	R.S.	5	$10^{-11}$ / rad	Accumulated dose : 10 krad Max rate : 0.1 rad / mn

#### IV. FINAL PRODUCTION TESTS AND BURN-IN : PHASE 1

According to the P.I.D. formally approved by CNES and DGA in January 1994, 51 resonators issued from 5 series have been fabricated and accepted by CNES after precap. Each serie corresponds to a particular quartz material origin. Each quartz block has been qualified according to our own procedure to quantify the impurities and dislocations contents and to verify the quality of the sweeping.

21 resonators have been finally chosen and tested as indicated below :

- \* electrical parameters measurements (Table 1 and table 2, except "ageing")
- \* environmental tests (see below table 3)
- \* final electrical parameters (Tables 1 and 2)
- \* Burn-in (85 °C during 240 Hours)
- \* tables 1 and 2.

Table 3 : environmental conditions

N°	Environmental tests
1	<b>Shocks</b> : 50 g, 6 ms, 3 axes 1/2 sinus, 1 shock per axis
2	<b>Random vibrations</b> : 1 min / axis, 3 axes 20 - 100 Hz , +6 dB/oct 100 - 600 Hz , 0.08 g <sup>2</sup> /Hz 600 - 2000 Hz , -6 dB/oct
3	<b>Sine vibrations</b> : 4 oct/min, 3 axes 5 - 21 Hz , 11 mm 0-c 21 - 2000 Hz , 10 g
4	<b>High temp. storage</b> : 85 °C, 240 H
5	<b>Change of temperature</b> : 2 cycles Min : -55 °C Max : 125 °C Temp. variation : 2 to 5 °C/min Duration of exposures : 30 min
<b>Sanctions</b> : at T <sub>0</sub> $\Delta F/F \leq \pm 2.10^{-6}$ and $\Delta R/R \leq \pm 10\%$	

#### IV.1. Electrical parameters and frequency-temperature characteristics :

The resonators are tested under temperature up to 95 °C during a "step-by-step" characteristic : with a desk computer, we check the phase difference between the input and the output levels on a  $\pi$ -bridge (according to

the IEC 444 method) within which the quartz resonator is inserted. We store the resonance frequency and the motional resistance. first **just after bake-out** (MPE1) and second a few **weeks later after thermal cycles** (MPE2).

For each characteristic, the three coefficients of the best fitted third degree polynomia as :

$$\frac{\Delta F}{F} = a.\Delta T + b.\Delta T^2 + c.\Delta T^3$$

are calculated and stored.

The frequency is only stored when a given temperature of the oven is stabilized within 0.1°C (and checked by a quartz thermometer put close to the resonator under tests).

During measurements, we can estimate the stability of the oven is better than  $\pm 0.05^\circ\text{C}$  and consequently we evaluate the error on the T.O. point at  $\pm 0.2^\circ\text{C}$ . The other performances of our measurement system are, in the worst case :

Frequency :  $\pm 5.10^{-9}$

Motional resistance : 1%

Q-factor : 2%

The following figure presents an example of the temperature-frequency curve plotted for each resonator of the batch after mechanical tests or burn-in.

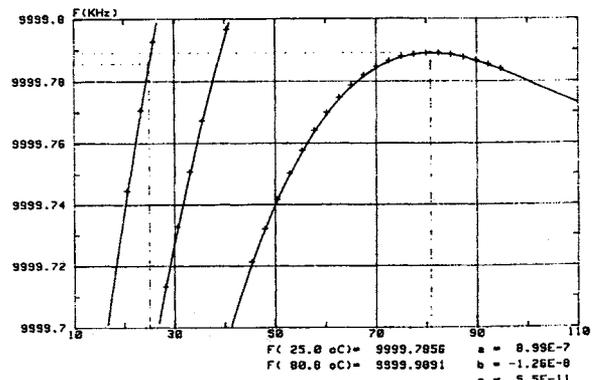


Fig. 2 : Frequency versus temperature  
+ : experimental points - : best fit

So, as to better appreciate the evolution of electrical parameters, we present here a table which indicates their mean values and standard deviations (which are in brackets) :

1. just after back-out (MPE1)
2. after thermal cycles (MPE2)

- 3. after mechanical tests
- 4. after burn-in (and about 1 year after back-out)

**Table 4 : electrical parameters evolution**

	1	2	3	4
10 MHz - ΔF (in Hz)	16.8 (7.7)	15.9 (8.7)	15.6 (8.8)	15.2 (8.8)
Turn-over point: T <sub>0</sub> (°C)	79.4 (2.1)	78.6 (2.7)	79.2 (2.5)	79.2 (2.5)
Motional resistance : R <sub>T</sub>	96.2 (2.1)	97.1 (2.7)	99.2 (1.9)	99.5 (2.1)
Q (* 10 <sup>6</sup> )	1.3 (± 0.02)			
L <sub>1</sub> , C <sub>1</sub>	2 H and 0.16 fF			

This table shows that frequency increases (particularly after thermal cycles which appear as indispensable to stabilize the resonator) whereas the turn-over point is slightly stable. At least, we prove that the environmental tests have no particular incidence on the resonance frequency and the motional resistance.

**IV.2. Short term stability and ageing :**

Though the short term stability of a quartz oscillator is mainly determined by additional noise of sustaining circuit, we prefer, to save time, to measure the power spectral density of the phase fluctuations (Sφ(f)) of

a 10 MHz oscillator. Then, a software, developed by LCEP and called "Phi-noise" allows to plot the entire representation of the real data recorded and gives the corresponding σ<sub>y</sub>(τ) for 10 second.

As the end of this evaluation program phase 1, the mean value of the 21 Allan variances has been :

$$3.1 \pm 2.1 * 10^{-13} \text{ for } \tau = 10 \text{ seconds.}$$

As the long term drift, measured in the same oscillators than previous ones, it is almost positive (except for 2 resonators) and consistent with the guaranteed specification presented in table 2.

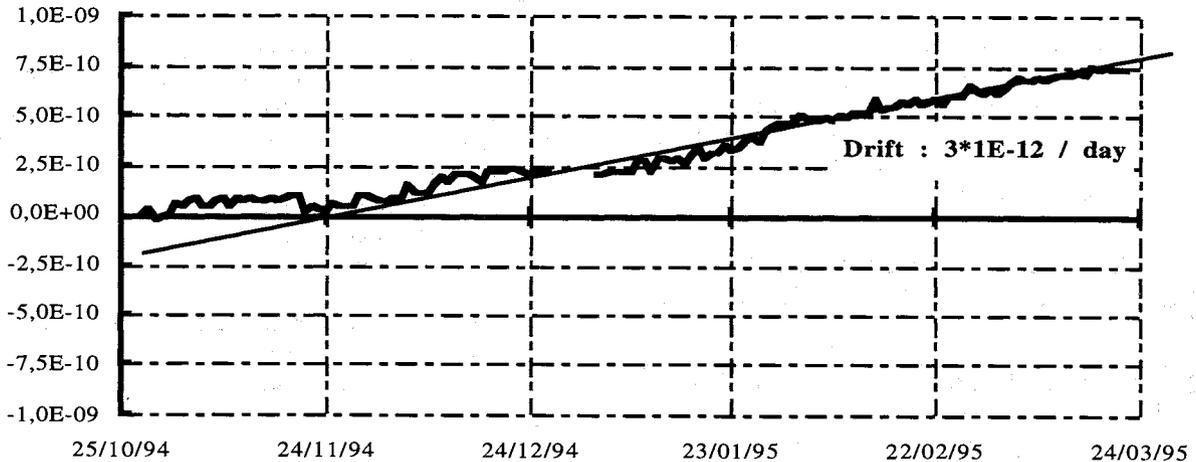
Furthermore, half the resonators of the batch exhibits ageings better than

$$1 * 10^{-11} \text{ per day,}$$

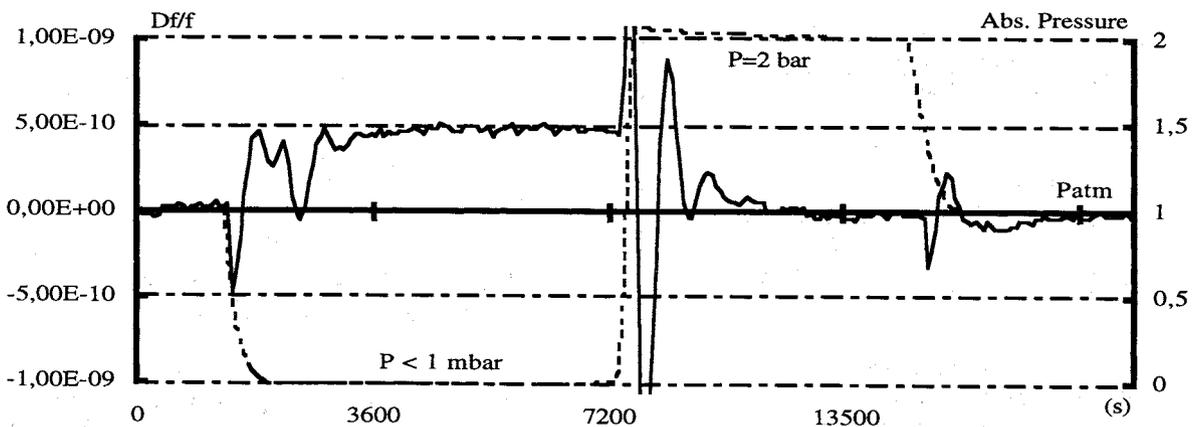
calculated by linear regression over 30 days at the end of the phase 1 (See example in figure 3).

**IV.3. Barometric sensitivity :**

The same oscillators have been used to estimate the barometric sensitivity of each resonator. The figure 4 presents a typical plot of the recorded data during a 4-hours test. All the units are consistent to the guaranteed specification, i.e. better than 5\*10<sup>-10</sup> / bar.



**Fig. 3 : Long term drift on 30.03 resonator**



**Fig. 4 : Barometric sensitivity in the range [0, 2 bar]**

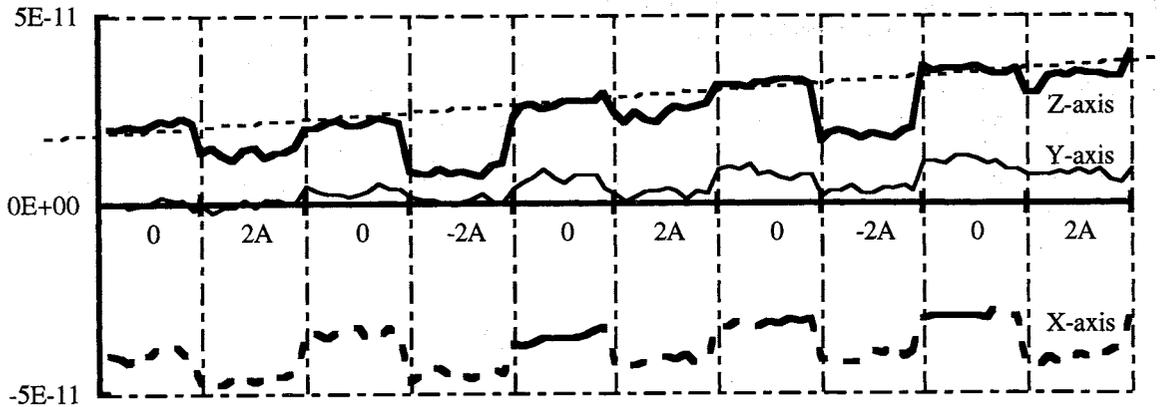


Fig. 5 : Typical measurement of the magnetic sensitivity

#### IV.4. Magnetic sensitivity :

As it has been presented in [7], the frequency variations due to a magnetic field is measured with an oscillator which for the resonator is deposed of about 20 cm. So, only the resonator and its oven are in the center of three pairs of Helmholtz coils allowing to insure a magnetic field application in three perpendicular directions. The electronics, protected by a thin shielding of *Mumetal*, is quasi-insensitive to the magnetic field which is created by a loop of current. The current power is driven by a computer which can inverse its polarity in the coils. A 2A current leads to a measured magnetic field in the center of the Helmholtz coils of :

- || 8.8 Gauss along X
- || 8 Gauss along Y
- || 10 Gauss along Z.

We store 10 difference frequencies by step, the magnetic field being applied on each axis following the cycle :

0 / 2A / 0 / -2A / ... performed as indicated in Fig. 5. The mean magnetic sensitivity is also calculated as the modulus of a vector in the reference frame X, Y, Z.

Our main results are presented in the table 5 which for we indicate a maximum of magnetic sensitivity.

Table 5 : Magnetic sensitivity of a few units

	X	Y	Z	$\Delta f/f$ ( $10^{-12}/\text{gauss}$ )
30.01	1	1.2	1	1.8
30.03	2.1	2.5	1.8	3.5
30.12	2	2	2	3.4
31.03	2	1.9	0.8	2.9
31.06	2	2.2	1.2	3.1
31.08	3.2	2.6	1	4.2
33.05	2	2.5	2	3.7
33.06	1.6	1.9	1.8	3.0
34.03	1.2	1.5	0.5	2.0
34.06	2.2	2.2	1.6	3.5
Mean : $3.1 \cdot 10^{-12}/\text{Gauss}$				

#### IV.5. Accelerometric sensitivity :

Here too, we will recall the results presented last year in [7] just completed by the mean values obtained for the all resonators of the batch.

As for the other electric parameters, it has been necessary to repeat the measurements according to the following sequence :

1. after thermal cycles : under random vibrations in the range [20 - 100 Hz]
2. after mechanical tests : 2G-tipover test
3. after burn-in : 2G-tipover too.

Table 6 : accelerometric sensitivities (mean values and standard deviations)

$\Delta F/F$ in $10^{-10} / g$	1	2	3
Along X	0.93 (0.3)	0.65 (0.4)	0.46 (0.3)
Along Y	1.02 (0.6)	0.71 (0.4)	0.41 (0.3)
Along Z	0.97 (0.5)	0.82 (0.5)	0.78 (0.3)
$\Gamma = \sqrt{\gamma_x^2 + \gamma_y^2 + \gamma_z^2}$	1.77 (0.6)	1.26 (0.5)	0.99 (0.4)

The previous table shows that the 2G-tipover test is more "optimistic" than the other one. This difference can be due to two reasons :

- the measurements under vibrations are realized with an oscillator not really made for this test. Indeed, the step level between 20 and 100 Hz increases at the high frequencies. So, the mean value is slightly higher than in a really correct test.
- the very low sensitivities are measurable with difficulty on the phase noise spectrum (the flicker noise is not sometimes enough low).

Although it is, all these sensitivities are largely better than the guaranteed specifications (less than  $2.10^{-10} / g$ ).

#### V. LIMIT EVALUATION TESTS : PHASE 2

After this first phase of the evaluation program, half resonators have been submitted to severe conditions, according to the next table (Table 7). After each item of

the level 1, one unit has been tested under vibrations and shocks at limit of destruction. Our design is now capable of standing, at least, the following mechanical tests :

\* **sinusoidal vibrations** of 50 g (20 Hz - 2,000 Hz) during 15 minutes per axis on 3 axes after standing same treatment at 20 g, 30 g and 40 g.

\* **random vibrations** : 3 units have been measured after the following conditions :

20 - 100 Hz , +6 dB/oct  
100 - 600 Hz , 0.2 g<sup>2</sup> / Hz  
600 - 2000 Hz , -6 dB/oct

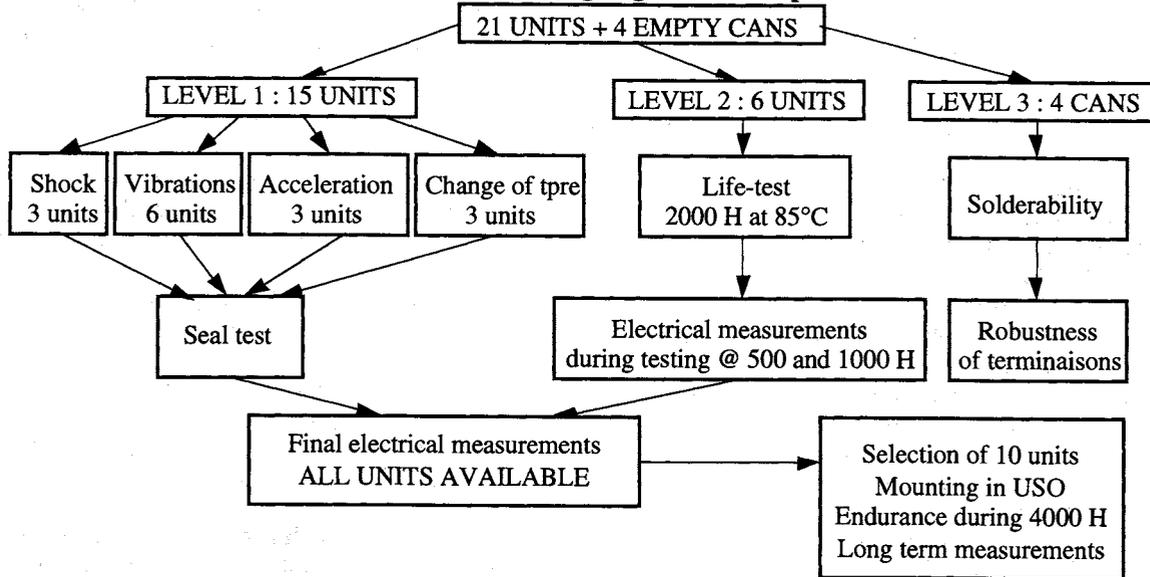
Furthermore, one unit has been tested with success at :

**1.0 g<sup>2</sup> / Hz.**

\* **Shocks** : 1 unit has stood without damages shocks of 500 g during 2 ms on all 3 axes, in each sense.

\* **Constant acceleration** : 3 units are consistent to the sanctions after an acceleration of 20 g (3 axes) during 3 min per axis. For one resonator the acceleration was of **100 g.**

**Table 7 : Organigram of the phase 2**



## VI. CONCLUSION

We show that the performances goals have been obtained for 21 resonators fabricated according to a P.I.D. formally approved by CNES and DGA. Furthermore, results are largely better than the guaranteed specifications. We summarize in the last table the mean values obtained :

Characteristics	Results	Unit
Short term stability $\tau = 10$ s	3	10 <sup>-13</sup>
Ageing :		
a. per day	1	10 <sup>-11</sup>
b. per month	3	10 <sup>-10</sup>
Accelerometric sensitivity	1 worst axis	10 <sup>-10</sup> / g
Barometric sensitivity	1	10 <sup>-10</sup> / bar
Magnetic sensitivity	3 worst axis	10 <sup>-12</sup> / gauss

The irradiation tests will performed by CNES this year. We will test two units in oscillator which for the resonator (and its oven) is deported of about 20 cm. The goal is 1\*10<sup>-11</sup>/rad with a dose rate of 1 rad/H.

The previous table (and the results obtained during the limit evaluation tests) indicates that all conditions are fulfilled with a large success during this

2.5 years evaluation program. All the problems are resolved in a very satisfactory way. The manufacturing and measurements procedures are available so that the technology of the *bva*<sub>2</sub> SC-cut 10 MHz 3rd Overtone is actually in position to answer to future needs for space programs.

## REFERENCES

- [1]. R.J. BESSON, J.J. BOY, M. MOUREY : "BVA resonators and oscillators : a review. Relation with space requirements and quartz material characterisation", 49th Ann. Symp. Freq. Cont. (1995)
- [2]. R.J. BESSON, R. DELAITE and Al., French patent n°8509097 (June 1985).
- [3]. R.J. BESSON, R. DELAITE, G. RENARD and Al., French patent n° 8814197 (Oct. 1988).
- [4]. J.J. BOY, A. BERTHAUT, G. RENARD and Al. : "Characterisation of BVA in industrialization process", 7th Europ. Freq. Time Forum (1993), pp. 57, 62.
- [5]. M. MOUREY, R.J. BESSON : "A space oscillator with cylindrical oven and symmetry", 45th Ann. Symp. Freq. Cont. (1991), pp. 431, 441.
- [6]. J.R. NORTON, R.J. BESSON : "Tactical BVA quartz resonator performance", 47th Ann. Symp. Freq. Cont. (1993).
- [7]. J.J. BOY, F. DEYZAC, A. BERTHAUT : "Accelerometric sensitivities of BVA resonators", 9th Europ. Freq. Time Forum (1995), pp. 9, 15.

# COMPARISON OF PASSIVE AND ACTIVE AGING OF SC-CUT AND AT-CUT CRYSTALS

Bernd W. Neubig  
TELE QUARZ GROUP  
D-74924 Neckarbischofsheim / Germany

## 1. ABSTRACT

This paper reports on the results of pre-aging of four different types of OCXO crystal units. It compares the aging behaviour for a passive aging at +80°C over 28 days and 12 days at different drive levels and the active aging in the following 34 days with the crystals continuously operating in oscillators. For the aging rates good correlation was found. One crystal of each group was additionally aged for 240 days. The measured long-term aging results are compared to the aging prediction derived from the first 30 days.

## 2. INTRODUCTION

Pre-aging of OCXO crystal units is commonly done in order to remove the initial aging and to verify the long-term stability of the device. This is either done in test oscillators or in the final oscillator unit as a so-called active pre-aging, where the crystal is vibrating continuously and the output frequency is observed over time.

In the passive aging method the crystal unit is measured periodically in a suitable measurement system, and thus the crystal is not "active" most of the time and is only vibrating for a short instance during the measurement. The advantages of the passive pre-aging method are:

- only the crystal itself (without the influence of a test oscillator) is measured.
- the crystal behavior over time can not only be characterized by the resonance frequency but also by the resistance and - if necessary - the other crystal parameters.
- the crystal can be operated at arbitrary drive levels.
- other effects of the crystal, e.g. drive level dependency (DLD), which can cause irregular aging, can be easily detected.
- no soldering and de-soldering of the crystal (source of reliability risks) is necessary.

Common measurement systems do not have sufficient accuracy required for passive aging tests of high precision units. The micro-bridge technique developed by Erich Hafner [8] in connection with a precision temperature chamber with mK-stability was meanwhile improved to such an extent, that this system (XOTEX QXMS-A) allows an accuracy and reproducibility of parts in  $10^{-10}$  for the frequency measurement and of  $10^{-3}$  for the motional crystal parameters.

The focus of our experiments was to find out, if and how the results of passive pre-aging correlate with those of active aging and if irregularities in the active aging can be identified also in the passive pre-aging. From previous publications such as [10]

and [13] only very few data are available about such a correlation.

## 3. EVALUATION METHODS FOR AGING DATA

The aging of quartz crystal units as seen by the manufacturer can be subdivided into three intervals:

- stabilization period or initial aging,
- aging test period and
- extrapolated period.

The stabilization period starts after power-on, when the OCXO / resonator has reached its operation temperature equilibrium and lasts about one or a few days. This is the period of initial aging, which is determined by physical processes with shorter "time constants". The initial aging rate is usually stronger than the aging rate observed later, and the slope of the initial aging rate may be positive or negative, and does not necessarily correlate with the longer-term aging rate observed thereafter.

In a production environment the aging test period usually lasts several weeks. A logarithmic shape of the frequency vs. time aging curve is expected in well-behaved crystal units, and it is also observed quite frequently, when one aging process clearly dominates over the others.

Several different aging functions are reported in literature, which stand for different aging mechanisms. Some of them are purely mathematical approaches without any assignment to physical processes. Very often linear combinations of functions are used in order to consider more than one aging mechanism. The most commonly used functions for the frequency change over time are:

- pure log function [5],[6],[11],[19]
  - (1)  $\Delta f/f_0 = A_0 + A_1 \log t$  or
  - (2)  $\Delta f/f_0 = A_0 + A_1 \ln t$  where  $t \geq 1$
- modified log function [3],[4],[5],[9],[11],[20],[22]
  - (3)  $\Delta f/f_0 = A_0 + A_1 \log (1 + A_2 t)$

This function - which is also used in MIL-C-49468 [25] and MIL-O-55310 [26] - is the most frequently used approach.
- exponential function [5],[11]
  - (4)  $\Delta f/f_0 = A_0 + A_1 (1 - \exp(-A_2 t))$
- polynomial functions [5],[11],[22]
  - (5)  $\Delta f/f_0 = A_0 + A_1 t^{A_2}$
  - (6)  $\Delta f/f_0 = A_0 + A_1 (t - A_2)^N$  ( $N=0,5$  in [5],[11])
  - (7)  $\Delta f/f_0 = A_0 + A_1 t + A_2 t^{0,5}$
- Kalman filtering technique [12],[14],[16], a recursive computation based on a weighted sum of the modified log function
  - (8)  $\Delta f/f_0 |_{k+1} = \Delta f/f_0 |_k + \sum_j A_1^{j,k} \log \frac{1 + A_2 t_{k+1}}{1 + A_2 t_k}$

The coefficients  $A_1$ ,  $A_2$  may include other variables, which have an impact on aging, such as temperature or temperature gradients etc. [23].  $A_2$  in equations (3), (4) has the dimension  $s^{-1}$ , therefore  $A_2^{-1} = \tau_2$  stands for a time constant.

At TELE QUARZ we are currently using the modified log-function (3) and the polynomial function (7) and choose the significantly better fitting one (if comparable, we choose the modified log-fit).

It is common practice to derive the expected long-term aging behavior from the curve-fitted aging data in the test period by extrapolation. The aging rates per day, per month, per year are computed from the fitted curve for 30 days of operation.

#### 4. AGING TEST CONDITIONS

The aging tests were performed with 12 crystals of 4 different types of industrially manufactured AT-cut and SC-cut OCXO crystals (see table 1).

TABLE 1: Survey of crystals under test

freq. [MHz]	cut	over tone	enclo- sure	nom. $C_1$ [fF]	$P_Q$ [ $\mu$ W]	TOP [ $^{\circ}$ C]
16,384	AT	3rd	HC-27/U	2,0	50	85
16,384	SC	3rd	HC-27/U	0,35	200	78
10,0	AT	fund	HC-26/U	7,0	100	78
13,0	AT	fund	HC-26/U	7,0	10	75

The passive aging was done in an automatic test systems using the micro-bridge technique developed by Erich Hafner [8] in connection with a precision temperature chamber with mK-stability. This system (XOTEX model QXMS-A) allows an accuracy and reproducibility of parts in  $10^{-10}$  for the frequency measurement and of  $10^{-3}$  for the motional crystal parameters. The passive aging test was performed at  $+80^{\circ}$ C in two subsequent periods of 28 days and 12 days. While in the first period all crystals were measured with the same RF generator output level, in the second period the generator was set such that each crystal operated at its nominal drive level as indicated in table 1. The crystals were measured approximately every four hours and therefore vibrated only during the measurement time for a few minutes. In each measurement a complete set of crystal data, i.e. series resonance frequency  $f_s$ , series resonance resistance  $R_1$ , motional capacitance  $C_1$  etc. was determined and was stored together with the values of drive level, chamber temperature and date/time of measurement. Only frequency and resistance data were used in our evaluation.

The crystals were then removed from the XOTEX system and were built into OCXOs. The active aging test over 34 days was done afterwards in an automatic aging system (PRA model 2350), which operates at room temperature. The oscillators were continuously operating at the individually set crystal turn-over temperature with the crystals driven at their nominal drive level. The system measured every two hours the output frequency and the current consumption of the OCXO and stores the data.

## 5. TEST RESULTS

### 5.1 AT-CUT 16,384 MHz / 3RD OVERTONE

Figures 1a to 3a show the frequency aging for the two passive periods and the active period. While the passive measurements are referred to nominal frequency and thus can be compared between each other, the active measurement is referred to the first measurement after 1 day stabilization time. The step between the two passive periods is due to the DLD sensitivity of the crystal. Figures 1b to 3b show the resistance changes in the passive aging periods referred to the initial measurement. The  $R_1$  of all 3 crystals is approx.  $10,5\Omega$ , the  $Q$  is approx. 460 000.

Crystal #25942 (Fig.1) shows a constant aging rate of  $-0,34$ ppb/day in the 1<sup>st</sup> period, which repeats exactly ( $0,36$ ppb/day) in the active aging. In the 2<sup>nd</sup> passive period a flat aging rate is reached after 3 days stabilization after a DLD step of  $+26$ ppb.  $R_1$  is constant through both periods. The other two crystals also show a fairly constant aging rate in the 1<sup>st</sup> passive aging ( $1,12$ ppb/d and  $0.53$ ppb/d), which is flatter in the 2<sup>nd</sup> period. The active aging displays a logarithmic shape, which ends in the same aging rate as in the first aging period. It is remarkable, that the resistances are much noisier, particularly in the 2<sup>nd</sup> period, and are increasing with time, which may be related to the stronger aging. The DLD effect of frequency is  $-24...+42$ ppb in both directions, while the DLD effect of resistance is  $+10\%$ .

### 5.2 SC-CUT 16,384 MHz / 3RD OVERTONE

Figs. 4 to 6 show the aging curves of this crystal type, which has the same  $Q$  ( $\approx 460$  000) as the AT-cut above. Initial passive aging is negative and weaker than above. The aging rates of passive and active aging are well comparable in all three cases and are in the order of  $0.1$ ppb/day, i.e. by a factor of  $5...10$  better than the AT-cut. The initial log aging approaches a slope, which continues in the subsequent periods. The  $R_1$  curves ( $R_1 \approx 60\Omega$ ) are smooth with a DLD-effect of  $-5,8\%$  (#488) or  $<1\%$ , the DLD of  $f_s$  is  $+53...-74$ ppb with both signs occurring.

### 5.3 AT-CUT 10 MHz / FUNDAMENTAL

Figs. 7 to 9 show the results for this crystal type ( $Q \approx 270$  000). Again the aging rate of the active aging reproduces that of the passive aging very well and is in the order of  $0,1$ ppb/day. The DLD effect of  $f_s$  for this crystal is strong ( $\approx 400$ ppb), while the DLD of  $R_1$  is very small. Resistance ( $R_1 \approx 8,5\Omega$ ) has a weak trend over time, the two  $R_1$  dips coincide with  $f_s$  dips and are measurement errors.

### 5.4 AT-CUT 13 MHz / FUNDAMENTAL

These crystals ( $Q \approx 200$  000) show a strong initial passive aging (see Figs. 10 to 12), which continues in the 2<sup>nd</sup> passive period without any DLD. After the intermission, the active aging starts with an initial aging response, afterwards the aging curves seem to approach those of the prior periods (if the frequency offset between 2<sup>nd</sup> and 3<sup>rd</sup> period is

removed). The  $R_1$  curves ( $R_1 \approx 7,5 \dots 12,8 \Omega$ ) are smooth and show no DLD.

### 5.5 SUMMARY

In Table 2 the results of all four crystal types are summarized.

TABLE 2: Comparison passive vs. Active aging

freq	cut	xtal #	aging	aging	DLD	DLD
			passive	active	of $f_s$	of $R_1$
MHz			ppb/d	ppb/d	ppb	%
16,384	AT 3	25942	0,34	0,36	26,3	-0,2%
		25945	1,14	1,12	-24,5	9,6%
		25946	0,58	0,53	42,4	2,1%
16,384	SC 3	488	0,07	0,07	-74,3	-5,8%
		536	-0,17*	-0,17	53,2	0,9%
		513	-0,14	-0,12	69,3	0,5%
10,000	AT 1	25334	-0,06	-0,06	446	0,3%
		25335	-0,13	-0,02	419	0,1%
		25336	0,09	0,18	384	-0,5%
13,000	AT 1	26594	0,85	0,4	7,6	-0,1%
		26595	3,67	1,16	33,4	0,2%
		26596	3,94	1,52	35,3	-6,3%

mean value 1<sup>st</sup> and 2<sup>nd</sup> period  
italic: log shape

- The aging rates observed with passive aging are in general very close to those observed during active aging.
- In case of strong logarithmic aging - as for the 13 MHz AT 1 - the aging curve of the active aging follows the curve of the preceding passive aging after a stabilization interval. I.e. strong initial aging at passive aging repeats also with active aging.
- The reproducibility is also excellent in cases of strong DLD of frequency and/or resistance. No systematic dependancy of the aging rate on the drive level could be found.
- Coincidence of  $f_s$ - and  $R_1$ -discontinuities can be used to identify wrong measurements.
- These results demonstrate, that passive aging measurements are capable to deliver reliable results for the selection of well aging crystals prior to their usage in the OCXO.

## 6. PREDICTIBILITY OF LONG-TERM-AGING

Aging predictions for quartz crystals have a long history. At the 1959 Frequency Control Symposium Mulvihill [1] showed a simple way to compare long-term aging with short-term aging by setting up a „binary correlation chart“ of „passed“ and „failed“ crystals in long-term vs. short-term aging. Ray Filler mentioned already in 1980 [19]:

„Extrapolating of aging data is usually not a reliable method for determining the long-term aging because the observed aging is usually the sum of the aging produced by various mechanisms.“

Nevertheless aging data extrapolation has to be done every day in the industrial OCXO production because the customers (i.e. the applications) require a guaranteed maximum aging (rate) over the specified operation time of the device. Every

manufacturer has its own methods to assure this. It has to be noted, that the uncertainty of the correlation between passive and active aging (as the uncertainty of any comparison of aging results under different conditions) cannot be less than the uncertainty of long-term extrapolations as such.

To get an indication of the validity of our results, the aging period for the crystal oscillators of this experiment was extended to 240 days, and the long-term aging was compared with the logarithmic extrapolation derived from the data of the first 30 days of active aging. The results for one crystal of each group are given in Figs. 13 to 16. In all four cases, the aging prediction based on 30 days data and fitted by the modified logarithm (eq.[3]) delivers too optimistic results. The active aging measured after 240 days is in average twice the value predicted from 30 days data as shown in table 3.

TABLE 3: Predicted and measured active aging over 240 days

freq	cut	xtal #	predicted aging	measured aging
MHz			[ppb]	[ppb]
16,384	AT3	25946	70	130
16,384	SC3	513	-6	-35
10,000	AT1	25336	82	120
13,000	AT1	26594	120	220

In table 4 the measured and the predicted aging rates per day for passive and active aging are compared. While the measured 30 day aging data for passive aging are comparable (columns 4 and 5), the extrapolated (predicted) aging rate after 240 days (column 7) is only one third of that determined from the data of 240 days (col. 8).

TABLE 4: Computed aging rates [ppb/day]

freq	cut	xtal #	passive aging	active aging after 30 days		active aging after 240 days	
			30day data	30d data	240d data	30d data	240d data
MHz							
16,384	AT3	25946	0,58	0,53	0,91	0,07	0,20
16,384	SC3	513	-0,14	-0,12	-0,19	-0,04	-0,11
10,000	AT1	25336	0,09	0,18	0,87	0,05	0,14
13,000	AT1	26594	0,85	0,40	1,66	0,10	0,36

This means that the uncertainty of long-term prediction is much larger than the observed difference between passive and active aging

## 7. CONCLUSIONS

- Within the limits of these tests it was proofed, that the passive aging of OCXO crystals correlates well with active aging. The correlation is at least better than the uncertainty of aging predictions based on logarithmic curve fitting.
- Passive pre-aging is a powerful tool for selecting well-aging OCXO crystals before assembling them into the oscillator.
- No correlation between aging rate and drive level was found over the test periods.

- Crystals with strong DLD did not show higher aging rates than others.
- Aging predictions computed by log-fitting of aging data over 30 days have shown an uncertainty of -50%/+100%.
- Further verification tests on different kinds of OCXO crystals are necessary to proof these statements.

## 8. ACKNOWLEDGEMENT

The author wants to thank Erich Hafner/XOTEX for performing the passive measurements on his test system. We appreciate the fruitful discussions on aging evaluation we had with Dr. William Hanson / Piezo Crystals.

## 9. REFERENCES

- (1) Mulvihill, P.E.: Aging Characteristics of Quartz Crystal Units; Proc. 13th Frequency Control Symposium (FCS)(1959), p.109-112
- (2) Mason, W.P.: Influence of Lattice Parameters on the Properties of Crystal Resonators; Proc. 14th FCS (1960), p.35-52
- (3) Armstrong, J.H., Blomster, P.R., Hokanson, J.L.: Aging Characteristics of Quartz Crystal Resonators; Proc. 20th FCS (1966), p.192-207
- (4) Hafner, E., Blewer, R.S.: Quartz Crystal Aging; Proc. 22nd FCS (1968), p.136-154
- (5) Dybwad, G.L.: Aging Analysis of Quartz Crystal Units with Ti Pd Au Electrodes; Proc. 31st FCS (1977), p.144-146B.
- (6) Beetley, D.E., Blicht, B.R., Snowden, T.M.: The Quartz Resonator Automatic Aging Measurement Facility; Proc. 35th FCS (1981), p.263-270
- (7) Filler, R.L., Kosinski, J.A., Rosati, V.J., Vig, J.R.: Aging Studies on Quartz Crystal Resonators and oscillators; Proc. 38th FCS (1984), p.225-231
- (8) Hafner, E., Jackson, H.W.: Aging Measurements on Quartz Crystals in the Batch Mode; Proc. 40th FCS (1986), p.306-312
- (9) Feinberg, A.A.: Parametric Failure Rate Model for Quartz Crystal Device Aging with Application to Surface Wave Filters; Proc. 41st FCS (1987), p.360-364
- (10) Filler, R.L., Lindenmuth, R., Messina, J., Rosati, V.J., Vig, J.R.: The Aging of Resonators and Oscillators under Various Test Conditions; Proc. 41st FCS (1987), p.444-451
- (11) Miljkovic, M.R., Trifunovic, G.Lj., Brajovic, V.J.: Aging Prediction of Quartz Crystal Units; Proc. 42nd FCS (1988), p.404-411
- (12) Filler, R.L., Stein, S.R.: Kalman-Filter Analysis for Real Time Applications of Clocks and Oscillators; Proc. 42nd FCS (1988), p.447-452
- (13) Vig, J.R., Meeker, T.R.: The Aging of Bulk Acoustic Wave Resonators, Filters and Oscillators; Proc. 45th FCS (1991), p.77-101
- (14) Wei Su, Filler, R.L. : Application of Kalman Filtering Techniques to the Precision Clock with non-constant Aging; Proc. 46th IEEE FCS (1992), p.231-237
- (15) Vanier, J., Gagnepain, J.J., Riley, W.J., Walls, F.L., Granveaud, M.: Aging, Warm-Up time and Retrace, important Characteristics of Standard Frequency Generators (proposal for IEEE Standards, Project P1193); Proc. 46th IEEE FCS (1992), p.807-815
- (16) Wei Su, Filler, R.L.: A New Approach to Clock Modeling and Kalman Filter Time and Frequency Prediction; Proc. 47th IEEE FCS (1993), p.331ff.
- (17) Warner, A.W., Fraser, D.B., Stockbridge, C.D.: Fundamental Studies of Aging in Quartz Resonators; IEEE Trans. Sonics Ultrasonics, SU-12 (June 1965), 52-59
- (18) Vig, J.R.: Resonator Aging; Proc. Ultrasonic Symp. (1977), 848-850
- (19) Filler, R.L., Keres, L.J., Snowden, T.M., Vig, J.R.: Ceramic Flatpack Enclosed AT and SC-Cut Resonators; Proc. Ultrasonics Symp., (1980), Vol2, p.819-824
- (20) Gerber, E.A.: Long-Term Stability and Aging of Resonators; in: Precision Frequency Control, Vol.1, (Ed. Gerber, E.A. & Ballato, A.), Academic Press (1985), p.271-284
- (21) Filler, R.L., Vig, J.R.: Long term Aging of Oscillators; Proc. 46th IEEE FCS (1992), p.470-484
- (22) TELE QUARZ Internal Standard for Aging Evaluation (1995)
- (23) Filler, R.L.: Aging Specification, Measurement, and Analysis; Proc 7th Quartz Devices Conf., (1985), p.93-104
- (24) Grata, K.E.: Long term Crystal Stability Study; Proc. 5th Quartz Crystal Conf. (1983), p.214-221
- (25) MIL-C-49468, Military Specification, General Specification for Precision Quartz Crystal Units
- (26) MIL-O-55310, Military Specification, General Specification for Crystal Oscillators
- (27) Beetley, D.E.: Nine-Year Aging Behaviour of the Ceramic Flatpack Resonator; Proc. 3rd European Time and Frequency Forum (1989), p.139-142

16.384MHz AT/3.OT B1Nr. 25946

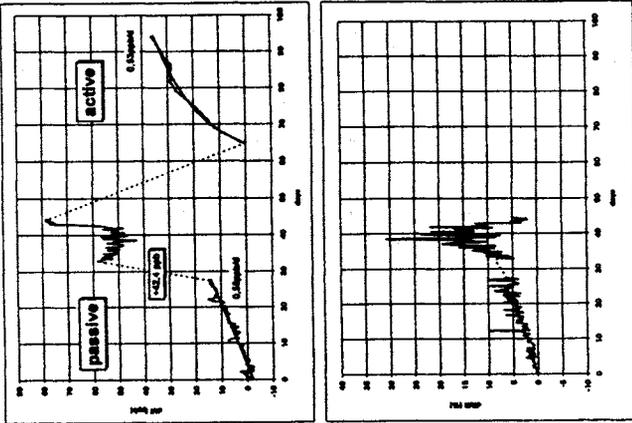


Fig. 3: passive and active aging

16.384 MHz SC/3.OT B1Nr 513

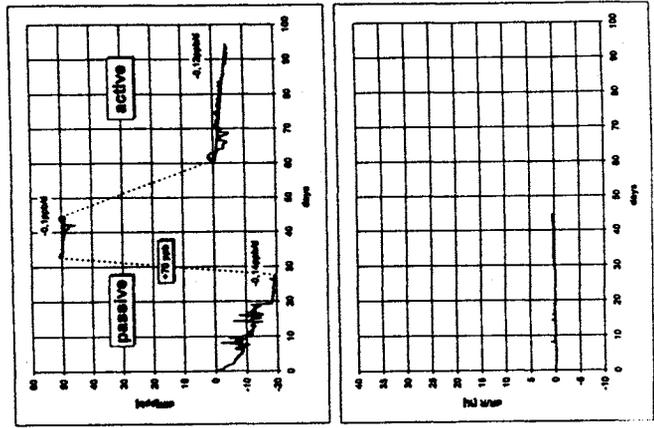


Fig. 6: passive and active aging

16.384MHz AT-Schnitt/3.OT B1Nr.25945

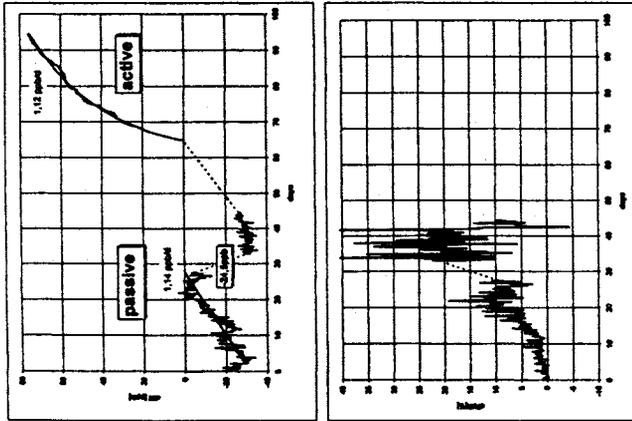


Fig. 2: passive and active aging

16.384 MHz SC/3.OT B1Nr. 536

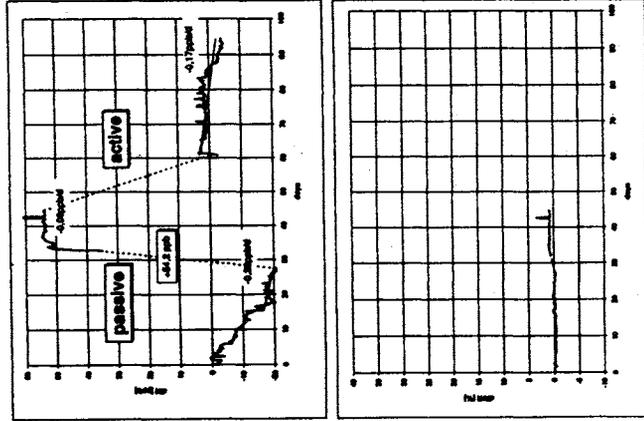


Fig. 5: passive and active aging

16.384MHz AT/3.OT B1Nr.25942

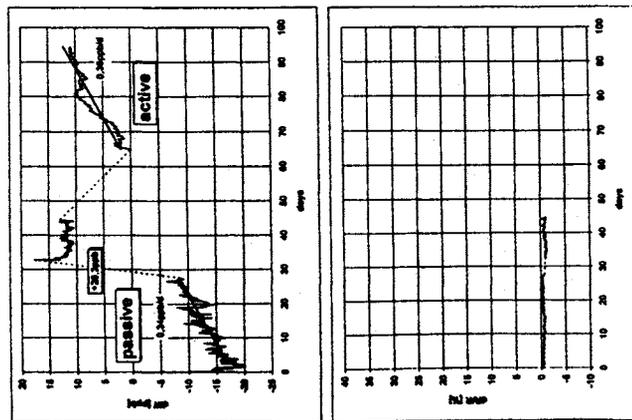


Fig. 1: passive and active aging

16.384 MHz SC/3.OT B1Nr. 488

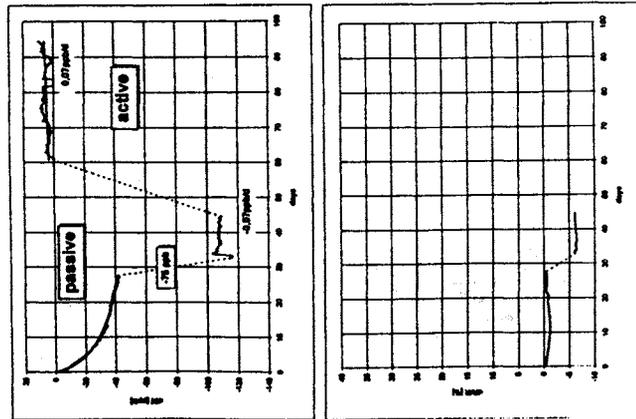


Fig. 4: passive and active aging

10.000 MHz AT/1.0T Bitr. 25334

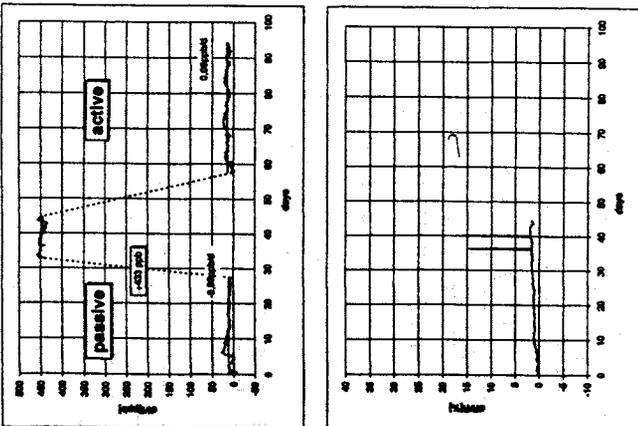


Fig. 7: passive and active aging

10.000 MHz AT/1.0T Bitr. 25335

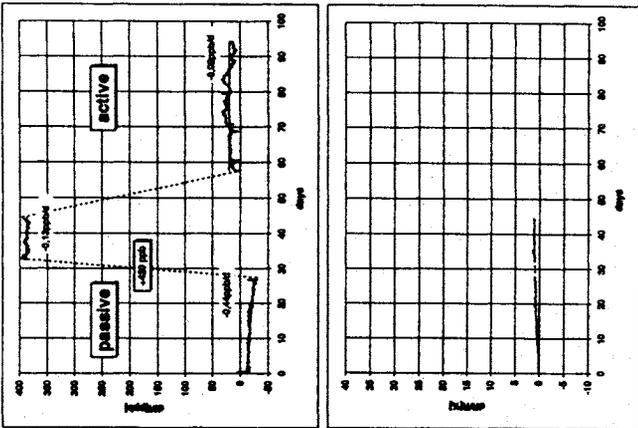


Fig. 8: passive and active aging

10.000MHz AT/1.0T Bitr. 25336

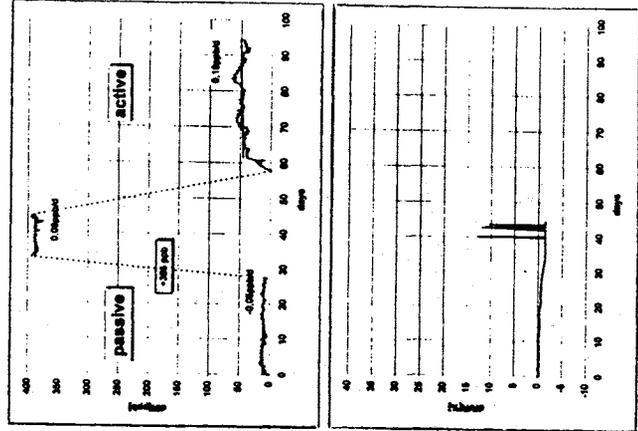


Fig. 9: passive and active aging

13.000 MHz AT/1.0T Bitr. 26594

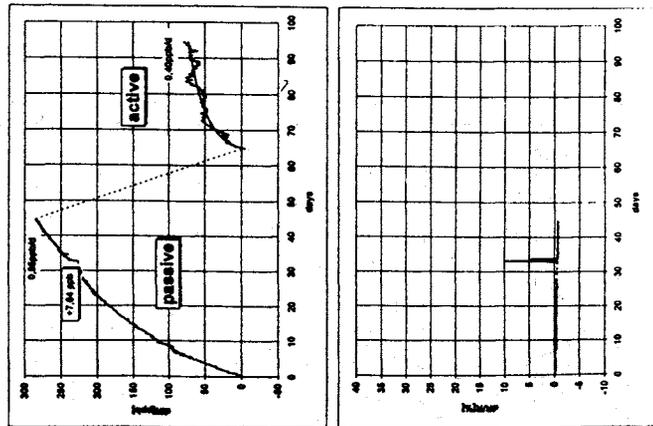


Fig. 10 passive and active aging

13.000 MHz AT/1.0T Bitr. 26595

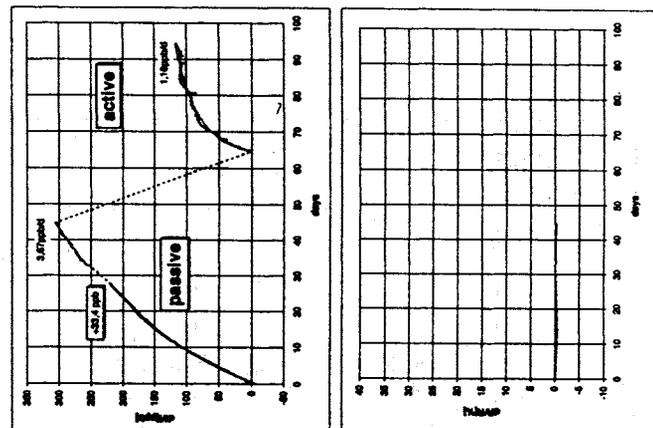


Fig. 11 passive and active aging

13.000 MHz AT/1.0T Bitr. 26596

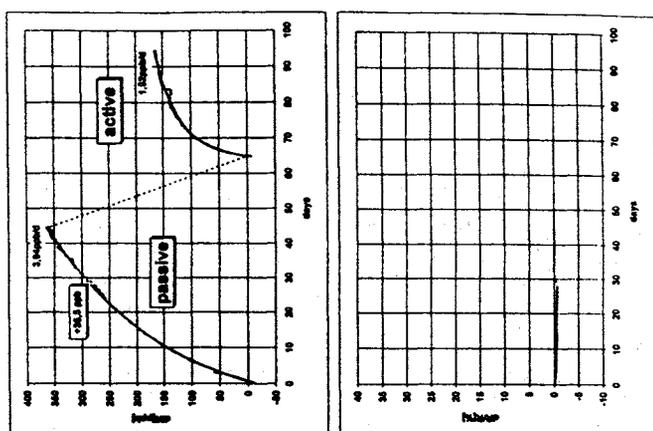


Fig. 12 passive and active aging

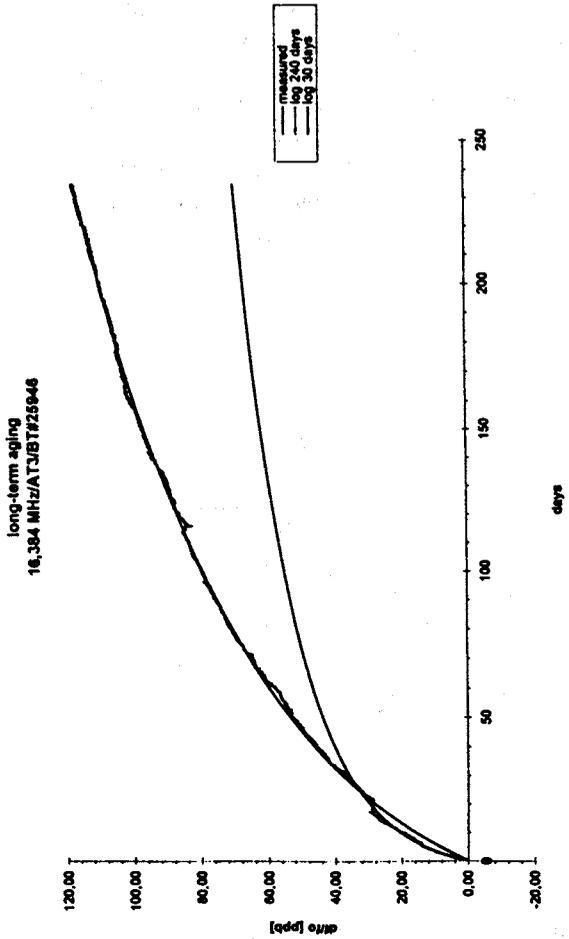


Fig. 13: predicted and measured long term aging

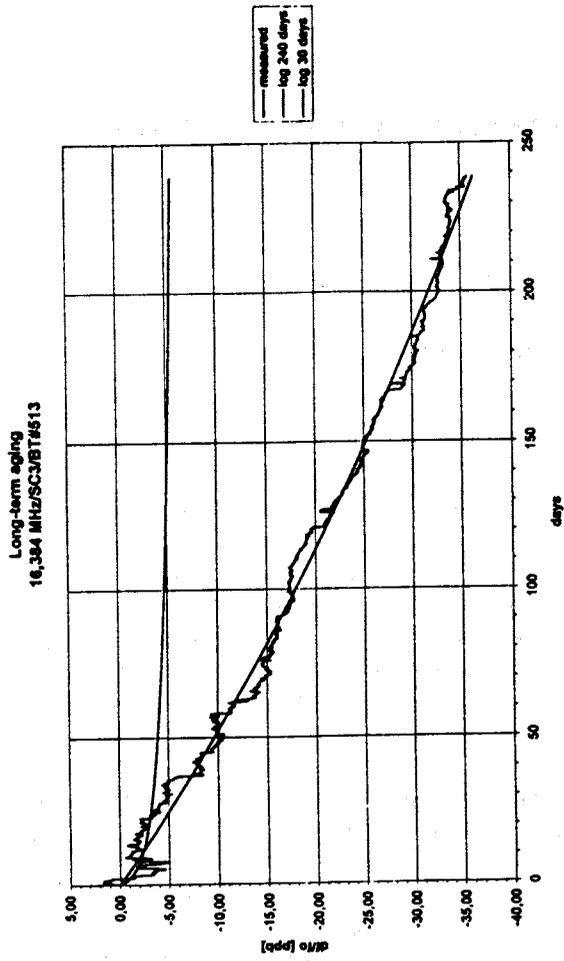


Fig. 14: predicted and measured long term aging

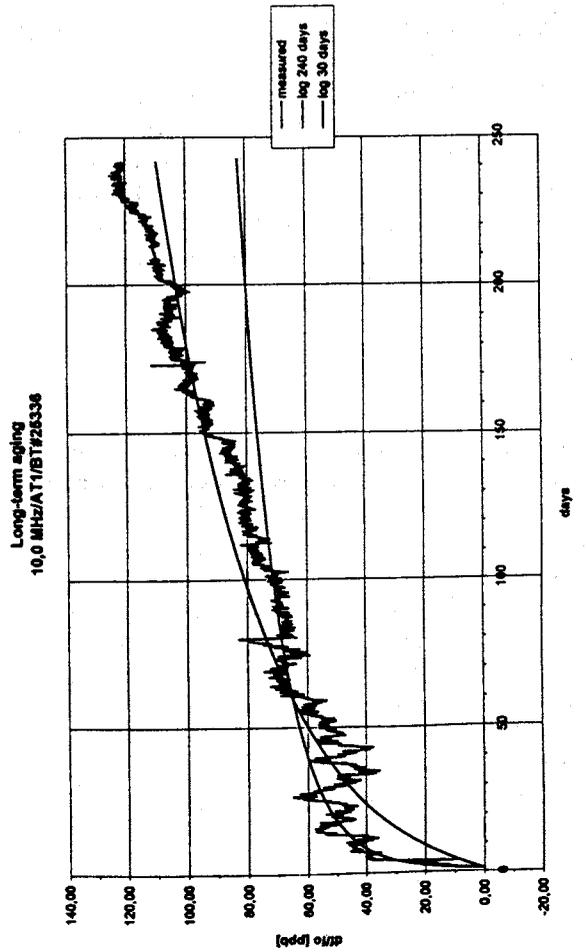


Fig. 15: predicted and measured long term aging

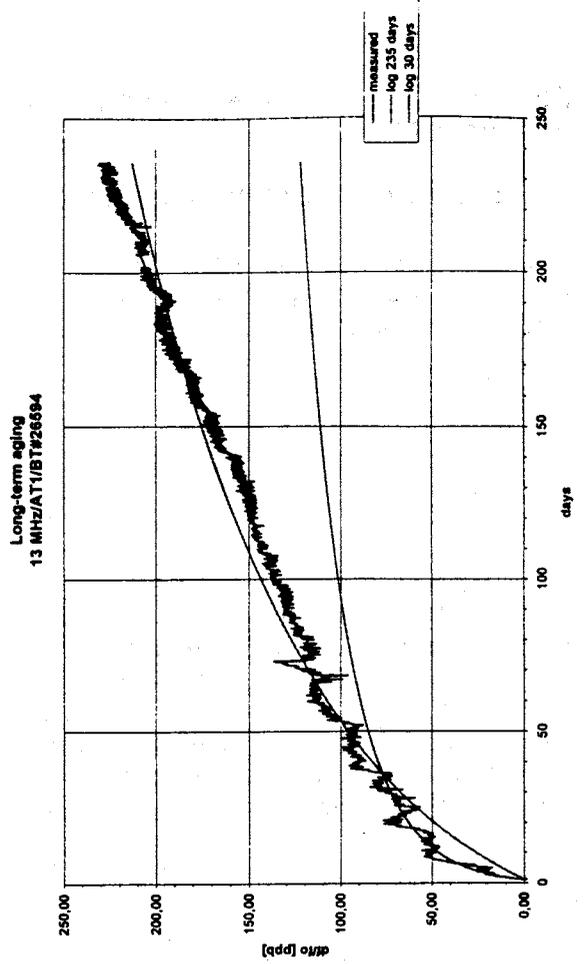


Fig. 16: predicted and measured long term aging

## RELATIONSHIP BETWEEN AMPLITUDE AND RESONANT FREQUENCY IN QUARTZ CRYSTAL RESONATORS\*

E.S. Ferre-Pikal\*\*+, F.L. Walls\*, and J.F. Garcia Nava\*

\*National Institute of Standards and Technology, 325 Broadway, Boulder, CO 80303, USA

+University of Colorado, Boulder, CO 80310, USA

•Centro Nacional de Metrologia, km. 4.5 Carr. A Los Cues, Municipio del Marques, C.P. 76900, Queretaro, Mexico

### ABSTRACT

In this paper we report results of measurements on the resonant frequency of crystal resonators as a function of drive current (amplitude). Previous studies indicate that the resonant frequency increases as the square of the driving current. If we assume that the limiting factor in the flicker frequency noise of crystal resonators is the noise in the drive current, then the quadratic dependence suggests that crystal resonators should be driven at low current for better frequency stability. Frequency versus amplitude measurements were made on SC-cut, 5th overtone, 100 MHz crystal resonators using a network analyzer. As expected, the measurements show a general quadratic dependence of frequency versus drive current. Nevertheless, some crystals exhibit phase (frequency) jumps at certain drive currents and certain temperatures. Phase modulation (PM) noise measurements were made in test oscillators at several currents to see if there is a correlation between the amplitude-frequency effect and flicker of frequency noise. Our results indicate that the flicker of frequency noise varies with current, but the current at which the flicker of frequency noise is the lowest is not necessarily the lowest current (as the quadratic relation of  $\nu_0$  versus drive current suggests).

### INTRODUCTION

The goal of this research was to investigate in fine detail the frequency of 100 MHz quartz resonators as a function of crystal drive. It has long been known that the resonant frequency of quartz resonators depends slightly on the level of excitation [Gagnepain and Besson (1), Gagnepain (2), Kusters (3), Tiersten (4), Filler (5)]. This effect is commonly called the amplitude-frequency effect. The data and analysis of Kusters (3), Filler (5), and others indicate that the frequency depends (approximately)

quadratically on crystal current. Several questions arise. Is the curve smooth, or are there jumps in crystal frequency? Is the frequency monotonic with increasing crystal current, or is there an extremum at some low current? If there is an extremum, how does this affect the flicker of frequency noise? If there are jumps how do they affect the flicker of frequency noise?

We made this study at 100 MHz because the resonators are less expensive than 5 or 10 MHz resonators. They are also smaller which allows for more resonators to be made from a given quartz bar. Furthermore, the flicker of frequency noise of such resonators should be less dependent on temperature fluctuations than 5 or 10 MHz resonators because the fractional flicker of frequency noise in the 100 MHz resonators is approximately 20 dB higher than in comparable 10 MHz resonators. The general characteristics of the resonators reported in this study are given in Table 1. All resonators were cut from the same bar of synthetic, unswept quartz, and fabricated with the manufacturer's standard polishing, electrode size, and electrode deposition techniques. Later studies will examine the effect of changing electrode size, polishing technique, dislocation density, and sweeping on the amplitude-frequency effect and flicker of frequency noise [Ferre-Pikal et al., (6)].

TABLE 1. General characteristics of crystal resonators.

Crystal Resonators used in this Study	
Frequency	100 MHz
Overtone	5th
Blank diameter	6.3 mm
Blank thickness	90 $\mu$ m
Electrode Diameter	3.05 mm
Geometry	plano-plano
Turnover temperature	60 - 72 °C

\* U.S. Government work not protected by U.S. copyright.

**MEASUREMENT OF THE AMPLITUDE-FREQUENCY EFFECT AT 100 MHZ**

Figure 1 shows the measurement technique used to determine the amplitude and phase of the signal transmitted through the resonators at series resonance. The variable capacitor is used to cancel the holder capacitance of the resonator.

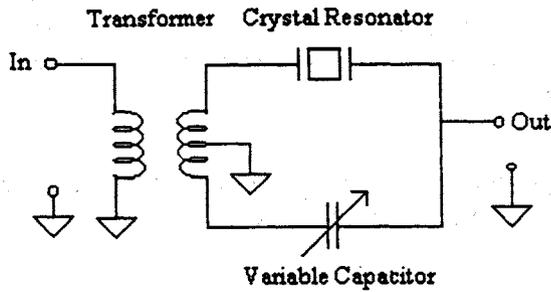


Figure 1: Technique used to measure the amplitude and phase of the transmitted signal through the resonators.

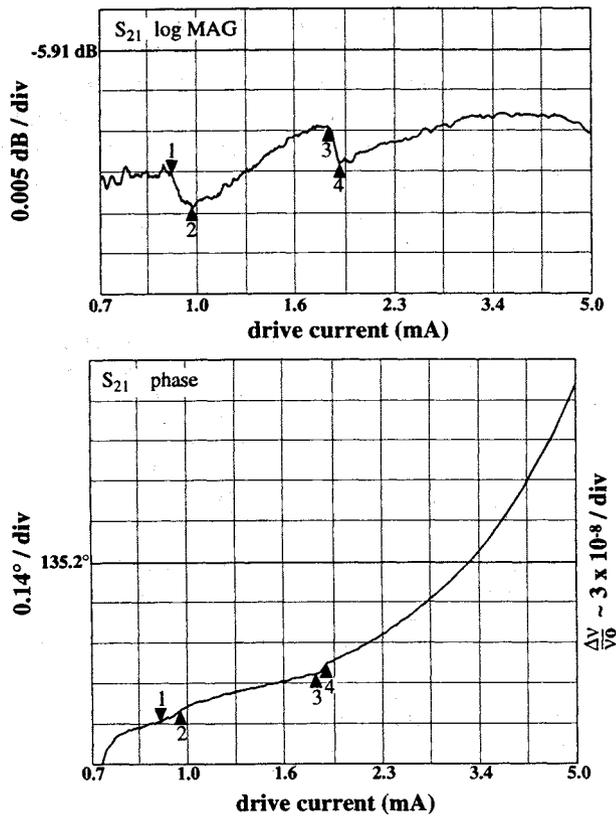


Figure 2: Amplitude and phase for the transmitted signal in resonator 2. The fractional frequency axis in the phase plot was calculated assuming a loaded  $Q$  ( $Q_L$ ) of  $0.4 \times 10^5$  and using the relation  $\Delta v/v_0 = \Delta\phi/(2Q_L)$  [7].

The upper part of Fig. 2 shows the amplitude of the transmitted signal for resonator 2 at the frequency-temperature turnover point. The vertical scale is 0.005 dB per division. The horizontal scale is linear in drive current squared. The left corner (origin) corresponds to approximately 0.7 mA, while the right corner approximately corresponds to 5.0 mA. Two step changes in the transmitted amplitude are clearly visible. The lower trace shows the phase shift across the resonator as a function of drive current. Changes in the phase shift across the resonator lead to frequency changes in the oscillator [Leeson, (7)], therefore the approximate quadratic dependence of phase on drive current translates into quadratic frequency dependence on drive current. [The initial increase in phase is due to startup effects in the resonator.] The markers are at the same position as in the upper trace. The phase steps are not as easily seen as in the upper trace (magnitude changes), but nonetheless are visible.

In Fig. 3 we compare the magnitude of the transmitted signal in resonator 2 at 20°C and 65°C. Clearly there is significant change in the current at which the jumps occur between 20°C and 65°C in this resonator.

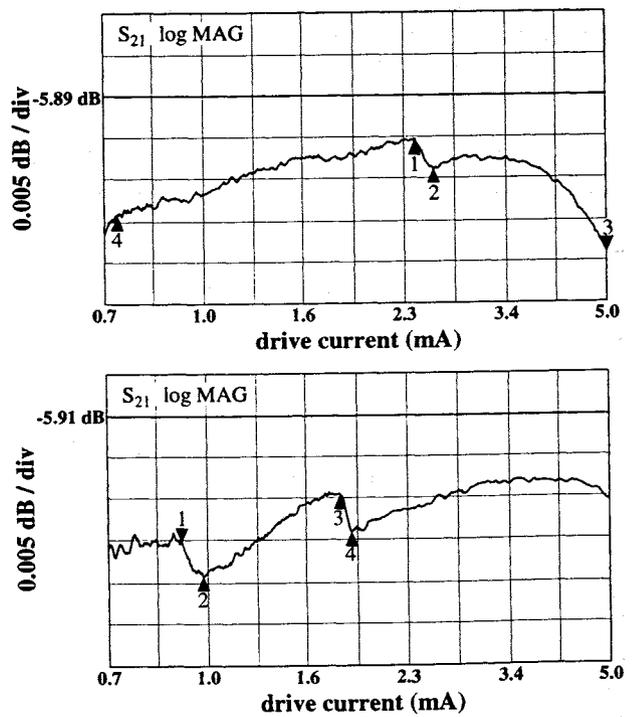


Figure 3: Magnitude of the transmitted signal in resonator 2 at 20°C (upper plot) and 65°C (turnover temperature) (lower plot).

In Fig. 4 we show the amplitude and phase for resonator 15 at 20°C. The upper trace clearly shows an abrupt jump in amplitude between markers 1

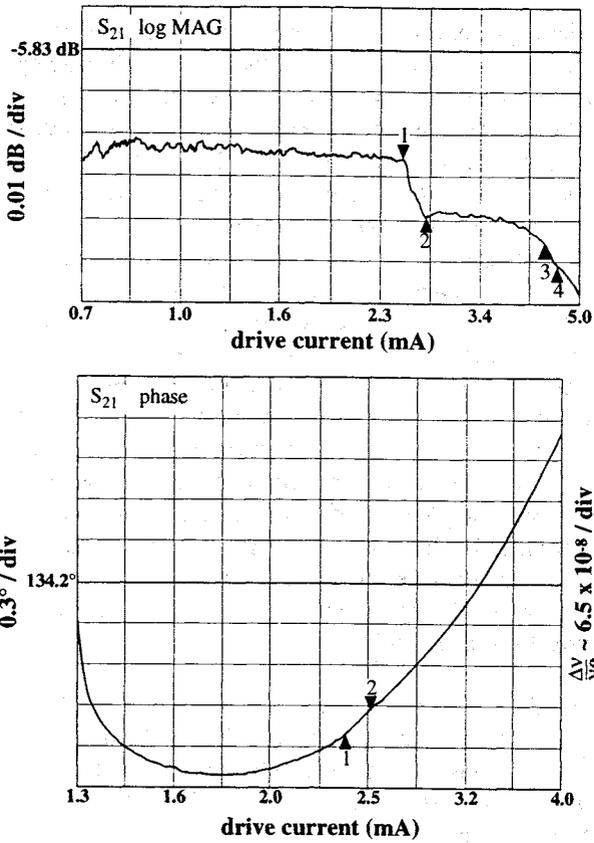


Figure 4: Magnitude and phase of the transmitted signal for resonator 15 at 20°C. The fractional frequency axis in the phase plot was calculated assuming  $Q_L = 0.4 \times 10^5$ .

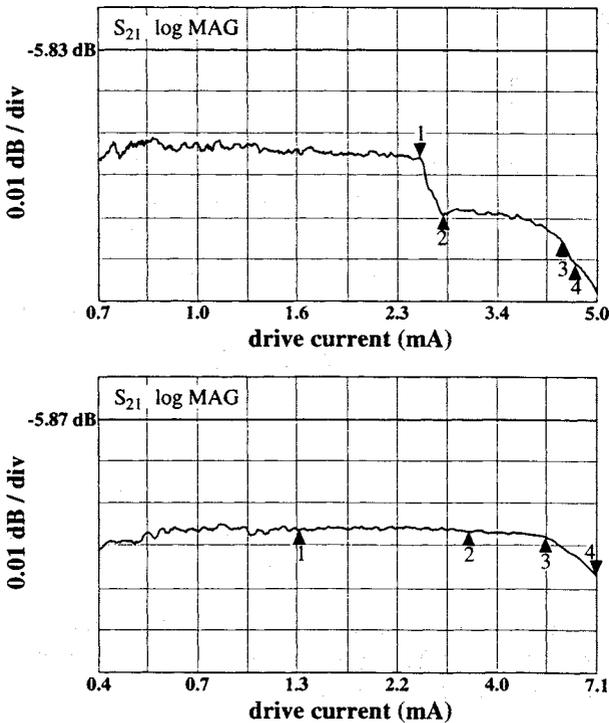


Figure 5: Amplitude of the transmitted signal in resonator 15 at 20°C (upper plot) and 72°C (lower plot).

and 2. The lower trace, however, does not show a corresponding jump in the phase.

In Fig. 5 we compare the amplitude of the transmitted signal in resonator 15 at 20°C and the frequency-temperature turnover point. At turnover no jumps are apparent. [The initial drop in phase is due to startup effects in the resonator.]

A survey of 12 resonators made from the same quartz material at the same time by the same manufacturer yielded 6 resonators with jumps in the transmitted signal at either room temperature or at turnover point.

### MEASUREMENT OF PHASE MODULATION (PM) NOISE VERSUS DRIVE CURRENT

Figure 6 shows the block diagram of the cross-correlation technique used to measure the PM noise of the test resonators incorporated into test oscillators [F.L. Walls (8), W.F. Walls (9)]. This approach provides a noise floor for measuring the PM noise of the test oscillator which is at least  $\mathcal{L}(10 \text{ Hz}) = -118 \text{ dBc/Hz}$  [10]. So far this is below that of any of the resonators tested. The PM noise in the 100 MHz crystal oscillators was measured at Fourier frequencies of 5 to 200 Hz, where the PM noise exhibits flicker of frequency noise ( $f^{-3}$  power law). Here we use  $\mathcal{L}(10 \text{ Hz})$  as an indication of the magnitude of the flicker of frequency noise in different resonators.

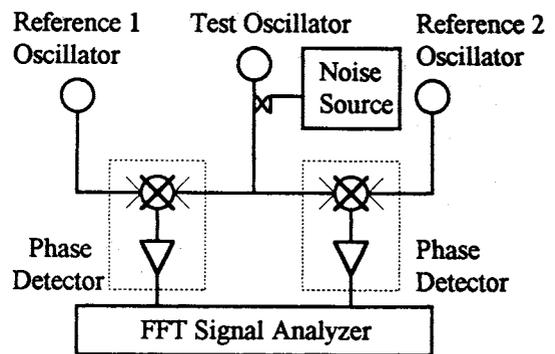


Figure 6: Block diagram of the cross-correlation PM noise measurement system

In Fig. 7 we show  $\mathcal{L}(10 \text{ Hz})$  for resonator 2 measured in a test oscillator as a function of crystal drive current. The arrows show the direction of the changes in crystal drive. There is a definite change in  $\mathcal{L}(10 \text{ Hz})$  with crystal drive and a very clear

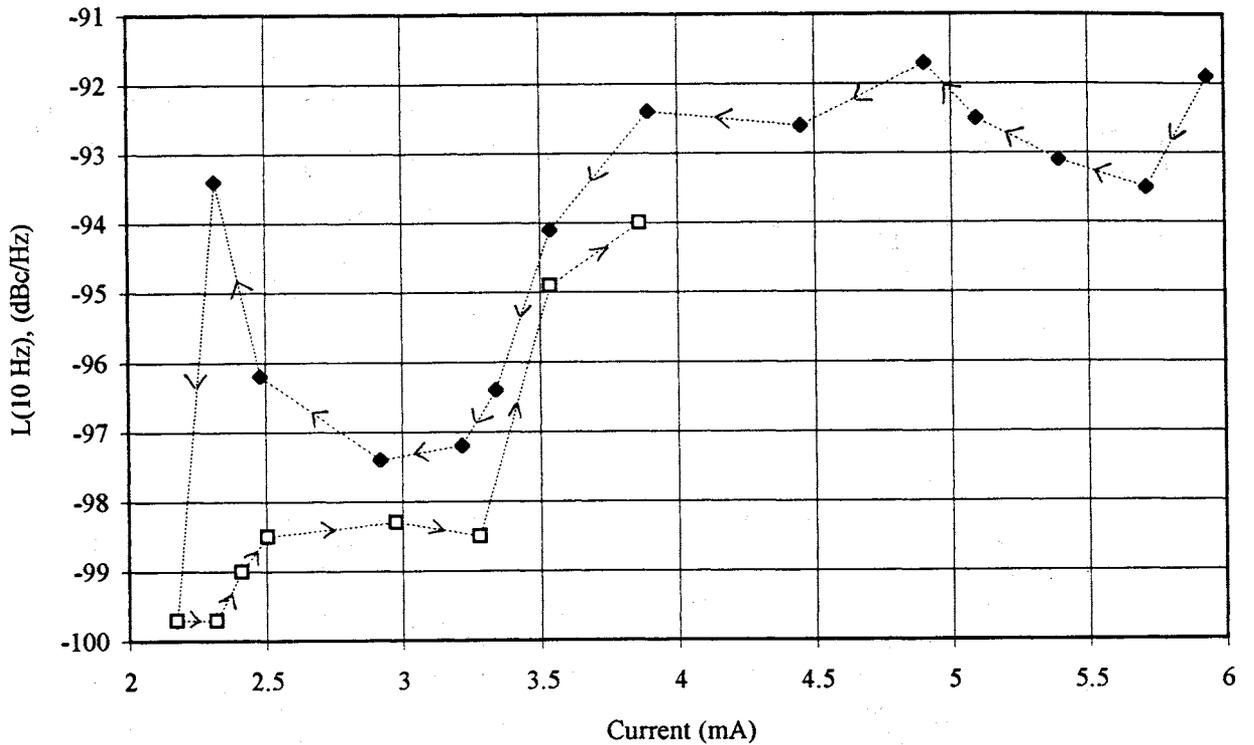


Figure 7:  $L(10\text{ Hz})$  of resonator 2 as a function of drive current

hysteresis of  $L(10\text{ Hz})$  versus direction of change in in crystal current. Referring to Fig. 2, the hysteresis occurs close to the point of the upper jump in the transmitted amplitude. In general, the flicker of frequency noise changed considerably with drive current. Measurements of  $L(10\text{ Hz})$  versus crystal current for all 12 resonators are shown in Fig. 8, trace labeled "Initial".

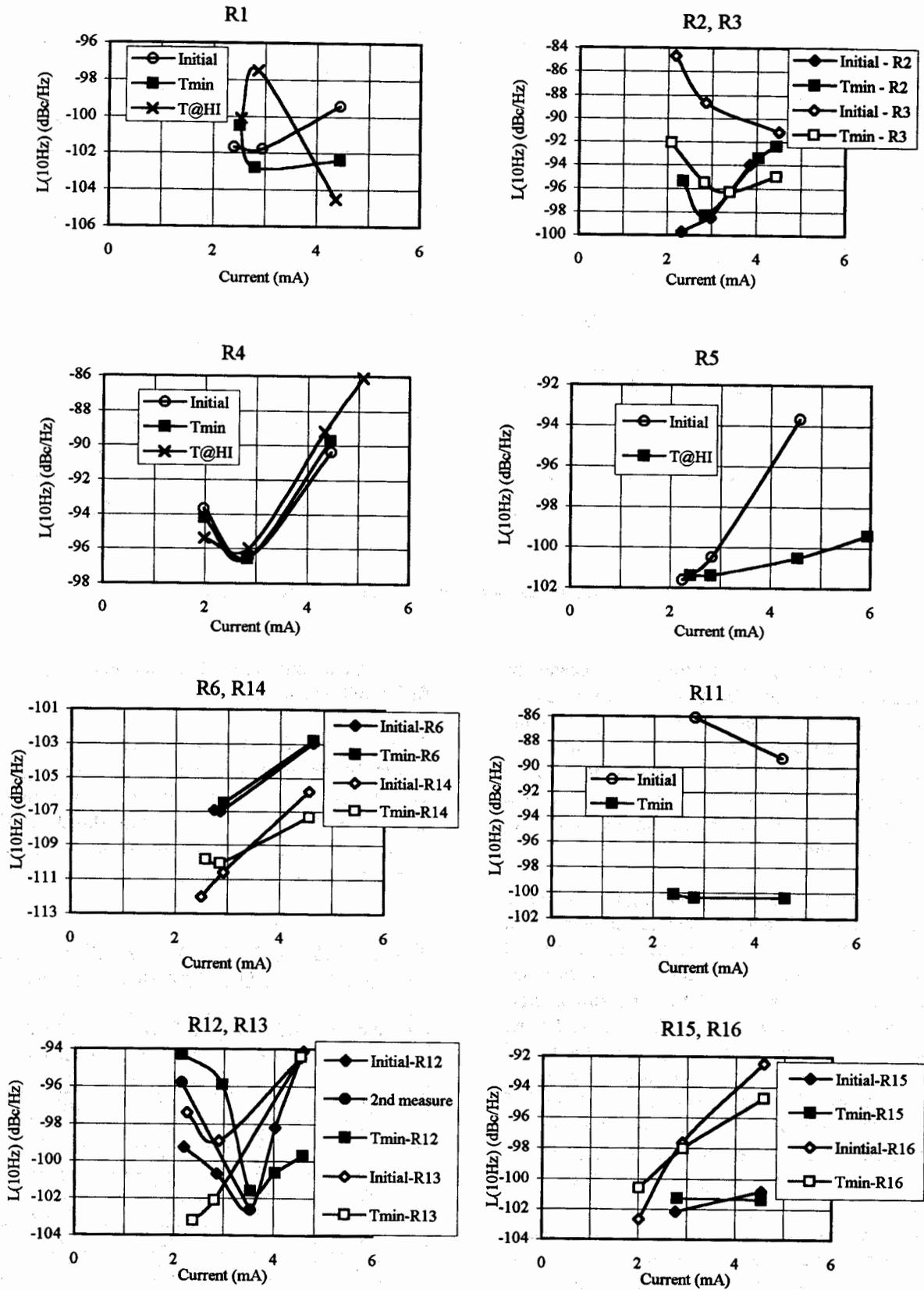
We also considered the effect of baking the resonators on the flicker of frequency component of the PM noise. After initially measuring the flicker of frequency noise ( $L(10\text{ Hz})$ ) at several drive currents, the resonators were baked at  $90^\circ\text{C}$  for approximately 2 days and then baked at  $100^\circ\text{C}$  for at least 3 more days. Measurements after baking show substantial improvement in some resonators and no effect in others. See Fig. 8, where the trace labeled "Initial" refers to the initial  $L(10\text{ Hz})$  measurements, and the trace labeled "Tmin" refers to  $L(10\text{ Hz})$  after baking the resonator at the current with best noise performance. For example,  $L(10\text{ Hz})$  of R1, R3, R11, and R13 improved considerably when baked. Most of the changes in the flicker of frequency noise occur in the first 2-3 days of the baking cycle. After baking, the number of resonators with  $L(10\text{ Hz})$  less than  $-100\text{ dBc/Hz}$  was 9 (out of 12). The effect of baking the resonators at high currents (6 mA) was also investigated (trace "T@HI", Fig. 9), but there was

not a strong indication that this had an effect on flicker of frequency noise.

Table 2 shows  $L(10\text{ Hz})$  of the resonators at  $100^\circ\text{C}$  and at the frequency turnover. It is quite surprising that the  $L(10\text{ Hz})$  at  $100^\circ\text{C}$  is equal or better than the  $L(10\text{ Hz})$  at the turnover temperature (for 9 out of 12 resonators) since at  $100^\circ\text{C}$  the frequency changes very rapidly with temperature.

TABLE 2. Comparison of  $L(10\text{ Hz})$  at turnover temperature and at  $100^\circ\text{C}$  for crystal resonators.

Resonator #	Current (mA)	$L(10\text{ Hz})$ (dBc/Hz) Turnover	$L(10\text{ Hz})$ (dBc/Hz) $100^\circ\text{C}$
R1	2.8	-102.8	-100.8
R2	2.2	-95.1	-96.2
R3	4.5	-95	-98
R4	2.8	-96.6	-98.4
R5	2.7	-101.3	-101.2
R6	2.8	-106.5	-106
R11	2.4	-100.4	-103.2
R12	3.5	-102.0	-94
R13	2.9	-102	-104
R14	3	-111	-110.3
R15	2.7	-101.6	-104.1
R16	2	-102.3	-102

Figure 8.  $L(10\text{ Hz})$  for crystal resonators as a function of drive current

## DISCUSSION AND CONCLUSION

This study shows that the amplitude-frequency effect is very complicated in some resonators. These units show features that are not explained by a simple quadratic dependence on crystal drive current. The step changes in frequency with changes in current strongly suggest that there is nonlinear coupling between the primary mode and other modes. This conclusion is supported by the hysteresis in the  $\Delta f$  (10 Hz) data from resonator 2 and the data of Table 2, which show that some resonators of this design have lower flicker of frequency noise at 100 °C than at temperature turnover. Figure 8 indicates that the flicker of frequency component of the PM noise can be significantly improved in some resonators by a post-processing bake with the crystal current set to the final operating value (lowest PM noise).

These data raise many questions. What mediates the coupling between the mode of interest and other modes? Is it strain caused by electrode plating, mounting, and/or dislocations and impurities? Can the geometry be altered to optimize the energy trapping and reduce coupling to other modes? What is changing during the high temperature processing? Is it strain or perhaps is it the movement of impurity ions to more stable positions which depend on the drive current? If it is related to impurity ions, would the results be different if swept quartz were used? What is the role of amplitude modulation (AM) noise in the oscillating loop? Does it affect the resulting flicker of frequency noise of the oscillator?

We hope to address some of these questions in the future by comparing the effect of changing electrode size, polishing technique, dislocation density and sweeping on amplitude-frequency effect and flicker of frequency noise.

## ACKNOWLEDGMENTS

We thank J.R. Vig for his suggestion to investigate PM noise as a function of crystal current.

## REFERENCES

1. J.J. Gagnepain and R. Besson, 1975, Physical Acoustics, vol. XI, W.P. Mason and R.N. Thurston, editors, Academic Press, New York.

2. J.J. Gagnepain, 1981, "Nonlinear properties of quartz crystal and quartz resonators: a review," Proc. 35th Ann. Symp. Freq. Control, pp. 14-30.
3. J.A. Kusters, 1981, "The SC-cut crystal - an overview," Proc. 1981 IEEE Ultras. Symp., 1981, pp. 402-402.
4. H.F. Tiersten, 1984, "An analysis of nonlinear resonance in electroded contoured AT- and SC-cut quartz crystal resonators." Proc. 38th Ann. Symp. Freq. Control, pp. 132-140.
5. R.L. Filler, 1985, "The amplitude-frequency effect in SC-cut resonators," Proc. 39th Annual Freq. Control Symp., pp. 311-316.
6. E.S. Ferre-Pikal, F.L. Walls, J.R. Vig, and J.F. Garcia Nava, "Experimental studies on flicker noise in quartz crystal resonators as a function of electrode volume, drive current, type of quartz, and fabrication process," to be presented at the 1996 IEEE Freq. Control Symp.
7. D.B. Leeson, 1966, "A simple model of feedback oscillator noise spectrum," Proc. IEEE, 54, no. 2, pp. 329-330.
8. F.L. Walls, 1992, "Secondary Standard for PM and AM noise at 5, 10, and 100 MHz," IEEE Trans. Instrum. Meas., 42, 136-143.
9. W.F. Walls, "Cross-correlation phase noise measurements," Proc. 46th Ann. Freq. Control Symp., pp. 257-261.
10. dBc/Hz refers to dB below the carrier in a 1 Hz bandwidth.

## DURABILITY OF SPECTRAL LAMPS WITH VAPORS $^{87}\text{Rb}$ FOR OPTICALLY PUMPED FREQUENCY STANDARDS

A Gevorkyan, G Smirnova, V Khutorshchikov  
Russian Institute of Radionavigation and Time, St.-Petersburg, Russia

### THE SUMMARY

The electrodeless spectral lamps with vapors of  $^{87}\text{Rb}$  in many respects determine the accuracy and service life of optically pumped frequency standards. The service life of frequency standards should be of 7-15 years at instability of intensity of light source in required limits during the specified period. Change  $^{87}\text{Rb}$  and buffer gas quantity, determining stability of radiation intensity and reproducibility of characteristics of spectral lamps is investigated. Object of research are lamps by a diameter up to 13 mm, filled with vapors of  $^{87}\text{Rb}$  and buffer gas. Dependence of change of buffer gas pressure in time from discharge power, as well as speed of absorption  $^{87}\text{Rb}$  by glass wall is investigated.

A technique of accelerated check of lamps durability and appropriate equipment, enabling to predict service life of lamps at their manufacturing with high accuracy are developed.

### INTRODUCTION

The beginning of development of quantum gas cell frequency standards was more than 35 years ago in Russian Institute of Radionavigation and Time (RIRT, till 1990 LNIRTI). It has appeared possible due to occurrence of new high-intensity compact and reliable light sources based on high-frequency electrodeless spectral lamps with vapors of metals [1].

But in spite of their qualitative superiority over other light sources the high-frequency electrodeless lamps limited achievable sensitivity and durability of quantum optically pumped frequency standards. Therefore systematic researches were carried out at RIRT aimed at perfection of light sources and first of all increasing of their durability and stability [2].

During the work the model of the electrodeless discharge in the mixture of alkali metal vapor and buffer gas was offered confined in the explanation the features of lamps spectral characteristics by concentration rubidium atoms in ground and excited states near the walls of a lamp [3]. This model, justified originally theoretically, was confirmed further by direct measurement of distribution of atoms in lamp's volume [4,5] and was used at the

solution of a problem of lamps reliability. Creation later [6,7] the mathematical model have allowed to determine operation modes ensuring high reliability of a lamp, to develop techniques of accelerated tests and to evaluate limiting opportunities of similar light sources. It should be noted that actuality problem of increasing the reliability in spite of obvious successes (the guaranteed durability of lamps has grown from a beginnings of 70 ties more than 50 times from 1000 till 55000 hours, is still exist now in view of continuous growth of the requirements to reliability and durability of quantum frequency standards.

### STABILITY OF RADIATION INTENSITY OF LAMPS

The main parameter of a light source is the radiation intensity and therefore its durability is characterized by time of maintenance of constant intensity in required limits.

The radiation intensity of a lamp (without the calculation of cascade transitions and step excitation) is determined by expression [2]:

$$I_{i0}(T) = A_{i0} \cdot \hbar \omega_{i0} \cdot \int_V n_0(T) \cdot n_e \cdot dV \cdot \int_0^\infty \sigma_{0i}(v) \cdot f(v) \cdot v^2 dv,$$

where  $n_e$  is local concentration of electrons;  $n_0(T)$  is local concentration of atoms of rubidium in ground state;  $T$  is temperature of condensed rubidium;  $A_{i0}$  is factor of Einstein;  $f(v)$  is function of the electrons distribution on velocities  $v$ ;  $\sigma_{0i}(v)$  is crosssection of excitation process of atoms from a main state to a state  $i$ . For the calculation of intensity one should solve the system of kinetic equations and balance of power, to determine space and power distribution of electrons and space distribution of atoms [6,7]. Thus it is essential that the concentration of rubidium atoms is calculated under the formula

$$n_0(T) = \frac{P}{kT},$$

where  $k$  is Boltzmann constant,  $P$  is pressure of saturated vapor over condensed metal, which is determined by formula of Klaperyon-Klausiz

$$\frac{dP}{dT} = \frac{q}{T^2(V_2 - V_1)},$$

where  $q$  is output energy of atom from a condensed condition to vapor;  $V_2 - V_1$  is the change of volume at the change of a aggregation.

Actually the vapor is in balance in a stationary condition with atoms in a drop and distributed on whole surface of a lamp. Because of distinction of temperature on a surface of a lamp, unhomogeneity availability, the distribution of atoms on a surface are not similar and will vary during the service because of interaction of atoms and ions with a material of a lamp's envelope, change of the output energy and reduction of a surface, occupied with a condensate of rubidium. Thus the transparency of a material of an envelope will vary for light and exciting field. The condensation of rubidium film on a wall will cause the change of boundary conditions for an equation of the electrons balance because of the photoeffect phenomenon, ionic atomization etc. Thus radiation intensity which will gradually vary, can grow or decrease depending on prevailing processes. However, the strongest change occurs after burn-out of whole rubidium.

Excessively amount of metal dose into the lamp is inadmissible because of arising instability owing to migration of rubidium drops on a lamp's surface. If speed of change of metal amount is designated  $G(T,t)$ , as at qualitative manufacturing, rather high initial pressure of gas in a lamp, the service life is limited by burn-out of rubidium:

$$\tau = m_2 - m_1 / G(T,t),$$

where  $m_2$  is the initial quantity of the metal contain,  $m_1$  is the final permissible amount of the metal contain in the lamp.

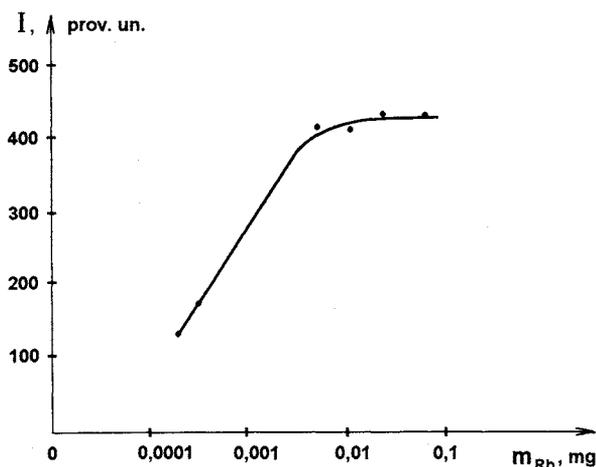


Fig. 1. Dependence of radiation intensity on amount of  $^{87}\text{Rb}$  in a lamp.

## CONTROL AND DOSING OF RUBIDIUM

To determine the allowable metal amount in a lamp, the velocity of change of metal amount a thermal-sorption technique of measurement of rubidium amount [8,9] was developed and introduced in manufacture process, the error of which has appeared from 5 up to 50% depending on metal amount in a lamp and material of an envelope. This technique has allowed to carry out comprehensive research of influence of metal amount on service life of a lamp.

First of all it was established, that the radiation intensity remains approximately constant at reduction of metal amount up to the value of the order 0.003 mg (fig. 1), and then rather quickly decreases. Further it was established [9], that the heaviest change of metal amount occurs in the first 100-200 hours of training, and then the velocity of change of metal amount in a lamp sharply drops (saturation of walls of a lamp with the metal). Thus in the first hours of absorption up to 0.01 ... 0.015 mg of rubidium occurs. In this connection, taking into account measured velocity of rubidium absorption, for maintenance of guaranteed function of 10000 hours in conditions of the intensive H-discharge with the power  $2-3 \text{ W/cm}^3$  was established minimum amount of metal 0.03 mg and maximum 0.2 mg [9]. These parameters are controlled in manufacturing process of lamps for quantum frequency standards and indicated in the literature [9,5,10,2].

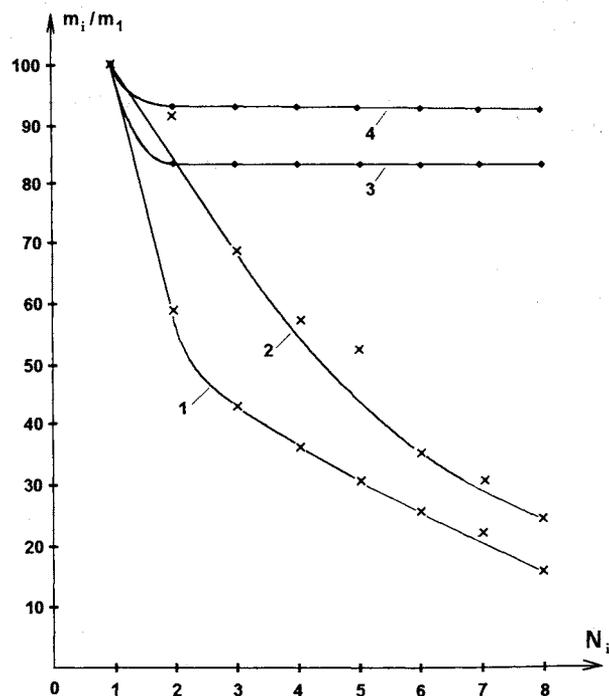


Fig. 2. Change of amount  $^{87}\text{Rb}$  in a lamp. 1,2 are for lamps, made without enrichment of B; 3,4 are for lamps, made with enrichment.

Some years later a dosing technique of metal in lamps [5,11] was introduced being by development of a technique of the rubidium amount control and enabling to receive in a lamp given amount of metal and, thus, to decrease losses by their manufacture.

During research of the burn-out speed of metal amount, spread of this value at different copies of lamps and yet greater spread at different series of lamps was found out, though all lamps were produced from the same consignment of alume-bor-silicate glass, designated in Russia as S51-1. This problem has become a subject of special research. With the help of the X-ray spectroscopic research of the internal surface of the bulbs [10] it was found out, that the rubidium and krypton absorption by glass depends on the contents of B in it, which can vary at fire work-out process of lamp's envelopes.

Absorption of the glass S51-1 was investigated at various manufacturing of preparations (fig. 2). The glass enriched by B (3,4 on fig. 2) absorbs rubidium much poorly and consequently for maintenance of required durability one should supervise the contents of B in glass, that was entered in the technology process of lamps manufacturing.

One should note, that the similar researches of optimum metal amount carried out in USA [13,14] have shown identical results - recommended metal amount has appeared from 0.03 up to several tenth's mg of rubidium. It was noted [13], that 0.4 mg and more amount leads to the light source noise increase. The rubidium amount in two lamps was measured [15]. By comparison of results [9,12,13,14,15,18] one should mean, that they concern to concrete investigated light sources and lamps, materials of an envelope and modes of the discharge. Nevertheless, by obtained results [9,12, 13,14,17] general

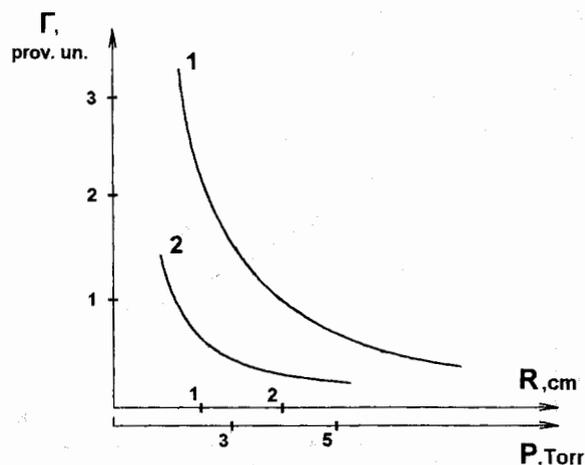


Fig. 3. Dependence of an ions flow on a wall of a lamp on pressure of gas (1) and from radius of a lamp (2).

conclusion about minimum and maximum allowable rubidium amount can be made.

## CONTROL OF GAS PRESSURE IN LAMPS

The intensity of radiation, the ignition voltage, the impedance of lamps depends on the gas pressure. For the control of pressure it was also developed using special technique [16], confined in determination of pressure on the frequency shift of the radiooptical resonance, which is measured by means of a quantum frequency standard. First measurements have shown strong difference of filling from expected and huge spread of real pressure in the manufactured lamps. Therefore the technique was introduced into manufacture of lamps at once, that has allowed to increase essentially their reproducibility [5,10,12, 16,17]. The error of the technique has appeared not more than 1%.

The calculation of dependence of ions flow on the wall of a lamp due to gas pressure (1 on fig. 3) and radius of lamp (2 on fig. 3), as well as experimental detection of reduction of gas pressure while in service was shown necessity of increasing of gas pressure in lamps up to  $3.2 \pm 0.2$  Torr, as well as opportunity of essential decrease of lamps aging speed at increase the pressure up to 5 ... 10 Torr. The efficiency of light sources at such increasing of pressure is not reduced, since selecting the mode of operation it is possible to save lamp's radiation intensity (fig. 4.) (in detail about such light sources in [2]).

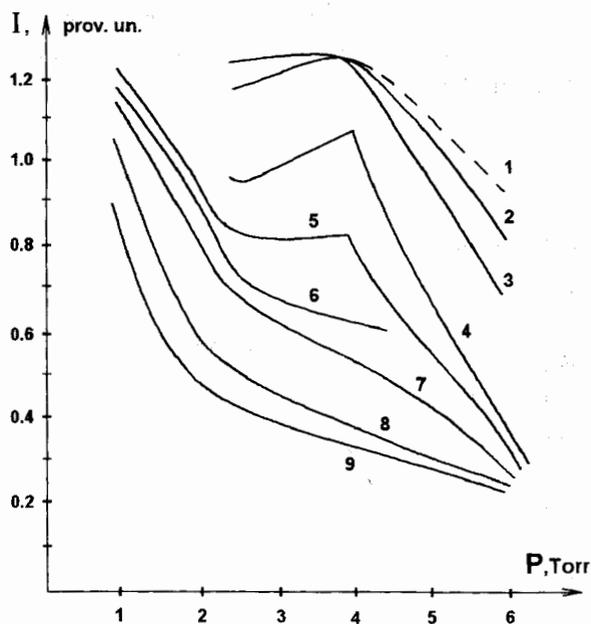


Fig. 4. Dependence of radiation intensity on gas pressure at various temperature of a lamp: 1 - 150°C, 2 - 143°C, 3 - 141°C, 4 - 137°C, 5 - 130°C, 6 - 125°C, 7 - 121°C, 8 - 116°C, 9 - 110°C.

The table

Power of the discharge, W	1.7	1.0	0.7	0.3	0.2	0.1
Durability, $\tau$ , hours	10000*	22165	37846	134894	247816	700927

\* guaranteed durability

### ACCELERATED DETERMINATION OF LAMPS RELIABILITY AND DURABILITY

Researches of gas pressure change in lamps depending on mode of operation allowed to develop technique of accelerated test of lamps at increase of power. The results of measurements for this technique (other techniques of acceleration and their substantiation are indicated in [2,10]) are submitted on fig. 5.

By studying of experiment results dependence of gas absorption speed on the discharge power (fig. 6) was constructed, which within the limits of an error has coincided with settlement, received in [6,10] from the decision of a set of equations, describing kinetic of processes in a lamp at excitation of the discharge of gas and rubidium vapor in it. Thus, as a result of theoretical and experimental researches one can approximate dependence of lamp's service life by following expression [12]

$$\tau = (R^2/W)^{1.7}, \quad (3)$$

where  $R$  is radius of a lamp,  $W$  is power in the discharge.

Using (3), it appeared to have opportunity to develop and include in lamps manufacture process a technique of operative control of lamps quality. From (3) follows the opportunity of quantitatives evalu-

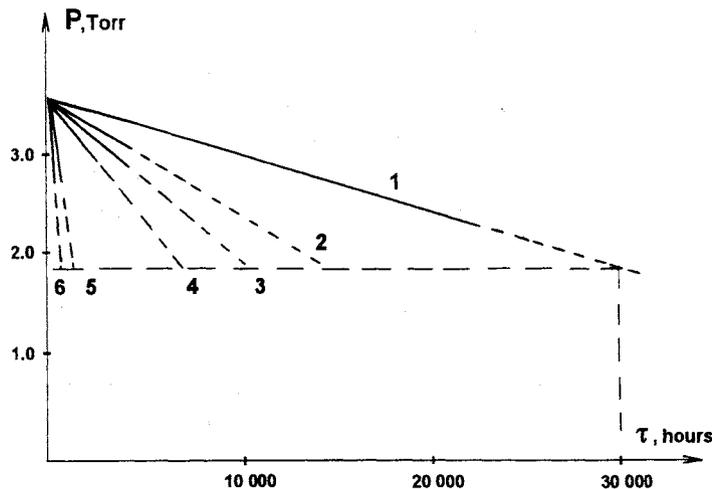


Fig. 5. Change of gas pressure in lamps at different discharge power.  
1 -  $W = 1,7W$ , 2 -  $W = 3W$ , 3 -  $W = 3,5W$ ,  
4 -  $W = 5,8W$ , 5 -  $W = 9W$ , 6 -  $W = 10W$ .

ation of expected lamp's durability at decrease of the discharge power and change of the lamp's size. In contrast of empirical results presented in [18] we received theoretically and experimentally weaker dependence on discharge power, which seems more reasonable.

The calculated data of expected durability of spectral lamps depending on discharge power under condition of certain guaranteed durability at maintainance of the discharge mode, at the decrease of the discharge power are indicated in the table.

Thus, at decrease of the discharge power it is possible to reach practically any given durability. One must note, that the decrease of consumed power up to certain limits does not reduce parameters of a light source (restriction on decrease of power are discussed in [2]) and does not worsen characteristics of a quantum frequency standard, as the factor quality of the discriminator near to an optimum rather poorly depends on the radiation intensity. Besides at the decrease of discharge power there are new opportunities at designing the quantum discriminator (see, for example, [19]), enabling to increase efficiency of use of light.

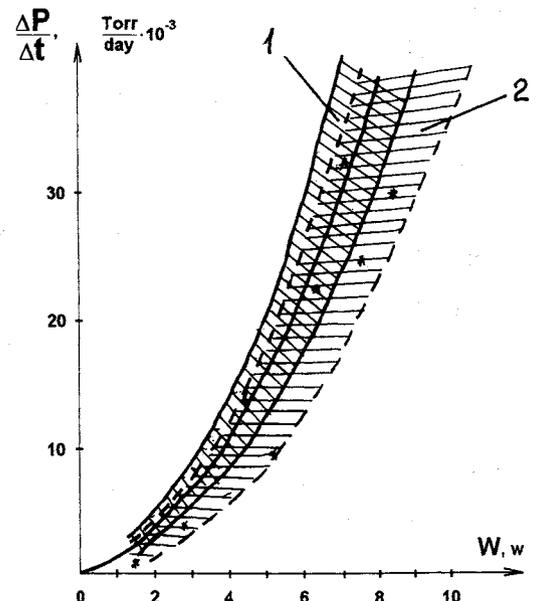


Fig. 6. Dependence of krypton absorption speed in a lamp on the discharge power.  
1 is settlement, 2 is experimental.

## CONCLUSION

Thus, as a result of carried out researches, technological process is developed, which is successfully used by spectral lamps manufacture process. In result the guaranteed durability of lamps makes at serial production corresponds to 5 years in modes of the intensive discharge and 15 years in low-power conditions.

Developed technological process with introduced techniques of the control and dosing permits to create light sources practically with any required durability.

## REFERENCES

- [1] Bell V, Bloom A, Linch J, "Spectral lamps of alkaline metals filled by vapor". The review of scientific instruments. 1961. No 32. p. 588-583.
- [2] Kazantsev S A, Khutorshchikov V I, "High frequency light source of linear spectra". Saint-Petersburg. 1995. 326 p. (russian).
- [3] Semenov S V, Smirnova G M, Khutorshchikov V I, "Physical processes in high-frequency electrodeless spectral lamps with vapor of metals". Referative information about transfers, reviews, deposited manuscripts and theme indexes. N<sup>o</sup> 17. Moscow, 1975.
- [4] Khutorshchikov V I, Luizova L A et al. "Space distribution of atoms in high-frequency electrodeless spectral lamps with vapor of metals" *Optika i spektroskopiya*. 1984. Vol. 56. Vyp. 2. p. 238-243.
- [5] Khutorshchikov V I. "Electrodeless lamps for optical pumping". Thesis. Leningrad. 1976. 176 p.
- [6] Agapov A S, Khutorshchikov V I. "High-frequency electrodeless spectral lamps with vapor of metal. The theory and experiment". *Voproc radioelektroniki. Ser. Obsshii voproc radioelektroniki*. Vyp. 9. Leningrad, 1983. p. 111-118.
- [7] Agapov A S, Matveev A A, Khutorshchikov V I. "Mathematical modeling of processes in high-frequency electrodeless spectral lamps with vapor of metals". *Processes of energy transfer in vapor of metals*. Riga, 1985. p. 89-98.
- [8] Nikolaeva T I, Semenov S V, Khutorshchikov V I. "About measurement of quantity of metal in spectral devices". *Voproc radioelektroniki. Ser. Obshchetekhnicheskaya*. Vyp. 4. Leningrad, 1974. p. 164-167.
- [9] Khutorshchikov V I. "About durability of high-frequency electrodeless spectral lamps with vapour Rb". *Voproc radioelektroniki. Ser. Obshchetekhnicheskaya*. Vyp. 2. Leningrad, 1975. p. 115-118.
- [10] Smirnova G M. "Development of the physics package with improved characteristics based on the rubidium cell". Thesis. Leningrad, 1985. 185 p.
- [11] Khutorshchikov V I et al. A.S. USSR N<sup>o</sup> 1050506.
- [12] Khutorshchikov V I. "About measurement of pressure in high-frequency electrodeless spectral lamps". *Voproc radioelektroniki. Ser. Obshchetekhnicheskaya*. Vyp. 2. Leningrad, 1975. p. 110-114.
- [13] Volk C N, Frueholz R P. "The role of long-term lamp fluctuations in the random walk of frequency behavior of the rubidium frequency standard: a case study". *J. Appl. Phys.* 1985. Vol. 57. p. 980-983.
- [14] Volk C N, Frueholz R F, English T C. a.o. "Lifetime and reliability of rubidium discharge lamps for use in atomic frequency standards". *Proceedings of 38th Annual. Control. Sympos.* 1984. p. 387-400.
- [15] Roshat P, Couplet C, Mileti G, Schweda H, Thoman P, Busca G. "Miniaturized rubidium clocks for space and industrial applications". *Proc. of 49th Annual. Control. Sympos.* 1995.
- [16] Khutorshchikov V I. "Radiospectroscopic method of measurement of gas pressure". Moscow, 1992. 71 p. (russian)
- [17] Kazantsev S A, Smirnova G M, Khutorshchikov V I. "Presized radiospectroscopic method of gas pressure measurement". *Optika i spektroskopia*. 1996. N<sup>o</sup> 1. p. 14-26.
- [18] Oyamada H, Takahashi K, Sato I, Ushida H. "A considaration of rubidium lamp stability for rubidium frequency standard". *NEC Res. and Development*. 1975. p. 340-343.
- [19] Gevorkyan A G, Smirnova G M, Khutorshchikov V I. "A miniaturized rubidium physics package of atomic clock for space air and industrial applications". *Actes proceedings of Ninth EFTF, Besancon, France*. p. 380-383.

## IMPROVEMENT OF A DIGITAL AUTOMATIC CAVITY TUNING SYSTEM IN THE ACTIVE HYDROGEN MASER CH1-75

*N.A.Demidov, V.A.Logachev, G.M.Chernov*

Institute of Electronic Measurements "Kvarz"  
Prospect Gagarina, 176, Nizhny Novgorod, Russia

Possibilities of the active hydrogen maser CH1-75 frequency stability improvement for time intervals 1 h - 1 d due a digital automatic cavity tuning system are discussed. This system uses a hydrogen atom spectral line quality modulation method and a reversible counter measuring two masers beat period for the time interval T at two spectral line quality significances. Two hydrogen maser CH1-75 frequency instability at T = 100 and 1000 s and at different modulation methods (by beam density, by unhomogeneous magnetic field bringing into a storage bulb and by mixed method) was investigated. Frequency instability  $\sigma_y(2, \tau) = (5 - 7) * 10^{-15}$  at T = 100 s and  $(2 - 3) * 10^{-15}$  at T = 1000 s for 1 h (for all modulation types) was got. Results for 1 d frequency stability are showed in the table.

T, s	100	1000	1000	1000
Modulation type	beam	beam	magnetic field	magnetic field + beam
$\sigma_y(2, \tau) * 10^{-15}$	4,1	2,0	2,5	1,5

So frequency stability improvement for  $\tau = 1$  h and 1 d by more than two times is got. This results allowed to change CH1-75 specification and to fix frequency stability  $3 * 10^{-15}$  for 1 h ( instead  $1 * 10^{-14}$  ) and  $3 * 10^{-15}$  for 1 d ( instead  $5 * 10^{-15}$  ).

## Collisional effect with optically pumped atomic clock

Masatoshi Kajita

Communications Research Laboratory

Nukui-kitamachi, Koganei, Tokyo 184, JAPAN

### I. Introduction

Two types of atomic clocks based on the cesium ( $^{133}\text{Cs}$ ) hyperfine transition have been developed, the first type is based on the primary standard employing an atomic beam [1], and the more recent type is based on the atomic fountain [2-4]. The atomic fountain is expected to improve the accuracy and stability of the frequency standard up to  $\delta\nu/\nu < 10^{-16}$ , as the atoms are cooled to several microkelvin.

Recently it has been shown that the collisional frequency shift is a rather serious problem with the atomic fountain. Tiesinga *et al.* have given a theoretical prediction [5] and Gibble *et al.* have observed the frequency shift experimentally [4]. For ultracold collisions, the following novel effects are observed: (a) the broadening of wave packet in three direction becomes large compared to the scale size of the interatomic potential, and (b) the thermal energy is small compared with level shifts caused by collisional interactions. In particular, the broadening of wave packet makes the absolute value of the collisional cross section large, and makes the collisional frequency shift significant.

To our knowledge, there has been no discussion of the case in which the kinetic temperatures for the three spatial degrees of freedom are different. The uncertainty principle of position and momentum is independent in three directions. If the variance of velocity component in one or two directions are small, the atomic wave packet is broadened in those directions even if the variance of total velocity is large. The collisional cross-section is a size on the plane transverse to the relative motion, so it should be large if the broadening of the atomic wave packet is large only in two directions transverse to the relative motion. In this paper, we show that the collisional frequency shift may be a serious problem for the atomic clock based on the atomic beam also; particularly with optically pumped atomic clock. In an atomic beam that passes through two widely separated small holes, the variance of velocity component perpendicular to the beam direction is small. The atomic wave packet is broadened in two directions perpendicular to the atomic beam. Then the collisional cross-section is large, as the relative motion is almost parallel to the atomic beam (effect (a)). On the other hand, the thermal energy is larger than the level shift. The collision dynamics are quite different from those of an ultracold collision and the sign of the frequency shift can be different from that observed with the atomic fountain. It is useful to study the collisions in an atomic beam and compare them with those of the atomic fountain, as only effect (a) is observed with an atomic beam.

### II. Previous collisional cross-section formulas

The Cs frequency standard is based on the transition frequency between  $\alpha = |3,0\rangle$  and  $\beta = |4,0\rangle$  hyperfine states  $|F, M_F\rangle$ . The collisional frequency shift  $(\Delta\nu_{\alpha\beta})$  has been obtained by [5]

$$\Delta\nu_{\alpha\beta} = \sum_{\gamma} (1 + \delta_{\alpha\gamma} + \delta_{\beta\gamma}) n_{\gamma} v \sigma_{\alpha\beta,\gamma} \quad (1)$$

where  $n_{\gamma}$  denotes the density of atoms in the quantum energy state  $\gamma$ , and  $v$  is mean relative velocity.  $\sigma$  is a parameter defined as "collisional cross-section". Collisional cross-section have been given as [6]

$$\sigma_{\alpha\beta,\gamma} = \frac{\pi}{k^2} \sum_{L=0}^{\infty} (2L+1) \sin[\eta_{\alpha\beta,\gamma}(L,k)] \quad (2)$$

where  $\eta(b)$  is the phase shift.  $L$  denotes the quantum number of angular momentum and  $k$  is the wavenumber of the atomic relative motion. When  $(\pi/k)$  is much larger than the scale size of atomic potential ( $d$ ), Eq. (2) corresponds to a semiclassical formula [7]

$$\sigma_{\alpha\beta,\gamma} = \int_0^{\infty} 2\pi b \sin[\eta_{\alpha\beta,\gamma}(b)] db \quad (3)$$

employing an impact parameter  $b (=L/k)$ .

Previously, the phase shift has been obtained +assuming that the angular momentum is deterministic value. Employing Eq. (3) with the same assumption, the phase shift is obtained taking  $b$  as a deterministic value. Actually the angular momentum is not always a deterministic value, and it can be a deterministic value only when the interatomic potentially is spherical symmetrical and the angular momentum can not be changed by the collision. The effect of interatomic potential depends on the form of the atomic wave packet, if the broadening of wave packet is larger than  $d$ . As shown below, the interatomic potential cannot be spherically symmetric in the atomic beam whose transverse velocity component is small enough.

The effective potential ( $P_{av}(x_0, y_0, z_0)$ ), where  $(x_0, y_0, z_0)$  is the position of center of wave packet at a certain time ( $t=0$ ), is obtained as follows.

$$P_{av}(x_0, y_0, z_0) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} V(x, y, z) |C(x_0, y_0, z_0, x, y, z, t=0)|^2 dx dy dz \quad (4)$$

Here  $C(x_0, y_0, z_0, x, y, z, t)$  is the form of wave packet.  $C(x_0, y_0, z_0, x, y, z, t)$  is obtained as follows, assuming that it propagates parallel to  $z$ -axis of  $x, y, z$ -coordinate.

$$\begin{aligned} & C(x_0, y_0, z_0, x, y, z, t) \\ &= \iiint A_x(k_x) A_y(k_y) A_z(k_z) e^{i[\omega t + k_x(x-x_0) + k_y(y-y_0) + k_z(z-z_0)]} dk_x dk_y dk_z \\ & \quad \omega(k_x, k_y, k_z): (\text{atomic energy}) / \hbar \end{aligned} \quad (5)$$

$\omega(k_x, k_y, k_z)$  can be expressed as

$$\begin{aligned} \omega(k_x, k_y, k_z) &= \omega_0 + v_x k_x + v_y k_y + v_z(k_z - k) \\ \omega_0 &= \omega(0, 0, k) \end{aligned}$$

$$v_{x,y} = \left( \frac{\partial \omega}{\partial k_{x,y}} \right)_{k_{x,y}=0}$$

$$v_z = \left( \frac{\partial \omega}{\partial k_z} \right)_{k_z=k} \quad (6)$$

Actually,  $v_{x,y} = 0$  and  $v_z = v$  are obtained. Here we assume

$$\begin{aligned} A_{x,y}(k_{x,y}) &= \frac{1}{\sqrt{2\pi K_{x,y}}} & |k_{x,y}| \leq K_{x,y} \\ &= 0 & |k_{x,y}| > K_{x,y} \\ A_z(k_z) &= \frac{1}{\sqrt{2\pi K_z}} & |k_z - k| \leq K_z \\ &= 0 & |k_z - k| > K_z. \end{aligned} \quad (7)$$

Equation (5) is calculated as

$$\begin{aligned} C(x_0, y_0, z_0, x, y, z, t) \\ = \sqrt{\frac{8}{\pi^3 K_x K_y K_z}} e^{i(\omega_0 t + kz)} \frac{\sin K_x(x - x_0)}{(x - x_0)} \frac{\sin K_y(y - y_0)}{(y - y_0)} \frac{\sin K_z(z - z_0 - vt)}{(z - z_0 - vt)}. \end{aligned} \quad (8)$$

Equation (8) shows that the wave packet propagates only in the  $z$ -direction with velocity  $v$ , so the assumption that the relative motion is only in the  $z$ -direction is not violated even if we consider velocity distribution in  $x, y$ -directions. The broadening of wave packet in  $x, y, z$ -directions are given by  $(\pi/K_{x,y,z})$ , respectively.

Here we consider the atomic collision in an atomic beam flowing in  $Z$ -direction of  $X, Y, Z$ -coordinate. Considering the variance of atomic velocity component in  $X, Y, Z$ -direction ( $V_{x,y,z}$ ), the angle between  $x$ -axis and  $Z$ -axis is in the order of  $\tan^{-1}(V_{x,y}/V_z)$ . With  $(V_{x,y}/V_z) \rightarrow 0$ ,  $x$ -axis is parallel to the  $Z$ -axis. Taking  $x$ -axis parallel to  $Z$ -axis, we assume the following condition.

$$K_x = K_y (=K) \ll \pi/d \ll K_z \approx k \quad (9)$$

The probability distribution is obtained by

$$|C(x_0, y_0, z_0, x, y, z, t = 0)|^2 = D(x_0, y_0, x, y) \delta(z, z_0) \quad (10)$$

$$D(x_0, y_0, x, y) = \frac{4}{K^2 \pi^2} \frac{\sin^2 K(x - x_0)}{(x - x_0)^2} \frac{\sin^2 K(y - y_0)}{(y - y_0)^2} \quad (11)$$

Considering a certain value  $r$ , which is much larger than  $d$ , but much smaller than  $(\pi/K)$ ,  $P_{av}(r, 0, 0)$  is in the same order of  $P_{av}(0, 0, 0)$ . On the other hand,  $P_{av}(0, 0, r)$  is a negligible value. So the effect of interatomic potential is not spherically symmetric and the previous assumption is violated. We need a novel formula of collisional cross-section, calculating the phase shift taking the uncertainty of angular momentum into account. The broadening of wave packet in  $x, y$ -directions are much larger than  $d$ , so the collisional cross-section is much larger than  $\pi d^2$ .

III. Derivation of a formula of the collisional cross-section, that accounts for the form of the incident wave packet

With the condition shown in Eq. (9), the collisional cross section should be obtained with a novel formula, which corresponds to to Eq. (3) with  $K \rightarrow \infty$ . The form of Eq. (3) can be obtained even

without the concept of angular momentum, so it is reasonable to consider that the collisional cross-section is obtained by

$$\sigma = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sin \eta'(x_0, y_0) dx_0 dy_0 = \int_0^{\infty} 2\pi b \sin \eta(b) db, \quad (12)$$

where  $\eta'$  is the phase shift taking the broadening of the wave packet into account, which is obtained by

$$\eta'(x_0, y_0) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \eta(x, y) D(x_0, y_0, x, y) dx dy. \quad (13)$$

Since  $b^2 = x_0^2 + y_0^2$ , Eq. (11) can be rewritten with the polar coordinate [  $(x, y) = (r, \theta)$ ,  $(x_0, y_0) = (b, \theta_0)$  ],

$$D(b, \theta_0, r, \theta) = \frac{4 \sin^2 Kr'}{\pi^2 r'^2} r$$

$$r' = [r^2 + b^2 - 2rb \cos(\theta - \theta_0)]^{1/2}. \quad (14)$$

$\eta'(b)$  is obtained by

$$\eta'(b) = \int_0^{\infty} \int_{-\pi}^{\pi} \frac{4 \sin^2 Kr'}{\pi^2 r'^2} \eta(r) r d\theta dr \quad (15)$$

As  $K \rightarrow \infty$ ,  $\eta'(b)$  and Eq. (12) reduce to the corresponding classical limits given by  $\eta(b)$  and Eq. (3), respectively.

To estimate the collisional cross-section, consider the following simple model.

$$\eta(b) = \eta_0 \quad b \leq d$$

$$= 0 \quad b > d \quad (16)$$

It is considered that  $\eta_0$  is much larger than  $\pi$ . Considering that  $d \ll (\pi/K)$ , Eq. (15) is calculated as

$$\eta'(b) = \eta_0 \frac{4d^2 \sin^2 Kb}{\pi b^2} \quad (17)$$

Then Eq. (12) is calculated as

$$\sigma = \int_0^{\infty} 2\pi b \sin \left( \eta_0 \frac{4d^2 \sin^2 Kb}{\pi b^2} \right) db$$

$$= \frac{2\lambda^2}{\pi} \int_0^{\infty} X \sin \left( \eta_0 \frac{4\pi d^2 \sin^2 X}{\lambda^2 X^2} \right) dX \quad \begin{matrix} X = Kb \\ \lambda = \pi/K \end{matrix} \quad (18)$$

Eq. (18) can be approximated with

$$\sigma = \frac{2\lambda^2}{\pi} \int_0^{\pi/2} X \sin \left( \eta_0 \frac{4\pi d^2}{\lambda^2} \right) dX$$

$$= \frac{\pi\lambda^2}{4} \sin \left( \eta_0 \frac{4\pi d^2}{\lambda^2} \right) \quad (19)$$

If  $\eta_0$  is large enough that  $(4\eta_0 d^2 / \lambda^2)$  is greater than the order of the unit, the collisional cross-section is on the order of  $\pi\lambda^2/8$ .

#### IV. Estimation of collisional frequency shift with an actual primary atomic clock and optically pumped atomic clock

As an example, let us estimate the collisional cross section in an atomic beam, which starts from a beam nozzle (diameter 0.8 mm) and passes through holes given on two microwave cavities (diameters are 3 mm). The frequency is measured with Ramsey spectrum. So, we should discuss the collision in the beam between the two microwave cavities, which has passed the beam nozzle and the hole on the first microwave cavity. The distance between beam nozzle and the first microwave cavity is assumed to be 50 cm. The mean atomic velocity is in the order of 200 m/s, so the velocity component perpendicular to the atomic beam is less than 65 cm/s. Assuming that the wavenumber distribution is shown with Eq. (7),  $K$  is estimated to be  $6.5 \times 10^7 \text{cm}^{-1}$  and  $\lambda$  is  $4.8 \times 10^{-7}$  cm. Considering that  $d$  is in the order of  $10^{-8}$  cm and  $(\pi k)$  is in the order of  $10^{-9}$  cm, Eq. (9) is filled and  $\sigma$  is estimated to be on the order of  $10^{-13}$  cm<sup>2</sup>. The frequency shift is obtained by Eq. (1). The mean relative velocity ( $v$ ) is estimated by the velocity variance. With the primary atomic clock,  $v$  is estimated to be in the order of 30 m/s, as the beam is velocity selected by quadrupole or hexapole magnet. For an atomic density of the order of  $10^7$  cm<sup>-3</sup>, the observed frequency shift is expected to be in the order of 1 mHz. Employing both of quadrupole and hexapole magnet, the velocity variance can be reduced upto several m/s. Then, the collisional frequency shift can be negligible.

With the optically pumped atomic clock, the variance of atomic velocity is in the order of 150 m/s, as there is no velocity selection. So, the frequency shift is estimated to be in the order of 5 mHz with the same atomic density.

Employing two lasers, it is possible to transfer all atoms into either  $|3,0\rangle$  or  $|4,0\rangle$  level. In this case,  $(1 + \delta_{\alpha\gamma} + \delta_{\beta\gamma})$  is 2 (otherwise it is almost 1.1). Then the frequency shift is in the order of 10 mHz. So, the collisional effect is much more serious with optically pumped atomic clock than with primary atomic clock.

#### V. Conclusion

An atomic clock based on an atomic beam passing through two widely separated small holes can be significantly affected by a collisional frequency shift caused by the broadening of the wave packet in the two directions perpendicular to the atomic beam. This effect cannot be derived from previous formulas of collisional cross-sections, so a new collisional cross-section formula that takes the actual form of incident wave packet into account was developed.

With optically pumped atomic clock, this effect is much more serious than with primary atomic clock.

#### Acknowledgment

I thank Dr. I. Shimamura (Riken, JAPAN), Prof. K. Nakagiri (Kinki Univ., JAPAN), Dr. E. Tiesinga (N.I.S.T., U.S.A.) , and Prof. B. J. Verhaar (Eindhoven Univ., Netherlands) for their invaluable comments.

## References

- [1] J. Vanier and C. Audoin, *The Quantum Physics of Atomic Frequency Standards*, Adam Hilger (1989).
- [2] M. Kasevich, E. Riis, S. Chu, and R. DeVoe, *Phys. Rev. Lett.* **63**, 612 (1989).
- [3] A. Clairon, C. Salomon, S. Guellati, and W.D. Phillips, *Europhys. Lett.* **16**, 165 (1991).
- [4] K. Gibble and S. Chu, *Phys. Rev. Lett.* **70**, 1771 (1993).
- [5] E. Tiesinga, B.J. Verhaar, H.T.C. Stoof, and D. van Bragt, *Phys. Rev. A* **45**, R2671 (1992).
- [6] J. Schmiedmayer, M. S. Chapman, C. R. Ekstrom, T. D. Hammond, S. Wehinger, and D. E. Pritchard, *Phys. Rev. Lett.* **74**, 1043 (1995).
- [7] J. F. Kielkopf, *J. Phys. B* **9**, L547, (1976).

## A REGULARIZATION APPROACH FOR THE DETERMINATION OF THE TIME OF FLIGHT DISTRIBUTIONS IN THE CESIUM BEAM STANDARDS

A. Makdissi and E. de Clercq

*Laboratoire Primaire du Temps et des Fréquences  
Bureau National de Métrologie - Observatoire de Paris  
61, Avenue de l'Observatoire  
75014 Paris - France  
E-mail: Alaa.Makdissi@obspm.fr*

### ABSTRACT

In this paper we describe a new method to directly determine from the shape of a Ramsey pattern the real microwave level present in the cavity of a beam frequency standard. Then a new method to evaluate the atomic time of flight distribution using a regularization approach. The validity of this method is confirmed by experimental results.

### I - INTRODUCTION

In thermal beam frequency standards a precise knowledge of the atomic Time Of Flight (TOF) distribution and of the real microwave (MW) level in the cavity is essential to properly evaluate most of the effects which shift the clock transition [1]. The methods presented in this paper allow a straightforward determination of the actual MW level and of the TOF distribution from experimental records.

The observed signals in a caesium beam standard are of the form :

$$s(\omega, b) = \int_0^{\infty} P(\tau, \omega, b) f(\tau) d\tau \quad (1)$$

where :

- $\omega$  is the detuning from the atomic resonance pulsation.
- $b$  is the Rabi pulsation (rad/s), related to the MW magnetic induction amplitude  $B$  by
 
$$b = \mu_B B / \hbar \quad (2)$$
- $\tau$  is a random variable representing the TOF of one atom across one arm of the Ramsey cavity.
- $f(\tau)$  is the TOF distribution or the Probability Density Function of  $\tau$ .
- $P(\tau, \omega, b)$  is the Ramsey transition probability given in [1] by:

$$P(\tau, \omega, b) = \frac{4b^2}{\Omega^2} \sin^2\left(\frac{\Omega}{2}\tau\right) \left[ \cos\left(\frac{\Omega\tau}{2}\right) \cos\left(\frac{a}{2}\omega\tau\right) - \frac{\omega}{\Omega} \sin\left(\frac{\Omega\tau}{2}\right) \sin\left(\frac{a}{2}\omega\tau\right) \right]^2 \quad (3)$$

where :

- $\Omega^2 = b^2 + \omega^2$ .
- $a$  is the ratio of the drift region length  $L$  to the excitation region length  $l$ .

The curve  $s(\omega, b)$  versus  $\omega$  at fixed MW level is called the Ramsey pattern.

### II - ESTIMATION OF $b$

Equation (2) shows that the Rabi pulsation  $b$  is proportional to the MW amplitude  $B$  inside the cavity which can not be measured directly. Instead,  $B$  is proportional to the square root of injected MW power which can be measured. Then we can write :

$$b = \alpha \sqrt{\text{The MW power injected into the cavity}} \quad (4)$$

where  $\alpha$  is an unknown coefficient which can be calculated if we know at least one value of  $b$  and the corresponding measured MW power.

At high MW level ( $b \gg \omega$ ), equation (1) may be approximated by :

$$s(\omega, b) = \frac{1}{2} \int_0^{\infty} \sin^2(b\tau) \cos(a\omega\tau) f(\tau) d\tau + C_1(b) \quad (5)$$

where  $C_1(b)$  is a constant versus frequency and it depends on the value of  $b$ .

Let  $CT[g(x)]$  denote the *Cosine Transform* operator of a function  $g(x)$ ,  $x \in IR^+$  defined by :

$$G(\omega) = CT[g(x)](\omega) = \int_0^{\infty} g(x) \cos(\omega x) dx \quad (6)$$

Then equation (5) can be written using the properties of the Cosine transform (which are similar to those of the Fourier transform) as :

$$s(\omega, b) = \frac{1}{2} CT \left[ f\left(\frac{\tau}{a}\right) \sin^2\left(b\frac{\tau}{a}\right) \right] (\omega) + C_1(b) \quad (7)$$

$$= \frac{1}{2} F(a\omega) * \left[ \delta(\omega) - \frac{1}{2} \delta\left(\frac{2b}{a} - \omega\right) - \frac{1}{2} \delta\left(\frac{2b}{a} + \omega\right) \right] + C_1(b)$$

where :

- $F(\omega)$  is the Cosine Transform of the TOF distribution  $f(\tau)$ .

- $\delta(\omega)$  is the Dirac function.
- $*$  is the convolution operator.

Equation (7) shows that the Ramsey pattern is the convolution of  $F(a\omega)$  with three Dirac pulses at  $\omega_0 = 0$ ,  $\omega_{-1} = -2b/a$  and  $\omega_1 = 2b/a$ . Estimating the positions of these Dirac pulses allows us to determine the Rabi pulsation in the MW cavity by :

$$b = a \frac{\omega_1 - \omega_0}{2} \quad (8)$$

This relation was mentioned by Ramsey [2] without mathematical demonstration.

Figure 1 shows an example of an experimental Ramsey pattern from the LPTF Cs beam clock. The three arrows show the position of the Dirac pulses. The horizontal line represents the value of the constant  $C_1(b)$ . In this case, we calculate  $b = 205500$  rad/s which is about six times  $b_{opt}$ , the value of  $b$  maximizing the signal at resonance.

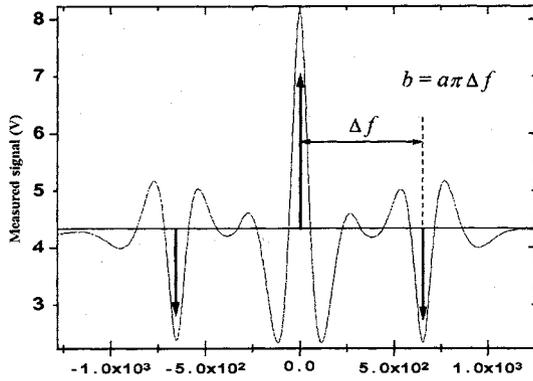


Fig. 1 : Experimental Ramsey Pattern

In the case of a not very large value of  $b$  compared to  $\omega$  the value of  $b$  can be estimated using the *Cepstrum* method [3].

### III - INVERSE PROBLEM

Boulinger [4] showed that the TOF distribution can be evaluated from the dependence of the amplitude of the central maximum of the Ramsey resonance on the Rabi pulsation  $b$ , called the  $I(b)$  function ( $I(b) = s(0, b)$ ). He used a FFT algorithm and a fit to an experimental Ramsey pattern. Here we use the approach of inverse problem to evaluate the TOF distribution from an  $I(b)$  function.

Let  $\omega = 0$  in (1), then the equation of  $I(b)$  is :

$$I(b) = \int_0^{\infty} \frac{1}{2} (1 - \cos(2b\tau)) f(\tau) d\tau. \quad (9)$$

This could be seen as a *Fredholm* integral equation of the first kind [5]:

$$I(b) = \int_0^{\infty} K(b, \tau) f(\tau) d\tau \quad (10)$$

where the left-hand side  $I$  and the kernel  $K$  are known, and where  $f$  is the unknown solution. Equation (10) is *ill-posed* in the sense that the solution  $f$  is potentially very sensitive to perturbations in the measured signal  $I$ . The discretization of (10) over a finite support for  $\tau$  ( $0 \leq \tau \leq \tau_{max}$ ) in the Riemann sense leads to :

$$I(b) = \sum_{i=1}^n K(b, i\Delta\tau) f(i\Delta\tau) \Delta\tau \quad (11)$$

Putting measured data  $I(b)$  in a vector  $\mathbf{y}$  we obtain a linear system of equations

$$\mathbf{y} = \mathbf{A}\mathbf{x}, \quad \mathbf{A} \in \mathbb{R}^{m \times n}, \quad m > n \quad (12)$$

where :

- $\mathbf{x}$  is a vector of  $n$  unknowns representing the values of  $f$  in a finite set of times of flight :

$$x_i = f(i\Delta\tau), \quad i \in \{0, 1, \dots, n-1\}$$

- $\mathbf{y}$  is a vector of  $m$  measured data at  $m$  different Rabi pulsations ( $y_i = I(b_i) + e_i$ , where  $e_i$  is the noise contribution to the measure  $y_i$ ).

- $\mathbf{A}$  is a real matrix so that :

$$a_{i,j} = K(b_i, (j-1)\Delta\tau) \Delta\tau, \quad 1 \leq i \leq m, 1 \leq j \leq n.$$

Solving (8) in the sense of least-squares leads to solutions that minimise the residual norm :

$$\mathbf{x}_{lsq} = \arg \min \{ \|\mathbf{A}\mathbf{x} - \mathbf{y}\| \} \quad (13)$$

which could lead to non meaningful solutions because of noisy  $\mathbf{y}$  and of the ill-posedness nature of the problem. Hence, it is necessary to incorporate further information about the desired solution in order to stabilise the problem and to single out a useful and stable solution. This is the purpose of regularization.

### IV - REGULARIZATION APPROACH

In order to reduce the effect of noise which will yield oscillating components on the solution we require from the solution to be smooth. Hence, the additional information of smoothness is incorporated by adding a side constraint  $C(\mathbf{x})$  to (13) which involves the minimisation of the first or higher order derivative of the solution to prevent its rapid and oscillatory variations. For example, in the case of a side constraint representing the first order derivative of the solution  $\mathbf{x}$   $C(\mathbf{x})$  is written as :

$$C(\mathbf{x}) = \sum_{i=1}^{n-1} (x_{i+1} - x_i)^2. \quad (14)$$

This is the most common form of regularization and is known as the *Tikhonov regularization* [6]. Here, the regularised solution  $\mathbf{x}_\lambda$  is the minimizer of the following weighted combination of the residual norm and the side constraint:

$$\mathbf{x}_\lambda = \arg \min \{ \|\mathbf{A}\mathbf{x} - \mathbf{y}\| + \lambda C(\mathbf{x}) \} \quad (15)$$

where the regularization parameter  $\lambda$  controls the weight given to the minimisation of the side constraint relative to the minimisation of the residual norm. Clearly,  $\lambda$  should be chosen with care and a variety of choices strategies have been proposed in the literature [7]. The most obvious strategy (in the case of knowledge of the variance of the noise signal or the norm of the noise vector  $\mathbf{e}$ ) consists of choosing  $\lambda$  such that the residual norm of the desired solution  $\mathbf{x}_\lambda$  satisfies

$$\|\mathbf{A}\mathbf{x}_\lambda - \mathbf{y}\| = \|\mathbf{e}\|. \quad (16)$$

Another type of side constraints  $C(\mathbf{x})$  could be incorporated. When a solution with positive elements is sought, the following non linear function is used as side constraint :

$$C(\mathbf{x}) = \sum_{i=1}^n x_i \log(w_i x_i) \quad (17)$$

where  $w_1, \dots, w_n$  are  $n$  weights. This is called the *Maximum Entropy regularization* (because  $-C(\mathbf{x})$  measures the entropy of  $\mathbf{x}$ ) and it has less immunity to noise than the Tikhonov regularization.

Numerically, computing a solution of (15) may be done either by iterative methods like the *conjugate gradient* (CG) algorithm or by nearly direct methods like the *generalized singular value decomposition* (GSVD) [8].

## V - EXPERIMENTAL RESULTS

We have applied the previous approach to evaluate the TOF distribution in the LPTF caesium beam standard. At first we have deduced the value of the coefficient  $\alpha$  in (4) from the Ramsey pattern of Fig. 1 as explained in section II.

The experimental Ramsey signal of the 0-0 transition at resonance versus the Rabi pulsation  $b$  (the  $I(b)$  function) is shown on Fig. 2. The values of  $b$  are calculated from the measured MW power using equation (4). In order to eliminate contributions of neighbouring transitions ( $m \neq 0$ ) to  $I(b)$  at high MW power a static magnetic field larger than that for normal operation is required. Here the different MW  $\Delta m = 0$  transitions are spaced by 250 kHz.

The TOF distribution extracted from the previous curve using the regularization approach in the case of a first order derivative side constraint (14) and a regularization parameter  $\lambda = 10^{-4}$  is shown in Fig. 3 (dashed line). In our standard [9] the length of the interaction region is  $l = 10.2$  mm and the ratio  $a$  is  $L/l = 100$ . The oven temperature is  $93.5^\circ\text{C}$ . The Cs beam is optically pumped and a cycling transition ( $F = 4 \Leftrightarrow F = 5$ ,  $D_2$  line) is used for the detection. The expected TOF distribution (solid line in Fig. 3) is then a Maxwellian weighted by  $\tau$  distribution [10].

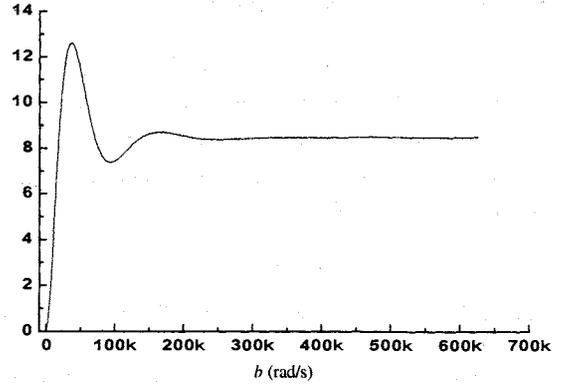


Fig. 2 : Experimental  $I(b)$  function.

The TOF distribution obtained by Boulanger method is also shown in Fig. 3 (dotted line) where all the curves are normalized so that :

$$\int_0^{\infty} f(\tau) d\tau = 1. \quad (18)$$

Zooming in these curves shows that the TOF obtained by the regularization approach is null for  $\tau$  greater than  $220 \mu\text{s}$ , while the Boulanger method gives a flat function for high values of  $\tau$ . This flatness is related to the FFT of the noise (white noise).

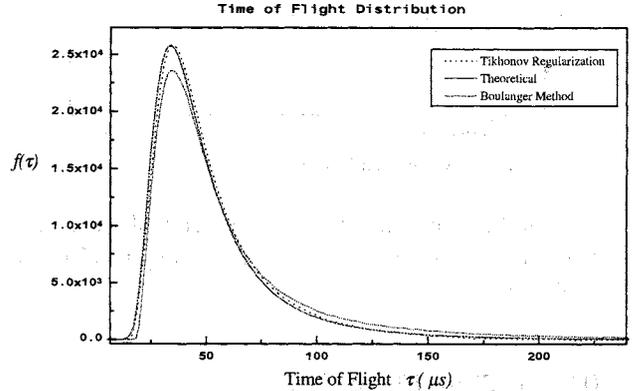


Fig. 3 : The TOF distribution calculated by regularization approach compared to the theoretical Maxwellian weighted by  $\tau$  distribution and that calculated by Boulanger method

## VI - CONCLUSION

A new and precise method for estimation of the Rabi pulsation in caesium beam frequency standards has been presented. Its accuracy is related to the precision of the instrument used to measure the MW power.

The *regularization* approach for evaluation of the TOF distribution is introduced. It has more immunity to noise than existing methods and does not require uniformly sampled (versus  $b$ ) data, and hence it avoids the interpolation of noisy data.

**VII - REFERENCES**

1. J. Vanier and C. Audoin, 1989, "The Quantum Physics of Atomic Frequency Standards", Adam Hilger, Bristol, UK
2. N. F. Ramsey, 1956, "Molecular Beams", Oxford University Press, Oxford, UK
3. V. Oppenheim & R. W. Schaffer, 1989, "Discrete-Time Signal Processing", Prentice-Hall
4. J.-S. Boulanger, 1986, "A New Method for the Determination of Velocity Distributions in Cesium Beam Clocks", Metrologia, 23, 37-44
5. C. W. Groetsch, 1984, "The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind", Pitman, Boston
6. A. N. Tikhonov & V. Y. Arsenin, 1977, "Solutions of Ill-Posed Problems", Winston & Sons, Washington, D.C
7. H. W. Engl & J. Gfrerer, 1988, "A posteriori parameter choice for general regularization methods for solving linear ill-posed problems", App. Numerical Math., 4, 395-417
8. P. C. Hansen, 1989, "Regularization, GSVD and truncated GSVD", BIT 29, 491-504
9. E. de Clercq et al., 1993, "The LPTF Optically Pumped Primary Frequency Standard", IEEE Trans. Instrum. Meas., vol. 42, NO. 2, 457-461
10. A. Hamel et al., 1990, "Velocity distribution measurement in an optically pumped cesium beam resonator", in Proc. 4<sup>th</sup> Europ. Freq. And Time Forum, 169-174

# THEORETICAL DESCRIPTION AND EXPERIMENTAL EVALUATION OF THE EFFECT OF THE INTERROGATION OSCILLATOR FREQUENCY NOISE ON THE STABILITY OF A PULSED ATOMIC FREQUENCY STANDARD.

G. Santarelli, Ph. Laurent and A. Clairon.

BNM/Laboratoire Primaire du Temps et des Fréquences,  
Observatoire de Paris, 61, Av. de l'Observatoire, 75014 Paris, France.

Giorgio.Santarelli@obspm.fr

G. J. Dick, C. A. Greenhall.

Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, CA 911109,  
U.S.A.

C. Audoin.

Laboratoire de l'Horloge Atomique, Unité Propre du CNRS associée à l'Université de Paris Sud, Bat. 221,  
Université de Paris Sud, 91405 Orsay Cédex, France.

## ABSTRACT.

In this paper we evaluate the effect of the interrogation oscillator frequency noise on the stability of a pulsed atomic frequency standard such as an atomic fountain or ion trap frequency standard. The atomic response to a phase perturbation in Ramsey and multi-Rabi interrogation schemes has been calculated using the density matrix formalism. The results of these calculations are used to obtain a simple model for the limitation of the frequency standard stability. An experimental evaluation of this effect has been performed by using the Cs atomic fountain frequency standard. Possible means to reduce this effect are considered.

## 1 INTRODUCTION.

The development of new passive frequency standards using trapped ions or cold atoms has produced devices with a potential fractional frequency stability of the order of  $10^{-13}\tau^{-1/2}$  or better. In this new type of standards, the internal interrogation process and the control of the interrogation oscillator are periodic, with period  $T_c$ . The frequency of this oscillator is compared to that of the atomic resonance during a part of the operating cycle only and its frequency is controlled at the end of each cycle.

In the late eighties, one of us [1] at the Jet Propulsion Laboratory derived the atomic response to the oscillator frequency fluctuations using a geometrical approach. Furthermore, he has shown that the oscillator frequency noise at Fourier frequencies which are close to multiples of  $1/T_c$  is down-converted, leading to a degradation of the long-term frequency stability.

This work has been the motive for new thoughts on this subject and the present paper reports briefly a part of the results achieved.

The atomic response to a frequency variation in the interrogation oscillator is deduced from a quantum mechanical treatment for different interrogation schemes using the Ramsey and the multi-Rabi [2] methods.

The equation giving the frequency stability limitation arising from the oscillator frequency noise down-conversion is derived from simple reasoning. The validity of this formula is verified experimentally. Finally, possible means of reducing this spurious effect are considered.

## 2 SENSITIVITY FUNCTION TO OSCILLATOR FREQUENCY FLUCTUATIONS: A RECALL.

Let  $\delta P$  be the change of the probability that a transition occurred, at the outcome of the atom interaction with the R.F. field. If this change is the result of a fluctuation,  $\delta\alpha(t)$ , of the frequency of the field during the interaction, we have [1]:

$$\delta P = \frac{1}{2} \int_{\text{int.}} g(t) \delta\omega(t) dt \quad (1)$$

This equation defines  $g(t)$ , the sensitivity function to frequency fluctuations. The integration holds during the time  $T_i$  of the interaction.

It is worth noting that the R.F. field is frequency modulated. Its (angular) frequency is  $\alpha(t) + \omega_m$  or  $\alpha(t) - \omega_m$  according to the half period of modulation considered.  $\alpha(t)$  is very close to the atom resonance frequency  $\omega_0$  and  $\omega_m$  is the modulation depth. In Eq. (1),  $\delta\alpha(t)$  is the fluctuation of  $\alpha(t)$  and  $g(t)$  is the sensitivity function at the frequency  $\omega_0 \pm \omega_m$ .

It has been shown that  $g(t)$  can be calculated by introducing an infinitesimally small phase step  $\varepsilon$  at time  $t$  in the oscillator frequency. It produces the change  $\delta P(t, \varepsilon)$  of the probability that a transition occurred and  $g(t)$  is given by:

$$g(t) = 2 \lim_{\varepsilon \rightarrow 0} \delta P(t, \varepsilon) / \varepsilon \quad (2)$$

In frequency standards considered here the function  $g(t)$  is not a constant during each cycle ( $T_c$ ). As explained in the following this causes the degradation of the frequency stability of the locked oscillator.

### 3 FUNDAMENTAL LIMITATION OF THE FREQUENCY STABILITY DUE TO SAMPLING.

#### 3.1 Origin of the effect.

The control loop being closed, frequency corrections are applied to the interrogation oscillator at discrete times  $t_k$  at the end of each cycle. Immediately after this adjustment, the frequency offset is  $\Delta\omega_k(t_k)$ . Between the instants  $t_k$  and  $t_{k+1}$ , the frequency of the oscillator varies freely and the offset  $\Delta\omega(t) = \omega(t) - \omega_0$  of the R.F. field is given by :

$$\Delta\omega(t) = \Delta\omega_s(t_k) + \Delta\omega_f(t) - \Delta\omega_f(t_k) \quad t_k < t < t_{k+1} \quad (3)$$

The subscripts  $s$  and  $f$  stand for slaved and free, respectively. The frequency offsets considered are those of the R.F. field when the deviation due to the modulation is not taken into consideration.

The variation  $\Delta\omega(t)$  of the frequency of the interrogation field induces a change in the probability that the transition occurred and this adds a term to the error signal. This term can be interpreted as being produced by a false frequency offset  $\delta\omega_f(t_k)$  of the applied field, which is given by :

$$\delta\omega_f(t_k) = \frac{1}{g_0 T_c} \int_{t_k}^{t_k + T_c} g(t) [\Delta\omega_f(t) - \Delta\omega_f(t_k)] dt, \quad (4)$$

where  $g_0$  is the mean value of  $g(t)$  during the same cycle. The frequency control loop includes a numerical integration. It can be shown that its equation is the following :

$$\begin{aligned} & \Delta\omega_s(t_k) - (1 - \beta)\Delta\omega_s(t_{k-1}) + \beta \Delta\omega_s(t_{k-2}) \\ & = \Delta\omega_f(t_k) - \Delta\omega_f(t_{k-1}) - \beta [\delta\omega_f(t_{k-1}) + \delta\omega_f(t_{k-2})] \end{aligned} \quad (5)$$

where the optical detection noise is not considered, and  $\beta$  is the open loop gain. For slow frequency fluctuations, Eq. (5) shows that we have :

$$\Delta\omega_s(t_k) = -\delta\omega_f(t_k) \quad (6)$$

In Eq.(4),  $g(t)/g_0$  is a periodic function of time with frequency  $1/T_c$ . Therefore the spectral components of  $\Delta\omega_f(t)$  around frequencies  $m/T_c$  will be translated to the very low frequencies. This is the equivalent of an aliasing phenomenon in a sampling process.

#### 3.2 Derivation of the formula for the long term frequency stability limitation.

We will consider the limitation of the frequency stability, for observation times  $\tau$  sufficiently larger than the servo-loop time constant. For this analysis we

assume that the bandwidth of the down-converted noise,  $\Delta f$ , is of the order of magnitude of  $1/\tau$ , and thus smaller than  $1/T_c$ .

The low frequency noise in the bandwidth  $\Delta f$  originates from spectral components of the oscillator noise around Fourier frequencies  $m/T_c$ . Then, it suffices to consider that part of the oscillator frequency noise which is filtered in a set of spectral windows centred around frequencies  $m/T_c$  and having a noise bandwidth  $2\Delta f$ . The Rice representation [3] of this narrow band limited noise is the following :

$$\begin{aligned} \Delta\omega_f^F(t) = \\ \sum_{m=1}^{\infty} \left[ p_m(t) \sin 2\pi m \frac{t-t_k}{T_c} + q_m(t) \cos 2\pi m \frac{t-t_k}{T_c} \right] \end{aligned} \quad (7)$$

where  $p_m(t)$  and  $q_m(t)$  are slowly variable random amplitudes. The constant phase,  $-2\pi m t_k/T_c$  is introduced for convenience and does not affect the final result. The one-sided power spectral density of  $p_m(t)$  and  $q_m(t)$  is related to that of  $\Delta\omega_f(t)$  around  $m/T_c$  by:

$$S_{p_m}(f \leq \Delta f) = S_{q_m}(f \leq \Delta f) = 2S_{\Delta\omega_f}(m/T_c) \quad (8)$$

The frequency offset  $\delta\omega_f(t)$  is calculated by substituting  $\Delta\omega_f^F(t)$  for  $\Delta\omega_f(t)$  in Eq. (4). According to the assumptions made,  $p_m(t)$  and  $q_m(t)$  vary very little during the time interval  $T_c$ . Therefore, they can be taken out of the integral sign and we obtain :

$$\delta\omega_f(t_k) = -\Delta\omega_f^F(t_k) + \frac{1}{g_0} \sum_{m=1}^{\infty} [g_m^s p_m(t_k) + g_m^c q_m(t_k)] \quad (9)$$

with

$$\begin{pmatrix} g_m^s \\ g_m^c \end{pmatrix} = \int_0^{T_c} g(\theta) \begin{pmatrix} \sin 2\pi m \theta / T_c \\ \cos 2\pi m \theta / T_c \end{pmatrix} d\theta \quad (10)$$

where  $g(\theta)$  is defined during the time interval  $[0, T_c]$ . Referring to Eq. (6), the second term of the right-hand side of Eq. (9) represents the low frequency noise of the slaved oscillator resulting from the aliasing. The power spectral density of these fluctuations is easily derived from Eqs. (8) and (9), and the related Allan variance is given by:

$$\sigma_{y\text{lim}}^2(\tau) = \frac{1}{\tau} \sum_{m=1}^{\infty} \left( \frac{g_m^c{}^2}{g_0^2} + \frac{g_m^s{}^2}{g_0^2} \right) S_y^f(m/T_c) \quad (11)$$

where  $S_y^f(m/T_c) = S_{\Delta\omega_f}^f(m/T_c) / \omega_0^2$  is the one-sided power spectral density of the relative frequency

fluctuations of the free running interrogation oscillator at Fourier frequencies  $m/T_c$ .

#### 4 EXAMPLE OF THE FREQUENCY SENSITIVITY FUNCTION AND OF THE LEVEL OF FREQUENCY STABILITY LIMITATION.

##### 4.1 Principle of the calculation of $g(t)$ .

As shown in [4], the change of the quantum state of the atoms interacting with a R.F. field of given amplitude and phase can be represented in a matrix form. We have, in general :

$$\begin{pmatrix} a_1(\theta) \\ a_2(\theta) \\ a_3(\theta) \end{pmatrix} = R \begin{pmatrix} a_1(0) \\ a_2(0) \\ a_3(0) \end{pmatrix} \quad (12)$$

where  $a_1$  and  $a_2$  denote the atomic coherence and  $a_3$  the population difference of the two levels involved in the transition. The column matrices at the right and at the left represent the atom properties at the beginning and at the end of an interaction of duration  $\theta$ , respectively.  $R$  is a  $3 \times 3$  matrix whose elements depend on  $\theta$ , the R.F. amplitude and the phase.

Therefore, the population difference of atoms submitted to various amplitude and phase conditions during their interaction with the magnetic R.F. field can be calculated from matrix products. The effect of the small phase step  $\varepsilon$  occurring at a given instant during the interaction can be expressed easily. The subsequent change  $\delta P(t, \varepsilon)$  of the probability that a transition took place during the interaction follows and  $g(t)$  is obtained using Eq. (2). In the event that the R.F. amplitude is not a constant during the interaction, two different methods can be implemented. One may divide the interaction time into elementary intervals during which the amplitude is assumed a constant. Or, the differential equations describing the evolution of  $a_1$ ,  $a_2$  and  $a_3$  (see Eq. 5.2.20 of [4]) may be integrated numerically.

##### 4.2 Ramsey method of interrogation

This method is applied in the cesium fountain at LPTF [5]. The atoms experience successively two R.F. fields, each for the time interval  $\tau$ , and there is no R.F. field applied between these two partial interactions, for the time  $T$ . Assuming  $T \gg \tau$  and that the magnetic R.F. field is a constant, represented by the Rabi frequency  $b$ , we have :

$$g(\theta) = \begin{cases} 0 & 0 \leq \theta \leq t_p \\ d \sin b\theta & t_p \leq \theta \leq t_p + \tau \\ d \sin b\tau & t_p + \tau \leq \theta \leq t_p + T + \tau \\ d \sin b(T + 2\tau - \theta) & t_p + T + \tau \leq \theta \leq t_p + T + 2\tau \\ 0 & t_p + T + 2\tau \leq \theta \leq T_c \end{cases} \quad (13)$$

where  $d = -\sin \Omega_0 T \sin b\tau$  and  $\Omega_0 = \pm \omega_m$  according to the half period of modulation considered. Fig. 1 shows the variation of  $g(\theta)$  (assumed centred during the cycle, solid line) and of the related coefficients  $g_m/g_0$ , for  $b\tau = \pi/2$  which provides the optimal condition of interrogation. It is assumed  $\tau = 0.015$  s,  $T = 0.5$  s and  $T_c = 1$  s. The function  $g(\theta)$  for  $b\tau = 3\pi/2$  is shown also (dashed line).

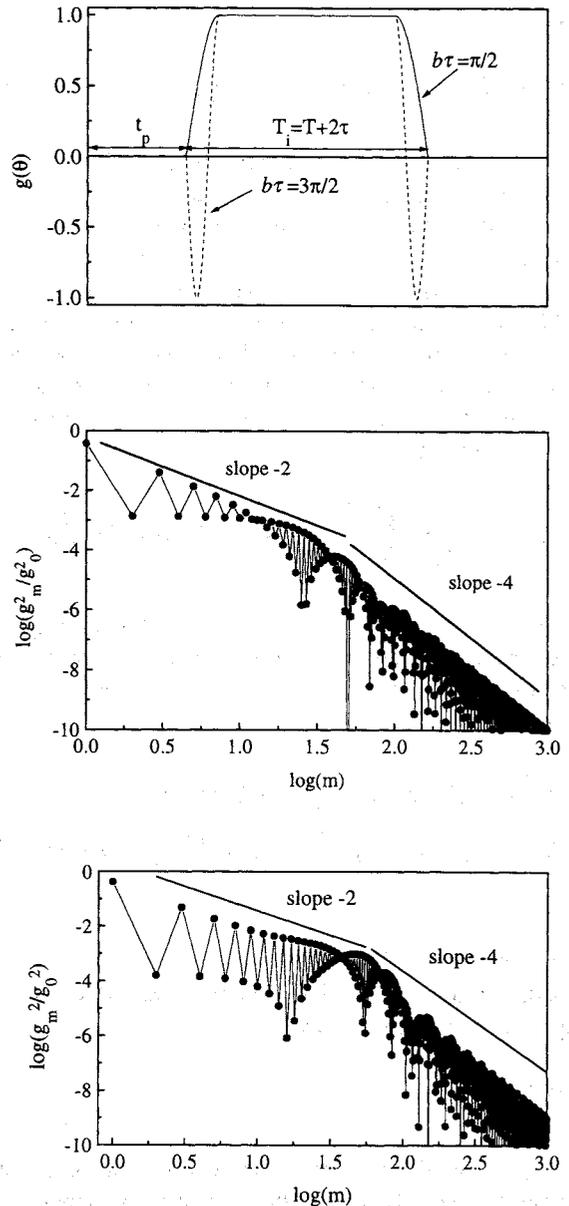


Fig. 1 The function  $g(\theta)$  and the related spectrum in the case of a Ramsey interrogation scheme for  $b\tau = \pi/2$  and  $b\tau = 3\pi/2$ .

In order to assess the level of the frequency stability limitation due to the aliasing, we assume that the oscillator is a VCXO and that the power spectral density of its relative frequency fluctuations is given by:

$$S_y^f(f) = 3.2 \cdot 10^{-29} f^2 + 1.0 \cdot 10^{-27} f + 3.2 \cdot 10^{-26} / f \quad (14)$$

This hypothetical oscillator shows a flicker floor at the level  $2.1 \cdot 10^{-13}$ .

For the Ramsey method of interrogation, we have:

$$\sigma_{y,\text{lim}}(\tau) = 1.17 \cdot 10^{-13} \tau^{-1/2} \quad (15)$$

for the given values of the parameters.

#### 4.3 Interrogation in a $TE_{01n}$ cavity

In the PHARAO project [6], it is planned to launch balls of cold cesium atoms along the axis of a cylindrical cavity tuned to a  $TE_{01n}$  resonant mode. During their interrogation, the atoms experience a magnetic R.F. field whose amplitude is proportional to  $\sin(n\pi\theta/T_i)$ , where  $T_i$  is the total interaction time. In that case,  $g(t)$  is calculated numerically.

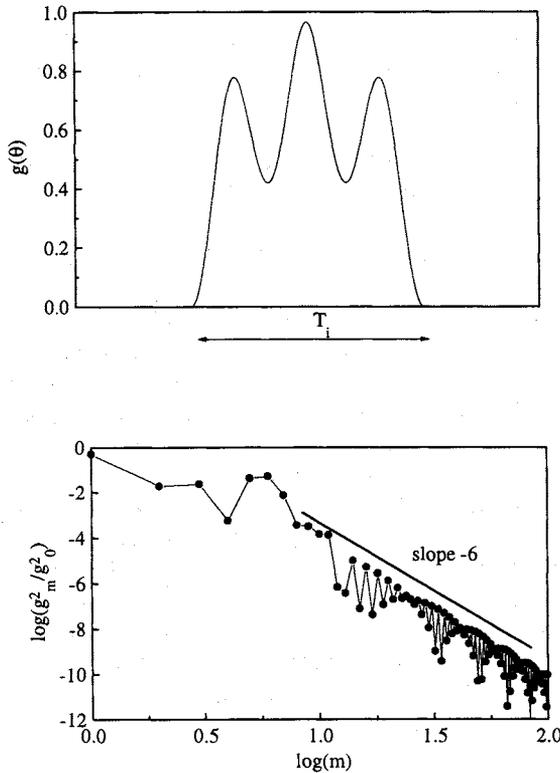


Fig. 2 The function  $g(\theta)$  and the related spectrum in the case of a multi-Rabi interrogation scheme.

Fig. 2 shows, for  $n=3$ , the variation of  $g(\theta)$  (assumed centred during the cycle) and of the related coefficients  $g_m/g_0$ . Here, the operating parameters are chosen so as to provide the steepest slope of the resonance curve. This is achieved for  $b_c T_i/n = 3.66$  and  $\omega_n T_i = 2.31$ , where  $b_c$  is the Rabi frequency at an anti-node of the magnetic field.

It is worth noting that, for  $m$  larger than about 10,  $(g_m/g_0)^2$  decreases as  $m^{-6}$ . This property provides a very

good immunity against the white phase noise of the oscillator. For  $T_i = 0.53$  s and  $T_c = 1$  s, we have:

$$\sigma_{y,\text{lim}}(\tau) = 1.35 \cdot 10^{-13} \tau^{-1/2} \quad (15)$$

which is similar to the value given in Eq. (13).

#### 5. EXPERIMENTAL EVALUATION OF THE FREQUENCY STABILITY DEGRADATION.

In order to verify the model and provide evidence of the aliasing effect, we have made various measurements with an oscillator voluntarily degraded with different types of frequency noise. For this, we use the LPTF Cs atomic fountain standard [5].

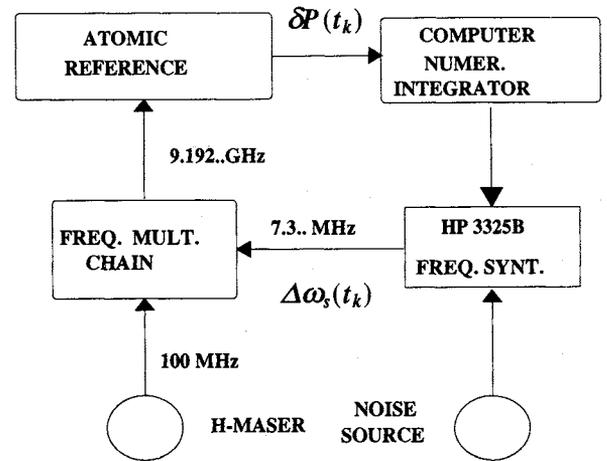


Fig. 3 Schematic of the atomic fountain frequency

We have used three different sources of noise. Firstly a white noise voltage in the range 0.1 Hz-1600 Hz ( $f^0$ ) with the possibility of using different low-pass filters. Secondly a flicker noise generator  $f^{-1}$  in the range 0.5-100 Hz and thirdly a generator with spectral density proportional to  $f^{-3}$ , for Fourier frequencies from 0.5 to 100 Hz.

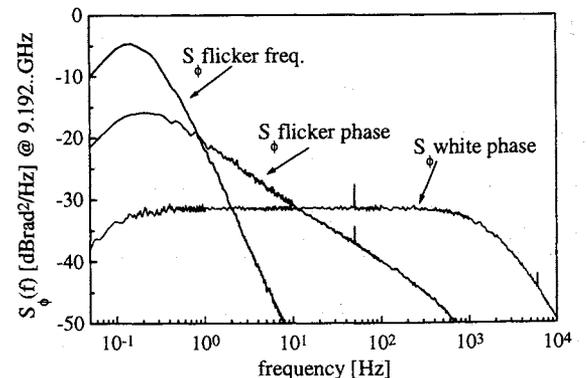


Fig. 4 Measured phase noise spectral densities of the degraded interrogation oscillator.

We use the noise generators to drive the phase modulation input of a HP3325B synthesiser. The

HP 3325B is used as an external generator for the fountain frequency multiplication chain (Fig. 3). The phase noise added to the HP 3325B is transferred to the interrogation oscillator spectrum. The added noise is high-pass filtered (0.1Hz) to avoid the degradation of the medium and long-term stability of the oscillator (Fig. 4).

The Allan standard deviation of the locked oscillator is calculated taking the frequency output of the synthesiser at the end of each cycle. As shown in Fig. 5 the fractional frequency stability of the free running oscillator, behaves as  $\tau^{-1}$  whereas the stability of the locked oscillator is proportional to  $\tau^{-1/2}$ . This clearly shows that the frequency stability of the locked oscillator is dominated by the aliasing noise for integration times longer than 10-20s.

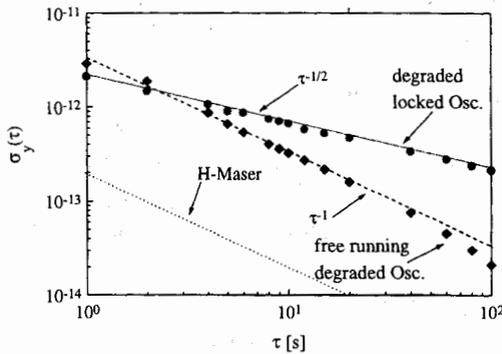


Fig. 5 Fractional frequency stability of the free running and locked oscillator

We measured the stability for two conditions:  $b\tau=\pi/2$  and  $b\tau=3\pi/2$ . Tables 1 and 2 report the calculated (Eq. (11)) and measured values for the flicker phase and flicker frequency noises.

Table 1,  $b\tau=\pi/2$

Type of noise	$\sigma_{y,meas.}(1s)$	$\sigma_{y,calc.}(1s)$
Frequency Flicker	$2.4 \cdot 10^{-12}$	$2.3 \cdot 10^{-12}$
Phase Flicker	$3.0 \cdot 10^{-12}$	$2.9 \cdot 10^{-12}$

Table 2,  $b\tau=3\pi/2$

Type of noise	$\sigma_{y,meas.}(1s)$	$\sigma_{y,calc.}(1s)$
Frequency Flicker	$2.8 \cdot 10^{-12}$	$2.4 \cdot 10^{-12}$
Phase Flicker	$4.8 \cdot 10^{-12}$	$4.6 \cdot 10^{-12}$

For these types of noise the frequency stability limitation mainly depends on the ratio between interrogation time and cycle time. In order to verify precisely the model we need a measurement which is more sensitive to the shape of  $g(\theta)$ . In the case of white phase noise the frequency noise spectrum behaves as  $f^{-2}$  and consequently the down-conversion is strongly dependent on the harmonic content of  $g(\theta)$ . The possibility to

change the frequency noise spectrum by low-pass filtering the noise source with different cut-off frequencies improves the sensitivity of the measurements. Fig. 6 reports the calculated and measured results, which agree to the limit of the measurements errors (20%). Measurements performed with the filtered white phase noise confirms the validity of the model even for different interrogation oscillator levels ( $b\tau=\pi/2$  and  $3\pi/2$ )

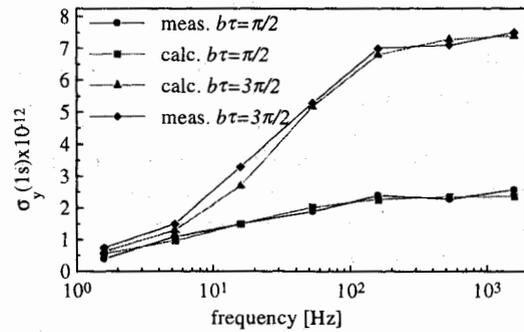


Fig. 6 Measured and calculated frequency stability versus the cut-off frequency of the white phase noise.

## 5. POSSIBLE MEANS FOR THE MINIMIZATION OF THE FREQUENCY STABILITY LIMIT

In a cesium fountain, in the PHARAO set-up [6] or in a trapped ions frequency standard, it will not be possible to use notch filters to reject the oscillator noise at Fourier frequencies  $m/T_c$  since  $T_c$  is equal to 1s or less.

The level of the annoying effect would obviously be reduced with an oscillator, such as a cryogenic sapphire oscillator [7] showing a much improved spectral purity. However, we will limit ourselves here to exemplify the beneficial effect of the increase of the interrogation duty cycle [1] or in the case of cold atoms frequency standards of the release of several balls during each cycle, assuming that the VCXO characterised by Eq. (14) is slaved to the atomic resonator. In the case of the Ramsey method of interrogation, the total interrogation time is  $T_i=T+2\tau$ . Fig. 7 shows the variation of  $\sigma_{y,lim}$  versus  $T_i/T_c$  for three sets of parameter values.

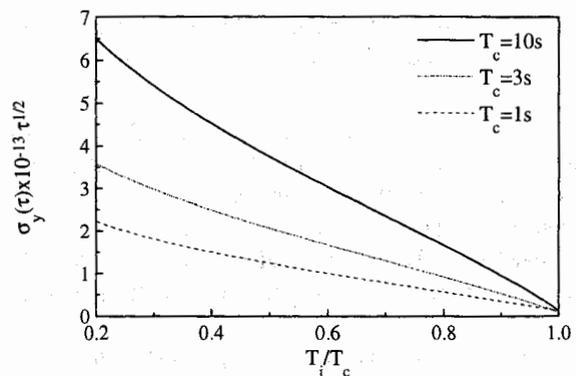


Fig. 7 Frequency stability versus the interrogation duty cycle for the Ramsey interrogation.

Several balls of cold cesium atoms can be launched during a cycle. Fig. 8 shows, as an example, the variation of  $\sigma_{y\text{lim}}$  versus the number of balls. It is assumed that the interrogation occurs in a TE<sub>013</sub> cavity and that the time interval between two successive ball releases is equal to  $T_i/6$ . Again, three sets of parameter values are considered.

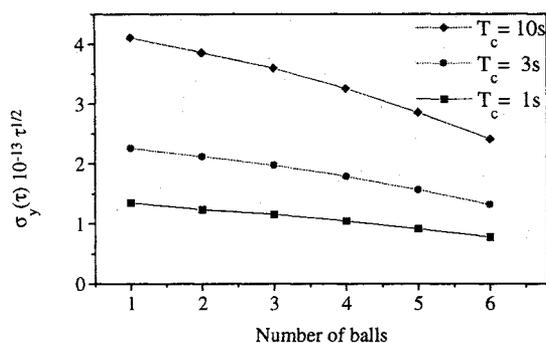


Fig. 8 Calculated frequency stability versus the number of launched atomic balls for the multi-Rabi case.

Thus, it is possible, in an atomic resonator based on the interrogation of atoms launched sequentially, to minimise the limiting effect considered. This may be accomplished by a proper design of the resonator leading to a duty cycle as large as possible and/or by launching a cloud of atoms several times during a cycle. In the case of the ion traps it has been proposed to use two parallel traps [8].

One may also note that in the two given examples,  $\sigma_{y\text{lim}}$  is the smallest for  $1/T_c=1\text{Hz}$ . This value is the closest to the Fourier frequency,  $f_Q$ , for which  $S_y^f(f)$  of the VCXO, shows a minimum. With the data of Eq. (14), we have  $f_Q=5\text{Hz}$ . This suggests that, whenever possible, the characteristics of the atomic resonator and of the VCXO should be matched. This is achieved when the condition  $f_Q \approx 1/T_c$  is fulfilled.

## CONCLUSIONS.

In this paper we have developed a simple model for the degradation of the frequency stability in a pulsed atomic frequency standard due to aliasing of the frequency noise of the interrogating oscillator. We have also compared results of the calculations based on this model with experimental values obtained by using the LPTF Cs atomic fountain with a voluntary degraded oscillator. The theory and the experiments agree in the limits of the measurement errors. The frequency stability limitation is about  $10^{-13}$  if we use a state of the art 5-10 MHz VCXO BVA. Better results can be achieved using cryogenic sapphire oscillators. The model shows that the limitation comes primarily from the flicker frequency noise of the oscillator and the typical white phase floor does not

affect the results. It seems that there are no obvious signal processing techniques which would be able to reduce the consequences of the detrimental aliasing.

## ACKNOWLEDGMENTS.

This work forms a part of the program of scientific developments of the Bureau National de Metrologie (BNM). The research programme has the financial and logistic support of the French Space Agency (CNES) via the Pharao project. We also acknowledge the financial participation of the Conseil Général de la Région Ile de France through the Sesame programme.

## REFERENCES.

1. Dick G. J., 1987, "Local oscillator induced instabilities in trapped ion frequency standards," in Proc. of *Precise Time and Time Interval*, Redondo Beach, 133-147.
2. De Marchi A., Bava E., Tavella P., 1993, "Accuracy analysis of the atom-microwave interaction in a TE<sub>017</sub> cylindrical Rabi Cavity for a high C-field cesium beam resonator," in Proc. of *7th European Frequency and Time Forum*, Neuchatel, 489-493.
3. Rice S.O., 1945, "Mathematical analysis of random noise". *Bell System Tech. J.*, **23**, 1944, 282-332 and **24**, 46-156.
4. Vanier J. and Audoin C., 1989, "The quantum physics of atomic frequency standards", Adam Hilger Ed.
5. Clairon A., Laurent Ph., Santarelli G., Ghezali S., Lea S.N., and Bahoura M., 1995, "A cesium fountain frequency standard: preliminary results," *IEEE Trans. Instrum. Meas.*, **44**, 128-131.
6. Clairon A., Laurent Ph., Nadir A., Drewsen M., Grison D., Lounis B. and Salomon C., 1992, "A simple and compact source of cold atoms for cesium fountains and microgravity clocks" in Proc. of 6th European Frequency and Time Forum, Noordwijk (The Netherlands), 27-33.
7. Luiten A. N., A. G. Mann, M. E. Costa and Blair D. G., 1995, "Power stabilized cryogenic sapphire oscillator," *IEEE Trans. Instrum. Meas.*, **44**, 132-135.
8. Dick G. J., Prestage J. D., Greenhall C. A., and Maleki L., 1990, "Local oscillator induced degradation of medium-term stability in passive atomic frequency standards," in Proc. 22nd Ann. Precise Time and Time Interval (PTTI), 487-508.

## STABILIZED LASER DIODE AT 729 NM FOR A $\text{Ca}^+$ OPTICAL FREQUENCY STANDARD

B. Fermigier, M. Houssin, F. Plumelle and M. Desaintfusien

Laboratoire de l'Horloge Atomique, Unité propre de Recherche du CNRS,  
associée à l'Université Paris-Sud  
Bât. 221 Université Paris-Sud 91405 ORSAY France

### INTRODUCTION

Experiments involving laser cooled ions confined in electromagnetic traps are very attractive in the domain of optical frequency standards. Expected performances specially the stability are extremely promising and these experiments remain a reliable alternative to the cold atom ones.  $\text{Ca}^+$  is a particularly interesting candidate because it has an optical transition with a very narrow natural linewidth which can be actually observed with very small perturbations in a Paul trap [1,2,3]. Moreover all the useful transitions involved in the experiment can be reached with laser diodes yielding a reliable and compact set up.

To take advantage of the line Q of the clock transition, three points must be satisfied simultaneously: stabilization of the center frequency, drastic linewidth reduction of the field spectrum and stable frequency sweeping. A stabilization on an ultra-stable cavity using the optical Pound-Drever method is suitable.

In the project in progress at LHA, the clock transition at 729 nm will be probed by a cooled commercial laser diode [4]. This laser diode is placed in a self-aligned external cavity yielding a preliminary linewidth reduction. We first present this mounting.

This external cavity must then be stabilized on a Fabry-Perot high finesse cavity used in reflection. We describe this set up. The error signal is sent to the diode current. So it is important to have a continuous wavelength tuning. However we faced some difficulties because some diodes were not efficiently AR coated. In order to understand this problem we study how the diode output facet reflectivity perturbs the external cavity accordability.

### SELF ALIGNED EXTERNAL CAVITY

External cavity is attractive to obtain better spectral purity and frequency stability than solitary laser diode. The dispersive element (grating) in the system permits the selection and tunability of the emission wavelength. Unfortunately classical extended cavity lasers (ECL) show a great sensitivity to misalignments. In this part, we

present the scheme and performances of a self aligned external cavity.

In self-aligned cavity lasers, sensitivity to vertical misalignments is reduced with a 1D cat's eye [5]. The laser diode is properly fed back along the vertical direction while the cat's eye has no effect along the horizontal one which corresponds to the dispersion plane of the grating.

Two different configurations have been tested (fig 1).

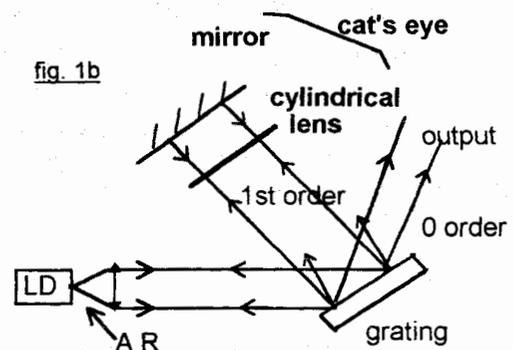
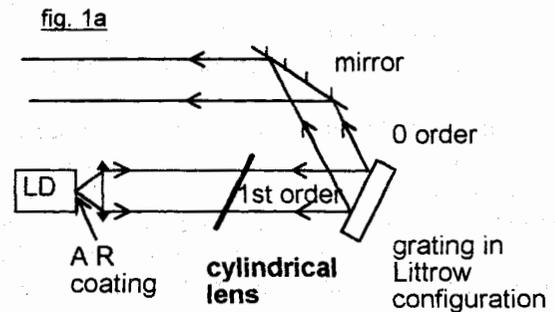


Fig.1: Two configurations of self aligned extended cavity laser (a: folded ; b: unfolded).

In both systems, the output facet of the diode is AR coated. The diode is mounted with the junction plane parallel to the grating lines for two reasons:

- less sensitivity to vertical misalignment
- better grating selectivity.

The unfolded cavity has been described and tested in ref 5. In the folded extended cavity, the grating is in Littrow configuration as in a classical external cavity. The cylindrical lens is inserted between the LD and the grating. The cat's eye is composed of a cylindrical lens and the grating separated by a focal distance. The cylindrical lens can be put :

- perpendicular to the optical axis
- parallel to the grating.

The cylindrical lens placed perpendicularly to the optical axis reduces the coupling efficiency in the junction because the grating (which is tilted against the optical axis) is not in the cylindrical lens focal plane.

When placed parallel to the grating, the cylindrical lens leads to very satisfying results in terms of mechanical stability. In particular the vertical misalignment tolerance varies from 1  $\mu\text{m}$  for a classical external cavity to several tens of  $\mu\text{m}$  (cf fig 2).

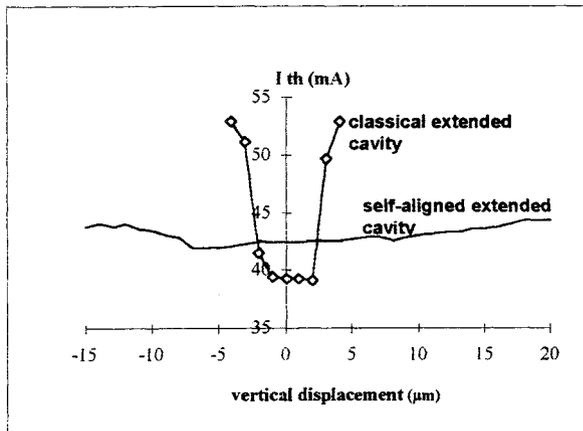


Fig.2 : threshold current versus collimating lens vertical misalignment.

Unfortunately the output beam quality is very poor: the important cylindrical lens inclination (about 30 degrees) introduces aberrations.

Because of these problems we have turned to the unfolded external cavity. In this configuration the output beam direction (0 order grating) is wavelength independant. Moreover the cat's eye translation instead of the grating rotation permits a continuous wavelength tuning over a 3 GHz range corresponding to one free spectral range of our extended cavity (cavity length: 5 cm). In conclusion, the unfolded extended cavity is a better configuration and satisfies our requirements.

## INFLUENCE OF THE DIODE FACET REFLECTIVITY ON THE CONTINUOUS WAVELENGTH TUNABILITY

We want to stabilize this external cavity on a high finesse cavity using the optical Pound-Drever method (cf fig. 3) [6].

Our frequency modulation is around 40 MHz.

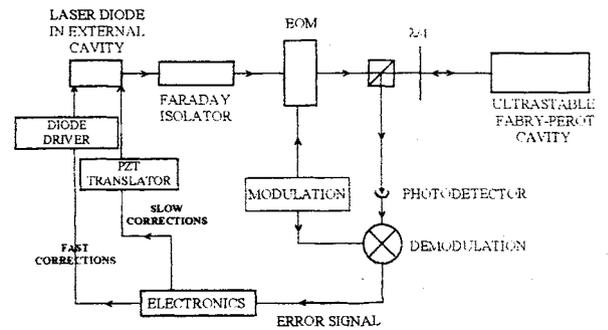


Fig. 3 : Pound-Drever servo-loop technique.

The corrections can be separated in two ways: slow one's are sent to the piezotranslator that translates the cat's eye; fast one's are sent to the laser current. Naturally these corrections must lead to continuous wavelength tuning. That is why we study the ECL wavelength tunability.

The reflection coefficient  $r_d$  of the diode facet has a significant influence on the continuous wavelength tunability. Understanding of this dependence of laser continuous tunability is important. Using the theoretical description from ref. 7, the oscillation condition for the ECL model can be written:

$$\omega - \omega_q^0 = (\alpha \ln( | r_{\text{eff}} | / r_d ) - \text{Arg}(r_{\text{eff}}) ) / \tau_0$$

where :

- $\omega_q^0$  is the  $q^{\text{th}}$  longitudinal mode of the chip cavity,
- $\tau_0$  is the chip cavity round trip transit time (without feedback),
- $\alpha$  is the linewidth enhancement factor,
- $r_{\text{eff}}$  is the effective reflectivity of the composite mirror formed by the diode output facet and the grating.

This equation gives a set of resonant frequencies. The lasing frequency is the one for which the threshold gain is minimum. Fig. 4 shows mode hopping when the diode current is tuned.

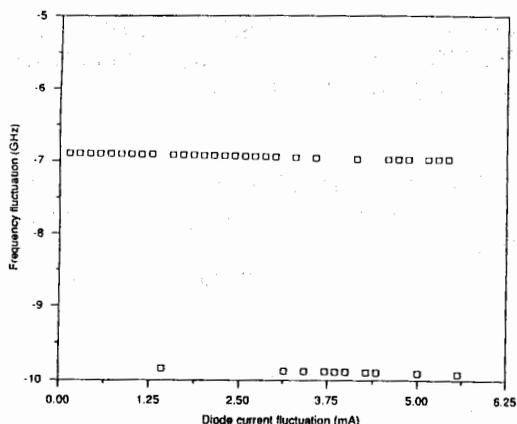


Fig. 4: example of frequency tuning for  $r_{\text{ext}}=0.3$  and  $r_d=0.1$

Equivalent cavities have been made with different laser diodes:

- SDL 5412 (800 nm),
- Sharp LT030 and LT031 (around 750 nm),
- Sharp LT030 with AR coating ( $r_d \sim 10^{-2}$ ).

All these LD's are single transverse mode and single longitudinal mode.

By cat's eye translation we obtain the same continuous wavelength tunability (3 GHz) with the different settings.

On the contrary, by varying the diode current, only the AR coated diode gives good results: 1 GHz for  $\Delta I=20$  mA. That gives a variation of:

$$(\Delta\nu/\Delta I)_{\text{ECL}}=0.05 \text{ GHz/mA}$$

In the non AR coated solitary diode, we have:

$$(\Delta\nu/\Delta I)_{\text{diode}}=2.8 \text{ GHz / mA}$$

In a ECL with an insufficient AR coating quality, this ratio is not constant and fluctuates between these two values.

## CONCLUSION

The self aligned cavity presents significant improvements, in particular much less sensitivity to mechanical perturbations. Among the two studied configurations, the unfolded cavity is without doubt the more suitable regarding beam quality and reliability.

Continuous tunability depends greatly on the quality of the AR coating of the diode output facet.

So we plan to send the fast corrections of the error signal to an external electro-optic modulator.

## References

1. F. Plumelle, M. Desaintfuscien, M. Houssin, *IEEE Trans. Instr. Meas* **42** n°2 (1993)

2. S. Urabe, K. Hayasaka, M. Watanabe, H. Imajo, R. Ohmukai, R. Hayashi, *Appl. Phys.* **B57**, pp 367-371 (1993)
3. M. Knoop, M. Vedel, F. Vedel, *J. Phys. II France* **4**, pp 1639-1650 (1994)
4. B. Fermigier, M. Desaintfuscien, F. Plumelle and M. Houssin, *Proc. of 9<sup>th</sup> EFTF* (Besançon, France, 1995), p.349
5. N. Dimarcq, E. Aucouturier, P. Petit and P. Cerez, *Proc. of 9<sup>th</sup> EFTF* (Besançon, France, 1995), p.111
6. D. Hils and J. L. Hall, *Frequency Standards and Metrology, Proc. of the 4th Symposium*, A. de Marchi, Ed. Belin: Springer-Verlag (1989)
7. P. Zorabedian, W. R. Trutna and L. S. Cutler, *IEEE J. Quantum Electron.*, vol **QE-23**, pp. 1855-1860 (1987).

**ABSTRACT**

Frequency modulated laser light, when synchronously detected after reflection from a Fabry-Perot cavity, give rise to well known discrimination curves, if the the source is tuned across the resonance. In lineshape calculations, the laser field is normally assumed to be monochromatic. Several different models describe the laser phase/frequency noise, namely phase diffusion model (PDM), FM or phase modulation from a band-limited gaussian noise (GNM), and a generalized phase jump model (PJM). Our work deals with the analysis of the dependence of the static frequency response on the parameters pertinent to each model. Moreover, in the PSD case, approximate expressions for the intensity spectrum are given.

**Authors:** Elio Bava and Fabio Massari

**Title:** Frequency response of a Fabry-Perot cavity excited by a frequency modulated stochastic laser field.

Dipartimento di Elettronica e Informazione  
Politecnico di Milano  
Piazza Leonardo da Vinci 32  
20133 Milano

# ON THE ESTIMATING OF POWER SPECTRAL DENSITY OF FRACTIONAL FREQUENCY FLUCTUATIONS FROM THE SAMPLED PHASE DEVIATIONS

Mihai P. Dinca

Physics Department, University of Bucharest, PO Box MG-11, Bucharest-Magurele, 76900, Romania

## INTRODUCTION

The correct characterization of the noise effect is extremely important to both manufacturers and users of high quality oscillators because their precision is ultimately limited by the noise.

Five different types are considered to perturb oscillators and are described in the Fourier frequency domain by the power spectral density (PSD) of the fractional instantaneous frequency fluctuations,  $S_y(f)$ . The standard power-law model assumes that, for these noise processes,  $S_y(f)$  is proportional to  $f^\alpha$  with  $\alpha$  an integer ranging from -2 to 2, [Barnes (1)]

However, in the frequency stability measurements made in the time domain, generally used for characterizing the oscillator behavior over a long time interval, the quantity  $y(t)$  is not directly observable. Upon the status of the experimental set-up [Allan and Daams (2)], samples of  $\bar{y}_i(t)$  or  $x(t)$  are obtained. Here, the average is performed over the time interval  $\tau$ ,  $x(t)$  is the phase-time deviation and the samples are taken separated by the interval  $\tau$ . The identification of the noise type is done by using the Allan variance  $\sigma_y^2(\tau)$  whose slope on a log-log plot is connected with the value of  $\alpha$  [Allan(3)]. In order to increase the determination accuracy, the multi-variance method was developed [Vernotte(4)] for the case when many noise types are present simultaneously.

The variance definition can be regarded as a digital filtering applied on the sequence  $x_n$  which represents the sampled phase-time deviations. Thus, the dependence  $\sigma_y^2(\tau)$  results to be the dependence of the power at the filter output upon the parameter  $\tau$  characterizing the implemented filter. Therefore, it appears more natural to identify the noise type by the classical spectral analysis techniques.

Unfortunately, for all noise processes considered, the PSD of the phase-time deviations,  $S_x(f)$ , does not exist in the strict sense. Only for the white

phase noise, when it is band limited, the random process which models  $x(t)$  is stationary.

For the power-law model, the PSD of the phase-time deviations result to be proportional to  $f^\beta$  where  $\beta$  is an integer ranging from 0 to -4. The estimation of the sample spectrum [Kasdin and Walter (5)], by FFT estimators (periodograms) was addressed by Waals et al (6), emphasizing the prime importance of the time window used in analysis.

On the other hand, difficulties in spectral analysis of the sequence  $\bar{y}_n$ , due to the lack of the last element, were reported [Vernotte (4), Vernotte et al (7)] an a method to overcome them, by considering the sequence  $x_n$  as periodic, was proposed.

In this paper we investigate the possibility to identify the noise type by FFT spectral estimators with different time windows, performed on the sequences  $x_n$  and  $\bar{y}_n$ . The noise sequences  $x_n$ , with  $\beta = 0, -1, -2, -3$  and  $-4$ , was generated by using a computer program [Dinca (8)] fulfilling the simulation criteria stated by Kasdin (5). The results obtained by Waals et al (6) are verified and extended for  $\beta = -3$ . No significant spectral distortion similar with that reported by Vernotte (4,7) was observed.

Thus, The noise type can be successfully identified by spectral analysis on  $x_n$  or  $\bar{y}_n$  if proper time windows are used.

## BASIC QUANTITIES AND RELATIONSHIPS

An ideal oscillator delivers an output signal

$$V(t) = A_0 \cos(2\pi v_0 t) \quad (1)$$

where  $A_0$  and  $v_0$  are the nominal amplitude and frequency, respectively. The "real" oscillator is perturbed by the phase noise  $\varphi(t)$  and provides the signal

$$V(t) = A_0 \cos[2\pi v_0 t + \varphi(t)] \quad (2)$$

exhibiting an instantaneous frequency fluctuation

$$\Delta v(t) = \frac{1}{2\pi} \frac{d\varphi(t)}{dt} \quad (3)$$

The most used quantities are the instantaneous fractional frequency fluctuations

$$y(t) = \frac{\Delta v(t)}{v_0} \quad (4)$$

and the phase-time deviations

$$x(t) = \frac{\varphi(t)}{2\pi v_0} \quad (5)$$

The averaged value of  $y(t)$ , over the interval  $\tau$ , can be expressed as

$$\bar{y}_\tau(t) = \frac{1}{\tau} \int_t^{t+\tau} y(t') dt' = \frac{x(t+\tau) - x(t)}{\tau} \quad (6)$$

One can write the standard power law model

$$S_y(f) = \sum_{\alpha=-2}^2 h_\alpha f^\alpha \quad (7)$$

or, in terms of PSD of  $x(t)$

$$S_x(f) = \sum_{\beta=-4}^0 g_\beta f^\beta \quad (8)$$

## ESTIMATION OF PSD USING THE FFT ALGORITHM

Although the PSD does not exist for non stationary processes, the sample spectrum expressed as

$$\hat{S}_{x,T} = \frac{1}{T} \left| \int_{t_0}^{t_0+T} x(t) \exp[-2\pi jf(t-t_0)] dt \right|^2 \quad (9)$$

can be convergent and in the following will be used for PSD estimation.

Even for stationary process, the PSD estimated by this method from the sequence which samples the continuous function is affected by errors. Firstly, only sampled values of the function  $x(t)$ , namely the sequence  $x_n$ , are available. Thus, the spectral estimation has to use the techniques developed

for digital signal processing. The sampling rate,  $f_s = 1/\tau$ , is very important because if the PSD of  $x(t)$  has significant power distributed over frequencies greater than  $f_s/2$  aliasing occurs and can change the spectrum shape. For  $\beta = -1, -2, -3$  and  $-4$  the aliasing can be neglected, but for the white phase noise this effect is certainly important, providing that the upper cut frequency  $f_h$  is greater than  $f_s/2$ . In the case when  $f_h \gg f_s/2$  aliasing produces a level of the PSD of  $x_n$  much greater than  $S_x(f)$  but the shape are practically unchanged.

Secondly, the available sequence  $x_n$  has finite length. The expected value of the PSD of  $x_n$  is then the convolution of the true  $x_n$  PSD with the Fourier image of the Bartlett window (Fig. 1), providing that a rectangular time window was used [Oppenheim and Schaffer (9)].

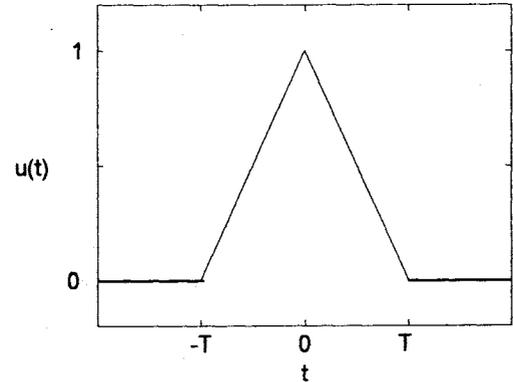


Fig. 1. Bartlett window

## NOISE GENERATOR

The sequence  $x_n$  is produced by the linear convolution

$$x_n = \sum_{i=0}^{n-1} h_i w_{n-i} \quad (10)$$

where  $w_n$  is an independent identical distributed sequence. The sequence  $h_n$  is expressed as

$$\begin{aligned} h_0 &= 1 \\ h_n &= \frac{n-1-\beta/2}{n} h_{n-1} \end{aligned} \quad (11)$$

providing that the digital filter has the Z transform

$$H(z) = (1 - z^{-1})^{\beta/2} \quad (12)$$

which is a generalization of the Z transform associate with the Brownian process ( $\beta=-2$ ). The PSD of the generated sequence  $x_n$  is obtained proportional to  $\sin(\pi f\tau)^\beta$  which for  $f \ll f_s/2$  becomes proportional to  $f^\beta$ .

The linear convolution (9) is computed by a circular one completing the sequence  $h_n$  and  $w_n$  with zero elements [Oppenheim and Schaffer (9)]. Thus, the linear convolution can be simply performed by direct FFT, multiplication and inverse FFT. In order to allow longer sequences handling, the feature that the sequences are real was exploited (Fig. 2).

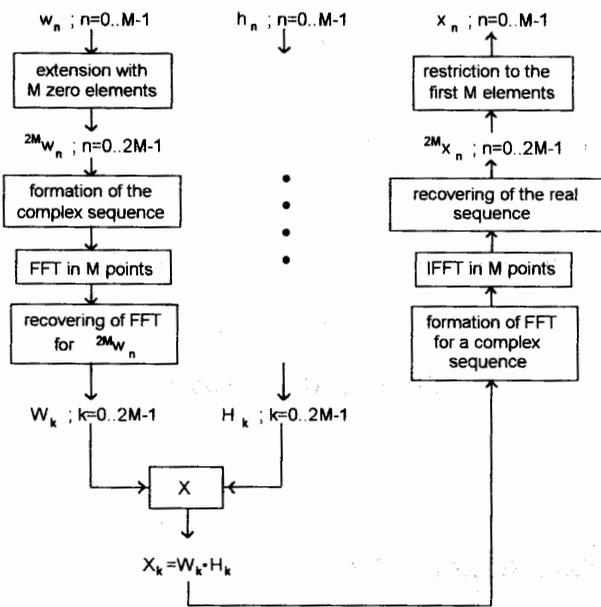


Fig. 2. Algorithm for noise generation.

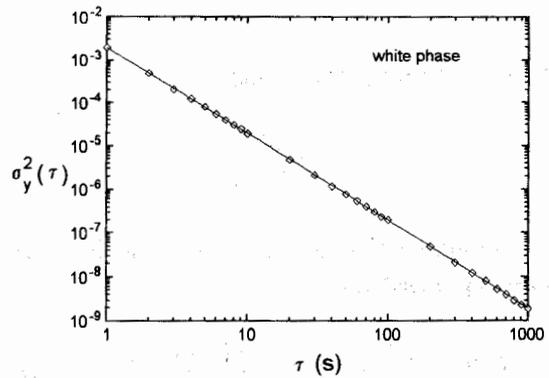
The sequences generated by the program were tested by computing the Allan variance and the results agreed with the well known behavior of the variance.

**RESULTS**

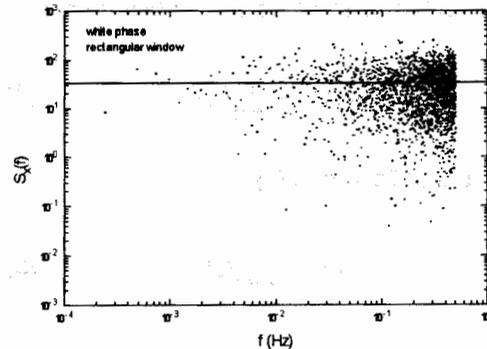
Five sequences  $x_n$ , each of them consisting of 4096 elements, with  $\beta=0,-1,-2,-3$  and  $-4$ , were generated. The corresponding sequences  $\bar{y}_n$  were calculated using (6). Only 4095 elements were obtained and the sequences were completed with

a zero value. For each  $x_n$  sequence, the Allan variance was computed. Also, the PSD of the sequence was estimated by a FFT algorithm with rectangular and Hanning time window

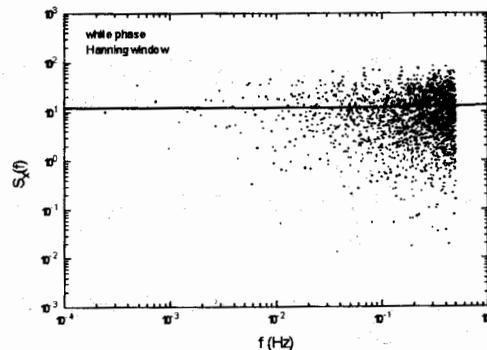
$$\begin{aligned} \text{rect} u_n &= 1 \quad \text{for } 0 \leq n \leq N-1 \\ H u_n &= \left( 1 - \cos \frac{2\pi(n+0.5)}{N} \right) \quad \text{for } 0 \leq n \leq N-1 \end{aligned} \quad (13)$$



a)

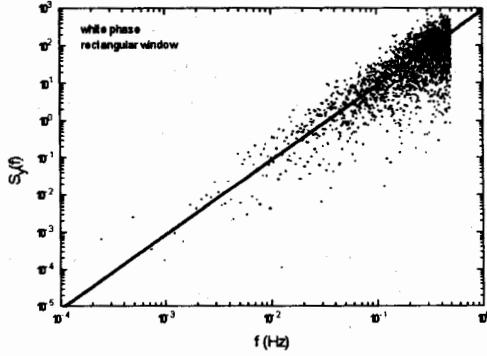


b)

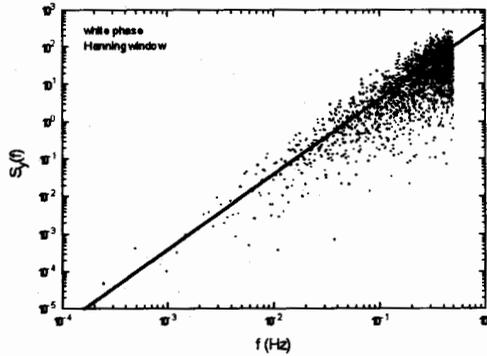


c)

Fig. 3 Simulated white phase noise: a) Allan variance, b),c)  $S_x(f)$  with rectangular and Hanning window,



d)



e)

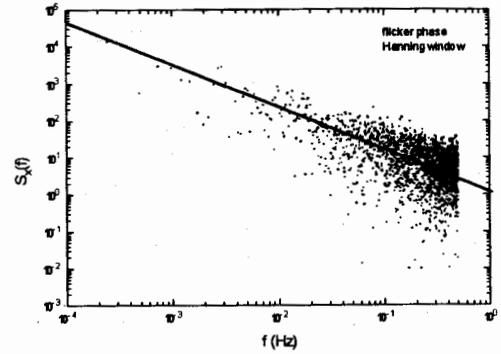
Fig. 3. Simulated white phase noise: d), e)  $S_y(f)$  with rectangular and Hanning window.

In the same way the PSD for the  $\bar{y}_n$  sequences were calculated.

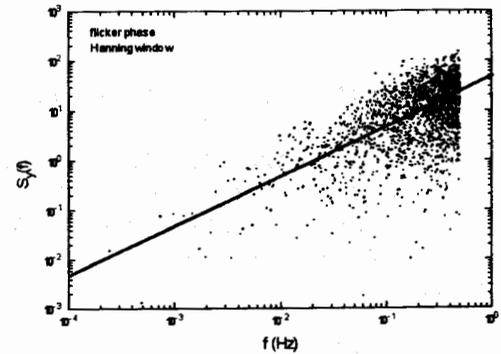
For the white phase noise the results are presented in Fig. 3 a, b, c, de and e. One can observe that for  $S_x(f)$  both windows provide the correct slope of the PD, while for  $S_y(f)$  little alterations are observed at low frequencies in the case of rectangular window.

For flicker phase noise both windows allow correct identification of noise type for both  $S_x(f)$  and  $S_y(f)$ . In Fig 4 are represented the results obtained for Hanning window. Similar results regarding identification possibility were obtained for the white frequency noise.

However, for flicker frequency noise the window type is essential for identification by  $S_x(f)$  computation as can be seen in Fig 5. The same conclusion can be outlined for the random walk frequency noise (Fig. 6). On the other side, for these last noise types both windows provides the correct slope for  $S_y(f)$ .



a)



b)

Fig 4 Simulated flicker phase noise: a)  $S_x(f)$  with Hanning window, b)  $S_y(f)$  Hanning window.

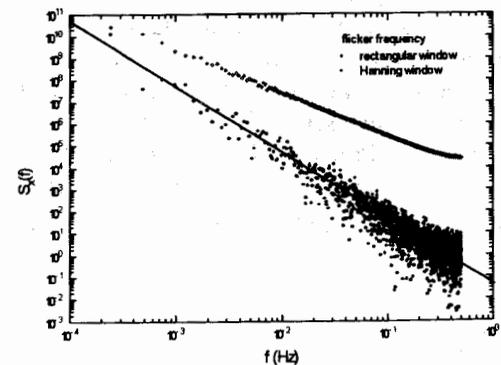


Fig. 5.  $S_x(f)$  for simulated flicker frequency noise.

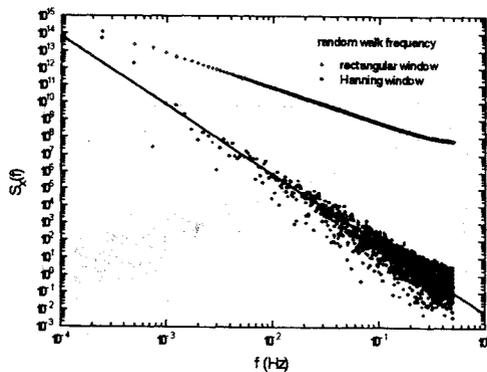


Fig. 6.  $S_x(f)$  for simulated random walk frequency noise.

## CONCLUSION

Rectangular and Hanning windows were used in spectral analysis, using the FFT estimator, applied on the sampled phase-time deviations and sampled averaged fractional frequency fluctuations. All five noise types were investigated, the sequences being produced by a computer noise generator.

Despite a little difference in the spectrum level, both windows gave the correct slope for  $S_y(f)$ . Slight distortion was observed at low frequencies for the white phase noise when a rectangular window is used.

For  $S_x(f)$  the window type does not affect the spectrum slope for white phase, flicker phase and white frequency noises. In the case of flicker frequency and random walk frequency noises, the rectangular window produced an incorrect spectrum slope, while the Hanning window allows the correct identification of the noise type. It must be emphasized that the slight distortion observed near  $f_s/2$  is caused by the noise generator which provides a PSD proportional to  $\sin(\pi ft)^\beta$ .

## References

1. Barnes J. A. et al, 1971, "Characterization of frequency stability", *IEEE Trans. Instrum. Meas.*, IM 20, 2, 105-120.
2. Allan D. W. and Daams H. 1975, "Picosecond time difference measurement system", Proc. 29th Ann. Frequency Control Symp., 404-411.

3. Allan D. W., 1966, "Statistics of atomic frequency standards", *Proc. IEEE*, 54, 2, 221-230.

4. Vernotte F, 1991, "Stabilité temporelle des oscillateurs: nouvelles variances, leurs propriétés, leurs applications", Thesis, Université de Franche-Comté.

5. Kasdin N. J. and Walter T, "Discrete simulation of power law noise", 1992, Proc. 1992 IEEE Frequency Control Symposium, 274-283.

6. Waals F. L. et al, 1989, "Biases and variances of several FFT estimators as a function of noise type and number of samples", Proc. 43rd Ann. Symp. on Frequency Control, 336-341.

7. Vernotte F. et al, 1991, "Practical calculation methods of spectral density of instantaneous normalized frequency deviation from instantaneous time error samples", Proc 5th European Frequency and Time Forum, 449-455.

8. Dinca M. P., 1995, "Contributions on the study of frequency stability of atomic standards", Thesis, University of Bucharest (in Romanian).

9. Oppenheim A.V. and Schaffer R. V., "Digital Signal Processing", Prentice-Hall, Inc., Englewood Cliffs, USA.

## CLOCK PREDICTION AND CROSS-SIGMA

A Lepek

Tech Projects, PO Box 9769, Jerusalem 91091, Israel

### ABSTRACT

The purpose of this paper is to describe relations between the prediction of clock time series  $X_m$ , the prediction error standard deviation  $\sigma$  and,  $\#\sigma$  which is used to estimate  $\sigma$  in the presence of noise.

We confine the discussion to linear predictors. These predictors are sufficient to optimally predict clocks whose stability is described by an extension of the common oscillator model, including drifts.

We compare the 3-corner-hat estimate of  $\sigma$  of a single clock with  $\#\sigma$  and show that in the presence of noise  $\#\sigma$  is a better estimate.

If all measurements are not made simultaneously,  $\#\sigma$  becomes an estimate of an auto-correlation function rather than of  $\sigma$ . As a result the power spectrum of the prediction errors of a single clock may be estimated.

### DEFINITIONS

We start with a time series of time differences between two clocks,  $X_m$ , known at sampling intervals  $\tau$ . A linear predictor  $P$  is fully defined by its coefficients

$$P = [a_0, \dots, a_{N-1}] \quad [1]$$

The predicted value of  $X_m$  is

$$P_m = \sum_{i=0}^{N-1} a_i X_{m-N+i} \quad [2]$$

After measuring  $X_m$ , the predicted  $m$ th error is

$$D_m = X_m - P_m = \sum_{i=0}^N -a_i X_{m-N+i}, \quad a_N = -1 \quad [3]$$

Similarly to [1] we can symbolize the error function as

$$D = \{-a_0, \dots, -a_{N-1}, 1\}, \quad a_N = -1 \quad [4]$$

Assuming that the average  $D_m$  is zero (non trivial), the standard deviation of  $D_m$  is

$$\sigma = \sqrt{\langle D_m^2 \rangle} \quad [5]$$

Where  $\langle \rangle$  denote average over  $m$ .

In optimum prediction,  $\langle D_m \rangle = 0$ ,  $D_m$  is random and its  $\sigma$  is unbiased and minimum.

Normally,  $D_m$  results from comparison of two clocks. If such comparison is made using simultaneously two measurement systems, 1 and 2, we obtain two time series  $D1_m$  and  $D2_m$ . See figure 1 for possible measurement system configurations.

We define cross-sigma as

$$\#\sigma = \text{sign} \sqrt{\langle |D1_m \cdot D2_m| \rangle} \quad [6]$$

where "sign" is the sign of the average.

This definition follows the approach described in Lepek and Walls (1).

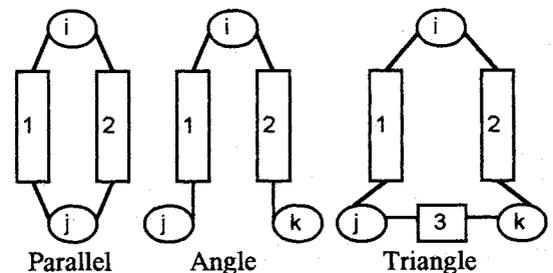


Figure 1: measurement configurations.  $i, j$  and  $k$  are clocks, 1, 2 and 3 are measurement systems from which the time series  $X_{ijm}$  etc. are obtained.  $D1_m$  is obtained from  $X_{ijm}$ ,  $D2_m$  from  $X_{ikm}$  in the angle configuration or from  $X_{ijm}$  in the parallel configuration.

### OPTIMUM PREDICTION

The straightforward approach to define the "a" coefficients, is to use the least squares method to minimize  $\sigma$  and then to test with a time series for which  $\sigma$  is known. With finite number of coefficients it is possible to have optimum predictors for some types of clock instability and only near optimum for other. Once we have coefficients for a predictor we can manipulate them using simple rules to obtain new predictors. In the following we impose constraints on the coefficients which impart desired properties to the predictors.

We want the errors to average to zero, then

$$\sum_{i=0}^N a_i = 0, \quad a_N = -1 \quad [7]$$

For  $N=1$  we obtain  $P=[1]$  and  $D=\{-1,1\}$ . This predictor is optimal for white frequency modulation instability.

Generalizing equation [7] to allow for prediction about a straight line or frequency drifts, we have,

$$\sum_{i=0}^N (i+1)^Z \cdot a_i = 0, \quad a_N = -1 \quad [8]$$

For  $N=2$ ,  $Z=0$  (equation [7]) and  $Z=1$  (straight line), we have  $D=\{1,-2,1\}$  and  $P=[-1,2]$ . It is clear that  $\sigma/\tau = \sqrt{2} \cdot Adev$ , where  $Adev$  is the Allan deviation. This predictor is optimum for random walk frequency modulation instability. In this context, for a given  $\tau$ , the modified Allan deviation,  $Mdev$ , can be interpreted as originating from a predictor which uses and predicts the average  $X_m$  rather than the next point,  $\tau$  being the time interval for the averaging. Adopting this approach, the modified  $\sigma$  of the modified  $D=\{1,-2,1\}$ , is  $Mod \sigma = \tau \sqrt{2} \cdot Mdev = \sqrt{6} \cdot Tdev$ . For the definitions of  $Tdev$ ,  $Mdev$  and  $Adev$  see Allan et al (6).

For white phase modulation we expect that

$$\sum_{i=0}^{N-1} a_i^2 = \text{minimum} \quad [9]$$

will result in minimum  $\sigma$ . Equations [7] and [9] result in  $D=\{-1/N, \dots, -1/N, 1\}$  and  $P=[1/N, \dots, 1/N]$ , that is, the predictor is the average of  $N$  previous points. It converges to optimum with  $N \rightarrow \infty$  for white phase instability. With  $N=8$ ,  $\sigma$  is 6% above optimum, which is satisfactory for many practical applications. The dependence is  $\sigma \propto (\sqrt{N+N^2})/N$ . A trivial case is when it is known a priori that  $\langle X_m \rangle = 0$ . Then the optimum predictor is  $P=[0]$ ,  $D=\{1\}$ ,  $\sigma = \text{RMS}$ .

### Predictor Manipulation

Starting with a time series which simulates white phase instability (by a series of random numbers) it is possible by half integrations or differentiations to simulate any power law instability within the common oscillator model. If a predictor has a specific  $\sigma$  value for one time series, given a second time series obtained from the first (e.g. by integration), we can transform the predictor to have the same value of  $\sigma$  by an opposite operation on its "a" coefficients. Below we describe the algorithms.

**Coefficient differentiation.** Denoting the coefficients of the present error function  $B$  by  $b_i$  and of the new one,  $C$  by  $c_i$  we have

$$c_0 = -b_0$$

$$c_i = b_{i-1} - b_i \quad [10]$$

$c_N = b_N = 1$   
 $C$  has one coefficient more than  $B$   
 If the original  $B$  obeys equation [7] then the transformed  $C$  obeys both equations [7] and [8].

### Coefficient Integration.

$$c_{N-1} = b_N = 1$$

$$c_i = b_i + b_{i+1}$$

$$c_0 = b_1 + c_1 \quad [11]$$

$C$  has one coefficient less than  $B$ .  
 If the original  $B$  obeys equation [8] then the transformed  $C$  obeys equation [7].

### Coefficient weighted averaging of two predictors.

Usually the two predictors have a different number of coefficients. The shorter is padded with zeros at the lower index side before the averaging. In many cases this operation is equivalent to partial differentiation or integration.

Table 1 summarizes some useful predictors for an extension of the common oscillator model (that is, more than 5 instability types). "I+n" indicates the number of integrations ("- " for differentiations). The  $\sigma$  values are normalized with respect to optimum (that is, the  $\sigma$  value of  $D=\{1\}$  for white phase modulation). The predictors derived from  $P=[1/N, \dots, 1/N]$  are given with 8 coefficients. It is easy to see which predictors comply with equations [7] or [8].

$\sigma$  depends on  $\tau$ , therefore we give in table 1 the slopes of  $\text{Log}(\sigma)$  vs.  $\text{Log}(\tau)$  rounded to the nearest 0.5. The slopes of  $Tdev$  and  $Mod \{1\}$  (that is, for which the  $D_m$  are averaged over  $\tau$ ) are also given for comparison. The change in slopes in  $Tdev$  and  $Mod \{1\}$  at the low "I" region, is due to the fact that the errors are averaged over the time interval  $\tau$  so that the number of samples used in the averaging is proportional to  $\tau$  and is not constant. For each  $\tau$ ,  $Mod \{1\}$  and generally  $Mod D$ , is interpreted as before, that is, as originating from a predictor that predicts the average next point.

The slopes can be used to choose the best  $\tau$  for prediction. If the slope  $< 0.5$  then it is advantageous to use the largest  $\tau$  as compared to a smaller  $\tau$  but with more prediction steps to predict the same point in time. If the slope  $> 0.5$  the opposite is true.

It is easy to compute many of the  $\sigma$  values of the various error functions. To do so we note that the values can be obtained from each other by shifting table 1. Also, the  $\sigma$  value for I+0 (white phase instability) is proportional to the root sum squares of the coefficients of the error function. Other values were obtained from simulations. Values above 10 are uncertain.

TABLE 1 - Some Predictors, their normalized-to-optimum  $\sigma$  values for different number of integrations (1st lines) and Log( $\sigma$ ) vs. Log( $\tau$ ) slopes (2nd lines)

Integrations:	I-1	I-0.5	I 0 White PM	I+0.5 Flicker PM	I+1 White FM	I+1.5 Flicker FM	I+2 Random walkFM	I+2.5	I+3	I+3.5	I+4
{1,2,3,4,5,6,7,8,9}/9 slope:	1.05 0	1.21 0	1.87 0	3.39 0	159 0						
{-1,0,1,2,3,4,5,6,16}/16 slope:	1.38 0	1.03 0	1.17 0	1.80 0	71 0						
{-1,-1,-1,-1,-1,-1,-1,-1,8}/8 slope:	1.51 0	1.19 0	1.06 0	1.19 0	1.80 0.5	3.19 0.5	216 1				
{-1,-1,-1,-1,-1,-1,-1,-9,14}/14 slope:	2.00 0	1.54 0	1.20 0	1.08 0	4.73 0	14 0					
{1,0,0,0,0,0,8,7}/7 slope:	2.62 0	2.00 0	1.52 0	1.25 0	1.07 0.5	1.19 1	1.70 1.5	2.96 1.5	192 2		
{1,0,0,0,0,7,-22,14}/14 slope:	3.48 0	2.64 0	1.93 0	1.46 0	1.16 0.5	1.06 1	1.22 1.5	1.92 1.5	120 2		
{-1,1,0,0,0,7,-13,6}/6 slope:	4.85 0	3.68 0	2.67 0	2.00 0	1.54 0.5	1.27 1	1.08 1.5	1.19 2	1.60 2.5	2.72 2.5	168 3
{1} RMS slope: slope of Mod{1}:	1.41 0 -1	1.12 0 -0.5	1.00 0 -0.5	1.16 0 0	32 0 0						
{-0.5,1} slope:	1.87 0	1.41 0	1.12 0	1.00 0	24 0						
{-1,1} slope:	2.45 0	1.87 0	1.41 0	1.15 0	1.00 0.5	1.16 1	32 1				
{0.5,-1.5,1} slope:	3.38 0	2.57 0	1.87 0	1.41 0	1.12 0.5	1.00 1	24 1				
{1,-2,1} slope: slope of Tdev:	4.47 0 -1	3.38 0 -0.5	2.45 0 -0.5	1.87 0 0	1.41 0.5 0.5	1.15 1 1	1.00 1.5 1.5	1.16 2 2	32 2 2		
{-0.5,2,-2.5,1} slope:	6.31 0	4.47 0	3.38 0	2.45 0	1.87 0.5	1.41 1	1.12 1.5	1.00 2	24 2		
{-1,3,-3,1} slope:	8.37 0	6.31 0	4.47 0	3.38 0	2.45 0.5	1.87 1	1.41 1.5	1.15 2	1.00 2.5	1.16 2.5	32 3
{0.5,-2.5,4,-3.5,1} slope:	12 0	8.37 0	6.31 0	4.47 0	3.38 0.5	2.45 1	1.87 1.5	1.41 2	1.12 2.5	1.00 2.5	24 3
{1,-4,6,-4,1} slope:	16 0	12 0	8.37 0	6.31 0	4.47 0.5	3.38 1	2.45 1.5	1.87 2	1.41 2.5	1.15 2.5	1.00 3.5

If one wishes to characterize clocks by the performance of a set of predictors, it is suggested to use the average of a forward (normal predictor) and a backward predictor (having coefficients in reverse order). The error function of such a characterizer has an odd number of coefficients and the middle one is 1. For example, {1,-2,1} becomes {0.5,-1,1,-1,0.5}.

For each type of oscillator instability we would like to have an optimum predictor. Having such a set of predictors, then for a given time series and chosen  $\tau$ , we devise an algorithm to choose the best predictor for the next predicted point. The decision is based on the minimum  $\sigma$  for this set of predictors, computed from  $D_m$  points immediately preceding the predicted point. The type of instability is not a concern because the algorithm chooses the best predictor for the given  $\tau$  and

may choose another one when in due time the statistics changes. This algorithm was used in Allan et al (2,3) by the author to generate time scales from sub-sets of clocks contributing to TAI ( $\tau=10$  days), as a means to analyze TAI and the clocks' performance.

Currently, such an algorithm is used by the author to generate UTC(INPL) with  $\tau=1$  day.

Table 2 lists  $\sigma$  values from one iteration of this algorithm for two commercial Cs clocks. The deviations from the expected values for white FM indicate that there is contamination of flicker FM. Predictors 1 and 12 fail to predict reasonably because they do not comply with equation [8] ( $Z=1$ ) and cannot predict points about a straight line. The  $\sigma$  with predictor 10 (average of predictors 4 and 11) for clock "A" is equivalent to  $y=2.8 \cdot 10^{-14}$  for 24 hours.

TABLE 2 - Prediction  $\sigma$  obtained as a part of an algorithm choosing in real time the best predictors for Cs clocks with  $\tau=1$  d.

	error function D	clock A $\sigma$ / ns	clock B $\sigma$ / ns
1	{-1,1}	89.	123.
2	{1,0,0,0,0,0,0,0,-10,9}/9	2.45	3.04
3	{1,0,0,0,0,0,0,0,-9,8}/8	2.49	3.11
4	{1,0,0,0,0,0,0,-8,7}/7	2.49	3.13
5	{1,0,0,0,0,-7,6}/6	2.52	3.13
6	{1,0,0,0,0,-6,5}/5	2.67	3.19
7	{1,0,0,0,-5,4}/4	2.59	3.29
8	{1,0,0,-4,3}/3	2.51	3.57
9	{1,0,-3,2}/2	2.72	3.59
10	{3,0,0,0,0,0,7,-38,28}/28	2.42	3.07
11	{1,0,0,0,0,0,7,-22,14}/14	2.50	3.17
12	{0.5,-1.5,1}	44.	67.
13	{1,-2,1}	2.90	3.73
14	{-0.5,2,-2.5,1}	3.71	4.68
15	{-1,3,-3,1}	4.84	6.06

Prediction as part of time scale generation or analysis is used for many years. An approach aimed at the prediction of UTC based on predictors constructed of  $n$ th order differences and especially the 3rd order, is described in Barnes and Allan (4). A prediction based on filtered 2nd difference is used currently to generate AT2 NIST time scale. This scale is described in Weiss and Weissert (5). The  $\tau$  for which Allan deviation is minimum, is used to define the filter parameters.

### $\sigma$ , $\#\sigma$ AND NOISY MEASUREMENTS

We have defined  $\sigma$  as the standard deviation of prediction error  $D_m$ . In real systems the added noise may be high and  $\sigma$  no longer characterizes the time prediction of the clocks. If it is possible to compare the output of two measuring systems, then it may be possible to filter out the uncorrelated excess noise by cross correlating the two series  $D_{1m}$  and  $D_{2m}$  and so restore  $\sigma$ .  $\#\sigma$  was defined in equation [6] by such cross correlation. Three possible measurement configurations are given in figure 1.

Cross correlations were investigated also by Grosslambert et al (7).

In the following we classify the results concerning  $\#\sigma$  according to the measurement configurations.

#### Parallel Configuration

we can disassemble the  $D_m$  time series into  $q_m+n_m$ , where  $q$  is the error due to the clocks and  $n$  is the error due to the added noise. If  $n_1$  and  $n_2$  are uncorrelated and have zero mean, then for sample size  $N \rightarrow \infty$ ,

$$\#\sigma_{1,2}^2 = \langle (q_m + n_{1m})(q_m + n_{2m}) \rangle = \langle q_m^2 \rangle = \sigma_q^2 \quad [12]$$

$$\sigma_1^2 = \langle (q_m + n_{1m})^2 \rangle = \sigma_q^2 + \langle n_{1m}^2 \rangle \quad [13]$$

where  $\sigma_q$  is the standard deviation of the clock prediction error. This means that  $\sigma_1$  is limited by the added noise while  $\#\sigma$  is not.

To obtain  $\sigma_q$  with equation [12] we note that the non correlated terms in  $\#\sigma$  converge to zero with the sample size  $N$ , as  $N^{-0.25}$ . Therefore large samples may be needed to reveal clocks buried deeply in measurement noise. However, such large samples are readily available for small  $\tau$ .

For small samples to reach the low level of  $\sigma_q$  we require that the uncertainty in  $\#\sigma$ ,  $U(\#\sigma)$  be smaller than the uncertainty in  $\langle n^2 \rangle$ ,  $U(\langle n^2 \rangle)$  plus the noise level  $\langle n^2 \rangle$ . This condition may be fulfilled in many cases giving  $\#\sigma$  an advantage. For example, if  $n_1$  and  $n_2$  are random numbers in the range  $(-0.5, 0.5)$  then with  $N=1$  we have  $U(\#\sigma)=0.25$ ,  $\langle n^2 \rangle=0.29$ ,  $U(\langle n^2 \rangle)=0.14$  and  $2U(\#\sigma) < 2U(\langle n^2 \rangle) + \langle n^2 \rangle$ . This small advantage improves with  $N$ . For  $N=81$  it is better by 70%.

If the measurements in systems 1 and 2 are not simultaneous, such as when using one time interval counter which is switched between the systems, then equation [12] becomes

$$\lim_{N \rightarrow \infty} \#\sigma^2 = \langle q(t) \cdot q(t+T) \rangle = R(T) \quad [14]$$

where  $R(T)$  is the auto correlation function of the error  $q$  and  $T$  is the delay between the measurement systems. If  $P$  is an optimum predictor for the clocks then  $q$  has a white noise statistics and  $R(T>0)=0$ . Controlling the delay  $T$  between the measurements, the complete function  $R(T)$  is obtained from which the power spectrum of the error  $q$  can be computed.

Figure 2 shows  $R(T)$  obtained with  $D=\{1,-2,1\}$  and  $\tau=1$  for several instabilities of table 1 indicated by  $I+n$  as before. We note that for random walk FM ( $I+2$ ), as much as the resolution allows,  $R(T>0) \ll R(T=0)$  as expected. This is not the case with other instabilities. For  $I+3$ , which is beyond a reasonable prediction for this predictor,  $R$  is a straight line.

#### Angle Configuration

In this case all clocks are compared to one clock  $i$ .  $D_{jkm}=D_{ikm}-D_{ijm}=D_{2m}-D_{1m}$  by definition since  $X_{jkm}$  it is not measured. Using this relation we have

$$\begin{aligned} \# \sigma_{12}^2 &= \langle D_{1m} \cdot D_{2m} \rangle = \\ & \langle D_{1m}^2 + D_{2m}^2 - D_{jkm}^2 \rangle / 2 = (\sigma_{ij}^2 + \sigma_{ij}^2 - \sigma_{jk}^2) / 2 \end{aligned} \quad [15]$$

In general, the 3-corner-hat expression does not equal to  $\sigma_i$ , the single clock  $\sigma$ , because  $\sigma_{jk}$  is not independent by definition. However if the added noise is negligible then  $\# \sigma_{12} = \sigma_i$  because without noise  $q_{jk} = q_{ik} - q_{ij}$  always.

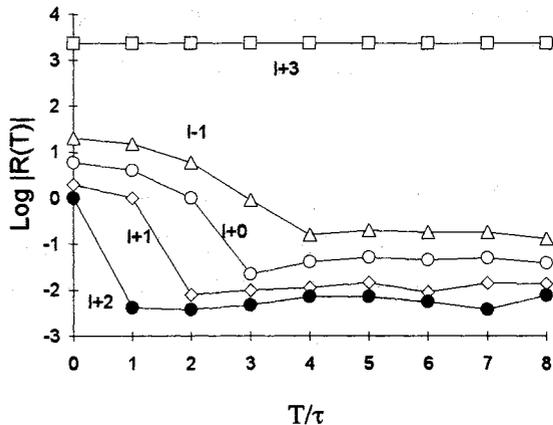


Figure 2:  $\text{Log } |R(T)|$  as a function of  $T$  for  $D=\{1,-2,1\}$  and  $\tau=1$  for several instabilities  $I+n$  defined in table 1.

Disassembling  $D=q+n$  as before and noting that  $q_{ij}=q_i-q_j$  ( $q_i$  is the prediction error of clock  $i$  against an ideal clock), then for  $N \rightarrow \infty$

$$\begin{aligned} \# \sigma_{1,2}^2 &= \langle D_{ijm} \cdot D_{ikm} \rangle = \\ & \langle (q_{im} - q_{jm} + n_{ijm})(q_{im} - q_{km} + n_{ikm}) \rangle = \\ & \langle q_{im}^2 \rangle = \sigma_i^2 \end{aligned} \quad [16]$$

That is,  $\# \sigma$  estimates clock  $i$  standard deviation from prediction, as before, given a large enough sample size.

As before, if channels 1 and 2 are delayed by  $T$ , we have for  $N \rightarrow \infty$ ,  $\# \sigma^2 = \langle q_i(t) \cdot q_i(t+T) \rangle = R(T)$  and the power spectrum of  $q_i$  can be computed by scanning  $T$ .

### Triangle Configuration

Equation [16] is, of course, true as with the angle configuration. Here  $D_{jk}$  is obtained from measurements. If the predictor is optimum then the prediction errors are of white noise. The 3-corner-hat is meaningful and (for  $N \rightarrow \infty$ ) it equals to  $q_i^2 + n_i^2$ , where  $n_i$  is a 3-corner-hat noise term. The 3-corner-hat estimates clock  $i$  error standard deviation only if the noise is negligible.

### SUMMARY

We have analyzed linear predictors. For each type of instability within the common oscillator model we have

indicated an optimum or close to optimum predictor  $P$  whose performance can be characterized by the standard deviation  $\sigma$  of its error function  $D$ . The predicted time series  $X_m$  can be generalized to any time series such as  $y_m$  (fractional frequency) or  $\langle X_m \rangle$  (averaged over  $\tau$ ) as with modified predictors.

We have described an algorithm for the generation of time scales where the best predictor is chosen by comparing the various  $\sigma$  in real time. This method is simpler and faster than solving for the "a" coefficients in real time using the least square method.

We show that it is possible to improve the estimate of  $\sigma_i$  of a single clock  $i$ , even if masked by added noise, by computing  $\# \sigma$  in an appropriate measurement configuration.

### ACKNOWLEDGMENTS

I gratefully acknowledge helpful discussions with David W Allan and Fred F Walls.

### REFERENCES

1. Lepek A and Walls F L, "Cross Correlation Analysis improves time domain measurements", 47th IEEE International Frequency Control Symposium, 1993
2. Allan D W, Lepek A, Cutler L, Giffard R and Kusters J, "A  $10^{-15}$  International Frequency Reference Available Now", 5th Symposium On Frequency Standards and Metrology, 1995
3. Allan D W, Lepek A, Cutler L, Giffard R and Kusters, "A  $10^{-15}$  International Frequency Reference Available Now", 27th Precision Time and Time Interval Planning Meeting, 1995
4. Barnes J A, and Allan D W, "An approach to the prediction of coordinated universal time", Frequency, 5, Nov./Dec. 1967
5. Weiss M and Weissert T, "AT2, A new Time Scale Algorithm: AT1 Plus Frequency Variance", Metrologia, 28, p65, 1991
6. Allan D W, Weiss M A and Jespersen J L, "A Frequency-Domain View of Time Domain Characterization of Clocks and Frequency Distribution Systems", 45th Annual Symposium on Frequency Control, p 667, 1991
7. Grosslambert J, Fest D, Olivier M and Gagnepain J J, "Characterization of Frequency Fluctuations by Crosscorrelations", 35th Annual Frequency Control Symposium, p 458, 1981

**Preliminary Result of Time Scale Comparison Uncertainty  
by Means of the Computer Controlled A-724M-02 GLONASS Receiver.**

\*Al'shina A. , \*\*Bogdanov P. , \*\*Gevorkyan A. , \*Koshelyaevsky N.,  
\*\*\*Koucherov V. , \*\*Ovchinnikov V. , \*\*\*\*Medvedev Yu. , \*Pushkin S.

\* Institute of Metrology for Time and Space , GP VNIIFTRI, 141570, Mendeleevo, Russia, \*\* Russian Institute of Radionavigation and Time, 193124, St.-Petersburg, Russia, \*\*\* State Time and Standard Frequency Commission, 117049, Moscow, Russia, \*\*\*\* ATZUS, 143040, Golitzino, Russia.

The A-724M-01 type of GLONASS receivers are widely used for time-scale comparison of remote clocks inside and abroad RUSSIA since the end of 1989. Till now it operates under manual control. That's why it is too difficult to realise actual "common view" mode of operation and to get completely the same data acquisition, short-term processing and output data format at each site of clock location. The above mentioned problems limit precision of time synchronization.

Starting fall 1995 in Institute of Metrology for Time and Space (IMVP) in collaboration with Russian Institute of Radionavigation and Time (RIRT) the investigations of A-724-02 type of GLONASS receiver are conducted. This type of receiver differs slightly from previous one by ROM unit and some hardware changes which permit one to get directly satellite - local time scale difference without additional time interval meter. The most striking difference is availability of computer control of this type receivers.

The ensemble of A-724-02 type of GLONASS receivers were investigated and results are presented in common-view mode first of all in the same location in IMVP, then at the distance of about 60 km between Mendeleevo and Golitzino, and then between these two laboratories and St.-Peterburg.

## **RIRT activity in clock synchronization area by means of meteor bursts radio link**

**Prof. Yuri G.Gouzhva, Prof. Arvid G.Gevorkyan,  
Dr. Igor M.Bolotov, Sergey N.Pluzhnikov**

Presented are short history of development the technique of clock synchronization using meteor bursts channel in Russia. Features of the meteoric channel are considered and evaluation of clock synchronization limited accuracy is given. First results are presented and equipment developed in RIRT at early 80-ties, ensuring accuracy of clock synchronization at a level  $\sim 200$  ns on distance up-to 1800 km is described. In early 90-ties an equipment of second generation with an error of  $\sim 15$  ns is developed. Characteristics of a modern equipment of synchronization are indicated.

Prospects of use of meteoric complexes in a LW PPRNS "Chayka" (Loran-C) system for synchronization of secondary stations and transfer of the service information, and as redundant means for metrological support and control of functioning synchronization system of "GLONASS", are considered.

## TRANSFER OF THE PHASE PROCESS THROUGH THE BORDER BETWEEN TWO TIME SCALES

Andrzej Dobrogowski

Institute of Electronics and Telecommunication  
Poznań University of Technology, Poland

### INTRODUCTION

A digital signal (a digital stream) is a bearer of digital (or converted to digital) information when information is transported through digital communication network. A digital signal may be regarded as digital data and discrete instants (discrete points in time). Instants are discrete as the neighbouring ones are separated by some time interval. These discrete instants are provided by a clock generating timing pulses. To achieve undistorted (without slips) transport of data across communication network each digital element (bit) has to find its timing pulse (i.e. timing pulse not carrying another bit) within the whole network.

There are two types of clocks involved in data transport, the regular clock and the gapped clock. A regular clock has all its timing pulses evenly distributed in time. The regular clock supplies time scale for digital processes. It is rather common situation that regular clocks being in plesiochronous relationship provide time scales for different regions of communication network. This is the case when a number of Primary Reference Sources is used for timing the network. Generally, there is no one-to-one correspondence between timing pulses of regular clocks at both sides of the time scales' border. Therefore it is not possible to use all timing pulses of each regular clock for carrying the transported data.

Generally, transported data (payloads) are timed using gapped clocks (especially in Synchronous Digital Hierarchy). A gapped clock is derived from a regular clock by removing a number of clock pulses (or leaving "gaps" in the clock where the regular clock would have a pulse; Sexton and Reid (1)). Adapting the size of gaps it is possible to establish one-to-one correspondence between the pulses of gapped clocks at both sides of the time border, what we can interpret as transfer of some phase process across the border between two time scales or we can consider this mechanism in terms of synchronization of the phase processes characterizing gapped clocks at both sides of the border.

In the paper the phase processes of regular and gapped clocks are defined. The synchronization of gapped clocks is considered. Mechanism of transferring the phase process of Virtual Container level 4 (VC-4) across the time border is studied.

### CLOCK PHASE PROCESS

An example of timing pulses of a regular clock and its phase process is given in Fig.1. Successive count of clock pulses (and usually multiplying the count by  $2\pi$ ) gives us the phase of clock (observable phase time) at discrete (characteristic) instants. A specific course of the clock phase between characteristic instants depends on accepted manner of interpolation. Stepwise and linear interpolation are shown in Fig.1.

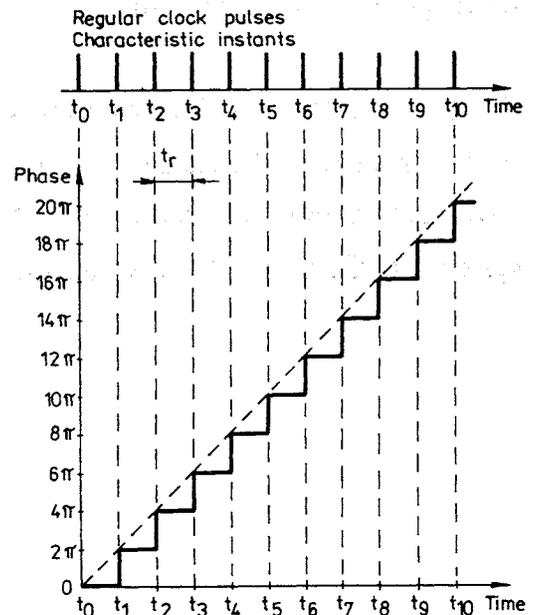


Fig. 1. A regular clock

In Fig.2 a gapped clock derived from regular clock of Fig.1 is presented. Linear approximation for the phase clock process is assumed. A slope of dashed line in Fig.2 represents the average frequency of the gapped clock in the corresponding time interval. Controlling the size and number of gaps the desired average frequency of the clock can be set.

It should be well understood that any digital signal of significant bit rate has to be transmitted across any distance using a regular clock (1). The timing pulses of a regular clock carry payload as well as overheads and dummy bits (bytes). These dummy bits can be replaced by data when it is necessary. The size and number of gaps (within a given time interval, e.g., frame period)

of gapped clock depend mainly on necessary overheads. On the other hand, variation of the size of gaps depends on changing payload rate.

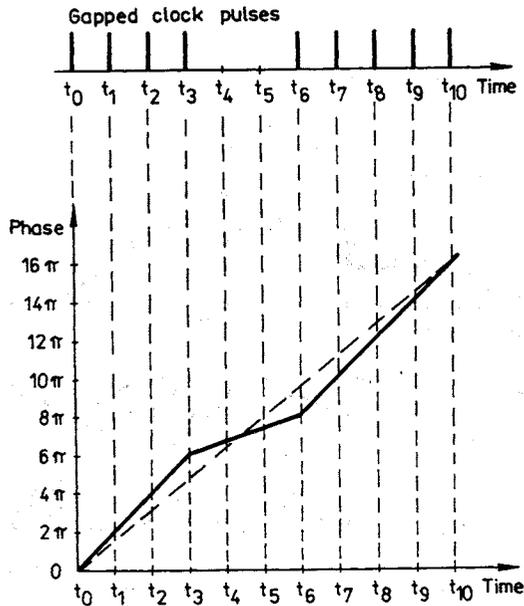


Fig.2. A gapped clock

### SYNCHRONIZATION OF GAPPED CLOCKS

Synchronization is a mapping of phase process of a synchronizing gapped clock into phase process of a synchronized gapped clock. Domains of different time scales are interfaced by an elastic store (FIFO) shown in Fig.3. Transferring data (payload) are written into the elastic store using the gapped clock of incoming

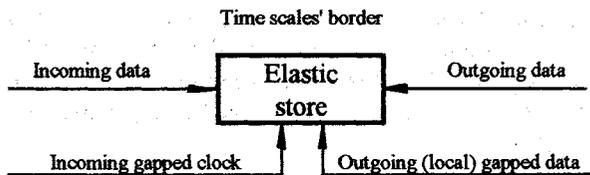


Fig.3. An elastic store at time scales' border

payload (domain of one time scale). Then the data are read from the elastic store using the outgoing gapped clock (domain of another time scale). The elastic store can absorb minor variation of phase difference between gapped clocks. The elastic store's capacity and resulting delay is the price for that ability. Phase variation is minor if the elastic store fill ranges between lower threshold and upper threshold. Major phase variation (variation which results in crossing the elastic store threshold) forces a control logic (not shown in Fig.3) to change the reading pattern from the elastic store. In Synchronous Digital Hierarchy (SDH) environment an elastic store and its control logic is called pointer processor.

When the outgoing clock is subject to excessive delay in relation to the incoming gapped clock, a designed number (24 in SDH VC-4 gapped clock) of timing pulses is inputted into the intended gap (making this gap shorter) which results in advancing the phase of the outgoing gapped clock. When the delay of the outgoing gapped clock in relation to the incoming gapped clock becomes too small, the intended gap is lengthened removing a designed number (24 in SDH VC-4 gapped clock) of timing pulses immediately following the intended gap which results in retardation of the outgoing gapped clock. The described procedures are called negative justification and positive justification correspondingly.

The possible relationship between phase processes of the incoming and outgoing gapped clocks clocking the elastic store is presented in Fig.4; Dobrogowski (2). The incoming regular clock is slower than the outgoing regular clock thus the positive justification occurs to keep both gapped clocks in synchronism (to keep gapped clocks' phase processes close enough). The lines representing phase processes of gapped clocks must not cross in order to avoid emptying the elastic store (bits would be read twice). On the other hand vertical distance between the lines must not exceed the capacity of the elastic store (some bits would be omitted during read out).

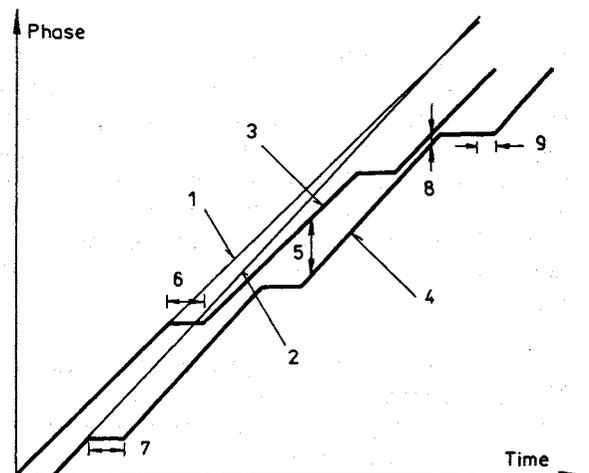


Fig.4. The relationship between phase processes at input and output of an elastic store:

- 1 - incoming regular clock; 2 - outgoing (local) regular clock; 3 - incoming gapped clock; 4 - outgoing (local) gapped clock; 5 - elastic store fill; 6 - clock's gap: writing disabled; 7 - clock's gap: reading disabled; 8 - the decision to make positive justification; 9 - positive justification

Generally, the transported digital stream is structured in frames. Sequential transmission of bits results in

corresponding shift of frame phase when justification occurs. Negative justification results in positive frame phase increment and positive justification results in negative one, which is clear in Fig. 5.

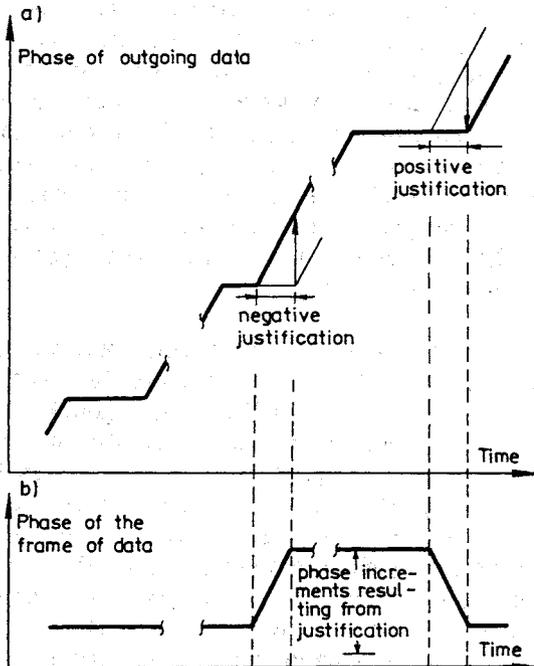


Fig. 5. Phase of outgoing data:

a) increments of data stream instantaneous phase resulting from justification; b) phase of the frame of data stream

It is also clear that the movement of frame phase coincides with justification timing pulses (let us notice, that positive justification means no timing pulses in a gapped clock). Usually the change of the frame phase resulting from justification is shown as a step change. This is because the duration of the time interval involved in justification is much shorter than the duration of the time interval between justification events.

Synchronization of gapped clocks, timing the Virtual Container level 4 (VC-4 is some signal structure of SDH) transferring across time border, is demonstrated in Fig. 6; (2). It was assumed that the frequency of outgoing (local) regular clock is a reference. Within the time interval  $0 \rightarrow t_1$  the frequency of incoming regular clock displays a positive offset with respect to the reference. From  $t_1$  to  $t_2$  the offset is negative. Frame phases of incoming and outgoing VC-4 are referenced to the value they have at  $t=0$ .

In lower part of Fig. 6 variation of elastic store fill is presented. The presentation corresponds to the most frequent situation when the gaps of incoming gapped clock do not overlap the gaps of outgoing gapped clock. In Fig. 6 sudden changes of the elastic store fill are suggested (arrows). Obviously, the elastic store fill is a single-valued function of time. The decision to

present the fill as a double-valued function comes from the fact, that the value of nominal ratio of the time interval with gaps to the time interval without gaps is 72/2088.

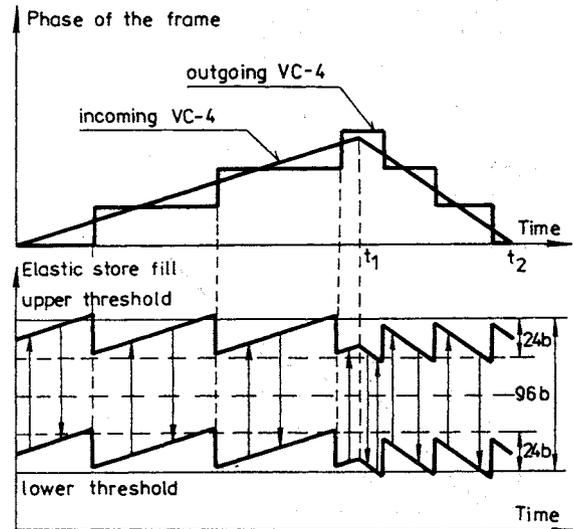


Fig. 6. Relationship between frame's phase of incoming and outgoing VC-4 and elastic store fill

Fig. 6 clarifies a good short-term stability condition imposed on SDH equipment clocks (ITU-T G.81s). After necessary justification (Pointer Justification Event - PJE) the clock's phase noise might cause an undesired PJE of the opposite sign because the elastic store fill is periodically close to the opposite elastic store threshold.

A course of the frame phase of outgoing VC-4 is a quantized version of a course of the frame phase of incoming VC-4, where the quantum is a phase increment which is equal to the number of justification timing pulses.

Using Fig. 6 helps us to predict the results of justification in incoming VC-4. Justification in incoming gapped clock influences a lower fill of elastic store. E.g., negative justification adds 24 bits to the lower fill of the store. Thus there is a great chance that the next gap in outgoing gapped clock, lasting 72 timing pulses, will result in crossing the upper threshold of the elastic store and negative justification at the nearest justification opportunity in outgoing gapped clock (and outgoing VC-4).

From Fig. 6 it is straightforward to draw input-output characteristic of the elastic store (pointer processor in SDH), Fig. 7. It is easy to notice that for adopted pattern of gaps in VC-4 and the elastic store capacity suggested by ITU-T G.783 the transfer of phase process features

in small hysteresis, but it may results in large number of PJE.

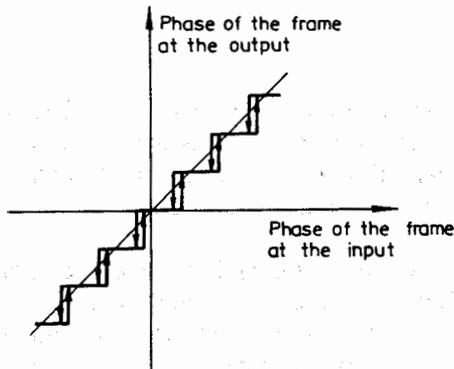


Fig. 7. Input - output relationship of an elastic store

## CONCLUSIONS

The transfer of a digital signal across the border between different time scales can be effectively analysed in terms of synchronization of corresponding gapped clocks clocking the data stream at both sides of the border. Analysis is strongly supported by graphs presented in the paper. Parameters of synchronization process (e.g., PJE rate, hysteresis) depend on the elastic store capacity and pattern of gaps (including justification) of gapped clocks.

## REFERENCES

1. Sexton M., and Reid A., 1992, "Transmission networking. SONET and the Synchronous Digital Hierarchy", Artech House, Boston - London.
2. Dobrogowski A., 1995, "Timing and synchronization in SDH multiplexers" (in Polish), Przegląd Telekomunikacyjny, 10, 562-572.

## A COMPUTER CONTROLLED MULTI-REFERENCE CLOCK SYSTEM

Eskelinen P., Engelberg J., Salpala A. and Matala M.  
 Kotka Institute of Technology  
 Department of Electrical Engineering/Research laboratory  
 Finland

### ABSTRACT

This paper first describes an end-user attempt for a computerized, multi-reference clock system and its main components, then deals with the efforts of providing automatically one physical time and frequency output with the least possible historical delay and finally discusses the achieved accuracy. External references are provided through separate GPS clocks and a TV-line receiver. The primary internal clock is a rubidium oscillator assisted by an ovenized crystal unit and another rubidium, which are both controllable through a computer interface. Some of the microprocessor-based GPS receivers show unexpected behaviour and have frequently lost sync. TV-line frequency reception is scrambled by live transmissions extended to the mid day-time causing unpredictable phase jumps. The high quality rubidium, after several weeks of operation below the  $10^{-12}$  level, may show abrupt frequency deviations. Main problems are associated with the low-drift, high resolution D/A-conversion and 10 MHz EMI from the rubidium unit. The needed automatic frequency correction is less than  $2 \times 10^{-10}$  / day.

### SYSTEM CONFIGURATION

The configuration described here was originally developed for the project outlined in Eskelinen et al [1.], but has turned out to be well worth continuing as such and is currently under rapid development. The clock system produces separate, controlled 1 pps and 5 MHz outputs, the long term stability of which is based on 1 Hz signals from autonomous GPS receivers and a modified TV-linesync device, the latter being locked to the national standard frequency with a known error. Two rubidium units are utilized, with Rb1 as a flywheel and Rb2 as a controlled source. An OCXO operates parallel to Rb2, but with a control algorithm of its own. The necessary measurements are performed with three HP time interval counters and an 8½-digit DVM, which are all controlled by a 9000/300-

series workstation. This also has a DC-link to the OCXO, Rb2 and to a temperature probe. A multichannel matrix is used to combine test instruments and signal outputs.

The system lay-out seen in the block diagram of Fig.1 has three main functional parts. Most of the sensors are mounted in the rack, which also has the back-up batteries for about 24 hours and an additional frequency counter for sporadic measurements. Most test instruments are stacked on each other and connected together with the IEEE-bus. The controlling computer and associated coaxial switching and D/A devices are mounted separately due to operator convenience.

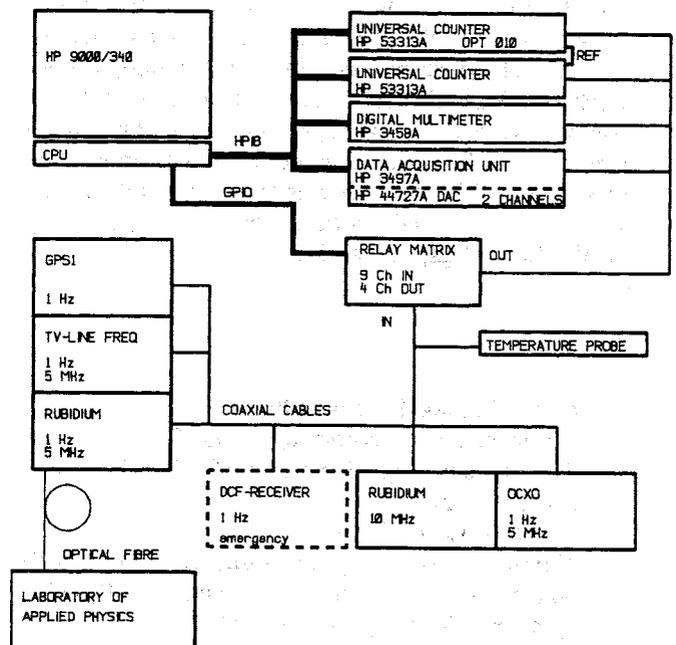


Fig.1. The clock system with sensors, test instruments and controlling equipment

The computer gathers data and performs a live statistical analysis, which yields on-line information on the differences between various sources and is also used to define the estimate for current time and frequency and to adjust the the Rb2-unit and the OCXO so, that a real physical 1 Hz pulse plus the reference frequency are provided with the least error.

The correction cycle is currently 4 hours and modified linear regression is utilized for both TV- and GPS-based results. In the background, the program stores test data on disk, which is analyzed later on with another computer for further improvements.

## SENSORS

### Gps

Three different GPS receivers have been tested in the system, which provide a sub-microsecond error when operating properly. The unavoidable S/A impairing procedure as observed in Kotka is shown in the histogram of Fig.2. In this case an HP58503A sensor was used. The nice behaviour of this particular HP unit can also be seen in the time domain recordings with no sync losses during a three weeks continuous measurement despite only the western half of the sky was visible from the antenna mounting position.

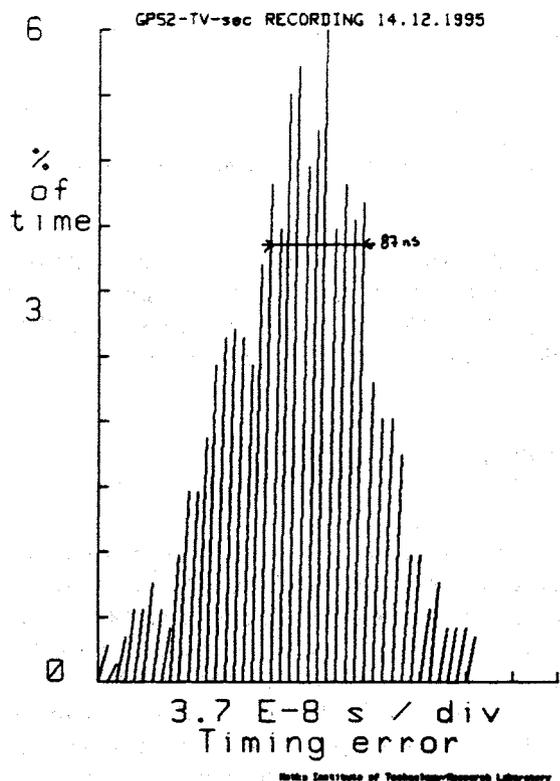


Fig.2. The observed effects of S/A as measured in Kotka show a timing uncertainty of 87 ns for a test period of 150 hours.

Less satisfactory results have been obtained with the Navstar XR-receiver. A performance similar to that reported e.g. by Kalliomäki and Mansten [2.], Azoubib and Thomas [3.] or

Hartmut and Klische [4.] has not been available. Identical total test periods, when compared with the HP unit, show several losses of all satellites, sometimes once an hour, sometimes once per day, see Fig.3. The statistics over half a year show a cut every 2.5 hours (mean value) or a unit being out of service for 28 hours (present worst-case value), but the duration of most cuts is, however, only a couple of minutes. The problem would not be such annoying if the recovery process were not extremely hard to automate. Some preliminary results exist, which suggest the total sync-loss to be also dependent on RF-interference. The Navstar antenna is mounted on top of the institute's main building and has thus a complete view of the sky.

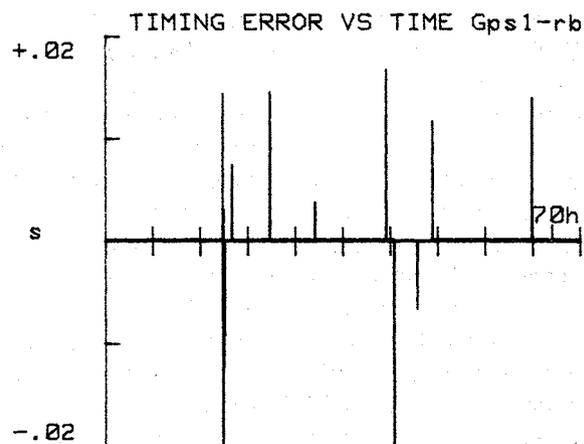


Fig.3. The Navstar XR-series receiver is not capable of providing a continuous 1 pps clock output but suffers from frequent cuts, which normally last for some minutes.

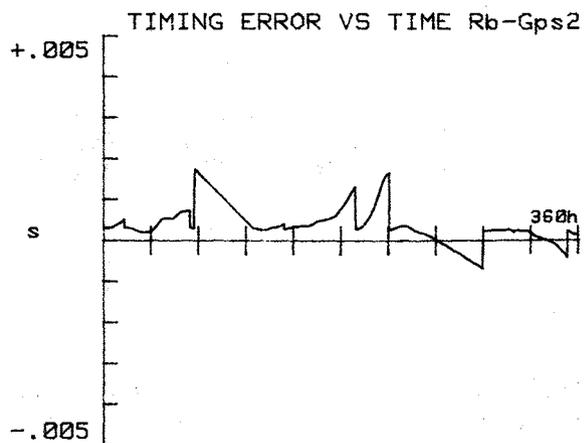


Fig.4. The Collins GPS-receiver seems to be out of lock all the time despite the operator interface shows a healthy state.

The Collins GPS receiver does not provide any useful timing information as can be seen in the

respective plot of Fig.4. The unit achieves lock for only a couple of seconds and generally is out of lock for several tens of hours per a one week test period. Sometimes an attempt to lock is visible, but obviously the acquisition algorithm of the unit does not match the present satellite constellation. The self-test routine and the terminal display do not, however, inform the user of any possible problems. No improvement was noticed when different antenna positions were tried.

#### Tv-line

The TV-transmission chain is locked by the video switching center in Helsinki to an R&S rubidium oscillator, which is monitored by the Technical Research Center of Finland.

The commercial TV-sync separator device, mounted on top of a TV set, provides originally a standard frequency output, the behaviour of which is plotted in Fig.5 over a 60 hours test period with a 10 seconds averaging time per sample. Severe short term instabilities of the order of  $4E-8$  are visible due to sudden changes in the video path and as such the output is not suitable for direct control purposes. Obviously, the more stable portions of the recording with sigma values below  $5E-9$ , are observed at night, when a stable video connection is established at the switching center. For the actual computations a modified 1 pps output is used.

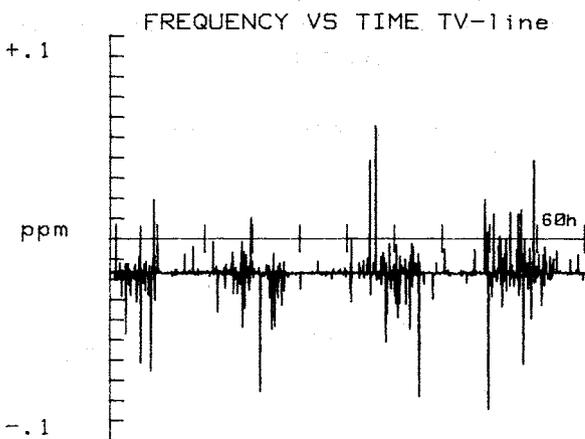


Fig.5. The TV-sync signal suffers from several abrupt phase jumps due to changes in program or microwave chain. The variation between day and night is clearly visible.

Some information of the general observed performance of the total TV chain was obtained during the HP GPS-tests. A measurement over three weeks, plotted in Fig.6 shows a mean

frequency difference of  $6E-12$  between the GPS clocks and the TV-sync system and daily variations within  $2E-12$ . The strange phenomenon of MJD 50064 has not yet been solved, but it might be due to a switching action of the microwave system. A discrepancy between the GPS-units was not visible.

From the two results of Figs. 5 and 6 we can conclude that an integration time of roughly 24 hours is suitable, if the desired, controlled frequency accuracy is of the order of  $10E-11$ . The question is, if commercial oscillators can be left running alone for such a long time.

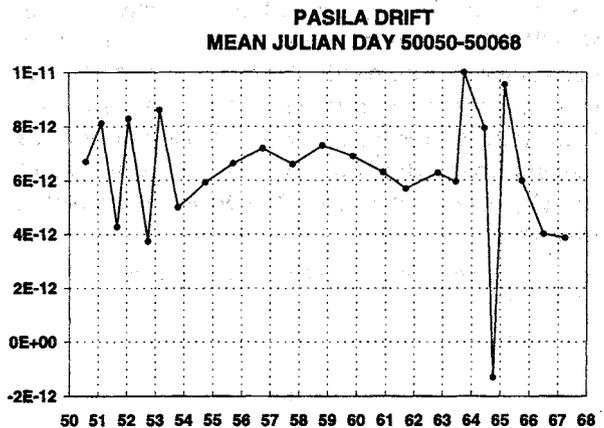


Fig.6. A recording over three weeks shows, that the TV-chain up to Kotka is running at an offset of  $6E-12$  with reference to GPS and that typical daily variations are around  $2E-12$ .

#### Rubidium

The clock system uses two different rubidium oscillators (Rb1 and Rb2). A fairly long test of 360 hours, visible in Fig.7, with an Efratom unit shows some ugly behaviour as the generally stable frequency profile of the flywheel unit, which stays within  $1E-10$ , suddenly drops to  $2E-10$  and below for a couple of days with no simultaneous indications of either temperature, pressure or supply changes and even more amazingly re-establishes with no external intervention or control applied. Similar bounces have been observed randomly along the year.

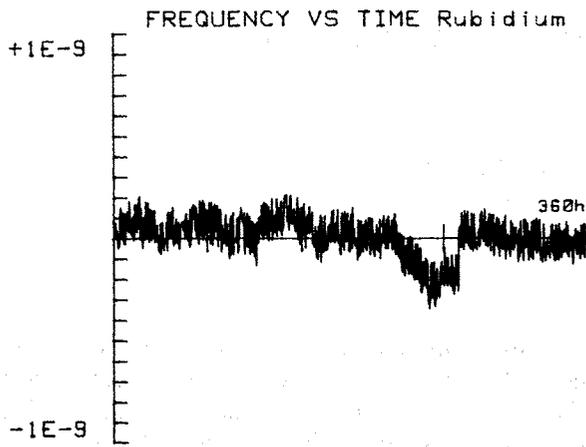


Fig.7. The rubidium oscillator shows occasionally frequency jumps of unknown origin.

A minor bug is the factory made buffer circuitry, which is not capable of isolating the rubidium oscillator from its load. With a 1 meter coaxial cable at the 10 MHz output, the frequency offset is  $3E-11$  and without it only  $1.7E-11$ .

SYNTHESIZED OUTPUT

The present algorithm uses the 1 pps outputs of the GPS receiver and the TV-sync system for a comparison with the 1 pps signal from the Rb1 flywheel as pairs GPS-Rb1 and TV-Rb1. To minimize the effects of S/A, an average over 50 seconds is first calculated. In case both the GPS and TV agree, the GPS is selected to numerically correct for the long term offset of the Rb1 frequency (no actual adjustment is performed) and simultaneously for the offsets of the counters. If the GPS-Rb1 shows, due to loss of sync, a difference greater than 10 us per 4 hours, the TV-chain is selected. If both comparisons show an error greater than 10 us, the numerical correction values are not updated. Unless the Rb1 unit had its current unpredictability, the correction cycle would be extendable up to 100 hours or more.

The screen shows a live display of current observed timing offsets with also the 1 pps output of the OCXO as Xtal-Rb1. In the frequency table, a photo of which is shown in Fig.8, the Rb1 offset is the numerical correction determined by 1 pps comparisons. Rb2 and OCXO frequencies are controlled through an HP 44727A Data Acquisition Unit against the Rb1 frequency. An averaging time of 10 seconds is normally used for these through an internal loop in the program.

The Rb2 unit is controllable to at least  $1E-11$  if a 24 hour integration time is allowed for the 1 pps process. The OCXO unit suffers from the tuning sensitivity of the crystal and is presently maintained within  $2E-9$ . It also produces the 1 pps output of the clock system with the same relative uncertainty. The controlled OCXO is improving its performance all the time as the software parameters are being optimized.

SENSORS:		SECOND PULSE DIFFERENCES (seconds)			FREQUENCY DIFFERENCES (ppm)			ENVIRO
	Rb1	GPS1	TV_Rb1	Xtal_Rb1	Rb1	Rb2	Xtal_0	TEMP
2	-5.81E-8	9.6E-8	0.001E-8		-0.17012	-0.10120E-9	0.070002	25.33
3	1.074E-7	1.0E-8	1.410E-8		-0.17022	-0.00000E-9	0.000000	25.30
4	1.09E-7	2.2E-8	2.40E-8		-0.17011	-0.01000E-9	0.000000	25.34
5	0.7E-8	5.41E-8	3.00E-8		-0.17015	-0.04000E-9	0.001104	25.32
6	1.00E-7	4.7E-8	0.00E-8		-0.17032	-0.00000E-9	0.000001	25.10
7	1.20E-7	4.0E-8	0.00E-8		-0.17010	-0.00000E-9	0.000000	25.27
8	-4.431E-7	1.70E-7	0.00E-8		-0.17020	-0.02000E-9	0.000000	25.31
9	6.74E-8	1.4E-7	0.00E-8		-0.17014	-0.00000E-9	0.000007	25.28
10	1.70E-7	2.10E-7	0.00E-8		-0.17009	-0.00000E-9	0.000000	25.32
11	0.84E-8	1.12E-7	0.00E-8		-0.17020	-0.00000E-9	0.000000	25.32
12	1.51E-7	2.410E-7	0.00E-8		-0.170	-0.00000E-9	0.000000	25.28
13	-1.70E-7	2.33E-7	0.00E-8		-0.17006	-0.00000E-9	0.001100	25.28
14	1.00E-7	2.30E-7	0.00E-8		-0.17020	-0.00000E-9	0.000000	25.28
15	4.84E-8	1.70E-7	0.00E-8		-0.17001	-0.04000E-9	0.000001	25.28
16	1.2E-8	1.77E-7	0.00E-8		-0.17001	-0.00000E-9	0.000000	25.27
17	-2.2E-8	0.00E-8	0.00E-8		-0.17003	-0.00000E-9	0.001100	25.28
18	0.84E-8	1.12E-7	0.00E-8		-0.17020	-0.00000E-9	0.000000	25.33
19	1.30E-7	0.00E-8	0.00E-8		-0.17001	-0.00000E-9	0.001100	25.14
20	-4.2E-8	0.00E-8	0.00E-8		-0.17073	-0.00000E-9	0.000000	25.27
21	-0.2E-8	0.00E-8	0.00E-8		-0.17020	-0.00000E-9	0.000000	25.27
22	1.74E-8	0.00E-8	0.00E-8		-0.17008	-0.00000E-9	0.000000	25.1
23	0.70E-8	0.00E-8	0.00E-8		-0.17008	-0.00000E-9	0.000000	25.21
24	1.40E-8	0.00E-8	0.00E-8		-0.17020	-0.00000E-9	0.000000	25.42
25	-0.00E-8	0.00E-8	0.00E-8		-0.17007	-0.00000E-9	0.000000	25.42

Fig.8. The controller display shows the live situation of the two primary sensors (GPS and TV-line) and the calculated relative frequency errors of Rb2 and OCXO.

REFERENCES

1. Eskelinen P., Engelberg J., Frilander J. and Matola M., 1995, "A Low Cost Ovenized Crystal Oscillator", Proc. 9th EFTF
2. Kalliomäki K. and Mansten T., 1995, "Phase Stability of XR4GPS-receivers at VTT and University of Oulu", Proc. 9th EFTF
3. Azoubib J. and Thomas C., 1995, "The use of hydrogen masers in TAI computation", Proc. 9th EFTF
4. Hartmut N. and Klische W., 1995, "Highly Accurate Realtime Clocksynchronization with GPS Time Receivers", Proc. 9th EFTF

## A METHOD FOR QUARTZ OSCILLATOR SYNCHRONIZATION BY GPS SIGNAL

A. Lisowiec \*, A. Czarnecki \*, Z. Rau \*\*

\*) Tele and Radio Research Institute, Ratuszowa 11, 03-450 Warsaw, Poland

\*\*\*) Telecommunication Institute, Szachowa 1, 04-894 Warsaw, Poland

### ABSTRACT

In the paper a GPS controlled frequency reference source has been described. The source combines the advantages of very good short term stability of high quality quartz oscillator with advantages of GPS signal over the long term. Details of the internal workings of the source have been presented. The deviation of the local oscillator from the ideal reference is determined by counting the pulses of the multiplied local oscillator frequency between successive GPS pulses. Special arrangement of counting circuitry eliminates discrete error. The correction voltage at the oscillator's control input consists of two terms. One term is computed from averaging the counted pulses over 1s interval and is thus responsible for minimizing the short term frequency deviation. The other term is computed from the accumulated phase time drift. The control algorithm ensures that the frequency accuracy averaged over one day is significantly better than  $10^{-11}$  and the peak accumulation of phase time drift does not exceed 200 ns.

### INTRODUCTION

The wide availability of GPS signal enables construction of synchronized frequency sources/clocks of  $10^{-12}$  frequency accuracy for one day average and timing accuracy  $< 100$  ns with respect to UTC (SA on). These frequency sources/clocks find application in telecommunication networks as alternatives or complements to cesium sources and in measurement laboratories. The aim was to build an inexpensive frequency reference disciplined to GPS around in house manufactured VCXO and test its suitability for synchronizing SDH networks. The aim was to achieve frequency accuracy better than  $10^{-11}$  when averaged over one day period and peak phase time drift relative to ideal reference less than 40 ns.

### FUNCTIONAL DESCRIPTION

#### Principle of operation

Block diagram of the GPS controlled source is presented in fig. 1. The local oscillator VCXO is a high stability ovenized quartz oscillator. The VCXO frequency is multiplied to 100 MHz and the 100 MHz pulses are counted by two counters in time intervals between successive GPS pulses. The counters count the pulses alternately with zero dead time, which means that none of the pulses are omitted or counted twice.

Thus the number of counted pulses after subtracting ideal reference marks the accumulated phase time drift. The deviation of the oscillator's frequency from nominal is determined by averaging the number of pulses counted over 1s interval. Based on this deviation the voltage at the control input of the local oscillator is modified so as to minimize the long time drift. The control algorithm is written into the EPROM memory and executed by the 8051 microcontroller. Alternatively the source can be fully remote controlled via the RS-232 interface by the IBM PC computer. The source is equipped with an output for the recorder. The voltage at this output is proportional to the residue of dividing the accumulated phase time drift by 1  $\mu$ s.

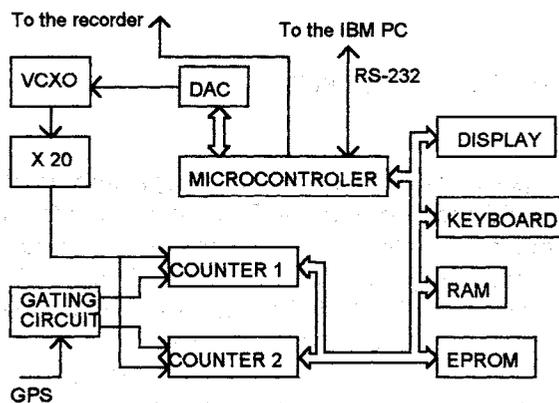


Fig. 1. Block diagram of the GPS controlled source.

#### The control algorithm

The equation from which the deviation of the local oscillator frequency from nominal is computed has the form (fig. 2 demonstrates the summation scheme):

$$\overline{df} = \frac{1}{f_N p N (A + 1)} \sum_{j=0}^A \sum_{i=1}^N S_{k+j+i-1} \quad (1)$$

where:

- p is the measurement interval in seconds,
- $f_N$  is nominal frequency multiplied by 20,
- $S_{k+j+i-1}$  is the remainder from subtracting  $p \cdot f_N$  from the number of pulses counted over p seconds interval starting at  $k+j+i-1$  seconds,
- N and A are constants.

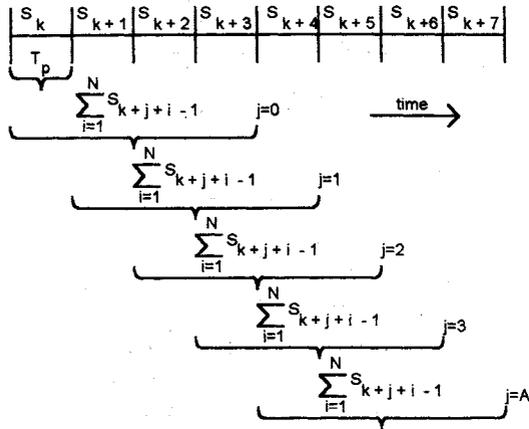


Fig. 2. Illustration of the summation scheme of equ. (1).

From the deviation determined by the formula (1) a correction voltage

$$\Delta U_1 = \overline{df} k \varepsilon \quad (2)$$

is applied to the control input of the VCXO oscillator where  $k$  is the slope of the frequency control characteristic and  $\varepsilon$  is a coefficient dependent on  $N$  and  $A$  from equation (1).

In the summation formula (1) the highest weights have measurements lying in the middle of the summation interval, the weights vector being of the form

$$\left[ 1, 2, \dots, A, \underbrace{A+1, A+1, \dots, A+1}_{N-A}, A, \dots, 2, 1 \right] \quad (3)$$

Rearranging this vector to

$$\left[ 1, 1, 2, 2, \dots, A, A, \underbrace{A+1, \dots, A+1}_{N-A} \right] \quad (4)$$

gives highest weights to the most recent results and improves the dynamic properties of the algorithm.

The behavior of the control algorithm for various values of  $N$ ,  $A$  and  $\varepsilon$  was simulated on a computer. The local oscillator was modeled as a source with an initial frequency inaccuracy and linear drift in time. The tool well suited for such simulations is a *Mathematica* software package. The *Mathematica* program used in simulation trials is given in the appendix.

Fig. 3 shows the simulated response of the algorithm for the initial frequency inaccuracy where the ordinate shows frequency deviation as computed by equation (1) while Fig. 4 shows the plot of phase time drift over the same time period. The values of  $N=14$ ,  $A=14$  and  $p=60$  were chosen from the examination of the spectrum of GPS signal ( $p(N+A)$  seconds is the time constant of control loop) and the value of  $\varepsilon=N/2$  was chosen for critical damping.

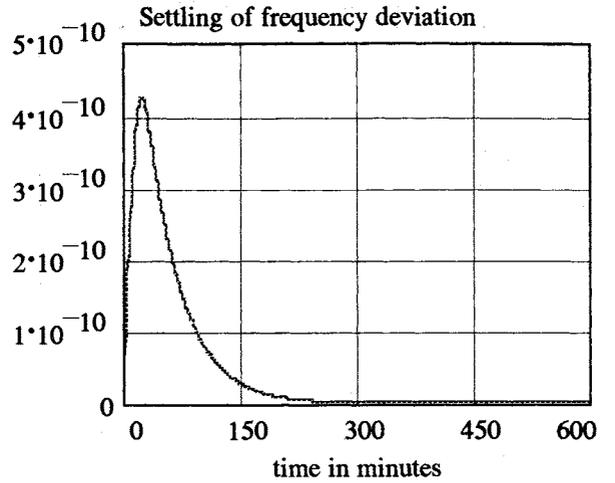


Fig. 3. Response of the control algorithm for the initial frequency inaccuracy.

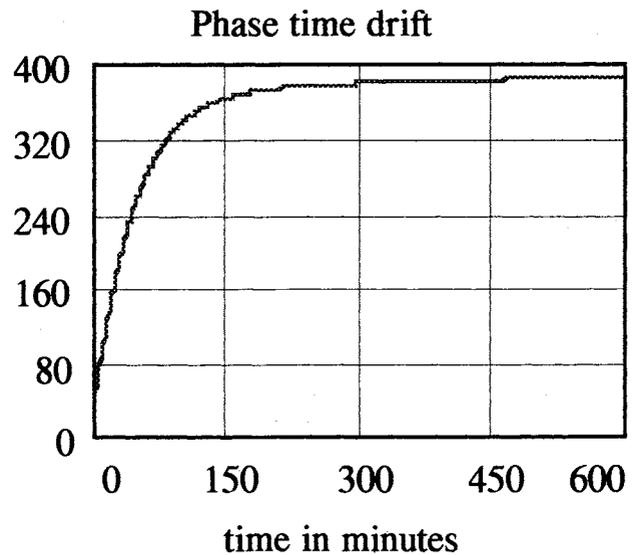


Fig. 4. Phase time behavior of the control algorithm without phase time correction term.

It was found that the phase time drift couldn't be eliminated without including into the control voltage the correction term depending directly on phase time drift. The form of this correction term worked out experimentally was

$$\Delta U_2 = a \delta \quad (5)$$

where  $a$  is the slope of the regression line fitting the phase time drift curve over several most recent measurement results and  $\delta$  is the coefficient equal to  $1/5N$ .

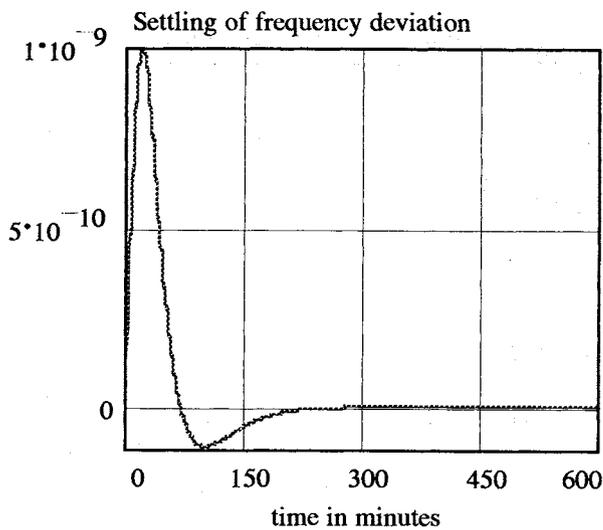


Fig. 5. Response of the control algorithm for the initial frequency inaccuracy.

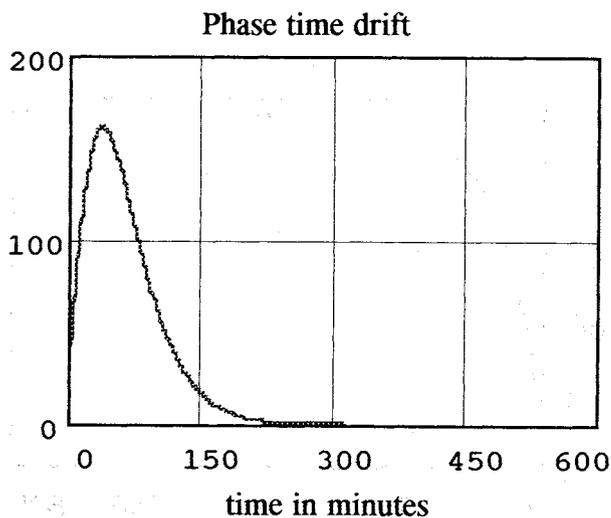


Fig. 6. Phase time response of the control algorithm with phase time correction term.

Figures 5 and 6 show the response of the control algorithm when the voltage applied at the control input of the local VCXO oscillator was the sum of  $\Delta U_1$  and  $\Delta U_2$ . The phase time drift settled over time to zero which means that the long term accuracy approached the long term accuracy of GPS. Comparing figures 3 and 5 it could be seen that correcting for zero phase time drift disturbed slightly the frequency control. In the response of the control algorithm for the initial frequency inaccuracy there appeared an undershoot and the response took longer time to settle to zero.

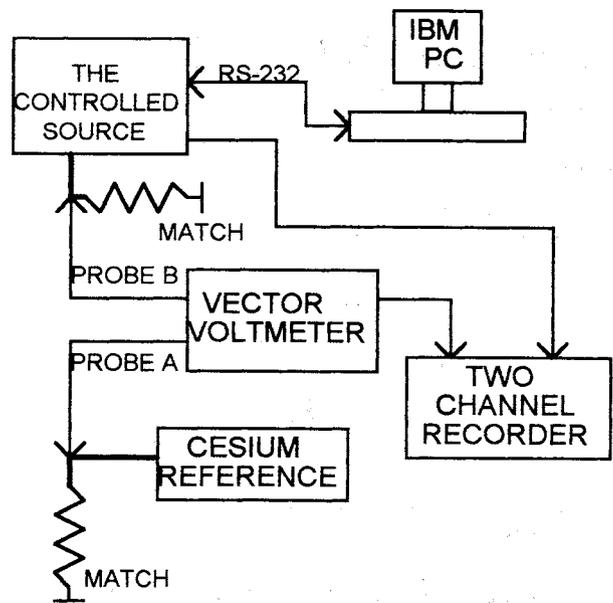


Fig. 7. Measurement setup.

The control algorithm described above was implemented on IBM PC computer controlling the source via RS-232 and tested in a setup of fig. 7. Data obtained during the testing was evaluated by the methods of the least sum of squares LSQ and the sequential LSQ, Rau(2). It was confirmed that the loop time constant of  $p(N+A) > 1200$  was long enough to filter out the GPS jitter.

In the absence of GPS pulses the source entered the holdover mode in which the local oscillator's drift was compensated for using the data obtained when the source was locked to GPS.

It must be said however that from time to time the local oscillator exhibited some unexpected behavior which demonstrated itself as fluctuations of a few  $10^{-11}$  over a period of about an hour. These fluctuations were of too short duration to be filtered out by the control loop. They contributed to the temporary increase of the phase time drift to 200 ns. It is believed that the source of these fluctuations was temperature.

## CONCLUSION

Testing the GPS controlled frequency reference confirmed the effectiveness of the method of measuring the deviation of the local oscillator frequency relative to GPS signal. The control algorithm assures the long term accuracy of better than a few  $10^{-12}$  when averaged over several days period. However due to fluctuations of the local oscillator over a period of several minutes to a couple of hours, temporary frequency deviation and phase time drift exceed limits imposed on frequency references by SDH requirements. The solution to this problem is to include into the control algorithm the correction for temperature dependence of the local oscillator frequency or/and to improve the temperature control of the local oscillator. At present some work is

carried out at ITR on improved temperature chamber for VCXO used primarily in synchronized frequency sources.

## APPENDIX

The list **data** contains frequency deviation of the local oscillator as computed by equ. 1. The list **cf** contains accumulated phase time drift. The list **dT** contains random numbers simulating GPS jitter. The list **wsp** contains the weights list of the form from equ. (4). **Wspcz** and **wspfz** are the coefficients  $\varepsilon$  and  $\delta$  from equ. (2) and (5).

```
fnom=1.0*10^8;          (*nominal frequency*)
dryft=10^-10/(60*24);  (*drift per day*)
dfpocz=10^-9;         (*initial inaccuracy*)
slope=1.5*10^-8;      (*tuning coefficient*)
wspcz=0.5;wspfz=0.2;  (*coefficients*)
p=60;                 (*measurement int. in sec*)
m=p*600;              (*test period in minutes*)
faza[0]=30;          (*initial phase inaccuracy*)
N=14;                 (*loop parameter*)
A=14;                 (*loop parameter*)
wsp=Join[Flatten[Table[{i,i},{i,Min[N,A]}]],
          Table[Min[N,A]+1,{i,Abs[N-A]}]];
dT=Table[0*(0.5*10^-7*(Random[]-0.5)),{i,m}];
data={};cf={};Utable={};izf=1;f[0]=0;U[0]=0;
dU[0]=0;
ZT=Table[0,{i,N+A}];ZF=Table[0,{i,10}];
dfpoprz=0;dfakt=0;
For[j=1,j<=m,j++,f[j]=p*fnom*(j*dryft+
  U[j-1]*k+dT[[j]]/p + dfpocz);
  For[jk=1,jk<N+A,jk++,ZT[[jk]]=ZT[[jk+1]]];
  ZT[[N+A]]=f[j];
  faza[j]=faza[j-1]+f[j];
  ZF[[izf]]=faza[j];izf=izf+1;If[izf==11,izf=1,izf=izf];
  dfpoprz=dfakt;dfakt=(Apply[Plus,ZF])/10;
  korfaz=dfakt;
  dU[j] = -(Dot[ZT,wsp]*wspcz)/(N*(A+1))+
    (korfaz*wspfz)/(N+A)/(slope*p*N*fnom);
  U[j]=U[j-1]+dU[j];
  AppendTo[data,-dU[j]*slope*N*(1/wspcz)];
  AppendTo[cf,faza[j]];]
```

## REFERENCES

1. Lisowiec A., Czarnecki A., 1994, "Digital time comparator of 1 ns resolution for 10 s observation time", 8th Piezoelectric Conference PIEZO'94 Proc., 257-259
2. Rau Z., 1995, "The application of the least sum of squares method in the examination of the standards and cloks", (in polish), Works Of The Telecommunication Institute, 104

## INFLUENCE OF THE IONOSPHERIC DELAYS ON THE SUBNANOSECOND ACCURACY OF KU-BAND TWO-WAY SATELLITE TIME AND FREQUENCY TRANSFER.

Pierre J-M. Uhrich

BNM-LPTF, Observatoire de Paris,  
61 Avenue de l'Observatoire, F-75014 Paris, France.

### ABSTRACT

Even if the subnanosecond accuracy of clock comparisons is not reached yet by the Two-Way Satellite Time and Frequency Transfer (TWSTFT), this technique appears to be the best available today for the frequency comparisons of the new generation of primary atomic frequency standards. It is generally assumed that the path non-reciprocity due to the influence of the ionosphere can be neglected for TWSTFT using Ku-Band carriers within an accuracy of 100 ps (1  $\sigma$ ).

This paper presents some estimations of ionospheric delays for future frequency comparisons by TWSTFT between the primary standards of the Laboratoire primaire du temps et des fréquences du Bureau national de métrologie at Paris Observatory, France, of the National Institute of Standards and Technology, United States of America, and of the Physikalisch Technisches Bundesanstalt, Germany. The computations are made for different approximate configurations of solar activities. The results show that specific schedules should be prepared for the comparisons of primary standards, where the frequency stability of the phase measurements should stay within  $10^{-15}$  or better.

### INTRODUCTION

The recent accuracy estimation of the primary standard of the Laboratoire primaire du temps et des fréquences du Bureau national de métrologie (BNM-LPTF), a cold atoms Cesium fountain, is at the level of 3 parts in  $10^{-15}$  (1  $\sigma$ ) as shown by Clairon et al (1). The last published estimations of the primary standards of the National Institute of Standards and Technology (NIST) at Boulder, in the United States of America, and the Physikalisch Technisches Bundesanstalt (PTB) at Braunschweig, in Germany, are at the level of 1 to a few parts in  $10^{-14}$  (1  $\sigma$ ). These results show that the frequency stability of the phase measurements made by the time transfer technique chosen to compare these primary standards should stay within  $10^{-15}$  or better.

It is anticipated that the BNM-LPTF at Paris Observatory (OP), France, will in a near future be equipped with an earth station to achieve Two-Way Satellite Time and Frequency Transfer (TWSTFT). Even if the subnanosecond accuracy of clock comparisons is not reached yet by TWSTFT, this technique appears the best available today for the short term frequency comparisons of the new generation of

primary atomic frequency standards. It has been shown by Hackman et al (2) that the stability of the earth stations equipment able to perform TWSTFT could reach the level of  $10^{-15}$  over one day. And it has been shown by Kirchner et al (3) that the stability could stay below 100 ps over 2 to 10 days, when the measurements are corrected by modelling the temperature and humidity (or water vapour) behaviour of the transmit and receive delays of the earth station. But these measurements were made locally, which means without taking into account the path non-reciprocity due to the influence of the ionosphere along the path between two remote stations.

The primary frequency standards are not operated continuously at the BNM-LPTF. They are used to calibrate the frequency output of an H-maser which gives the reference signals used for the time transfer. The purpose of this paper is to give some estimations of the influence of the ionosphere on two future comparisons of primary standards by TWSTFT, OP-NIST and OP-PTB, that should ideally be performed twice a day or more during 7 to 10 days.

### THE IONOSPHERE

The ultraviolet radiation of the Sun induces the ionization of the highest layers of the atmosphere, generating an area only made up of free electrons and ions. That area, located at an altitude between 70 and 1000 km, is called the ionosphere, and the consequence of the ionization is that all radio signals are refracted. It is not the purpose of this paper to provide an extensive theoretical study on the ionosphere: this can be found for example in Jespersen (5). There is no model, even accurate at the ns level, available for the computation of the ionospheric delays due to the refraction of the signals, for any techniques of time and frequency transfer by satellite: the ionosphere is a medium which is dispersive, unhomogeneous and unisotropic. But the extensive use of the Global Positioning System (GPS) has provided many studies that describe in a qualitative way the behaviour of the ionosphere.

One will consider in the following that the path extension of the signal propagation due to the ionosphere is given by:

$$\Delta_r(f) = 40.3 \text{ TEC} / f^2 \quad (1)$$

where  $f$  is the carrier frequency of the signal in Hz, and TEC means 'Total Electron Content' along the path in  $e/m^3$  ('e' for 'electron'). In the ionosphere surrounding

the earth, the TEC may vary between  $10^{16}$  to  $10^{19}$  e/m<sup>3</sup>, depending typically on the following parameters.

- The solar activity, by a factor of 1 to 4. The period of the solar activity is around 11 years, the next minimum is scheduled for 1997.
- The seasons of the year, by a factor 1 (July) to 4 (November), as given by Wells et al (6).
- The moment of the day, by a factor of 1 (night) to 5 (day), the maximum being reached at 2 p.m. local time.
- The elevation (El) of the satellite, by a factor 1 (El = 90°) to 3 (El = 0°).
- The latitude of the station: the TEC values are 4 times higher at the equator and the poles than at mean latitudes.

Leick (7) shows that the formula given in (1) is accurate at the level of 5 cm at the GPS frequencies, 1.2 and 1.5 Ghz. Let us assume that the formula (1) remains accurate at the subcentimetric level for frequencies ten times higher in the Ku-Band, between 11 and 14 GHz. Of course, to evaluate the time delay due to the propagation of the signal through the ionosphere,  $\Delta t$ , has to be divided by the speed of light  $c$ .

## INFLUENCE OF THE IONOSPHERE ON TWSTFT

TWSTFT is a well known time transfer technique which uses spread-spectrum pseudo-noise coded signals referred to the clocks outputs, and a telecommunication satellite transponder. The frequencies of the carriers are usually in the Ku-Band (11-14 Ghz) for occasional use on international links. The description of the technique and an estimation of the time transfer accuracy of TWSTFT can be found in Kirchner et al (4).

The signal propagation delay in the TWSTFT equation, or path delay, can be given by:

$$I = [ (\tau_1^u - \tau_1^d) - (\tau_2^u - \tau_2^d) ] / 2 \quad (2)$$

where 1 and 2 are for the remote ground stations, 'u' means 'uplink' and 'd' means 'downlink'. One consider here that the path non-reciprocity due to the earth rotation, the Sagnac effect, or to the satellite transponder are not included in this term. Moreover, as the effect of the troposphere on radio signals do not depends on the frequency, it is also considered that the tropospheric delays cancel at the ps level. In that way,  $I$  represents here the path delay only due to the ionosphere.

A regular TWSTFT measurement session lasts usually 2 min, with one measurement per second. The ionosphere is varying with time slowly enough to be considered constant over one TWSTFT session. It is also generally assumed that the path non-reciprocity due to the influence of the ionosphere can be neglected for TWSTFT using Ku-Band carriers (11-14 Ghz) within an uncertainty of 100 ps (1  $\sigma$ ). But for frequency comparisons between remote primary standards, this value is the upper limit of the requested daily global stability. What are the fluctuations of the ionospheric

delays, from one measurement session to the other, that could be expected along the paths OP-NIST or OP-PTB? The characteristics of the links between these remote stations lead to the following simplifying assumptions.

- All the stations are located at mean latitudes.
- The uplink carrier frequency is 14 Ghz for all sites.
- The downlink carrier frequency is 12 Ghz over Europe, and 11 Ghz over the United States of America.
- The elevation of the satellite is very low. This is a consequence of the choice of one transponder only for all the international clock comparisons.

The only accurate way to evaluate the TEC values is by measuring them along the path of the radio waves simultaneously carried by two frequencies, different but close. This is not the case for TWSTFT. In order to investigate some of the limits of the ionospheric delays, all the computations are going to be made for 3 different periods of solar activity. The proposed average values of the TEC in July are given in the following table, where 'High', 'Mean', and 'Low' are applied to the solar activity.

TABLE 1 - Approximate average TEC values in July for TWSTFT between OP, NIST, and PTB.

TEC (July) 10 <sup>17</sup> e/m <sup>3</sup>	High	Mean	Low
Day	6.25	3.125	1.563
Night	1.25	0.625	0.313

## A TRANSATLANTIC LINK: OP-NIST

There is a time lag of 8 hours between Paris and Boulder. Two major effects on the path delay are related to the ionosphere: the difference of the downlink frequencies, and the TEC values that could only by chance be equal over Europe and the United States of America. Given the carrier frequencies of this particular link, and considering the formula (1), the terms of equation (2) are:

$$\begin{aligned} \tau_1^u &= 40.3 ( \text{TEC}_1 / f_1^{u2} ) / c \\ \tau_1^d &= 40.3 ( \text{TEC}_1 / f_1^{d2} ) / c \\ \tau_2^u &= 40.3 ( \text{TEC}_2 / f_2^{u2} ) / c \\ \tau_2^d &= 40.3 ( \text{TEC}_2 / f_2^{d2} ) / c \end{aligned}$$

where 1 and 2 are for the remote ground stations, 'u' means 'uplink' and 'd' means 'downlink'. It gives easily:

$$I = 40.3 ( \text{TEC}_1 - \text{TEC}_2 ) / 2c f^{u2} + 40.3 ( \text{TEC}_2 / f_2^{d2} - \text{TEC}_1 / f_1^{d2} ) / 2c \quad (3)$$

Incidentally, one can see that the ionospheric delays cancel each time the TEC values for both stations are in the following ratio:

$$\text{TEC}_1 / \text{TEC}_2 = ( f_1^{d2} / f_2^{d2} ) ( f_2^{d2} - f^{u2} ) / ( f_1^{d2} - f^{u2} )$$

### Computation of the Ionospheric Delays

For the case where the moment of the day is similar in both stations, for instance daytime at OP and at NIST, one propose the very simplifying assumption that the TEC value is the same for both locations. It is equivalent to consider only the effect of the different downlink frequencies. The results are given in the following tables, in ps, where 'High', 'Mean', and 'Low' are applied to the solar activity.

**TABLE 2 - Ionospheric delays in TWSTFT between OP(day) and NIST (day).**

OP - NIST			
Day / Day	High	Mean	Low
November	222	111	56
July	56	28	14

**TABLE 3 - Ionospheric delays in TWSTFT between OP(night) and NIST (night).**

OP - NIST			
Night/Night	High	Mean	Low
November	44	22	11
July	12	6	3

**TABLE 4 - Ionospheric delays in TWSTFT between OP(day) and NIST (night).**

OP - NIST			
Day / Night	High	Mean	Low
November	204	102	51
July	50	25	13

**TABLE 5 - Ionospheric delays in TWSTFT between OP(night) and NIST (day).**

OP - NIST			
Night / Day	High	Mean	Low
November	470	235	116
July	118	59	30

The table 5 shows that, during a period of maximum solar activity, the influence of the ionospheric delays in the phase measurements of a TWSTFT session taking place in November could reach 470 ps. This slightly exceeds the limits of  $\pm 300$  ps which are generally neglected for the estimation of the time transfer accuracy, but it remains consistent, considering the numerous assumptions that were made.

### Accuracy of the Frequency Comparison

For the accuracy estimation of the frequency comparison of primary standards, only the variations of the usually neglected delays have to be considered. It can be seen on the previous tables that two TWSTFT sessions taking

place in November, separated by 12 h (considering only the option day/night, followed by the option night/day) could exhibit a fluctuation in the ionospheric delays of 65 to 266 ps, depending on the period of solar activity. During the more common periods of mean solar activity, the variations in the ionospheric delays could stay within 19 to 213 ps, considering all the possible options presented here.

The problem is not the same for different periods of measurement: primary frequency standards do not run continuously, but continuous measurements could be carried out over long periods by using the H-masers only. From a TWSTFT measurement made in July to a measurement made in November, the variations of the ionospheric delays could stay within 8 to 352 ps, considering all the above mentioned options. This represents less than  $4 \cdot 10^{-17}$  over 100 d. But over such a long period, other instabilities could also occur in the equipment of the earth stations.

In many cases, the instabilities in the TWSTFT measurements due to the ionospheric delays could nevertheless exceed the accuracy limit of  $10^{-15}$ , which means for instance 100 ps over 1 d. One can conclude that the comparison between the primary standards of OP and NIST, that should ideally be performed twice a day during periods of 7 to 10 days, should be scheduled preferably during summertime, when the daily variations of the ionospheric delays stay under 70 ps anytime. Or it could also be scheduled in winter, but preferably during night time in both stations, to keep the variations of the ionospheric delays at the level of a few ps.

### A CONTINENTAL LINK: OP-PTB

These two european stations are located about 750 km apart. In that case, the uplink (14 Ghz) and downlink (12 Ghz) carrier frequencies are considered here to be the same for both remote stations. The major effect on the ionospheric delays is in the difference of the TEC values due to the different paths of the radio signal to the satellite, as seen by the remote stations. In a similar way as in the previous chapter, it is very easy to obtain:

$$I = 40.3 (1/f^{n2} - 1/f^{d2}) (TEC_1 - TEC_2) / 2c \quad (4)$$

### Computation of the Ionospheric Delays

When considering excellent atmospheric conditions, like minimum solar activity and night time observations made in July, for remote stations not more than 2000 km apart, the horizontal gradient of the vertical TEC, 'VTEC' in the following, is about  $10^{15}$  e/m<sup>3</sup> every 100 km, as given by Blewitt (8). Over the OP-PTB link, the change in the VTEC is then about  $7.5 \cdot 10^{15}$  e/m<sup>3</sup>. That VTEC value has to be expressed as a regular TEC value by using a transfer fonction, which could be of various shapes depending on the chosen parameters. The transfer fonction mentioned in a CNES report (9) (CNES means 'Centre national d'études spatiales', the French Space Agency) depends on the elevation of the

satellite, which is roughly  $10^\circ$  for OP or PTB, and on the latitude of the remote stations, about  $45^\circ$  in that case. The multiplicative factor given by that transfer function is then a little less than 4. That figure is used for the simplified computation of the TEC difference between OP and PTB, which, for low solar activity and night time observations made in July, is given by:

$$\text{TEC}_1 - \text{TEC}_2 \approx 3.0 \cdot 10^{16} \text{ e/m}^3$$

It then becomes very easy to use the equation (4) to compute the ionospheric delays of a TWSTFT between OP and PTB, considering that there could only be daytime, or night time for both stations simultaneously. The results of the computations are given in the following tables in ps, with 'High', 'Mean', and 'Low' applied as before to the solar activity.

**TABLE 6 - Ionospheric delays for TWSTFT sessions at night between OP and PTB.**

OP-PTB			
Night	High	Mean	Low
November	60	30	15
July	15	7	4

**TABLE 7 - Ionospheric delays for TWSTFT sessions in the day between OP and PTB.**

OP-PTB			
Day	High	Mean	Low
November	297	148	74
July	74	37	19

Again, one can see in table 7 that the usually neglected delays could reach the limits of  $\pm 300$  ps for some particular configuration of the link between those remote stations.

### Accuracy of the Frequency Comparison

Over such a short baseline, the situation appears much simpler than in the previous case. The highest variations of the ionospheric delays would come from the fact that the two daily sessions of TWSTFT could occur one in the day and the other one at night. In that case, the tables 6 and 7 show that the daily variations could stay within 15 to 237 ps depending on the solar activity and the season. During the more common periods of mean solar activity, the variations in the ionospheric delays could stay within 30 to 118 ps.

But in fact, two daily sessions between OP and PTB could be scheduled either both at night in winter, and the daily variations in the ionospheric delays should remain within a few ps, or anytime during summertime, where the maximum daily variation that could occur between one session in the day and one session at night stays within 60 ps.

### CONCLUSION

All the computations made in this paper are very simple, with the use of numerous simplifying assumptions, because the estimation of a TEC value along the path of a radio signal is not obvious at the subnanosecond level. Nevertheless, for the frequency comparison of primary standards, it appears that the date of the TWSTFT sessions, that should ideally be performed twice a day or more during 7 to 10 days, should be chosen carefully in order to maintain the variations of the usually neglected ionospheric delays widely under 100 ps from one session to the other.

The accuracy of the frequency comparison between the primary standards of the BNM-LPTF and the NIST depends as well on the stability of the satellite transponder, for which no data seem to be available. This is not the case for a frequency comparison between the primary standards of the BNM-LPTF and the PTB, for which, as for the other link anyway, the stability of the earth stations equipment remains the major issue.

It is anticipated that the new generation of primary standards, based on a design of a cold atoms Cesium fountain, could in a very near future reach accuracies at the level of a few parts in  $10^{-16}$ . At this level, any variations in the delays of the time transfer technique used to compare remote standards should remain within  $10^{-16}$ , or for instance 10 ps over a 1 d period. This means that even for Ku-Band carrier frequencies, the ionospheric delays should be measured along the path of the radio signal.

### ACKNOWLEDGEMENTS

The author is grateful to Dr Bertrand Juompan (formerly at the BNM-LPTF) for extensive discussions about the ionosphere.

### REFERENCES

1. Clairon A., Ghezali S., Santarelli G., Laurent Ph., Lea S. N., Bahoura M., Simon E., Weyers S. and Szymaniec K., 1995, "Preliminary Accuracy Evaluation of a Cesium Fountain Frequency Standard", Symposium on Frequency Standards and Metrology, Woodshole, MA, October 15-19.
2. Hackman C., Jefferts S. R. and Parker T. E., 1995, "Common-Clock Two-Way Satellite Time Transfer Experiments", Proceedings of the IEEE International Frequency Control Symposium, San Francisco, CA, USA.
3. Kirchner D., Ressler H. and Robnik R., 1995, "An Automated Signal Delay Monitoring System for a Two-Way Satellite Time Transfer Terminal", Proceedings of the 9th European Forum on Time and Frequency, Besançon, France.

4. Kirchner D., Ressler H., Grudler P., Baumont F., Veillet Ch., Lewandowski W., Hanson W., Klepczynski W. and Urich P., 1993, "Comparison of GPS Common-view and Two-way Satellite Time Transfer Over a Baseline of 800 km", *Metrologia*, **30**, 183-192.

5. Jespersen J., 1989, "Impact of Atmospheric Non-reciprocity on Satellite Two-way Time Transfer", 43rd Annual Symposium on Frequency Control.

6. Wells D. (Prepared under the leadership of), 1986, "Guide to GPS Positioning", Canadian GPS Associates, Fredericton, New Brunswick, Canada.

7. Leick A., 1990, "GPS Satellite Surveying", John Wiley & Sons.

8. Blewitt G., 1989, "Carrier Phase Ambiguity Resolution for the Global Positioning System Applied to Geodetic Baselines up to 2000 km", *Journal of Geophysical Research*, Vol. 94, n° B8, pp. 10187-10203.

9. CNES, 1993, "Modélisation du contenu ionosphérique à l'aide d'un calibrateur GPS", Rapport CNES/CLS/DT/NT/93.025 du 24 février 1993.

## GPS DISCIPLINED RUBIDIUM OSCILLATOR

Dr Cosmo Little, Mr Clive Green

Quartzlock(UK)Ltd - England

Phase 1 Preliminary report 16/11/95

### 1/- INTRODUCTION

Quartzlock Instruments have a number of frequency standards in their product range, including a GPS controlled quartz crystal oscillator. This frequency standard has very good long term (averaged for greater than 24 hours) performance, as the oscillator is phase locked to GPS time. The medium term performance is poor (averaging times between 1 second and 24 hours) as the GPS signals are deliberately degraded by the US dept of defence (Selective Availability, or S.A.), and the control loop in the GPS receiver has a short time constant (estimated to be about 1 second). Also the frequency standard relies on continuous tracking of satellites. If the signal fails for any reason, the control of the oscillator is broken, and it free runs with a large frequency offset. This is undesirable in a highly stable frequency/time source which may be expected to run for years without slipping cycles (incurring a time offset which is never corrected.)

### 2/- PRODUCT PROPOSAL

The proposed product is an addition to the existing GPS frequency standard (although housed in the same case) which uses a rubidium oscillator as its main frequency standard. A rubidium oscillator has very good short and medium term stability. The maximum drift rate is expected to be less than 4 parts in  $10^{11}$  per month. There may be a frequency offset of up to 1 part in  $10^{10}$  after initial calibration.

By comparing the phase difference between the GPS 10MHz output, and the 10MHz output from the rubidium oscillator, it is possible to correct the initial frequency offset and remove the drift. By using sufficiently long time constants in the control loop, it is possible to remove the effects of S.A. completely, and frequency accuracies equivalent

to GPS time, which is itself locked to UTC within +/- 100nS, may be obtained.

It is vital that the system is tolerant of GPS signal failures. The rubidium is sufficiently stable to free run for hours or days provided that the control voltage is held at the point where the GPS control was interrupted. This requirement will be discussed in much greater detail later in this report.

A further output from the GPS receiver is a time mark that has a known relationship to, or may be coincident with, GPS time. This time mark will have time jitter on it that is a function of SA and other random errors that occur from the uncertainties in the satellite signals themselves, and the measurement process in the receiver. It may also have a systematic error that may not average to zero, however long the averaging time. A greatly improved time mark can be made available by generating a mark (typically one edge of a 1pps pulse) from the rubidium signal, and steering it into coincidence with the time mark from the GPS receiver, again using a very long time constant loop. By using a sufficiently long time constant, errors that appear systematic in the medium term may in fact average to zero in the long term. The time mark from the rubidium will share the exceptional short and medium term stability of the 10MHz frequency output. Time jitter of less than 1 nS can be expected.

A further use of the product is as a frequency comparator. If an auxilliary 10 MHz input is provided, with a small amount of extra hardware, long term comparisons and measurements may be made between the controlled rubidium and an external frequency standard. The resolution will depend upon the hardware design of the product, and will be discussed later in this report.

It is highly desirable that the product has a well designed user interface that informs the user as to the expected time and frequency accuracy of the outputs. These should be in the form of  
 (1) frequency :- current fractional frequency offset and standard deviation  
 (2) time current time offset and standard deviation (including systematic errors)

(3)comparator current fractional frequency offset and standard deviation

Information should also be provided as to GPS status.

### 3/- SCOPE

This report will consider the design of the hardware down to block diagram level. The control algorithm will be considered in detail, however, and performance analysis of the selected method will be carried out on simulated data.

### 4/- BACKGROUND

#### 4.1 Definitions

Before starting a discussion of time and frequency, some definitions are necessary:

**a) TIME** . Time can be measured as time difference from some reference point, or as phase or time difference between two periodic signals. For example, a pulse edge can be 250nS ahead of GPS time (the reference). Equivalently, we could say that one signal at 1MHz was  $\pi/2$  ahead of another signal at 1MHz. This would also be a time difference of 250nS. Note that here we have converted a phase difference to a time difference by dividing by angular frequency. Phase is always measured as a difference. The phase of a 1 MHz signal is actually advancing at 6.28 million radians per second, not in itself very useful.

Sometimes the phase is compared to that of an imaginary perfect signal at a nominal rate. For example, if we say that an oscillator had a discrete sinusoidal phase modulation of 1 radian peak, we would actually mean that the phase difference between the real oscillator and an imaginary perfect oscillator of the same mean frequency was 1 radian peak.

**b) FREQUENCY**. Frequency is rate of change of phase, averaged over any interval we like. In order to measure frequency, we need to make a time difference measurement. If a frequency counter measures the time taken for 10 complete cycles of a periodic waveform to be 100uS, we can say that

the frequency is 100KHz, or 200000Pi radians per second.

We often refer to frequency as a fraction of a nominal frequency. For example a Rubidium oscillator may have a fractional frequency error of 1 part in  $10^9$ . This is a correct definition as it stands. If we know that the nominal frequency is 10MHz, then we know that the absolute frequency error is 0.01Hz.

Just as frequency is rate of change of phase, frequency difference is equal to rate of change of phase difference. For example if the phase difference between two 10MHz oscillators was changing at a rate of Pi radians per second, then the frequency difference between them is 0.5 Hz. The fractional frequency difference is  $5 \cdot 10^{-9}$ . Obviously when calculating fractional frequencies consistent units must be used, either radians per second or Hz.

Frequency difference is also equal to rate of change of time difference multiplied by nominal frequency. If two 10pps edges have a relative time drift of 10nS per second, then the frequency difference is  $10^{-7}$  Hz. The fractional frequency difference is therefore 1 part in  $10^8$ . Thus fractional frequency difference is equal to rate of change of time difference, without any need for a reference frequency. Thus an oscillator which has a fractional frequency error of 1 part in  $10^{12}$ , will gain or lose  $10^{-12}$  seconds per second, or 1 second in 31710 years.

**c) FREQUENCY DRIFT**. Frequency drift is the rate of change of frequency, or the second time differential of phase. If an oscillator has a constant frequency drift, then its time behaviour is exactly specified for all times in the future by a cubic equation:

where a is the initial time offset, b is the initial fractional frequency offset, and c the drift rate.

If the frequency drift is a function of time, then the equation becomes:

Obviously we could consider higher derivatives of time and extend the above expressions as far as we like.

### 5/- PHASE ONE DETAILED REPORT

## 5.1 Hardware

The hardware design of the product may be summarised in figure 1.

First, it should be noted that the frequency standard part of the instrument is generally separated from the time part. This has been done for two reasons:-

a) Some customers may only wish to purchase the frequency standard instrument. The time standard may be an option.

b) It is felt to be important that the standard frequency output should not be affected by the time correction loop, i.e. it would not be permissible to slew the time pulse by deliberate offsetting of the output frequency. The frequency output may be distributed to many laboratories on a site, and users may not be able to refer to the display on the instrument to see if the frequency is currently accurate.

The frequency system will be described first.

The 10MHz output from the GPS receiver is compared with the 10MHz from the Rb oscillator. These signals will be digital signals with unknown mark/space ratios. The phase detector, shown in detail in fig 2, is a quadrature phase detector operating at 2.5MHz. The phase detector will have an unambiguous range from  $-Pi$  to  $+Pi$ , obtained by dividing the complete range into 8 sectors, and only operating the I and Q phase detectors over their most linear regions. The processor will synthesise a linear phase detector by weighting and summing the I and Q outputs appropriately in each sector. Unlimited phase range will be achieved by adding or subtracting  $2Pi$  (400ns) each time the phase vector crosses the -I axis. (see figure)

Each phase detector (I and Q) will have a 12 bit AtoD channel. A third channel will be used for a zero phase reference.

Due to eventual very long loop time constant, the initial sampling rate will be 10 S/s on each channel. It is proposed that samples are decimated in the processor by a factor of 100, so the final input to the analysis algorithm will be 1 phase value every 10 seconds.

The time resolution of the phase detector will be about 50pS. This will be adequate for the comparison between the Rb oscillator and the GPS signal as the GPS signal will be very noisy, ensuring that quantisation noise is well hidden. However the comparison channel may be used to compare two very clean sources (one of which is the Rb). A fractional frequency difference of 1 in  $10^{12}$  will mean that a quantisation step will only be crossed every 50 seconds. In this case performance can be improved by dithering the convertor, i.e. by adding a small amount of random noise. This could be derived from a low resolution DAC updated with random numbers by the processor.

After analysis of the frequency error, the processor will update the Rb control DAC on a periodic basis (estimated to be about once every 100 seconds). The DAC is a 16 bit convertor, which will give control resolution of about 30 parts in  $10^{15}$ . This assumes that the total pulling range of the Rb osc C field adjustment is  $\pm 1$  part in  $10^9$ .

The controlling microprocessor will communicate with the GPS receiver over a serial link. The GPS receiver has available a number of useful parameters such as satellites tracked, and expected GDOP, which may be of use in adjusting the control algorithm adaptively.

The user interface will comprise a keyboard and LCD dot matrix alphanumeric display.

The time system can be less well defined at the moment due to lack of information regarding the 1 pps available from the GPS receiver. It may be that there is a fixed time relationship between the 10MHz output and the 1pps output. If this is the case, then the GPS 1pps may be used to reset the divide by  $10^7$  circuit, which will align the Rb 1pps to within 100ns of the GPS 1pps. The final alignment can be made using a phase detector of exactly the same type as used in the frequency system. A possible problem with this system is that if the GPS 1pps happens to have jittered more than 100nS when it is used for the divider reset, an undetected time error of 100nS will result.

A more certain and reliable system is to operate the time phase detector at 1pps directly. This avoids all ambiguities. The design of such a phase detector has not been considered at the present time, but it is possible that a tri state phase frequency comparator type would be suitable. Alternatively, a frequency counter technique using a high frequency clock, unsynchronised to the Rb

to avoid quantisation effects that would not average to zero. With a 100MHz clock, the time resolution will only be  $\pm 10\text{nS}$ , but this will be improved by averaging. Systematic errors are likely to be up to  $100\text{nS}$ . The main contribution will be uncertainty in user position. It is important that the GPS receiver should average its position fix over a long period of time to reduce this systematic error.

The recommended method for slewing the Rb 1pps, is to vary the phase of the 10MHz input to the divider by using an I/Q modulator. The principle of this is shown in fig 3. The circuit is an image rejection mixer using a low frequency quadrature signal generated by the processor. An exact phase relationship exists between the output at 10 MHz, and the processor quadrature signal, which may be DC. If the processor rotates the output vector (represented by the sine and cosine of an angle held as a variable by the processor) by  $2\text{Pi}$  radians, then the phase of the output 10MHz will also change by  $2\text{Pi}$  radians. This will slew the 1pps by  $100\text{nS}$ . The resolution is a non linear function of the DAC resolution, but 8 bit bipolar DACs should give better than  $0.1\text{nS}$  time resolution. I/Q modulators can be obtained as sealed modules from several manufacturers.

## 5.2 Software

It is in the software aspects of this project that most progress has been made, and this largely concerns the control algorithms for the frequency control of the Rb oscillator, and the phase control of the 1pps time output.

Initially it was thought that a simple phase lock loop would be adequate. This would be a type 2 second order loop (i.e. with an integrator in the loop filter) with a zero to give suitable phase margins for optimum dynamic performance. However one problem with a phase lock loop is that it must reduce the initial phase error to zero by changing the frequency of the VCO. With the very long loop time constant necessary to remove the effect of the SA, the eventual settling of the loop could take several days. It is also difficult to extract measures of performance from the loop, for example it is difficult to estimate the current frequency error of the VCO. It was felt that a frequency control loop would settle quicker. For a frequency standard we do not mind operating with a fixed phase offset, and there is no need to reduce this to zero. The first idea to be tried with simulated noisy data, was a quadratic least

squares fit on a block of data. This is a standard method for extracting phase offset, frequency offset, and frequency drift from phase difference information. Having extracted the offset frequency, we can then make a correction to the Rb oscillator to remove the offset. If the Rb control constant was known exactly, there would be no under or overshoot. The problem with this method is that we do not know how large to make the block of data that we analyse. The reliability of the fit is given by the correlation coefficient, and ideally this should be monitored on a continuous basis. What is required is a continuous least squares process. This is of course, a Kalman filter, and this was the eventual method selected for implementing the control algorithm.

The Kalman filter will be briefly described in general in a (hopefully) simple way, and then the specific implementation for our problem will be described in more detail.

A block diagram of a Kalman filter is given in figure 4. It is basically a recursive estimation, based on noisy measurements, of the future "state" of a system. The system is defined as a "state vector" and a "state transition matrix". The system in our case would be the Rb oscillator that we wish to predict, and the state vector would contain the phase offset, frequency offset, and frequency drift variables. The "state transition matrix" defines the differential relationship that exists between the state variables over one time increment. The concept of a system driven by noise processes is important here. If our Rb had absolutely constant drift, its output phase would be known for all time once the initial drift, frequency offset and phase offset had been determined. Data gathered a year ago would have as much validity as data an hour old. If the Kalman filter is given this model of the Rb, the results are identical to the least squares fit of all the data. (See document KALMAN5.MCD for comparison with least squares fit) Of course the quadratic least squares fit assumes that the Rb can be modelled by three constants.

A more realistic physical model would allow the drift to vary. If this varied in a deterministic way, we should add a further term to the state vector to reflect this deterministic process. However if the variation was random, we can tell the Kalman filter that this is so. Note that the filter is only optimum for white gaussian noise processes. However in our case we can model the noise of the Rb oscillator more accurately by adding white gaussian noise to each term in the state vector. If we add some uncorrelated noise to each term in

the state vector, we end up with white phase noise, white FM noise, and random walk FM noise due to the single and double integration in the model. This is shown in figure 5.

The measurements are also assumed to be contaminated with gaussian white noise. In our case we only have one measurement, that is phase offset. We do not know that the main contributor to measurement noise, SA, is either white or gaussian. However this is a limitation of the simple Kalman filter that we intend to use. If we are sure of the characteristics of the measurement noise, we can include this knowledge by adding more terms to the state vector. We are then essentially including the known aspects of the measurement in the system model. This much more complicated procedure will not be attempted until real data is available, and has been analysed, and only if the basic filter performs inadequately.

As well as the state vector, the Kalman filter maintains a matrix that gives the current variances (mean square error) of the quantities in the state vector. These give us current estimates of the likely errors in the state vector, in our case variances of phase offset, frequency offset, and frequency drift. These will be very useful for display to the user. They also have another use, which will be demonstrated later. In effect they control the "bandwidth" of the filter. As more data comes in, the variances decrease, and the filter gives more weight to the current estimate (which represents the complete history of the data), and less to the current measurement. The measurement variance, which we have to tell the filter, also affects the "bandwidth". If we tell the filter that the measurement is noisy, it reduces the bandwidth.

So far we have considered the Kalman filter as a device for analysing the incoming data in an optimum way. However we need to control the Rb oscillator, and reduce the frequency offset to zero. An elementary method would be to write periodic corrections to the Rb control DAC, and wait for the Kalman filter to track out the resulting discontinuity in the measurements. However there is a much better way. If we adjust the frequency offset term in the state vector at the same time that we correct the Rb, the filter will ignore the correction, and no extra settling time will be required. In effect we are defining the model of the Rb to have a frequency discontinuity at a particular time, and provided the real Rb has that discontinuity, the Kalman filter will see no difference between the model, and what the measurements are telling it about the real

system. This behaviour may be seen simulated in document KALMAN7.MCD. Note especially that there is no change in the variance of the frequency estimate.

Using this technique, we can correct the Rb as often as we like. However if we are uncertain as to the exact value of the control constant, then the correction will undershoot or overshoot the model. In this situation there is little point in updating more often than about every 1000 seconds (maybe initially every 100 seconds so the customer can see something happening!), as the filter will need some settling time. The control constant can be established by the stand alone instrument during a calibration procedure, described later.

Another trick that can be useful is if we know that there is a measurement discontinuity, but we do not know how large it is. An example would be if the GPS signal disappears for any reason. When satellites were reacquired, there could be a phase discontinuity between the GPS 10MHz and the Rb 10MHz. Although we cannot tell the filter the amount or direction of the discontinuity, we can tell it that its current estimate of phase is completely unreliable. We do this by adding a large number to the appropriate term of the error covariance matrix. The filter then gives maximum weight to the measurements to reacquire the phase as quickly as possible, however as it thinks its frequency is still accurate, it does not give excessive weight to the rate of change of phase measurement, and the frequency covariance hardly rises. (see document KALMAN9.MCD) This trick would be quite impossible with a phase lock loop.

Another even more exotic trick is to control the Rb oscillator at a deliberate frequency offset from GPS time. In this case there will be a continuous phase ramp at the phase detector. As our phase detector has an infinite phase range, this need not matter. We now need to change the state model of the Rb so the filter expects the phase to be a ramp. This can be done by adding the expected incremental phase offset into the state vector every iteration of the filter. For example, if we wish to offset 1 part in  $10^9$  low, we would subtract  $10^{-8}$  from the state vector, i.e. 10nS every 10 seconds. If the numbers coming out of the phase detector get too large, we can make a known phase correction in the same manner as a frequency correction was made (except this is only a "housekeeping" adjustment, and does not involve any change to the Rb oscillator). Both these effects can be seen in document

**KALMAN13.MCD.** This feature, assuming any customer was interested, gives us a frequency synthesiser of arbitrary resolution, with no degradation in stability or spurious output levels from the normal instrument. Unfortunately we are limited in range by the pulling range of the Rb.

**Calibration.** The instrument will work best, with minimum settling times, if the rubidium control constant is known exactly. Each Rb module will probably vary, so some calibration procedure is necessary. The instrument can achieve this by setting various DAC outputs, and measuring the frequency offset against GPS time. 10 readings should give an adequate table of values of fractional frequency offset against DAC value. A polynomial fit can then be made to the data and the table stored in nonvolatile memory. If each reading was allowed 5000 seconds to settle, the calibration would be an overnight process. Obviously the normal frequency correction procedure can be used to minimise settling time in changing from 1 point to the next.

**Prediction.** The Kalman filter can, of course, predict ahead if measurement data fails. In this case both the state vector and the error covariance matrix will be updated. The previously estimated value of drift will update the frequency offset automatically. Frequency corrections can be made in the usual way. The error covariances will rise to reflect the lower confidence in the predictions as time passes. When measurements resume, the filter will automatically recover and the error covariances will start to fall. Thus the user is always aware of the reliability of the frequency output. The document KALMAN12 shows the filter predicting for 5000 seconds half way through the data. If an unknown phase step is expected on resumption of measurements, then the phase variance should be augmented as previously described.

**Calculations.** Although the formulation of the Kalman filter involves matrix algebra, in our example the equations reduce to quite simple scalar equations. 21 equations are involved, using only normal arithmetical operations. Floating point representation will be necessary because of the large dynamic range of the quantities. As floating point must be used anyway, it is worth keeping fundamental units of seconds throughout, as then all the quantities are ready for display. It is more usual to display RMS error estimates (otherwise called standard deviations) rather than variances,

so a square root function will be required. The Kalman filter algorithm is given in full in Appendix A.

### 5.3 Further work

Some constants in the filter, such as the model noise variances and the measurement noise variance, can only be determined by using real data. Thus phase 2 of the project will consist of acquiring several days of continuous phase measurements between a real Rb oscillator and the GPS receiver to be used. The data will be analysed in various ways, including conversion to the frequency domain to examine for coloured noise. Least squares fits will also be used to extract the Rb drift and frequency offset. The Kalman filter will then be used to process the data, with various guess values for the variances, and the values manually adjusted until reasonable agreement with the least squares fit is obtained. Of course it will not be possible to tell the difference between measurement noise and Rb noise with only one set of phase comparisons. Another method of determining the model noise sources (figure 5) is by calculation from known Allan variances of a typical Rb. It might also be worth trying to compare two similar Rb oscillators to try and measure the model noise variances. In this case the measurement noise could be set at zero, and the Kalman filter should converge to give direct measurement of the combined Rb noise variances.

## 6/- CONCLUSIONS

It is hoped that this report has set the project on a firm basis of theory. The design of the frequency part of the instrument is now well defined, and unless SA has nasty surprises in store, should converge to within 1 part in  $10^{13}$  of GPS time over 100000 seconds (24 hours). After phase 2 of the project, when real data is available, it will be possible to correctly predict the performance using the MathsCad simulations before the prototype is constructed.

The time part of the instrument needs more work, especially the estimation of systematic errors. These are very much a function of the GPS receiver, such as unknown delays in the receiver RF path, and certainty of user position. Caution should be exercised as to claims of time accuracy at the present stage of development.

**FREQUENCY JUMPS ON BVA & OTHER PRECISION QUARTZ CRYSTAL RESONATORS  
AND  
BURST-NOISE ON OVERTONE MODE HIGH-FREQUENCY QUARTZ CRYSTAL RESONATORS.**

**Kanak K. Tuladhar and Gérald Jenni  
OSCILLOQUARTZ SA, CH - 2002 NEUCHATEL**

**ABSTRACT**

Frequency jumps have been observed on BVA and other precision quartz crystal resonators/oscillators during prolonged and continuous operation. To date frequency-jump phenomena have not been understood well, as such these jumps have been designated as micro-jumps or abrupt frequency jumps irrespective to the nature of the jumps. For the first time in this paper, these frequency jumps will be classified into three distinct and separate categories and these are :

- (i) Intermittent Frequency Jumps
  - (ii) Discrete Frequency Jumps
- and
- (iii) Abrupt Frequency Jumps

To characterise and distinguish these frequency jumps, apart from continuous and prolonged operation, other resonator/oscillator parameters, such as g-sens, retrace, drive-level-dependence, bandbreaks (i.e. dips), phase-noise etc. have also to be measured.

Intermittent and discrete frequency jumps can be correlated to one or more resonator/oscillator parameters. However, to understand troublesome abrupt frequency jumps a conventional crystal resonator equivalent circuit is not adequate and a new resonator model is required.

Finally burst-noise observed on overtone-mode high-frequency quartz crystal resonators will be reviewed and explanations will be given to demonstrate why this noise process cannot be correlated to known crystal resonator parameter variations. Again, as in the case of abrupt frequency jump, a new resonator model is required to explain this burst - noise phenomenon.

## 1. INTRODUCTION

Frequency jumps have been observed on BVA(bva) and other precision quartz crystal resonators/oscillators during prolonged and continuous operation. Currently these frequency - jump phenomena have not been understood well as such these frequency jumps are either completely ignored (see figs. 1,9,10 of ref. 1) or

classified as micro-jumps (ref. 2) or abrupt frequency jumps (see fig. 10 of ref. 3, fig. 4 of ref. 5, and figs. 5,6 of ref. 6). Frequency jumps could be negative or positive and are generally in the region of  $1E-11$  to  $1E-09$  for the precision resonators/oscillators and greater than  $1E-09$  for the non-precision resonators.

Note that BVA resonators are generally fifth(third) overtone-mode AT(SC) - cut devices at 5.0 & 5.12MHz and packaged inside a special BVA housing, whereas bva resonators are SC-cut third overtone mode 10MHz and are packaged inside an well known HC - 40 / U holder.

Note also that the conventional precision AT(SC) - cut resonators are fifth(third) overtone mode devices and could be operating in the frequency range 4 to 10MHz, could be packaged inside a variety of holder styles, HC - 36 / U, HC -27 / U, HC - 40 / U, and near equivalent glass package.

## 2. DISCUSSION ON FREQUENCY JUMPS

Consider the frequency jump described by J. R. NORTON, fig. 10 of ref. 3, currently screening for these jumps at the resonator level is not possible as these jumps cannot be attributed to any known resonator parameters. However one interesting observation is that the net frequency change is near zero, once taken into account the established ageing rate, after the jump duration of about 6 hours, and the frequency jump is about  $1E-10$ .

Similarly considered the jump described by Koyama, et. al. , figs 5 & 6 of ref. 5, firstly this jump is of completely different nature and can be easily screened by proper screening procedure. Note that there is a discrete net frequency change of  $2.1E-10$ , (see Fig. 1), once taken into account the established ageing rate, after the jump duration of about 2 hours, and the frequency jump is about  $1.1E-09$ . This net change in frequency is due to Drive - Level - dependence(DLD) & HYSTERESIS effects.

Note also a new type of frequency jump, Fig. 2 observed recently on BVA fifth-overtone mode AT - cut resonator/oscillator during prolonged ageing

measurement. Here the nature of the jump is rather intermittent(also see ref. 4).

### 3. CLASSIFICATION AND PROPERTIES OF FREQUENCY JUMPS

As mentioned in sections 1 & 2 of this paper frequency jump phenomena have been bundled together and have been designated as micro-jumps or abrupt frequency jumps irrespective to the nature of the jumps. However in this section observed frequency jumps will be critically analysed and classified into three distinct and separate categories assuming resonators are driven in the linear region(i.e. no isochronism).

Note also that occasionally jump phenomena have been linked to Second - Level - Drive(SLD) effects, however these effects are well understood(ref. 8) and the resonators having SLD effects should be rejected as these resonators will suffer from severe hysteresis effects(ref. 9).

#### 3.1 Intermittent Frequency Jumps

Intermittent frequency jumps are usually caused due to passive and / or active component faults. Frequency jumps observed due to noisy active devices, regulators, transistors, varactors, etc. are usually of intermittent in nature and patently obvious. These can be cured by pre - screening active devices using known techniques.

Other possible causes for intermittent frequency jumps are,

i) micro - leak in the resonator, this in turn degrade the Q - factor of the resonator over time. When the Q - factor of the resonator falls below the value required to maintain the oscillation, intermittent frequency jump will be observed just before oscillation ceases, as the Q - factor will be in a critical region for the oscillation.(i.e. loop - phase shift might be zero, but loop - gain might be just around one.)

ii) crack on the mounting paste, this in turn also produces an intermittent frequency jump as the crack on the paste produces unstable bonding to the resonator. This can be easily screened by testing crystal resonators under random vibration as the resonators having cracks on the paste are likely to fail during this test.

iii) BVA(bva) sandwich assembly, occasionally minute problem could arise during sandwich assembly of BVA resonators, this in turn gives a kind of unstable gap - widths between the electrodeless resonator & associated condensers during prolonged and continuous operation and hence intermittent frequency jumps. Therefore it is also a good idea to screen BVA resonators under random vibration before assembling into oscillators.

Note that during intermittent frequency jump period, frequency change is rather spontaneous and phase - noise of the resonator / oscillator is stable however might have side - bands.

#### 3.2 Discrete Frequency Jumps

Discrete frequency jumps are caused due to non - optimum resonator design and / or oscillator circuit configuration. A number of possible causes for observing discrete frequency jumps are,

i) bandbreaks(i.e. activity - dips), although almost all precision oven controlled crystal oscillators are operated at their turn - over - temperature, if there happened to be activity - dips in the neighbourhood of the turn - over - temperature, it is patently obvious that there would be some sort of frequency jumps during prolonged and continuous operation. Note that even if the activity - dips are far from the operating temperature, due to subtle effects in the resonator / oscillator, jumps will be observed and the frequency offset due to the jump will depend on two parameters. These are severity of the activity - dips and the temperature difference between the oven temperature and temperature at which the activity - dip occurs.

These jumps are usually discrete and can be easily observed by measuring activity - dips over a wider temperature range at a resonator level or by measuring retrace or warm - up behaviour of the complete oscillator and can be easily cured by designing optimum activity - dip free resonator, although difficult but possible.

ii) g - sensitivity, 2g tip - over - test is usually used to determine g - sensitivity of the quartz crystal resonators. During such test for a non - optimum resonator a discrete frequency jump of any thing from 1E-10 to 1E-09 has been observed. A typical example is given in fig. 3. This shows at a certain acceleration vector the resonator has two stable frequency modes which are about 2E-10 apart, zones a) and b) of fig. 3. This is usually caused due to non - optimum diameter of the quartz crystal blank(Dblank) relative to the thickness of the quartz crystal blank, or non - optimum electrode diameter(Delectrode), or inadequate radius of curvature(R), i.e. to say energy trapping is not adequate to operate the resonator properly. This can again be cured simply by optimising, Dblank or Delectrode or R.

iii) drive - level - dependence(DLD), DLD of the quartz crystal resonator is well known, and suppose due to some reasons drive level to the resonator is changed during prolonged and continuous operation then there would be a frequency change too. This change is also discrete and can be cured by having stable and reliable circuit components which will maintain the drive level constant to the resonator. However it should be noted that, if there happened to be discontinuity in the drive - level versus activity curve, the root of the problem is likely to be due to activity - dips or SLD

effects or hysteresis of the quartz crystal resonator. Therefore before proceeding any further these parameters should be measured and investigated and the resonators should be rejected accordingly.

Above description clearly indicates that discrete jumps can be induced as and when necessary by changing appropriate resonator / oscillator operating conditions. The duration of the jump can also be defined and during the entire jump period, phase - noise of the resonator / oscillator is stable and can be defined properly.

Once again consider, fig. 1, i.e. the jump described by Koyama et. al., now it is clear that this kind of jump is in fact discrete as this is due to the combination of change in drive level and the activity - dips. Their recommendation to operate the resonator at a drive level above or below the value where there is no jump, is not satisfactory as these jumps tend to move over time due to stress relaxation and change in load capacitor value. The root of the problem must be found and cured by designing optimum resonator.

### 3.3 Abrupt Frequency Jumps

Now consider once again the jump described by J. Norton, i.e. fig. 10 of ref. 3, unlike intermittent and discrete jumps this jump is of completely different type. It has been attributed to thermal effect in ref. 7 without any foundation, if it is really due to thermal effect then one should be able to induced this kind of jump with right sort of thermal stimulus, this is not possible in real situation. It should also be noted that currently there is not any known resonator and oscillator stimuli which will generate abrupt frequency jumps. This phenomena is unpredictable, cannot be triggered ON / OFF as and when required and the abrupt frequency jump period cannot be defined as well. Furthermore phase - noise behaviour during the entire jump period is unstable and cannot be defined properly. Therefore it is believed that that commonly known simple Butterworth - Van Dyke oneport single resonance equivalent circuit of the quartz crystal resonator is probably not valid during the jump duration. A new equivalent circuit to describe abrupt frequency jump will be given in part II(ref. 11) and bulk of the paper(ref. 11) will be devoted to describe methods of minimising abrupt frequency jumps via optimum resonator design and improved processing techniques.

## 4. BURST - NOISE A REVIEW

Short - term frequency changes similar to the chaotic burst - noise or pop - corn noise observed in operational amplifier integrated circuit have also been observed on high - frequency overtone - mode crystal resonators/oscillators. This burst - noise phenomenon is distinct in character from the quartz crystal resonator  $1/f$  noise. Although a method of screening for the burst - noise has been developed and described in

ref. 10, the origin of this noise is not really known and not understood at all.

Currently a good high - frequency overtone - mode crystal resonator has been defined as the one which has ten or less burst during first two minutes after the oscillator switch on( see fig. 21 of ref. 10)

Attempts to correlate this burst - noise to the known resonator motional parameter variations proved unsuccessful(see figs. 17,18,19,20 of ref. 10). This indicates probably again the simple resonator equivalent circuit is not valid during the burst period as in the case of abrupt frequency jump duration.

## 5. FURTHER DISCUSSION ON BURST-NOISE

It has been speculated in ref. 10 that this burst - noise phenomenon could be related to *Second - Level - of - Drive*(SLD) effects(ref. 8) due to surface defects on the resonator blanks during polishing. However one of the authors(KKT) of this paper likes to point out following:

- i) Effects such as SLD will result in poor but stable phase - noise( i.e. no burst). Note also that during the burst mode, phase - noise of the resonator/oscillator cannot be defined.
- ii) Occasionally SLD effects are also designated as sleeping sickness and the resonators are labelled as sleepy - crystals. These resonators sometimes can be cured by driving the crystal resonator momentarily at a higher drive - level whereas the burst - noise phenomenon cannot be cured. Furthermore the observed bursts are persistent, unpredictable and cannot be triggered ON / OFF as & when required. Note also that the burst duration cannot be defined.

Above facts clearly demonstrate that the burst - noise on the high - frequency overtone - mode crystal resonator is in a way related to the abrupt - frequency - jump observed on the precision quartz crystal resonators. Therefore this is in fact a new fundamental problem on the quartz crystal resonator. Possible causes and minimisation techniques will be given in part II(ref. 11).

## 6. CONCLUSIONS

In this introductory paper for the first time frequency jump phenomena have been classified into three distinct and separate categories. These are intermittent, discrete and abrupt frequency jumps.

It has been demonstrated that that intermittent frequency jumps are usually caused due to either passive(i.e. quartz resonator, and other RLC circuits) and / or active(i.e. transistors, varactors, regulators and integrated circuits) component faults. These jumps are usually spontaneous and phase - noise behaviour

during the jump period is stable but might have side - bands. These intermittent jumps can be cured simply by replacing faulty components.

Discrete jumps are caused due to non - optimum resonator and / or oscillator circuit configuration. These jumps can be cured by appropriate resonator design or oscillator circuit optimisation procedures. In general phase - noise behaviour of the resonator / oscillator during the entire jump duration is stable and can be defined properly. Discrete jumps can be induced as and when required and jump duration can also be defined. However the jumps are not instantaneous but are built up slowly.

Unlike intermittent & discrete frequency jumps, abrupt frequency jumps on the precision resonators (these are in general contoured crystals) & burst-noise on high - frequency overtone - mode quartz crystal resonators (these are usually polished plano - plano crystals) are of completely different nature. These in general cannot be induced (i.e. switching ON / OFF) by changing known resonator and/or oscillator parameters. Phase - noise of the resonator or oscillator is in fact unstable during the entire jump period and burst duration cannot be defined. As such these two phenomena are believed to be due to the same subtle fundamental problems associated to the resonator design and processing techniques, which will be explored at a later date in part II (ref. 11)

## 7. ACKNOWLEDGEMENT

The authors wish to thank Dr. E.P Graf for his helpful comments and suggestions.

## 8. REFERENCES

1. Filler, R. L, and Vig, J. R, "Long Term Ageing of Oscillators", Proc. 46th AFCS, May 1992, pp. 470-484.

2. Hafner, E, Private Communication at the 49th AFCS, 1995, San Francisco, CA.

3. Norton, J, R, "BVA - Type Quartz Oscillator for Spacecraft", Proc. 45th AFCS, May 1991, pp. 426-430.

4. Norton, J.R, "Tactical BVA Quartz Resonator Performance", Proc. 47th AFCS, June 1993, pp. 609-613.

5. Koyama, M, Watanabe, Y, Sekimoto, H, and Oomura, Y, "A Novel Technique for Measuring Drive Level Dependence of Quartz Units", Proc. 9th EFTF, March 1995, pp. 16-19.

6. Koyama, M, "An Experimental Study of Long Term Ageing of Quartz Oscillators", Proc. 49th AFCS, May/June 1995, pp. 620-622.

7. Walls, F.L, "The Quest to Understand and Reduce 1/f Noise in Amplifiers and BAW Quartz Oscillators", Proc. 9th EFTF, March 1995, pp. 227-240.

8. Knowles, J. E, "On the Origin of the Second - Level - of - Drive Effects in Quartz Crystal Oscillators", Proc. 29th AFCS 1975, pp. 230 - 236.

9. Tuladhar, K.K, Adams. B. J, and Fletcher, E. D, "Hysteresis Effects in Quartz Crystal Resonators", Proc. IERE, International Conference on Frequency Control and Synthesis, University of Surrey, Guildford, UK, April 1987, pp. 19-27.

10. Moulton, G. E. "Burst Noise and 1/f Noise in Quartz Crystals and Oscillators", Proc. 42nd AFCS, June 1988, pp. 389-396.

11. Tuladhar, K.K, and Jenni, G, "Frequency Jumps on BVA & Other Precision Quartz Crystal Resonators and Burst - Noise on Overtone - Mode High - Frequency Quartz Crystal Resonators, Part II", to be published in Proc. 50th AFCS, 1996.

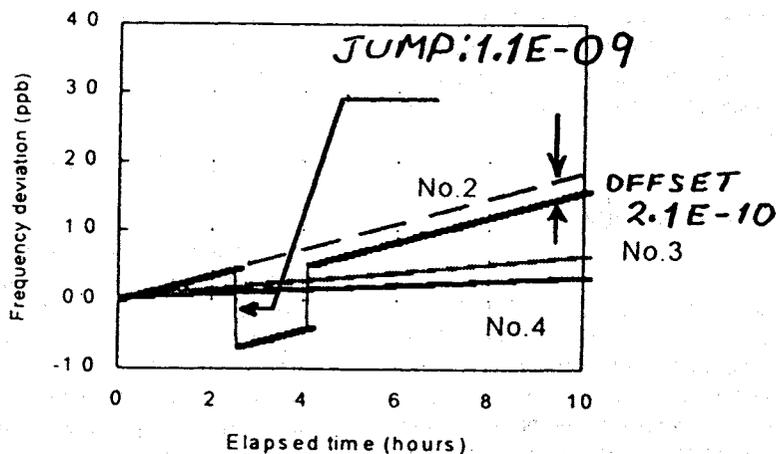


Fig. 1. fig.6 of ref. 6 redrawn to detail frequency jump,  $1.1E-09$  and frequency offset of  $2.1E-10$ .

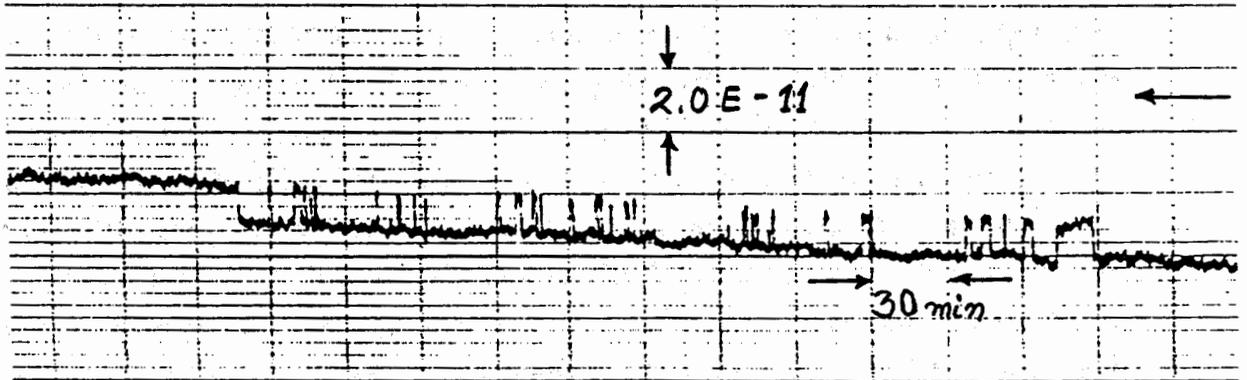


Fig. 2. Intermittent frequency jumps observed on 5.0MHz fifth - overtone mode BVA AT - cut based oscillator.

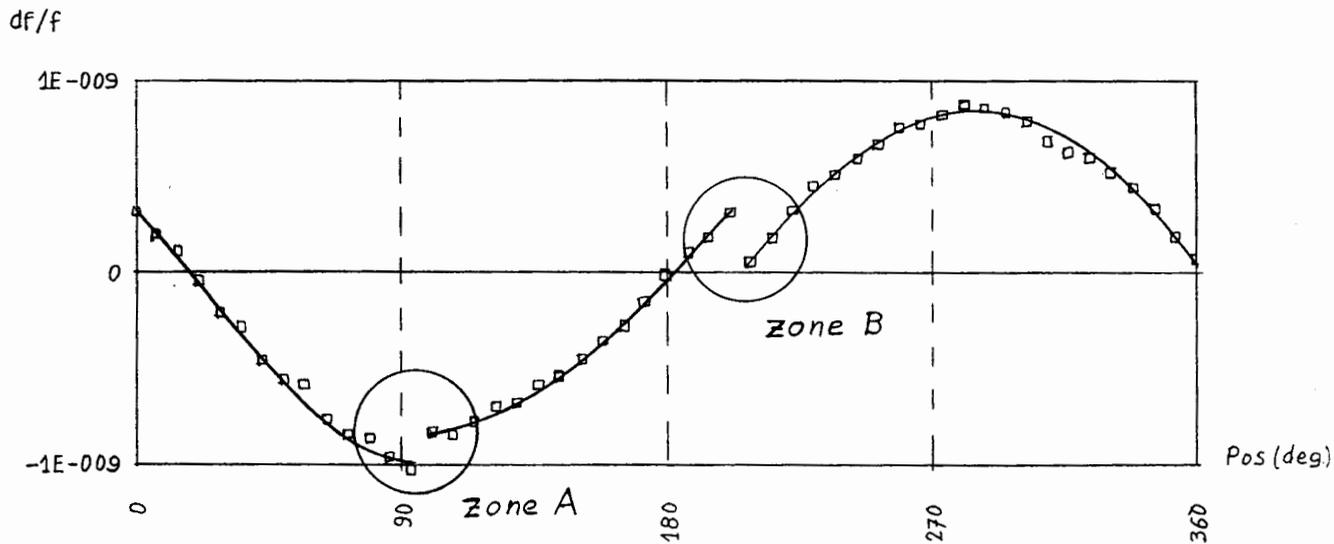


Fig. 3. 2g - tip - over test for non - optimum resonator having jumps in the regions, a) and b).

## QUARTZ CRYSTAL UNITS WITH BAR PIEZOELEMENTS VIBRATING IN DOUBLE-SHEAR MODE

M.E. Bogoush, P.G. Pozdnyakov

Scientific Research Institute "Phonon"  
44 Krasnobogatyrskaya str., Moscow 105023, Russian Federation

Investigations aimed at improving crystal unit characteristics over the range 400÷1000 kHz resulted in designing quartz crystals operating in a double-shear mode. The investigations were realized on the base of quartz crystal bars having a rectangular cross section of close cross sizes, with its length being along X-axis. Vibrations of such bars may be approximately interpreted as a superposition of two shears - in thickness and in width of the piezoelements. Crystal units operating in the above mentioned mode have high electro-physical parameters: dynamic resistance 5÷100 Ohm, motional inductance 2÷20 H, quality factor  $10^5 \div 2 \times 10^6$ , capacitance ratio up to 350, high temperature - frequency stability and low aging. Calculation results showed that quartz crystal units may have high temperature stability by changing the cut angle and across dimensions relations of piezoelements vibrating in double-shear mode.

## Theoretical and Experimental Investigations of Frequency Transient Process Caused In Crystal Plates Under Local Heating.

Alexander N Lepetaev, Anatoly V Kosykh  
Omsk State Engineering University, Russia

It is known that nonuniformly heating produced an additional frequency shift in anisotropic crystal plates. This affect is most pronounced when Crystal is spotly heated. Temperature gradient appearing along the plate causes mechanical stresses that leads to frequency changing. Theoretical investigations of local heating effect on resonator frequency are carried out. Dependence of frequency control efficiency was investigated in relation to:

- Size of heater;
- number of excited harmonic;
- cut angle;
- piezoelement thickness;
- coefficient of heat exchange with environment.

As a result we have theoretical equations, described a process of frequency setting and a lot of experimental investigations. By using MATLAB and MATCAD software maps of heat and stresses fields were calculated for different heating conditions. With allowance for plate zonde characteristic the corresponding frequency shift was calculated. The entirely new result is frequency transient process calculation. This process is essential for many practical applications such as frequency control in low-noise PLL system.

It is shown that the form of transient process depends on the ambient temperature and has two time constant. In conclusion most parameters of local heating controlled oscillator was shown:

- frequency setting process;
- controlling characteristics;
- phase noises.

obtained theoretical results are close to experimental one.

## INFLUENCE OF LOSS ON THE CRITICAL FREQUENCIES OF CERAMIC RESONATORS

Joseph Kelly,<sup>#</sup> Arthur Ballato,<sup>+</sup> and Ahmad Safari<sup>#</sup>

<sup>#</sup> Ceramic Engineering, Rutgers University, Piscataway, NJ 08855-0909, USA

<sup>+</sup> US Army Research Laboratory, Fort Monmouth, New Jersey 07703-5601, USA

### ABSTRACT

Piezoresonators are often characterised by equivalent circuits. Loss then usually appears as an *ad hoc* addition to the circuit. Future stability and manufacturing reproducibility requirements dictate a more detailed understanding of the types of loss mechanisms and their disposition in equivalent networks. The effects of a lossy piezoelectric coupling coefficient,  $k^*$ , on the complex resonance frequencies of a plate resonator driven in a simple thickness mode are explored.

### INTRODUCTION

Acoustic plate resonators are used extensively in telecommunications applications such as cellular radio. The design and characterisation of these components usually proceeds with the aid of equivalent electrical circuits, the element values of which are derived, for the most part, from the resonator's material constants and geometry. The most notable exception is loss, which often is added in an unjustified, superficial manner.

Future commercial requirements on frequency stability and reproducibility in manufacture will necessitate a much more thorough knowledge of material constants, and their variability, as well as a detailed understanding of the types of loss mechanisms and their disposition in equivalent electrical networks of various kinds.

We explore the effects of loss, as manifested in a lossy piezoelectric coupling coefficient,  $k^*$ , on the critical frequencies arising from roots of the complex equation  $\tan(X^*) = +X^*/(k^*)^2$ .

### SIMPLE THICKNESS PLATE MODES

The mechanical motion of a laterally unbounded piezoelectric crystal plate that depends solely on the

thickness coordinate is described as a simple thickness mode. There are generally three such modes, with each possessing harmonics. Piezoelectric excitation of these modes is of two canonical types: thickness excitation (TE), where the driving field is parallel to the thickness coordinate, and lateral excitation (LE), where the driving field is perpendicular to this coordinate. We limit discussion in this paper to TE, and for the sake of clarity, to a single mode; the material coefficients are treated, therefore, simply as scalars.

Reference [1] derives the input admittance for such a plate, assuming no loss, to be:

$$Y_{in} = j\omega C_0 / [1 - k^2 T(X)];$$

where  $T(X) = \tan(X)/(X)$ , and  $X = \kappa_0 h$ . The parameter  $\kappa_0 = \omega/v$  is the acoustic wavenumber,  $v$  is the acoustic velocity, and  $2h$  is the plate thickness. The exact network synthesis [1] of  $Y_{in}$  is shown in Figure 1.

In the case of piezoelectric ceramics, it is unrealistic to neglect the presence of loss [2,3]. Figure 2 shows the position occupied by ceramics *vis-à-vis* other classes of piezoelectrics in the plane of  $Q$  and  $r = (\pi/k)^2/8$ . Ceramics are seen to occupy a region of high coupling and moderate loss. Low cost and high coupling factors make ceramics very attractive for many applications; the effects of loss, however, must be regarded [4,5].

### SOME LOSS MECHANISMS

Among the contributors to loss in the acoustically vibrating plate are: 1) DC conductivity,  $\sigma$ ; 2) dielectric loss,  $\epsilon''$ ; 3) acoustic viscosity,  $\eta$ ; and 4) piezoelectric loss  $e''$ . With reference to Figure 1,  $\sigma$  and  $\epsilon''$  show up primarily as a resistance,  $R_0$ , shunting  $C_0$ ;  $e''$  makes the piezo-transformer turns ratio,  $n$ , complex. The presence of viscosity,  $\eta$ , affects the acoustic transmission line by making both the mechanical admittance,  $Y_0$ , and wavenumber,  $\kappa_0$ , complex through the velocity,  $v$ . In addition, because the acoustic velocity,  $v$ , is piezoelectrically stiffened,  $\sigma$ ,  $\epsilon''$ , and  $e''$  all appear in the formula for velocity as well.

These loss mechanisms all enter the expression for the piezoelectric coupling factor. In the lossless case this

factor is defined as  $k = |e|/\sqrt{\epsilon c}$ ; the piezoelectrically stiffened elastic stiffness is  $c$ . The coupling factor is a dimensionless combination of material coefficients that alone determines the locations of the poles of the input admittance,  $Y_{in}$ . It is natural, therefore, to extend the definition of  $k$  to the general and natural condition of loss. We define  $k$  in the lossy state as  $k^* = |e^*|/\sqrt{\epsilon^* c^*}$ , where the asterisked quantities are the lossy values corresponding to  $e$ ,  $\epsilon$ , and  $c$  in the lossless case.

In this paper we do not consider further the genesis of  $k^*$ , but simply assume its complex nature in order to explore the effects on the resonance frequencies determined from the poles of  $Y_{in}$ .

### COMPLEX RESONANCE DISPLACEMENTS

The denominator of the expression for  $Y_{in}$  equals zero in the lossy case when  $\tan(X^*) = +X^*/(k^*)^2$ . Here, the frequency variable  $X$  has been replaced by  $X^*$  in anticipation of its complex nature. When dealing with harmonics in addition to the fundamental ( $M = 1$ ), it is more convenient to use the frequency displacement [1], defined as  $\delta = M - 2X/\pi$ , or as  $\delta^* = M - 2X^*/\pi$  for the lossy case.

Resonance displacements,  $\delta^* = \delta_{re} + j\delta_{im}$ , can now be computed simply for a desired harmonic from the relation  $\cot(\pi\delta^*/2) = \pi(M - \delta^*)/2(k^*)^2$ .

To ensure passivity, it is necessary that  $k^* = k_{re} - jk_{im}$ , with both  $k_{re}$  and  $k_{im} \geq 0$ . In addition, we define the quotient  $\kappa$  (not to be confused with the wavenumber  $\kappa_0$ ) as  $\kappa = (k_{im}/k_{re})$ , so that  $k^* = k_{re}(1 - j\kappa)$ .

### BEHAVIOUR OF $\delta^*(k_{re}, \kappa, M)$

In Figure 3 are plotted the variations of  $\log |\delta|$  versus harmonic number  $M$ , for  $k_{re} = 0.50$  and  $\kappa = 0(0.25)1$ ; the graphs are only meaningful for  $M$  an integer. Figure 4 graphs the loci of minus  $\delta_{im}$  versus  $\delta_{re}$  for  $k_{re} = 0.50$  and  $M = 1(2)11$  and  $\kappa = 0(0.2)1$ . Notice that  $\delta_{re}$  is usually  $> 0$  (i.e., the real resonance frequency is below the antiresonance frequency occurring at  $\delta_{re} = 0$ ), but not always; for large comparative loss ( $\kappa = k_{im}/k_{re}$  measures comparative loss), the real resonance frequency can be brought above the antiresonance frequency, especially at the lower harmonics.

In Figure 5, minus  $\delta_{im}$  is plotted versus  $\delta_{re}$  for  $M = 1$ , the fundamental harmonic;  $k_{re} = 0(0.2)1$  and  $\kappa = 0, 0.1, 0.5, 0.8$ , and  $1$ . Here is clearly seen how large  $\kappa$  values lead to negative values of  $\delta_{re}$ . Minus  $\delta_{im}$  is graphed in

Figure 6 versus  $\delta_{re}$  for  $M = 1$  and  $\kappa = 0(0.2)1$  for  $k_{re} = 0.2, 0.5, 0.8$ , and  $1$ . From these plots one may determine the character of the complex resonance frequencies,  $f_R^*$ , and their variation with  $k^*$ . These follow from  $X^* = \pi(M - \delta^*)/2 = (\pi/2)(f_R^*/f_0)$ , where  $f_0 = v/4h$  is the fundamental lossless antiresonance (mechanical) frequency.

### CONCLUSIONS

The resonance frequencies of plate resonators executing simple thickness motions depend solely upon the piezocoupling factor. Loss mechanisms, e.g., acoustic viscosity, DC conductivity, dielectric relaxation, etc., require that the coupling factor be a complex number,  $k^*$ . The roots of the frequency equation thereby also become complex.

The complex frequency roots are represented as normalised displacements from the positions of the lossless antiresonances. The complex displacements are constrained to negative imaginary parts by the passivity of the resonator, while the real part may have either sign, depending on the extent of loss, the real part of the coupling factor, and the harmonic number.

### REFERENCES

1. Ballato A, 1977, "Doubly rotated thickness mode plate vibrators," in Physical Acoustics: Principles and Methods, (W P Mason and R N Thurston, eds.) **13**, Ch. 5, pp. 115-181, Academic Press, New York
2. Ballato A, Kelly J, and Safari A, 1996, "Dissipation in ceramic resonators and transducers," IEEE Intl Frequency Control Symp Proc, Honolulu, HI, June, forthcoming
3. Ballato A, and Ballato J, 1996, "Accurate electrical measurements of modern ferroelectrics," Ferroelectrics, forthcoming
4. Jaffe B, Cook W R Jr, and Jaffe H, 1971, "Piezoelectric Ceramics," Academic Press, New York
5. Landolt-Börnstein, Numerical Data and Functional Relationships in Science and Technology, New Series, Group III: Crystal and Solid State Physics, Volumes III/1, 1966; III/11, 1979; and III/29a, 1992. Springer-Verlag, Berlin, New York.

THIS WORK IS IN THE PUBLIC DOMAIN.

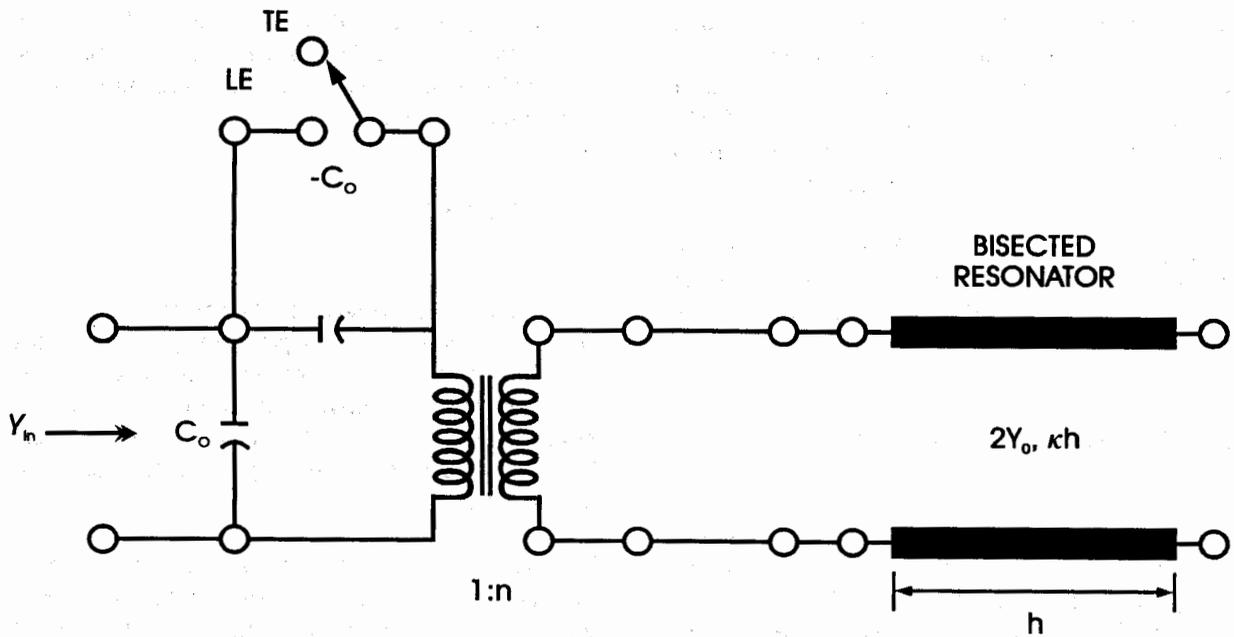


Figure 1. Transmission line equivalent network leading to the resonance frequency equation  $\tan(X) = X/k^2$

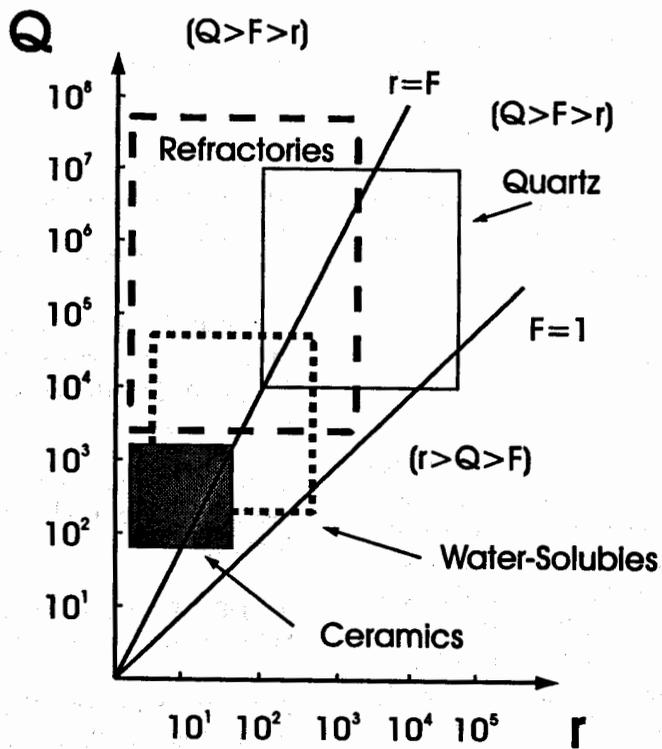


Figure 2. Material quality factor,  $Q$ , versus  $r = (\pi/k)^2/8$ , for various piezoelectric classes of substances. Ceramics are seen to occupy a region of high coupling and moderately high loss.

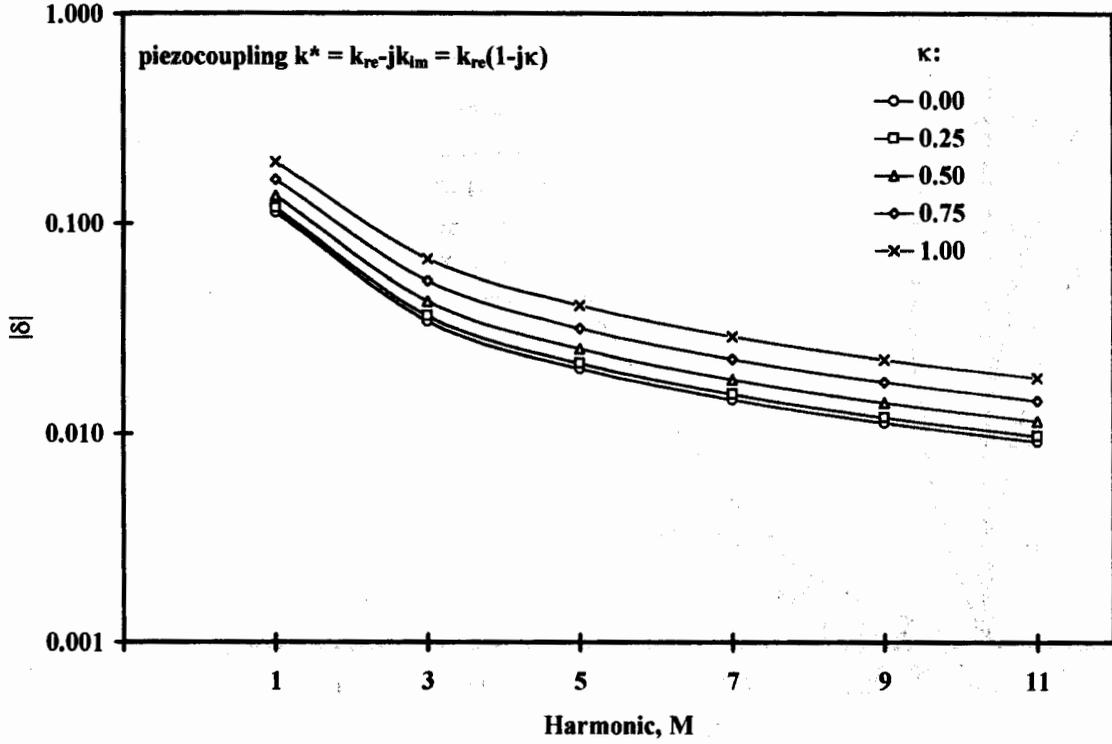


Figure 3. Log  $|\delta|$  versus harmonic number M, for  $k_{re} = 0.50$  and various values of  $\kappa = (k_{im}/k_{re})$ .

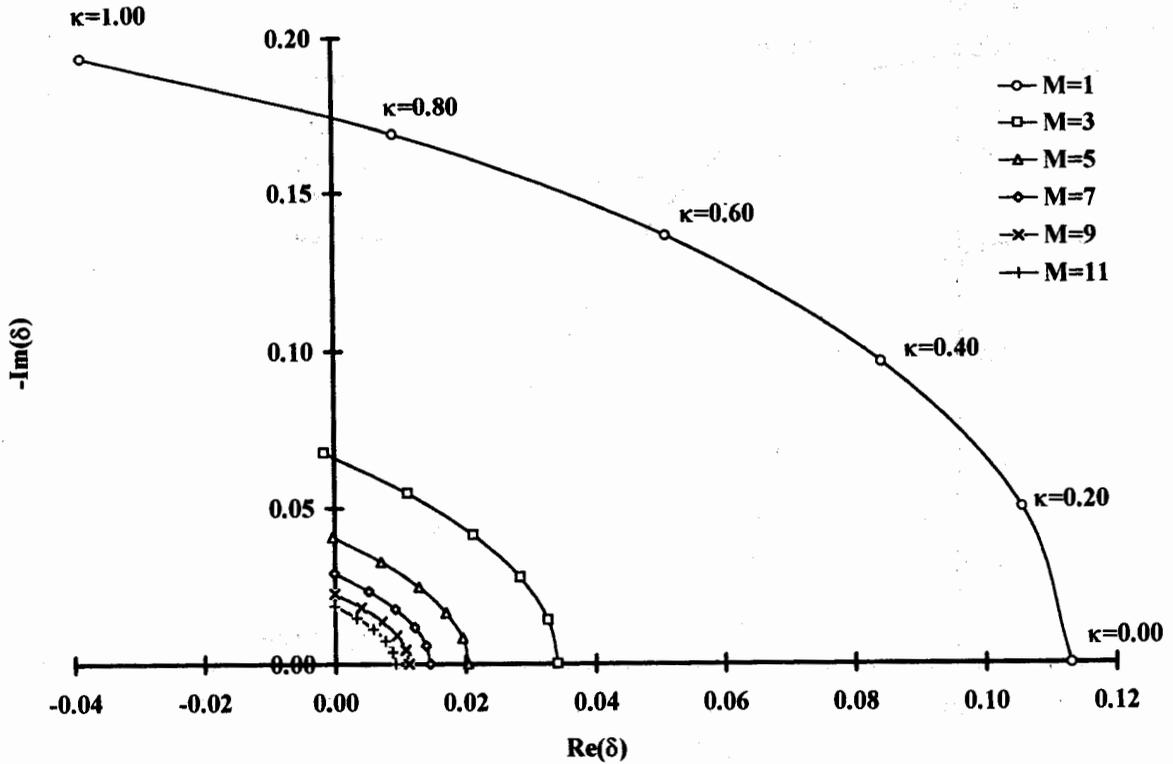


Figure 4. Minus  $\delta_{im}$  versus  $\delta_{re}$  for  $k_{re} = 0.50$  and various values of M and  $\kappa$ .

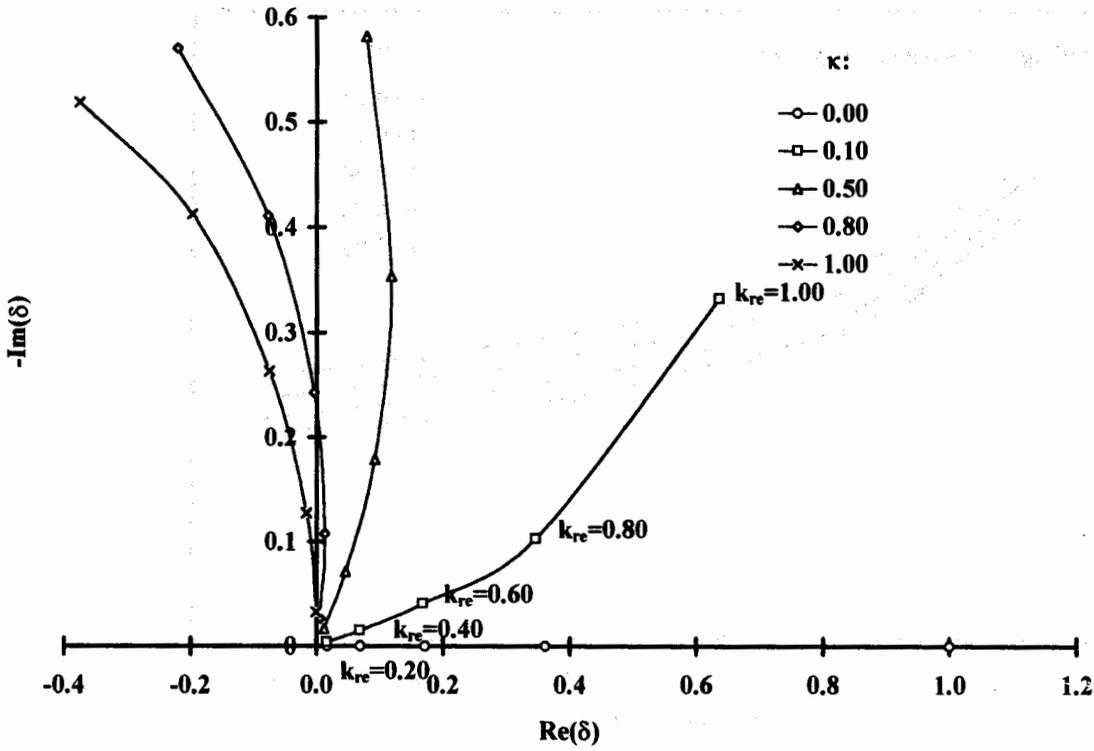


Figure 5. Minus  $\delta_{im}$  versus  $\delta_{re}$  for  $M = 1$ ;  $k_{re}$  varied for fixed values of  $\kappa$ .

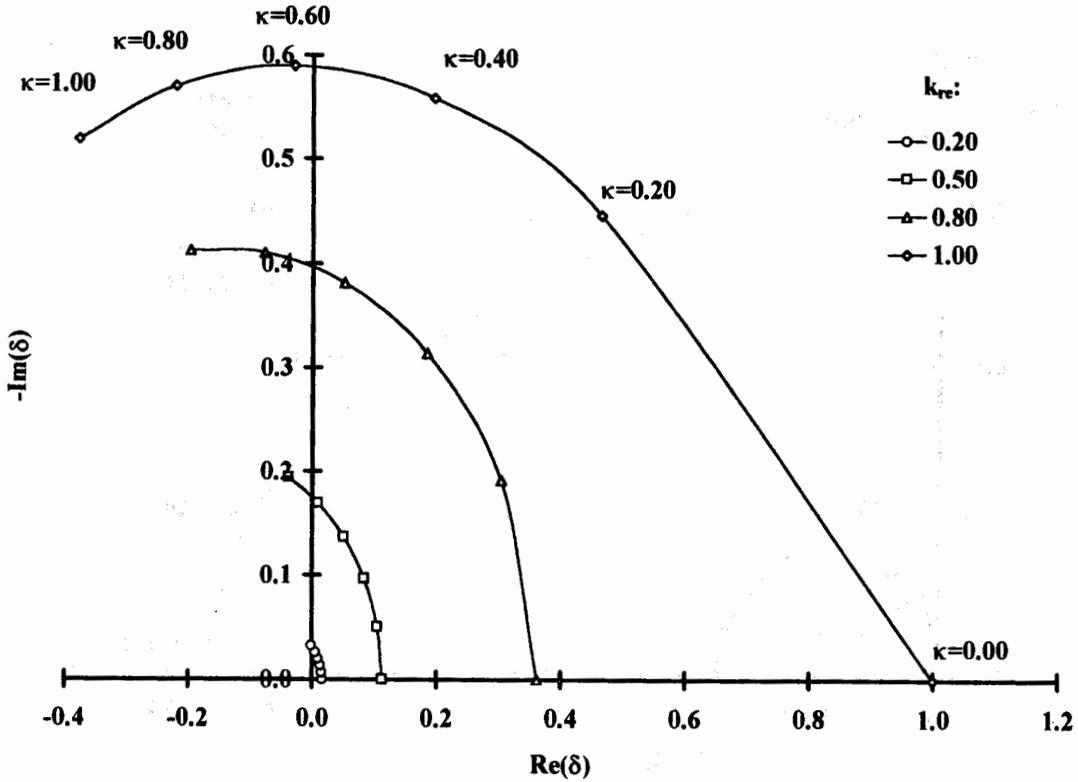


Figure 6. Minus  $\delta_{im}$  versus  $\delta_{re}$  for  $M = 1$ ;  $\kappa$  varied for fixed values of  $k_{re}$ .

## ELECTROMAGNETIC AND MECHANICAL BEHAVIOURS OF WHISPERING GALLERY MODE RESONATORS

O. Di Monaco, B. Piranda, J.S. Armand, Y. Gruson, I. Lajoie, M. Chaubet<sup>#</sup>, J. Piranda<sup>\*</sup>, V. Giordano

Laboratoire de Physique et de Métrologie des Oscillateurs du C.N.R.S.  
associé à l'Université de Franche-Comté Besançon  
32, avenue de l'Observatoire- 25044 Besançon Cedex - France

<sup>#</sup>Centre National d'Etudes Spatiales  
18, avenue Edouard Belin, 31055 Toulouse Cedex - France

<sup>\*</sup>Laboratoire de Mécanique Appliquée R. Chaleat  
28, rue de l'Epitaphe- 25000 Besançon

### INTRODUCTION

Today, frequency and time metrology, modern telecommunication systems and space applications require X-band signal sources with increasing phase noise performances. In this field, the sapphire Whispering Gallery Mode Resonator (WGMR) technique seems to be very attractive. Indeed, exceptional phase noise performances have been recently reported for X-Band sources based on such type of resonator [1,2].

Although the WGMR structure seems relatively simple, the design of an operational resonator brings out a number of technological problems, especially for space applications where in-board systems are submitted to severe environmental stresses.

This paper resumes our investigations on the design of an X-Band WGMR for space applications. We present the measurement results of sapphire permittivity, electromagnetic resonance frequency determination and studies of vibrational behaviour of WGMR.

### SAPPHIRE PERMITTIVITY MEASUREMENT

In order to determine the resonance frequencies of the Whispering Gallery Modes at ambient and liquid nitrogen temperatures, we need precise determination of the sapphire tensor permittivity between 300 and 77K. We use the well known Courtney resonator method to measure  $\epsilon_t$ , the permittivity component in a direction perpendicular to the sapphire disk axis Oz. In figure 1, we reported the results of  $\epsilon_t$  measurements done at two different frequencies 5.4 and 8.5 GHz using respectively the  $TE_{011}$  and the  $TE_{012}$  mode of the same dielectric resonator. It appears that in the range of our measurements the sapphire permittivity could be considered non dependant of frequency.

The measurement of  $\epsilon_z$  is more difficult. In principle,  $\epsilon_z$  could be determined with the resonance frequency of

the  $TM_{011}$  mode of the same Courtney cavity. The  $TM_{011}$  resonance is more affected by the geometrical imperfections of the cavity. And we rather used a Whispering Gallery Mode resonance to determine  $\epsilon_z$ . In this case, the high degree of electromagnetic energy confinement in the sapphire disk itself reduces noticeably the influence of the metallic cavity dimensions. Nevertheless, the formal relation between the resonance frequency and  $\epsilon_z$  is relatively complicated. Then the extraction of  $\epsilon_z$  from the frequency measurements which is done by dichotomy is computer time consuming. In order to get reasonable delay we limited intentionally the accuracy of the dichotomy algorithm.

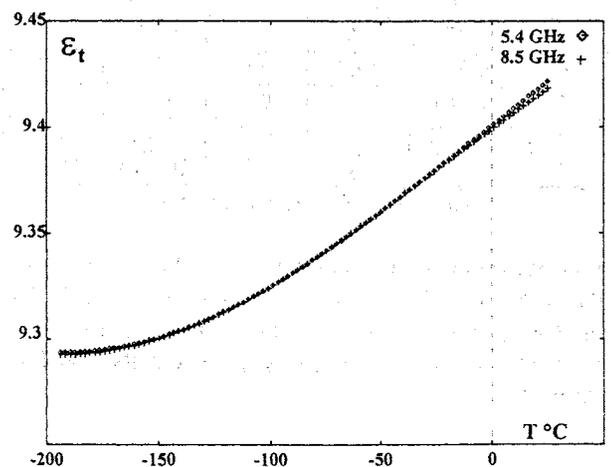


Figure 1:  $\epsilon_t$  variations vs temperature

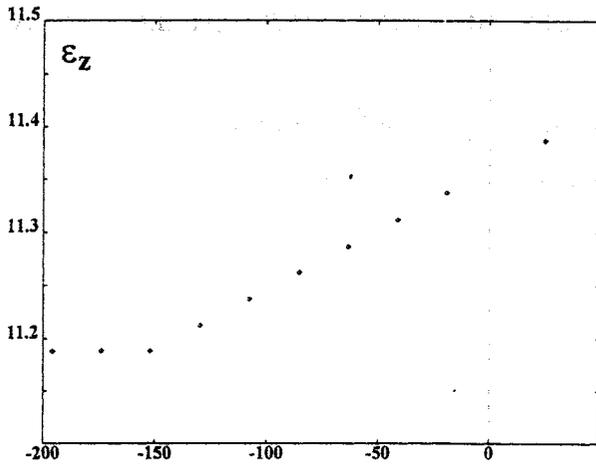


Figure 2 :  $\epsilon_z$  variations vs temperature

### DETERMINATION OF THE RESONANCE FREQUENCIES.

Two theoretical models have been developed to evaluate the resonance frequencies of the Whispering Gallery Mode Resonator, i.e. Mode Matching Method (MMM)[3] and 2D Finite Elements method (FE). Computation results obtained with these two models are compared with the experimental observations (exp.) in table 1 for different azimuthal numbers. The sapphire disk is 36mm diameter and 9mm thick.

m	exp. (GHz)	F.E. (GHz)	$\Delta f/f$ (%)	MMM (GHz)	$\Delta f/f$ (%)
3	6.182	6.23	0.8	6.235	0.9
4	7.07	7.14	1	7.014	-0.8
5	7.855	7.9	0.6	7.778	-1
6	8.612	8.68	0.8	8.549	-0.7
7	9.381	9.45	0.7	9.33	-0.5
8	10.142	10.22	0.8	10.121	-0.2
9	10.911	11	0.8	10.918	0.1

Table 1: resonance frequencies

It is interesting to note that the F.E. method gives about the same error for all azimuthal numbers  $\Delta f/f=1\%$ . On the contrary the accuracy of the MMM increases with the azimuthal number and gives better results than F.E for  $m>6$ .

### MECHANICAL BEHAVIOUR OF WGMR

Regarding the phase noise performance potentialities, the X-band WGMR oscillator compares with the standard synthesis chain based on low-frequency quartz oscillators. WGMR could find some applications in some specific space missions where ultra-low noise microwave source are needed. Nevertheless, whether WGMR survives to spacecraft environment remains to be proved especially during

the launch where the in-board instruments are submitted to severe random vibrations between 0-2000Hz.

In this preliminary works, we analysed the first mechanical resonances of different WGMR structures with the help of a 3-D finite element simulator.

#### Isolated sapphire resonator:

In a first approach, we considered only the sapphire resonator with its supporting structure. Different configurations have been analysed:

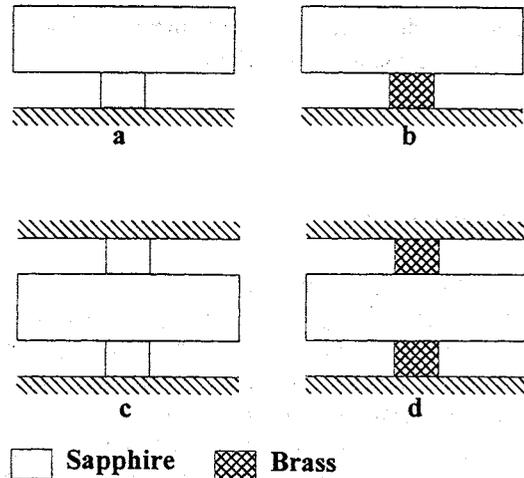


Figure 3. Mechanical Structures of WGMR

The sapphire disk is 36mm diameter and 9mm thick, the supporting post is 10mm diameter and 4,5mm height. All the mechanical bonds are supposed perfect embeddings. The table 2 gives the first mechanical resonance frequencies for these structures:

a (Hz)	b (Hz)	c (Hz)	d (Hz)	
7329	6144	10331	8672	torsion
9198	7624	25887	22769	deflection
29439	25060	34755	29000	-

Table 2. Mechanical resonance frequencies for isolated sapphire resonator

The first resonant mode is a torsion mode. The following figure represent this mode for the structures a and c. The dark region represents the region where the displacement is maximum.

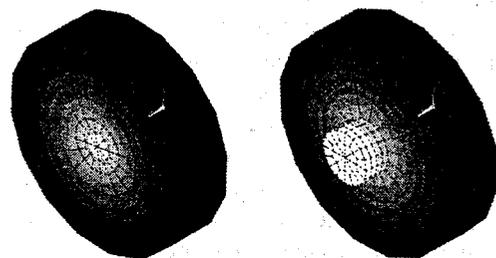


Figure 4. The torsion mode of the dielectric resonator.

We can see :

- as expected the embedded-embedded structure is more rigid than the embedded-free one.
- The Young module of the sapphire being higher than the brass or copper one, the use of metallic post decreases the mechanical resonance frequencies.

Nevertheless, from these results, a good mechanical behaviour could be presaged. Indeed the mechanical resonance frequencies for the 4 configurations are greater than 2000Hz.

Sapphire disk in the metallic cavity:

In the actual configuration, the sapphire disk is placed on the centre of a metallic (copper) cavity as shown in figure 5:

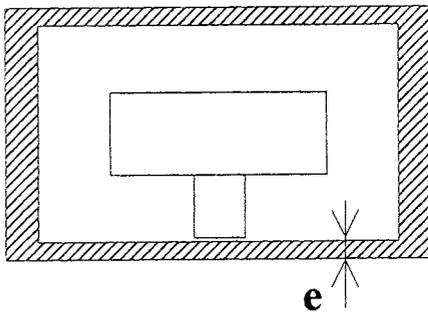


Figure 5 Sapphire disk in the copper cavity

The rigidity of the copper being not infinite, some departures from the preceding model could appear, especially when the thickness  $e$  of the cavity base is decreased. Table 3 and 4 compare for the configurations a and c, the first mechanical resonance frequencies for two different values of the copper cavity thickness.

Embedded-free configuration			
e=15mm		e=6mm	
6061 Hz	sapphire torsion	4809 Hz	sapphire deflection
6268 Hz	sapphire deflection	5071 Hz	cavity deformation
8089 Hz	sapphire deflection	5920 Hz	sapphire torsion
10407 Hz		6997 Hz	cavity deformation

Table 3. Mechanical resonance frequencies for the configuration a in the copper cavity

Embedded-embedded configuration			
e=15mm		e=6mm	
8542 Hz	sapphire torsion	6446 Hz	cavity deformation
10714 Hz	sapphire deflection	8345 Hz	sapphire torsion
11557 Hz	sapphire deflection	8891 Hz	sapphire deflection
		10618 Hz	sapphire deflection

Table 4. Mechanical resonance frequencies for the configuration c in the copper cavity

When the sapphire disk is embedded in the copper cavity, the mechanical resonance frequencies decrease but remain higher than 2000Hz. Moreover when the thickness of the cavity base is decreased to 6mm, cavity deformation modes appear in the spectrum.

Temperature compensated sapphire WGMR

The last structure analysed is due to J.Dick [2]. It allows to compensate the temperature dependence of the sapphire permittivity in order to obtain a noticeable improvement of the frequency stability compared with the traditional structure studied previously. Such WGMR is constituted from 2 sapphire disks separated by a copper post. The increase of the electromagnetic resonance frequency with the temperature is mechanically compensated, at a certain temperature, by the dilatation of the copper post which increases the air-gap between the two sapphire pieces. It has been demonstrated by the JPL group that this configuration could provide a frequency stability of the order of a few  $10^{-13}$  for integration times of  $1 \leq \tau \leq 100$  seconds. For the mechanical analysis, we took the dimension of the resonator described by J.Dick in its publication. The first mechanical resonance frequencies are:

1627 Hz  
1680 Hz  
3862 Hz  
3944 Hz

In this case, the first mechanical modes are below 2000Hz which could be detrimental for spatial applications.

**CONCLUSION:**

This paper resumes our investigations on the design of an X-Band WGMR for space applications. Precise determination of the sapphire permittivity allows us to determine with a good accuracy electromagnetic resonance frequencies of the WGMR. In the second part of this paper, we focused our attention on the mechanical behaviour of WGMR submitted to vibration. The preliminary results presented here presage a good behaviour of the traditional WGMR structure. Indeed, the calculated resonance frequencies are greater than 2000Hz. However, it should be noted that the mechanical models used in this paper assumed perfect mechanical bonds between the different parts of the resonators. Then, further studies are needed to determine the influence of looses, non-perfect mechanical fixing and frictions.

**REFERENCES:**

[1] *"Advanced phase noise suppression technique for next generation of ultra low-noise microwave oscillators"*

E.N. Ivanov, M.E. Tobar and R.A. Woode.

Proceedings of the 1995 IEEE International Frequency Control Symposium, 31 May- 2 June 1995, San Francisco Cal. US, pp 314-320.

[2] *"Improved performance of a temperature compensated LN<sub>2</sub> cooled sapphire oscillator."*

D.G. Santiago, R.T. Wang and G.J. Dick.

Proceedings of the 1995 IEEE International Frequency Control Symposium, 31 May- 2 June 1995, San Francisco Cal. US, pp 397-400.

[3] *"Approximat approach to the design of shielded dielectric disk resonators with Whispering-Gallery modes."*

E.N. Ivanov, D.G. Blair and V.I. Kalinichev.

IEEE Trans. on MTT, Vol 41, n°4, April 1993, pp 632-638.

## POISSON'S RATIO FOR PIEZOCERAMICS

Arthur Ballato

US Army Research Laboratory, Physical Sciences Directorate  
Fort Monmouth, New Jersey 07703-5601, USA

### ABSTRACT

Formulas for Poisson's ratio are given for piezoelectric ceramics of arbitrary orientation and electrical immittance boundary conditions.

### INTRODUCTION

Poisson's ratio finds application in a number of frequency control areas, for example, as indication of the mechanical coupling between various modes of motion in crystal resonators for filters and oscillators. Future applications will involve mechanically resonant microstructures integrated with electronic and optical circuitry [1], and novel quantum devices [2,3]. These require extension of Poisson's ratio considerations to a variety of piezoelectric substances, such as high-coupling piezoceramics.

The poling state of piezoceramics is nearly always either lateral, or parallel, to the major surfaces of the device. This maximizes the effective piezoelectric coupling by the electrode placements. Configurations for microelectromechanical structures (MEMS) that are currently under development utilise more general orientations that take advantage of the achievable differences in Poisson's Ratios in different directions, and their alteration by electrical loadings.

Poisson's ratio,  $\nu$ , is defined for isotropic media as the quotient of lateral contraction to longitudinal extension arising from application of a simple tensile stress.

In most materials, the dimensionless number  $\nu$  is positive. In crystals and poled electroceramics,  $\nu$  takes on different values, depending on the directions of stress and strain chosen. The maximum value of  $\nu = +1/2$  is obtained in the incompressible medium limit, where volume is preserved; for ordinary materials, values of  $+1/4$  to  $+1/3$  are typical, but in crystals  $\nu$  may vanish, or take on negative values. In order to provide a synoptic, yet relatively uncomplicated picture, Figure 1 sketches the bounds on  $\nu$  as function of the traditional Lamé constants of an isotropic medium. Table 1 relates

various elastic measures for substances or conditions indicated in the figure, or discussed in the sequel.

Analytical formulas for Poisson's ratio (in the absence of piezoelectricity) are expressed in terms of elastic moduli alone [4]. For the case of crystals of general anisotropy, these expressions are quite unwieldy, but for substances in the hexagonal system the symmetry elements reduce the complexity considerably. Many of the materials being considered for future microdevices are characterized by hexagonal symmetry. Inclusion of piezoelectricity makes the analysis somewhat more complex, but the magnitude of the effect in many poled ceramics renders its consideration mandatory.

The hexagonal crystal system [5,6,7] comprises seven point groups (6-bar m2, 622, 6mm, 6/m mm, 6-bar, 6, and 6/m), and includes a number of the binary semiconductor systems, and their alloys. These have the piezoelectric wurtzite structure; examples are GaN and AlN. The family of poled electroceramics, including BaTiO<sub>3</sub>, PZT, and related alloys are characterized by symmetry  $\infty$ mm, that is, they are transversely isotropic. However, this symmetry is equivalent to 6mm for all tensor properties up to and including rank five [8]; this includes piezoelectricity and elasticity.

### POISSON'S RATIO FOR CRYSTALS - DEFINITION

Poisson's ratio for crystals [5] is defined in general as  $\nu_{ji} = -s_{ij}' / s_{ij}'$ , where  $x_j$  is the direction of the longitudinal extension,  $x_i$  is the direction of the accompanying lateral contraction, and the  $s_{ij}'$  and  $s_{ij}'$  are the appropriate elastic compliances (including the piezoelectric contribution), referred to this right-handed axial set. It suffices to take  $x_1$  as the direction of the longitudinal extension; then two Poisson's ratios are defined by the orientations of the lateral axes  $x_2$  and  $x_3$ :  $\nu_{21} = -s_{12}' / s_{11}'$  and  $\nu_{31} = -s_{13}' / s_{11}'$ . Application of the definition requires specification of the orientation of the  $x_k$  coordinate set with respect to the crystallographic directions, transformation of the compliances, piezoelectric, and dielectric coefficients accordingly, and determination of the effective

compliances from a knowledge of the electrical immitance boundary conditions.

### ROTATED COMPLIANCES - GENERAL

The unprimed compliances are referred to a set of right-handed orthogonal axes related to the crystallographic axes in the manner defined by the IEEE standard [9]. Direction cosines  $a_{mn}$  relate the transformation from these axes to the set specifying the directions of the applied longitudinal extension ( $x_1$ ), and the resulting lateral contractions ( $x_2$  and  $x_3$ ). General expressions for the transformed hexagonal compliances that enter the formulas for  $v_{21}$  and  $v_{31}$  are:

$$s_{11}' = s_{11} [a_{11}^2 + a_{12}^2]^2 + s_{33} [a_{13}^4] + (s_{44} + 2 s_{13}) [a_{13}^2] [a_{11}^2 + a_{12}^2]$$

$$s_{12}' = s_{11} [a_{11} a_{21} + a_{12} a_{22}]^2 + s_{33} [a_{13}^2 a_{23}^2] + s_{44} [a_{13} a_{23}] [a_{11} a_{21} + a_{12} a_{22}] + s_{12} [a_{11} a_{22} - a_{12} a_{21}]^2 + s_{13} [a_{23}^2 [a_{11}^2 + a_{12}^2] + a_{13}^2 [a_{21}^2 + a_{22}^2]]$$

$$s_{13}' = s_{11} [a_{11} a_{31} + a_{12} a_{32}]^2 + s_{33} [a_{13}^2 a_{33}^2] + s_{44} [a_{13} a_{33}] [a_{11} a_{31} + a_{12} a_{32}] + s_{12} [a_{11} a_{32} - a_{12} a_{31}]^2 + s_{13} [a_{33}^2 [a_{11}^2 + a_{12}^2] + a_{13}^2 [a_{31}^2 + a_{32}^2]]$$

One similarly obtains, in a straightforward manner, transformations for the piezoelectric and dielectric matrices.

### TRANSFORMATION MATRIX FOR ROTATIONS

Poisson's ratio for the most general case may be derived by considering the transformation matrix for a combination of three coordinate rotations: a first rotation about  $x_3$  by angle  $\phi$ , a second rotation about the new  $x_1$  by angle  $\theta$ , and a third rotation about the resulting  $x_2$  by angle  $\psi$ . When these angles are set to zero, the  $x_1, x_2, x_3$  axes coincide respectively with the reference crystallographic directions. The condition of transverse isotropy, however, renders all results independent of azimuthal angle  $\phi$ , which is henceforth not considered. For nonzero angles  $\theta$  and  $\psi$ , the direction cosines  $a_{mn}$  are as follows, with the abbreviations  $c(\theta)$  and  $s(\theta)$  for  $\cos(\theta)$  and  $\sin(\theta)$ , etc.:

$$\begin{array}{lll} [c(\psi)] & [s(\theta)s(\psi)] & [-c(\theta)s(\psi)] \\ [0] & [c(\theta)] & [s(\theta)] \\ [s(\psi)] & [-s(\theta)c(\psi)] & [c(\theta)c(\psi)] \end{array}$$

Substitution of these  $a_{mn}$  into the expressions for  $s_{11}'$ ,  $s_{12}'$ , and  $s_{13}'$ , and thence into the formulas  $v_{21} = -s_{12}'$

and  $v_{31} = -s_{13}' / s_{11}'$  formally solves the problem for specified values of  $\theta$  and  $\psi$  for the nonpiezoelectric case.

### INCLUSION OF PIEZOELECTRICITY

Four sets of elastic-piezoelectric-dielectric constitutive equations are commonly used. Poisson's ratio is defined using a single strain  $T_1 = T$  and therefore the constitutive equations of greatest utility are:

$$\begin{array}{ll} S = s^E T + d^t E & S = s^D T + g^t D \\ D = d T + \epsilon^T E & E = -g T + \beta^T D \end{array}$$

Boundary conditions (BCs):

$$\begin{array}{ll} -D = \epsilon^b E & -E = \beta^b D \\ \epsilon^b \beta^b = I \end{array}$$

One could add mechanical BCs as well, if necessary; we discuss only the purely electrical BCs given. These BCs can be realized in some cases by use of another (external) crystal with electrodes, or, more generally, by attachment of an external active or passive electrical network. The  $\epsilon^b$  and  $\beta^b$  can be complex numbers.

In the general equations given above, and subsequently, a superscript "t" denotes transpose. Because of the single stress component, it will be convenient below to use a prime somewhat unconventionally to denote a transpose for certain portions the piezoelectric matrices. The components of the field matrices are:

$$S = (S_1, S_2, S_3)^t; T = (T_1 = T); D = (D_1, D_2, D_3)^t; E = (E_1, E_2, E_3)^t.$$

The components of the matter matrices are:

$$s^* = (s_{11}, s_{12}, s_{13})'; \epsilon^+ = (\epsilon_{11}, \epsilon_{12}, \epsilon_{13}); (\epsilon_{21}, \epsilon_{22}, \epsilon_{23}); (\epsilon_{31}, \epsilon_{32}, \epsilon_{33}); \text{ and ditto for the dielectric impermeabilities } \beta.$$

In the above,  $s^*$  is  $s^E$  or  $s^D$  and  $\epsilon^+$  is  $\epsilon^T$  or  $\epsilon^b$  and ditto for  $\beta$ .

For the piezoelectric matrices, we define:

$d' = (d_{11}, d_{21}, d_{31}), (d_{12}, d_{22}, d_{32}), (d_{13}, d_{23}, d_{33})$ ; that is,  $d'$  consists of a (3x3) matrix where  $(d_{11}, d_{21}, d_{31})$  is the first row, etc. Likewise:  $g' = (g_{11}, g_{21}, g_{31}), (g_{12}, g_{22}, g_{32}), (g_{13}, g_{23}, g_{33})$ . But, for the quantity  $d$ , we have the definition  $d = (d_{11}, d_{21}, d_{31})'$ ; that is,  $d$  is a (3x1) single column matrix consisting of the elements of the first row of  $d'$  rather than the transpose of the (3x3) matrix  $d$ . Likewise:  $g = (g_{11}, g_{21}, g_{31})'$ ; that is,  $g$  is a (3x1) single column matrix consisting of the elements of the

first row of  $g'$  rather than the transpose of the (3x3) matrix  $g$ .

Then, for the general case, we let

$$\varepsilon^T + \varepsilon^b = \varepsilon \quad \text{and} \quad \beta^T + \beta^b = \beta$$

So that:  $D = d^T T + \varepsilon^T E = -\varepsilon^b E$ ; and  $d^T T + \varepsilon E = 0$ ; thus  $E = -\varepsilon^{-1} d^T T$ ; this is inserted into  $S = s^B T + d^T E$  to give:  $S = s^B T + d^T E = s^B T - d^T \varepsilon^{-1} d^T T$ , or  $S = s^o T$ , where  $s^o = [s^B - d^T \varepsilon^{-1} d]$

Also:  $E = -g T + \beta^T D = -\beta^b D$ ; so  $-g T + \beta D = 0$ ;  $D = +\beta^{-1} g T$ ; this is inserted into  $S = s^D T + g^T D$  to give:  $S = s^D T + g^T D = s^D T + g^T \beta^{-1} g T$ , or  $S = s^d T$ , where  $s^d = [s^D + g^T \beta^{-1} g]$

Equating  $s^o$  and  $s^d$ , it is found that  $\varepsilon^b \beta^b = I$ , as asserted above. Thus the two equivalent general solutions are:

$$S = s^o T, \text{ where } s^o = [s^B - d^T \varepsilon^{-1} d], \text{ or}$$

$$S = s^d T, \text{ where } s^d = [s^D + g^T \beta^{-1} g].$$

For certain simple limiting cases of boundary conditions, i.e., short circuit (SC) or open circuit (OC), either  $\varepsilon^{-1}$  or  $\beta^{-1}$  may not exist. In these cases one must proceed with caution. Depending upon the BCs, it is advantageous to use one or another of the piezoelectric ( $d$  or  $g$ ) matrices.

## POISSON'S RATIO FOR PIEZOELECTRICS

The Poisson's ratio for piezoelectric substances is given formally by  $\nu_{21} = -s_{12}'/s_{11}'$ ;  $\nu_{31} = -s_{13}'/s_{11}'$  as before. But now the primed (rotated) compliances stand for the pertinently rotated components of either of the quantities  $[s^B - d^T \varepsilon^{-1} d]$  or  $[s^D + g^T \beta^{-1} g]$ .

## APPLICATION TO PIEZOCERAMICS

The poling state of piezoceramics nearly always encountered in present commercial practice is either parallel, or lateral, to the major surfaces of the device. This is because the effective piezoelectric coupling is thereby maximized by the electrode placements. Newer configurations, currently under development for microelectromechanical structures and devices utilize more general orientations that take advantage of the achievable differences in Poisson's ratios in different directions. Table 2 gives some representative examples, based on the relations derived above; entries are computed from data in Refs. [10] and [11].

## CONCLUSIONS

Poisson's ratio, with respect to rotated coordinate axes, for hexagonal materials, and particularly, poled ferroelectric ceramics, has been obtained. All results are independent of rotations about the six-fold axis. A number of simple cases are of particular interest; these obtain for open-circuit boundary conditions:

- For longitudinal extension in the basal plane,  $\psi = 0$ :

$$\text{When } \theta = 0: \nu_{21} = -s_{12}/s_{11}; \nu_{31} = -s_{13}/s_{11}$$

$$\text{When } \theta = \pi/4: \nu_{21} = \nu_{31} = -(s_{12} + s_{13})/2s_{11}$$

- For longitudinal extension at an angle  $\psi$  from the basal plane:

$$\text{When } \theta = 0; \psi = \pi/4: \nu_{21} = -2(s_{12} + s_{13})/(s_0 + s_{44})$$

$$\nu_{31} = -(s_0 - s_{44})/(s_0 + s_{44}); s_0 = (s_{11} + s_{33} + 2s_{13})$$

$$\text{When } \theta = \pi/4; \psi = \pi/4: s_2 = (s_{11} + s_{33} - (s_{44} + 2s_{13}))$$

$$\nu_{21} = -[4s_{12} + 12s_{13} + 2s_2]/[12s_{11} + 4s_{33} - 3s_2]$$

$$\nu_{31} = -[8s_{12} + 8s_{13} + s_2]/[12s_{11} + 4s_{33} - 3s_2]$$

$$\text{When } \psi = \pi/4; \chi = \pi/4: s_2 = (s_{11} + s_{33} - (s_{44} + 2s_{13}))$$

$$\nu_{21} = \nu_{31} = -[s_{12} + 3s_{13} + s_2/2]/[2(s_{11} + s_{33}) - s_2]$$

- For longitudinal extension along the six-fold axis:  $\nu_{21} = \nu_{31} = -s_{13}/s_{33}$

## REFERENCES

1. Ballato A, 1995, 'Piezoelectricity: old effect, new thrusts,' IEEE Trans UFFC, **42**, 916-926
2. Strosio M A, Kim K W, Yu S, and Ballato A, 1994, 'Quantized acoustic phonon modes in quantum wires and quantum dots,' J Appl Phys, **76**, 4670-4675
3. Yu S, Kim K W, Strosio M A, Iafate G J, and Ballato A, 1994, 'Electron-acoustic-phonon scattering rates in rectangular quantum wires,' Phys Rev B, **50**, 1733-1738
4. Ballato A, 1996, 'Poisson's ratio for tetragonal, hexagonal, and cubic crystals,' IEEE Trans UFFC, **43**, 56-62
5. Cady W G, Piezoelectricity, McGraw-Hill, New York, 1946; Dover, New York, 1964
6. "IRE Standards on Piezoelectric Crystals, 1949," Proc IRE, **37**, 1378-1395

7. Nye J F, Physical Properties of Crystals, Clarendon Press, Oxford, 1957; Oxford University Press, 1985

8. Juretschke H J, 1974, Crystal Physics, W A Benjamin, Reading, MA

9. "IEEE Standard on Piezoelectricity," ANSI/IEEE Standard 176-1987, Inst Elec Elx Eng, New York

10. Jaffe B, Cook W R Jr, and Jaffe H, 1971, Piezoelectric Ceramics, Academic Press, New York

11. Landolt-Börnstein, Numerical Data and Functional Relationships in Science and Technology, New Series, Group III: Crystal and Solid State Physics, Volumes III/1, 1966; III/11, 1979; and III/29a, 1992. Springer-Verlag, Berlin, New York

THIS WORK IS IN THE PUBLIC DOMAIN.

TABLE 1. RELATIONS AMONG ISOTROPIC ELASTIC PARAMETERS.

SUBSTANCE OR CONDITION	$\nu$	$\lambda$	$\mu$	$Y$	$\kappa$	$v_{sh}/v_{long}$
Ideal fluids	1/2	$\lambda$	0	0	$\lambda$	0
Many metals	1/3	$2\mu$	$\mu$	$8\mu/3$	$8\mu/3$	1/2
Poisson relation	1/4	$\mu$	$\mu$	$5\mu/2$	$5\mu/3$	$1/\sqrt{3}$
Pure rigidity	0	0	$\mu$	$2\mu$	$2\mu/3$	$1/\sqrt{2}$
Perfect compressibility	-1	$-2\mu/3$	$\mu$	0	0	$\sqrt{3}/2$
Incompressible liquids	1/2	$\infty$	0	0	$\infty$	0
Incompressible solids	---	$\infty$	$\infty$	$\infty$	$\infty$	---

[ $v_{sh}/v_{long}$ ] is ratio of shear to longitudinal wave velocities

TABLE 2. POISSON'S RATIO, YOUNG'S MODULUS, AND COMPRESSIBILITY OF SELECTED PIEZOCERAMICS; OPEN CIRCUIT CONDITION ON ALL FACES.

COMPOSITION	$\theta^\circ/\psi^\circ$	$\nu_{21}$	$\nu_{31}$	$Y$	$\kappa$
Ba Ti O <sub>3</sub>	0/0	0.305	0.333	117.0	106.3
PZT-4	45/0	0.380	0.380	81.3	92.9
PZT-5A	0/45	0.380	0.392	58.6	89.1
PZT 52/48	45/45	0.401	0.354	63.7	93.5
PZT 65/35 9.067	0/0	0.290	0.395	110.3	95.6
Pb <sub>0.76</sub> Ca <sub>0.24</sub> Ti O <sub>3</sub> 7.35	45/0	0.399	0.399	136.1	69.6
Pb <sub>0.96</sub> La <sub>0.04</sub> Ti O <sub>3</sub> 7.21	0/45	0.232	0.163	147.2	81.4
Pb <sub>0.89</sub> Nd <sub>0.11</sub> Ti O <sub>3</sub> 6.84	45/45	0.235	0.262	149.4	94.0

[Y and  $\kappa$  in GPa]

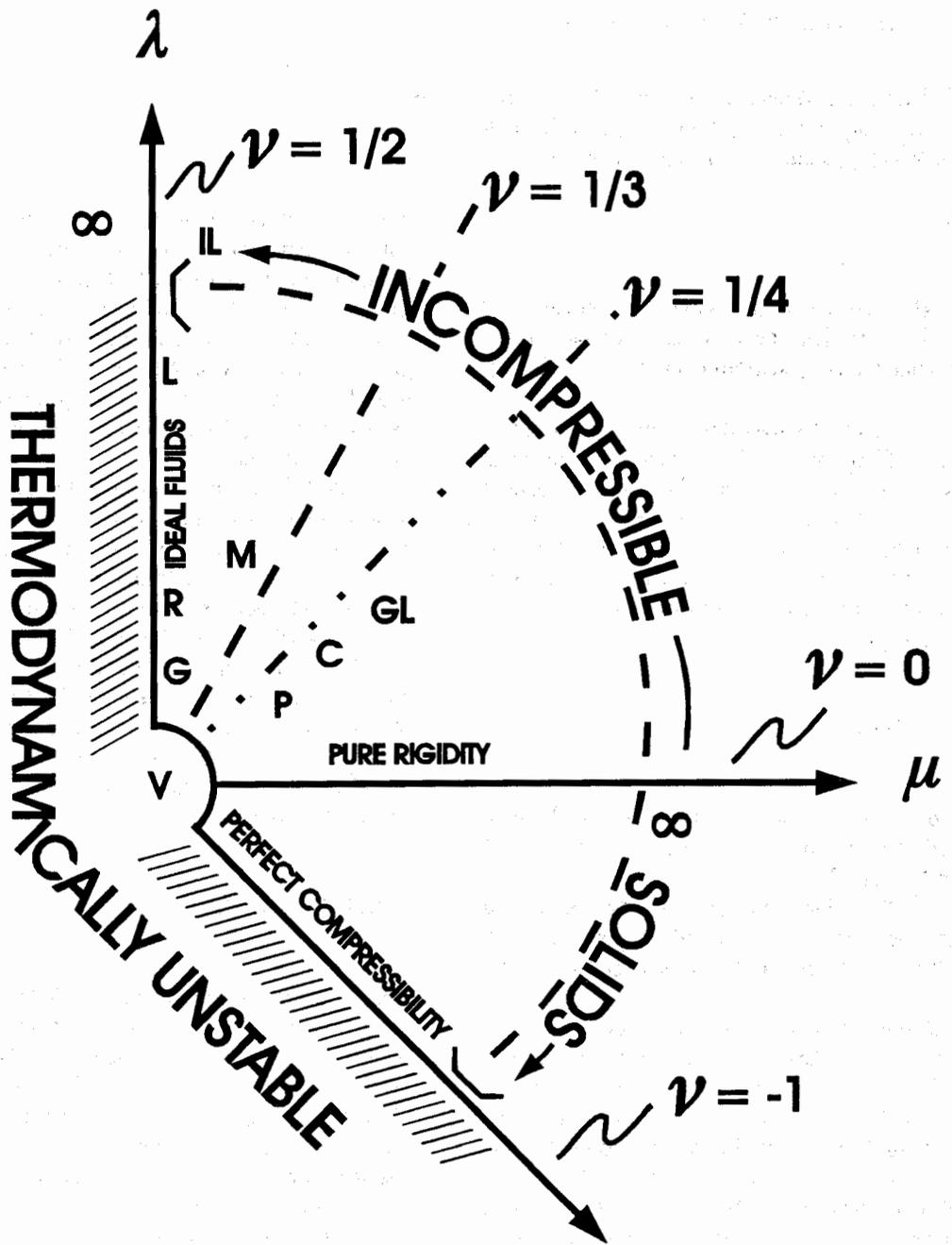


Figure 1. Limits on the Lamé constants of isotropic solids. Symbols employed are: V, vacuum; G, gas; R, rubber; L, liquid; IL, incompressible liquid; M, metal, ceramic; P, plastic; C, cement; GL, glass.

## CONTRIBUTION TO THE DESIGN OF THE WAVE FILTERS COMPOSED FROM THE ELASTICALLY COUPLED TWO RESONATORS STRUCTURE

Irina Mateescu, Jiri Zelenka\*, Alexandru Serbanescu\*\*

Institute of Physics and Tehnology of Materials, Bucharest, Romania

\*Technical University of Liberec, Liberec, Czech Republic

\*\*Technical Military Academy, Bucharest, Romania

### ABSTRACT

The large electromechanical coupling coefficient of langasite makes possible the production of monolithic filters with much wider bandwidth than are achievable on quartz.

In this paper we present a new design method for eight-pole band-pass filter using an Y-cut langasite substrate. The computation of exact values of the inter-resonator coupling coefficient is performed by using the Tiersten's description of the monolithic structure. The results are compared with those deduced from the relation between coupling coefficient and dimensions of two-pole monolithic structure.

### INTRODUCTION

Crystalline quartz is a weak piezoelectric material which cannot fulfil the requirements of modern piezo- and acousto-electronics (1).

The evolution of the electronic equipment has led to new piezoelectric materials for filters with larger bandwidth and oscillators with better frequency stability than quartz.

Many applications of the piezoelectric devices impose new requirements for these devices in order to enhance the achievable system performances or simply to allow for the implementation of new systems (2). The most important among them are probably the land mobile and space digital radiocommunication systems (2,3,4).

New filtering devices with much larger bandwidth at increased centre frequencies, require resonators having larger coupling coefficient and improved thermal stability. These applications also require a drastic reduction of the volume of the devices in order to allow for the decrease of the equipment size.

The new materials of crystal class 32 can greatly surpass quartz for all applications requiring a high coupling coefficient, lower losses, reduced phase noise, very low dimensions, lower impedance level, high temperature operations (5,6).

Recently has been observed that berlinite, lithium tetraborate, gallium phosphate and langasite can fulfil these conditions, being similar or better than quartz for other very important properties such as acoustic losses. They are particularly useful for new applications in telecommunication equipment (2).

The use of lithium tantalate and berlinite, piezoelectrics with zero temperature frequency coefficient, is limited by the growth difficulties and high cost.

The larger coupling coefficients of langasite (LGS) allow to obtain much larger bandwidth for the filters and permit also to use overtones with much higher ranks than for quartz. LGS is very promising for getting extremely high Q factor devices in applications requiring very low phase noise.

Monolithic filters on Y-cut langasite crystal can operate in shear mode over the frequency range from 5 MHz to 18.5 MHz with relative pass bandwidth of 0.3% to 0.8% from the central frequency of the filter (1,3,7).

The basic characteristics of the langasite crystal compare favourably with those of quartz. Their resonance interval are 2-3 times larger, the dynamic inductance and dynamic resistance decrease by 6-18 and 2-6 times respectively, the squared electromechanical coefficient grows by 2-3 times, the losses by SAW propagation drop by 2 times, the dimensions of LGS elements being smaller by 20-30%.

### FILTER DESIGN

The design process of the monolithic filters, based on the acoustic controlled coupling of two or more resonators realised on the Y-cut langasite plate, follows three steps: synthesis, analysis and two-pole filter design.

This design method is useful for the band-pass filters having any pole numbers, for central frequencies between 5 and 18.5 MHz and bandwidth from 0.3% to 0.8% from central frequency.

We propose an eight-pole monolithic filter realised by the capacitive interconnection of four two-pole monolithic filters (8,9).

The polynomial design method was chosen for this application due to the rigorous control of the reflection coefficients at the input and output with respect to the terminating impedance. This method also allows a smaller pass-band ripple for a given attenuation characteristic. In order to rich the desired shape factor, the Chebyshev approximation was used for the attenuation characteristic (10,11,12).

The structure of the eight-pole wide-band filter composed of capacitively coupled two-pole, symmetrical monolithic crystal filters is shown in Figure 1 and the modified eight-pole ladder filter is shown in Figure 2.

**Analysis**

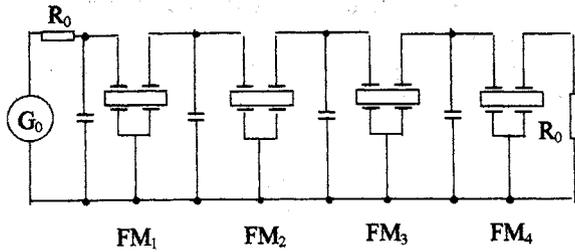


Figure 1. Eight-pole filter structure.

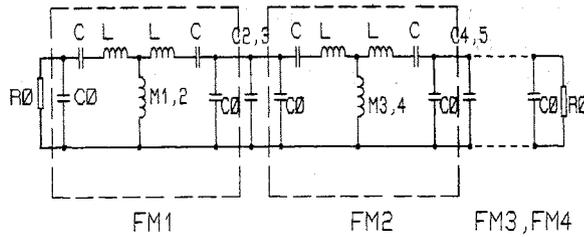


Figure 2. Eight-pole ladder filter.

The design specifications used in this paper, proper to a wide-band band-pass filter, are:

Center frequency	$f_0$	10.7 MHz
3 dB bandwidth	$df$	50 kHz
Insertion loss	$A_0$	2 dB
Passband ripple	$A_{max}$	2 dB
Stop band/attenuation		$\pm 65$ kHz/80 dB
Load resistance	$R_0$	910 $\Omega$
Pole number	$n$	8
Capacitance ratio (Y-cut langasite)	$r$	68.6

**Synthesis**

The synthesis of the monolithic crystal filters, identical to that used for the design of discrete crystal filters, has the following steps:

- convert the given band-pass filter requirement into a normalized low-pass specification;
- the low-pass filter prototype is transformed into a narrow-band band-pass filter with equal inductances and with frequency independent reactances in the series arms;
- the narrow-band band-pass filter is changed into a polynomial filter;
- the normalized low-pass parameters are transformed into the required band-pass parameters.

The relations used for these steps of synthesis of eight-pole monolithic filter are presented in (8),(10),(11).

The response attenuation characteristics of the synthesized band-pass filter were examined by using the values of the circuit elements obtained from the filter synthesis in the PSpice circuit analysis program.

The transforming a prototype low-pass filter in a band-pass filter may lead to insertion loss distortion of the attenuation characteristic of the filter. In these terms the optimization of the available band-pass structure is a better solution instead of a succession of the re-designs of the filter.

The "OPTATEN" program, shown in Figure 3, changes iteratively the coupling capacitor values and / or the source and the load resistors of the filter through a minimisation of error criterion (13).

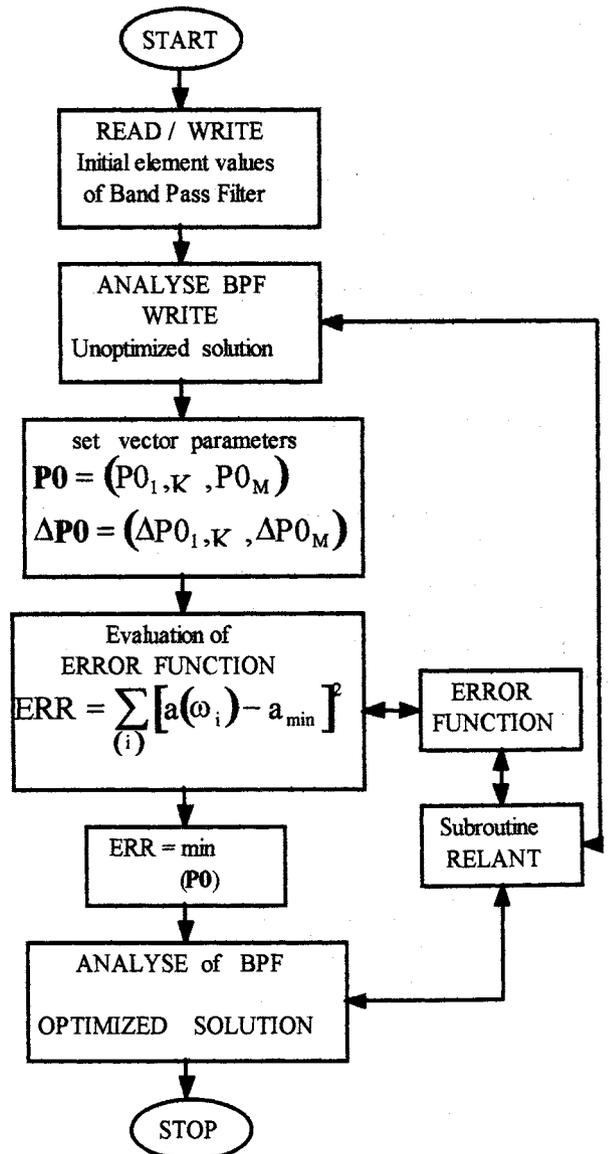


Figure 3. The main steps in OPTATEN program

The parameter  $\mathbf{PO}$ , whose components  $\mathbf{PO}(i)$ ,  $i = \overline{1, M}$  are the nominal values of variable elements of the filter, is iteratively modified, so that  $\mathbf{PO} = \mathbf{PO} + \Delta\mathbf{PO}$ . The initial modifications  $\Delta\mathbf{PO}(i)$ ,  $i = \overline{1, M}$  of the elements starting from their nominal values are controlled by the minimisation process of the error-function

$$\text{ERR} = \sum_{(i)} [a(\omega_i) - a_{\min}]^2.$$

In Figure 4 the optimisation results of the insertion loss characteristics of the monolithic band-pass filter with  $n=8$  are presented.

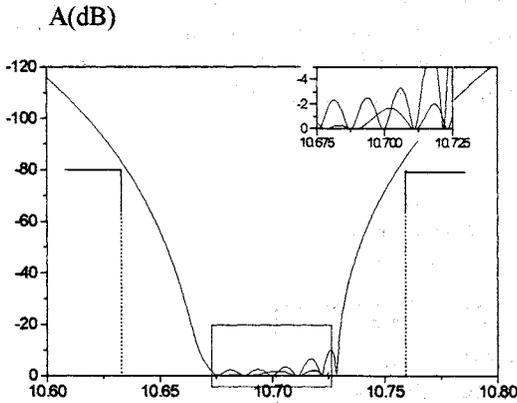


Figure 4. Attenuation characteristic of eight-pole monolithic filter

### Two-pole structure design

From the relations between the elements of the equivalent electrical circuit and the physical properties of the monolithic structure, we calculated the geometry of the individual two-pole filters ( see Figure 3 ), namely, the plate thickness  $t$ , electrodes dimensions  $l_x$  and  $l_z$ , and electrode spacing  $d$ .

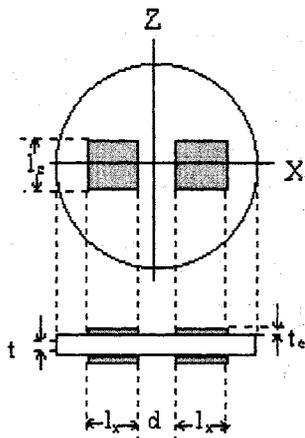


Figure 5. Two-pole monolithic filter layout.

The inter-resonator coupling has been chosen along the X-axis due to the higher coupling coefficient in X direction (7).

The resonance frequency of the plate and the plate thickness are related by

$$f_0 = \frac{N_m}{t}$$

where  $f_0$  is the resonance frequency of the plate ( MHz ),  $N_m$  is the frequency constant of the  $m$ th harmonic ( 1.381 MHz mm for fundamental Y-cut langasite ), and  $t$  is the plate thickness ( in mm ) (7). Using this equation we find that the required plate thickness is  $t=0.129$  mm.

Optimum electrode length along X-axis coupling direction (7) is

$$l_x = (8 \div 10) \cdot t$$

relation established from the energy trapping criterion and the condition of unwanted response suppression.

For  $l_x=9t$  we get 1.161 mm. For the case of Y-cut langasite, the expression of the resonator motional inductance  $L$  is given by Sakharov et al (7)

$$L = 4.5 \cdot \frac{2 \cdot t}{A} \quad (\text{mH})$$

where  $A=l_x \cdot l_z$  is electrode area ( in  $\text{mm}^2$  ). Using the inductance value obtained by synthesis, the electrode area  $A$  and  $l_z$  are determined. For our specification  $l_z=1.88$  mm.

**Calculation of the electrode spacing.** The dependences between coupling coefficient and electrode spacing can be deduced with the relation, currently used in literature, between inter-resonator coupling and the dimensions of the two-pole monolithic filter structure (7), established by analogy with Beaver's equation:

$$k_{i,i+1} = \alpha \cdot \left( \frac{2l_x + d}{t} \right)^{-2} \exp \left( -\beta \sqrt{\frac{\Delta}{f_0}} \cdot \frac{d}{t} \right)$$

where  $\alpha=3.05$ ;  $\beta=2$

$$\frac{\Delta}{f_0} = \frac{4}{\pi^2} \cdot k_{26}^2 + \frac{c_{11}}{c_{66}} \cdot \frac{\rho_e \cdot t_e}{\rho \cdot t}; \quad R = 2 \cdot \frac{\rho_e \cdot t_e}{\rho \cdot t}$$

$k_{26}$  - electromechanical coupling coefficient

$\rho_e, \rho$  - electrode and substrate material density

$t_e$  - one half the electrode thickness

$t$  - langasite plate thickness

In Figure 6 is presented the dependence between the coupling coefficient and electrode spacing.

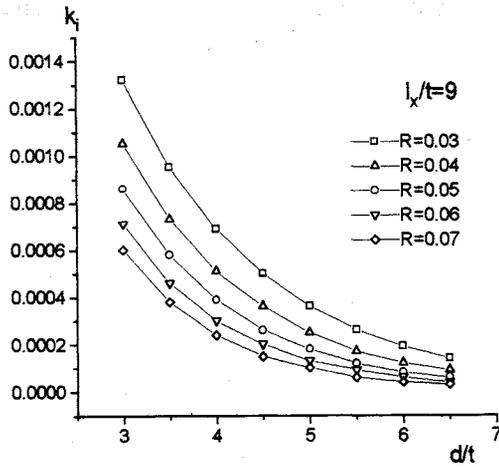


Figure 6. Inter-resonator coupling versus electrode spacing

The relation between electrode spacing and plate-back  $R$  is shown in Figure 7.

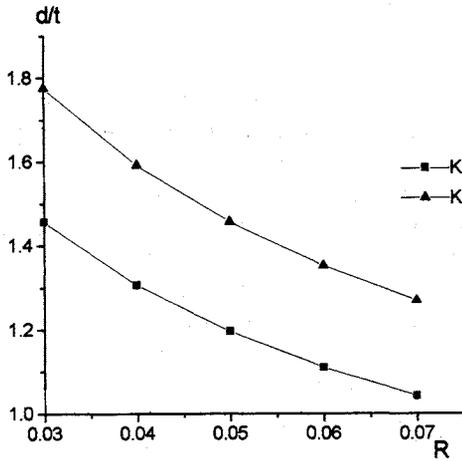


Figure 7. Electrode spacing versus plate-back

The previous semi-empirical relations have been compared with more elaborated calculations based on original approach.

We computed the inter-resonator coupling coefficient of the elastically coupled two-pole filter using the Tiersten's description of the monolithic structure (14,15). The method is based on the solution of the wave propagation along  $X$ -axis in an infinite piezoelectric plate and takes into account the influence of the electroded and unelectroded part of the plate on the wave propagation. We considered the narrow frequency range near the cut-off frequency of the thickness-shear fundamental mode.

The inter-resonator coupling coefficient can be calculated as a relative distance between the two lowest frequencies  $\Omega_1$  and  $\Omega_2$ :

$$k_i = 2 \cdot \frac{\Omega_2 - \Omega_1}{\Omega_1 + \Omega_2}$$

where  $\Omega_1$  and  $\Omega_2$  are the two lowest roots of the equation

$$\begin{vmatrix} H_{11} & H_{12} & 0 & 0 & 0 & H_{16} & 0 & 0 \\ 0 & H_{22} & H_{23} & 0 & 0 & H_{26} & H_{27} & 0 \\ 0 & 0 & H_{33} & H_{34} & 0 & 0 & H_{37} & H_{38} \\ 0 & 0 & 0 & H_{44} & H_{45} & 0 & 0 & H_{48} \\ H_{51} & H_{52} & 0 & 0 & 0 & H_{56} & 0 & 0 \\ 0 & H_{62} & H_{63} & 0 & 0 & H_{66} & H_{67} & 0 \\ 0 & 0 & H_{73} & H_{74} & 0 & 0 & H_{77} & H_{78} \\ 0 & 0 & 0 & H_{84} & H_{85} & 0 & 0 & H_{88} \end{vmatrix} = 0$$

in the range of

$$1 - R - 4 \frac{\bar{k}_{26}^2}{\pi^2} < \Omega < 1$$

where

$$\begin{aligned} H_{11} &= 1 & H_{12} &= -\cos\left(\xi \frac{2l_x}{t}\right) \\ H_{16} &= \sin\left(\bar{\xi} \frac{2l_x}{t}\right) & H_{22} &= \cos\left(\bar{\xi} \frac{2l_x}{t}\right) \\ H_{23} &= -\cosh\left(\xi \frac{2d}{t}\right) & H_{26} &= \sin\left(\bar{\xi} \frac{2l_x}{t}\right) \\ H_{27} &= \sinh\left(\xi \frac{2d}{t}\right) & H_{33} &= \cosh\left(\xi \frac{2d}{t}\right) \\ H_{34} &= -\cos\left(\bar{\xi} \frac{2l_x}{t}\right) & H_{37} &= -\sinh\left(\xi \frac{2d}{t}\right) \\ H_{38} &= \sin\left(\bar{\xi} \frac{2l_x}{t}\right) & H_{44} &= \cos\left(\bar{\xi} \frac{2l_x}{t}\right) \\ H_{45} &= -1 & H_{48} &= \sin\left(\bar{\xi} \frac{2l_x}{t}\right) \\ H_{51} &= -\xi & H_{52} &= -\bar{\xi} \sin\left(\bar{\xi} \frac{2l_x}{t}\right) \\ H_{56} &= -\bar{\xi} \cos\left(\bar{\xi} \frac{2l_x}{t}\right) & H_{62} &= -\bar{\xi} \sin\left(\bar{\xi} \frac{2l_x}{t}\right) \\ H_{63} &= \xi \sinh\left(\xi \frac{2d}{t}\right) & H_{66} &= \bar{\xi} \cos\left(\bar{\xi} \frac{2l_x}{t}\right) \\ H_{67} &= \xi \cosh\left(\xi \frac{2d}{t}\right) & H_{73} &= \xi \sinh\left(\xi \frac{2d}{t}\right) \\ H_{74} &= -\bar{\xi} \sin\left(\bar{\xi} \frac{2l_x}{t}\right) & H_{77} &= -\xi \cosh\left(\xi \frac{2d}{t}\right) \\ H_{78} &= -\bar{\xi} \cos\left(\bar{\xi} \frac{2l_x}{t}\right) & H_{84} &= -\bar{\xi} \sin\left(\bar{\xi} \frac{2l_x}{t}\right) \\ H_{85} &= -\xi & H_{88} &= \bar{\xi} \cos\left(\bar{\xi} \frac{2l_x}{t}\right) \end{aligned}$$

$\bar{\xi}$  is the normalized wave number in the electroded part of the plate

$$\bar{\xi}^2 = \frac{\pi^2}{2(\bar{\gamma}_{11} + \bar{\kappa}_6^2(1+2R))} \cdot \left(1 + 2R - \frac{4\bar{k}_{26}^2}{\pi^2}\right) \cdot \left(\Omega - 1 + R + \frac{4\bar{k}_{26}^2}{\pi^2}\right)$$

and  $\xi$  is the normalised wave number in the unelectroded part of the plate

$$\xi^2 = \frac{\pi^2}{2(\gamma_{11} + \kappa_6^2)}(1 - \Omega)$$

where

$$\kappa_6^2 = \frac{\pi^2}{12}$$

$$\bar{\kappa}_6^2 = \frac{\pi^2}{12} \left(1 + R - \frac{8\bar{k}_{26}^2}{\pi^2}\right)$$

$$\bar{k}_{26}^2 = \frac{e_{26}^2}{c_{66}\epsilon_{22} + e_{26}^2}$$

$$\bar{\gamma}_{11} = \frac{\gamma_{11}}{c_{66}} = \frac{1}{c_{66}} \left(c_{11} + \frac{c_{12}^2}{c_{22}} + \frac{c_{14}^2}{c_{44}}\right)$$

$$\Omega = \frac{\omega}{\omega_0}$$

where

$$\omega_0 = \frac{\pi}{t} \sqrt{\frac{c_{66}}{\rho} \left(1 + \frac{e_{26}^2}{c_{66}\epsilon_{22}}\right)}$$

The dependence of the inter-resonator coupling coefficient  $k_x$  on the ratio between electrode spacing and plate thickness with plate-back  $R$  as parameter is presented in Figure 8.

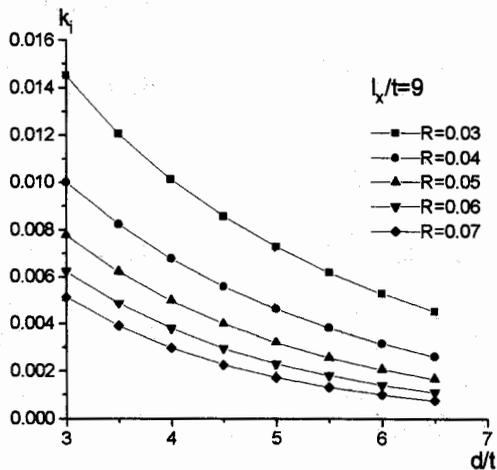


Figure 8. Inter-resonator coupling versus electrode spacing

In Figure 9 the relationship between the electrode spacing and plate-back  $R$  is shown for the two cases of

$K_{1,2}=K_{7,8}$  and  $K_{3,4}=K_{5,6}$  corresponding to the two cases of  $FM_1, FM_4$  and  $FM_2, FM_3$  respectively.

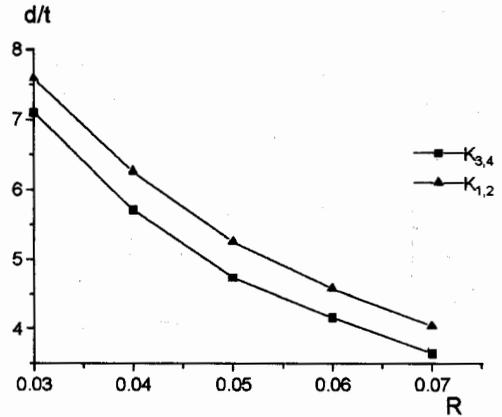


Figure 9. Electrode spacing versus plate-back

### CONCLUSIONS

We have developed the design methods for langasite monolithic filter composed from the elastically coupled two resonators structure.

The main features are the optimisation of the attenuation characteristic and a calculation of the coupling coefficient using the Tiersten's theory.

The comparative analysis of the results (Figure 6,7 and Figure 8,9) shows a good qualitative agreement between the semi-empirical description and the theoretical model.

A quantitative agreement could be obtain by adjusting of the theoretical model parameters.

For a further evaluation of the two alternative methods the analysis is to be extended for usual geometrical parameters of the structure and especially electrode thickness.

Final quantitative conclusions will be done when experimental results available.

### ACKNOWLEDGMENTS

We are indebted to Constantin Bora and Ioan Mateescu for the continuous assistance during elaboration of this paper, to Gabriela Piroasca for the help with data processing and to Dr. Mihai Popescu for the useful discussions.

### REFERENCES

1. M. F. Dubovnik, I. A. Andreyev, Yu. S. Shmaly, 1994, "Langasite ( $La_3Ga_5SiO_{14}$ ) an optical piezoelectric: growth and properties", Proc. of IEEE Frequency Control Symposium, 43-47

2. J. Détaint, J. Schwartzel, A. Zarka, B. Capelle, J. P. Denis, E. Philippot, 1994, "Bulk wave propagation and energy trapping in the new thermally compensated materials with trigonal symmetry", Proc. of IEEE Frequency Control Symposium, 58-71
3. A. Ballato, J.G. Gultieri, 1994, "Piezoelectric resonator materials", Proc. of IEEE Frequency Control Symposium, 22-34
4. N. Gotalskaya, D. I. Drezin, V.V. Bezelkin, V. N. Stassevich, 1993, "Peculiarities of technology, physical properties and applications of new piezoelectric material Langasite ( $\text{La}_3\text{Ga}_5\text{SiO}_{14}$ )", Proc. of IEEE Frequency Control Symposium, 339-347
5. J. Detaint, J. Schwartzel, 1994, "Filtering and frequency control for the next generations of radiocommunication systems", Proc. of European Frequency and Time Forum, 175-189
6. S.A. Sakharov, I.M. Larionov, I.V. Issaev, 1991, "Monolithic filters using strong piezoelectrics", Proc. of 45th AFCS, 181-183
7. S. A. Sakharov, I. M. Larionov, A. V. Medvedev, 1992, "Application of langasite crystals in monolithic filters operating in shear mode", Proc. of IEEE Frequency Control Symposium, 713-723
8. Irina Mateescu, John Kosinski, 1993, "Eight-pole monolithic filters using lithium tantalate", Proc. of IEEE Frequency Control Symposium, 620-625
9. Jiri Zelenka, Irina Mateescu, John Kosinski, 1994, "The effective coupling of the elastically coupled Y-cut  $\text{LiTaO}_3$  two and three resonator structure", 786-791
10. A.I. Zverev, 1967, "Handbook of filter synthesis", John Wiley, New York, USA
11. V.M. Kantor, 1977, "Monolitniie piezoelectriceskie filtri", Moskva, Russia
12. R.G. Kinsman, 1987, "Crystal filters", John Wiley, New York, USA
13. Alexandru Serbanescu, 1994, "Computer-aided design in transmission technics", Tehnical Military Academy, Bucharest, Romania
14. H. F. Tiersten, 1969, "Linear piezoelectric plate vibration", Plenum Press, New York, USA
15. Jiri Zelenka, 1986, "Piezoelectric resonators and their applications", Academia Prague, Czech Republic

## THERMOLUMINESCENCE OF SWEEPED AND DOPED IRRADIATED QUARTZ CRYSTALS

M. SMAALI\*, J. J. BOY\*\* and P. ZECCHINI\*

\*Université de Franche-Comté, UFR des Sciences et Techniques, Laboratoire de Cristallographie et Chimie Minérale, 16 route de Gray, 25030 Besançon CEDEX, France.

\*\*Ecole Nationale Supérieure de Mécanique et Microtechniques, Laboratoire de Chronométrie, Electronique et Piezoélectricité, Chemin de l'Epithaphe, 25030 Besançon CEDEX, France.

### ABSTRACT

X-ray or  $\gamma$ -radiations product paramagnetic defect centers in quartz material, which are associated to aluminum as impurity by substitution of silicon atoms, during the crystal growth.

Annealing the material after irradiation restores the as-grown defect centers ( $[AlO_4/M^+]$  where  $M^+$  is an interstitial alkali ion and / or  $[AlO_4/H^+]$ ). In order to understand and to explain this phenomenon of restoration, the study of the thermoluminescence is used. It consists to measure the current of the emitted radiations during the annealing process of irradiated quartz crystals. These radiations are due to electron releases which induce an emission peak at a characteristic temperature.

Many samples of natural and synthetic quartz crystals are submitted to different pre-treatments such as sweeping (into and out), or/and  $\gamma$ -irradiation doses. Investigations by M.I.R spectroscopy and I.C.P analyses are also applied, before and after each treatment.

Strong differences have been observed between unswept and hydrogen swept crystals of synthetic material. No effect of alkali exchange under high electric field, when it has been performed in air. So thermoluminescence leads to characterize two families of peaks, one at 550 K for natural quartz crystals, one at 470 K for synthetic quartz crystals.

### INTRODUCTION

Thermoluminescence is used as a characterization technique of minerals, and more particularly for the study of the kind of defects and their state of recombinations in minerals. For several years, quartz material has been the subject of many studies using various techniques in order to characterize the different defects within the quartz lattice. Thermoluminescence is one of these techniques which are applied in quartz research domain (McKeever (1), Bohm et al (2)). This phenomena was presented at first by Randall and Wilkins (3); it was observed during the heating of an irradiated material.

In our study, this technique has been used in order to determine at first which defects are responsible of the glow peaks. Many natural and synthetic samples were submitted to electrodiffusion exchange along the z-

axis, or by sodium, or by lithium, or by hydrogen ions. Then, an attempt to determine both the activation energies of each trap and the lifetime of defects created during irradiation, by using the models which are proposed in the literature. Before recording TL spectra, all samples are irradiated with  $\gamma$ -ray dose until they are saturated. Infrared spectroscopy was used in order to check the effect of all the pretreatments, sweeping and irradiation.

### EXPERIMENTAL PROCEDURE

Synthetic quartz crystals from three different growths, named as QS-A, QS-D and QS-H, and natural crystals from Brazil, are used in this study. All the crystals but one of each origin, have been submitted to an electric field. The QS-A and QS-D samples have been swept in air or in hydrogen atmosphere. In the QS-H and natural crystals, lithium or sodium ions are introduced in order to replace interstitial impurities (except hydrogen atoms). All the experiments are performed along the z-axis at 740 K in a 1000 V/cm electric field. For more details about the equipment used for the electrolytic exchange experiment see for instance Stegger (4). The end of the ion exchange is detected when a constant intensity of current is measured. After each run, the z-faces of the samples are ground and polished. Infrared spectroscopy measurements are also made before and after each electrolytic exchange, (Figure 1).

Prior to the TL measurements, the crystals were irradiated with a  $\gamma$ -ray saturated dose. The results of irradiation are shown by liquid nitrogen temperature infrared measurements, Figure 2.

The TL measurements are performed from room temperature to 700 K, using a 1 degree /s heating rate. The TL emission is detected by an R1104 photomultiplier tube. A corning heat absorbing filter is placed between the sample and the photomultiplier tube in order to reduce the blackbody radiations which are not negligible near 700 K. The photomultiplier out put signal is amplified and recorded by means of a high sensibility digital multimeter which is connected to an IBM computer.

The final TL spectra are obtained by subtracting the measured background from the measured spectra.

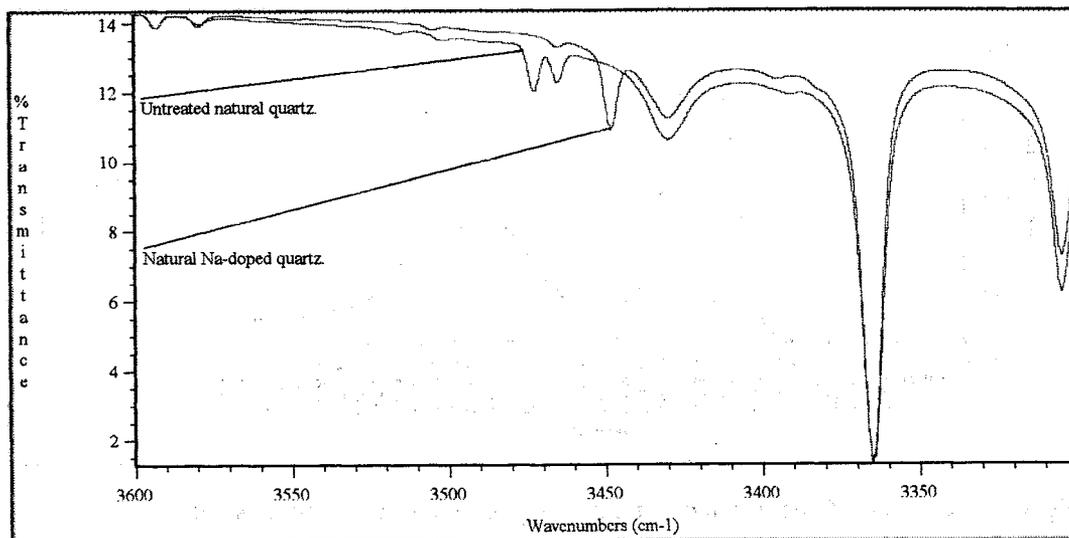


Figure 1: Infrared spectra of natural quartz before and after Na electrolytic exchange.

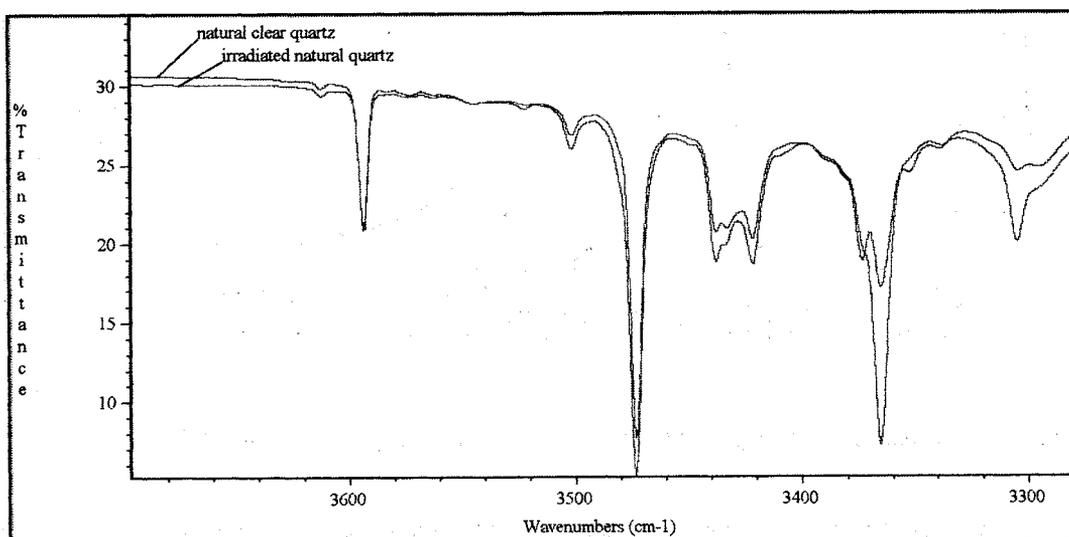


Figure 2: Infrared spectra of natural quartz before and after a  $\gamma$ -irradiation.

## RESULTS AND DISCUSSION

Figures 3 and 4 show the TL emission spectra of cultured quartz respectively from QS-D and QS-A growth runs. On each figure, the TL spectra of unswept and swept sample are presented. As can be seen, the TL spectra of unswept samples show two main emission peaks at approximately 470 K and 610 K, when the swept samples give just only one emission peak at 470 K with a low response intensity in regard with the unswept ones. Castiglioni et al (5) have studied the TL emission of unswept and swept cultured quartz, and they have obtained the same results, in contrast with other authors (Yang and McKeever (6),

Martini et al.(7)), who have shown an increase of the TL intensity after sweeping treatments. The differences they observed were due to the different conditions of their treatments (atmosphere and/or temperature). In this study, the electrodiffusion is performed or in air, or in hydrogen atmosphere, so the majority of the interstitial impurities are replaced. After irradiation, the density of holes is much lower than the one of the other centers such the hydrogen centers. Since the TL emission is a consequence of an electronic transition, from the excited state to the ground state, then the hydrogen centers which are created during irradiation cannot participate in this emission. As a consequence of that, the low intensity detected can be explained.

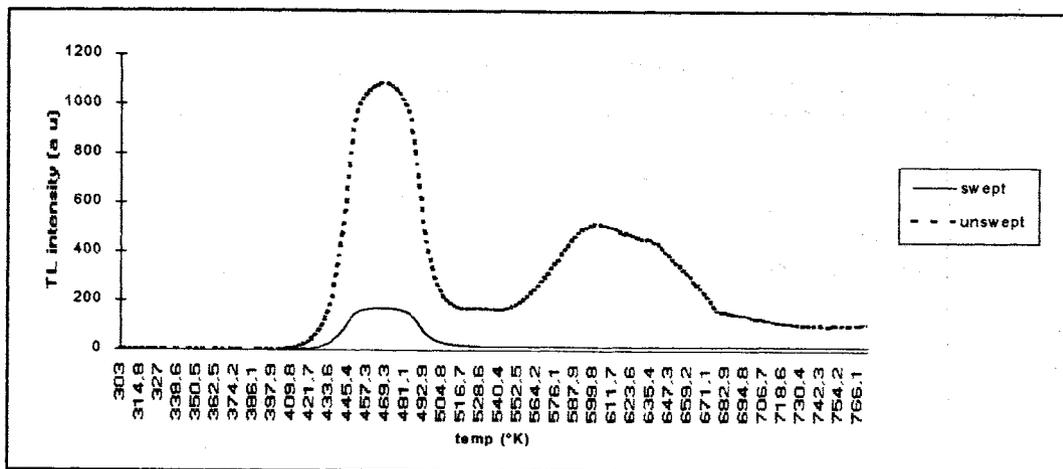


Figure 3: TL spectra of unswept and hydrogen swept synthetic quartz crystals (QS-D).

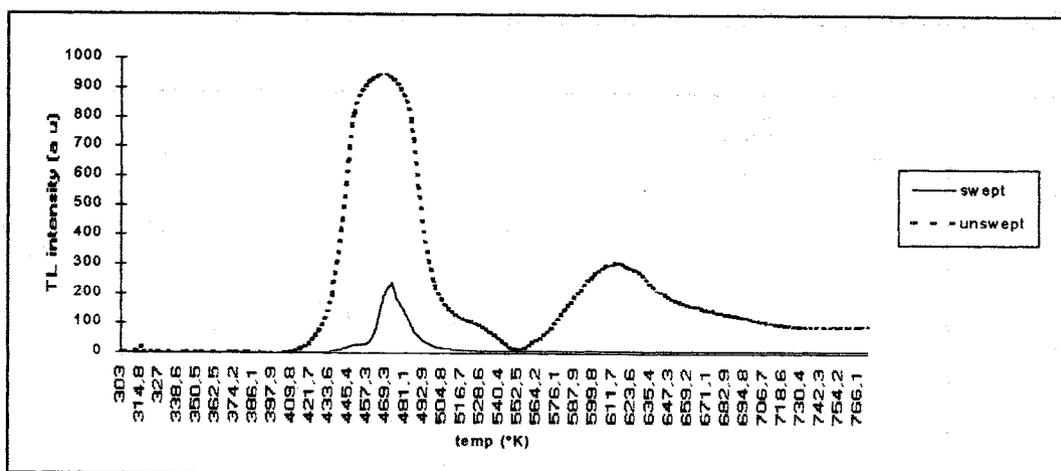


Figure 4: TL spectra of unswept and hydrogen swept synthetic quartz crystals (QS-A).

Figure 5 shows the TL spectra of two samples from the QS-H growth run: one is Li-doped and the other is Na-doped. The curves obtained in this case are very similar to those given by the hydrogen swept samples. This similarity is also observed in the intensity order. When electrolytic exchange is achieved in air, hydrogen cannot be substituted by alkali ions. Nevertheless, it is possible that one part of the hydrogen within the as-grown crystal can replace alkali atoms, especially in the case of Na doped crystals.

Figure 6 shows the spectra of two natural quartz samples which have not received any treatment except

an irradiation dose. The two spectra give a common peak at 560 K with the same intensity, but the spectrum related to natural quartz 1 gives a second peak at 600 K which is like a shoulder on the highest peak when the TL spectrum of the natural quartz 2 gives the second peak at approximately 620 K, and can be easily distinguished from the first peak at 520 K.

The high response intensity obtained with these samples of natural quartz is due exclusively to the high impurities content. The ICP (inductively coupled plasma) analyses have given the concentration values shown in table 1.

TABLE 1: -Concentration of the principal impurities in quartz given in atomic ppm.

quartz origin	Al	Fc	Li	Na
synthetic quartz	1 - 4	—	0.5 - 2	0.5 - 1
natural quartz	15 - 40	0 - 0.5	10 - 15	—

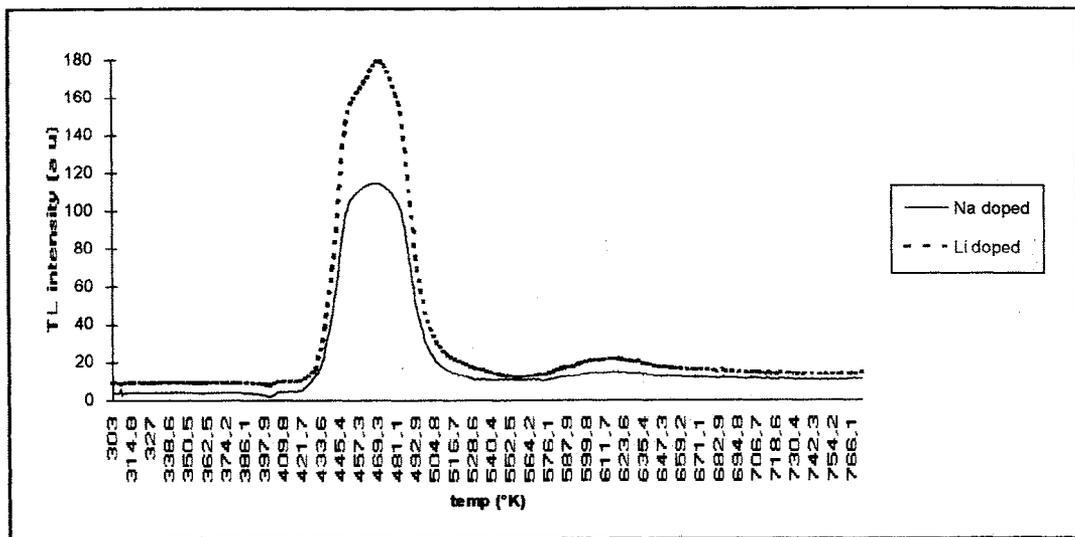


Figure 5: TL spectra of the QS-H cultured doped quartz crystals.

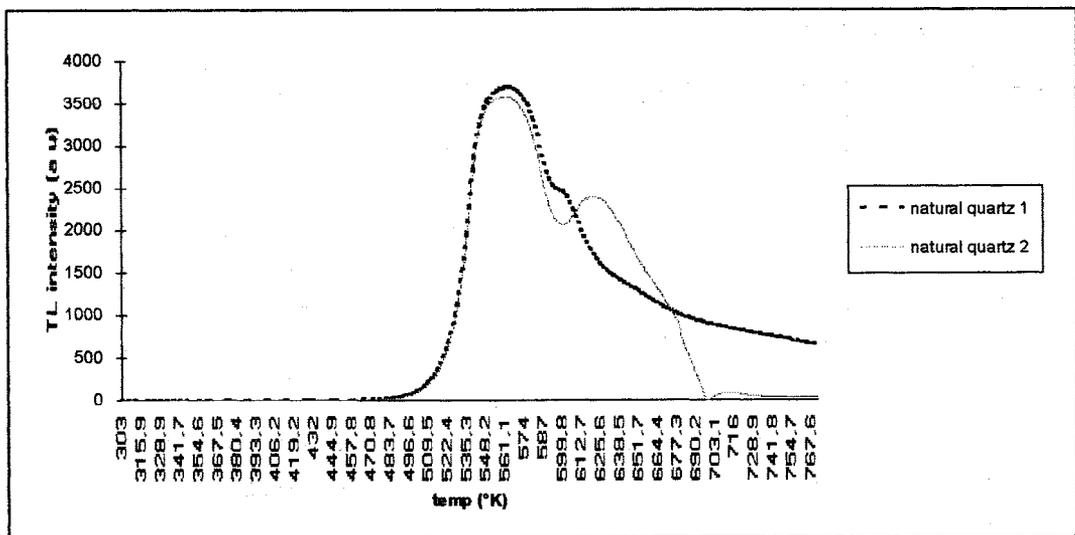


Figure 6: TL spectra of two Brazilian natural quartz crystals.

Figure 7 shows the TL emission spectra obtained with the two natural Li- and Na- doped quartz crystals. An examination of these spectra shows at first a decrease of the intensity by comparison with the spectra in figure 6. The second observation is that the spectra in figures 5 and 7 related respectively to synthetic and natural crystals have the same aspect. In each case Li doped crystal shows a high TL intensity.

An attempt is made to calculate the characteristic factors such activation energy, frequency factor and lifetime of traps responsible on the appeared peaks. The further method, called "peak shape method", uses two remarkable points of the TL spectrum: the

temperature value corresponding to the maximum of the peak ( $T_m$ ), and the highest value of the two temperatures corresponding to the half-peak intensity ( $T_{1/2}$ ). Equations that relate activation energy  $E$  to these parameters, have been developed by Grossweiner (8).

The results obtained are shown in table 2. The values obtained by other authors are reported too.

The low TL peaks mentioned in the literature at 358, 383 and 453 K do not appear in our TL spectra. In this study, the TL measurements are performed few days after the irradiation treatment, and because their short lifetime, they cannot appear here.

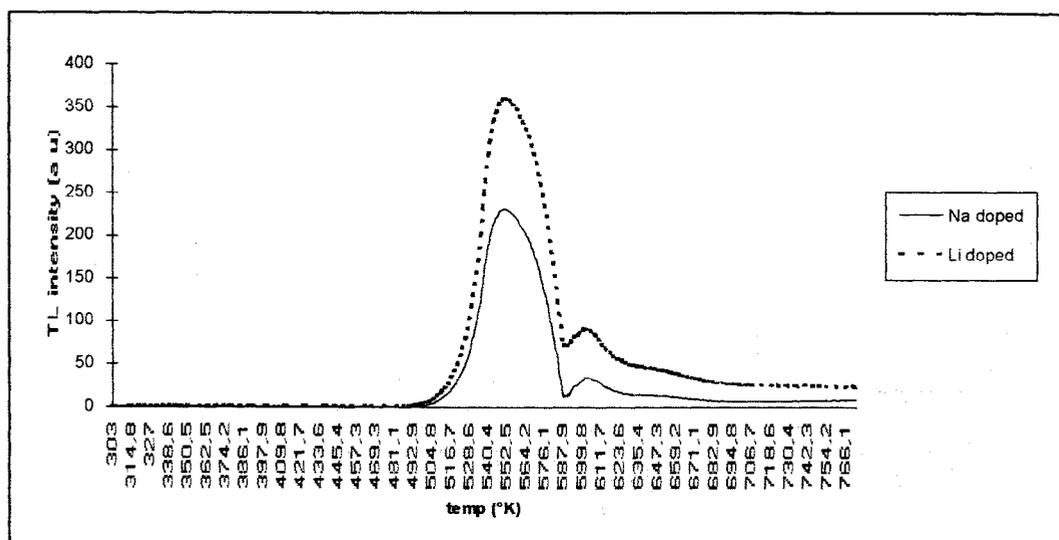


Figure 7: TL spectra of the doped natural quartz crystals.

TABLE 2 -Some values for trapping parameters of quartz TL peaks.

Peak temp (K)	Trap depth (eV)	Frequency factor ( $s^{-1}$ )	Lifetime (years)	Source
358	0.84	$1.9 \cdot 10^{11}$	$0.13 \cdot 10^{-3}$	Aitken (9)
383	0.80	$2.7 \cdot 10^9$	$1.2 \cdot 10^{-3}$	Aitken (9)
383	0.98	$8 \cdot 10^{12}$	$0.8 \cdot 10^{-3}$	Fleming (10)
383	0.99	—	—	Wintle (11)
453	1.19	$1.8 \cdot 10^{11}$	$0.1 \cdot 10^3$	Levy (12)
463	1.42	$3.4 \cdot 10^{14}$	$0.7 \cdot 10^3$	Aitken (9)
470	1.62	$3.4 \cdot 10^{12}$	$4.1 \cdot 10^3$	present study
503	1.79	$5 \cdot 10^{18}$	$130 \cdot 10^3$	Aitken (9)
550	1.49	$2.5 \cdot 10^{12}$	$5.6 \cdot 10^3$	present study
553	1.45	—	$> 3 \cdot 10^6$	Hütt et al (13)
598	1.7	$1 \cdot 10^{14}$	$160 \cdot 10^6$	Wintle (11)

## CONCLUSION

Hydrogen sweeping has strong influence on TL spectra. Strong decreases in TL intensity have been observed, when a comparison is made between those of untreated and those of hydrogen swept quartz crystals. Doping quartz with alkali ions in air gives the same results as hydrogen sweeping.

Now, it needs to complete this study, by changing the electrolytic conditions: sweeping and doping must be made in vacuum or in inert atmosphere.

## REFERENCES

1. McKeever S.W.S., 1984, Thermoluminescence of quartz and silica, Rad.Prot.Dosim. 8, 81-89.
2. Bohm M., Peschke W. and Scharmann A., 1980, Thermoluminescence and thermally stimulated conductivity in quartz induced by thermal treatment after X-irradiation, Rad. eff. 53, 67-72.
3. Randall J.T. and Wilkins M.H.F., 1945, Phosphorescence and electron traps, Proc.R.Soc.Lond. A184, 336-407.
4. Stegger P., 1986, Elektronenspinresonanzuntersuchungen an paramagnetischen Eisen (III)-defecten in  $\alpha$ -quarz, Dissertation, Westfälische Wilhelms-Universität Münster.
5. Castiglioni M., Martini M., Spinolo G. and Vedda A., 1994, Thermally stimulated luminescence (TSL) and conductivity (TSC) of synthetic crystalline quartz, Rad.Meas. 23, (Nos 2/3), 361-365.

6. Yang X.H., and McKeever S.W.S., 1988, Characterization of the pre-dose effect using ESR and TL, Nucl.Tracks.Rad.meas. 14, 75-79.
7. Martini M., Spinolo G. and Vedda A., 1987, Defects dynamics in as grown and electrodiffused quartz: an interpretation of the pre-dose effect, J.Appl.Phys. 61, 2486-2488.
8. Grossweiner L.J., 1963, J.Appl.Phys. 24, 1306.
9. Aitken M.J., 1985, Thermoluminescence dating, Academic Press, London.
10. Fleming S., 1979, Thermoluminescence techniques in archaeology, Clarendon Press, Oxford.
11. Wintle AG., 1975, Thermal quenching of thermoluminescence in quartz, Geophys.J.R.Astron.Soc. 41, 107-113.
12. Levy P.W., 1979, Thermoluminescence studies having applications to geology and archaeometry, PACT, 3, 466-480
13. Hütt G., Smirnov A., and Tale I., 1979, On the application of thermoluminescence of natural quartz to the study of geochronology of dedimentary deposits, PACT, 3, 362-373.

# TEMPERATURE DEPENDENCE OF SH-APM DELAY-LINES ON QUARTZ : DESIGN RULES FOR (BIO)CHEMICAL SENSORS

Corinne DÉJOURS, Dominique REBIÈRE, Jacques PISTRÉ and Roger PLANADE \*

IXL - Université Bordeaux I (URA CNRS n° 846)  
351, Cours de la Libération - 33405 TALENCE Cedex - France  
Phone : (33) 56 84 65 40 - Fax : (33) 56 37 15 45

\* Centre d'Études du Bouchet (DRET/DGA)  
Le Bouchet BP n°3  
91710 Vert le Petit - France

## Abstract

Acoustic wave delay-lines with interdigital transducers (IDTs) have been studied for many years, and nowadays they are used in more and more applications, as various as device signal processing units [1] or chemical sensors [2]. Whatever the field of application, temperature variations are a critical parameter for these devices. Consequently, it is important to know the device behaviour with temperature, in order to define its best characteristics. This was already achieved for surface acoustic wave (SAW) devices [3], and it has been done partially for shear horizontal acoustic plate mode (SH-APM) devices [4].

In this paper, we propose new results in the case of SH-APM delay-lines on quartz substrate. These modes propagate at discrete frequencies that are defined, in the case of free surfaces.

## INTRODUCTION

Acoustic wave delay-line have demonstrated utility in a variety of sensing applications including gas or vapor detection [5] as well as specific species in liquid medium [6]. An acoustic wave delay-line consists of an input transducer for piezoelectrical elastic wave generation, an acoustic path along which the propagating wave interacts the environment and an output transducer for wave detection. A sensitive film, coated on the acoustic propagation path, modifies the wave characteristics. The physical and chemical modifications of the sensitive coating, due to gas sorption for instance, involve variations of the propagation characteristics. These interactions result in changes in wave amplitude and phase delay.

We propose to describe a class of acoustic wave sensors based on a shear horizontal acoustic plate mode (SH-APM). This shear horizontal acoustic plate mode delay-line is very rich, because a great number of modes can propagate, it also allows more applications than SAW devices, as it can work

in a liquid medium. Indeed, SH modes have particle displacement quasi parallel to the device surface and normal to the direction of propagation. The absence of a surface normal component of displacement allows each SH plate mode to propagate in contact with a liquid without coupling excessive amounts of acoustic energy into the liquid.

Another main advantage of the SH-APM devices lies in the existence of a reflection surface free of transducers and free of electrical connections, which allows the generating function to be dissociated from the perturbation function of the waves.

In this paper we describe in details the SH-APM, its propagation characteristics and its responses to temperature variations in order to design biochemical sensors; theoretical studies of temperature dependence are examined and experimental results with various frequency modes for different plate geometries are presented. They show the frequency-temperature sensitivities and bring to light temperature turn over of the various modes.

## 2. THEORETICAL STUDIES

### 2.1 Analysis of APM propagation characteristics

The devices discussed in this paper utilize quartz crystal plates which act as acoustic waveguides. Acoustic waves are generated and received by two interdigital transducers (IDTs) sputtered on one surface of the crystal plate. In the case of simple rotated quartz plate, the generation of various acoustic waves depends on the transducers orientation. When the propagation direction goes along the crystallographic X axis, the IDT generates three bulk waves in addition to the Rayleigh wave (SAW mode). When the direction of propagation is normal to the X axis, only the SH wave is piezoelectrically generated. A sine wave, applied to IDT, provides a shear horizontal bulk wave, which is emitted in a direction  $\theta_e$  from the surface (figure 1). The frequency relation is given by the following expression :

$$f_m = \frac{V(\theta_e)}{\lambda \cos(\theta_e)}$$

where  $\lambda$  is the transducer periodicity,  $V(\theta_e)$  is the SH bulk wave velocity.

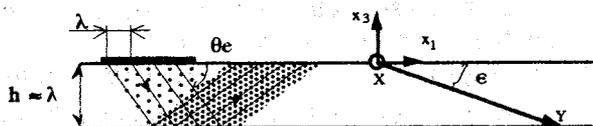


Fig.1 Schematic of SH guided waves.

So an IDT can generate, according to the dispersion relation, SH bulk waves. However, coherent bulk waves can only occur above a cut-off frequency  $f_c$  defined by :

$$f_c = \frac{C_a}{\lambda_0} \sqrt{\frac{C_{44}}{\rho}}$$

$$\text{with } C_a^2 = \left( \frac{C_{66}}{C_{44}} \right) - \left( \frac{C_{46}}{C_{44}} \right)^2$$

where  $C_a$  is the anisotropic factor,  $\rho$  is the piezoelectric density and  $C_{44}$ ,  $C_{66}$  and  $C_{46}$  are stiffnesses in the orthogonal reference system [7].

The SH waves can be guided by successive reflections on the two parallel surfaces of the quartz plate without changing polarization. For a thin plate, the interference between incident and reflected waves produces shear horizontal plate modes. It was also demonstrated [8] that, in a delay line, these modes propagate at discrete frequencies as shown by the relation :

$$f_m = f_c \sqrt{1 + \left( \frac{m\lambda}{2hC_a} \right)^2}$$

where  $h$  is the quartz plate thickness.

### 2.2 Influence of temperature on SH-APM propagation

The temperature dependence of a delay-line is a very important parameter in the design of an acoustic wave device, and in particular with SH-APM sensor applications. Indeed, what is often need for a useful acoustic delay-line sensor is a crystalline orientation, a plate thickness and a propagation direction, having a vanishing first order temperature coefficient of frequency (or time delay). The temperature dependence of SH-APM is derived from the analytical expression of the  $m^{\text{th}}$  frequency. This calculation is similar to the method used for temperature studies of SAW devices [9]. The relation between frequency and temperature can be expressed by the following relation :

$$f_m(T) = \frac{C_a(T)}{\lambda(T)} \sqrt{\frac{C_{44}(T)}{\rho(T)} \left[ 1 + \left( \frac{m \cdot \lambda(T)}{2 \cdot h(T) \cdot C_a(T)} \right)^2 \right]}$$

Using a Taylor development, the temperature coefficients of elastic constants, density, thickness and periodicity permit to predict temperature dependence of frequency. In the case of Quartz, material coefficients have been extracted from Bechmann [7], Jaillet [9] and Michel [10]. The temperature coefficient values ( $i^{\text{th}}$  order,  $i = 1..3$ ) are given in table 1 where  $\alpha_{C_i}^{(i)}$  are the elastic temperature coefficients,  $\alpha_{\rho}^{(i)}$  are the density temperature coefficients,  $\alpha_X^{(i)}$  is the thermal expansion of the X crystallographic axis,  $\alpha_Y^{(i)}$  is the thermal expansion of the Y crystallographic axis,  $\alpha_Z^{(i)}$  is the thermal expansion of the Z crystallographic axis.

i	$\alpha_{C11}^{(i)}$	$\alpha_{C12}^{(i)}$	$\alpha_{C13}^{(i)}$	$\alpha_{C14}^{(i)}$	$\alpha_{C33}^{(i)}$	$\alpha_{C44}^{(i)}$	$\alpha_{C66}^{(i)}$
1 ( $10^{-6}$ )	-48,5	-3 000	-550	101	-160	-177	178
2 ( $10^{-9}$ )	-107	-3 050	-1 150	-48	-275	-216	118
3 ( $10^{-12}$ )	-70	-1 260	-750	-590	-250	-216	21

i	$\alpha_{\rho}^{(i)}$	$\alpha_x^{(i)}$	$\alpha_y^{(i)}$	$\alpha_z^{(i)}$
1 ( $10^{-6}$ )	-34,9	13,71	13,71	7,48
2 ( $10^{-9}$ )	-15,08	6,5	6,5	2,9
3 ( $10^{-12}$ )	-49,91	-1,9	-1,9	-1,5

Table 1 : Temperature coefficients of elastic constants, thermal expansion coefficients and temperature coefficients of density - Quartz [7, 8, 9].

Temperature dependence of elastic constants is described by the expression :

$$C_{ij}(T) = C_{ij}(T_0) \left[ 1 + \alpha_{C_{ij}}^{(1)}(T - T_0) + \alpha_{C_{ij}}^{(2)}(T - T_0)^2 + \dots \right]$$

where  $T_0$  is the reference temperature ( $25^\circ\text{C}$ ),  $I = 1..6$  and  $J = 1..6$ .

A similar development permits to obtain the temperature dependence of density :

$$\rho(T) = \rho(T_0) \left[ 1 + \alpha_{\rho}^{(1)}(T - T_0) + \alpha_{\rho}^{(2)}(T - T_0)^2 + \dots \right]$$

Our study is based only on singly rotated quartz, with  $x_1$  perpendicular to the X axis. So, we can notice that  $x_1$  and  $x_3$  axis are in the plane YZ. A direct consequence is that  $\varphi$  and  $\psi$  (Euler angles) are always equal to zero.

In this calculation, we assume that  $\Theta$  is temperature variations independant. In conclusion crystal cuts are supposed to be invariable to temperature variations.

According to table 1, the thermal expressions of plate thickness  $h(T)$  and periodicity  $\lambda(T)$  are given by table 2.

$$\lambda(T) = \lambda(T_0) \left[ 1 + \alpha_{\lambda}^{(1)}(T - T_0) + \alpha_{\lambda}^{(2)}(T - T_0)^2 + \dots \right]$$

$$\text{with } \alpha_{\lambda}^{(1)} = \alpha_Y^{(1)} \cos^2 \Theta + \alpha_Z^{(1)} \sin^2 \Theta$$

$$\alpha_{\lambda}^{(2)} = \alpha_Y^{(2)} \cos^2 \Theta + \alpha_Z^{(2)} \sin^2 \Theta + \frac{1}{2} \left( \alpha_Y^{(1)2} + \alpha_Z^{(1)2} \right) \sin^2 \Theta \cos^2 \Theta$$

and

$$h(T) = h(T_0) \left[ 1 + \alpha_h^{(1)}(T - T_0) + \alpha_h^{(2)}(T - T_0)^2 + \dots \right]$$

$$\text{with } \alpha_h^{(1)} = \alpha_Y^{(1)} \sin^2 \Theta + \alpha_Z^{(1)} \cos^2 \Theta$$

$$\alpha_h^{(2)} = \alpha_Y^{(2)} \sin^2 \Theta + \alpha_Z^{(2)} \cos^2 \Theta + \frac{1}{2} \left( \alpha_Y^{(1)2} + \alpha_Z^{(1)2} \right) \cos^2 \Theta \sin^2 \Theta$$

Table 2 : Thermal expressions of plate thickness and periodicity - Quartz.

### 2.3 Theoretical results

Frequency shift variations, induced by temperature are represented on Fig. 2. The evolution is close to parabolic shape. Different turnover temperature points can be seen when the considered mode changes. In order to understand the sensitivity of SH-APM propagation, we have represented the evolution of the anisotropic factor (Ca) versus temperature with  $\Theta$  as parameter (Fig. 3). These representation shows that there is a weak dependence of anisotropic factor with temperature.

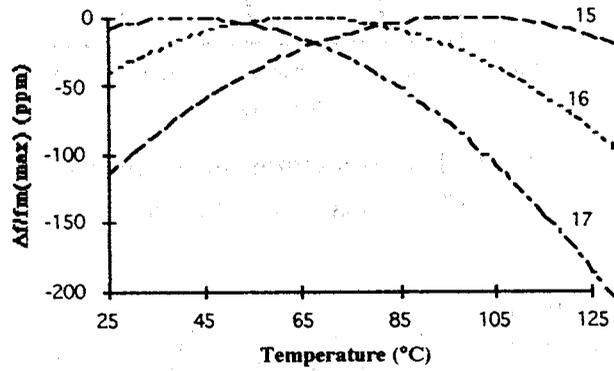


Fig. 2 : frequency shift versus temperature BT-cut ( $\lambda_0=34.4 \mu\text{m}$  and  $h=0.5 \text{ mm}$ ).

But if we are looking for the derivative expression ( $dCa/dT$ ), we can notice that there is two particular points where the derivative is equal to zero. These two points correspond to  $\Theta$  equal to  $40^\circ$  and  $125^\circ$  respectively.

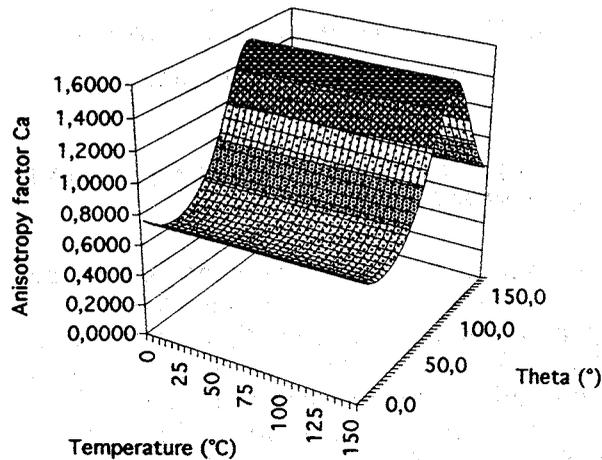


Fig. 3 : Anisotropic factor of quartz versus temperature with  $\Theta$  as parameter.

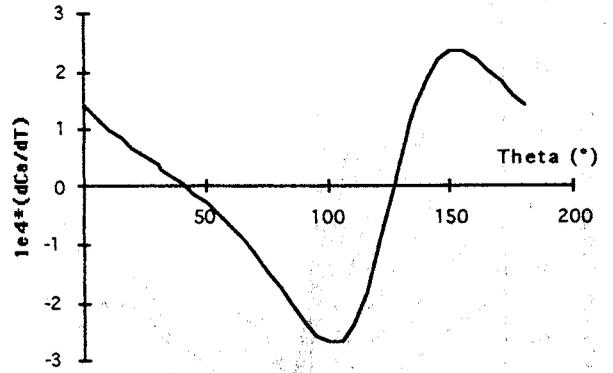


Fig. 4 :  $dCa/dT$  versus  $\Theta$ .

In the case of a particular crystal cut (quartz-ST,  $\lambda=32 \mu\text{m}$ ,  $h=0.5 \text{ mm}$ ), Fig. 5 shows the influence of the mode number on the turnover temperature.

It is interesting to generalize this approach to a large number of singly rotated quartz cuts. The results are reported in Figs. 6 and 7. In Fig. 6 it can be seen that for a specific device geometry ( $\lambda=32 \mu\text{m}$ ,  $h=0.5 \text{ mm}$ ) a particular point appears near  $\Theta=40^\circ$  and  $\Theta=125^\circ$  where all modes have a turnover temperature value about  $10^\circ\text{C}$ . These particular cuts correspond to  $dCa/dT$  equal to zero.

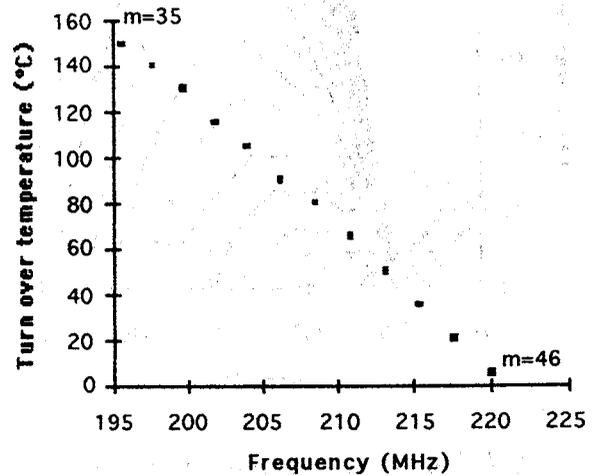


Fig. 5 : Turnover temperature versus mode frequencies ST cut-quartz,  $\lambda=32 \mu\text{m}$ ,  $h=0.5 \text{ mm}$ .

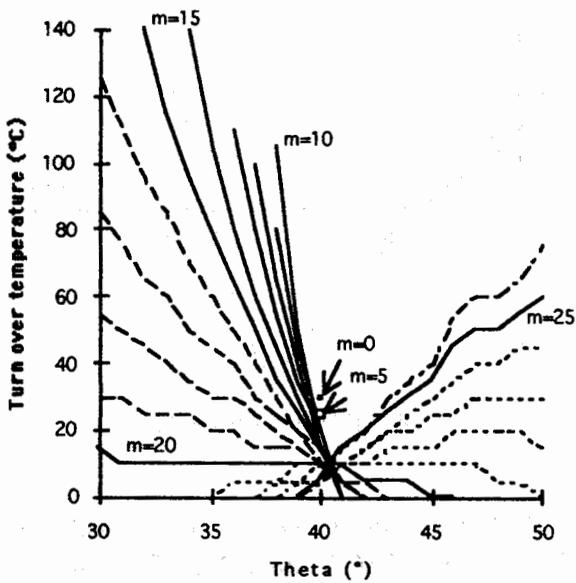


Fig. 6 : Turnover temperature versus  $\Theta$   
ST cut - quartz,  $\lambda=32 \mu\text{m}$ ,  $h=0.5 \text{ mm}$ .

temperature for the considered mode. Indeed, on Fig. 8, the slope of curves is about  $1^\circ\text{C}/\mu\text{m}$  ( $h : 100 \text{ to } 700 \mu\text{m}$ ) ; on Fig. 9, the slope is inferior to  $20^\circ\text{C}/\mu\text{m}$  ( $\lambda : 20 \text{ to } 60 \mu\text{m}$ ).

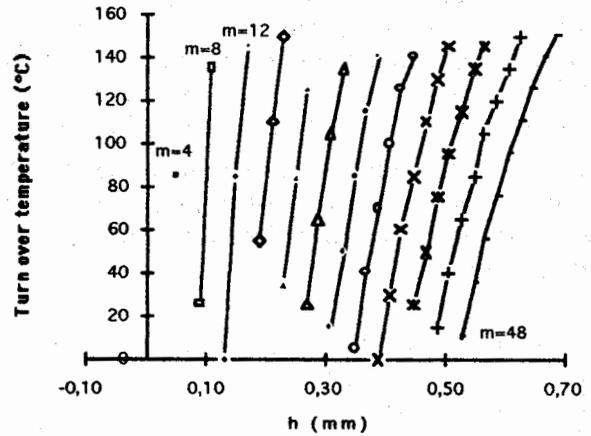


Fig. 8 : Turnover temperature versus  $h$   
ST cut-quartz,  $\lambda = 32 \mu\text{m}$ .

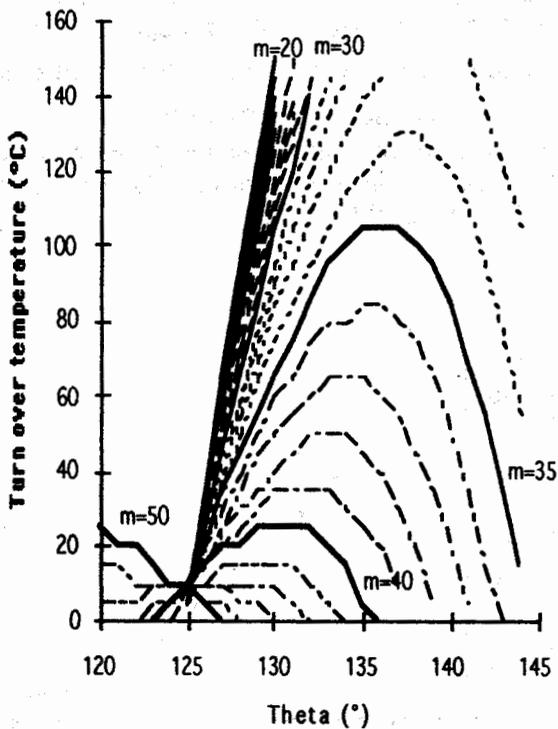


Fig.7 : Turnover temperature versus  $\Theta$   
ST cut - quartz,  $\lambda=32 \mu\text{m}$ ,  $h=0.5 \text{ mm}$ .

On Figs. 8 and 9 are represented the turnover temperature of several modes ( $m$ ) with the thickness  $h(T_0)$ : it can be seen that a little error (a few percents) on  $h$  or  $\lambda$  will not induce a great variation of the turnover

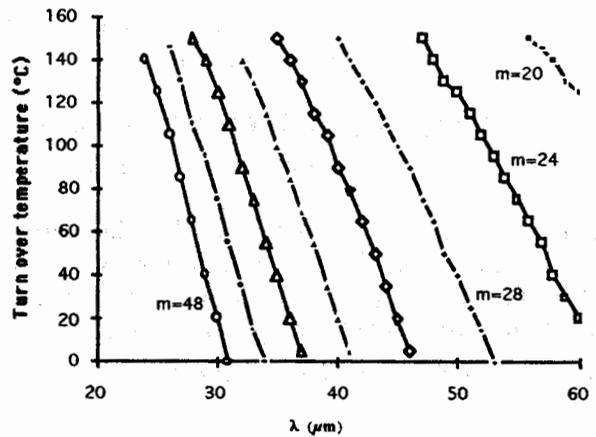


Fig. 9 : Turnover temperature versus  $\lambda$   
ST cut-quartz,  $h = 32 \mu\text{m}$ .

### 3. EXPERIMENTAL RESULTS AND DISCUSSION

The experimental device is based on a ST-cut quartz delay-line on which have been sputtered Aluminium interdigital transducers (IDTs). The quartz plate was 20 mm along the propagation direction, perpendicular to the X crystallographic axis, 15 mm wide, 0.5 mm thick. The IDTs had a periodicity equal to  $32 \mu\text{m}$ , with a centre to centre distance and an overlap of 9.632 mm and 3.5 mm respectively. In an oscillator configuration the mode is chosen thanks to a variable phase

shifter and a band-pass filter, also placed in the retroaction path.

Figure 10 shows the measured frequency-temperature curves. These results bring to light a slope of 5 ppm/°C for the higher mode (194.5 MHz), 13 ppm/°C at 176.5 MHz and 29 ppm/°C at 158.3 MHz (one of the first order modes:  $f_0 \approx 156$  MHz). In table 3, a comparison between theoretical results and experimental measurements obtained with a ST cut device ( $\lambda = 32 \mu\text{m}$ ,  $h = 0.5$  mm) is proposed. We can notice a very good agreement with predict and experimental values which confirm the validity of frequency-temperature behaviour model.

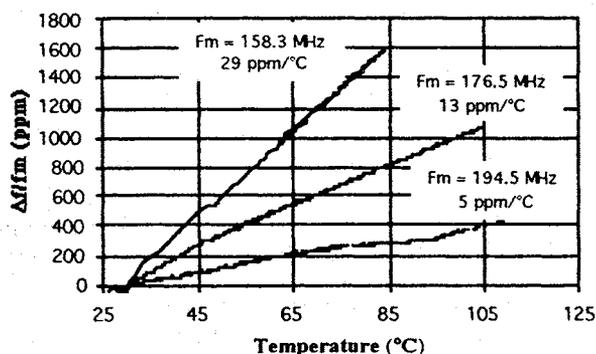


Figure 10 : Experimental frequency shift versus temperature of three modes ST-cut quartz,  $\lambda_0=32\mu\text{m}$ ,  $h=0.5\text{mm}$ .

Mode (MHz)	$\Delta f_m/f_m$ measured (kHz/°C)	$\Delta f_m/f_m$ measured (ppm/°C)	$\Delta f_m/f_m$ theoretical (ppm/°C)
158.3	+ 4.6	+ 29	+30
176.5	+ 2.3	+ 13	+ 16
194.5	+ 1.0	+ 5	+ 6

Table 3 : Comparison between theoretical and experimental temperature sensitivity of three modes of ST-device.

#### 4. CONCLUSION

This study allowed us to confirm theoretical and experimental results on SH-APM temperature behaviour. These results permit a knowledge of the sensivity of SH-APM oscillators to temperature variations and give some informations to design biochemical sensors which need a high frequency stability towards temperature variations.

#### ACKNOWLEDGEMENTS

This work was sponsored by the French Ministry of Defence, through a specific program from the Direction des Recherches et Etudes Techniques (DRET) and especially from the Centre d'Etudes du Bouchet (CEB).

#### References

1. D.P. Morgan, Surface-Wave Devices for Signal Processing, Elsevier (1985).
2. T.E. Edmonds, Chemical Sensors, Blackie (1988).
3. D. Hauden, G. Théobald, Dynamic thermal sensitivity of SAW quartz oscillators, Proc. IEEE Ultrasonic Symp., Chicago (IL, USA), 264-267 (1980).
4. M. Hoummady, E. Bigler, G. Martin, Characterization and temperature sensitivity of shear horizontal acoustic waves in quartz, Proc. EFTF'93, Neuchâtel (Switzerland), 235-238 (1993).
5. M.S. Nieuwenhuizen and A. Venenma, Surface acoustic wave chemical sensors, 1989, Sensors and Materials, 5, 261-300.
6. JC. Andle, JT. Weaver, DJ. McAllister, F. Josse and JF. Vetelino, 1993, Improved acoustic plate mode biosensor, Sensors and Actuators B, vol. 13-14, p.437-442
7. R. Bechmann, A.D. Ballato, T.J. Lukaszek, Higher order temperature coefficients of the elastic stiffness and compliances of alpha quartz", Proc. IRE, 1812, 1962
8. M. Hoummady, Ondes élastiques transverses horizontales émises par des transducteurs interdigitées. Thèse 202, Université de Franche Comté, France, 1991.
9. G. Jaillet, Ondes de surface sur substrats piézoélectriques ; étude et réalisation d'un capteur de température, Thèse, Université Pierre et Marie Curie - Paris VI, France, 1981.
10. M. Michel, Etudes des propriétés thermiques des lignes à retard à ondes de surface, Thèse, n°77, Faculté des sciences et des Techniques de l'Université de Franche-Comté, France, 1977.

# INFLUENCE OF QUARTZ CRYSTAL ORIGIN ON FERROBIELASTIC TWINNING OBSERVED BY X-RAY TOPOGRAPHY

**P.L. GUZZO, J.J. BOY**

ENSM / LCEP - 26, chemin de l'Épitaphe - 25030 BESANÇON Cedex - FRANCE

**B. CAPELLE**

LMCP - Université PARIS VI - 75252 PARIS Cedex - FRANCE

## ABSTRACT

In this paper, we report the ferrobielastic twinning in AT-cut samples issued from natural (Brazil and Madagascar) and synthetic (from Sawyer company) quartz crystals. The experiments have been carried out using an uniaxial stressing equipment for which the stress is applied in a suitable direction of the quartz sample. The stress level associated with the optical appearance of the twinned domains and coercive stress values are presented for each material. It has been observed that these stress levels are smaller for synthetic quartz than for natural one.

The observations of the twin wall propagation were made by X-ray topographs obtained with synchrotron radiations. Transmission Laue patterns due to the white X-ray beam were taken at several compressive stress levels. As visualisation of a twin domain is governed by contrasts generated in topographs of chosen lattice planes, we notice that the twinned domain does not disappear after putting off the pressure.

Additionally, the content of lattice impurities such as Al, Li, Na, K and broad OH band are given.

temperature) are limited due to pressure hysteresis and ferrobielastic twinning, until now, no conclusive results were given. It was just reported [4, 5] that quartz units fabricated from synthetic crystals exhibit different characteristics compared with those manufactured in natural quartz. Additionally, the ferrobielastic twinning under uniaxial stress has been mainly investigated in X-cut samples made on synthetic quartz. There are just few studies involving others crystallographic orientations and natural quartz crystals [6, 7]. Thus, in the present work, AT-cut samples of both synthetic and natural quartz were used to examine the influence of quartz origin on the twin stress level.

In this paper, white beam synchrotron radiation topography is employed to visualise nucleation and propagation of the alternate twin state. Indeed, domain contrasts of Dauphiné twins can appear on such reflection planes because the structure factors of twin states are different [8]. Additionally, infrared absorption (IR) spectroscopy is adopted to characterize OH - related point defects and to estimate the content of the broad OH band. With the Ion Coupled Plasma spectroscopy, we have measured the other lattice impurities contents.

## I. INTRODUCTION

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

The switching occurs when the free energy of the alternate twin state is smaller than this of the existing state. The calculated energy difference between free energies of each twin state depends only on the stress direction and the crystallographic orientation of the sample [2, 3]. On the other hand, we show here that the observed ferrobielastic twinning depends also on the quartz crystal quality and, particularly, if the crystal is synthetic or natural.

Although, previous studies show that the performances of quartz pressure transducers working in severe environmental conditions (high stress and

## II. EXPERIMENTAL PROCEDURE

The samples used in our experiments were obtained from the Z-growth region of a synthetic quartz grown by Sawyer Res. Products and natural quartz from Brazil and Madagascar. The samples are free of macroscopic imperfections as inclusions and twins. The AT-cut blanks are rectangular with the faces to be stressed rotated towards Y' by +45° (according to the 1949 IRE Standard). Their dimensions along X', Y' and Z" are respectively 9.5, 5.0 (or less than 3.0 for X-ray topographies) and 8.5 mm. All surfaces, ground and optically polished, are flat and parallel within  $\pm 2 \mu\text{m}$ .

The apparatus used to apply almost uniaxial compressive stress was described in a previous work [9]. So, we recall just here that the applied force is measured by a strain gauge load cell, whereas two pairs of displacement gauges indicate the vertical displacement on each side of the sample setting between two hardened steel plates. The axial strain is evaluated from the average of these two differential displacements.

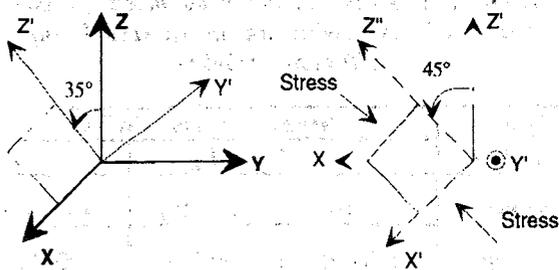


Fig. 1 : Stress-orientation in relation to the crystallographic axes.

X-ray topographs have been performed at L.U.R.E. (Orsay, France) using a white beam of synchrotron radiations. Transmission Laue patterns [10] have been taken at several compressive stress levels, the stressing equipment being settled in order to have the synchrotron beam parallel to the Y' axis of the sample. The height of the beam (6 mm) being smaller than the sample height, we cannot depict the entire volume of the blank. So, we decide to observe hardly always the bottom of this.

On the film, put at about 130 mm of the sample and perpendicular to the beam, we note several spots corresponding to diffractions obtained for radiation wavelengths belonging to the range [0.03, 0.2 nm]. To identify the most interesting lattice plane reflections, we use numerical simulations.

IR absorption spectroscopy has been carried out using a Nicolet 750 FT-IR spectrometer. The nominal resolution of the spectrometer is better than  $0.2 \text{ cm}^{-1}$ , in the  $3 \mu\text{m}$  region. IR spectra of the AT-cut samples (5 mm thick) have been obtained using an unpolarized beam at both, 78 K and room temperature.

The contents of Al, Fe, Li, Na and K have been measured by ICP spectroscopy in using a Perkin-Elmer ICP 6500.

III. RESULTS

The IR spectra of the samples used for our experiments are shown in Figure 2. Several sharp absorption bands due to OH vibrations are superposed upon a broad band absorption. The sharp bands detected in the IR spectrum of synthetic quartz are due to as-grown OH centers [11], whereas in natural quartz, they are mainly associated to either Al—OH or Li-dependent OH centers [12]. On the other hand, the defect structure corresponding to the broad OH band is not clear. But, by using numerical integration [13], we can estimate the content of this defect, presented in table 1.

It presents also the contents of the main metallic impurities. We note a large difference for Al and Li contents between synthetic and natural quartz. As IR spectrum of brazilian quartz exhibits higher absorption bands than madagascan quartz, we can predict more Al-OH and Li-OH centers for its than in the other.

Table 1 : Lattice impurities (in  $N/10^6 \text{ Si}$ )

Element	Sawyer	Madagascan	Brazilian
Al	3.5	39	-
Fe	0.2	1	-
Li	0.75	12.7	-
Na	< 0.01	< 0.01	-
K	< 0.01	0.06	-
$\alpha_{3500} \text{ (cm}^{-1}\text{)}$ at room tpre	0.025	0.030	0.043
OH	20	25	35

- : not measured

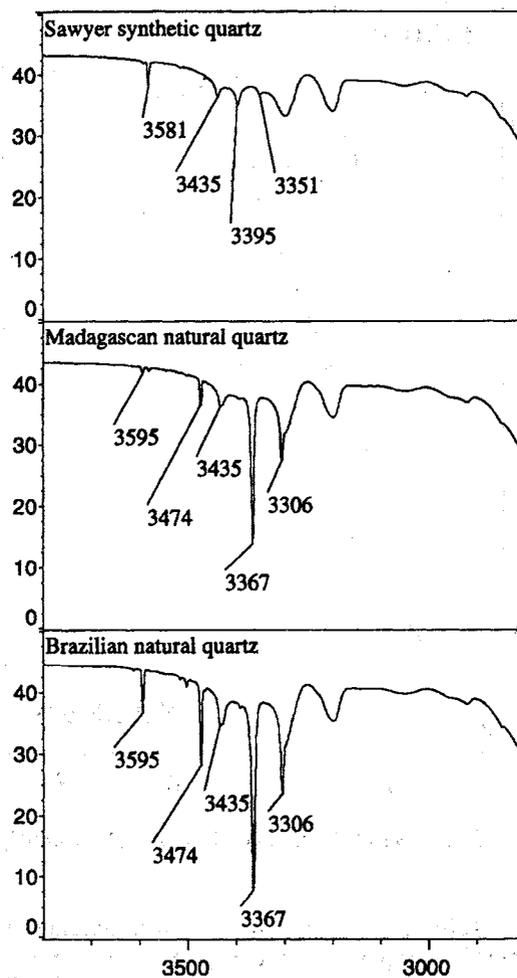


Fig. 2 : IR spectra of AT-cut samples (5 mm thick) measured at 78 K. Transmission (%) vs wavenumber ( $\text{cm}^{-1}$ )

Figure 3 shows a typical stress-strain curve (T-S) illustrating the switching from existing twin state (domain I) to the alternate one (domain II). Generally, the first indication of ferrobielastic twinning is observed through an optical contrast due to piezo-optical effect [2, 3], without change in T-S slope. The contrast between twin states observed by polarized light in AT-cut samples [6] is quite different of the well defined stripe-shaped domains of X-cut samples [2]. The stress level associated with this first observation of domain contrast is denoted as  $T_a$ . By increasing the stress, the optical contrast often

increases and the propagation of twin walls can be observed. The stress-strain slope is usually modified by twin wall propagation :  $T_b$  corresponds to the stress level in which the first change is observed in the slope of T-S curve. When the stress required to switch twin states (noticed  $T_c$ ) is attained, an abrupt jump is usually observed on the T-S characteristic. The switch-over is often followed by emission of an audible snap-sound. Upper  $T_c$ , the optical contrast is no more observed. In our experiments, the stress is increased even after the switch-over. When it is released neither domain contrast nor changes on stress-strain slope is observed.

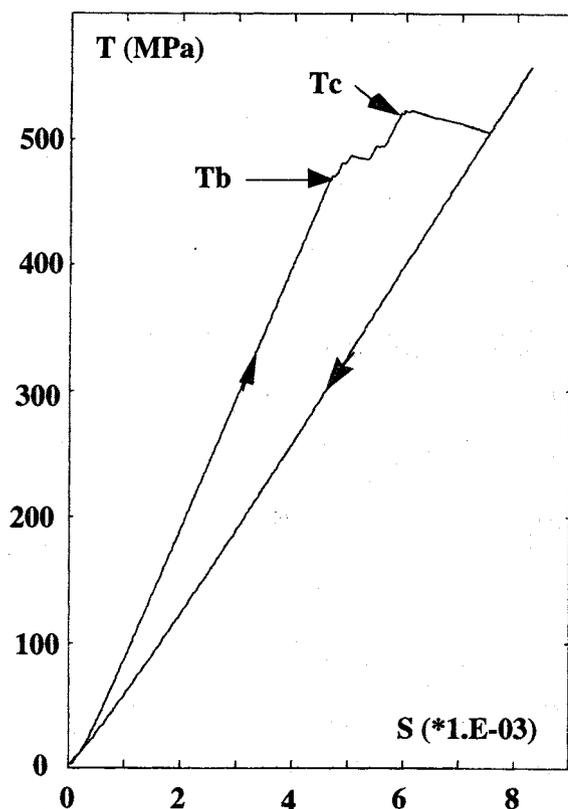


Fig. 3 : Typical stress-strain relationship

Table 2 shows the stress values obtained from compressive experiments using AT-cut samples of Sawyer synthetic quartz and natural quartz from Brazil and Madagascar. The average values for  $T_a$ ,  $T_b$  and  $T_c$  correspond to the measurements realized on five samples. The standard deviation is written between brackets. It is observed that Sawyer synthetic quartz exhibits the lowest values for  $T_a$ ,  $T_b$  and  $T_c$  compared with those for natural quartz. The largest difference is noticed in the stress level  $T_a$  which is 50% higher for natural quartz. The coercive stress  $T_c$  is about 25% higher for natural quartz.

Table 2 : Stress levels associated to the ferrobielastic twinning in quartz from different origins.

Quartz crystals	Stress levels (MPa)		
	$T_a$	$T_b$	$T_c$
Sawyer	185 (105)	365 (45)	400 (50)
Brazil	400 (30)	400 (30)	520 (60)
Madagascar	405 (75)	445 (50)	545 (45)

Figure 4 shows several topographs taken from an AT-cut sample of madagascan natural quartz submitted to a compressive stress cycle. We have particularly observed 2 reflections associated with two reticular planes : (112) and (331) (called P1 and P2, respectively).

The P1 reflection obtained at unstressed state shows a lot of surface defects which are probably due to grinding process. It is also observed that this sample is free of dislocations. The reflecting power of P1 becomes stronger as the stress increases.

At about 230 MPa, this reflection shows an intense contrast region on the right side of the sample and a lot of spots inside the crystal. For the same stress, several spots and an intense contrast region appear in the (331) reflection which were not observed previously. By increasing the stress, it is observed that these spots join with each others to form larger ones.

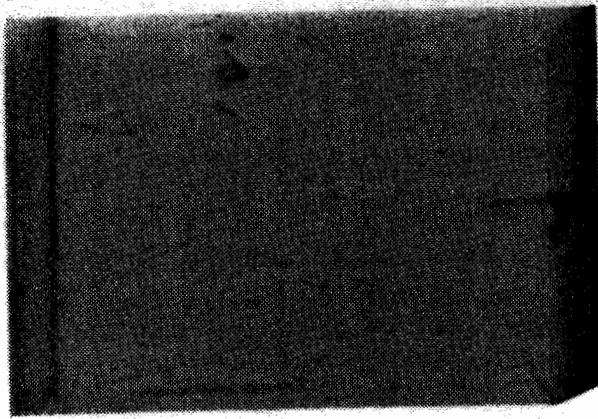
At 455 MPa, P2 reflection indicates that the spots are oriented along [100] and the area showing the strong reflection power is smaller than this observed on the (112) reticular plane.

At least, at about 485 MPa, an abrupt jump appears in the stress-strain relationship and a strong reflecting power is observed in the whole surface sample of both reflections. The stress is increased up to 530 MPa with any change in reflection patterns.

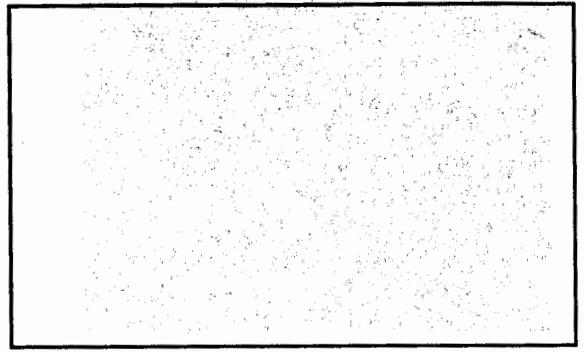
Then, the stress is slowly released. The intensity of both (112) and (331) reflections decreases with the decrease of stress. When the stress reaches zero, the P2 reflection is still observed. The other one is also observed but with a weaker intensity than at the beginning of experiment.

#### IV. DISCUSSION AND CONCLUSION

The effect of the quartz quality material on the ferrobielastic twinning has been studied by means of compressive experiments on AT-cut samples and X-ray topographies. The observation of the diffraction intensity on several reticular planes (particularly (112) and (331)) shows that there is not switch-back after removing the stress in synthetic and natural quartz if twinning has been completed. The following pictures show distinctive dissolution figures obtained after prolonged etching with ammonium bifluoride. The first one corresponds to the AT-cut sample before experiment (domain I), the second being the dissolution figure obtained on other sample after switch-over (domain II).



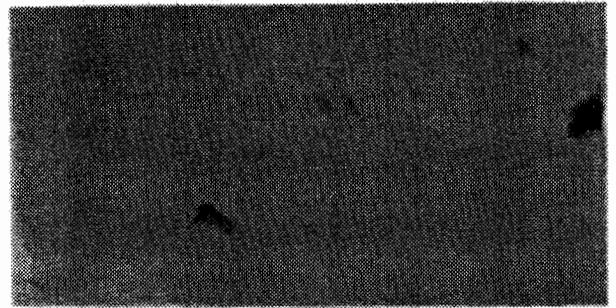
0 MPa



0 MPa



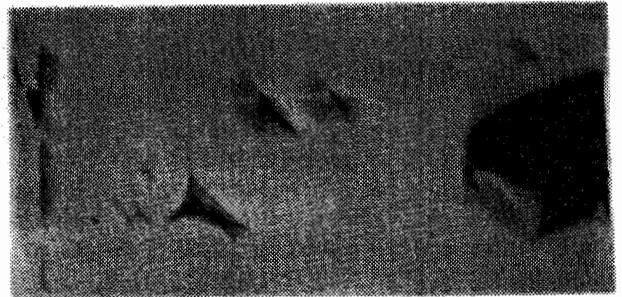
230 MPa



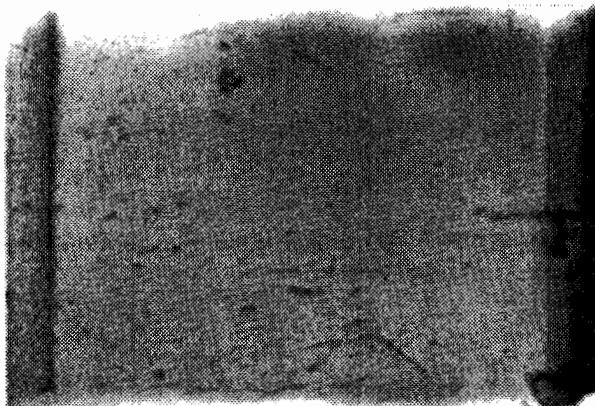
230 MPa



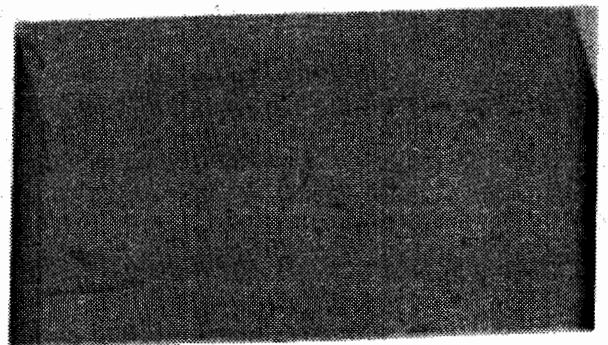
455 MPa



455 MPa

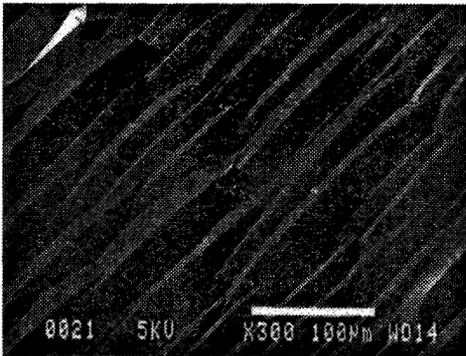
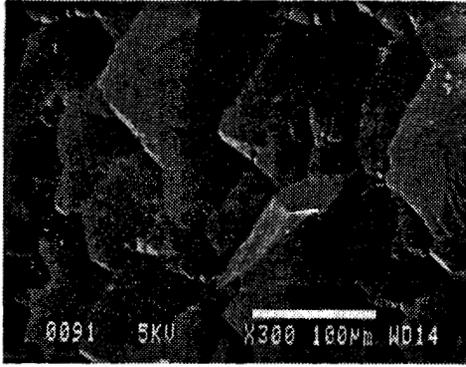


0 MPa



0 MPa

**Fig. 4 : Transmission topographs of an AT-cut madagascan natural quartz under compressive stress applied in the vertical direction (two reticular planes : 112 and 331)**



The compressive experiments show that the stress level required to optically observed and to reverse twin states are respectively 50 and 25 % higher in natural quartz than for synthetic. On the other hand, Cline et al. [7] reported that the coercive stress in X-cut natural quartz was about 30 % lower than measured in synthetic samples. These authors also believed that the impurity content of their brazilian natural quartz was probably lower than in Motorola synthetic quartz but no quantitative values were given. From our IR spectra and table 1 we notice, in contrary, that Sawyer synthetic quartz exhibits lower impurity content compared with our natural quartz. So, further investigations must be carried out in order to clarify whether impurity content and impurity-related point defects influence the stress level associated to ferrobielastic twinning.

At least, the dislocations density of Sawyer crystal is much higher than for natural crystals studied here. It is observed that the nucleation of alternate twin state in madagascan quartz appears in a region free of dislocations, but the surface defects (as micro-cracks) can play an important role in the nucleation of the alternate twin state. Until now, there is any evidence relating the dislocations density and the stress level required to twin. Additionnally, we think that the interaction between dislocations and twin walls is also to be investigated.

#### ACKNOWLEDGMENTS

The authors wish to thank Professors P Zecchini and B. Viard from LCCM laboratory of the Besançon University, for giving facilities for IR and ICP spectrometers. One of the author (PLG) is grateful for scholarship provided by CNPq (Brazil).

#### REFERENCES

- [1]. K. AIZU : "Possible species of ferromagnetic, ferroelectric and ferroelastic crystals", *Phys. Rev. B* 2 (1970), pp. 754-772.
- [2]. K. AIZU : "Second order ferroic state shifts", *J. Phys. Soc. Japan*, 34 (1) (1973), pp. 121-128.
- [3]. R.E. NEWNHAM, L.E. CROSS : "Symmetry of secondary ferroics II", *Mat. Res. Bull* 9 (1974), pp. 1021-1032.
- [4]. J.A. KUSTER, G.S. KAITZ : "Characteristic of natural, swept natural and cultured X- and Z-growth quartz material in high temperature, high stress applications", 39th Ann. Symp. Freq. Cont. (1985), pp. 223-229.
- [5]. R.W. WARD, E.P. EERNISSE : "A reduced hysteresis, extended range quartz pressure transducers", 41st Ann. Symp. Freq. Cont. (1985), pp. 344-349.
- [6]. P.L. GUZZO, J.J. BOY : "Influence of crystallographic orientation on ferrobielastic twinning in quartz", 9th Europ. Freq. Time Forum (1995), pp. 438-442.
- [7]. T.N. CLINE, J.W. LAUGHNER, R.E. NEWNHAM, L.E. CROSS : "Electrical and acoustic emission during ferrobielastic twinning in quartz", 32d Ann. Symp. Freq. Cont. (1978), pp. 43-49.
- [8]. H. KLAPPER : "X-ray topography of twinned crystals", *Prog. Crystal Growth and Charac.*, 14 (1987), pp. 367-401.
- [9]. P.L. GUZZO, J.J. BOY, B. DULMET, P. MAITRE : "Influence of the sample geometry on twinning generation in quartz : experimental measurements and finite elements analysis", 8th Europ. Freq. Time Forum (1994), pp. 802-810.
- [10]. J. MILTAT : "White beam synchrotron radiations topography" in "Characterization of crystal growth defects by X-ray methods" ed. by B.K. Tanner and D.K. Baven, Plenum Press, New-York (1980), pp. 401-420.
- [11]. R.N. BROWN, A. KAHAN : "Optical absorption of irradiated quartz in the near IR", *J. Phys. Chem. Solids* 36 (1975), pp. 467-476.
- [12]. A. KATS : "Hydrogen in alpha-quartz", *Philips Reas. Reports* 17 (1962), pp. 133-195, 201-279.
- [13]. P.L. GUZZO, F. IWASAKI, H. IWASAKI, C.K. SUZUKI, A.C.F. ARRUDA : "Impurity-related point defects and darkening grade in brazilian natural quartz", *Revista de Física Aplicada e Instrumentação* 8 (1993), pp. 68-83.

## SAW RESONATORS FREQUENCY TRIMMING BY PLASMA ETCHING

Krzysztof WEISS, Mariusz TUPAJ,

Tele &amp; Radio Research Institute, Warsaw, POLAND

## ABSTRACT.

SAW resonators need resonant frequency with defined tolerances. The manufacture process, depending on photolithographic forming of interdigital structures on piezoelectric substrate surface, gives frequency dispersion - in general - higher than required. This situation needs correction of resonator frequency after leads mounting. The SAW resonators frequency trimming is possible by plasma etching of quartz surface between interdigital structure fingers. The laboratory equipment for plasma etching has been worked out at Tele & Radio Research Institute. Using this equipment the experiments with frequency trimming of one port SAW resonators on quartz substrate have been carried out. The hexafluoroethan as etching agent was used. In this paper construction of plasma etching laboratory equipment and results of SAW resonators frequency trimming are presented.

## 1. THE SAW ONE PORT RESONATOR.

One port SAW resonators operate by reflecting a propagating wave an itself many times. It is desirable that the reflectors be efficient (i.e. high reflectivity) in order to obtain a sharp high Q resonance. Surface wave reflection gratings have the special property that they can be made efficient and at the same time frequency selective. This latter feature allows such devices to operate in the UHF frequency range where bulk resonators would have to very thin or be operated in an undesirable overtone mode.

The surface wave resonator illustrated in fig. 1. makes use of two reflection gratings to form a surface wave cavity. An interdigital transducer placed inside the cavity couples to the standing surface waves. Unlike a filter resonator is a one port device whose special characteristic is rapidly varying terminal impedance in the vicinity of the resonant frequency. The reflective gratings can be in form of identical periodic arrays of simple groves or metal strips.

The maximum of reflection is achieved if the simple formula is satisfied.

$$f = \frac{v}{2\lambda}$$

Where: f - resonant frequency of cavity

v - surface acoustic wave velocity under reflective structures

$\lambda$  - reflective grating period.

If the cavity length between the inside edges of two gratings is an integral number of half-wavelengths, waves multiply reflected between the two gratings have the same phase, add coherently and the cavity resonates at the frequency of maximum reflection [1]. The interdigital transducer wide band with central frequency the same as cavity resonant frequency positioned exact in the center of cavity is used for surface acoustic wave excitation and collection.

In the case of aluminium strips in reflective gratings the surface acoustic wave velocity is expressed by formula [4]:

$$v = v_p \cdot \left[ 1 + \frac{2d}{\lambda} \left( -5.8 \cdot 10^{-4} - 0.2 \frac{h}{\lambda} \right) - 3.5 \left( \frac{h}{\lambda} \right)^2 \right]$$

where:  $v_p$  - acoustic wave velocity on piezoelectric material

d - electrode width

h - electrode thickness

In the plasma etching process using the aluminium electrodes as mask it is possible to achieve groves between aluminium pattern [3]. In the case of aluminium where substrate and metal layer density are practically the same (quartz  $2.65 \text{ g/cm}^3$ ; Al  $2.7 \text{ g/cm}^3$ ) the groves etching gives effect similar to additional thickening of aluminium layer and in result the resonant frequency decrease. In this case the surface acoustic wave velocity is expressed by formula:

$$v_k = v_p \cdot \left[ 1 + \frac{2d}{\lambda} \left( -5.8 \cdot 10^{-4} - 0.2 \frac{h+h_1}{\lambda} \right) - 3.5 \left( \frac{h+h_1}{\lambda} \right)^2 \right]$$

where:  $h_1$  - depth of etched groves (see fig. 2.)

and the resonant frequency shift is expressed:

$$\frac{\Delta f}{f} = \frac{-\frac{h}{\lambda^2} (0.4d + 7h + 3.5h_1)}{1 + \frac{2d}{\lambda} \left( -5.8 \cdot 10^{-4} - 0.2 \frac{h}{\lambda} \right) - 3.5 \frac{h^2}{\lambda^2}}$$

In this formula the denomination value is near 1 (~0.9975). In practice we can use the simple formula:

$$\frac{\Delta f}{f} = -\frac{h}{\lambda^2} (0.4d + 7h + 3.5h_1)$$

For example for aluminium thickness 150 nm and resonant frequency 200 MHz the 100 nm groves etching will give according to above formula the frequency decrease about 300 kHz.

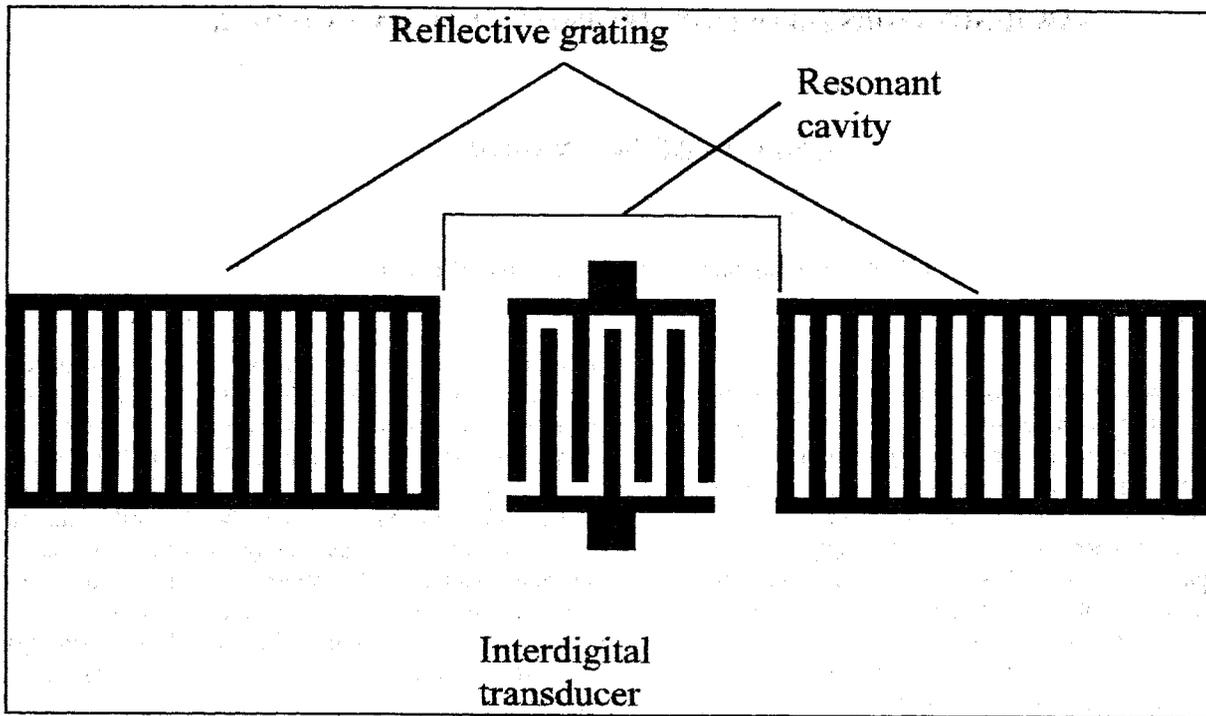


Figure 1. One port SAW resonator.

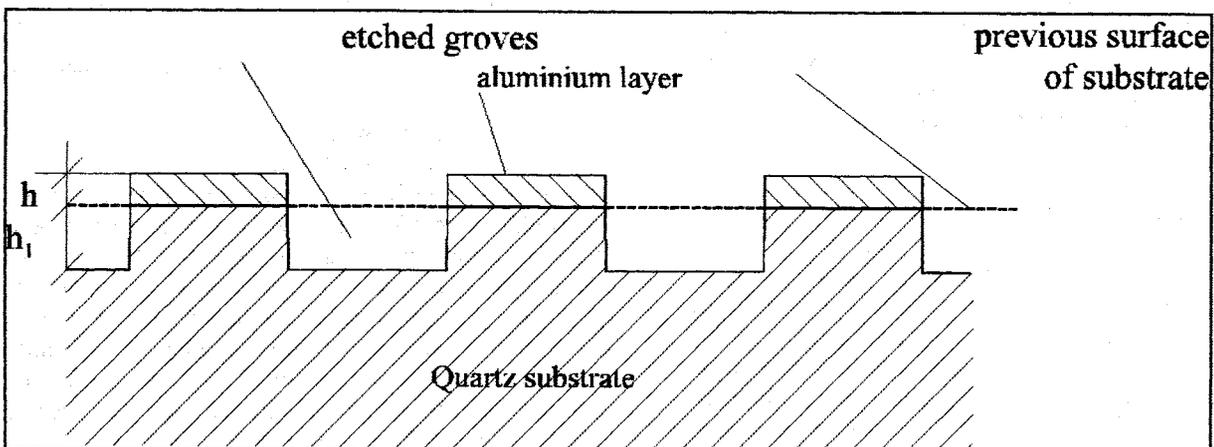


Figure 2. Cross-section of etched resonator structure.

## 2. PLASMA ETCHING PROCESS.

In low pressure the fluorocarbon gases dissociate in the high voltage electric field existence. Free ions of  $F^-$  reacts with  $SiO_2$  giving  $SiF_4$  gas. This process can be utilized for so called dry quartz etching. Fortunately the aluminum is etching about 100 times slower than quartz. It makes possible to etch grooves between the reflective grooves and interdigital transducers aluminium strips. The plasma etching process is carried out in vacuum chamber with continuous pumping and simultaneous supplying in reactive gas which is general one of fluoro carbon as  $CF_4$ ,  $C_2F_6$  or similar, mixed with oxygen and inert gas, or in simplest situation with air. The plasma ionization is generated usually by high frequency high voltage electric field. There are used planar or coaxial electrodes configuration [2].

## 3. LABORATORY PLASMA ETCHING EQUIPMENT.

In Tele and Radio Research Institute was build simply equipment for plasma etching presented in fig. 3. This equipment is built up of cylindrical glass chamber with planar aluminium electrodes, rotary vacuum pump, valves for vacuum height control and valves for reactive gas and air dosage. This unit is equipped in vacuum meter for vacuum measurement and rotameter for reactive gas flow measurement. The electrodes in chamber are supplied in high frequency (13,54 MHz) high voltage current from electronic oscillator with 200 W power. The vacuum in chamber is controlled from 1 to 10 Pa. Electrodes are 140 mm in diameter with changeable distance of  $10 \div 100$  mm. A  $C_2F_6$  gas mixed with air is used as a reactive agent.

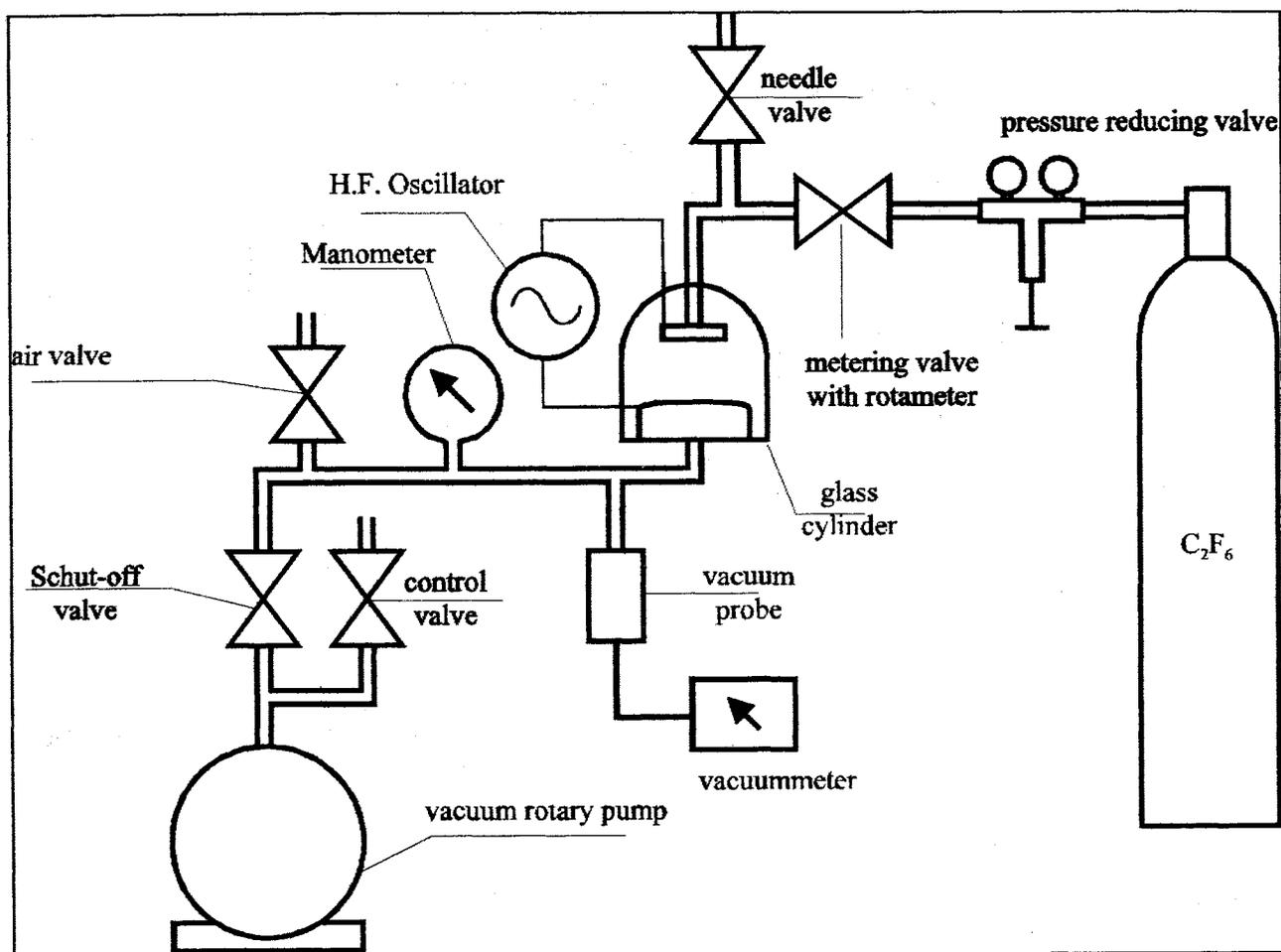


Figure 3 Equipment for plasma etching.

#### 4. THE RESONANT FREQUENCY PLASMA ETCHING TRIMMING EXPERIMENTS RESULTS.

For experiments were used SAW one port resonators with nominal frequency 140; 165 and 190 MHz. This resonators were made on ST quartz substrate with 250 nm aluminium layer thickness. Interdigital structure was consisting of 131 electrodes. The reflective array was consisting of 550 aluminium strips on both sides. The finally mounted resonators without covers were located on special aluminium holder on the table (down electrode) in the reactive chamber. The plasma etching process was carried out in short cycles from 3 to 6 minutes depending on the distance between initial and target frequency. The cycles longer than 6 minutes were not possible because the electrodes temperature arose very quick and etching rate was varying very strong for each process. The temperature measurement in the chamber showed for 6 minutes etching the temperature on down electrode exceeding  $100^{\circ}\text{C}$ . Post every cycle the resonators were removed from chamber and resonant frequency of each resonator was measured. The results are presented in table 1. In all process there were kept the same parameters of etching. The electrodes distance - 50 mm, pressure in chamber -  $\sim 10\text{Pa}$  and  $\text{C}_2\text{F}_6$  flow 20 ccm/min. The electrodes voltage was about 2000 V with electron tube anode

current about 100 mA. For the etching rate determining there was located in the chamber AT quartz plate. The etching rate calculated from AT quartz plate frequency shift was 6,3 nm/min. The resonant frequency before trimming was higher than required about  $1500 \div 2630$  ppm with dispersion 500 - 900 ppm. After frequency correction all resonators were in tolerance 50 kHz i.e.  $< 300$  ppm. The resistance of resonators was  $30 \div 60 \Omega$  and practically was not changing in all trimming process. In few cases past 900 ppm etching did occur additional resonance see fig. 4.

#### 5. CONCLUSIONS.

The plasma etching is easy and very convenient method for one port resonators with aluminium reflective grating frequency trimming. The 2500 ppm frequency decrease is possible without significant resistance increase. To high nonrepeatability of etching rate is caused probably be very strong temperature change in chamber in the etching process. The production equipment must be possessed in cooled down electrode for good process repeatability and long etching time using. The proposed frequency decrease formula can be used for frequency shift calculation.

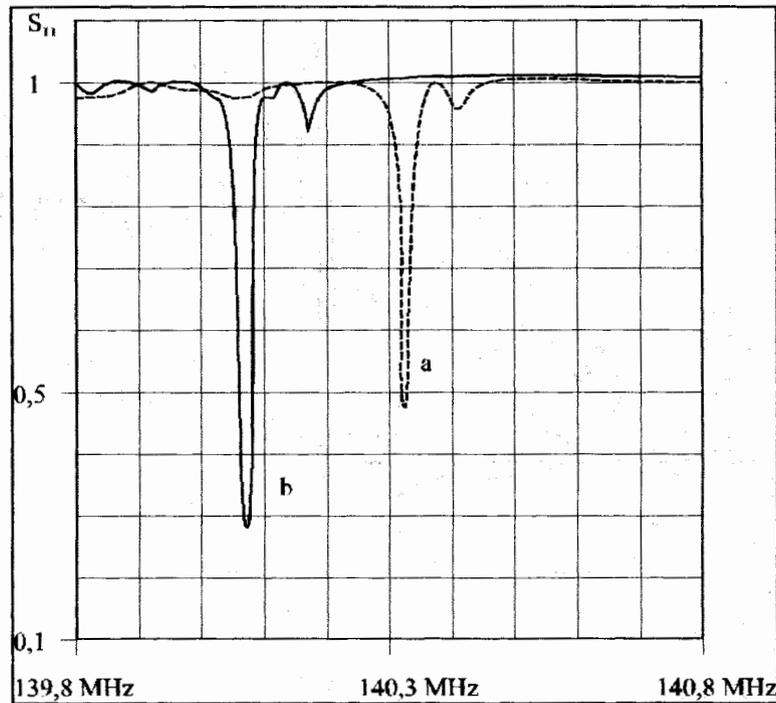


FIGURE 4. Frequency  $S_{11}$  characteristic of resonator before (a) and after (b) frequency trimming.

TABLE 1 -Frequency trimming by plasma etching results

$\Sigma t$ [min]	0	3	6	9	15	20	25	30	35	40	
Resonator no											
1	f [MHz]	165,257	165,245	165,205	165,180	165,054					
	$\Delta f$ [kHz]		11	40	25	126					
2	f [MHz]	140,336	140,330	140,310	140,290	140,245	140,210	140,180	140,145	140,110	140,075
	$\Delta f$ [kHz]		6	20	20	45	35	30	35	35	35
3	f [MHz]	140,260	140,255	140,235	140,205	140,115	140,085				
	$\Delta f$ [kHz]		5	20	30	90	30				
4	f [MHz]	190,330	190,315	190,275	190,230	190,065					
	$\Delta f$ [kHz]		15	40	45	165					
5	f [MHz]	190,502	190,485	190,425	190,380	190,265	190,195	190,130	190,050		
	$\Delta f$ [kHz]		17	60	45	115	70	65	80		

TABLE 2 -Etching rate in [nm/min]

t [min]	3	3	3	6	5	5	5	5	5	main value
Resonator no										
1	2,34	8,08	5,05	12,73						8,18
2	1,43	4,76	4,76	5,36	5,00	4,29	5,00	5,00	5,00	4,30
3	1,19	4,76	7,14	10,71	4,28					6,25
4	2,63	7,02	7,89	14,47						7,72
5	2,98	10,53	7,89	10,09	7,73	6,84	8,42			7,93
									main value	6,876

## 6. REFERENCES.

1. H. Matthews, 1977, "Surface Wave Filters", John Willey & Sons Inc. p. 412-413.
2. R. Beck, 1991, "Technologia Krzemowa" WNPWN Warszawa p. 180-218.
3. P. S. Cross. W. R. Shreve, 1982, "Frequency trimming of SAW resonators", proc IEEE
4. W. Soluch, private discussion and consultation

## PARASITIC BULK WAVES IN WIDE BAND SAW FILTERS DESIGNED ON 128° ROTATED Y-CUT OF LiNbO<sub>3</sub>

Piotr Nagłowski, Hanna Majewska

Institute of Electronic Materials Technology, ul. Wólczyńska 133, 01-919 Warsaw, Poland

### ABSTRACT

This paper presents the results of bulk wave transmission investigation in a wide band SAW filter designed on 128° YX LiNbO<sub>3</sub>. Bulk wave transmission in the filter was investigated in both frequency and time domain. Fourier transform and windowing techniques were used to identify the most important bulk wave modes and their contribution to the frequency response of the device. The influence of the substrate backside grooving on bulk wave propagation was examined. Application of grooving significantly reduces the level of bulk wave modes bouncing off the backside of the substrate. As a result substantial improvement in stopband rejection was obtained. It was found that in the devices with their backside grooved one of the most prominent bulk waves is the longitudinal wave travelling directly between the IDTs of the filter. The longitudinal wave limits the rejection in the upper stopband to approx. 40 dB.

### INTRODUCTION

Bulk wave transmission is a serious problem in SAW transversal filters. This transmission degrades filter stopband especially on the high-frequency side of the passband. One way to cope with the problem is the use of a multistrip coupler. However, this solution leads to the increase of both the size and the cost of the device. Some crystalline material cuts have been found for which the bulk wave transmission is relatively small. One of these cuts is the widely used 128° rotated Y-cut of lithium niobate. Comparatively low stopband degradation caused by bulk waves was demonstrated by Shibayama et al (1), Milsom et al (2) and Wagers (3) for low and medium bandwidth filters using this cut as a substrate.

Generally, both direct and indirect bulk wave transmission is possible in SAW filters. In the first case the waves are transmitted directly along the crystal surface and in the second one their path includes one or more bounce off the bottom of the crystal. Usually the bulk waves travelling between

the filter IDTs indirectly can be substantially reduced by an appropriate roughening of the crystal backside, so the filter performance is determined by the directly transmitted bulk wave modes.

It follows from the results presented in (1,2,3) that the coupling of directly transmitted shear waves in 128° rotated Y-cut of lithium niobate is relatively small. For filters of moderate passband width the directly transmitted longitudinal wave, due to its high velocity, manifests itself only for frequencies substantially higher than the frequencies of the filter passband and therefore usually the wave does not degrade the stopband within its specified frequency range.

The purpose of this work was to investigate bulk wave transmission in wide band filters designed on 128° rotated Y-cut of LiNbO<sub>3</sub>. Below we present the results obtained for a 40% bandwidth in-line filter consisting of one apodized and one unapodized IDT.

### EXPERIMENTAL PROCEDURE

The SAW filter employed in the investigation was designed to operate at approx. 70MHz center frequency. Double electrodes were applied in both transducers. Electrode center-to-center spacing of 14μm and metallization ratio of 0.5 were used. The input transducer (apodized) contains 80 electrodes (79 active overlaps). The output transducer (unapodized) contains 12 electrodes (5 active overlaps). The distance between the centers of transducers was 3630μm and the substrate was 510μm thick. Thin film metallic shield of 1400μm width was applied between the transducers to suppress the direct electromagnetic breakthrough signal. The acoustic aperture of the transducers was 5500μm. Schematic outline of the SAW filter is shown in Fig. 1.

To examine the influence of the substrate backside grooving on bulk wave propagation both grooved and ungrooved filters were investigated. The grooves were cut obliquely (the angle between them and the X axis was 70°). Other parameters of the grooves were as follows: depth - 70±15μm,

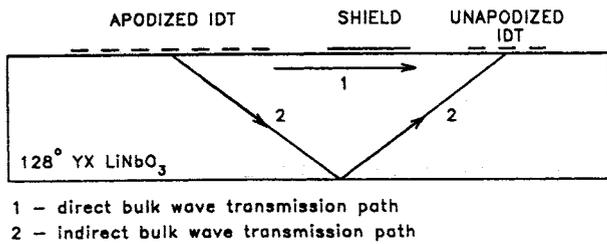


Fig. 1. Schematic outline of the SAW filter

center-to-center spacing -  $300 \pm 5 \mu\text{m}$ , width -  $100 \pm 20 \mu\text{m}$ .

Substantial amounts of bulk wave energy in the filter were transmitted within the time window corresponding to Rayleigh wave propagation and/or within the frequency range of the filter passband. To suppress Rayleigh wave propagation the upper surfaces of some filter chips were covered with an acoustic absorber (epoxy resin). As it follows from publications of many authors and from our own experiments the application of an acoustic absorber on the filter surface has little effect on bulk waves propagation. It is also well known that epoxy resins suppress Rayleigh wave effectively.

The bulk wave investigation was performed in both frequency and time domain. HP 8752A network analyzer was used to measure the frequency response. The frequency response was measured at 801 equidistantly spaced frequency points within the range from 10MHz to 130MHz (the frequency step was 150kHz). Both source and load impedances were  $50\Omega$ .

Discrete Fourier Transform (DFT) was used to compute the impulse response of the devices under investigation. Kaiser window function ( $\beta = 4.538$ , 50dB sidelobe level, see Rabiner and Gold (4)) was applied to reduce the Gibbs effect. Windowing improves the dynamic range of time domain measurements, however, at the expense of increased impulse width.

For comparison purposes the theoretical filter responses are also presented in the paper. The frequency response was computed assuming that all the energy in the filter is transmitted by means of Rayleigh wave (bulk waves and electromagnetic leakage were neglected). SAW filter mathematical model which was used to compute the frequency response takes into account SAW diffraction as well as the influence of external circuits. The parabolic approximation was used to simulate diffraction. Motional admittances of the IDTs were computed as proposed by Nagłowski in (5). A method similar to the one shown by Morgan (6) was used to evaluate static IDT capacitances. The

detailed description of the filter mathematical model is given in Nagłowski (7). The theoretical impulse response was determined by the application of DFT to the theoretical frequency response.

Some of the most important bulk wave modes were identified by comparison of the theoretically predicted time delays corresponding to bulk wave modes with the delays of actual peaks visible in the experimentally determined impulse response (3).

Gating technique was used to evaluate the effect of certain fragments of the experimentally determined impulse response on the device frequency response. This technique is based on mathematically removing (by nulling) selected parts of the impulse response and transforming the result back into the frequency domain.

## RESULTS AND DISCUSSION

Fig. 2 illustrates experimental amplitude response of the ungrooved device and the theoretical response. Large differences between the curves are visible in the stopbands especially on the high-frequency side of the passband. Impulse responses corresponding to the two curves are presented in Fig. 3.

The time window corresponding to Rayleigh wave can be determined through simple calculations from the filter dimensions and SAW velocities for both free and metallized surface of  $128^\circ \text{YX LiNbO}_3$ . Neglecting SAW reflections and TTE we obtained that this time window extends from 763ns to 1091ns. Substantial differences between the curves shown in Fig. 3 occurring outside this window indicate the presence of the mechanisms of energy transfer in the filter other than Rayleigh wave. One of this mechanisms is the electromag-

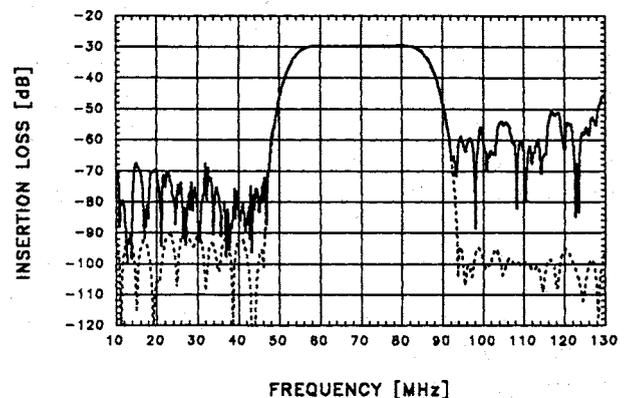


Fig. 2. Amplitude responses of the ungrooved filter  
— measured  
- - - theoretical

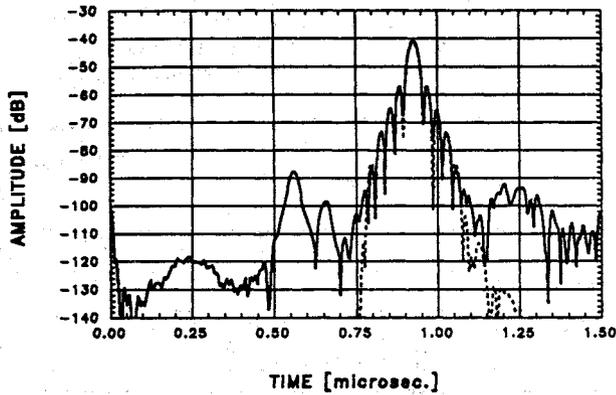


Fig. 3. Impulse responses of the ungrooved filter  
 ——— determined experimentally  
 - - - - - theoretical

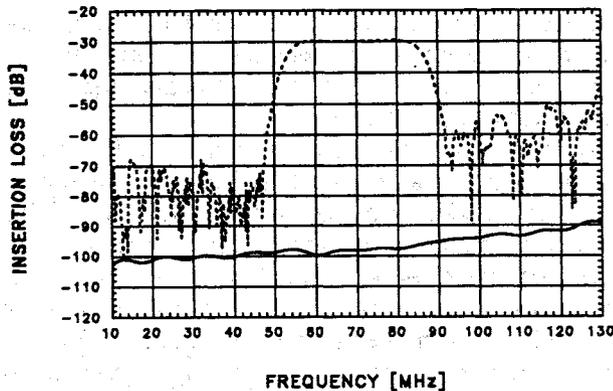


Fig. 4. Amplitude responses of the ungrooved filter  
 ——— corresponding to the direct breakthrough signal  
 - - - - - measured

netic feedthrough, which causes the impulse response peak at  $t = 0$ . To examine the influence of the electromagnetic feedthrough on the frequency response the fragment of the experimental impulse response extending from 0ns to 100ns was Fourier transformed back to the frequency domain. The result is presented in Fig. 4. Comparison of the amplitude response corresponding to the electromagnetic feedthrough with the measured amplitude response of the filter shows that the electromagnetic feedthrough has only small effect on the filter frequency response and is not responsible for low stopband rejection of the filter. To suppress the Rayleigh wave the upper surface of the chip was coated with epoxy resin. The amplitude responses of the filter before and after coating are presented in Fig. 5. Strong correlation between the two curves (particularly in the upper stopband) indicates that bulk waves are one of the main factors degrading the filter stopband rejection.

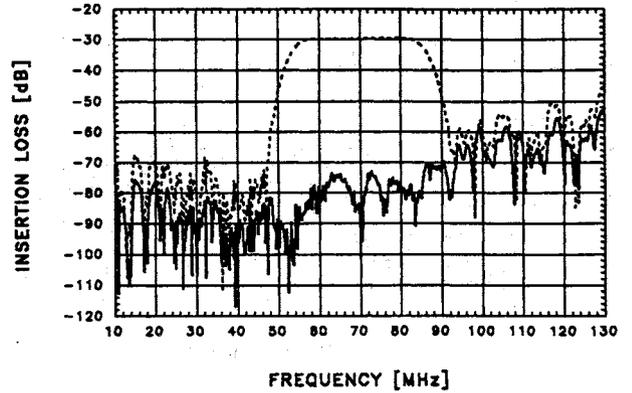


Fig. 5. Measured amplitude responses of the ungrooved filter  
 - - - - - before covering the chip with epoxy resin  
 ——— after covering the chip with epoxy resin

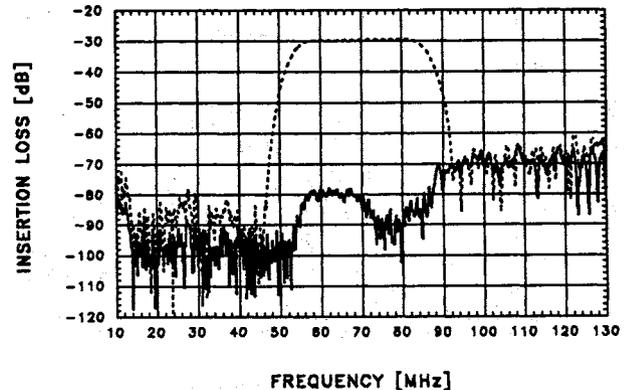


Fig. 6. Measured amplitude responses of the grooved filter  
 - - - - - before covering the chip with epoxy resin  
 ——— after covering the chip with epoxy resin

Analogous results, but obtained for the filter with grooved backside of the chip, are shown in Fig. 6. Strong correlation between the curves in the upper stopband is visible again. The effect of grooving on the filter amplitude response can be evaluated by comparison of Figs. 5 and 6. Application of grooving substantially improves stopband rejection of the filter. At some frequencies the improvement is greater than 20 dB. The improvement is observed on both sides of the passband. The frequency responses of Figs. 5 and 6 were Fourier transformed to time domain. The results are shown in Figs. 7 and 8 respectively. Solid line of Fig. 7 exhibits much more peaks of high amplitude than the analogous line of Fig. 8. This indicates that the peaks are the result of the bulk waves bouncing off the chip bottom and that the operation of grooving reduces the level of bulk waves travelling indirectly between the IDTs (it seems reasonable to assume that grooving has

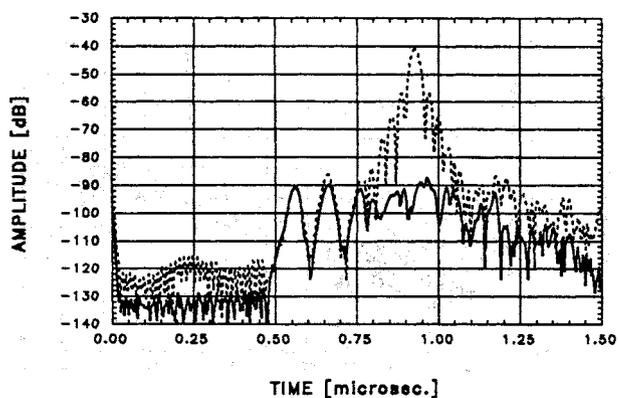


Fig. 7. Experimentally determined impulse responses of the ungrooved filter  
 - - - - - before covering the chip with epoxy resin  
 ——— after covering the chip with epoxy resin

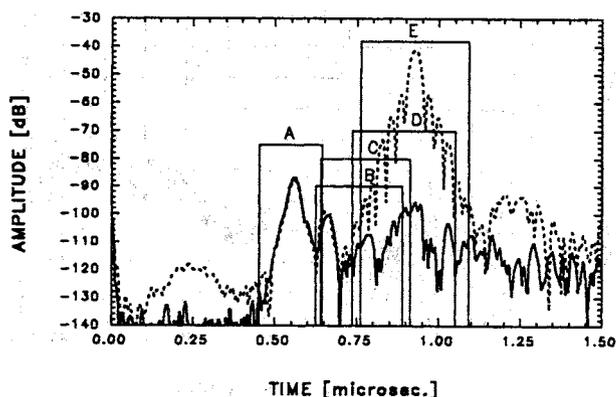


Fig. 8. Experimentally determined impulse responses of the grooved filter and time windows corresponding to directly transmitted bulk waves  
 - - - - - response before covering the chip with epoxy resin  
 ——— response after covering the chip with epoxy resin  
 A - longitudinal wave (L) time window  
 B - fast shear wave (FS1) time window  
 C - fast shear wave (FS2) time window  
 D - slow shear wave (SS) time window  
 E - Rayleigh wave time window

only little effect on the bulk waves travelling along the surface of the crystal).

Assuming that the indirectly transmitted bulk waves can be effectively scattered by appropriate grooving of the chip backside the filter performance strongly depends on the transmission of the bulk waves travelling along the substrate surface. Theoretical data for all four types of these waves are given in Table 1. The data concerning phase as well as group velocities were taken from (3). Main peak time delays and time windows corresponding to each of the bulk wave types were determined from appropriate group velocities and geometrical dimensions of the filter. The time windows from

TABLE 1: Data for slow shear (SS), fast shear (FS) and longitudinal (L) waves propagating along X axis in the SAW filter (phase and group velocity angles, phase and group velocities, main peak time delays and time window ranges)

Wave	$\theta_p$ (deg.)	$\theta_g$ (deg.)	$v_p$ (m/s)	$v_g$ (m/s)	T (ns)	Time window (ns)
SS	0	0	4063	4063	893	737 - 1050
FS1	19.6	0	4397	4668	778	641 - 914
FS2	0	0	4795	4795	757	624 - 890
L	0	0	6628	6628	548	452 - 644

Table 1 and the Rayleigh wave window are outlined in Fig. 8. The solid line in this figure, representing the impulse response of the grooved epoxy resin covered filter, exhibits its highest peak almost exactly in the middle of the longitudinal wave time window. The windows corresponding to the other bulk waves also contain peaks of considerable amplitudes.

To estimate the effect which each of the four types of bulk waves has on the filter frequency response Fourier transforms of appropriate fragments of the impulse response of the grooved epoxy resin covered filter were computed. Since the time windows of Fig. 8 overlap the windows of smaller width have been used in the computations. Each of the windows contained only the main part of its corresponding bulk wave impulse. Some peaks which are probably the result of bulk waves bouncing off the substrate backside were not included in the windows used in the computation. Due to similar group velocities it was impossible to separate the two fast shear wave pulse responses. Fast shear waves shared a common window in the computation. The windows used in the computation are shown in Fig. 9 and their exact range is given in Table 2.

Decreasing the width of the time windows corresponding to the investigated bulk waves is a source of error in determining the influence of these waves on the frequency response of the

TABLE 2: Time windows used to calculate the influence of various bulk waves on the filter frequency response

Wave	Time window (ns)
slow shear (SS)	825 - 998
fast shear (FS1 and FS2)	701 - 814
longitudinal (L)	452 - 630

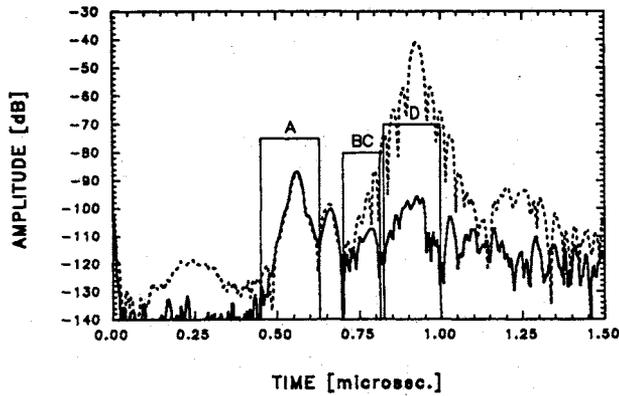


Fig. 9. Experimentally determined impulse responses of the grooved filter and the time windows used to calculate the influence of various bulk waves on the filter frequency response

- response before covering the chip with epoxy resin
- response after covering the chip with epoxy resin
- A - time window used for longitudinal wave
- BC - time window used for fast shear waves
- D - time window used for slow shear wave

filter. It is expected that this error is much larger for shear waves since the windows of shear waves were shortened in a much bigger degree than the longitudinal wave time window. Another possible source of error is the presence of the pulses being the result of the waves bouncing off the chip backside in each of the windows used in the computation. The multitude of peaks in the time window used for the slow shear wave suggests the presence of the indirectly transmitted bulk waves within this window. Fig. 10 illustrates the computation results. It follows from this figure that the longitudinal wave is the main factor degrading the filter stopband on the high-frequency side of the passband, limiting the stopband rejection to approx. 40dB. Slow shear wave substantially contributes to energy transmission in the passband (compare Figs. 6 and 10). According to the computation this wave also contributes to the upper stopband degradation. However, the results obtained for the slow shear wave should be interpreted with care due to dramatic narrowing of the time window and the possible distortion from the indirectly transmitted bulk waves.

Throughout this paper it was assumed that covering the upper surface of the chip with epoxy resin has small effect on the transmission of bulk waves. We verified this assumption for the case of longitudinal wave travelling along the crystal surface. The verification was possible because the longitudinal wave time window and the Rayleigh wave time window do not overlap. Small differences between the impulse responses shown

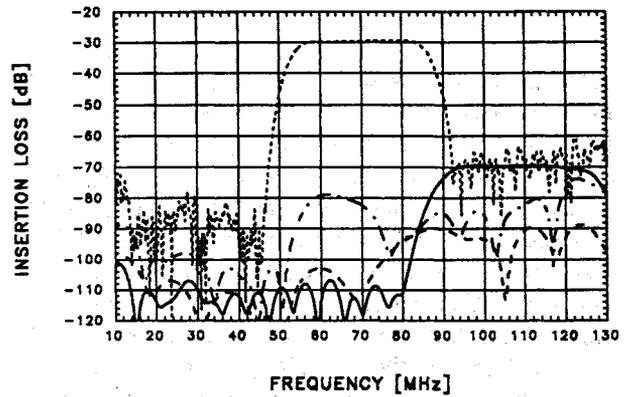


Fig. 10. Amplitude responses of the grooved filter

- measured
- corresponding to longitudinal wave time window (determined by gating)
- · - · - corresponding to fast shear waves time window (determined by gating)
- corresponding to slow shear wave time window (determined by gating)

in Fig. 9 observed within the longitudinal wave time window prove the validity of the assumption in time domain. To verify the assumption in frequency domain amplitude responses corresponding to the longitudinal wave time window were computed for epoxy resin coated and uncoated grooved filters. The computation results are shown in Fig. 11. Small differences between the two amplitude responses confirm that epoxy resin has small effect on the transmission of longitudinal bulk wave.

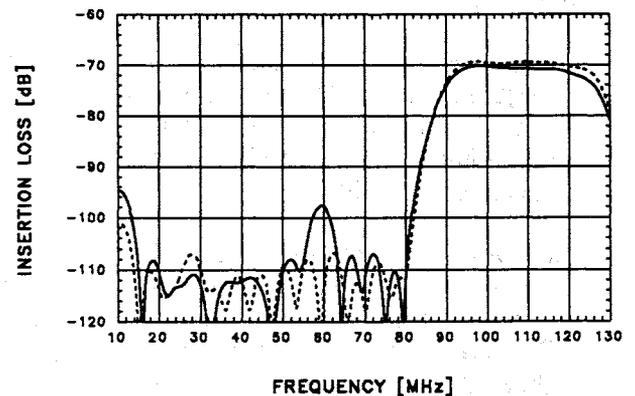


Fig. 11. Amplitude responses of the grooved filter - corresponding to the longitudinal wave

- before covering the chip with epoxy resin
- after covering the chip with epoxy resin

## CONCLUSION

Bulk waves in the wide band SAW filter of 40% bandwidth designed on  $128^\circ$  YX  $\text{LiNbO}_3$  were investigated. It was shown that grooving of the filter backside significantly improves the stopband rejection. In the devices with their chip backside grooved the most prominent bulk wave (in both frequency and time domain) is the longitudinal wave travelling along the surface of the chip. This wave limits the stopband rejection to approx. 40 dB. Contrary to the results presented earlier by other authors for filters of low and medium bandwidth designed on the same substrate material the longitudinal wave manifests itself in the frequency range adjacent to the filter passband. The resulting degradation of stopband rejection is unacceptable in many applications. Possible solutions to the problem are: applying a multistrip coupler in the design or developing an appropriate bulk wave compensation design algorithm.

## ACKNOWLEDGEMENT

The authors would like to express their gratitude to prof. Waldemar Soluch for his support of this project and for his encouragement. Thanks are also due to Mr. Marian Teodorczyk for fabrication of the filters.

## REFERENCES

1. Shibayama K., Yamanouchi K., Sato H. and Meguro T., 1975, "Optimum Cut for Rotated Y-Cut  $\text{LiNbO}_3$  Crystal Used as the Substrate of Acoustic-Surface-Wave Filters", Proceedings of the IEEE, **64**, 595-597
2. Milsom R.F., Murray R.J., Flinn I. and Redwood M., 1981, "New Orientation of Lithium Niobate for Low Bulk-Wave Degradation of SAW Filter Stopband", IEEE Ultrasonics Symposium Proc., 299-304
3. Wagers R., Cohn R., 1984, "Residual Bulk Mode Levels in (YX)  $128^\circ$   $\text{LiNbO}_3$ ", IEEE Trans. on Sonics and Ultrasonics, **SU-31**, 168-174
4. Rabiner L. and Gold B., 1975, "Theory and Application of Digital Signal Processing", Prentice-Hall, Englewood Cliffs, New Jersey, U.S.A.
5. Nagłowski P., 1988, "Fast Computation of Finger-Length-Weighted SAW Transducer Dynamic Admittance", Electronics Letters, **24**, 1494-1497
6. Morgan D., 1985, "Surface-Wave Devices for Signal Processing", Elsevier, Amsterdam, The Netherlands
7. Nagłowski P., 1992, "Using Constrained Optimization Techniques for Compensation of Second-Order Effects in SAW Filters" - Ph.D. dissertation (in Polish), Warsaw Technical University, Warsaw, Poland

# AN ANALYSIS OF BAW TO STW MUTUAL TRANSFORMATION ARISING FROM PERIODIC CORRUGATION ON THE SURFACE OF A RECIPROCAL COUPLER

Bernard DULMET and Hervé WATCHUENG

Laboratoire de Chronométrie, Electronique et Piézoélectricité  
Ecole Nationale Supérieure de Mécanique et des Microtechniques  
F - 25030 Besançon Cedex, France

*Abstract*-In this paper we present a theoretical analysis of BAW to STW transformation occurring in a reciprocal coupler consisting of periodic corrugations on the free surface of a singly-rotated cut of a class-32 crystal. The analysis is based on the coupling-of-modes model. We consider horizontal shear waves travelling in the cross-section of such a cut that contains  $Z'$  direction. In the considered frequency range, SSBW wavelength along  $Z'$  is close to the period of corrugations. Boundary conditions in the coupler's region are projected onto the average flat plane of the free surface with means of Brekhovskikh's well-known expansion[1], and the coupled solutions are sought as series of Bloch's partial waves. Since a previous paper[2], significant advances have been made in the treatment of the dispersion equation. We especially point out the necessity of taking at least five coupled modes to obtain a determinantal equation that actually vanishes. The problem of STWs launched by BAWs falling under normal incidence onto the corrugated surface is treated for a coupler of finite length.

## INTRODUCTION

Horizontally polarized surface waves recently became very popular according to their higher propagation velocity in compare with Rayleigh waves. A minimal understanding of the effect of periodic corrugations on STWs is required for their study since the existence as well as the structure of this kind of waves are mainly conditioned by these corrugations. So, the bulk scattering that may accompany STW propagation has been analyzed by researchers in the SAW field.

From another hand, researchers in the BAW field have essentially studied the scattering of plane waves under normal incidence onto a randomly rough surface. This scattering will not spontaneously generate any surface waves but will mostly decrease the Q factor of resonator devices. Nevertheless, if we consider the growing need

for higher and higher frequency devices, achievement of combined BAW/SAW resonant devices might attract some practical interest : for instance, the design and understanding of electrodes is much simpler for BAW devices than for SAW, also a reciprocal coupler consisting of simply etched gratings allows for various adjustment parameters such as the depth and shape of the surface profile.

The available literature in this field proposes theoretical analyses of BAW to SAW couplers in which the SAW waves of interest are either Rayleigh waves in isotropic media [3-5] or Love waves [6-7]. The scattering of Rayleigh waves in anisotropic and piezoelectric media is a fairly complicated phenomenon, due to the multi-component structure of Rayleigh waves. Thus, we judged more reasonable to devote some work to the study of bulk waves to surface waves conversions in the restricted case of horizontally polarized shear waves. Then, we have to consider only one component, although there is some price to pay for this convenience : the existence and structure of STW itself outside of the coupler's region relies on corrugations, whereas Rayleigh wave can propagate on a flat free surface. As a consequence, in the framework of the Coupling of Modes model, the mono-component simplicity comes together with tricky case studies to consider degeneracy of harmonics into SSBW waves.

## I EQUATIONS OF COUPLED MODES

As in our previous paper [2], we consider horizontal shear modes travelling in a sagittal plane which contains  $Z'$  axis, in a  $Y + \theta$  singly-rotated cut of quartz or another crystal with similar symmetry, and we neglect piezoelectric effect. Then, it is easily shown that, for any propagation in the ( $Y'Z'$ ) sagittal plane,  $u_2$  and  $u_3$  components of the mechanical displacement can be ignored since they are completely uncoupled from  $u_1$ , so we will boldly use «  $u$  » in the followings to denote the

displacement component. The system of propagation equation, constitutive equations and boundary conditions reduces to a very simple set :

$$\left. \begin{aligned} \rho \frac{\partial^2 u}{\partial t^2} &= \frac{\partial T_{12}}{\partial y} + \frac{\partial T_{13}}{\partial z} \\ T_{13} &= c_{56} \frac{\partial u}{\partial y} + c_{55} \frac{\partial u}{\partial z} \\ T_{12} &= c_{66} \frac{\partial u}{\partial y} + c_{56} \frac{\partial u}{\partial z} \\ T_{12} - f'_z T_{13} &= 0 \quad \left( \begin{array}{l} \text{on the external surface} \\ \text{defined by } y = f(z) \end{array} \right) \end{aligned} \right\} \quad (1)$$

We assume that the semi-infinite substrate lies on the  $y < 0$  side of the half-space. It is well-known that a self-sufficient solution whenever the profile of the free external surface (located at  $y=0$ ) is perfectly flat consists of a SSBW, which is a plane wave propagating in a direction such that  $c_{66}k_2 + c_{56}k_3$  vanishes,  $\mathbf{k}$  being the wave vector. Such a simple wave satisfies the boundary condition when the profile function  $f$  identically vanishes and it carries away some energy in the  $Z'$  direction only, *i.e.* parallel to the outer surface. The velocity of guided SSBWs along  $Z'$  is given by simple formulas:

$$\begin{aligned} V_S &= V_t \sqrt{1 - C_2^2 / C_1} \\ V_t &= \sqrt{c_{55} / \rho} \\ C_1 &= c_{66} / c_{55} \quad C_2 = c_{56} / c_{55} \end{aligned} \quad (2)$$

In case of purely sinusoidal corrugations, we can choose the following profile function :

$$f(z) = \frac{2\eta}{Q} \cos(Qz) \quad (3)$$

where the amplitude of profile  $2\eta$  is assumed to be very small with respect to the corrugation period  $Q=2\pi/\Lambda$ . Since the actual slope of the surface profile remains very small, we expand the boundary conditions as Taylor's series in terms of the local value of the profile function, as was first proposed by Brekhovskikh [1]:

$$\left. \begin{aligned} T_{12} - f'_z T_{13} \Big|_{y=f} &= T_{12} - f'_z T_{13} \Big|_{y=0} + y \frac{\partial (T_{12} - f'_z T_{13})}{\partial y^2} \Big|_{y=0} \\ &+ \frac{y^2}{2!} \frac{\partial^2 (T_{12} - f'_z T_{13})}{\partial y^2} \Big|_{y=0} + h.o.t. \end{aligned} \right\} \quad (4)$$

A second order expansion is required to simultaneously model bulk scattering and multiple reflections of surface transverse waves. According to the periodic nature of the corrugations the possible eigenmodes in the coupler's region are sought as a sum of Bloch's functions :

$$\begin{aligned} u(y, z) &= e^{-i\omega t} \sum_k A_k u_k(y, z) \\ u_k(y, z) &= e^{i(\beta_k y + Q_k z)} \quad ; \quad Q_k = q + kQ \end{aligned} \quad (5)$$

Each harmonic  $u_k$  appearing in the sum should individually obey the simple propagation equation (1-a). After substitution into Eq(1-a), which must be individually obeyed by each harmonic, the following generic propagation equation is easily derived :

$$C_1 \beta_k^2 + 2C_2 \beta_k Q_k + Q_k^2 - k_f^2 = 0 \quad \forall k \quad (6)$$

The obvious effect of the sum (5) is to turn Brekhovskikh's expansion of the boundary conditions into a series of  $\exp(iQ_k z)$  functions. Each individual term of this series must identically vanish, then the generic boundary conditions can be written as follows :

$$\left. \begin{aligned} -\eta^2 a_{k+2} A_{k+2} - \eta b_{k+1} A_{k+1} + i t_{1k} \left( 1 - \eta^2 t_{3k}^2 \right) A_k \\ + \eta d_{k-1} A_{k-1} + \eta^2 e_{k-2} A_{k-2} = 0 \end{aligned} \right\} \quad (7)$$

where we use the following intermediate quantities :

$$\left. \begin{aligned} a_k &= i t_{3k} \left( t_{2k} + \frac{1}{2} t_{1k} t_{3k} \right) \\ b_k &= t_{2k} + t_{1k} t_{3k} \\ d_k &= t_{2k} - t_{1k} t_{3k} \\ e_k &= i t_{3k} \left( t_{2k} - \frac{1}{2} t_{1k} t_{3k} \right) \end{aligned} \right\} \quad (8)$$

None of the  $a_k, b_k, d_k, e_k$  terms explicitly depends on  $\eta$  and the reduced stress terms  $t_k$  are defined by the following formulas :

$$\left. \begin{aligned} t_{1k} &= \beta_k C_1 + Q_k C_2 \\ t_{2k} &= \beta_k C_2 + Q_k \\ t_{3k} &= \beta_k / Q \end{aligned} \right\} \quad (9)$$

Although all expansions come from very simple equations the detailed expressions remain relatively cumbersome, especially when the wavenumber  $q$  is

complex, which must be considered here since we study waves that intrinsically leak energy into the bulk.

## II DISPERSION CURVES

We consider a restricted frequency range, such that the SSBW wavelength along  $Z'$  is very near the corrugation period. Under such assumptions, we found convenient to turn the wavenumber  $q$  and angular frequency  $\omega$  into dimensionless quantities defined by the following formulas :

$$\begin{aligned} q &= Q(1 + \xi) \\ \omega &= QV_S(1 + \Omega) \end{aligned} \quad (10)$$

By substituting into the generic propagation equation for the  $k$ -th harmonic, we easily obtain the corresponding wavevector component along the  $Y'$  direction, normal to the average free surface :

$$\begin{aligned} \beta_k^\pm &= \frac{Q}{C_1} [C_2(1+k+\xi) \pm KG_k] \\ G_k(\xi, \Omega) &= \left[ (1+\Omega)^2 - (1+k+\xi)^2 \right]^{1/2} \\ K &= \sqrt{C_1 - C_2^2} \end{aligned} \quad (11)$$

Because  $K = \sqrt{C_1 - C_2^2}$  is always positive, no problem is ever encountered in the definition of  $K$ . Since we restrict our interest to a small frequency domain in the vicinity of the structure resonance,  $\xi$  and  $\Omega$  are very small with respect to 1, and, for any set of values  $(k, \xi, \Omega)$ , only one of the couple of  $\pm$  signs must be retained to get either a  $k$ -th STW-like harmonic that exponentially damps in the depth of the substrate (if  $k \neq -1$  and  $Im(Gk) \neq 0$ ) or bulk waves emitted from the surface towards the inner of the substrate (i.e.  $k = -1$  and  $Re(Gk) < 0$ ). At Ref.[2], we provided with a slightly more detailed discussion on the issue of choosing the most significant harmonics in a arbitrary frequency range. It was shown that, for the here-considered case, the most significant terms must be governed by  $k \in [-2, -1, 0]$ . Thus, we truncated the infinite suite of boundary conditions (7) into a 3 by 3 system, for which we proposed asymptotic analytical solutions (roots of a polynomial with real coefficients), provided that  $\eta \rightarrow 0$ . This approach led us to four well-organized solutions, in complex conjugate pairs, which « degenerated » into distinct real roots when the frequency increased.

Although this approach seemed altogether consistent, we now consider that it is not at all satisfactory, since, after many attempts and detailed program checking, we have found that the proposed 3 by 3 determinantal equation actually does not vanish. The determinant may only take minimal but non-zero values in the considered domain. Conversely, we now rely on completely numeric solutions that do make sense, provided that we take at least 5 modes (namely  $k \in [-3, -2, -1, 0, 1]$ ) or 7 modes (namely  $k \in [-4, \dots, 2]$ ) in the COM model. The dispersion curves, roots of an implicit equation  $F(\Omega, \xi) = 0$ , are then obtained with help of Newton's method in the bivariate case. They are totally similar in shape for both cases (5 or 7 harmonics), although very small differences are observed for the most significant amplitude coefficients  $A_k$  which are still obtained for  $k = -2, -1, 0$ . We give on Fig 1a and Fig 1b the set of numerically obtained dispersion curves (respectively for the real and imaginary part of  $\xi$ ), in solid lines, simultaneously with our former results in dashed lines. In order to get a meaningful comparison, we stick to the same input data as in Ref[2], indeed, i.e. a  $42.75^\circ$  ST cut of quartz, with corrugation period  $\Lambda = 10^{-5}$  m, and  $\eta = 0.08$ , which must be thought as rather high value. The two branches appearing on Figs 1 describe almost - but not exactly - opposite solutions, each one associated with a definite sign of  $Im(\xi)$ . The unsufficiency of our previously proposed solutions seems to come from the fact that the high order powers of  $\eta$  that were discarded, though negligible in compare with retained terms, are not negligible in compare with the combination of them. The computer programs leading to newly presented results have been submitted to sufficient debugging and cross-checking to ensure the reliability of here-presented results.

## III COUPLER WITH FINITE DIMENSIONS

We denote by  $\xi_+$  (resp.  $\xi_-$ ) the branch characterized by  $Im(\xi) > 0$  (resp.  $Im(\xi) < 0$ ), that exponentially damps in the  $Z' > 0$  direction (resp.  $Z' < 0$ ). The two branches must be understood as a pair of complementary eigen-solutions appearing in a corrugated region of semi-infinite length, beginning at  $Z = 0$  (the question of origin is somewhat arbitrary) and extending towards  $Z' \rightarrow \infty$  (resp.  $Z' \rightarrow -\infty$ ). Because the actual achievement of a SH-BAW  $\Leftrightarrow$  STW reciprocal transformer requires a corrugated finite region, it is natural and even more necessary to seek the solution in a finite coupler as a combination of both  $\xi_+$  and  $\xi_-$  branches in order to satisfy the boundary conditions at both ends of the transformer. This

approach is clearly outlined in Chapter 8 of of Ref[2], in which the authors give detailed insight on a Rayleigh waves  $\leftrightarrow$  BAW coupler in an isotropic medium. Following the same general approach, we propose some results for the problem of an incident SH bulk wave coming under normal incidence from the depth of substrate onto the corrugated surface. We must then consider the following linear combination in the coupler:

$$\begin{aligned}
 u(\Omega) = & \lambda \sum_k A_k^+ e^{iQ(1+k+\xi^+)z + i\beta_k^+ y} \\
 & + \mu \sum_k A_k^- e^{iQ(1+k+\xi^-)z + i\beta_k^- y} \\
 & + \sum_k A_k^\otimes e^{iQ(1+k+\xi^\otimes)z + i\beta_k^\otimes y}
 \end{aligned} \quad (12)$$

This linear combination is governed by two unknown coefficients  $\lambda$  and  $\mu$  that can be determined by two end-conditions at  $Z'=-L/2$  and  $Z'=+L/2$ ,  $L$  being the actual length of the corrugated region. Supposing a coupler centered around the origin does not imply any loss of generality. Formula (13) simply means that the solution is a separately-weighted sum over both  $\xi^+$  and  $\xi^-$  branches of the dispersion curves ( that provide with normalized amplitude ratios  $A_k^\pm$  for the eigenmodes together with  $\xi^\pm$  in terms of  $\Omega$  ), and over a « forced » solution, indicated by the  $\otimes$  sign. This latter solution is the particular series of Bloch functions that obeys the system of {propagation equations}/{boundary conditions on the outer surface}, with some peculiarities : in this solution, both kind of incoming to- and outgoing from-(the surface) SH waves occur, wherewhile the  $\xi^\otimes$  horizontal wavevector is fixed by the incidence angle of the SH incoming towards the free surface. As a more concrete example, if we consider 5 modes, the  $\otimes$  solution actually contains 6 partial waves, each associated with an amplitude coefficient  $A_k^\otimes$  : 4 STW-like harmonics, 1 SH bulk incident to the surface, and 1 SH emitted from the surface. The amplitude of the incident SH wave is supposedly known and thereby can be arbitrarily fixed to 1 and the corresponding term is put in the second member of the boundary conditions. In addition, the vertical components of the SH bulk harmonics are exactly  $\beta_k^\otimes = \pm c\delta/\rho$ , together with  $\xi^\otimes = 0$  in case of normal incidence for the incoming SH wave. Thus, other  $A_k^\otimes$  are easily obtained by ordinary solving of a linear system with second member. The  $\lambda$  and  $\mu$  coefficients are ultimately obtained stating that the net

amplitude of the whole combination of all STW-like harmonics travelling towards  $Z'>0$  (i.e. with  $k>-1$ ) vanishes at the left end  $Z'=-L/2$ , whereas the net amplitude of all ( $k<-1$ ) harmonics vanishes at the right end  $Z'=L/2$ . This means that no energy is injected in the coupler besides the incoming SH bulk wave. Fig 2 gives the amplitude of the most important -2 and 0 harmonics in terms of the reduced frequency with coupler lengths  $L=20\Lambda$  (solid) and  $L=5\Lambda$  (dashed). As is previously mentioned in Res [6,2], a relatively similar approach can be used to model the reverse transformation of a STW that enters a limited grating into transmitted and reflected STWs, and bulk-SH waves penetrating into the substrate.

Fig 3 shows the reflection coefficients of incoming STW waves on a semi-infinite coupler, in both possible cases (extension towards  $Z'>0$  and  $Z'<0$ ). We present only the curves concerning the -2 and 0 harmonics, though the analysis was performed with 7 retained modes. It can be observed that the reflection coefficients along  $Z+$  and  $Z-$  directions are not equal, due to crystal anisotropy.

The here-presented approach is submitted to limitations due to additional scattering effects involved at both ends of the coupler, especially when the structure of STWs possibly existing in the outer regions neatly differs from the structure of STW-like harmonics inside the coupler. We consider that these issues are out of the scope of this paper.

#### ACKNOWLEDGEMENTS

Although this work is presently still unsupported, B. Dulmet is thankful for useful early discussions and encouragements by V. Plessky.

#### NOTES AND REFERENCES

- [1] L.M.Brekhovskikh, Sov.Phys.Acoust.,5,p288,(1959)
- [2] B.Dulmet, H. Watchueng, Proc of IEEE Ultrasonics Symposium, Seattle, (1995).
- [3] R.F. Humphries,E.A.Ash, Electronic Letters, 5, p175, (1969).
- [4] I.D.Akhromeeva, V.V. Krylov, Sov. Phys. Acoust, 23, p292, (1977).

[5] S.R. Seshadri, J. Appl. Phys. 73,8,p3637 (1993).

[6] S.V. Biryukov, Y.V.Gulyaev, V.V.Krylov, V.P.Plessky, "Surface Acoustic Waves in Inhomogeneous Media", Springer Verlag, (1995).

[7] S.R.Seshadri, J.Appl.Phys., 54, 6, p3017, (1983)

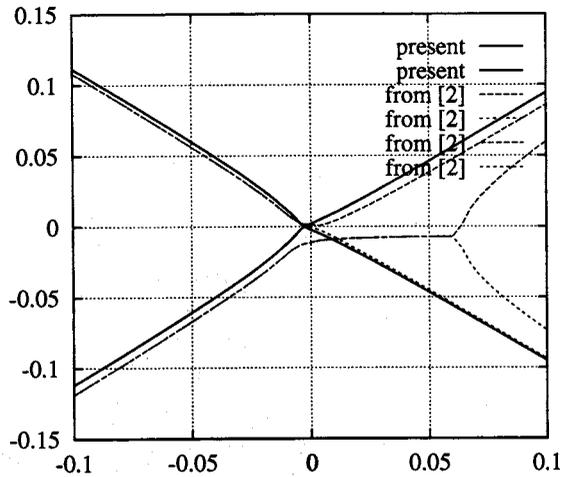


Fig.1a:  $\text{Re}(\xi)/\eta^2$  vs  $\Omega/\eta^2$   
Real part of the dispersion curves.

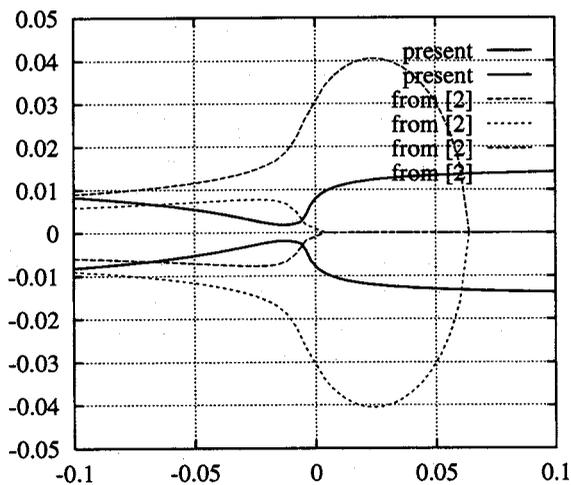


Fig.1b:  $\text{Im}(\xi)/\eta^2$  vs  $\Omega/\eta^2$   
Imaginary part of the dispersion curves.

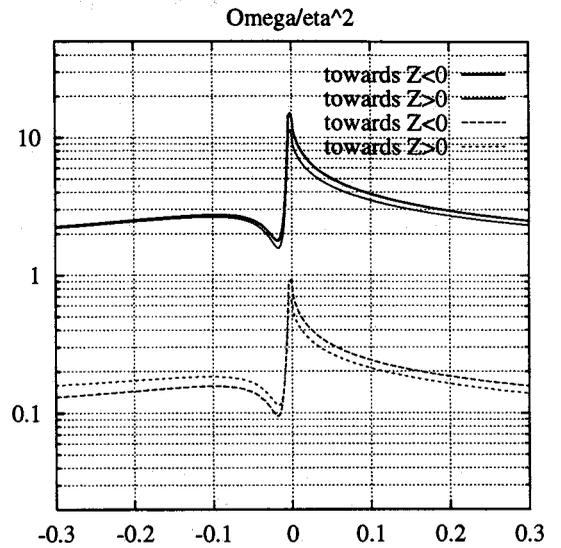


Fig.2: STW amplitude (module) launched by incident shear bulk wave on a coupler with finite length  $L=20\lambda$  (solid) and  $L=5\lambda$  (dashed)

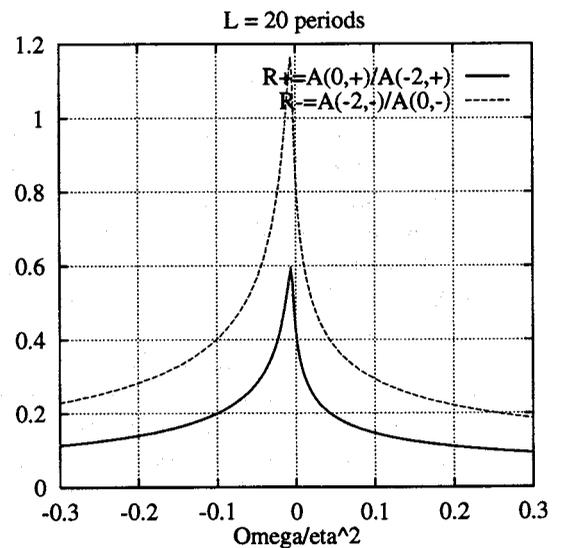


Fig.3: Module of reflection coefficients for the largest STW-like harmonics in a semi infinite coupler.

## SENSITIVITY OF INDUCTANCES TO MAGNETIC FIELDS

R. Brendel\*, G. Claudel\*, P. Bavière\*\*

\* Laboratoire de Physique et Métrologie des Oscillateurs du CNRS associé à l'Université de Franche-Comté  
32 avenue de l'Observatoire - 25044 Besançon Cedex - France

\*\* Centre National d'Etudes Spatiales 18 avenue Edouard Belin - 31055 Toulouse Cedex

### ABSTRACT

This paper presents the method used to measure the magnetic sensitivity of inductances. The measurement is performed by using an LC oscillator whose frequency is determined by the inductance to be measured. Because the records are often perturbed by noises and drift the data are processed so as to enhance the signal-to-noise ratio and to suppress the drift. A wide range of inductances have been measured. The results show that, according to their technology, for a magnetic field of 10 Gauss amplitude, the sensitivity may go from 0.1 ppm to 1400 ppm with sometimes a large spread of values in a given batch.

### 1. INTRODUCTION

For a long time designers have been concerned with the environmental sensitivity of the devices they have to build. The sensitivity of the electronic devices such as oscillators depends on the sensitivity of their components. Unfortunately, except for temperature dependence and sometimes for voltage dependence, the manufacturers are often discreet about the environmental sensitivity of their products. For several years, attention has been focused on the magnetic sensitivity of quartz crystal oscillators especially those developed for space applications (1-3). Recently, measurements have shown that environmental sensitivity mainly comes from either inductances used for mode selection or from the resonator itself (4-9).

Other experiments have shown that the sensitivity of inductances is more or less pronounced according to their technology. In order to clarify this particular point, 8 batches of inductances of various technologies have been measured. In all cases, their values were about 1  $\mu\text{H}$  enabling easy comparisons.

The experimental setup used is derived from the setup for the measurement of the magnetic sensitivity of quartz resonators (8).

### 2. EXPERIMENTAL SETUP

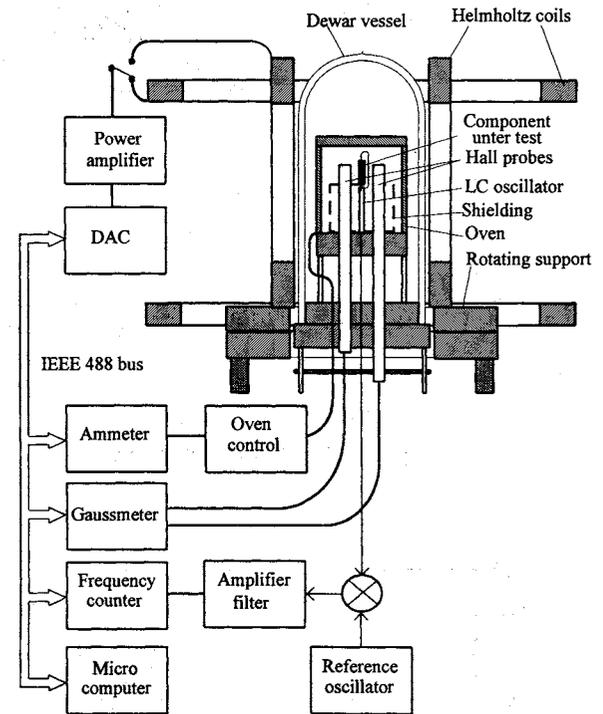


Figure 1. Experimental setup for magnetic sensitivity measurements

Figure 1 shows the experimental setup used for the measurement of the inductance magnetic sensitivity. The component under test sets the frequency of an LC oscillator in which it takes place. It can be shown that the frequency of the oscillator is given

by  $f = \frac{1}{2\pi\sqrt{LC}}$  where C is the capacitor of the

oscillator resonant circuit so that the relative variation of the inductance can be obtained from the relative frequency variation measurement by using

$$\text{the relation : } \frac{\Delta f}{f} = -\frac{1}{2} \frac{\Delta L}{L}$$

The oscillator and the inductance are put in an oven at a temperature about 40°C. When the equilibrium is reached, the temperature stability of the oven is better than  $10^{-4}$  °C over the experimental duration. To insure that only the magnetic sensitivity of the inductance is measured, the remaining part of the

oscillator can be shielded by a mu-metal enclosure which reduces the magnetic field by a factor about 50. Experiments have shown that there is no perceptible difference when the shielding is in action or not so that in most cases it was not used because of the strong magnetic field inhomogeneity induced by the mu-metal enclosure. In a given experiment, the component under test is held in a fixed position regardless of the magnetic field orientation as recommended by the IEEE standard (10). In order to give the magnetic field vector different orientations with respect to the component, the magnetic field is produced by using two pairs of Helmholtz's coils. One of them is fixed to produce a vertical magnetic field along the symmetry axis of the oven which is usually the symmetry axis of the component (Fig. 2), the other pair can go round the symmetry axis thus enabling any radial orientation of the magnetic field in a plane normal to the symmetry axis. The magnetic field is measured by means of a gaussmeter using two Hall effect probes located near the component under test, one of them is for axial field and the other one for radial field measurement. A microcomputer drives the current in the coils through a DAC and a power amplifier thus enabling various shape and amplitude of the magnetic field.

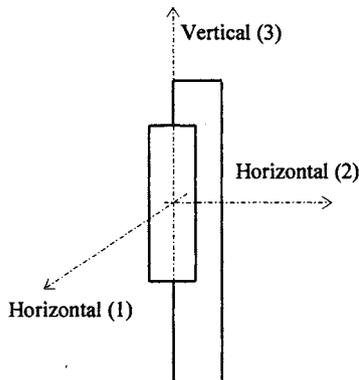


Figure 2. Convention for component orientation.

The frequency of the test oscillator is measured by using the popular method represented in Fig. 1. The test oscillator is compared with a reference oscillator by using a balanced mixer, a low noise amplifier and a low pass filter. The beat frequency is usually in the 1 kHz range. This frequency variation as well as the magnetic field and the oven heating current are simultaneously stored as a function of time in the microcomputer. All the experiment is automated by using a IEEE-488 bus interface.

### 3. EXPERIMENTAL PROCEDURE AND DATA PROCESSING

In order to improve the signal-to-noise ratio, the magnetic field is given a low frequency sinusoidal

modulation. The period of the magnetic field excitation 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 data processing. Some preliminary experiments performed in various conditions have shown that an excitation period of 100 s and a sampling time of 1 s constitute a satisfactory compromise.

Because of the periodic character of the magnetic field excitation and the frequency response, it is possible to average the data over a number of periods (usually several tens). In fact it is well known that in such a case, the signal-to-noise ratio increases as the square root of the cycle number. In addition, the low frequency variations due to the temperature drift can be partially removed by performing a first order regression on each period of the excitation signal.

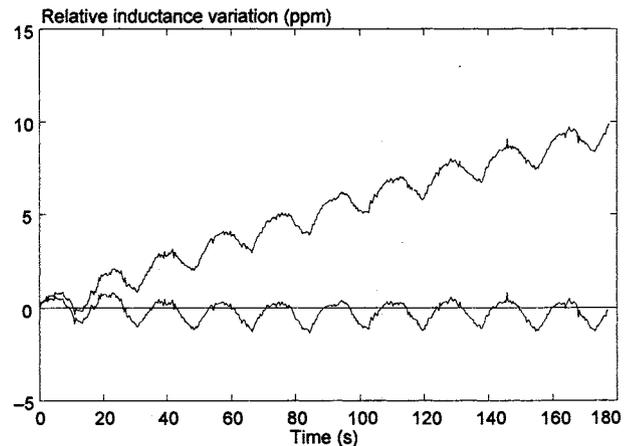


Figure 3. Effect of the drift removal by linear regression.

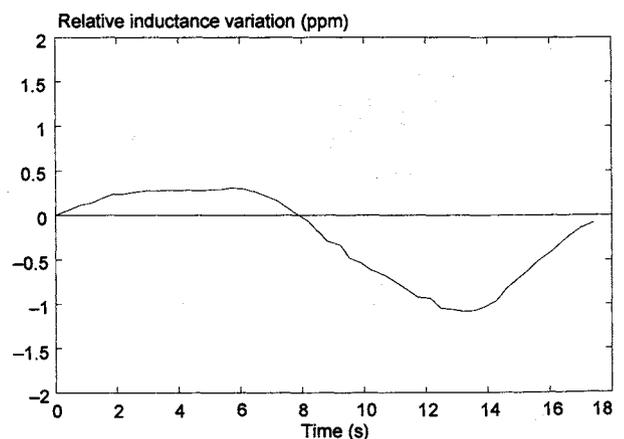


Figure 4. Enhancement of the signal-to-noise ratio by averaging over several tens periods.

Figure 3 shows the effect of the drift removal when the mean axis of each period is brought back on the horizontal axis and Fig. 4 shows the result obtained after the averaging process has been performed.

The sinusoidal magnetic field is successively given the amplitude 3, 5, 10 and 20 gauss.

#### 4. SAMPLES

Inductances of various shape and constitution have been measured, their main electrical characteristics are summarized in Table I and their geometric shape is shown in Fig. 5.

Table I. Electrical characteristics of the samples and results.

Batch #	Shape	Core	L ( $\mu$ H)	Turns	Rel. var. for 10 gauss (ppm)
1	rod	ferrite	1.01	6	750 - 1450
2	rod	neutral	1.13	25	< 0.2
5	pulley	iron powder	1.11	19	125 - 220
6	pot	ferrite	1.21	6	1200 - 1400
8	toroid	neutral	1.23	56	< 0.25
9	toroid	neutral	1.08	34	< 0.2

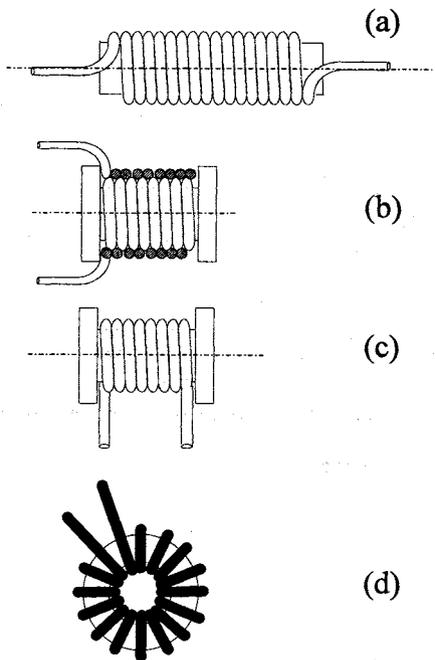


Figure 5. Various shapes of coils used in the experiments: (a) rod, (b) pulley, (c) pot, (d) toroid.

#### 5. RESULTS

Each kind of inductance has been measured by applying the magnetic field along the three main directions of the component as defined in Fig. 2 so as to determine the greatest sensitivity axis for each batch. Most of the results presented in this paper have been obtained with a magnetic field directed along the greatest sensitivity axis. Table I

summarizes the relative inductance variation measured for a 10 gauss amplitude magnetic field for each batch, when the magnetic field is directed along the greatest sensitivity axis.

Figures 6 to 11 show, for each kind of inductance, the typical behaviour of the component for 4 values of the magnetic field amplitude.

It should be noted that neutral core inductance exhibit a very low sensitivity, less than 0.2 ppm for an amplitude of 10 gauss. At the present time it is not obvious that the sensitivity observed with these components comes effectively from the inductance under test. It is possible that it comes from a residual sensitivity of the measurement setup. This point is not yet clarified.

Figures 12 to 16 show the spreading of the sensitivity observed for different batches.

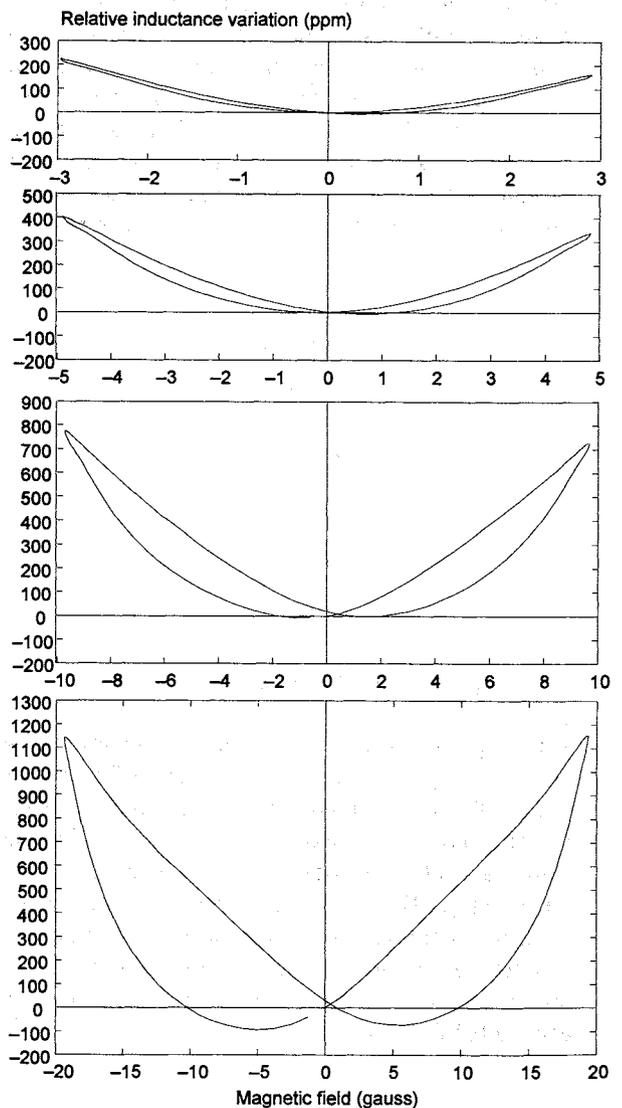


Figure 6. Ferrite core rod, batch #2, sample #2, magnetic field along rod axis.

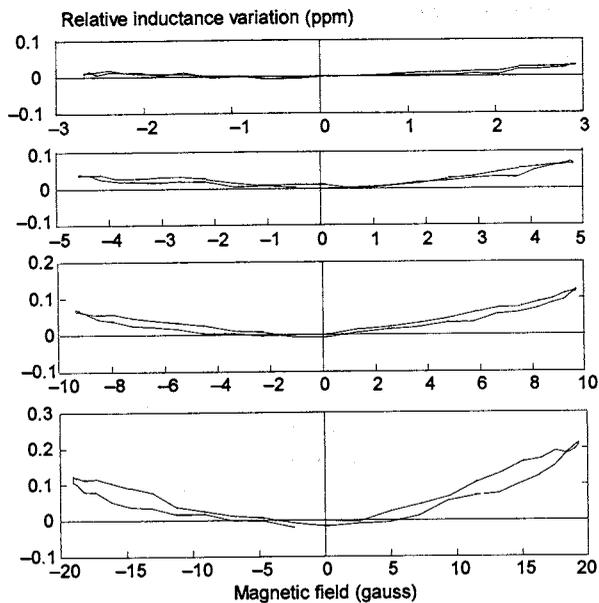


Figure 7. Neutral core rod, batch #2, sample #2, magnetic field along rod axis.

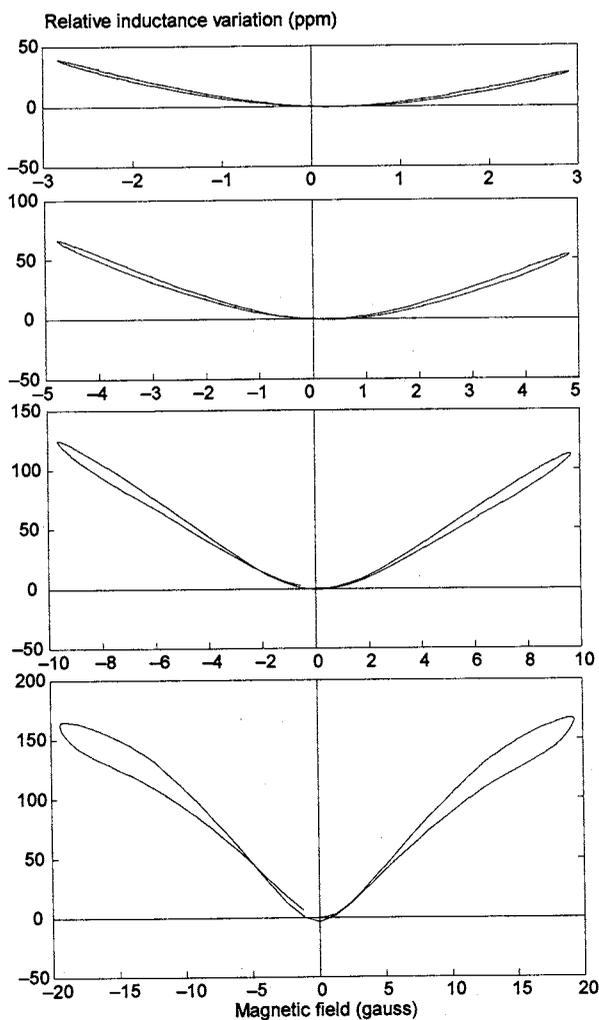


Figure 8. Iron powder core pulley, batch #5, sample #1, magnetic field along pulley axis.

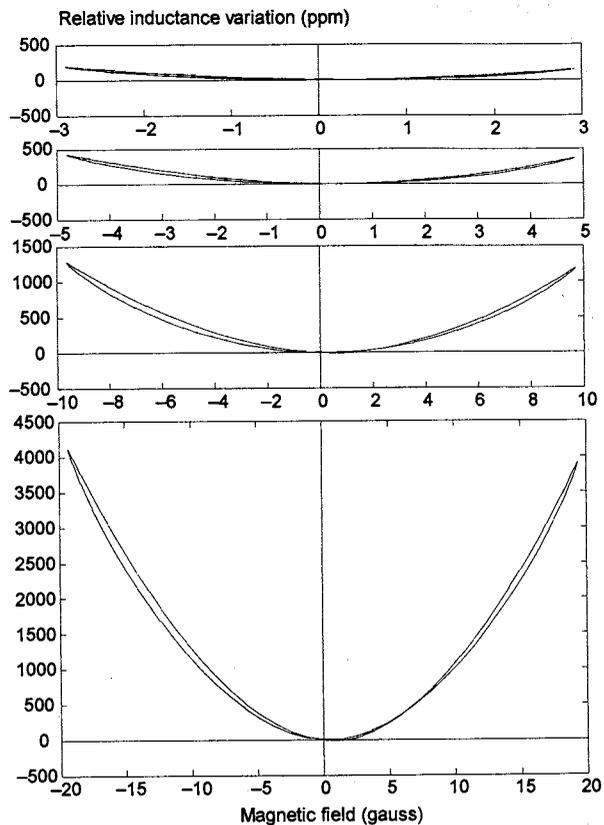


Figure 9. Ferrite core pot, batch #6, sample #1, magnetic field along pot axis.

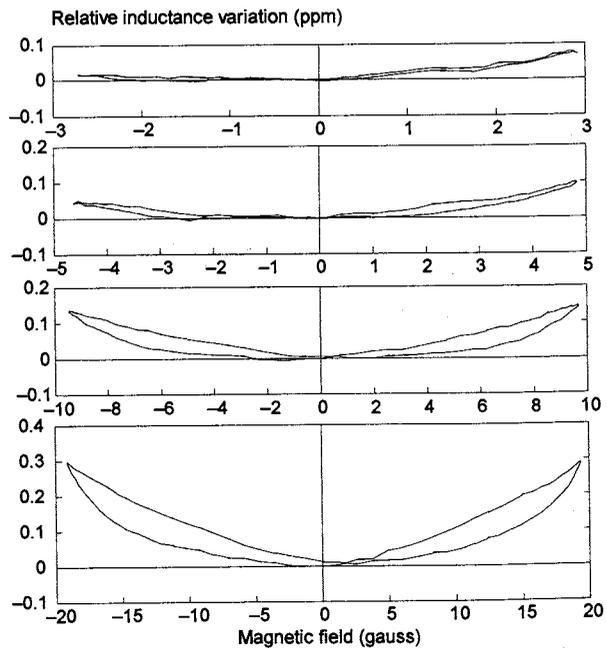


Figure 10. Neutral core toroid, batch #8, sample #3, magnetic field along toroid axis.

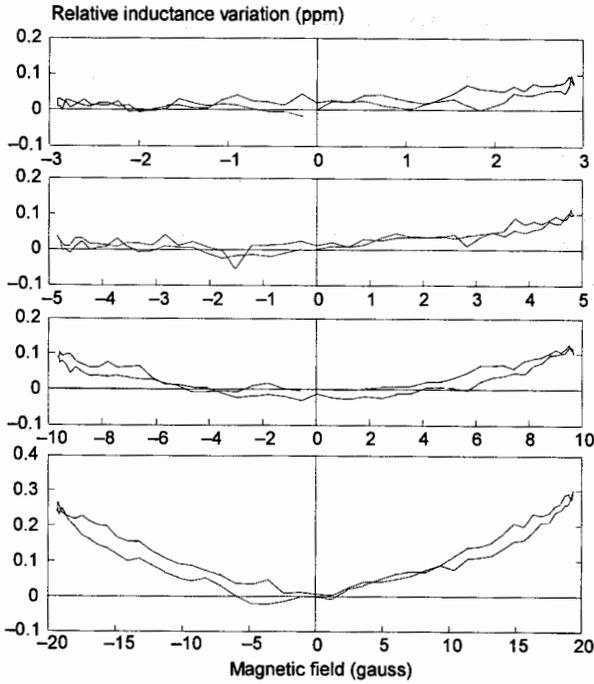


Figure 11. Neutral core toroid, batch #9, sample #3, magnetic field along toroid axis.

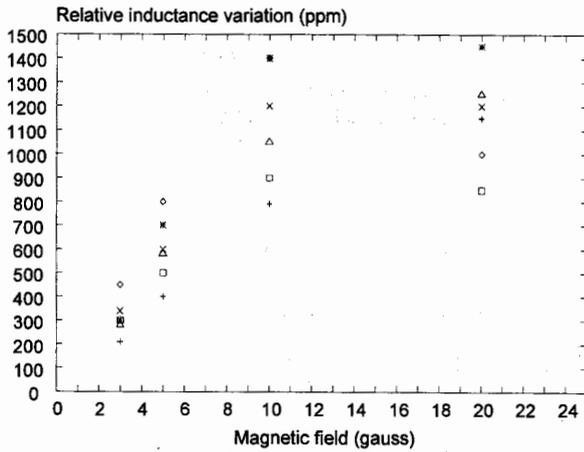


Figure 12. Ferrite core rods, batch #1. A large spread is observed among the different samples.

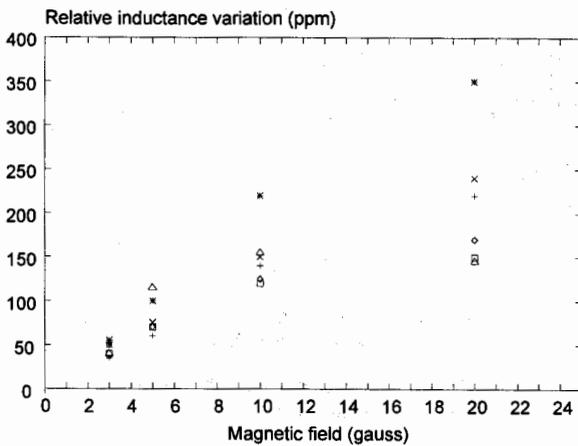


Figure 13. Iron powder core pulley, batch #5. The sensitivity is lower than ferrite core.

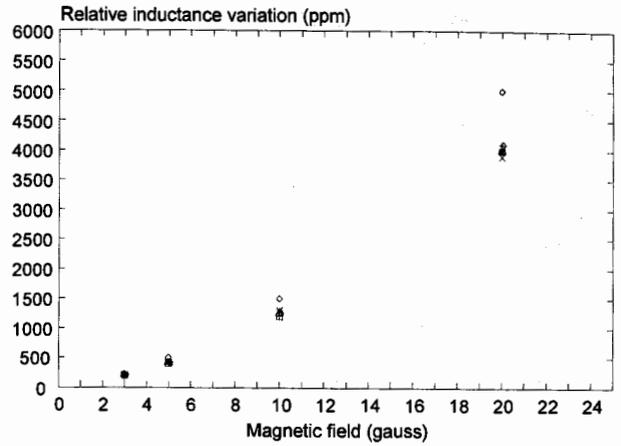


Figure 14. Ferrite core pot, batch #6. This batch exhibits the highest sensitivity.

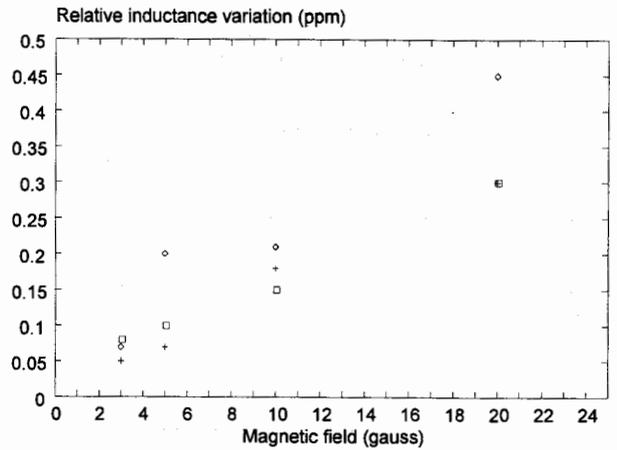


Figure 15. Neutral core toroid, batch #8. The sensitivity is quite low and the spread may be affected by the measurement setup resolution.

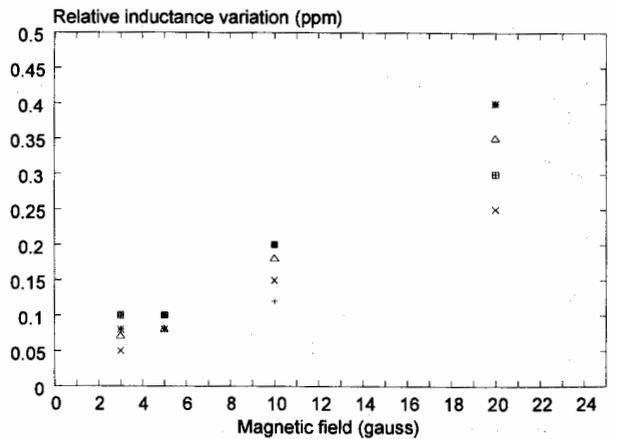


Figure 16. Neutral core toroid, batch #9. Same remarks as in figure 14.

## CONCLUSION

The figures and Table I presented in this paper show that ferrite core inductances present a very large sensitivity to magnetic field. This is not quite surprising, but the order of magnitude of this sensitivity (about 1000 ppm for 10 gauss) has to be taken into account when developing circuit exposed to magnetic field. Iron powder core inductances are somewhat less sensitive by a factor about 7 but remain much more sensitive than neutral core inductances which exhibit practically no magnetic sensitivity.

Obviously, these latter component should be preferred each time a system including inductances is susceptible to be exposed to magnetic field.

## ACKNOWLEDGEMENTS

The authors wish to thank the Company Microspire for providing the inductances.

This work was supported by CNRS and CNES under contract 844/2/94/CNES/0089.

## REFERENCES

1. Benezet A., Robert E., 1988, "Mesure de susceptibilité au champ magnétique de plusieurs OUS", Compte-rendu d'exécution d'essais n° L5160, CNES.
2. Robert E., Brunet M., 1988, "Susceptibilité magnétique des oscillateurs à quartz spatiaux, état des connaissances", Technical Report n° 88/354, CNES.
3. Deyzac F., 1988, "Oscillateurs à quartz de bord en environnement sévère. Sensibilité magnétique", Technical Report n° 37/7132 PY, ONERA.
4. Deyzac F., 1990, "Magnetic sensitivity of quartz oscillators", 4th EFTF, Neuchâtel, Switzerland.
5. Brendel R., El Hassani C., Brunet M., Robert E., 1989, "Influence of magnetic field on quartz crystal oscillators", 43rd Annual Frequency Control Symposium, Denver, USA.
6. Brendel R., El Hassani C., Krux P., Robert E., 1991, "Magnetic sensitivity of oscillator components", 5th EFTF, Besançon, France.
7. Brendel R., Robert E., 1992, "Weak magnetic field sensitivity of quartz crystal oscillators", 6th EFTF, Noordwijk, Netherlands.
8. Brendel R., Cretin B., El Habti A., 1994, "Origin and measurement of quartz resonator magnetic sensitivity", 8th EFTF, Munich, Germany.
9. Brendel R., "Influence of a magnetic field on quartz crystal resonator", submitted to IEEE Trans. on UFFC.
10. IEEE-SCC-27 Working group, 1994, "Guide for measurement of environmental sensitivity of standard frequency generators", Proposition PAR-P1193.

# HIGH-PRECISION IC QUARTZ OSCILLATOR

S Kovita, A Myasnikov

Russian Institute of Radionavigation and Time, St.-Petersburg, Russia

A Isakov

St.-Petersburg Elektrotechnical University, Russia

A Borisenko, O Sokolov

St.-Petersburg Technical University, Russia

## ABSTRACT

Development of satellite communication and navigation, along with adoption of synchronous techniques for data transmitting and processing, promotes an enhanced demand for compact high-stability quartz oscillators of TCXO and VCXO-types. Solving a topical problem - to minimize quartz oscillator dimensions - is presented based on development of dedicated IC. Considered is an approach to realization of self-excited oscillator using differential amplifiers and model of two-terminal networks with a negative impedance. Described are IC main stages.

The reduction of TCVCXO volume down to  $15 \text{ cm}^3$  is attained at instability for frequency of (3-20) MHz with Allan's variance up to  $(2-5) \cdot 10^{-12}$ .

## INTRODUCTION

In spite of significant progress in development of high precision quartz oscillators (HPQO), achievements in improvement of operational parameters (of which the most important are dimensions, consumption power and warm-up time) are quite conservative and restrict quality factors of up-to-date radio systems and complexes.

One of promising ways to solve this problem is to use dedicated analog IC. However the restrictions imposed by solid-state technology on components and parameters of HPQO circuits need a special approach to design such IC.

In this work it is proposed to develop a dedicated solid-state circuit of self-excited oscillator, in order to connect to it a crystal resonator without any inductance or other matching components, whose reproducibility within a solid-state circuit is limited by production. In so doing, it is proposed to use a complex nature of input impedance of active networks to realize a proper reactance for excitation of a quartz resonator.

IC self-excited oscillator version given in this work is

not based on conventional circuitry [1], but on building a circuit with differential amplifiers. Such devices are widely used in microcircuits due to its flexibility, versatility and convenience for realization in IC form [2].

A high-precision IC quartz oscillator consists of three assemblies:

- Oscillator Circuit, to which a crystal resonator is directly connected;
- Bias Controller providing an initial setting and stabilizing operation of all IC stages;
- Output Stage amplifying an output signal power and providing Oscillator Circuit isolation from load changes.

## OSCILLATOR CIRCUIT ANALYSIS

A self-excited crystal oscillator is realized as an asymmetric differential amplifier with two loops of voltage feedback (Fig.1): positive across the R7 resistor and negative across the R6 resistor. In order to decrease a gain of in-phase signal component, the Q3, a stable current generator, is introduced into emitters of Q1 and Q2 transistors. This generator regulates a current that enables to regulate an excitation power of crystal resonator.

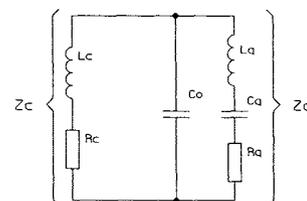


Fig. 1.

A crystal resonator is connected to the inverting input of differential amplifier in negative feedback loop. This makes it possible for an oscillator to operate at a frequency near to  $\omega_1$  frequency of series resonance without additional matching circuits.

It is possible to consider a connection of crystal resonator to the differential amplifier input as a connection to a two-terminal network with a negative resistance. A differential amplifier input impedance with proper interval feedbacks is used as a two-terminal network. In this case, such an equivalent network has S-type volt-ampere curve with a negative differential resistance and a reactance of an inductive nature at a falling part of the curve.

An equivalent circuit of a crystal resonator connected to the differential amplifier is given in Fig.2, where  $Z_c$  is self-excited oscillator impedance consisting of active component of  $R_c^{(-)}$  resistance and  $L_c$  equivalent input inductance;  $Z_q$  is equivalent impedance of crystal resonator including dynamic inductance  $L_q$ ,  $C_q$  capacitance,  $R_q$  loss resistance and  $C_o$  static capacitance.

$Z_q$  can be presented as:

$$Z_q = \frac{R_q}{(1-\varepsilon)^2 + D^{-2}} + j \frac{\varepsilon(1-\varepsilon) - D^{-2}}{\omega C_o [(1-\varepsilon)^2 + D^{-2}]}, \quad (1)$$

where  $\varepsilon = \frac{\omega^2 - \omega_1^2}{\omega_2^2 - \omega_1^2} \sim \frac{\omega - \omega_1}{\omega_2 - \omega_1}$

- generalized normalized detuning which sets a deviation of  $\omega$  generated frequency from  $\omega_1$  using a relation between frequencies of parallel  $\omega_2$  and series  $\omega_1$  resonances;

$D = (\omega_1 C_o R_q) - 1$  - is figure of merit.

Generalized detuning  $\varepsilon$  is connected with a fractal detuning

$$\frac{\Delta\omega}{\omega} = \frac{\omega - \omega_1}{\omega} \quad \text{through} \quad \varepsilon = 2 \frac{C_o}{C_q} \frac{\Delta\omega}{\omega}$$

From a condition for self-sustaining oscillations  $Z_c + Z_q = 0$  and considering for (1), the equations for amplitude and phase equilibrium will be as follows:

$$|R_c^{(-)}| > \frac{R_q}{(1-\varepsilon)^2 + D^{-2}}; \quad (2)$$

$$\omega_1 L_c = - \frac{\varepsilon(1-\varepsilon) - D^{-2}}{(1-\varepsilon)^2 + D^{-2}} \frac{1}{\omega_1 C_o}. \quad (3)$$

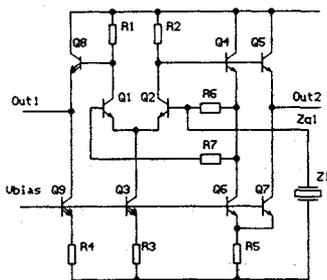


Fig. 2.

Expressions (2) and (3) define the values of differential amplifier negative resistance for providing and maintaining self-sustaining oscillations and its frequency (generalized detuning  $\varepsilon$ ) in steady-state behaviour.

Condition for stability of self-sustaining oscillations is  $\delta\phi_q/\delta\omega > 0$ , where  $\phi_q$ , a phase of impedance  $Z_q$ , restricts generalized detuning by an area with  $\varepsilon < 0.5$ . Frequency dependence of  $IC R_c^{(-)}$  and  $L_c$  has a trend to decrease as frequency increases. At a certain frequency  $\omega^*$  a nature of input reactance changes becoming capacitive.

Components of differential amplifier and its mode of operation are chosen to realize condition (2) at a frequency  $\omega \sim \omega_1 < \omega^*$  within an area with an abrupt change  $R_q^{(-)}$  near  $\omega^*$ . This enabled to eliminate an oscillator excitation at spurious frequencies exceeding  $\omega_1$  by (8-12) % and being typical for crystal resonators of two-rotation cuts, for example, SC-cut.

**BIAS CONTROLLER**

All the current sources for oscillator and output stage are biased from a single internal voltage reference. A schematic of the bias controller is shown in Fig.3.

The bias controller consists of a band-gap reference and turn-on circuit. Initial turn-on current for the band-gap reference transistors Q4-Q7 is provided by a collector Q3. When Q3 conducts, it turns on Q9 which supplies the bias bus. In this circuit, transistor Q4 is operated at a relatively high current density. The current density of the transistor Q5 is about 10 times lower and the emitter-base voltage differential  $\Delta V_{be}$  between the emitters of these devices appears across  $R_1$ . The emitter-base voltage differential between two transistors operating at different current densities is given by [1]

$$\Delta V_{be} = \frac{kT}{q} \ln \frac{J_4}{J_5},$$

where  $J$  is the current density,  $q$  is the charge of an electron,  $k$  is Boltzmann's constant,  $T$  is absolute temperature.

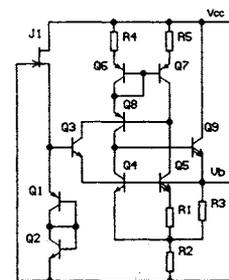


Fig. 3.

The bias voltage  $V_b$  is composed of emitter-base voltage,  $V_{be}$ , plus a voltage drop across which is  $R_2$  proportional to  $\Delta V_{be}$  and  $V_{be}$

$$V_b = V_{be} + \frac{R_2}{R_1} \frac{kT}{q} \ln \frac{J_4}{J_5} + \frac{R_2}{R_3} V_{be}, \quad (4)$$

where  $V_{be}$  is the emitter-base voltage of Q4. Referring to (4), the resistors  $R_1$ - $R_3$  determine a temperature drift bias voltage.

The bias voltage  $V_b$  will give a constant output current from an n-p-n current sources biased from it, compensating for temperature variation of diffused-base resistors. This voltage may be routed through the IC using  $n^+$  cross-unders diffused into the isolation wall to simplify layout.

### OUTPUT STAGE

Schematic of the output stage is shown in Fig.4. This circuit provides a differential input and output with gain only determined by a resistor ratio. It can amplify input signals with very low distortion and wide bandwidth. The cross-coupled transistors Q3 and Q4 keep the drop across  $R_3$  equal to the differential input voltage. When the base of Q1 is higher than Q2, the current in Q1 is higher than the current Q2, causing their emitter-base voltages to be mismatched. Transistors Q3 and Q4 have the same mismatch in  $V_{be}$ , but because they are cross-coupled, their mismatch cancels the mismatch in Q1 and Q2 leaving the drop across  $R_3$  equal to the differential input voltage.

The collector current difference between Q1 and Q2 is proportional to the voltage drop in  $R_3$ , since  $R_3$  is the only path for current to flow from one side of this circuit to the other, except for a small base current error. This means that the difference in collector current between Q1 and Q2 is proportional to the differential input voltage. This gives a differential output voltage proportional to input voltage.

The advantage of all this is that the non-linear transistors transfer characteristics,  $g_m$ , has been linearized. Also, the voltage gain  $K_v = (R_1 + R_2)/R_3$  is determined only by a resistor ratio.

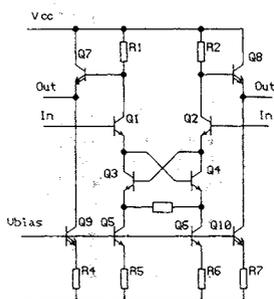


Fig. 4.

The output stage of this amplifier Q7 and Q8 is an emitter follower to obtain a high output current and a high load rejection.

### PERFORMANCE

A completed schematic of the IC quartz oscillator is provided in Fig.5. IC prototype for test conductivity within a crystal oscillator was manufactured using a basic array of KB1452XH1-4 type. IC tests have shown that as oscilltor maintained a stable excitation of AT- and SC-cut crystal resonators with  $Q$  from  $30 \cdot 10^3$  to  $2.5 \cdot 10^6$  at frequencies from 3 to 20 MHz when supply voltage changed from 4 V to 15 V. In this case, all the tested resonators were excited at fundamental resonant frequency without additional matching circuits.

Excitation power of a crystal resonator is regulated within a range of  $(0.5 - 4.5) \mu W$ , and an output signal level - within a range of  $(100 - 250) mV$ . IC consumption current is  $(5.5 - 7.0) mA$ .

Frequency test results for a quartz oscillator with the developed IC are given in Fig.6 for SC-cut crystal resonator with  $Q$  of  $2 \cdot 10^6$  at the the 5.0 MHz frequency.

A further IC improvement is based on introduction of device for automated control of crystal resonator excitation level [4]. In this case, an improvement of Allan's variance is expected due to decreasing a contribution into an instability of high-order harmonics of oscillator output signal spectrum.

### CONCLUSIONS

As a result of this work, a solid-state single-chip IC was designed and tested for compact TCVCXO for frequencies of (3-20) MHz with Allan's variance up to  $(2-5) \cdot 10^{-12}$  per 1s. Due to complex nature of input impedance, it was possible to provide a stable excitation of crystal resonators of two-rotated cuts at

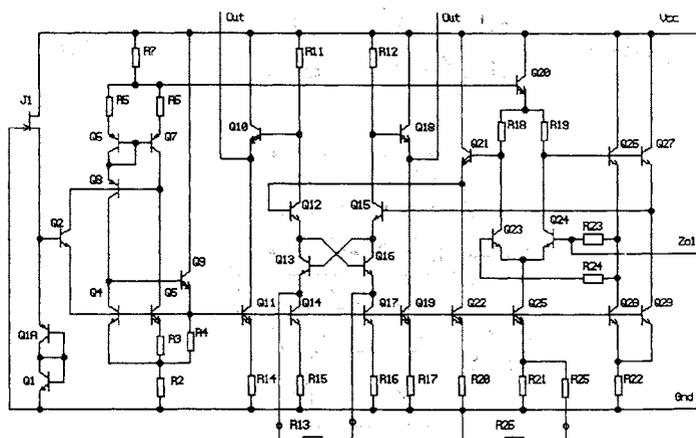


Fig. 5.

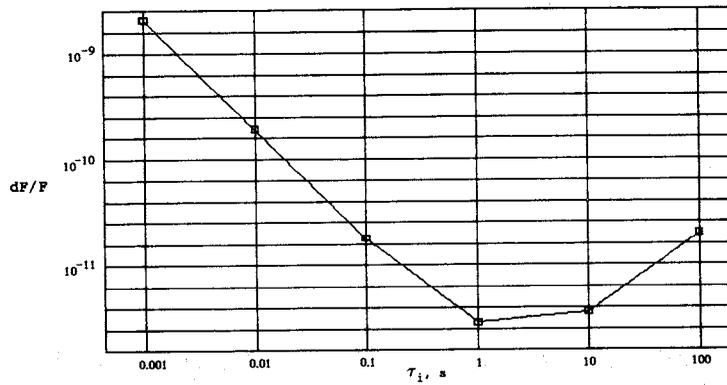


Fig. 6. Allan's variance of compact TCVCXO.

fundamental resonant frequency without special matching circuits for crystal resonator connection.

IC use enabled to reduce a volume of compact HPQO of TCVCXO up to  $15 \text{ cm}^3$ .

#### REFERENCES

- [1] Hildereth D E, "IC crystal oscillator". Ferranti Monochip Application Note, APN-4.
- [2] Borisenko A, Sokolov O, Kovita S, Myasnikov A, "Quartz oscillator". A.S. Russia No 1790026 from 09.22.92.
- [3] Greben A, "Bipolar and MOS analog integrated circuit design". A Wilay-Interscience Publication, New York, 1984.
- [4] Borisenko A et.al., "Quartz oscillator with automatic amplitude control", Patent Russia No 4940422 from 05.31.91.

## EFFECTS OF CRYSTAL MOUNT RESONANCE, UNDER RANDOM VIBRATION, ON THE CLOSE TO CARRIER PHASE NOISE.

Nigel D Hardy

C-MAC Quartz Crystals Ltd., UK

### ABSTRACT

The phase noise of quartz crystal oscillators under vibration is degraded, mainly due to the acceleration sensitivity of the quartz crystal element. It is shown that if a crystal mount resonance is excited by a random vibration profile then the phase noise close to the carrier, outside the random vibration profile, can be further substantially degraded. A consequence is that the vibration sensitivity would seem to increase rapidly at low vibration frequencies. For example, a crystal oscillator with a resonance at 1950 Hz subjected to a random vibration profile of 0.01 g<sup>2</sup>/Hz extending from 100 Hz to 3000 Hz compared with a profile of 100 to 500 Hz, where the resonance is not excited, can have over 40 dB degradation in the phase noise at 20 Hz from the carrier. The resonance level is shown to become non-linear with vibration power, possibly due to the large acceleration amplification of the resonance, 100 to 300 times. The vibration level where the resonance becomes non-linear varies by two orders of magnitude and starts as low as 0.0003 g<sup>2</sup>/Hz. This phenomenon is demonstrated for four oscillators and investigated under varying acceleration levels in an attempt to quantify the nonlinear behavior. The non-linear behaviour of the resonance is believed to produce intermodulation of the mechanical vibration which leads to the degradation of the close to carrier phase noise.

### INTRODUCTION

The acceleration sensitivity of quartz crystals has been well reported i.e. Fuller (1). The quartz crystal's resonant frequency moves when subjected to acceleration. This frequency shift, equation 1, is proportional to the magnitude of the acceleration and dependent on direction, giving rise to an acceleration sensitivity vector,  $\vec{\Gamma}$ .

$$f(\vec{a}) = f_0(1 + \vec{\Gamma} \cdot \vec{a}) \quad (1)$$

Under vibration the acceleration sensitivity causes frequency modulation of the oscillator. For a small modulation index the single sideband level  $\mathcal{L}$  is given by equation 2.

$$\mathcal{L}(f_v) = 20 \cdot \log \left[ \frac{(\vec{\Gamma} \cdot \vec{a}) f_0}{2 f_v} \right] \quad (2)$$

Where  $f_v$  is the vibration frequency,  $\vec{\Gamma}$  the acceleration sensitivity vector and  $\vec{a}$  the acceleration. Under random vibration the acceleration is described by its power spectral density *PSD*. The magnitude of the acceleration appropriate for substituting in equation 2 is:

$$|\vec{a}| = \sqrt{2 \cdot PSD} \quad (3)$$

When subjected to normal acceleration levels the acceleration sensitivity is independent of the acceleration level, i.e. the mechanism is linear. For a linear system the characterization of the acceleration sensitivity from a random vibration profile should be equivalent to the superposition of the acceleration sensitivity deduced from sinusoidal vibrations. A typical acceleration sensitivity would be fairly constant at low vibration frequencies extending until resonances in the crystal or oscillator mounting structures are encountered. At which point the apparent measured acceleration sensitivity can be many times higher than normal, due to the acceleration amplification of the resonance.

The phenomenon, described here of a resonance degrading the phase noise close to the carrier, was first observed by us on 46.72 MHz third overtone TCXO's. Good phase noise at 10 Hz when the device was subjected to random vibration from 20 to 2000 Hz was a requirement for these devices.

### METHOD

Four quartz crystal oscillators were measured under varying levels and types of random vibration profile. All measurements were made using the same random vibration controller, power amplifier and vibration table. The phase noise spectrum was measured using an HP3589 spectrum analyzer in its narrow band zoom mode. In this mode the spectrum analyzer mixes the

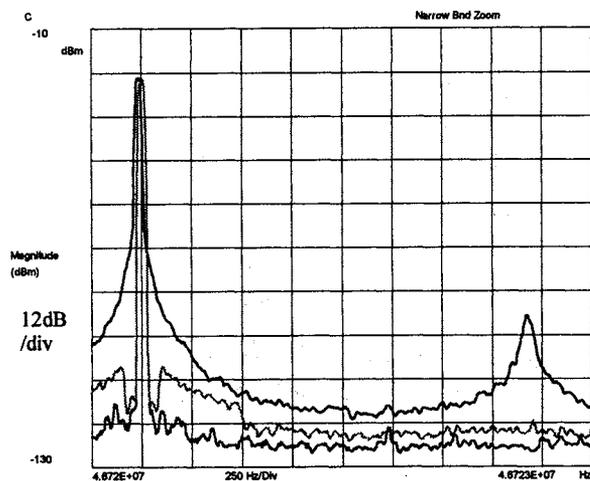
input down to baseband and performs an FFT. This gives good narrow band resolution and greater than 100 dB of dynamic range.

The quartz crystals contained in the oscillators consisted of: a 46.72 MHz third overtone in an HC45 holder, two supposedly identical 10.9493 MHz fundamentals in HC45 holders and a 16.8 MHz fundamental in a HC26 holder. All devices were compensated TCXO's and all measurements were made with the acceleration direction perpendicular to the plane of the quartz blank. No resonances in the vibration drive signal were observed indicating that any resonances observed in the oscillator phase noise were due to low mass items. All vibration profiles were a constant level between the frequencies specified and outside these frequencies the profile was attenuated as fast as the random vibration controller filters would allow.

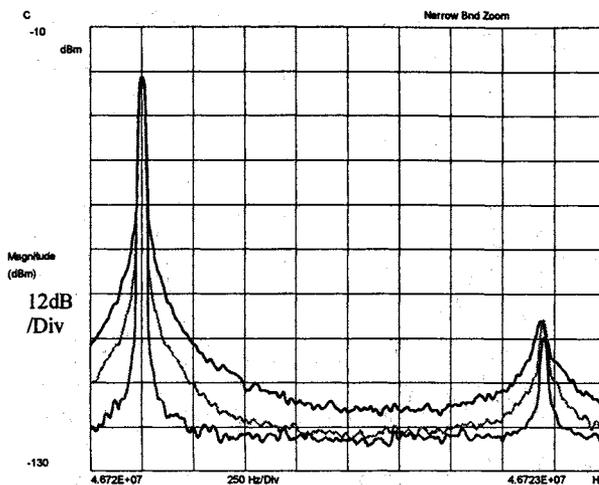
**EXPERIMENTAL RESULTS**

The 46.72 MHz oscillator was subjected to varying levels of vibration ranging from  $10^{-5}$  to  $10^{-1}$   $g^2/Hz$  and with two profiles 100 to 500 Hz and 100 to 3000 Hz. One to excite the resonance at 1950 Hz and one not. The comparison between the device not under vibration, the resonance excited and not excited is show in Fig 1. The degradation of the close to carrier phase noise and resonance level with vibration level is shown in Fig 2. The results show a single resonance at approximately 1950 Hz. When the vibration profile is limited to 100 to 500 Hz and the resonance is not excited the phase noise below 100 Hz is not degraded. But when the resonance is excited the phase noise below 100 Hz is degraded substantially, 40 dB or more.

No Vibration, 0.01  $g^2/Hz$  100 to 500Hz, 0.01  $g^2/Hz$  100 to 3000Hz  
 Figure 1: 46.72 MHz Phase noise under vibration

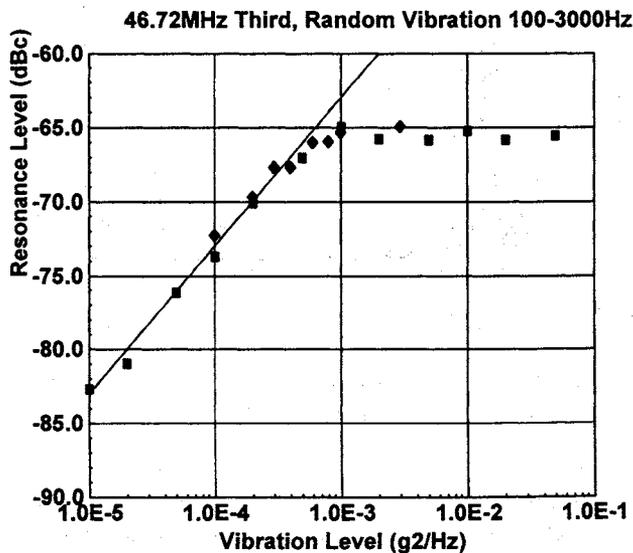


Vibration 100 to 3000Hz: 0.0002  $g^2/Hz$ , 0.002  $g^2/Hz$ , 0.02  $g^2/Hz$   
 Figure 2: 46.72 MHz with increasing vibration



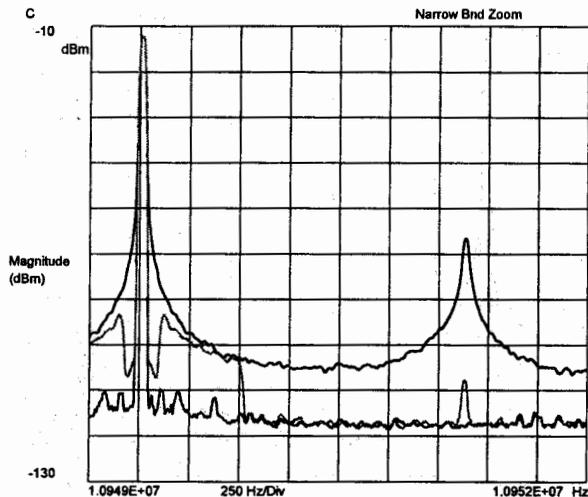
When the vibration level is increased, the resonance seems to saturate but the degradation of the carrier increases rapidly, Fig 2. The resonance level was measured as a function of vibration level, this resulted in linear behavior below  $\approx 3 \cdot 10^{-4}$   $g^2/Hz$ , a short non-linear region and then saturation above  $10^{-3}$   $g^2/Hz$ , Fig 3.

Figure 3: Resonance level vs. Vibration level

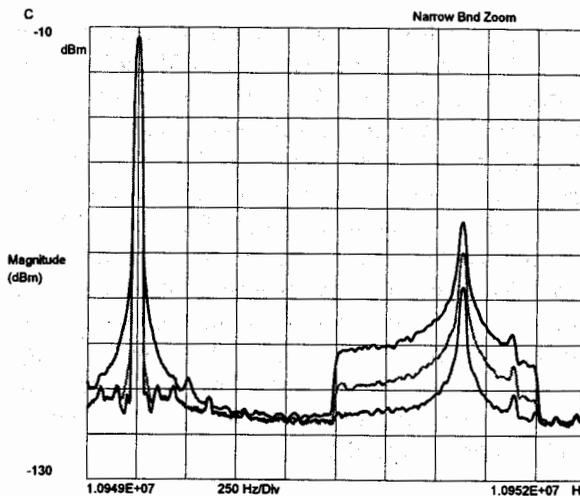


The measurements were repeated on an oscillator containing a 10.9493 MHz fundamental crystal again in an HC45 holder. The same phenomenon was observed, as shown in Fig 4, although the effect required over an order of magnitude increase in the vibration level to degrade the phase noise by a similar amount. To eliminate the possibility of any non-linearities in the measurement equipment due to differences in the total power applied, the device was measured using the same total power, in a 300 Hz band below, above and containing the resonance, Fig 5.

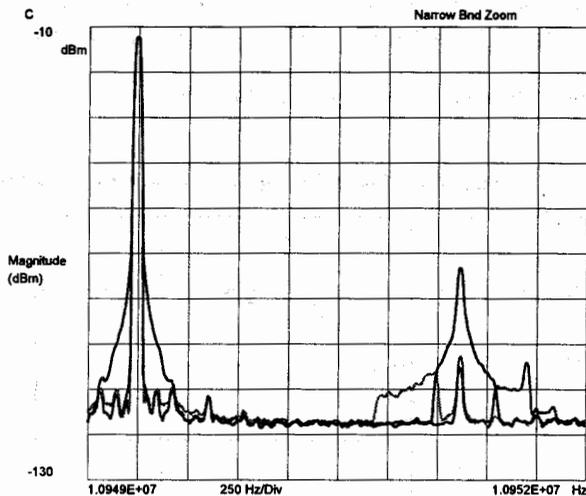
intermodulation of the resonances manifesting as a 200 Hz sideband at levels as low as  $0.0001 \text{ g}^2/\text{Hz}$ , Fig 8.



No Vibration,  $0.01 \text{ g}^2/\text{Hz}$  100 to 500 Hz,  $0.01 \text{ g}^2/\text{Hz}$  100 to 3000 Hz.  
Figure 4: 10.9493 MHz, Device #1



Vibration 1000 to 2000 Hz,  $0.3 \text{ g}^2/\text{Hz}$ ,  $0.03 \text{ g}^2/\text{Hz}$ ,  $0.003 \text{ g}^2/\text{Hz}$   
Figure 6: 10.9493 MHz, Device #2.



$0.01 \text{ g}^2/\text{Hz}$ : 1200 to 1500 Hz, 1500 to 1800 Hz, 1800 to 2100 Hz.  
Figure 5: 10.9493 MHz, Device #1

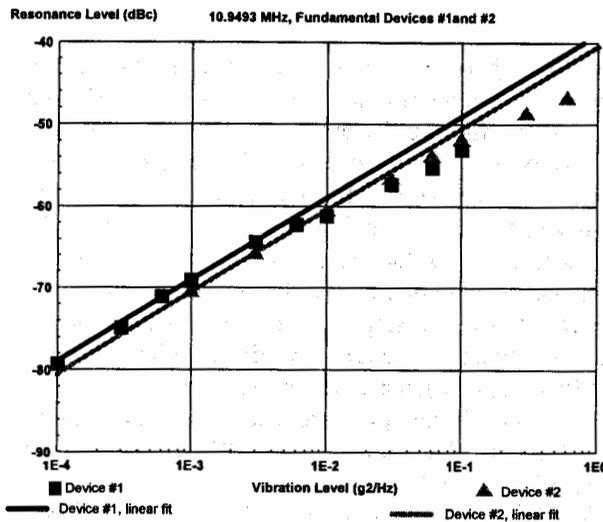
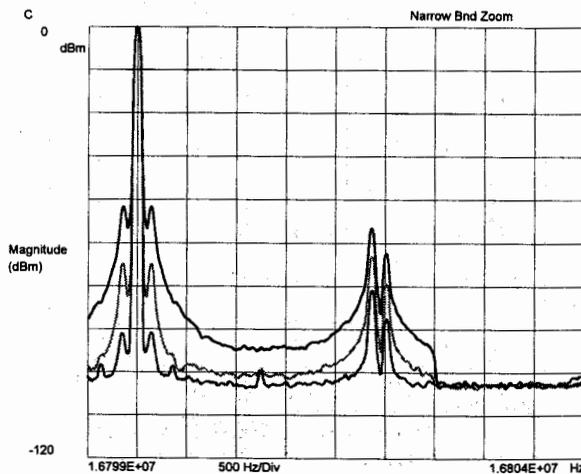


Figure 7: Resonance level vs. Vibration level

Again only when the resonance is excited is the degradation close to the carrier apparent. A second 10.9493 MHz device of identical construction was also measured. This however required nearly an order of magnitude increase in the vibration level to produce the same effect, Fig 6. The resonance level vs. vibration level for both devices is plotted in figure 7. Neither device shows the saturation effect of the 46.72 MHz device. The point at which the resonance leaves the linear response line is however higher for these devices and also differs by an order of magnitude between them.

Finally, a fourth device containing a 16.8 MHz fundamental crystal in a HC26 holder was measured. This showed a double resonance at 2350 and 2550 Hz. The degradation was again clearly evident with the



Random vibration, 500 to 3000 Hz:  $0.0001$ ,  $0.001$ ,  $0.01 \text{ g}^2/\text{Hz}$ ,  
Figure 8: 16.8 MHz Fundamental, HC26 Holder.

The presence of the double or split resonance allows the characterisation of the magnitude of the phenomenon as shown below, Fig 9.

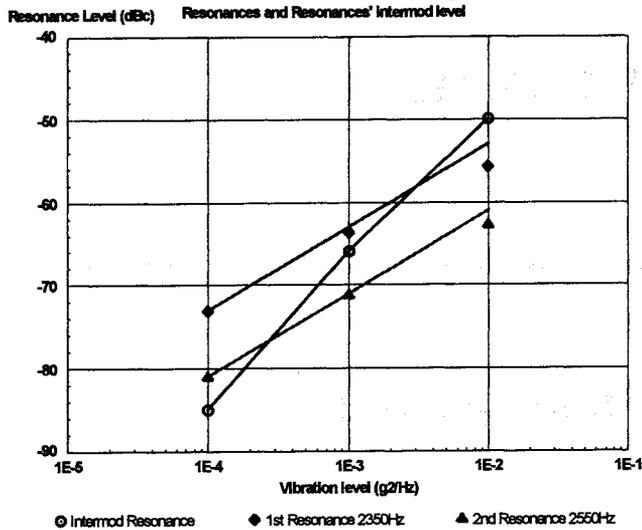


Figure 9: 16.8 MHz Fundamental, HC26 Holder.

The degradation of the close in phase noise follows the square of the resonance level i.e. 20 dB per decade change in vibration level.

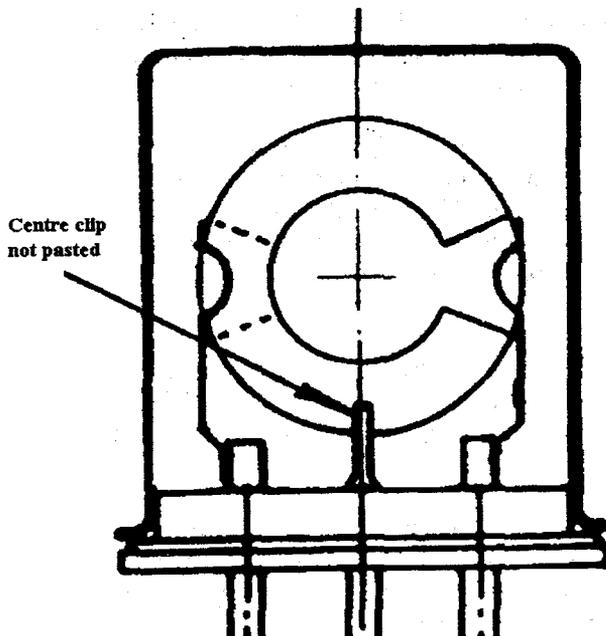


Figure 10: HC45 Holder.

The HC45 holders used had the mounting structure shown in Fig 10, which includes a third clip on the centre lead. This is used to hold the blank during mounting and pasting into the clips, it is not however pasted. This probably explains the saturation effect of the 46.72 MHz device as the centre clip could be restricting the movement of the crystal blank. The HC26 small glass holder has no centre clip to restrict movement or introduce any non-linearity. All the crystal mountings consisted of separate clips welded

onto the leads. A summary of the devices' characteristics is presented in table 1

TABLE 1: The characteristics of the crystals contained in the four oscillators:-

Device Freq. (MHz)	Mode and Blank design	Holder	G Sensitivity (ppb/g)	Resonance Frequency, Q and non-linearity level (g <sup>2</sup> /Hz)
46.72	3 <sup>rd</sup> , Flat, 5mm dia.	HC45	0.7	1950 Hz, Q ≈ 285, 3×10 <sup>-4</sup>
10.9493 #1	Fund., 6D plano-convex, 5mm dia.	HC45	2.0	1650 Hz, Q ≈ 255, 5×10 <sup>-3</sup>
10.9493 #2	Fund., 6D plano-convex, 5mm dia.	HC45	2.5	1650 Hz, Q ≈ 164, 5×10 <sup>-2</sup>
16.8	Fund., Flat, 7.5mm dia	HC26	3.6	2350 Hz, Q ≈ 190, 1×10 <sup>-3</sup> 2550 Hz, Q ≈ 80, 5×10 <sup>-3</sup>

SUMMARY

The degradation of the close to carrier phase noise when a crystal mount resonance is excited by a random vibration profile has been demonstrated. The degradation only starts when the resonance becomes non-linear. The point at which the resonance becomes non-linear varies by at least two orders of magnitude starting as low as 3.10<sup>-4</sup> g<sup>2</sup>/Hz.. When the resonance becomes non-linear, it causes the resonance vibration to be mixed down. This intermodulation and the strength of the intermodulation are both proportional to the resonance level.

REFERENCES

1. Filler R L, 1987, "The Acceleration Sensitivity Quartz Crystal Oscillators: a Review", 41<sup>st</sup> ASEC, 398-408.

# HIGH PERFORMANCE, CRYSTAL OSCILLATORS

## FOR SPACE APPLICATIONS

Authors: S E Morris, F S McClemon, A J Byrne, A L Haston, M G McGovern

GEC-Marconi Materials Technology Limited  
 Hirst Division  
 Elstree Way  
 Borehamwood  
 Herts, WD6 1RX

### Abstract

A family of space qualified, high performance, oven compensated, crystal oscillators have been developed employing stress compensated (SC-cut) resonators. The oscillators use rugged oven designs, robust enough to withstand launch vibration while reducing thermal losses to ensure low power operation. Electronic circuits are implemented using thick film hybrid technology, with interconnections via film wire flexible circuits, to achieve minimum size, weight and high reliability. Typical frequency stabilities of  $\pm$  few  $pp10^8$  over  $-40^\circ\text{C}$  to  $+75^\circ\text{C}$ , coupled with ageing rates of less than  $\pm 1pp10^{10}/\text{day}$  have been achieved with power consumptions of less than 450mW. Phase noise at 1kHz offset is  $-153\text{dBc}/\text{Hz}$ . Environmental testing includes vibration, thermal vacuum and radiation.

### 1. INTRODUCTION/GENERAL

Since 1990, the Hirst Division of GMMT (formerly the GEC-Marconi Hirst Research Centre), has been engaged in the development and manufacture of miniature, high stability low power crystal oscillators for space applications. This activity evolved from an existing capability for oscillators with similar electrical specifications,

which were required for military applications. The space-grade oscillators are conceptually the same as their military precursors, ie they use:-

- (i) Precision high overtone SC-cut resonators manufactured from Hirst-grown High Purity Quartz, which is relatively insensitive to the effects of ionising radiation.
- (ii) Miniature dewars for thermal insulation of the ovened components, which include the crystal, oscillator and oven control circuits.

However, the specific demands of the space applications we are addressing have necessitated some fundamental re-engineering of their military precursors as follows:-

- (i) The electronics is now implemented using hybrid circuit techniques, ie ceramic substrates, printed resistors, chip capacitors and bare die, encapsulated in seam-sealed metal flatpacks; discrete hybrids are interconnected by mounting them on alumina carriers and/or via film wires.

- (ii) The dewars used are of a composite glass/metal construction, with the glass inner tube suspended at both ends of a Kovar outer tube; this double-ended configuration was selected in preference to the thermally more efficient single-ended version because of its improved robustness; the ruggedness of the dewar is a key element in the overall ability of these oscillators to withstand the very high (rocket launch) levels of vibration (typically 30-40  $g_{rms}$ ).
- (iii) In addition to the essential careful selection of radiation-hard electronic components, the space-qualified versions incorporate additional radiation hardening measures including pre-irradiated crystals (where necessary) and radiation tolerant circuitry, eg AGC in the oscillator amplifier; the radiation environment anticipated for these devices is 10-13 years in geostationary orbit, simulated by 100krad Cobalt 60 gamma irradiation.

The space oscillator designs have been extensively analysed and modelled throughout their development, in particular finite element thermal modelling has been used to ensure that all components remain within their temperature and derating limits, and mechanical stress modelling has been used to predict resonant frequencies of sub-assemblies, as well as to ensure that the stresses in critical components are well within safety limits, eg during launch vibration.

We would like to present three oscillator designs in more detail below. Firstly, the Miniature Ovenised Crystal-Controlled Oscillator (MOXCO) is described. This is a miniature (85 x 25 x 20mm), high stability, low power (400mW @ 25°C) unit, which is now going into service as the master oscillator within frequency generator equipment aboard the KOREASAT and INMARSAT III satellites. Secondly, a 10MHz Ultra-Stable Oscillator (USO) is described - a miniature (100x40x30mm), ultra-stable, low power (450mW @ 25°C) unit with very low phase noise (-153dBc @ 1kHz offset). This oscillator is destined to serve as the master oscillator in the ARTEMIS frequency generator equipment.

Finally, a novel sub-miniature (25x25x25mm) oven controlled crystal oscillator, ie the High Accuracy

Reference Oscillator (HARO) is described. This is an ultra-low power (40mW @ 20°), high stability device which employs a direct heated resonator. Although originally perceived as a device for military applications, the HARO is eminently suitable for the space environment and could be considered for lower cost, higher volume, eg LEO communications satellite systems.

2. **MINIATURE OVENISED CRYSTAL-CONTROLLED OSCILLATOR (MOXCO)**

2.1 General

The MOXCO is a miniature oven-compensated oscillator with electrical and mechanical characteristics as detailed below:-

**PERFORMANCE AND PHYSICAL PARAMETERS**

Frequency: 10MHz standard or customer-specified in the range 8-20MHz  
 Adjustment Range:  $\pm 5 \times 10^{-7}$  by means of external select on test resistor or suitable telecommand system

Frequency Stability:	Condition	Maximum Frequency Deviation
	All causes, 13 years	$\pm 1 \times 10^{-6}$
	-25°C to 75°C	$\pm 2 \times 10^{-8}$
	Ultimate ageing	$< \pm 1 \times 10^{-10}$ per day

Phase Noise:	Offset	Maximum Frequency Content
	5Hz	-110dBc/Hz
	10Hz	-125dBc/Hz
	100Hz	-140dBc/Hz
	1kHz	-143dBc/Hz
	10kHz	-145dBc/Hz
	1MHz	-145dBc/Hz

Size: 85 x 25 x 20mm  
 Mass: less than 80g

With reference to the schematic diagram below, the main elements of the MOXCO design are discussed in more detail below:-

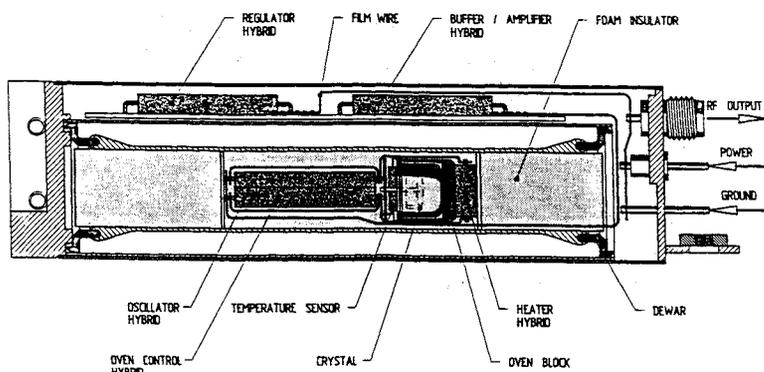


Figure 1. Schematic Diagram of MOXCO.

## 2.2 Crystal

The resonator is a 10MHz third-overtone SC-cut plano-convex element which is designed to be ovened at the lower turnover of its frequency temperature characteristic, ie 85°C. This strategy, coupled with the use of a miniature dewar as the oven wall, allows the crystal temperature to be controlled to within a fraction of a centigrade degree, over a wide range of ambient temperatures, and results in a low frequency temperature coefficient for the integrated oscillator.

The use of a third-overtone resonator provides a relatively thick quartz element which is favourable in respect of low ageing and close-in phase noise. The crystal is manufactured using Hirst-grown High Purity Quartz (HPQ), which is amongst the most radiation-resistant material available worldwide; this results in frequency/radiation sensitivity of the order of a few parts in  $10^8$  for a total dose of 100krad (Cobalt 60 gamma).

MOXCO crystals are subjected to a careful screening programme before selection for use in oscillators. This includes parametric, frequency/temperature and ageing tests.

## 2.3 Dewar

The dewar is a miniature composite glass/metal construction, comprising a Kovar outer tube and a glass inner tube, which forms the immediate oven wall. The outer tube is hermetically sealed to the glass inner tube at each end of the dewar. The resulting enclosure is evacuated via a copper pump-out tube to ultra-high vacuum, and subjected to a lengthy high temperature bake-out to ensure thorough de-gassing. The dewar is then sealed by crimping (cold welding) the pump-out tube.

The use of a dewar provides a highly compact and thermally efficient means of insulating the ovened segment of the oscillator and is the key to obtaining low power consumption within the small dimensions of the overall unit. An additional feature of the MOXCO dewar is its mechanical robustness - it can withstand very high levels of mechanical vibration, eg 40<sub>grms</sub>. The use of a double-ended design sets the first resonance outside the bandwidth of most input vibration spectra and increases the strength relative to a single ended design by sharing any mechanical load between the two ends. As a result of the low mass of the ovened segment, and the absence of any significant Q, the stresses in eg the glass during launch vibration remain well below the UTS of that material. MOXCO dewars are subjected to a demanding mechanical 'proof' test to validate their mechanical integrity before use.

## 2.4 Electronics

The MOXCO electronics is split between two basic segments. The ovened segment includes the crystal, heater transistor, oven temperature sensor (a silicon-based current source), and those circuits which exhibit significant temperature coefficients and which if not maintained at constant temperature would be deleterious to performance, ie the oscillator/amplifier hybrid, the oven control hybrid, and a voltage reference. The non-ovened segment includes a voltage regulator hybrid for the precision 5V rail and an output signal amplifier/buffer.

The MOXCO circuits are implemented using hybrid circuit technology which employs thick film technology on alumina substrates. Most resistors are implemented using printed components; capacitors are ceramic chip components and the active components are in the form of bare chips

bonded directly to the substrate. The hybrids are hermetically encapsulated using conventional seam sealing techniques. All hybrids are subjected to a searching screening and burn-in programme before integration into oscillators.

2.5. Packaging

The MOXCO is packaged in a metal enclosure which can be hermetically sealed if required.

2.6 Qualification

The MOXCO has been subjected to a comprehensive qualification programme including high level vibration, thermal vacuum and accelerated life tests. The MOXCO hybrids have been independently qualified.

Significant on-line screening of MOXCO Flight units is also carried out, including high level random vibration, thermal cycling, barometric testing and extended ageing.

2.7 Application

The MOXCO is scheduled for service in the KOREASAT and INMARSAT III frequency generator equipments.

3. ULTRA-STABLE OSCILLATOR

3.1 General

The Hirst USO is conceptually very similar to the MOXCO, although it is aimed at generally much more stringent requirements, particularly with respect to the all-causes frequency stability requirements ( $\pm 1pp10^{-7}$ , an order of magnitude more demanding than the MOXCO requirement), and very low phase noise.

PERFORMANCE AND PHYSICAL PARAMETERS

Frequency: 10MHz standard or customer-specified in the range 8-20MHz  
 Adjustment Range:  $\pm 1 \times 10^{-7}$  by means of external select on test resistor or suitable telecommand system

Condition	Maximum Frequency Deviation
All causes, 7 years	$\pm 1 \times 10^{-7}$
-25°C to 75°C	$\pm 1 \times 10^{-8}$
Ultimate ageing	less than $\pm 3 \times 10^{-11}$ per day

Phase Noise:

Offset	Maximum Frequency Content
5Hz	-121dBc/Hz
10Hz	-132dBc/Hz
100Hz	-145dBc/Hz
1kHz	-153dBc/Hz
10kHz	-155dBc/Hz
1MHz	-155dBc/Hz

Power Requirements:

Condition	Maximum Power
20°C	450mW
-25°C	670mW
Cold turn-on	2000mW

Size: 100 x 40 x 30mm  
 Mass: less than 180g

A schematic diagram of the USO is provided in Figure 2 below.

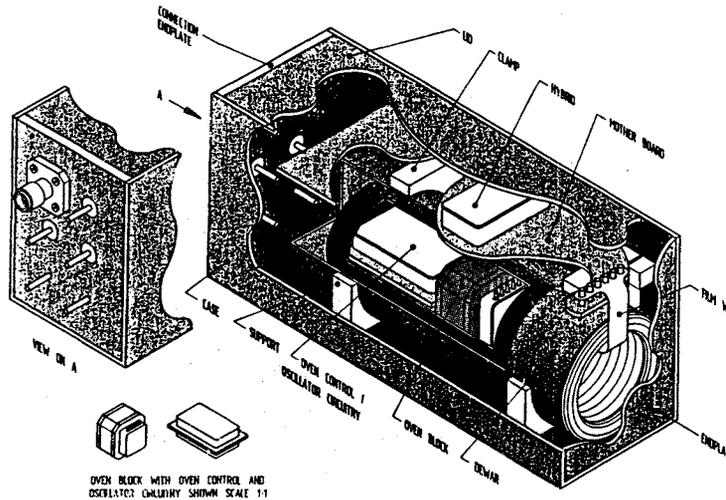


Figure 2. Schematic Diagram of Ultra-Stable Oscillator.

The main features of the USO design, with particular emphasis on the variations from the MOXCO configuration, are discussed below.

3.2 Crystal

The resonator in the USO is a 10MHz fifth overtone, SC-cut plano-convex element, with a

considerably increased diameter in relation to that of the MOXCO resonator. The high overtone results in a relatively massive quartz element, which is advantageous in respect of low ageing and close-in phase noise performance. The crystal is designed to be ovened at its lower turnover (85°C) such that, in conjunction with the temperature stability provided by the dewar, the overall frequency/temperature coefficient of the oscillator is very low.

The ultimate ageing rate for this design of resonator is in the low pp10<sup>11</sup>/day regime. The achievement of these rates relies upon manufacture of resonators under ultra-clean conditions through the encapsulation stage.

The radiation sensitivity of the as-manufactured USO resonator is low, ie in the few pp10<sup>8</sup>/100krad regime; however this is still excessive in relation to some applications and it has been found necessary to pre-irradiate completed resonators, thereby saturating the sensitivity and producing a residual frequency/radiation coefficient in the pp10<sup>9</sup>/100krad regime; moreover, the resonator will ultimately withstand high radiation doses, eg 1Mrad.

### 3.3 Dewar

The philosophy behind the USO dewar is identical to that described for the MOXCO above, ie it is a double-ended structure with high thermal efficiency resulting in low power operation of the oscillator. The dewar is of considerably wider bore than the MOXCO component, as a result of having to accommodate the larger diameter crystal. The dewar's mechanical resonance is outside the typical input vibration bandwidth and the structure is capable of withstanding high levels of (eg rocket launch) vibration.

### 3.4 Electronics

The USO electronics are distributed in a similar manner to that pertaining to the MOXCO but there are several key aspects of the USO configuration as follows:

- i) the oscillator/amplifier is based on a 'differential-mode' configuration for improved phase noise.
- ii) the oven temperature sensor is in the

form of a platinum resistance thermometer which is considered optimum in respect of long-term temperature stability of the oven.

- iii) the use of a regulated 10V rail in the output buffer/amplifier allows improved far-out phase noise to be obtained.

### 3.5 Packaging

The USO is configured in a compact aluminium case, with the overall mechanical structure being suitably compensated for thermal mismatches (compliant dewar mounting) and appropriately stiffened for elevation of mechanical resonance frequencies (10-point mounting footprint). The package is vented.

### 3.6 Qualification

The USO resonator has been qualified in its own right, essentially to ESA SCC/3501 standards. The USO hybrid build has been delta qualified (using the MOXCO hybrid qualification as the foundation). The oscillator itself is being qualified using EQM models and within the frequency generator equipment in which it is used.

### 3.7 Application

The USO is scheduled for service as the master oscillator within the frequency generator equipment for the ESA/ARTEMIS mission.

## 4. HIGH ACCURACY REFERENCE OSCILLATOR (HARO)

### 4.1 Description

The High Accuracy Reference Oscillator (HARO) is a novel oven-compensated unit (OCXO) employing a direct-heated quartz element. The crystal is a high overtone SC-cut unit. This configuration allows frequency stability and phase noise characteristics to be obtained which are comparable to those of bulky ultra-stable OXCOs, while size and power consumption are constrained to levels which are characteristic of non-ovened oscillators such as simple packaged crystal oscillators and temperature compensated crystal oscillators (TCXO's).

The HARO operates in a very low power mode

with modest phase noise performance, which is ideal for timekeeping applications. An alternative low phase noise mode, intended for frequency synthesis/LO requirements, can be activated by a suitable TTL command.

The device is designed for operation in a hostile military environment (-55°C to +85°C) and will withstand high level random vibration and 'man-in-the-field' radiation levels.

A summary of the anticipated performance of the HARO is given below.

#### HARO TARGET PERFORMANCE AND PHYSICAL PARAMETERS

Nominal Frequency: 10MHz

Frequency Stability:

Condition	Stability
-55 to +85°C*	± few pp 10 <sup>9</sup>
Ageing	< ±1pp 10 <sup>10</sup> /day
Voltage sensitivity	~ few pp10 <sup>9</sup> /volt
Short-term (Allan Variance)	< 1pp10 <sup>10</sup> (0.1 to 100 secs)

\*High temperature variant available, ie -55 to +115°C.

Phase Noise:

Offset	Standard Mode	Low Power Mode
1Hz	-95dBc	-85dBc
10Hz	-115dBc	-100dBc
100Hz	-135dBc	-120dBc
1kHz	-150dBc	-130dBc
10kHz	-160dBc	-130dBc

Power Consumption:

Condition	Standard Mode	Low power mode
-55°C	100mW	60mW
+20°C	80mW	40mW
+85°C	70mW	30mW
Cold turn-on	0.5 to 1W depending on warm-up time required	

Warm-up: (using 500mW)

To within ±1pp10 <sup>7</sup> in 30 seconds To within ± few pp10 <sup>9</sup> in 1 minute
--

Size: 25 x 25 x 30mm  
Mass: 25 grams  
Vibration: Survives 30g rms

The HARO is eminently suitable for space applications where small size/mass, and low power consumption are highly desirable. The radiation hardness at 'space' levels is as yet unproved but a variant is now being developed which we anticipate will exhibit sensitivity in the pp10<sup>8</sup> to pp10<sup>9</sup>/100krad regime.

#### 5. CONCLUSIONS

A world-competitive space oscillator capability has been developed at the Hirst Division of GMMT. We have direct access to a strategically advantageous material, ie High Purity Quartz, as well as the design and modelling software which is essential for the implementation of state-of-the-art space-qualifiable devices. This capability has already been demonstrated by means of three novel oscillator configurations exhibiting high stability, compactness and low power consumption. Two of these configurations are space-qualified and are about to enter service in demanding space applications, and the third is suitable for future satellite systems.

#### ACKNOWLEDGEMENTS

1. The MOXCO was developed in close collaboration with Matra Marconi Space (UK) Limited.
2. The initial development of the USO was funded by ESA; the final development for the ARTEMIS application was carried out in close collaboration with Matra Marconi Space (UK) Limited.
3. The development of the HARO was funded by DRA (EL), Malvern.

## DEVELOPMENT STATUS OF THE MICROCOMPUTER COMPENSATED CRYSTAL OSCILLATOR

Brian E. Rose, Howard Phillips

Q-Tech Corporation, U.S.A.

### INTRODUCTION

The design of the MCXO eliminates the three primary obstacles to stability of the conventional Temperature Compensated Crystal oscillator (TCXO). These are: 1. errors in measuring the crystal temperature. 2. use of a fundamental mode AT instead of an overtone SC cut crystal. and, 3. hysteresis and aging associated with "pulling" the crystal.

As a result, the MCXO achieves about 30 times better stability than the best TCXOs. In this paper we will review the MCXO design technology, describe the Q-Tech MCXO, QT2001, and describe some future developments.

### REVIEW OF MCXO TECHNOLOGY

#### Temperature Measurement

Conventional methods of crystal temperature measurement suffer from the problem that the temperature sensor cannot be in intimate contact with the crystal. Hence, there exist different thermal environments for the crystal and the sensor, including different thermal paths. Static and dynamic errors result. The dual mode temperature sensing method described by Schadowski (1) uses the difference between the fundamental and third overtone modes of a crystal, running simultaneously, to measure the exact crystal temperature. In the implementation of this method in the Q-Tech MCXO, the third overtone at 9.999 MHz is divided by 3 and subtracted from the fundamental signal in a mixer. The filtered difference frequency at about 50 KHz has an almost linear frequency-temperature curve with a slope of 77 PPM per degree C. i.e. an almost perfect thermometer for the crystal temperature. This frequency is counted and the digital number, N1, that results is correlated with the overtone and fundamental frequencies.

### The Crystal

The frequency correction in an MCXO is done by adding the error frequency to the crystal frequency, instead of the TCXO method of offsetting ("pulling") the oscillator frequency to counter temperature caused offsets. Another way to describe this difference is to say that the MCXO crystal is allowed to oscillate without any need for frequency correction or pulling. Consequently, a third overtone SC-cut crystal can be used in the MCXO. This type of crystal is very difficult to pull, and could not be used in a conventional TCXO. The third overtone crystal has superior aging characteristics to a fundamental mode AT-cut crystal, such as is typically used in a TCXO. In addition, the absence of a varactor diode pulling network eliminates a circuit source of aging, hysteresis, and noise.

### DESCRIPTION OF THE Q-TECH MCXO

#### Block Diagram

A simplified block diagram of the MCXO is shown in Fig. 1. The oscillator and phase-locked loop circuits use discrete surface-mount analog construction, while all other circuitry is digital. Most of the circuits are in an Application Specific Integrated Circuit. (ASIC) operating in conjunction with the microcomputer and its associated EEPROM.

The SC-cut crystal is a specially designed 10 MHz 3rd overtone unit, that is contoured and sized to provide a high Q on both fundamental and third overtones. The lower turning point of the third overtone mode is usually set between 10 deg. and 20 deg. Celsius to provide equal maximum slopes for the fundamental and overtone modes. Although the F vs T characteristic of the SC-cut is cubic, it appears to be parabolic because the inflection temperature is just above the upper limit of the operating temperature range. The Q-Tech MCXO crystal is plated so that the maximum frequency of the third overtone (Fo) is typically 1 KHz below 10 MHz. This produces a fundamental frequency (Ff) of about 3.384 MHz. The 3rd overtone frequency varies more than 700 Hz over the operating

temperature range but is compensated by the MCXO system to within less than 0.2 Hz, providing an improvement factor of 1 part in 3500.

Since both modes have essentially identical temperature characteristics, differing primarily in the magnitude of their first-order terms, the beat frequency between the overtone mode (divided by three) and the fundamental mode is a nearly linear function of temperature. In this way, temperature information is obtained directly from the resonator itself rather than from an external sensor, thus eliminating temperature offset and lag effects. Temperature velocity effects are also minimized by the Stress Compensation characteristics of the SC-cut resonator. A new oscillator circuit has been developed for the high resistance, third overtone C-mode of the SC-cut resonator. It has been called a True Bridged-Tee Oscillator (BTO) which produces a very high operating Q at low collector current through the use of a unique form of ALC. This design provides the very high stability and low phase noise required of the third overtone output.

Simultaneous oscillation on the fundamental frequency of the C-mode is sustained by an Impedance-Inverting Colpitts Oscillator circuit that is well suited to the low resistance of this mode. Both oscillators operate at the series resonant frequencies of the resonator in order to minimize crosstalk between them.

The beat frequency,  $F_b$ , which is generated in a digital mixer, is converted to digital number  $N_1$  in a period counter. During the initial calibration run (which each unit undergoes), five seventh-order polynomials are fitted to both sets of frequency data versus  $N_1$  and the coefficients stored in the EEPROM. The microcomputer uses this polynomial in conjunction with the  $N_1$  number to calculate and generate the required frequency correction ( $dF_o$ ) via the Direct Digital Synthesizer (DDS). This DDS uses only the MSB bit from its adder as the  $dF$  output. Neither sine look-up tables or the D/A converters usually associated with DDS's are used here.

The DDS resolution is better than 0.007 ppb with respect to the 10 MHz output, much finer than the repeatability of the system over temperature. A Phase-locked Loop (PLL) is employed to lock a VCXO to the sum of  $dF_o$  and  $F_o$  and to generate the MCXO's temperature-corrected output. This summing technique, which avoids the introduction of digital divider noise, forces the VCXO to assume the full stability of the summed signals within the 10 Hz bandwidth of the PLL's low-pass filter, while ensuring a smooth transition between DDS steps. Above 10 Hz, however, the spectrum of the VCXO gradually becomes independent of the summed

signals, and dependent only on the design of the VCXO. Through careful design and control of the power input to the VCXO, a low phase-noise spectrum can be obtained.

The phase noise measured at 1KHz. from the 10 Mhz. carrier using the square-wave output of the CMOS VCXO buffer is a maximum of -140 dBc. A sine wave output alternative results in lower noise.

While some spurs are visible in the spectrum, they are a minimum of 70 dB below the carrier and are offset from the carrier by at least 1 kHz. They are produced by the difference frequency between  $F_o$  and  $F_{out}$ .

### Mode Description and Selection

The MCXO has two operational modes: Frequency Mode and Timing Mode. The two modes have the following characteristics:

**Frequency Mode:** The frequency mode is the full function mode. The outputs are 10 MHz, 1 PPS (pulse-per-second), and 4 KHz. The MCXO makes temperature corrections every 285 milliseconds. The 4 KHz and 1 PPS outputs are derived from the 10 MHz by dividers. Both dividers are reset during 1 PPS synchronization. The maximum input power is 75 milliwatts.

**Timing Mode:** In the Timing mode the maximum DC input power is 25 milliwatts, 1/3 that of the Frequency mode. The power saving is accomplished by two means:

1. In the Timing mode, the 10 MHz output is shut off. The overtone oscillator is turned on only for temperature measurement. The outputs available in the Timing mode are 4 KHz and 1 PPS.
2. In the Timing mode, the microprocessor is allowed to power down, or "sleep", when the ambient temperature is constant or slowly changing.

In the Timing mode, the 4 KHz is generated by the DDS (direct digital synthesizer), and the 1 PPS by a divide-by 4000 from the DDS. Both DDS and divide-by-4000 are reset during 1 PPS synchronization.

Three standard outputs are: 10 MHz, 4 KHz, and 1 PPS. All three are at CMOS levels, driven by internal HCMOS gates.

All three outputs are available in the Frequency mode; only the 4 KHz and 1 PPS outputs are available in the Timing mode.

**Time Domain** The DDS changes frequency to compensate for temperature changes. The DDS frequency steps are phase continuous. Furthermore, the output PLL (phase-locked loop) acts as a low-pass filter for the correction steps. For these reasons, the output exhibits no phase or frequency jumps.

**Timing Mode Outputs** The timing mode outputs, 4 KHz and 1 PPS, are generated directly by the DDS. The DDS size is 32 bits, and the DDS clock is about 200 KHz. The frequency resolution is therefore 46 micro-Hertz, or 1.2 parts in  $10^8$  at 4 KHz. The maximum jitter on the output is equal to the period of the DDS clock, or 5 microseconds. Since the jitter averages out over time, it has no effect on normal timing applications. It does require long counts (>10 minutes) to accurately measure frequency.

**Status output** A status signal is used to indicate various MCXO conditions.

**Aging Correction** Aging refers to the slow frequency drift of a crystal over time. In the MCXO, the aging effect can become a significant portion of the total frequency accuracy after several months. This frequency shift can be removed in the field by the application of an external reference frequency to pin 10. This process is called aging correction.

When the external reference signal is applied to pin 10, the MCXO senses its presence, and begins an automatic aging correction which takes 102 to 104 seconds. The MCXO compares the external reference frequency to the MCXO's internal frequency, measures the error, and adds or subtracts to the constant term of the frequency-temperature correction curve, stored in an EEPROM.

If for any reason the measurement process produces an out-of-limit result, the correction cycle is abandoned, no correction is made, and the status pin is set.

### Synchronization

The MCXO 1 PPS output can be synchronized to an external 1 PPS signal. This is accomplished by applying the 1 PPS signal to the reference input, pin 10. This is the same pin that is used for the 10 MHz reference input for aging correction. The MCXO automatically detects the presence of a signal on that pin, and determines if it is 10 MHz or 1 PPS. If

it is 1 PPS, the MCXO resets all of the 1 PPS internal divider chains (and the DDS in Timing mode). This serves to align the output 1 PPS to the input 1 PPS. Synchronization will occur continuously as long as the external 1 PPS signal is present.

### Specifications Summary

#### Outputs:

Frequency mode: 10 MHz, 4 KHz, 1 PPS.

CMOS

Timing mode : 4 KHz, 1 PPS

Both modes : Status output

Frequency Stability -- Temperature  
+/-  $3 \times 10^{-8}$  from -55°C. to +85°C.

Aging - standard

$5 \times 10^{-10}$  per day, initial aging

Phase noise:

-145 dBc @ 1 KHz offset, 10 MHz CMOS output

Power consumption, @ 5 Volts DC.

Frequency mode: 75 milliwatts, average,  
100 milliwatts peak.

Timing mode: 25 milliwatts, average,  
50 milliwatts peak.

Start-up time:

10 seconds.

Inputs:

Aging correction reference signal - at 10 MHz (or customer specified frequency)

1 PPS synchronization - at 1 PPS

Aging correction:

Automatic, upon sensing external reference.

Time required - 102 to 104 seconds. An error aborts the correction and sets the status flag.

### Connections and Interface

MCXO connections are made by 10 feed-through pins. The pin connections are :

#### Pin Connections

1	+5V DC power input
2	DC power return. Circuit ground.
3	10 MHz CMOS square wave output
4	1 PPS CMOS square wave output
5	4 KHz CMOS square wave output
6	Reserved (serial port)

7	Reserved (serial port)
8	Mode select input
9	Status output
10	Reference input, aging correction (10 MHz) or 1 PPS sync.

## TESTING AND PERFORMANCE

As the MCXO project advanced from the development stage to pre-production, a number of test techniques were developed. Since the 3rd overtone 9.999 MHz SC-cut crystal is the key to performance, methods of measuring the crystal's hysteresis and residuals, before installation into the MCXO were needed.

### MCXO Hysteresis and Residual Errors

The Q-Tech MCXO system described in this paper now provides sufficient precision to reveal the limitations inherent in the resonator itself. Although crystal oscillators have been operated over wide temperature ranges for many decades, it was not possible to examine their performance with sufficient precision until the advent of the dual-mode oscillator, which now makes it possible to measure SC-cut crystal temperatures with high accuracy. The level of testing undertaken during the design and calibration of the MCXO has revealed the true hysteresis of the resonators that were tested, as well as the magnitude of the residuals that remain after (5) seventh-order least-squares fit to  $f_0$  vs.  $N_1$  data. These characteristics can be observed in Figure 2, where the open-loop of the "round trip" calibration indicates the degree of hysteresis, while the erratic deviations from the straight line are the residuals caused by the limitations of curve fitting.

The magnitude of both hysteresis and residuals is influenced by the crystal fabrication process, since each process shows a distinct "signature" on the crystal groups that have been tested.

## Crystal Testing

Several hundred crystals have been tested as part of the MCXO program. The individual crystals are installed in a small PC board which has a dual-mode oscillator circuit on it. 32 of these test boards are mounted on a frame, and installed in a computer-controlled oven. The outputs  $F_0$  and  $F_f$ , from each board, are multiplexed into two HP53131 frequency counters, also computer-controlled. Sets of data are taken and the curve fits are generated by a special program. The crystals are conditioned by repeated cycling until they are ready to be installed in MCXOs.

### MCXO Testing

The completed MCXOs are tested in a setup almost identical to that used for crystal test. 80 MCXOs are now in a long term aging test. A result of burn-in testing for 5 weeks is illustrated in figure 2.

## ADVANCED DEVELOPMENT

### Advanced MCXO for Satellite Terminals

Q-Tech has completed a design and breadboard demonstration of a version of the MCXO suitable for satellite terminal frequency and timing requirements. This contract was done in cooperation with MIT / Lincoln Laboratories. The objective of the contract is to develop the hardware and software for a miniature low power oscillator which will meet the frequency and timing requirements for a satellite terminal. The objective of the second phase of the contract will be to develop a miniature package for the unit.

The starting point for this design is the existing Q-Tech MCXO. The changes and additions to the performance of the QT2001 that were made for the this program are as follows:

1. Outputs at 200 MHz, 100 MHz, and 50 MHz.
2. Reduction in output spectrum spurious frequency levels to -70 dBc @ 200 MHz
3. A time-of-day (TOD) interface to the system computer ("real time clock")
4. An ultra low-power, 5 milliwatt, "Watch" Mode.

The complete specifications are shown in Table 1 - 200 MHz MCXO - Specifications

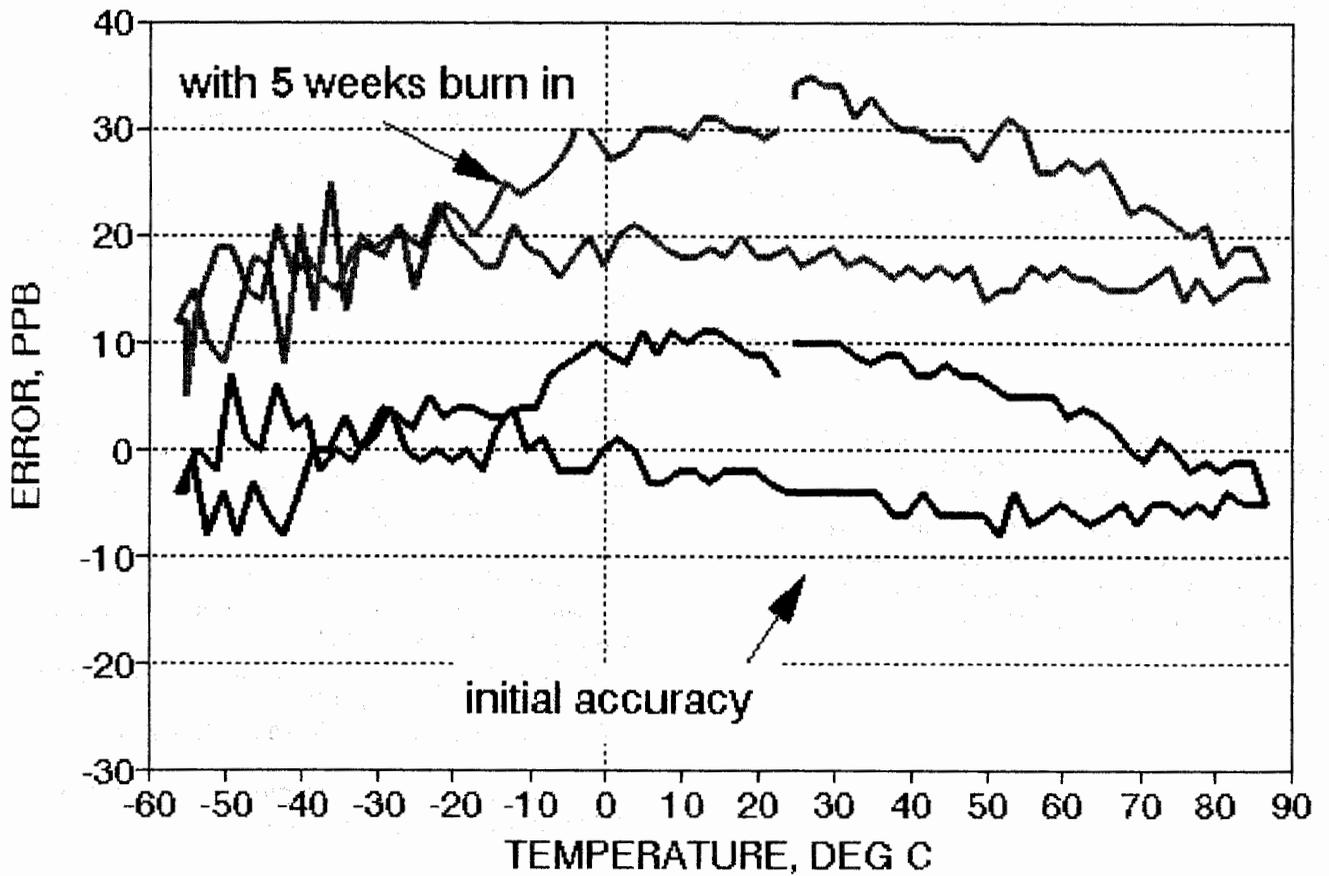
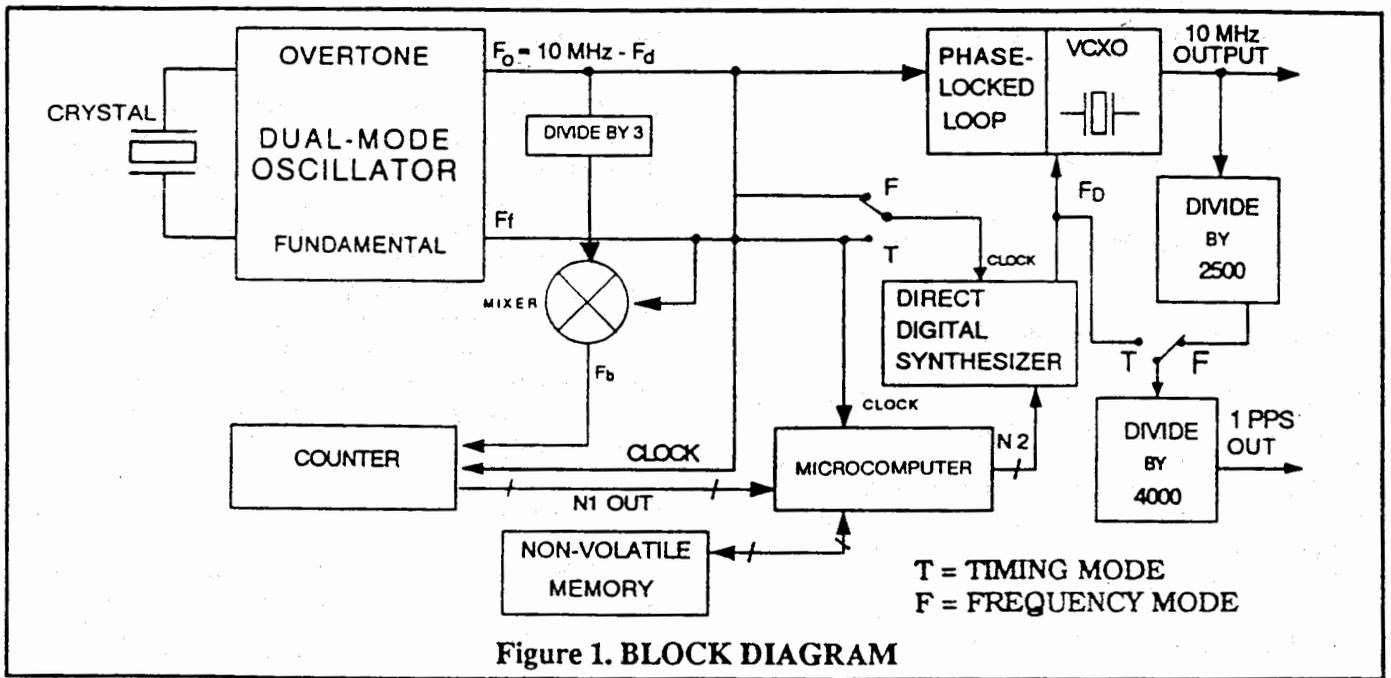


TABLE 1. Specifications for a 200 MHz MCXO

Outputs:	1.) 200 MHz primary, frequency mode only 2.) 100 MHz auxiliary, frequency mode only 3.) 50 MHz auxiliary, frequency mode only 4.) Time-of-day : 1 PPS
DC Power Consumption:	
Watch Mode:	0.5 mW to 2 mW
Frequency Mode:	250 mW
DC Supply Voltage	5 V
Frequency Stability vs. Temperature	Over full temperature range, including hysteresis
Watch Mode:	$< \pm 5 \times 10^{-7}$ to $\pm 2 \times 10^{-6}$ , user selectable
Frequency Mode:	$< \pm 3 \times 10^{-8}$
Frequency Accuracy vs. Time (Aging and Temperature)	$< \pm 5 \times 10^{-8}$ / year (see Note 3) (Aging correction can be done automatically at the depot level)
Phase Noise	@ 200 MHz Output
10 Hz offset:	Unspecified
100 Hz offset:	Unspecified
1 kHz offset:	Unspecified
3 kHz offset:	$< -114$ dBc/Hz
> 50 kHz offset:	$< -155$ dBc/Hz
Spurious Signals: 200 MHz	$< -70$ dBc @ > 1 kHz offset
Temperature Range:	-35°C to + 65°C
Packaging and Size:	< 0.7" height, minimize volume, hybrid or MCM package
Weight:	< 1 ounce

(1) S. Schodowski, 1989, "Resonator Self-Temperature Sensing Using a Dual-Harmonic-Mode Crystal Oscillator," in Proceedings of the 43rd Frequency Control Symposium 2-7

(2) R. Filler and J. Vig, 1989, "Resonators for the Microcomputer Compensated Crystal Oscillator," in Proceedings of the 43rd Frequency Control Symposium 8-15

(3) A. Benjaminson and B. Rose, 1991, "Performance tests on an MCXO Combining ASIC and Hybrid Construction.", in Proceedings of the 45th Frequency Control Symposium 393-397

## A SPECIFIC LSIC INTENDED FOR SYNTHESIZER-TYPE DTCXO

Anatoly V. Kosykh, Boris P. Ionov

Omsk State Engineering University, Russia

### SUMMARY

Thermocompensated crystal oscillators are widely used in cases when high frequency stability, low power consumption and short time of coming into readiness are needed. Digital methods of frequency compensation provide the highest frequency stability parameters. Popular as they are digital thermocompensated oscillators have the stability that usually does not surpass  $\pm 5 \times 10^{-7}$  in a wide temperature interval.

Much higher frequency stability can be achieved in a DTCXO of a synthesizer type (DSTCXO). As a matter of fact DSTCXO is a digital frequency synthesizer with thermocontrolled commutations. Taking into consideration that besides the synthesizer itself there are also a temperature meter and a nonlinear code converter in the DSTCXO circuit, the complexity of this technical solution becomes evident.

Several tens of microcircuit packages would be needed in order to construct a DSTCXO using discrete components. With the purpose of DSTCXO dimensions minimization a specific LSIC has been developed comprising nearly all oscillator elements. The microcircuit was intended for interaction with frequency sensor of temperature. The resolution of synthesizer is 10 ppb. When a double mode quartz oscillator is used temperature stability comes as high as  $\pm 3$  ppb.

### INTRODUCTION

Digital compensation oscillators stand out against a background of thermocompensated quartz oscillators (TCXOs). Higher thermocompensation accuracy and the possibility of completely automatic tuning are their indisputable qualities. Digital thermocompensation quartz oscillator usually comprises an analog thermosensor (thermoresistor or diode), an analog - to - digital converter (ADC), a nonlinear code transformer (ROM or micro-processor) and a digital - to - analog converter (DAC) controlling the oscillator frequency by varicap capacity variation. Sometimes frequency transducer are used in this circuit and then a pulse counter is applied as an ADC. Such converter is easier to realize by means of integration technology. There are two imperfections in the traditional circuit:

- thermocompensation accuracy dependence on the reference voltage stability of the ADC and DAC;
- higher aging of resonator quartz plates (to obtain wide frequency shift resonators are usually excited at the 1-st harmonics).

New quality level in frequency stabilization can be achieved if digital frequency synthesis method are applied. A simplified circuit-diagram of such oscillator is given in Fig. 1.

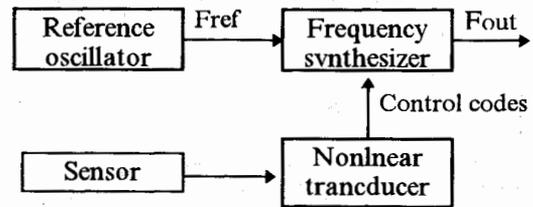


Fig. 1.

In this case there is no frequency change of the reference oscillator, but its output frequency  $F_{ref}$  is varied in temperature interval in according with its own  $f$ - $T$  curve. The output frequency  $F_{out}$  is correlated with the reference frequency according to the following expression

$$F_{out} = K(T) \cdot F_{ref} \quad (1)$$

where  $K(T)$  - synthesizer conversion ratio dependent on temperature  $T$

The  $K(T)$  functional dependence is selected so that  $F_{out}$  value would tend to constant. The nonlinear transducer transforms thermosensor signal into codes controlling the synthesizer. In general case the transducer dependence on temperature is nonlinear.

Several approaches to the synthesizer design are possible. The most suitable for LSI implementation are digital synthesizer based on the PLL principle.

### THERMOCOMPENSATION SYSTEM

The thermocompensation system is intended to process a signal from the thermosensor and to form controlling codes for the synthesizer. As our aim was to develop a microcircuit for high stability oscillator the compensation system should be acting with frequency transducers (thermosensitive quartz resonator or dual mode resonator). In this case a thermosensor signal converter into a code is essentially a frequency meter.

Functional code conversion for high stability oscillators can be performed either by a statistical piecewise interpolation [1] or by microprocessor [2]. (Step-function transformation would need too much space

of the ROM and instrumentation piece-wise transformation would need too high hardware cost). Both ways had been investigated during microcircuit design and implementation.

The processors available in Russia by the project beginning (80C48 and 80C51 Russian analogs) had rather large dimensions and power consumption therefore the version of piece-wise statistical interpolator was considered as basic one. The detailed description of this interpolator was given in [ ] so only aspects significant for microcircuit design will be explained here.

They are follows:

1. A temperature meter (frequency relation meter) should be realized in such a way that a "dead time" would be an integral quantity of measured pulses.
2. Every temperature measurement reading should be processed. Missing of readings can result in a statistic violation and the method failure. This item puts rigorous limitations on the synthesizer circuit. Calculation of pulses and synthesizer control should take place synchro. In particular, the well known synthesizer version [3] in this case turn out to be inapplicable.

3. The counter calculation time  $\tau_c$ , thermosensor transconductance  $\alpha_f$ , temperature range  $\Delta T$  and counter length "n" turn out to be related with the following expressions

$$\tau_c = 2^n / \alpha_f \Delta T$$

A schematic diagram of the frequency ratio meter is given in Fig. 2.

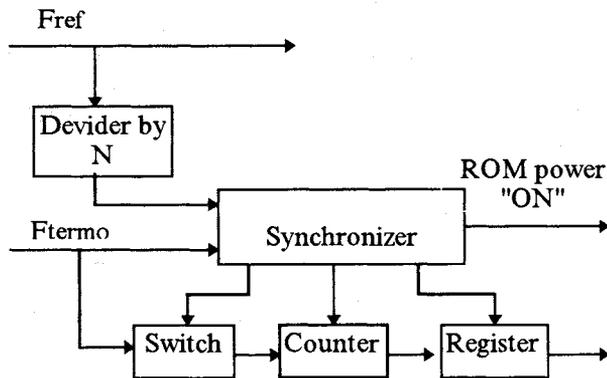


Fig.2

The inclusion of the synchronizer into the schematic diagram is obligatory. This little complication (2-3 triggers) permits to realize both statistical interpolation method and to interact with a microprocessor.

While operating with a processor an accurate temperature value is obtained as a result of digital filtration of the flow of roughly estimated readings. More precise (corrected value comes out after every 256 cycles in this case. Further temperature signal processing must include nonlinear transformation

and, possibly, dynamic correction forming [4]. Really, we have used piecewise-linear interpolation.

## FREQUENCY SYNTHESIZER

The solution presented in [3] was chosen as a starting point for construction of the SDTCXO. Its schematic diagram is given in Fig. 3.

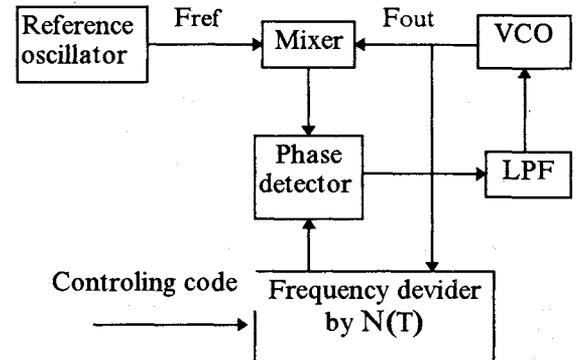


Fig. 3.

The stationary condition equation for this synthesizer is describe as follows:

$$F_{ref} - F_{out} = F_{out}/N(T)$$

where  $N(T)$  - frequency divider variable ratio.

There are several imperfections in such schematic diagram:

- nonlinear (hyperbolic) dependence of output frequency on the controlling code which lowers the accuracy;
- the phase detector performs at a variable frequency that can be very low and difficult for signal filtration;
- poor matching with a piecewise linear statistical interpolator which was chosen by the authors as a base for a nonlinear functional converter.

Instead of an integral ratio frequency divider we used a divider with fractional division ratio [5]. The circuit diagram is given in Fig. 4.

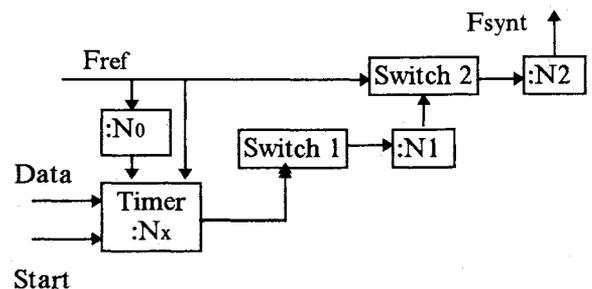


Fig. 4.

The divider principle of operation was based on cutting out definite number of pulses per period from

periodic sequence which results in slight medium frequency lowering. The output frequency  $F_{synt}$  at the output of the frequency divider is defined by the expression:

$$F_{synt} = \frac{F_{ref}}{N_2} \left( 1 - \frac{N_x}{N_0 N_1} \right)$$

Where  $N_x$  - controlling code incoming from the compensation system output.

The absence of skipped pulses results in phase front jitter in the output signal  $F_{synt}$ . The higher is the division ratio of the  $N_2$  divider the less will be the effect of jitter.

While operating with a microprocessor the synthesizer circuit may be altered. Synchronous action on synthesizer at the end of every temperature measurement cycle is not needed now. Low frequency synthesizer circuit diagram is given in Fig. 5.

5.

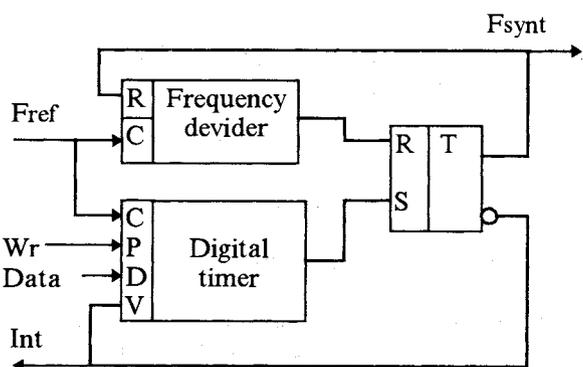


Fig. 5

Digital frequency synthesizer forms a signal of synthesized frequency  $F_{synt}$  out of a pulse sequence of  $F_{ref}$  frequency. This means that every frequency  $F_{synt}$  period includes an integral number of frequency  $F_{ref}$  periods. Multiple-fractional frequency division is performed by means of information accumulation on signal  $F_{synt}$  phase error in the microprocessor. This type of synthesizer was named in Russia "digital phase".

MICROCIRCUIT

At the stage of microcircuit design it was decided to combine in one chip the possibilities of operation both with a ROM and with a microprocessor. The 1-st version was intended for construction of medium temperature stability oscillators (up to 10 ppb), but having high radiation tolerance. The second version was intended for construction of high stability self-tuning systems. The choice of a configuration is performed with an external signal ROM/MP. The possibility of operation with thermosensor of different transconductance values (from 20 ppm/°C to

100 ppm/°C) is attained by variation of the frequency division ratios of the reference and thermosensitive oscillators.

When microprocessor configuration is operational a mode of self-tuning is actuated. A simplified circuit-diagram for this mode operation is given in Fig. 6.

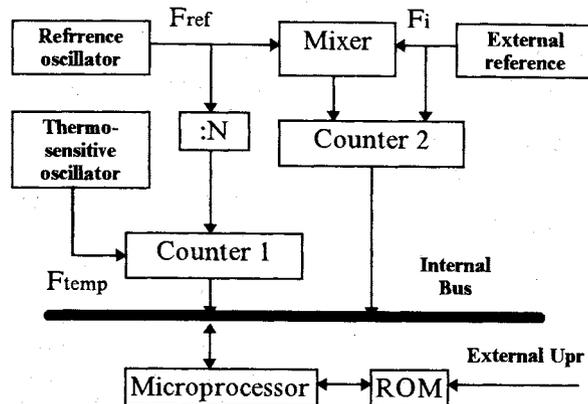


Fig. 6

When communication into self-tuning mode is performed a reference frequency from external source is fed at the oscillator output connector. Two counters perform calculation of the frequency ratios  $F_{ref}/F_{temp}$  and  $F_i/(F_{ref} - F_i)$ , where  $F_i$  input external reference frequency. The microprocessor handles these frequencies ratio and calculates transformation coefficient of transformation coefficient of the frequency synthesizer relevant to a given temperature. Running the oscillator in the temperature interval a table of ROM threading is obtained. ROM programming can be fulfilled even in a sealed oscillator which is very suitable for production. Triggers of the frequency synthesizer are used for the second counter realization. The complete block-diagram of LSIC is given in Fig. 7.

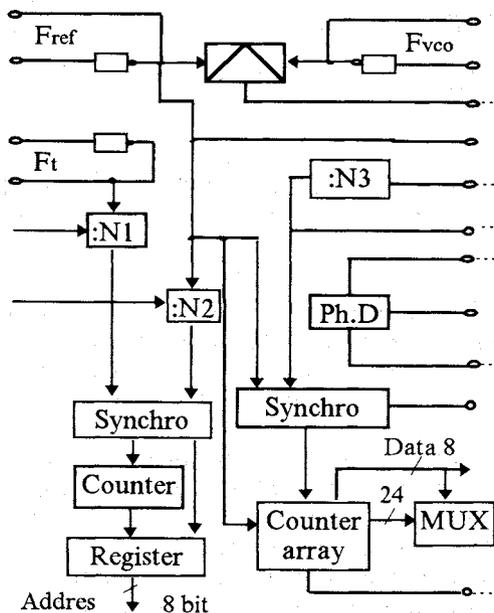


Fig. 7

A 16-bit synthesizer permits to set the frequency with accuracy of not less than 1 ppb. A phase detector has a characteristics with saturation and thus, can provide a wide domain of phase locking. The output of the detector has three states. 60% of the chip surface had to be occupied for the circuit realization. The chip comprises 3000 gates. The number of operational microcircuit outputs is 62.

## EMBODIMENT

An example of the microcircuit utilization in the most simple configuration is shown in Fig. 8.

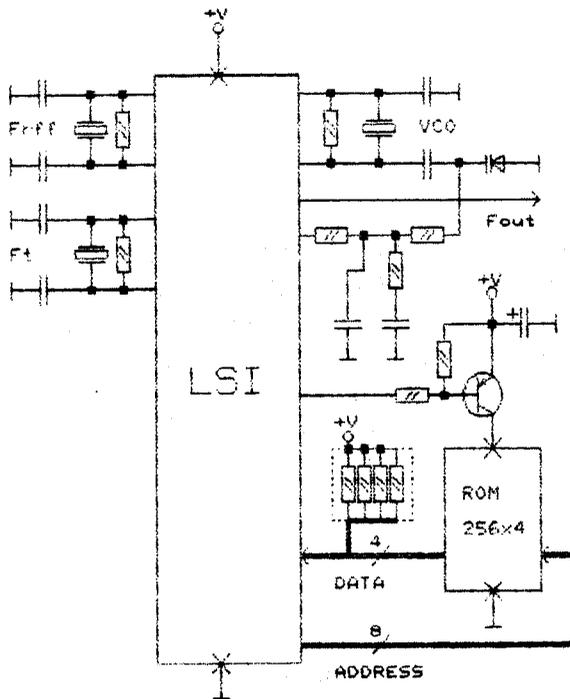


Рис. 8

In all three oscillators internal invertors serve as active parts. Only small number of elements is used as an external frame. It was shown in experiments that with proper thermal coupling between reference and thermosensitive oscillators a stability of  $\pm 0.2$  ppm can be achieved in the temperature interval of  $-60^{\circ}\text{C} \dots +85^{\circ}\text{C}$ .

More complex circuit (with microprocessor and double-mode oscillators) have not yet been tested, through their prototypes using small-scale integration devices showed very good results for that time [6]. One can be confident that large-scale integration will not deteriorate oscillator characteristic.

## CONCLUSION

Basic principles of microcircuit construction were put forth as far back as in 1989-91. Their realization

began in 1993 but positive results were obtained only in 1995. Such slow progress was due to the profound crisis raging in Russia's science and industry then. Many engineering solutions could be done now in another way and better, but the strength of mind has already exhausted and there are no customers for such high technology production in Russia now.

## REFERENCES

1. Patent №1206820 (USSR) "Stochastical piecewise-linear interpolator." A. Kosykh, V. Bagaev. Bulletin № 45, 1985.
2. Harbich D., Povasovich A, Vasilevich D.. 1995 . " The Design of a Microcontroller Temperature Compensated Oscillator ( $\mu\text{CTCXO}$ ) and Authomatic Compensation Line. IEEE Trans. on Ultrasonic, Ferroelectric and Freq. Control. Vol. 42, №4, July 1995, p. 509 - 516.
3. Warwick G, Gosling W, Prescott A, 1978. "A Digital Technique for temperature compensation of crystal oscillators". Proc. Conf. Radio Receivers and assoc. Syst. Southampton, 1978, London, p. 207 - 216.
4. Kosykh A, Ionov B, 1994, " Dynamic Temperature model and Dynamic Temperature compensation of Crystal Oscillator". IEEE Trans. on Ultrasonic, Ferroelectric and Freq. Control. Vol. 41, №3, May 1994, p. 370-374.
5. Patent №1659972A1 (USSR) "Pulse oscillator", Bagaev V, Kosykh A, Lepetaev A, Ionov B. Published in Bulletin № 24, 1991.
6. Kosykh A, Bagaev V, Ionov B, Lepetaev A, Zavalov S. "The new method of statistic piecewise-linear interpolation and its application to DTCXO creation". Proc. 1993 IEEE International Frequency Control Symposium. p. 687-697.

## ACKNOWLEDGMENT

The authors express deep appreciation to V.P. Bagaev, their teacher, who was an inspirer of this work.

# PRARE One resp. Two Years in Orbit – Performance Analysis of the Instruments' Oscillators onboard ERS-2 and Meteor-3/7

S. BEDRICH, F. FLECHTNER

GeoForschungsZentrum Potsdam, Div. 1, Telegraphenberg A17, D-14473 Potsdam  
currently c/o DLR, D-PAF, P.O.Box 1116, D-82230 Oberpfaffenhofen  
(bedrich@dfd.dlr.de)

## Abstract

The Precise Range And Range-Rate Equipment PRARE is a new space-borne microwave satellite tracking system, which is operational since end of April 1995 in redundant configuration onboard the European remote sensing satellite ERS-2. For system function verification, an identical flight model had been tested between February 1994 and September 1995 onboard the Russian meteorological satellite Meteor-3/7.

The system's primary function is to measure very precisely range and range-rate between the space segment and dedicated PRARE ground-station transponders which are distributed around the Earth. From these basic measurements, which have a precision below 3 cm and 0.1 mm/s respectively, several higher level products are derived, such as precise orbits, Earth rotation and gravity field parameters, and clock synchronization quantities [1].

The measurement signals are basically derived from one central oscillator inside the space segment. Analyses of the long-time behaviour of the two instruments' oscillators, which are generated on a routine basis at the PRARE Master Station Oberpfaffenhofen, the accurate modelling of their UTC-related time scales, and results of occasional remote clock synchronization experiments are presented. The excellent suitability of the system for alternative usage as for example highly precise time transfer between remote atomic standards becomes obvious.

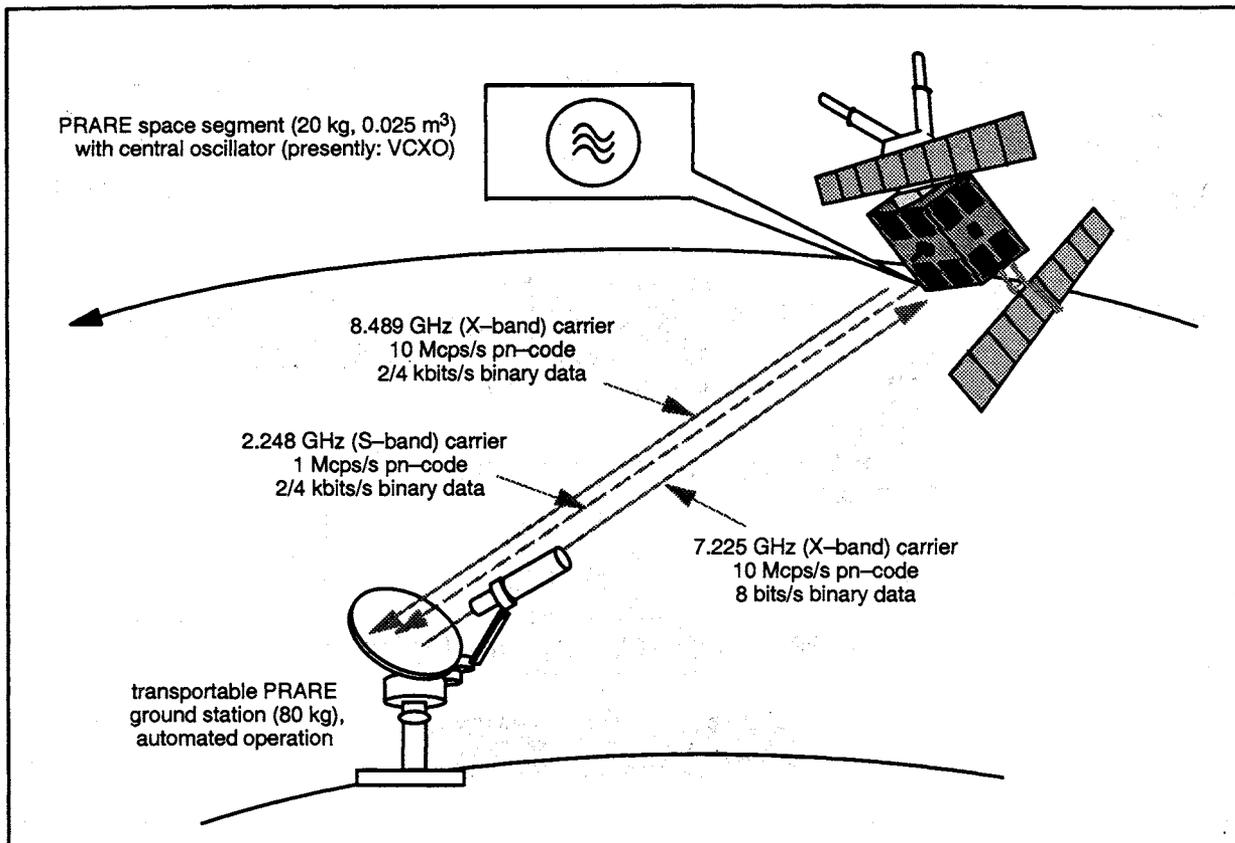
## 1 Introduction

The PRARE system has been developed during the last ten years on behalf of the German Space Agency DARA GmbH by the GeoForschungsZentrum Potsdam and the Stuttgart University's Institute of Navigation [2]. Its original main purposes were highly precise geodetic and geodynamic applications. The system works with low power microwave signals which are structured as a combination of high frequency carriers (X-/S-band), specific pn-codes (10/1 Mcps/s), and spread-spectrum binary data (4/2 kbits/s, fig. 1).

The measurement signals are generated inside the space segment and disseminated permanently to ground; any PRARE ground station in sight of the satellite receives the signals, and up to four preselected stations transmit the downconverted X-band signal simultaneously back to the space segment.

The four receiver channels of the space segment track the selected stations independently with an average duration of about 600 to 800 seconds and carry out autonomously the precise ranging (cf. signal travel time) and range-rate (cf. carrier doppler) measurements. Every second, one averaged two-way range and one two-way doppler value for each of the tracked ground stations is recorded.

Additional information is gained by individual ground station measurements of the ionospheric delay of the coherent X- and the S-band signals



**Figure 1:** The PRARE two-way measurement principle.

and by periodically closed internal transponder loops inside each ground station as well as inside the space segment for hardware delay calibration. Moreover, meteorological measurements are acquired at most of the ground station sites for an a-posteriori tropospheric signal refraction correction. These data are specifically coded and modulated onto the uplink signal to be received by the space segment.

All data are stored highly condensed in the space segment memory. During visibility of the satellite over the ground control segment in Germany, the data are dumped down in a compressed format by modulating the pn-code. This is initiated upon telecommand and controlled during transmission by the Monitoring and System Command Station Stuttgart.

## 2 PRARE System Accuracy

The space segment measurement signals themselves as well as the replica signals which are used

for correlation of the transmitted with the received signals after their two-way travel, are derived fully coherently from one ultra-stable 5 MHz-oscillator inside the space segment. This is the fundamental reason for the unique measurement precision of the system.

Therefore, the monitoring of the frequency stability of this central oscillator and the evaluation of its time correlation with UTC to time tag the measurement data is crucial. It is done at the PRARE Master Station Oberpfaffenhofen, which operates a DCF77-disciplined Rb-clock as system reference side by side with a high quality time interval measurement equipment.

The timing equipment compares routinely the 1-pps codes contained in the PRARE tracking signal against the time information received from the GPS satellites (fig. 2). The derived software clock model "PRARE time vs. UTC(USNO)" is the basis for the ensuing preprocessing of the measurement and the housekeeping data, which is a major tasks of the Master Station.

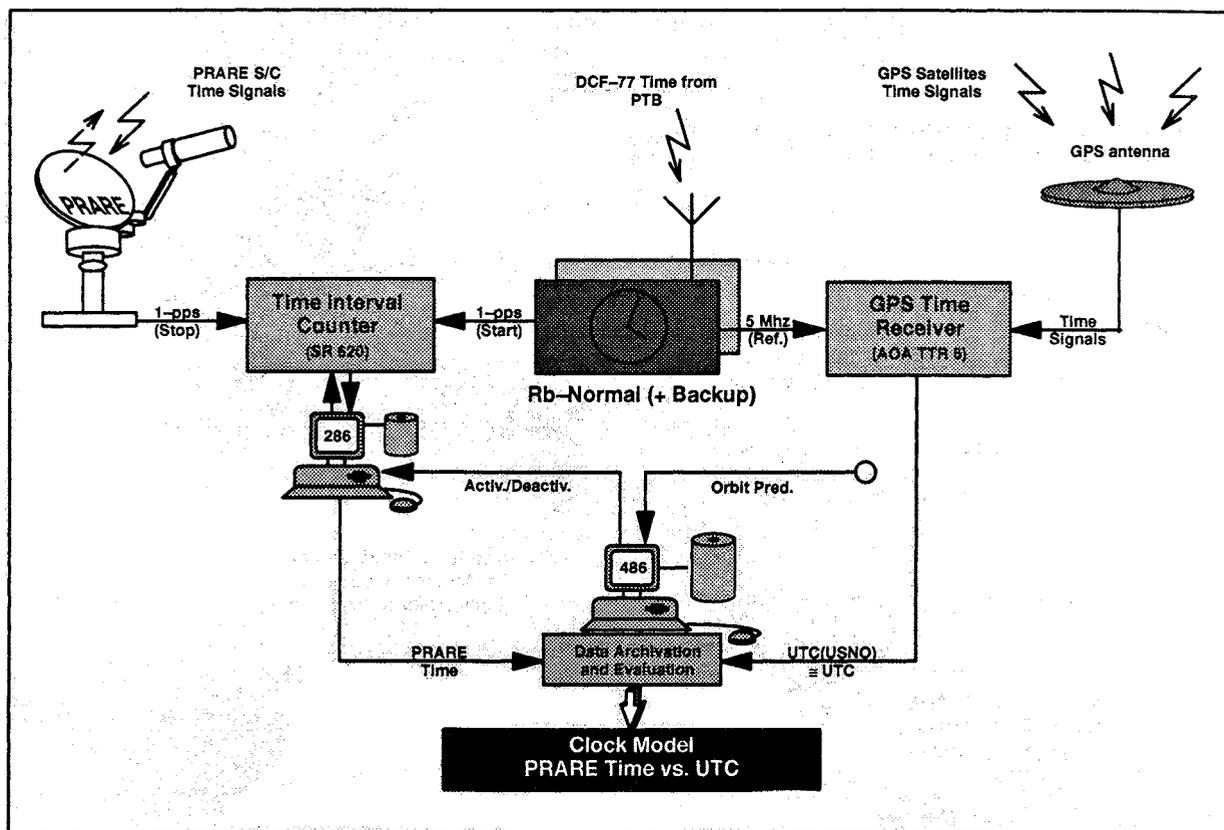


Figure 2: Time System of the PRARE Master Station Oberpfaffenhofen.

### 3 Master Clock Performance

The output oscillation of the Rb-normal operated at the PRARE Master Station is controlled by a periodic correction of its internal C-field to keep the time offset within several microseconds close to UTC(PTB). Under normal operation conditions, this is fulfilled by permanent reception of DCF-77 time signals and internal computation of a correction value which introduces a forced time drift every three days.

An evaluation of the clock data accumulated since beginning of operation of the time system (August 1993) is sketched in fig. 3. With the clock offset never exceeding  $5 \mu\text{s}$  referenced to UTC(PTB), the initial intention of the time system to have a reliable short and long time stability performance is reached. This could be fulfilled by the combination of a Rb-normal in permanent phase-lock (short time stability) with the superposed UTC(PTB) signal reception (long time stability of PTB Cs-standards).

### 4 Space Segment Oscillators

By combination of the computed Rb-clock model parameters with the received PRARE space segment 1-pps signals, a continuous PRARE clock model is generated. It includes permanently updated coefficients for time offset  $c_0$ , time drift/frequency offset  $c_1$ , and frequency drift  $c_2$  (2nd degree polynomial) of the space segment oscillator [3]:

$$t_{UTC} = T_{PRARE} + c_0 + c_1\Delta T + c_2\Delta T^2$$

This is achieved by reducing the 1-pps signals with individual their travel times, which are computed by an orbit prediction software in combination with several constant delays (atmosphere, hardware, etc.). Through this approach, the total accuracy of the model "PRARE time vs. UTC(USNO)" is around  $1 \mu\text{s}$ , which is by far enough for the geodetic purposes of the system.

Evaluation of the long term frequency offset

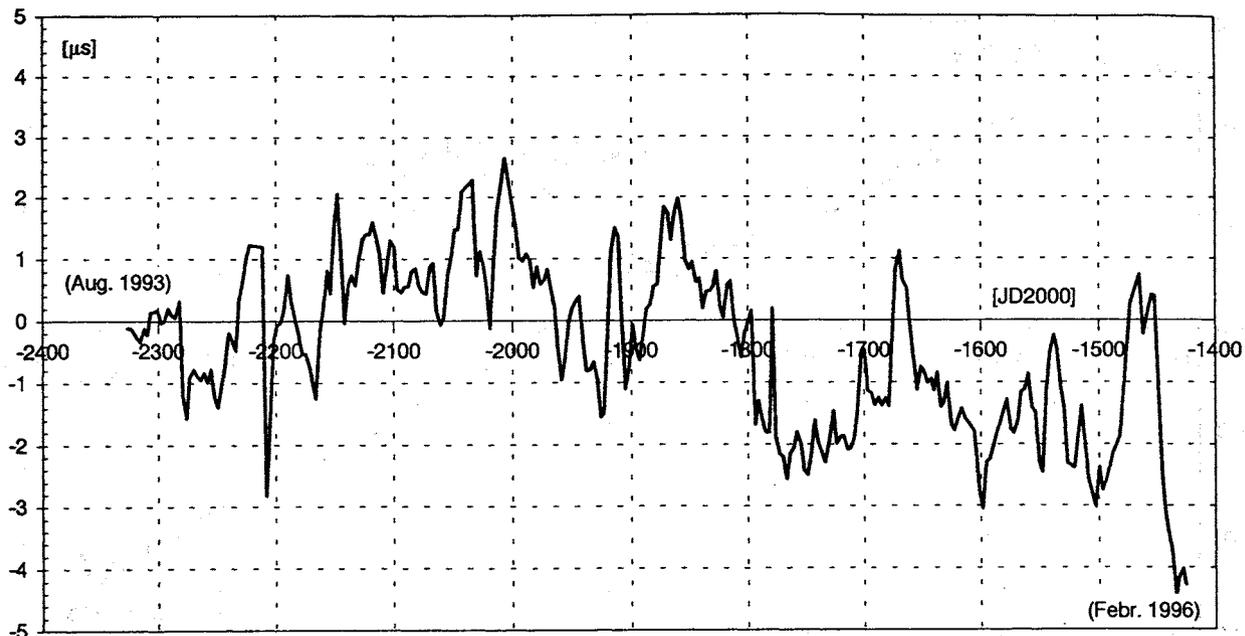


Figure 3: Long term evaluation of the PRARE Master Clock Time Offset (Rb with drift control).

developments of the space segment oscillators PRARE/Meteor-3/7 and PRARE/ERS-2 are shown in fig. 4 (Meteor-3/7 mission operation period: one and a half years) and fig. 5 (ERS-2 operation up to now: nearly one year). For PRARE/Meteor-3/7 there is a drift stabilization observable after first switch-on. After that period, the frequency drift is around  $-4 \cdot 10^{-11}$  Hz/day (referenced to the nominal operating frequency

8.489 GHz). For PRARE/ERS-2, this parameter is with a computed value of around  $-2.5 \cdot 10^{-11}$  Hz/day slightly better.

During operation, each space segment had to be switched-off several times due to diverse reasons. This always causes a several days period of frequency instability, what becomes apparent in the periods with slightly higher data noise.

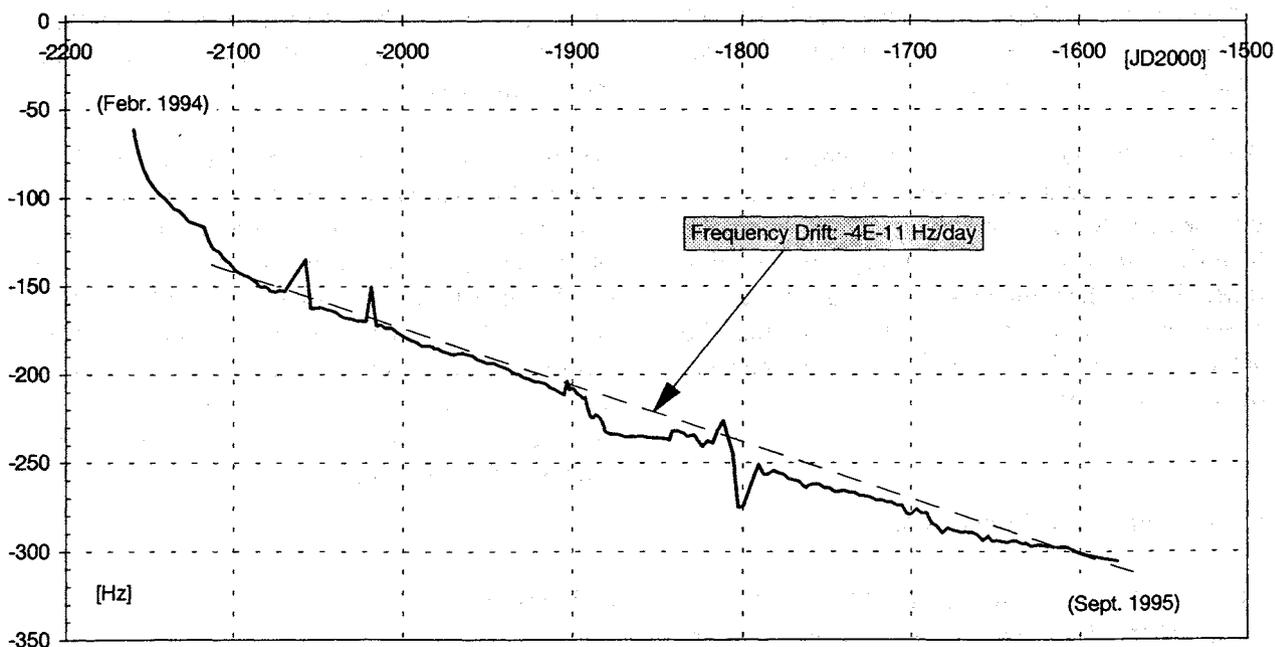


Figure 4: Long term evaluation of the PRARE/Meteor-3/7 oscillator frequency offset.

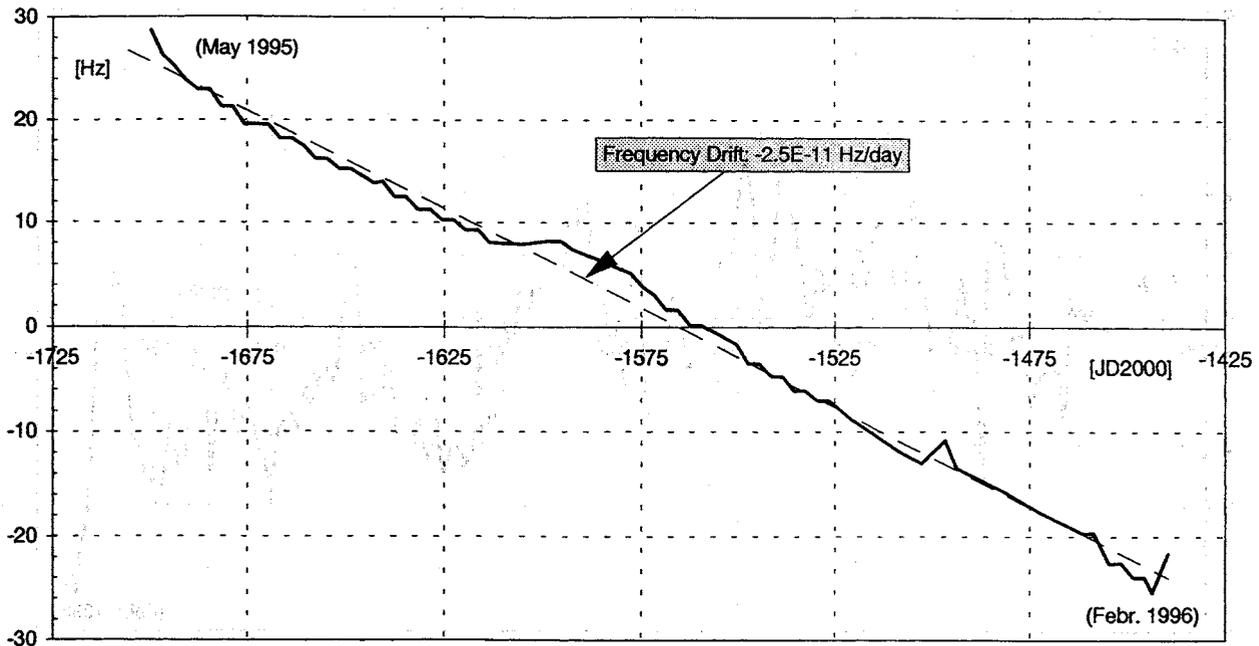


Figure 5: Long term evaluation of the PRARE/ERS-2 oscillator frequency offset.

Due to the different orbital parameters of the satellites (Meteor-3/7: orbit height 1200 km, inclination 82.5 degrees; ERS-2: orbit height 780 km, inclination 98.5 degrees, sun-synchronous), the oscillators are exposed to differing space environments.

A special influence has the Earth's magnetic field with its regions of high dose energetic radiation in the VanAllen belts. Above the Earth's poles, the inner one of these belts crosses the satellites' high inclination orbits, which is decisive for the oscillators' performance.

Although this radiation environment was even more unfavourable for PRARE/Meteor-3/7, no unexpected performance degradation could be identified during the one and a half years operation time. For further analysis, yet a longer lasting observation is necessary, which should be achieved for PRARE/ERS-2 (expected mission duration: 4 years).

## 5 Discussion and Conclusion

The availability of long-term PRARE oscillator data, observed at the PRARE Master Station Oberpfaffenhofen for PRARE/Meteor-3/7 and PRARE/ERS-2, which are based on a DCF77-

and GPS-disciplined Rb-normal as master reference, has led to evaluation of cross-correlation values for the the oscillators' characteristic parameters.

These are mainly dependent on the differing space environments and gravity potentials the instruments are exposed to. It could be shown that the oscillators are well inside their stability specification requirements. Short-time fluctuations are modelled precisely by a dedicated clock modelling algorithm, which is operated at the Master Station on a routine basis.

Oscillator noise has been proven to be very low, and the PRARE signal transmission and correlation technique is adding no further noise. Moreover, the two-way principle allows to synchronize the involved clocks without knowledge of orbital and ground station position informations and vice versa [4]. In view of these facts, the system seems to have the potential to play a central role in a future ultra-precise satellite time transfer concept [5, 6].

## References

- [1] REIGBER, Ch, S BEDRICH, F FLECHTNER, Ch FÖRSTE, T JOPEK, F H MASSMANN,

J C RAIMONDO: Results from PRARE on Meteor-3/7 and ERS-2. *Proc. XXI Assembly of IUGG*, Boulder, July 1995.

[2] HARTL, Ph: The Precise Range and Range-rate Equipment PRARE. *Proc. Workshop on ERS-1 Radar Altimeter Data Products* (ESA SP-221), Frascati, May 1984, 153-160.

[3] BEDRICH, S, A HÄRTING: Monitoring of the PRARE Clock on-board Meteor-3: Operations and Performance Analysis. *Proc. 9th EFTF*, Besançon, March 1995, 34-39.

[4] BEDRICH, S, J HAHN: Precise Two-way

Time Transfer to Space-borne Atomic Clocks. *Proc. 9th EFTF*, Besançon, March 1995, 485-491.

[5] REIGBER, Ch, E W GRAFAREND, W SCHÄFER, H NAU: *PRARETIME - Precise Time Transfer and Tracking for Earth Parameter and Orbit Determination*. Experiment proposal to ESA, GFZ Potsdam / Universität Stuttgart / DLR Oberpfaffenhofen, 1994.

[6] HAHN, J, S BEDRICH, H NAU: *Study on "H-Maser in Space" as Ultra-stable Clock for Time Transfer and Time Dissemination, Final Report*. DLR, Oberpfaffenhofen, 1995.

## DELAY STABILITY MEASUREMENTS MADE WITHIN A TWO-WAY TIME TRANSFER SYSTEM USING SATELLITE RANGING FROM SEVERAL LOCATIONS.

J A Davis.

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

### ABSTRACT

A method is developed to study the stability of a Two-Way Satellite Time and Frequency Transfer (TWSTFT) system during regular operations. Four or more earth stations simultaneously determine the range of the geostationary satellite. The redundant information obtained from the over determination of the satellite's position is used to estimate the delay stability of individual TWSTFT links.

### 1) INTRODUCTION

Two-Way Satellite Time and Frequency Transfer (TWSTFT) is being developed as a high-precision method for both time and frequency transfer (Kirchner (1)) (Davis et al (2)). The accuracy of the TWSTFT method is however limited by delay instabilities within the satellite earth station instrumentation. Recent experiments suggest that these instabilities may vary from 200 picoseconds up to several nanoseconds (3) (Kirchner et al (4)). To obtain the best possible performance from the TWSTFT system, it is important to study the delay stability of the earth station instrumentation.

Several methods have been used to study the delay stability of both the earth station instrumentation and the complete TWSTFT system. Time transfers have been performed between distant laboratories over extended periods of several hours. When commercial caesium clocks are compared, for example the HP5071A, the instabilities of the caesium clocks are relatively large. After only a few minutes averaging the time transfer  $\sigma_x$  is dominated by clock noise (2). With active hydrogen masers operating at both timing laboratories, the transfer  $\sigma_x$  is dominated by instabilities in the TWSTFT system. This is the case for averaging times of up to several days. These Hydrogen maser comparisons provide a good test of the TWSTFT system, however active hydrogen masers are available at only a relatively few laboratories. The use of colocated earth stations has provided valuable information on the delay stability of the TWSTFT earth station instrumentation (3). The construction of two or more TWSTFT earth stations at a single location is still relatively expensive, and any delay instabilities due to ionospheric and tropospheric delays will cancel due to the presence of a common atmosphere. More recently, satellite simulators have been successfully used to determine both the absolute delay asymmetry and the stability of the earth station instrumentation (Kirchner et

al (4)) (De Jong (5)).

During 1994 and 1995 six European primary timing laboratories have been performing TWSTFT experiments using the Intelsat satellite at 307°E (2). A single satellite transponder was used to carry all the time transfer links. Under these circumstances the earth stations have the capability to range the satellite simultaneously. By ranging the satellite from three locations, and knowing the earth station and transponder delays, the satellite's position may be determined. When four or more stations range the satellite, the satellite's position is over-determined. The information resulting from this redundancy may be used to provide information on the delay stability of the TWSTFT system. An initial study of the delay stability of the TWSTFT system, obtained by using ranging data is presented here.

In this paper the results of two experiments are reported. In the first, ranging measurements of five minutes duration are made during measurement sessions occurring three times per week, over a period of 18 months. In the second experiment, all six laboratories range the satellite simultaneously for a period of approximately eight hours. Linear combinations of ranging data are formed, the delays of these combinations are shown to a first approximation to be independent of the residual motion of the geostationary satellite (Davis and Pearce (6)). Any long term changes in these delays is due principally to the delay changes in the earth station instrumentation.

### 2) EUROPEAN TWSTFT EXPERIMENT

Six laboratories have been participating in the European TWSTFT experiment using the Intelsat satellite at 307°E. The characteristics of both the TWSTFT earth stations and atomic clocks have been given previously (2). A description of the TWSTFT method is given in (1). Regular time transfer sessions were performed three times per week starting at 10:00 UTC. Each session consisted of a series of five-minute time transfers made between each pair of European laboratories. The satellite ranging measurements are performed at the start of each session. Ranging data used in this work has been contributed by the Technical University of Graz (TUG), National Physical Laboratory (NPL), Van Swinden Laboratorium (VSL), Forschungs-und Technologiezentrum, (FTZ) and the Physikalisch-Technische Bundesanstalt (PTB).

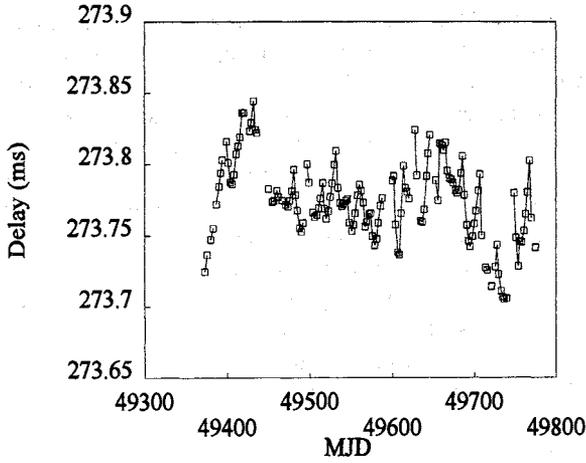


Figure 1: Ranging data measured at TUG

### 3) STUDY OF TWSTFT RANGING DATA

Long term variations of TWSTFT ranging measurements are shown in Figure 1. The mean values of the ranging data collected between 10:00 and 10:05 UTC on three occasions per week is plotted over a period of 18 months. Variations in the ranging delay measurements of 150 microseconds peak to peak are observed. These variations are almost completely due to the residual motion of the near-geostationary satellite. Delay stabilities better than one nanosecond have been obtained using the TWSTFT instrumentation (3). In order to successfully use ranging data to obtain measurements of the delay stability of the TWSTFT earth station instrumentation, the effects of residual motion of the satellite must be almost completely eliminated.

### 4) LINEAR COMBINATIONS OF RANGING DATA

The following method may be used to obtain information of the delay stability of the TWSTFT system from the ranging data collected during the experiments. Consider ranging measurements made at four distant locations, A, B, C and D. Following the terminology given in (1), (Davis et al (7)) and (Davis (8)), the counter timer readings recorded at station A when ranging the satellite is given by  $\Delta T_{AA}$  where:

$$\Delta T_{AA} = \tau_{AA}^{Tx} + \tau_{AA}^{Rx} + \tau_{AA}^U + \tau_{AA}^D + \tau_{AA}^S \quad (1)$$

where  $\tau_{AA}^{Tx}$  and  $\tau_{AA}^{Rx}$  are the up-link and down-link earth station instrumentation delays. These may be determined as described in (5).  $\tau_{AA}^U$  and  $\tau_{AA}^D$  are the up and down path components of the ranging delay of station A.  $\tau_{AA}^S$  is the satellite transponder delay. The total ranging path delay  $\tau_{AA}^P$  is defined as the sum of the path delays and  $(\tau_{AA}^U + \tau_{AA}^D)$ .

Neglecting tropospheric and ionospheric delay changes, the total path delay  $\tau_{AA}^P$  may be expressed in terms of coordinates of both the earth station and the satellite,

along with the free space speed of light  $c$ . Using earth-centred spherical polar coordinates  $(r, \theta, \phi)$ , fixed with respect to the earth's surface, the total path delay  $\tau_{AA}^P$  is given by:

$$\tau_{AA}^P = 2c |(\mathbf{R}_S(r_S, \theta_S, \phi_S) - \mathbf{R}_A(r_A, \theta_A, \phi_A))| = 2c(r_S^2 - r_A^2 - 2r_S r_A (\sin \theta_S \sin \theta_A \cos(\theta_S - \theta_A) + \cos \theta_S \cos \theta_A))^{0.5} \quad (2)$$

where  $\mathbf{R}_S$  and  $\mathbf{R}_A$  are the position vectors of the satellite and earth station.

Due to the limited motion of the near geostationary satellite, small changes in the total path delay  $\delta \tau_{AA}^P$  are assumed to depend linearly on the associated small changes in the components of the satellite's position  $\delta \mathbf{R}_S(r, \theta, \phi)$ . The validity of this approximation is discussed later. The constants of proportionality  $K_{r_{AA}}$ ,  $K_{\theta_{AA}}$ , and  $K_{\phi_{AA}}$  are obtained by differentiating  $\tau_{AA}^P$  with respect to the components of the satellite's position:

$$K_{r_{AA}} = d\tau_{AA}^P / dr_S \quad (3a)$$

$$K_{\phi_{AA}} = d\tau_{AA}^P / d\phi_S \quad (3b)$$

$$K_{\theta_{AA}} = d\tau_{AA}^P / d\theta_S \quad (3c)$$

A similar method has been used previously to eliminate the effects of satellite longitudinal drift from time transfers made using the common view of Direct Broadcasting Satellites (6).

Linear combination were formed from the ranging data obtained from four locations where:

$$C_{ABCD} = \alpha_A \Delta T_{AA} + \alpha_B \Delta T_{BB} + \alpha_C \Delta T_{CC} + \alpha_D \Delta T_{DD} \quad (4)$$

The following constraints were applied when choosing the coefficients  $\alpha_A$ ,  $\alpha_B$ ,  $\alpha_C$  and  $\alpha_D$ .

$$\alpha_A K_{r_{AA}} + \alpha_B K_{r_{BB}} + \alpha_C K_{r_{CC}} + \alpha_D K_{r_{DD}} = 0 \quad (5a)$$

$$\alpha_A K_{\theta_{AA}} + \alpha_B K_{\theta_{BB}} + \alpha_C K_{\theta_{CC}} + \alpha_D K_{\theta_{DD}} = 0 \quad (5b)$$

$$\alpha_A K_{\phi_{AA}} + \alpha_B K_{\phi_{BB}} + \alpha_C K_{\phi_{CC}} + \alpha_D K_{\phi_{DD}} = 0 \quad (5c)$$

$$|\alpha_A| + |\alpha_B| + |\alpha_C| + |\alpha_D| = 1 \quad (5d)$$

The last constraint was chosen so that the linear combination of ranging data  $C_{ABCD}$  would be calculated from the same total instrumentation delay as a two-way time transfer. Under these constraints, to a first approximation:

$$dC_{ABCD}/dr_S = 0, \quad dC_{ABCD}/d\theta_S = 0, \quad dC_{ABCD}/d\phi_S = 0 \quad (6)$$

The delay of the linear combination  $C_{ABCD}$  is independent of small changes in the satellite's position. From equations (1) and (4) and neglecting ionospheric and tropospheric effects:

$$\begin{aligned}
C_{ABCD} = & \alpha_A(\tau_{AA}^{Tx} + \tau_{AA}^{Rx}) + \alpha_A(\tau_{BB}^{Tx} + \tau_{BB}^{Rx}) + \\
& \alpha_C(\tau_{CC}^{Tx} + \tau_{CC}^{Rx}) + \alpha_D(\tau_{DD}^{Tx} + \tau_{DD}^{Rx}) + \alpha_A\tau_{AA}^P + \alpha_B\tau_{BB}^P + \\
& \alpha_C\tau_{CC}^P + \alpha_D\tau_{DD}^P + \alpha_A\tau_{AA}^S + \alpha_B\tau_{BB}^S + \alpha_C\tau_{CC}^S + \alpha_D\tau_{DD}^S \quad (7)
\end{aligned}$$

From geometry, the sum of the satellite delays  $\alpha_{11}\tau_{11}^S + \alpha_{22}\tau_{22}^S + \alpha_{33}\tau_{33}^S + \alpha_{44}\tau_{44}^S$  may be shown to approximate to zero. The sum of the propagation delays has been shown to a first approximation to be constant. Therefore neglecting tropospheric and ionospheric effects, delay changes of the linear combination  $C_{ABCD}$  are to a first approximation due only to delay changes in the earth station instrumentation.

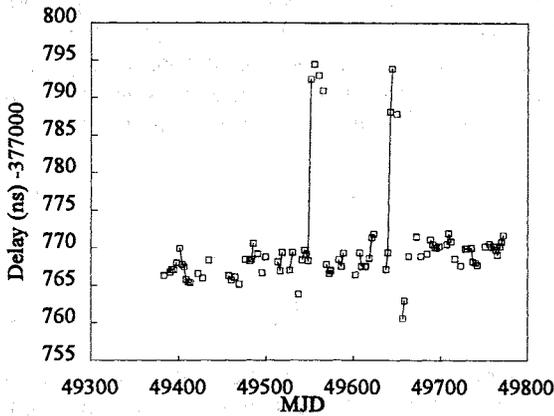


Figure 2: Linear combination of ranging data  
0.198TUG + 0.217VSL - 0.500FTZ + 0.085PTB

Assuming no correlation between the delay stabilities of the earth station instrumentation at different locations then:

$$\begin{aligned}
\sigma_x^2(C_{ABCD}) = & \alpha_A^2\sigma_x^2(\tau_{AA}^{Tx} + \tau_{AA}^{Rx}) + \alpha_B^2\sigma_x^2(\tau_{BB}^{Tx} + \tau_{BB}^{Rx}) + \\
& \alpha_C^2\sigma_x^2(\tau_{CC}^{Tx} + \tau_{CC}^{Rx}) + \alpha_D^2\sigma_x^2(\tau_{DD}^{Tx} + \tau_{DD}^{Rx}) \quad (8)
\end{aligned}$$

with similar expressions for  $\sigma_y^2(C_{ABCD})$ . Using five or more earth stations, a series of linear combinations  $C_{ABCD}$  may be formed. Using these combinations estimates of  $\sigma_y(\tau_{AA}^{Tx} + \tau_{AA}^{Rx})$  or  $\sigma_x(\tau_{AA}^{Tx} + \tau_{AA}^{Rx})$  may be made for the delay stability of the earth station instrumentation of stations with the higher weighting  $\alpha_A$ .

## 5) EXPERIMENTAL RESULTS

Linear combinations of ranging data have been formed from regular measurements made three times per week, during an 18 month period. The values of  $C_{ABCD}$  have been optimised by adjusting the weighting factors  $\alpha_A$ . This was performed to minimise the scatter on the data, and reduce effects due to errors in the earth station coordinates and linearization approximation. Examples are shown in Figures 2 and 3 for the TUG-VSL-FTZ-PTB and TUG-NPL-FTZ-PTB linear combination. A typical peak to peak scatter of four nanoseconds was

observed. This scatter corresponds to the maximum possible instrumentation delay instabilities effecting the two-way time transfers. This is because other effects, for example tropospheric delay changes or residual satellite motion may have been contributing to the scatter on the linear combinations  $C_{ABCD}$ . The weighting factors  $\alpha_A$  of the stations was much higher for those near the geographic centre of the group. In Figure 2 discrete delay steps have been observed on MJD 49543 and MJD 49630. By examining several different linear combinations, these delay changes were found to have occurred within the earth station instrumentation located at FTZ. A similar delay step occurring in Figure 3 has been removed by eliminating the effected data, in order to reveal the scatter on the major portion of the data. Standard deviations of 1.3 nanoseconds were determined for the delay spread of these linear combinations, after removing the data effected by the major delay step. Values of  $\sigma_y(C_{ABCD})$  have been calculated for  $\tau = 2, 5$  and 7 days, these are shown in Table 1. Values of  $\sigma_y(C_{ABCD})$  have been compared against  $\sigma_y$  values obtained from TWSTFT made over several days. The PTB-NPL, PTB-TUG and NPL-TUG time transfer was chosen for comparisons because there are high quality clocks used at each location, an active Hydrogen maser at NPL, a primary caesium clock at PTB, and a high quality commercial cesium clock at TUG. Values of  $\sigma_y$  obtained from the time transfers are higher than those obtained from the combinations of ranging data. This suggests that clock noise is making a significant contribution to the transfer  $\sigma_y$  calculated from the time transfers.

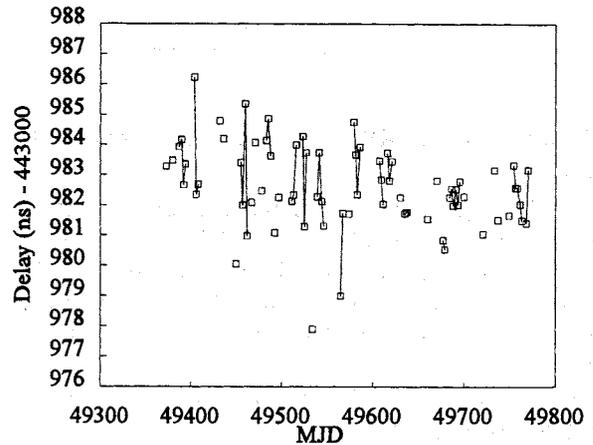


Figure 3: Linear combination of ranging data  
0.189TUG + 0.102NPL - 0.500FTZ + 0.209PTB

Linear combinations were formed from eight hours of ranging data. An example is shown in Figure 4 for the TUG-NPL-FTZ-PTB combination. Each point has been calculated from the mean of 60 measurements. A standard deviations of 0.32 nanoseconds was obtained over the eight hours of measurements, with peak to peak scatter of two nanoseconds. This is comparable to the delay stability obtained from co-located common clock measurements made individually at each of these sites. The ranging results may also be compared with time

Linear Combinations	$\sigma_y(\tau)$ $\tau=2$ days	$\sigma_y(\tau)$ $\tau=5$ days	$\sigma_y(\tau)$ $\tau=7$ days
FTZ-VSL-NPL-TUG	$1.4 \times 10^{-14}$	$4.7 \times 10^{-15}$	$7.0 \times 10^{-15}$
PTB-VSL-NPL-TUG	$1.5 \times 10^{-14}$	$7.2 \times 10^{-15}$	$4.8 \times 10^{-15}$
PTB-FTZ-NPL-TUG	$1.2 \times 10^{-14}$	$4.4 \times 10^{-15}$	$4.9 \times 10^{-15}$
PTB-FTZ-VSL-TUG	$6.3 \times 10^{-15}$	$4.2 \times 10^{-15}$	$4.8 \times 10^{-15}$
PTB-FTZ-VSL-NPL	$1.5 \times 10^{-14}$	$7.9 \times 10^{-15}$	$4.0 \times 10^{-15}$
Time Transfer	$\sigma_y(\tau)$ $\tau=2$ days	$\sigma_y(\tau)$ $\tau=5$ days	$\sigma_y(\tau)$ $\tau=7$ days
PTB-NPL	$1.8 \times 10^{-14}$	$1.0 \times 10^{-14}$	$1.0 \times 10^{-14}$
PTB-TUG	$2.9 \times 10^{-14}$	$1.6 \times 10^{-14}$	$1.3 \times 10^{-14}$
NPL-TUG	$2.4 \times 10^{-14}$	$1.5 \times 10^{-14}$	$1.2 \times 10^{-14}$

Table 1: Comparison of  $\sigma_y$  values calculated from linear combinations of ranging data and time transfers.

transfers made over extended periods (2), including those made between NPL and the United States Naval Observatory in Washington USA, where active hydrogen masers were used at each location. These time transfers show a peak to peak scatter of 1.5 ns over a period of 5 hours. Plots of  $\text{Log}_{10}(\sigma_x)$  against  $\text{Log}_{10}(\tau)$  are shown for several combinations in Figure 5. Curves show a minimum when  $\tau = 300$  seconds. These results show good agreement with plots of  $\text{Log}_{10}(\sigma_x)$  against  $\text{Log}_{10}(\tau)$  obtained from TWSTFT made over extended periods (2), including transatlantic time transfers. The principle cause of delay instability in the measurements made over

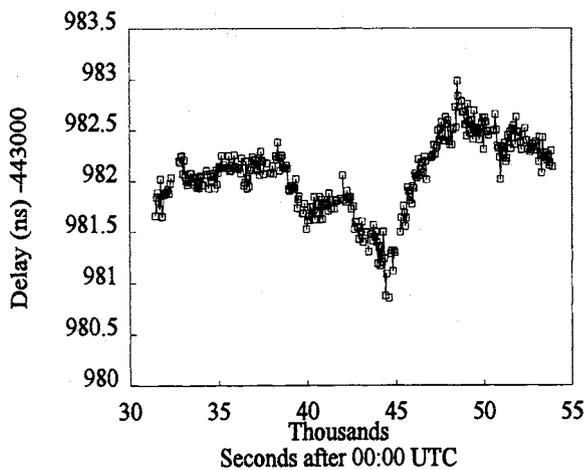


Figure 4: Linear combination of ranging data  $0.189\text{TUG}+0.102\text{NPL}-0.500\text{FTZ}+0.209\text{PTB}$

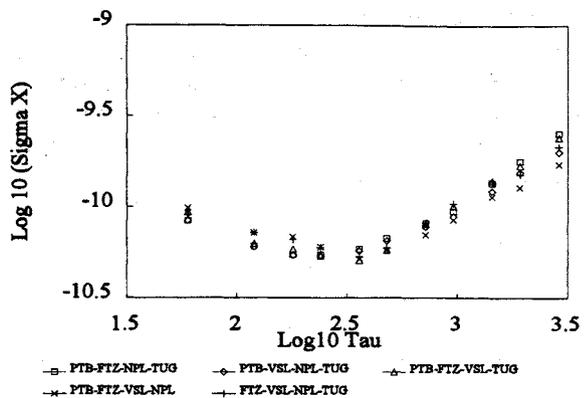


Figure 5: Plots of  $\text{Log}_{10}(\sigma_x)$  against  $\text{Log}_{10}(\tau)$  calculated from linear combinations of ranging data.

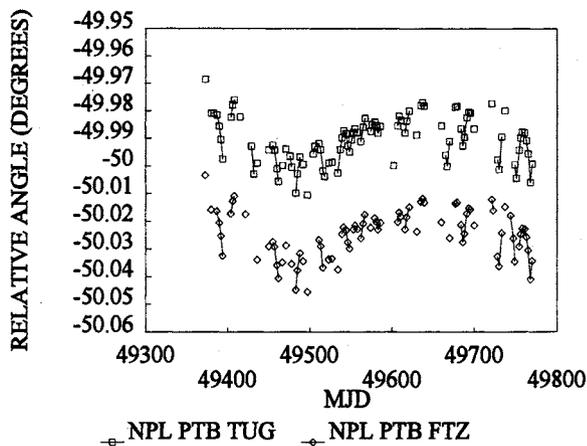


Figure 6: Changes in the longitude component of the satellite's orbit calculated from two combinations of ranging data (arbitrary offset).

extended periods up to eight hours appears to originate within the TWSTFT earth station instrumentation.

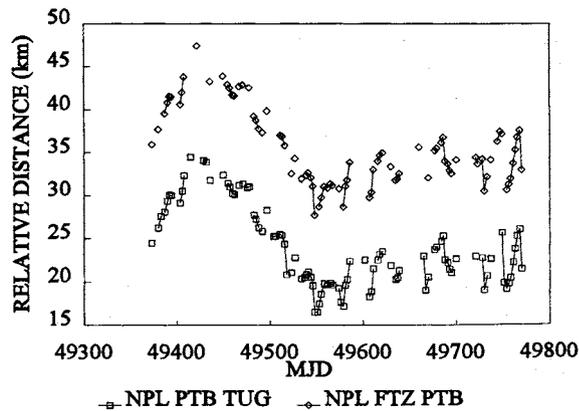


Figure 7: Changes in the radius component of the satellite's orbit calculated from two combinations of ranging data (arbitrary offset).

Plots of the variation of the components of the satellite's position at 10:00 UTC, over a period of 18 months are shown in Figures 6, 7 and 8. The satellite's positions have been determined using linear combinations of ranging data obtained from three locations. These combinations were constructed to be insensitive to variations in two components of the satellite's position. The components of the satellite's position were determined from two different linear combinations. The comparison of Figures 6, 7 and 8 with Figures 2 and 3 demonstrated that there is no observable correlation between changes in the satellite's position and the delay changes of the linear combinations of ranging data  $C_{ABCD}$ . This suggests that the residual motion of the near geostationary satellite is not contributing significantly to the delay instability of the linear combinations  $C_{ABCD}$ .

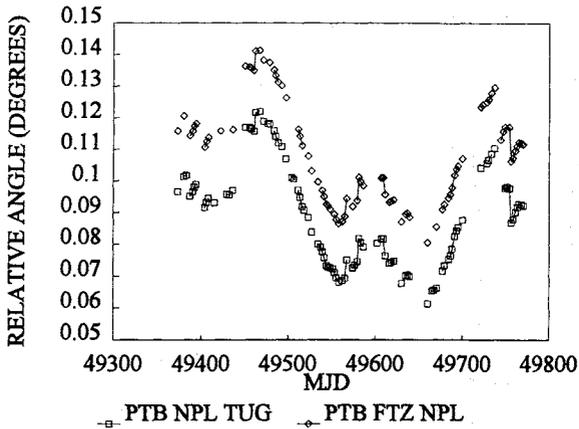


Figure 8: Changes in the latitude component of the satellite's orbit calculated from two combinations of ranging data (arbitrary offset).

## 6) LIMITATIONS OF THE LINEARIZATION APPROXIMATION

In the present analysis the changes in the ranging delays are to a first approximation assumed to vary linearly with changes in the satellite's position. The variation in the satellite's position may induce delay changes of typically 50-100 microseconds into the ranging data. In contrast the delays of the linear combinations of ranging data are required to be independent of satellite position at the nanosecond level. In the analysis described here the constants  $K_{TAA}$ ,  $K_{\theta AA}$ , and  $K_{\phi AA}$  are assumed to have fixed values. In reality these terms are dependent on the position of the satellite and may be determined from equations (2) and (3). Values have been calculated for changes of up to  $\pm 0.1^\circ$  in the satellite's longitude and latitude, and up to  $\pm 20$  km change in the orbit radius. These changes correspond to the maximum permitted station keeping manoeuvres of the Intelsat (VA-F13) satellite. The variation of the delay of the linear combination  $C_{ABCD}$  with satellite orbit radius has been calculated from equation (9) below. Similar expressions were used to determine the changes of the linear combinations resulting from changes of satellite

longitude and latitude:

$$d^2C_{ABCD}/dr_s^2 = \alpha_A dK_{TAA}/dr_s + \alpha_B dK_{rBB}/dr_s + \alpha_C dK_{rCC}/dr_s + \alpha_D dK_{rDD}/dr_s^2 \quad (9)$$

The values of  $d^2C_{ABCD}/dr_s^2$ ,  $d^2C_{ABCD}/d\theta_s^2$ ,  $d^2C_{ABCD}/d\phi_s^2$  remain almost constant over the distances involved in station keeping manoeuvres. By integrating these changes between the limits of the permitted station keeping it is possible to calculate the changes of the ranging delay  $\delta C_{ABCD}$  for small variations of satellite orbit components ( $\delta r_s, \delta \theta_s, \delta \phi_s$ ) where:

$$\delta C_{ABCD} = 0.5(dC_{ABCD}/d\theta_s)\delta\theta_s^2 \quad (10a)$$

$$\delta C_{ABCD} = 0.5(dC_{ABCD}/dr_s)\delta r_s^2 \quad (10b)$$

$$\delta C_{ABCD} = 0.5(dC_{ABCD}/d\phi_s)\delta\phi_s^2 \quad (10c)$$

for small changes in radius, latitude and longitude respectively. An example of the expected delay changes of the linear combinations ABCD with satellite orbit latitude is shown in Figure (9). Between the station-keeping limits variations of satellite latitude introduce the greatest changes to the delays of the linear combinations. In contrast the delays of the linear combinations are almost insensitive to changes in satellite orbit radius. In practice the observed changes in satellite latitude and longitude are only a fraction of their station keeping limits. In this study satellite longitude and latitude changed by only  $\pm 0.03^\circ$  from a nominal position (Figure 6,7 and 8). Under these circumstances the delay variation may amount to one nanosecond (peak to peak) in the most extreme cases. In the majority of cases it will be much less. The delay changes introduced by satellite position variation are close to the acceptable limits. If an inclined orbit satellite is used in these measurements, or the earth stations were located further apart, then the satellite position variations would introduce significant delay changes into the linear combination  $C_{ABCD}$ . It is possible to calculate correction terms for the variation of the satellite from its nominal position.

## 7) TROPOSPHERIC DELAY CHANGES

Tropospheric delays of electromagnetic signals are independent of the signal propagation frequencies. These delays will cancel when calculating a two-way time transfer. In contrast, tropospheric delays do not cancel when calculating ranging delays. Tropospheric delays of up to nine nanoseconds may occur when making transmissions to a satellite directly overhead (Kirchner et al (9)). Ranging measurements involve transmitting a signal through the troposphere twice. For satellites with inclination angles of approximately  $10^\circ$ , that is typical for the present European Intelsat experiment, the total troposphere delay may be as much as 60 ns. Variations of the troposphere delay may

significantly change the ranging measurements and hence the value of the linear combination  $C_{ABCD}$ . The occurrence of common weather patterns over Europe may act to reduce any seasonal troposphere effects. Further work is however required in this area, modelling the troposphere delay changes during a campaign of ranging measurements would improve the accuracy of the ranging method.

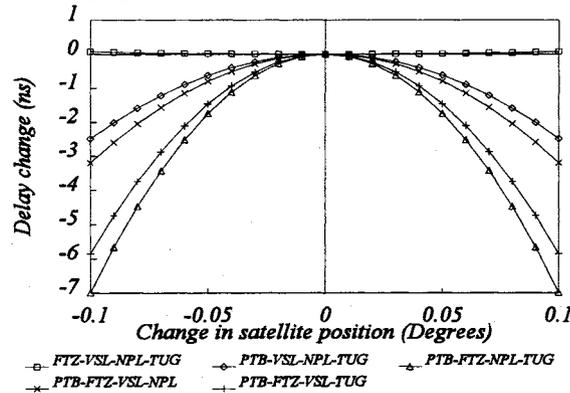


Figure 9: Variations in the delays of linear combinations of ranging data  $C_{ABCD}$  with satellite latitude.

## 8) IONOSPHERIC DELAY CHANGES

In contrast to the troposphere delays the Ionosphere delays are dependent on the signal propagation frequency. The Ionosphere propagation delay is proportional to  $1/f^2$  where  $f$  is the propagation frequency. A large proportion of the Ionospheric delay changes will cancel when calculating a two-way time transfer. This is due to the closeness of the transmit and received frequencies. The residual delay changes are proportional to  $(1/f_u^2 - 1/f_d^2)$ . When calculating ranging delays the ionospheric delays add and are proportional to  $(1/f_u^2 + 1/f_d^2)$ . In the case of the Intelsat satellite this may increase the effect of the ionospheric delays delay by a factor of nine. With the relatively low elevation angles used, a maximum ionosphere delay change of several nanoseconds may occur in the delay of the linear combination. This is unlikely to be a serious problem, the presence of a common ionosphere over Europe will greatly reduce the effect.

## 9) CONCLUSION

A method has been developed to obtain information on the delay instabilities occurring within a TWSTFT system by examining ranging data obtained from four TWSTFT earth stations simultaneously. Delay changes produced by the residual motion of the near geostationary satellite have been successfully eliminated using this method. Troposphere, Ionosphere, and cable delay changes may produce different effects on the delay stabilities of TWSTFT measurements and ranging measurements. Further work is required to both to develop the ranging method and to study origins of the delay instabilities within the TWSTFT system.

## 10) ACKNOWLEDGEMENTS

The author thanks D Kirchner (TUG), P Hetzel (PTB), G DeJong (VSL), and A Soering (FTZ) for the provision of ranging data. This work has been supported under the National Measurement System Program for Time and Frequency.

## REFERENCES

- (1) D. Kirchner 'Two way time transfer via communication satellites' IEEE Proc. No.79, pp.983-990, 1991.
- (2) J A Davis, P R Pearce, D Kirchner, H Ressler, P Hetzel, A Söring, G De Jong, P Grudler, F Baumont, L Veenstra, 'European two-way satellite time transfer experiments using the Intelsat (VA-F13) satellite at 307°E, IEEE J. Trans. Inst. and Meas. vol 44, No 2, April 1995.
- (3) Delay stability of satellite earth stations as used in two-way transfer, CCIR Radio-communications study groups, document 7-2/7-E.
- (4) D Kirchner, H Ressler and R Robnik, "An automated signal delay monitoring system for a two-way satellite time transfer terminal", Proc. 9th European Frequency and Time Forum, Besancon, pp75-79, March 1995.
- (5) G De Jong, D Kirchner, H Ressler, P Hetzel, J A Davis, P R Pearce, W Powell, A Davis McKinley, W Klepczynski, J A DeYoung, C Hackman, S R Jefferts, T E Parker, "Results of the calibration of the delays of earth stations for TWSTFT using the VSL satellite simulator method", Proceedings Precise Time and Time Interval (PTTI 95) Applications and Planning meeting, November 1995.
- (6) J A Davis, P R Pearce, 'Analytical techniques for improved time transfer using common view of direct broadcasting satellites', Proc. Sixth European Frequency and Time Forum, pp.195-199, March 1992.
- (7) J A Davis, P R Pearce, K A Bartle, P Hetzel, D Kirchner, H Ressler, W Powell, A McKinley, J A De Young, W Klepczynski, G De Jong, A Soering, F Baumont, P Claudon, P Grudler, C Hackman, L Veenstra, 'Closing errors observed during two-way time transfer experiments'. Proceedings of the 9th European Frequency and Time Forum, 1995.
- (8) J A Davis 'Theoretical treatment of closing errors occurring in two-way satellite time transfers' Proceedings of the 9th European Frequency and Time Forum, 1995.
- (9) D. Kirchner, C. Lentz, "Tropospheric Corrections to GPS Measurements Using Locally Measured Meteorological Parameters Compared with General Tropospheric Corrections," Proc. 25th PTTI, pp. 231-248, 1994.

# A MICROWAVE FREQUENCY STANDARD BASED ON TRAPPED, BUFFER GAS-COOLED $^{171}\text{Yb}^+$ IONS - SOME ACCURACY ISSUES

Peter T.H. Fisk, Matthew J. Sellars, Malcolm A. Lawn, Colin Coles

National Measurement Laboratory  
CSIRO Division of Applied Physics  
PO Box 218 Lindfield NSW 2070  
Sydney, Australia

## Introduction

Trapped  $^{171}\text{Yb}^+$  ions have been under investigation as a microwave frequency standard for a number of years (1,2,3,4,5). At the CSIRO National Measurement Laboratory (NML) two prototype frequency standards, IT-1 and IT-2, based on buffer gas cooled  $^{171}\text{Yb}^+$  ions confined in a linear Paul trap have demonstrated promising stability and accuracy characteristics (4,5,6).

In this paper we report briefly on the performance of the prototype microwave frequency standards, and discuss some of the issues concerning the accuracy of the standards.

## Operational Details

The principles of operation of the NML trapped ion standard have been described previously (4), and will be only briefly covered here.

The energy levels relevant to the NML trapped ion standard are shown in fig. 1. The standard operates by stabilising the frequency of a 12.6 GHz microwave source to the ground state hyperfine transition (the 'clock' transition) of  $^{171}\text{Yb}^+$  ions. When the microwave source is tuned precisely to the clock transition, the ions fluoresce under illumination by a 369.5 nm laser tuned to one of the hyperfine components of the optical resonance transition. Stabilising the frequency of the microwave source therefore involves, in principle, adjusting its frequency so as to maximise this fluorescence.

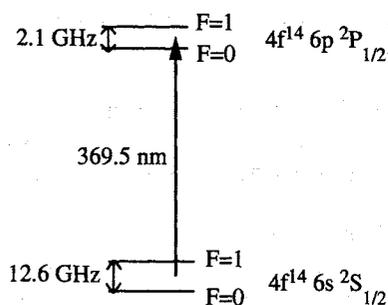


Figure 1: Partial energy level diagram of  $^{171}\text{Yb}^+$  ions.

The linear Paul trap (fig. 2) has also been described in detail previously (4). The stability results reported in this paper were obtained with 250 Vp-p at 510 kHz applied to the RF electrodes, and +10 V applied to the DC end-electrodes. Under these operating conditions the trap holds a cloud of radius 2 mm and length 10 mm, consisting of  $10^5$  ions. In the transverse direction, the RF pseudopotential of the empty trap is harmonic to a very good approximation over the region occupied by the ion cloud. The harmonic pseudopotential is characterised by the angular frequency  $\omega_s = 2\pi \times 19.8$  kHz of the transverse thermal harmonic motion executed by the confined ions when the trap is close to empty.

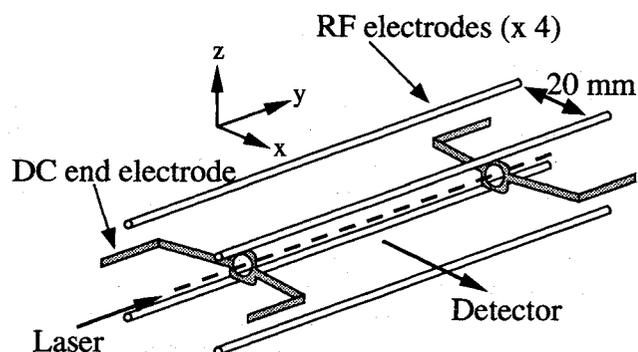


Figure 2: Schematic electrode structure of the linear ion traps IT-1 and IT-2. In this paper we refer to motion in the  $x$ - $z$  plane as transverse, and motion along the  $y$  axis as longitudinal. The diameter of the RF electrodes is 2.3 mm. The RF voltage is applied to the electrodes such that adjacent electrodes are  $\pi$  out of phase, and diagonally opposite electrodes are in phase. The resulting oscillating electric field has a node along the trap's longitudinal axis, and increases in amplitude away from the axis. This field exerts a ponderomotive force on ions towards the trap axis. The DC end electrodes, separated by 7 cm, prevent escape of the ions along the axis.

The vacuum system had a base pressure of  $4 \times 10^{-8}$  Pa. In operation, helium was admitted on a continuous flow basis through an electrically-heated quartz leak valve, and was pumped away by a partially baffled turbo-molecular pump. The He pressure was maintained at  $2.3 \pm 0.1 \times 10^{-4}$  Pa.

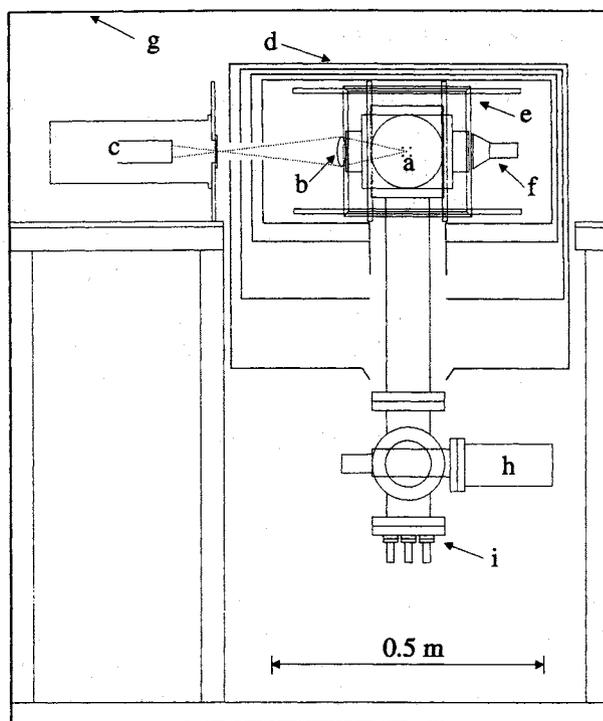


Figure 3: Schematic diagram of the ion trap system NML-IT-2. a) Trap RF electrodes, b) Fluorescence focussing lens, c) Photomultiplier tube, d) Magnetic shielding, e) Helmholtz coils, f) Microwave horn, g) Outer magnetic shield, h) Turbo-molecular pump, i) Electrical vacuum feed throughs.

The ion trap was enclosed in 5 layers of magnetic shielding (fig. 3). Helmholtz coils maintained a magnetic field of 56 mG in the region of the ion cloud. The ions were interrogated by a 12.6 GHz local oscillator (LO) microwave signal, synthesised from a superconducting cryogenic sapphire resonator (7), using Ramsey's method of separated oscillatory fields. A pair of  $\pi/2$  pulses of 1.6 s duration with pulse centres separated by 50 s yielded approximately sinusoidal Ramsey fringes separated by  $\Delta f_R = (50 \text{ s})^{-1} = 20 \text{ mHz}$ , corresponding to a Q factor of  $1.3 \times 10^{12}$ . Population of the  $F=1$  ground state hyperfine component was monitored after each Ramsey pulse sequence by detecting the 369.5 nm fluorescence signal resulting from illumination of the ions with 369.5 nm light generated by a frequency-doubled titanium:sapphire laser tuned to the  $F=1 \rightarrow F=0$  'cyclic' optical transition. At the end of each interrogation sequence the laser was tuned to resonance with the  $F=1 \rightarrow F=1$  optical transition to reset the ion population to the  $F=0$  ground state hyperfine level in preparation for the next sequence. Typically 30,000 fluorescence photons were counted per measurement sequence, plus a background signal of approximately 5000 photons.

The frequency of the LO was locked to the centre Ramsey fringe using an error signal generated by consecutive interrogation sequences at frequencies

separated from the peak of the central fringe by  $+\Delta f_R/2$  and  $-\Delta f_R/2$  (4).

### Stability Performance

The stability, characterised by the Allan deviation, of the resulting corrected clock frequency compared with a hydrogen maser (NML-Maser 2) is shown in fig. 4.

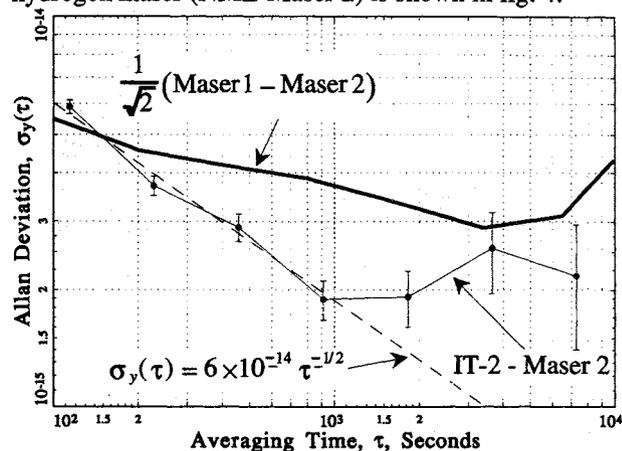


Figure 4: Fractional Allan deviation of IT-2 with respect to NML-Maser 2. Routine monitoring of the two masers over the last 12 months has shown that the performance of Maser 2 is superior to that of Maser 1.

The shot noise-limited performance of IT-2 is equal to or better than  $\sigma_y(\tau) = 6 \times 10^{-14} \tau^{-1/2}$  over averaging times  $\tau$  up to 1000 s. Beyond 1000 s the measured performance may be limited by that of Maser 2; however confirmation of this must await the completion of refurbishments to a second and similar ion trap system, IT-1.

### Clock frequency offsets

The environment within the ion trap results in a frequency offset of approximately 8 parts in  $10^{13}$  between the observed clock frequency, and the clock frequency of a hypothetical single  $^{171}\text{Yb}^+$  ion in a field-free vacuum. The offset consists of a number of contributions, as shown in fig. 5. The contributions fall into two classes; those that require a model of the spatial density dependence of the ion cloud (Magnetic inhomogeneity correction to second-order Zeeman shift, Stark shift and the second order Doppler shift) and those that do not (He pressure shift). The shifts which require a model of the ion cloud are not measured directly, but are calculated from measurements of the conditions experienced by the ion cloud. Consequently it is not easy to prove that the assumptions on which the calculations are based are correct. However, an indication of the validity of the calculated shifts may be obtained by varying the trap conditions, and comparing the observed clock frequencies, as corrected for the known

frequency offsets. In this section we briefly explain the calculation of each shift, and present the results of a preliminary investigation of the validity of the model upon which our calculation of the largest frequency offset, the second order Doppler shift, is based.

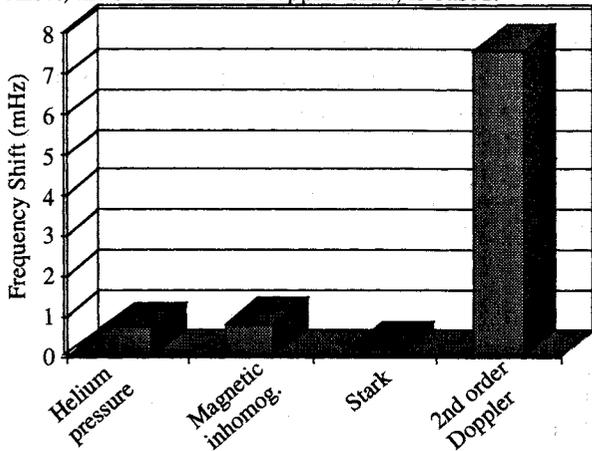


Figure 5: Magnitudes of the known frequency offsets due to the environment of the trapped ions, with the exception of the homogeneous second-order Zeeman shift (0.988 Hz). These offsets correspond to the trap conditions under which the data shown in fig. 4 was taken.

**He Pressure shift.** This is determined by extrapolating the He pressure dependence of the clock frequency to zero He pressure (4).

**Second order Doppler shift.** The second order Doppler shift  $\Delta f_D$  results from relativistic time dilation in the reference frame of each ion due to its motion with respect to the laboratory;

$$\frac{\Delta f_D}{f_0} = -\frac{v^2}{2c^2} \quad (1)$$

Here  $v$  is the velocity of the ion,  $c$  is the speed of light and  $f_0$  is the unshifted frequency of the clock transition.

There are two contributions to the second order Doppler shift, one from the thermal, or secular motion of the ions, and one from the micromotion of the ions at the trap RF frequency (510 kHz), so that in equation 1,  $v^2 = v_{\text{secular}}^2 + v_{\text{micro}}^2$ . Assuming thermal equilibrium between the three degrees of freedom of the secular motion, the secular contribution is obtained from the mean square speed of the ions;

$$\left. \frac{\Delta f}{f_0} \right|_{\text{secular}} = -\frac{3}{2} \frac{kT}{mc^2} \quad (2)$$

where  $T$  is the temperature of the ion cloud,  $k$  is Boltzmann's constant and  $m$  is the mass of a  $^{171}\text{Yb}^+$  ion.

The size of the ion cloud in the trap is determined by the mean energy of the ions, characterised by the temperature  $T$  of the cloud, and the mutual Coulomb

repulsion of the ions. For ion clouds of low density the Coulomb repulsion between the ions does not influence the cloud size, which is then determined purely by the confining potential and is independent of the ion density. The radial density profile of the ion cloud is then Gaussian (8). In this case, conservation of energy in the transverse dimensions requires that  $\langle v_{\text{Trans. micro}}^2 \rangle = \langle v_{\text{Trans. secular}}^2 \rangle$ , where the triangular brackets denote averaging over all ions in the cloud. Since the RF pseudopotential provides confinement only in the two transverse dimensions, we have

$$\langle v_{\text{micro}}^2 \rangle = \frac{2}{3} \langle v_{\text{secular}}^2 \rangle \quad (3)$$

so that

$$\left. \frac{\Delta f_D}{f_0} \right| = -\frac{3}{2} \frac{kT}{mc^2} \left( 1 + \frac{2}{3} \right). \quad (4)$$

In practice, due to signal-to-noise considerations, it is advantageous to operate a trapped ion microwave frequency standard with a relatively high density of ions such that their mutual Coulomb repulsion results in a larger cloud than would be expected from temperature considerations alone. This larger cloud results in equation 4 underestimating the second order Doppler shift, due to the larger number of ions at a greater radius from the trap axis where the amplitude, and consequently the velocity, of the micro-motion is greater. This expansion of the cloud is modelled by including the potential due to the space-charge of the ion cloud to the RF pseudopotential, to give the total confining potential. One effect of the expansion of the ion cloud is that the observed transverse secular frequency decreases as the density of ions in the trap increases (fig. 6).

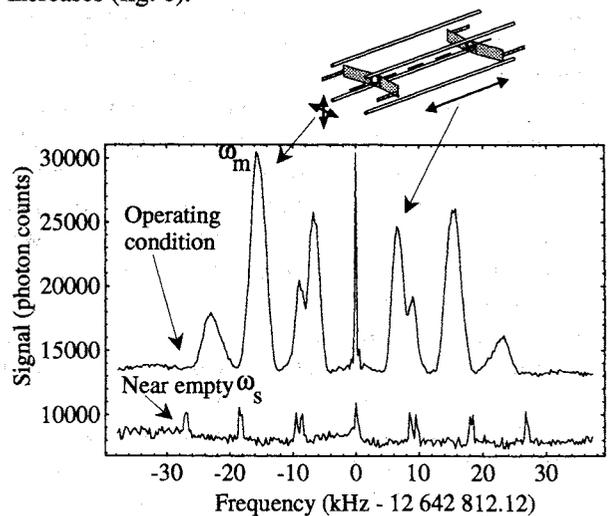


Fig. 6: Spectrum obtained by monitoring the laser-induced fluorescence at 369 nm from the trapped ions as the microwave frequency is scanned through resonance with the clock transition. Peaks due to microwave phase modulation resulting from the ions' secular motion in the longitudinal and transverse directions are marked. The other peaks visible are the sum and difference frequencies arising from nonlinear mixing.

If the confining potential, as modified by the space-charge of the ions, is assumed to remain harmonic then equation 4 becomes (8,5)

$$\frac{\Delta f_D}{f_0} = -\frac{3}{2} \frac{kT}{mc^2} \left( 1 + \frac{2}{3} \left( \frac{\omega_s}{\omega_m} \right)^2 \right), \quad (5)$$

where we now distinguish between the transverse secular frequencies  $\omega_s$  and  $\omega_m$  of the near-empty trap and loaded trap respectively.

In order to test the applicability of equation 5 to the conditions of the  $^{171}\text{Yb}^+$  frequency standard, the clock frequency was monitored as the number of ions in the trap was reduced to approximately 10% of normal. The corresponding trap-loaded secular frequencies  $\omega_m$  were recorded, and the corrections due to the second order Doppler shift were calculated. The results, plotted in fig 7, support the validity of equation 5, since the corrected clock frequencies are reasonably independent of the number of ions in the trap.

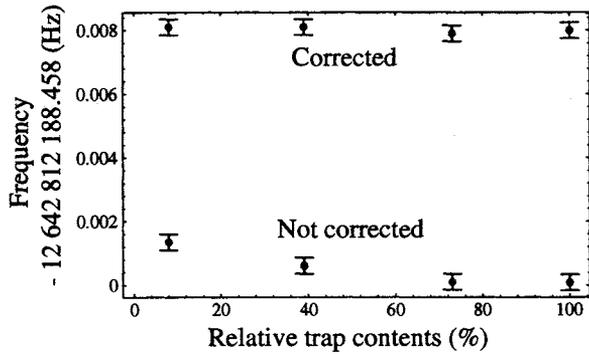


Figure 7: Dependence of the 12.6 GHz clock frequency on the number of ions in the trap, with and without correction for the second-order Doppler shift calculated according to equation 5. Here 100% contents represents the number of ions in the trap ( $\sim 10^5$ ) corresponding to the data shown in figs. 4 and 5. The end electrode voltage was 10 V. The data in this figure and figure 8 are not corrected for the magnetic inhomogeneity shift, and the correction for the frequency of the reference frequency standard vs TAI is only approximate since the relevant TAI data is not available at the time of writing.

The dependence of the clock frequency corrected according to equation 5 on the voltage applied to the end electrodes is shown in fig. 8. Increasing the voltage on the end electrodes from 1.5 V to 25 V shortens the cloud from approximately 20 mm to 6 mm (approximate cloud lengths inferred from longitudinal secular frequencies), and monitoring of the fluorescence signal from the ion trap indicates that the number of ions in the trap remains constant within approximately 10%. Thus, increasing the end electrode voltage 'squeezes' the cylindrical cloud into a shorter and fatter shape, and increases the ion density.

At end electrode voltages above 10 V, there is a marked increase in the clock frequency corrected for the second-order Doppler shift. Numerical calculations show that at end electrode voltages above 10 V, the central ion density increases such that the total radial confining potential differs significantly from the harmonic form assumed by equation 5 (6). This may be at least partly responsible for the deviation evident in fig. 8, however to take this matter further a full numerical calculation of the radial and longitudinal dependence of the ion density will be required. Until the results of that calculation are available, the present data indicates that it would be prudent to assign an extra uncertainty of  $\pm 1$  mHz to the second order Doppler shift calculated according to equation 5, for the case where the end electrode voltage is 10 V or less, and  $\pm 2$  mHz otherwise. This extra uncertainty is in addition to the effects of uncertainty in the experimental parameters inserted into equation 5.

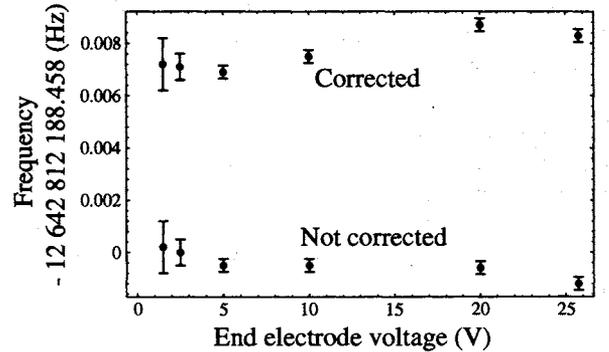


Figure 8: Dependence of the 12.6 GHz clock frequency on the voltage applied to the end electrodes, with and without correction for the second-order Doppler shift calculated according to equation 5. At low end electrode voltages the increased uncertainty in the frequencies arises from increased uncertainty in the measurement of the Larmor frequency, due to inhomogeneity of the magnetic field along the longitudinal trap axis.

**Second order Zeeman shift.** The second order Zeeman shift of the clock transition due to a magnetic field of strength  $B$  is (9)

$$\Delta f_{2oz} = -310.8 \langle B^2 \rangle \quad (6)$$

where the triangular brackets denote averaging over the ion cloud, and  $B$  is expressed in units of  $10^{-4}$  T. Measurement of the frequency difference (the Larmor frequency)  $f_L$  between the  $F=0 M_F=0 \rightarrow F=1 M_F=0$  and  $F=0 M_F=0 \rightarrow F=1 M_F=1$  transitions yields a value for  $\langle B \rangle$ ;

$$\langle B \rangle = \frac{f_L}{1.401 \times 10^6}, \quad (7)$$

where  $f_L$  is expressed in Hz. A complication arises (9) in that in general  $\langle B^2 \rangle \neq \langle B \rangle^2$ , the equality only being true for perfectly spatially homogeneous magnetic fields. This introduces a correction  $\Delta v_{M, \text{Inh}}$ , due to magnetic inhomogeneity, to the second order Zeeman shift,

$$\Delta f_{20z} = -310.8 \left( \frac{f_L}{1.401 \times 10^6} \right)^2 + \Delta f_{M.Inh}. \quad (8)$$

We refer to the first term in equation 8 as the homogeneous second order Zeeman shift. The magnetic inhomogeneity correction  $\Delta f_{M.Inh}$  is given by

$$\Delta f_{M.Inh} = 310.8 \left[ \frac{\left( \frac{\iiint_V n(x,y,z) |B(x,y,z)| dx dy dz}{\iiint_V n(x,y,z) dx dy dz} \right)^2 - \frac{\iiint_V n(x,y,z) |B(x,y,z)|^2 dx dy dz}{\iiint_V n(x,y,z) dx dy dz}}{1} \right] \quad (9)$$

where  $n(x,y,z)$  is the spatially dependent number density of the ion cloud and the integrals are taken over the volume  $V$  of the ion cloud.

The spatial dependence of the magnetic field was measured by monitoring the Larmor frequency as the ion cloud was displaced in the two transverse dimensions by DC biases applied to the trap RF electrodes. The dependence in longitudinal direction was then obtained from Maxwell's equations.

Assuming an ion cloud with a Gaussian density profile, consistent with the assumption that the space charge-modified confining potential remains harmonic as required by equation 5, yields the correction (0.77 mHz) shown in fig. 5. Violation of this assumption will result in the calculated value for  $\Delta f_{M.Inh}$  being in error, in a similar manner to the calculated value for the second order Doppler shift. However, due to the magnitude of  $\Delta f_{M.Inh}$  being small compared with the second order Doppler shift, the error so introduced has not yet been investigated in detail.

The magnetic inhomogeneity shift was inadvertently overlooked in reference (5).

**Stark shift.** The Stark shift  $f_{VS}$  due to the time-averaged RF electric field  $E_{AV}$  maintained by the trapping electrodes is given by

$$\Delta f_s = -2 \times 10^{-21} \frac{\iiint_V n(x,y,z) |E_{AV}(x,y,z)|^2 dx dy dz}{\iiint_V n(x,y,z) dx dy dz} \quad (10)$$

where  $E_{AV}$  is expressed in  $Vm^{-1}$  and the value for the numerical coefficient is given in reference (3).

As in the case of the magnetic inhomogeneity shift, the effect of deviation of  $n(x,y,z)$  from a cylinder with

Gaussian radial dependence has not yet been investigated due to the small magnitude (0.15 mHz) of  $\Delta f_s$ .

## Conclusion

The performance, as characterised by the Allan deviation, of the trapped  $^{171}Yb^+$  ion frequency standard NML-IT-2 is equal to or better than  $\sigma_y(\tau) = 6 \times 10^{-14} \tau^{-1/2}$ , for averaging times less than 1000 s. There is evidence that the measured performance is limited both in the short and medium ( $\tau > 1000$  s) term by the stability of the hydrogen maser used as a frequency reference.

A preliminary investigation of the dependence of the observed clock frequency corrected for the calculated second order Doppler shift on the trap conditions has yielded an approximate set of confidence limits on the calculated values of the second order Doppler shift.

## Acknowledgment

The authors are grateful to Dr A.G. Mann and Assoc. Prof. D.G. Blair, of the Department of Physics, University of Western Australia, for the loan of the sapphire-loaded superconducting resonator on which our microwave system is based.

## References

1. R. Casdorff, V. Enders, R. Blatt, W. Neuhauser, and P.E. Toschek, 1991, "A 12 GHz standard clock on trapped ytterbium ions", *Ann. Physik*, **48**, 41-55
2. D.J. Seidel and L. Maleki, 1991, "Progress toward the development of a ytterbium ion standard", *In Proc. 48th Annual IEEE International Symposium on Frequency Control*, 761-768.
3. C. Tamm, D. Schnier, and A. Bauch, 1995, "Radio-frequency laser double-resonance spectroscopy of trapped  $^{171}Yb$  ions and determination of line shifts of the ground-state hyperfine resonance", *Appl. Phys. B*, **60**, 19-29
4. P.T.H. Fisk, M.J. Sellars, M.A. Lawn, and C. Coles, 1995, "Performance of a microwave frequency standard based on laser-detected, trapped  $^{171}Yb^+$  ions", *Appl. Phys. B*, **60**, 519-527
5. , M.J. Sellars, P.T.H. Fisk, M.A. Lawn, and C. Coles, 1995, "Further investigation of a prototype microwave frequency standard based on trapped  $^{171}Yb^+$  ions", *Proc. 49th Annual IEEE International Symposium on Frequency Control*, 66-73

6. P.T.H. Fisk, M.J. Sellars, M.A. Lawn, and C. Coles, 1996, Proc. 5th Symposium on Frequency Standards and Metrology, (in press).
7. A.J. Giles, A.G. Mann, S.K. Jones, D.G. Blair, and M.J. Buckingham, 1990, "A very high stability sapphire-loaded superconducting cavity oscillator", Physica B, 165-166, 145-146
8. C. Meis, M. Desaintfuscien, and M. Jardino, 1988, "Analytical calculation of the space charge potential and the temperature of stored ions in an RF quadrupole trap," Appl. Phys. B, 45, 59-64
9. J. Vanier and C. Audoin, 1989, "The Quantum Physics of Atomic Frequency Standards", Adam Hilger

## THE LPTF PRELIMINARY ACCURACY EVALUATION OF CESIUM FOUNTAIN FREQUENCY STANDARD.

A.Clairon, S.Ghezali, G.Santarelli, Ph.Laurent, E.Simon, S.Lea\*, M.Bahoura, S.Weyers and K.Szymaniec.

BNM /LPTF, Observatoire de Paris, 61 Avenue de l'Observatoire, F-75014 Paris, France..

\* : N.P.L, Teddington. Middlesex TW11 0LW, UK

### ABSTRACT

Most of the frequency shifts which limit the accuracy of cesium atomic frequency standards are considerably reduced when using ultra-cold atoms and long interrogation times in a Ramsey geometry. These shifts are also more easily evaluated than in traditional cesium beam standards. Most of these shifts are reduced to  $10^{-16}$  or less and can be calculated accurately such as relativistic shifts, cavity pulling, Rabi pulling... Other effects have to be directly measured, for example cold collisions, residual first order Doppler shift, quadratic Zeeman shift...

At present, the accuracy of the fountain is estimated at  $3 \cdot 10^{-15}$ . A second fountain, using velocity selective Raman transitions in two dimensions, is being built leading to a potential improvement of a factor 10 of the fountain's accuracy evaluation.

### 1. Introduction.

The feasibility of Zacharias' proposal of an atomic fountain was demonstrated a few years ago using laser cooled sodium<sup>1</sup> and cesium<sup>2</sup> atoms. Cesium atoms can now be cooled to a few microKelvin ( $v_{RMS}=1.2\text{cm/s}$ ) by sub-Doppler cooling mechanisms in a low-pressure cell, near room temperature using only commercially available laser diodes. In a fountain geometry, the atoms follow a parabolic flight, passing twice through a single microwave cavity. The long time between the two interactions leads to a sub-Hertz Ramsey resonance, 100 times narrower than the Ramsey fringes obtained in conventional Cs frequency standards. The nearly monokinetic velocity distribution of the atoms and the pulsed character of our fountain simplify considerably the evaluation of its accuracy and allow us to perform tests and measurements that could not be realized with a continuous atomic fountain.

### 2. Experimental apparatus.

The scheme of our pulsed fountain prototype is shown in Fig.1, this apparatus has been described in detail in reference 3.

Six independent laser beams ( $\phi = 2\text{cm}$ ), tuned to the red of the cyclic  $F=4-F'=5$  transition capture and cool atoms in a low pressure Cs vapor cell ( $\approx 10^{-6}$  Pa). A few  $10^8$  atoms can be cooled in 0.3s when using a magneto-optical trap (MOT) and typically 10 times fewer when using  $\text{Lin} \perp \text{Lin}$

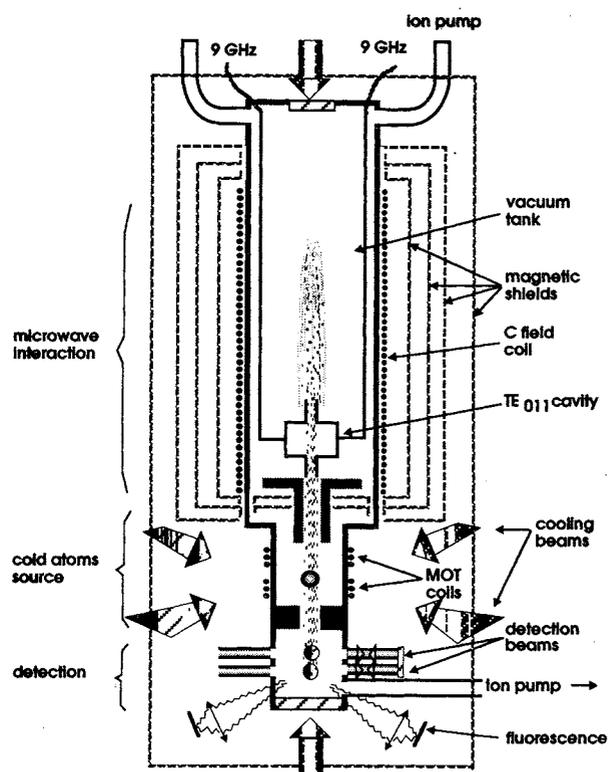


Fig.1: The atomic fountain.

optical molasses. The moving molasses method is used to launch atoms. In the C-field region, a highly homogeneous static magnetic field is produced by a solenoid. Three  $\mu$ -metal magnetic shields surround the C-field region. Additional compensation coils between the magnetic shields are used to improve

the magnetic field continuity along the atoms' path. The whole setup is surrounded by a fourth magnetic shield.

After the launching, the atoms pass through a cylindrical  $TE_{011}$  cavity with a loaded Q factor of  $10^4$ . The cavity is situated 33cm above the cooling region. This new cavity has been designed to reduce microwave leakage.

Less than 10% of the atoms pass two times through the microwave cavity because of the thermal expansion of the atomic cloud. The populations of both hyperfine levels  $F=3$  and  $F=4$  are independently measured by fluorescence. About 100 photons per atom are detected by a large area low noise photodiode, during the transit time of the atoms through the probe beams. The TOF signals are sampled, digitized and analyzed by a personal computer. The microwave resonance is then obtained by dividing the  $F=4$  level population by the total atom number ( $F=3+F=4$ ). This normalization procedure reduces the signal fluctuations due to shot-to-shot atom number fluctuations which are typically 1% r.m.s. When the atoms are prepared in a  $m_F=0$  state, this measurement is equivalent to a measurement of the microwave transition probability. Uncorrelated fluctuations from the two detection channels come mainly from the detection noise and from the fluorescence light of the weak diffusive atomic beam coming from the low pressure Cs loading zone. These two fluctuations limit the signal-to-noise ratio to 700.

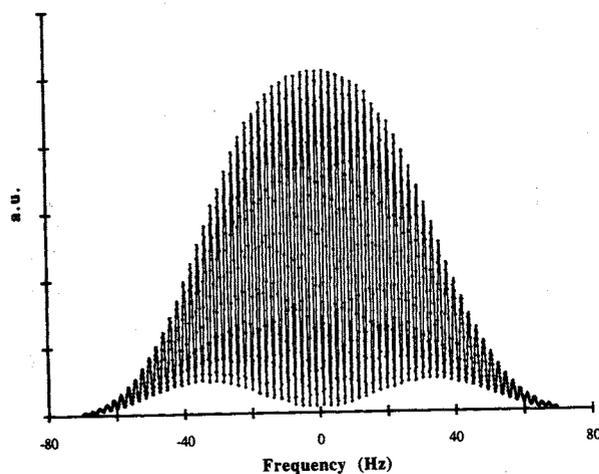


Fig.2: The Ramsey fringes

In order to compare the central fringe to the H-maser frequency or to record the Ramsey resonance pattern, a low noise multiplication chain using a BVA quartz oscillator is weakly locked to the H-maser<sup>4</sup>. Fig 2 shows the Ramsey fringes which are

obtained with a flight time of 500ms above the microwave cavity. The near monokinetic velocity distribution of the atoms is responsible for the large number of fringes. The width of the narrowest achievable fringe in our system is 700mHz.

### 3. Frequency stability.

When the peak to peak frequency modulation of the interrogation signal is equal to the fringe width, the two sample Allan fractional frequency stability is given by:

$$\sigma_y(\tau) = \frac{\delta\nu}{\pi\nu_0} \sigma_p \sqrt{\frac{T_C}{\tau}}, \quad \tau \gg T_C, \quad (1)$$

where  $\nu_0$  is the clock frequency,  $\delta\nu$  is the width of the Ramsey fringe,  $T_C$  is the cycle time,  $\tau$  is the integration time and  $\sigma_p$  is the standard deviation between successive measurements of the transition probability. For a large number of detected photons per atom, the measurement is in principle limited by the atomic projection noise<sup>5,6</sup>. In this case, we have  $\sigma_p = 1/\sqrt{2N_{at}}$  where  $N_{at}$  is the number of detected atoms. At present when we operate the fountain using molasses, we have about  $10^5$  atoms detected in  $m_F=0$ , indicating that a stability of  $8 \cdot 10^{-14} \tau^{-1/2}$  could be obtained. We find that  $\sigma_p \sim 2 \cdot 10^{-3}$  for  $10^5$  atoms detected in  $m_F=0$ . Taking into account these two noise contributions should lead to a fractional frequency stability of  $1 \cdot 10^{-13} \tau^{-1/2}$ , whereas we only measured  $2 \cdot 10^{-13} \tau^{-1/2}$  (even for a large number of atoms).

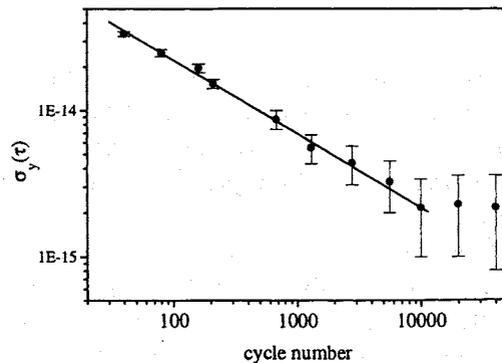


Fig.3: The fountain fractional frequency stability.

As pointed out by G. J. Dick<sup>7</sup>, the high frequency noise of the local oscillator is down converted by an aliasing phenomenon under pulsed operation. This degradation is directly related to the unavoidable

dead time between two successive cycles, due to the loading time, the detection, etc. Typically, we have 500ms of ballistic flight above the microwave cavity and 1.2s for the cycle time. An improvement of a factor two can probably be achieved by using the best SC-cut BVA quartz oscillators. Another way to reduce the contribution of the local oscillator frequency noise to the fountain standard stability is to use a multi-pulse fountain as proposed by S.Ohschima et al<sup>8</sup>.

Nevertheless, the stability obtained is three times better than the best obtained with atomic beam standards. For integration time longer than 10<sup>4</sup>s, the measured stability is limited by the H-maser reference.

#### 4. Frequency shifts.

When using cold atoms and long interrogation times, most of the frequency shifts which limit the accuracy of conventional Cs frequency standards are considerably reduced. Here, we will discuss the evaluation of some of them. For instance, the second order Doppler shift can be determined to within 10<sup>-18</sup> and the gravitational red shift is known to 10<sup>-16</sup> if the altitude of the standard with respect to the reference geoid is determined with 1m accuracy.

Cavity pulling, which depends on the width of the central Ramsey fringe, is considerably reduced. With a low atom number (10<sup>6</sup>) and a microwave cavity with a loaded quality factor  $Q_{cav}$  of 10<sup>4</sup> we are far from maser oscillation and the shift can be calculated as in a passive cesium beam standard:

$$\frac{\Delta\nu}{\nu_0} = \frac{\Delta\nu_{cav}}{\nu_0} (Q_{cav}/Q_{at})^2, \quad (2)$$

where  $\Delta\nu_{cav}$  is the cavity detuning and  $Q_{at}$  the atomic quality factor. With a detuning of 100kHz (which can be easily obtained by controlling the cavity temperature to 0.5K) and  $Q_{at}=10^{10}$ , we have a shift of only 10<sup>-17</sup>.

With a fountain geometry, the residual first order Doppler effect is only due to the transverse phase distribution of the microwave field resulting from the losses in the copper walls. When using monokinetic atoms, this effect becomes independent of the microwave power. This effect can be calculated as:

$$\frac{\Delta\nu}{\nu_0} = \frac{\langle \Phi_1 - \Phi_2 \rangle}{\pi Q_{at}}, \quad (3)$$

where  $\Phi_1$  and  $\Phi_2$  are the phases seen by the atoms during the two microwave interactions. For the second passage, the standard deviation  $\sigma_f$  of the atomic cloud distribution is much larger than the cavity aperture radius ( $R=5\text{mm}$ ). Using a 2D model, A. De Marchi and coworkers have calculated the phase distribution of the microwave field in a high Q TE<sub>011</sub> copper cavity<sup>9</sup>. Using these results and allowing for a possible 1mrd vertical misalignment of the launching direction, we have calculated an average phase difference of 4  $\mu\text{rd}$  leading to a first order Doppler shift of about 10<sup>-16</sup>. This result is probably too optimistic because the phase distribution also depends on the cavity end caps losses and on the effective conductance of the copper walls. From the same work, it seems possible to reduce the phase gradient by feeding the microwave cavity symmetrically. In our experimental setup, both microwave couplings can be used.

At present, we have only done some preliminary tests to estimate the maximum shift due to the residual first order Doppler effect. To increase the sensitivity to the phase distribution, these measurements have been performed using the MOT as a cold atom source. In this case, the atom distribution during the first pass through the cavity is much narrower than  $R$  ( $\sigma_i=2\text{mm}$ ). We verified that when the cavity is symmetrically or asymmetrically fed, the fountain frequency does not vary by more than  $\pm 2.10^{-15}$  (which is the resolution of the measurement due to the H-maser frequency instability).

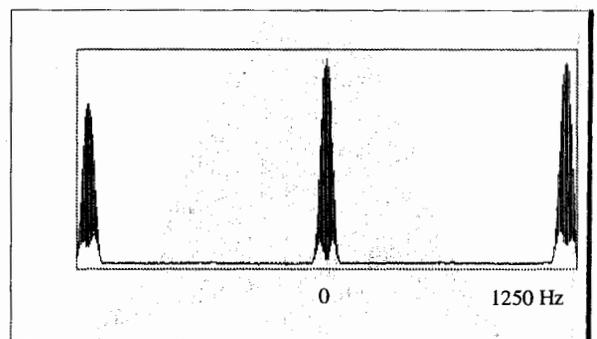


Fig. 4: The partial microwave spectrum

A second, more important, test is performed by partially blocking the detection beams in order to preferentially detect atoms that have crossed the cavity (during their second pass) in a given place. The maximum frequency shift observed in this

experiment is  $\sim 3\text{-}4 \cdot 10^{-15}$  when the central part of the detection beam is blocked so that only 50% of the atoms are detected.

From these measurements, we can safely deduce that the maximum shift is less than  $10^{-15}$ . A much better frequency resolution is required to really obtain a map of the transverse phase distribution of the microwave field.

As a consequence of the narrow linewidth, the C-field in a fountain can be small compared to that in a thermal beam device, leading to a reduction of the sensitivity to magnetic field fluctuations and inhomogeneities. A field of 1.7mG (40 times less than in conventional Cs atomic beam standards) is largely sufficient to completely resolve the field sensitive transitions (see fig.4). In the same figure, we can see that the contrast of the fringes for magnetic field sensitive transitions is better than 80%. This demonstrates that the average magnetic field for different atomic trajectories differs by  $\sim 1 \mu\text{G}$ .

A further advantage of the pulsed fountain is the possibility of mapping the static magnetic field by using the atoms as a probe. This can be done by measuring the frequency of the first order field sensitive transitions for different launching heights (when using only this method, it is sometimes difficult to identify confidently the central fringe). This method can be accompanied by an unambiguous technique to measure the local magnetic field. A microwave pulse of  $\sim 10\text{ms}$  duration is applied (with an antenna) when the atoms are at their apogee. Then, the frequency difference between the  $F=3, m_F=0 \Rightarrow F=4, m_F=0$  transition and a field sensitive transition is measured. A resolution of about  $0.1 \mu\text{G}$  is obtained when using this method (the present resolution is only limited by magnetic field fluctuations). By launching to different heights, we can record the magnetic map  $B(z)$  with a very high resolution. The Ramsey fringes have been recalculated by averaging the transition probability over several atomic trajectories because of transverse inhomogeneities estimated at about  $1.2 \mu\text{G}$ . Thus, we can evaluate

the quantity  $\overline{B^2} = \frac{1}{T} \int_0^T B^2(t) dt$  where  $T$  is the time of

flight above the cavity. The maximum uncertainty in determining this quantity is  $2\overline{B}\sigma_B$  where  $\sigma_B$  is the resolution of the measurement. Thus, we could in principle obtain an uncertainty of  $\sim 10^{-16}$ .

Actually, we have observed slow fluctuations of the magnetic field homogeneity, particularly when the MOT is used. To be cautious, an uncertainty of  $5 \cdot 10^{-16}$  is applied to the Zeeman correction. Of

course, this result can be improved 2 or 3 fold by simply reducing the field value.

The black body shift, due to the AC Stark shift induced by the thermal radiation, is  $1.76 \cdot 10^{-14}$  in our device. Because of the uncertainty (few %) of the DC Stark shift determination which is used in the calculation of ref.9 and our imperfect knowledge of the effective radiation temperature in the microwave region, we consider that the total uncertainty in this correction is  $1.5 \cdot 10^{-15}$ .

Currently, the spin exchange collisional shift is expected to be one of the limiting factors for the accuracy of the pulsed atomic fountain. Calculations made by B. J. Verhaar and coworkers estimate this shift at  $10^{-22} \text{atoms/cm}^3$ . For s scattering waves, the shift is independent of the temperature of the atoms. For a tall fountain and a nearly point-like cold atom source (MOT), this shift can be approximately written as:

$$\frac{\Delta v_c}{v} = \frac{-KN_i}{2(2\pi)^{3/2} \sigma_i^2 \sigma_f} \propto -K\bar{n}, \quad (4)$$

where  $\bar{n}$  is the average density if  $\sigma_f \gg \sigma_i$  and  $\sigma_f \gg R$ ,  $N_i$  is the initial atom number,  $\sigma_i$  and  $\sigma_f$  are respectively the initial and final rms atom cloud sizes in the cavity and  $R$  is the radius of the holes in the cavity. With our geometry, when using the MOT, only s scattering waves are expected to contribute. The average density  $\bar{n}$  that can be obtained by using pure optical molasses is too low to allow an accurate measurement of this shift; thus, we used the MOT as a cold atom source. The density is varied by changing the loading time of the trap and by changing the Cs pressure in the cell.

Atoms are prepared in  $F=3$  (the Zeeman sublevels being nearly equally populated) by a pulse of light resonant with the  $F=4 \Rightarrow F'=4$  transition. To prepare atoms in the  $m_F=0$  sublevel, just after the launching phase, we apply a microwave  $\pi$  pulse resonant with the  $F=4, m_F=0 \Rightarrow F=3, m_F=0$  transition; then, the atoms remaining in the  $F=4$  level are pushed away as in the previous case. In this way at least 97% of the atoms are in the  $m_F=0$  sublevel (see fig.5). The absolute uncertainty of these measurements is not better than 50% due to the density determination. Nevertheless, an extrapolation to zero density, with 10% uncertainty can be performed since only relative uncertainties are relevant in this case. In typical molasses operation, we have an average density (for atoms in  $m_F=0$  only) of  $\sim 5 \cdot 10^5 \text{atoms/cm}^3$  for  $\sim 10^5$  detected atoms. The collisional frequency shift is then only

$\sim 10^{-15}$ , for which we carefully consider a  $5 \cdot 10^{-16}$  uncertainty. With this number of atoms and a very good local oscillator, it will be possible to obtain a short term frequency stability better than  $10^{-13} \tau^{-1/2}$  and a long term stability of  $10^{-16}$  limited by fluctuations (10%) of the atomic density. By modifying the geometry of the molasses or the trap, a significant reduction of this shift could be obtained. Another way to reduce the collisional shift is to use 2D transverse Raman cooling or selection<sup>13</sup>.

We have also evaluated the frequency shift due to a spurious line in the microwave spectrum. From the measured spectrum of the microwave generator, we conclude that the maximum frequency shift due to spectral impurities is  $2.5 \cdot 10^{-16}$ .

The presence of other transitions near the clock transition or nonadiabatic transitions between Zeeman sublevels can induce frequency shifts. They are called Rabi pulling, Ramsey pulling and pulling due to Majorana transitions. All these shifts depend on the line Q and on the ratio Rabi frequency/Zeman frequency, these two factors are more favorable in our fountain than in conventional Cs beams. In addition, when the atoms are prepared only in the  $m_F=0$  level, the populations of the  $m_F=\pm 1$  sublevels are very small ( $\sim 0.4\%$ , see Fig.5). In the same figure, we can also see that  $\Delta m_F=1$  transitions are also very weak ( $\leq 0.5\%$ ) at optimum power. With these conditions, the Rabi pulling is negligible ( $10^{-18}$ ). The Ramsey pulling and the pulling due to Majorana transitions are more difficult to evaluate; a very preliminary estimate gives a maximum shift of  $\sim 5 \cdot 10^{-16}$  for both effects when atoms are prepared in the  $m_F=0$  sublevel.

As a consequence, we can operate the fountain with very high microwave power, without large loss of signal. Since these two shifts are functions of the microwave power, we can try to evaluate their order of magnitude by measuring the frequency variation

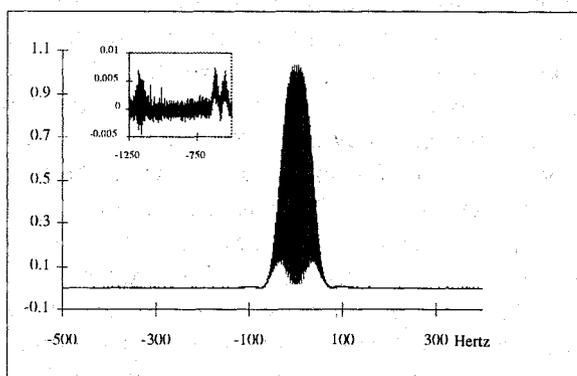


Fig.5: Microwave spectrum when atoms are prepared in  $m_F=0$ .

of the standard when the microwave power is changed. The maximum measured power shift is  $4 \cdot 10^{-16} P/P_{\text{opt}}$  (where  $P_{\text{opt}}$  is the microwave power to obtain a  $\pi/2$  pulse). Of course, other shifts also depend on the microwave power level: microwave leakage, spectral impurities, etc. Most of these can be independently evaluated. For instance, to evaluate the shift due to the microwave leakage, from the cavity or from the feeding circuitry, we can take advantage of the pulsed operation of the fountain to feed the cavity with a very high power (90dB more than the optimum power) when atoms are outside the cavity (no shift is observed in this case with our present resolution).

The summary of these preliminary evaluations is shown in table I. The present uncertainty of our fountain standard,  $\sim 3 \cdot 10^{-15}$ , is much better than that obtained with traditional Cs beam tubes. This evaluation is mainly limited by hydrogen maser frequency instabilities.

## 5. Conclusion.

Our recent results, obtained with the laser cooled fountain standard, demonstrate that this new type of device is working reliably, allowing very accurate measurements to be made. The frequency stability is  $2 \cdot 10^{-13} \tau^{-1/2}$ , limited by the frequency noise of the local oscillator. It is at least 3 times better than the stability obtained with the best conventional cesium standard. A long term stability of  $10^{-16}$ , limited by atomic density and magnetic field fluctuations, seems realistic. Our preliminary accuracy,  $3 \cdot 10^{-15}$ , is also much better than the accuracy of any other frequency standard. A second fountain, which is being built, will substantially improve the resolution of our measurements and future accuracy evaluations. At present, it seems difficult to predict which physical effect will be the limiting one for a room temperature fountain. During the evaluation, we have compared the fountain frequency with the frequency of some Cs clocks determining the TAI frequency. No effective frequency difference larger than the uncertainties has been observed.

Table I: Error budget in molasses operation with atoms prepared in  $m_F=0$ .

Physical effect:	Uncertainty $\times 10^{15}$ .
First order Doppler effect:	$\leq 1$
Black body radiation effect:	$\leq 1.5$
Magnetic field inhomogeneities:	$\leq 0.5$
Electronics:	$\leq 1$
Cold collisions:	$\leq 0.5$
Pulling due to other levels:	$\leq 0.5$
Microwave spectral impurities:	$\leq 0.25$
Microwave leakage:	$\leq 0.2$
Background gases:	$\leq 0.5$
Total shift	$\leq 3$

## 6. Acknowledgments.

This work forms a part of the program of scientific developments of the Bureau National de Metrologie (BNM). The assistance of M. Lours, A.H. Gerard and M. Dequin, is gratefully acknowledged. The research programme has the financial and logistic support of the French Space Agency (CNES) via the Pharaon project. We also acknowledge the financial participation of the Conseil Général de la Région Ile de France through the Sesame programme.

## 7. References.

- 1.M.Kasevich, E.Riis, S.Chu and R. De Voe, *Phys. Rev. Lett.*, **Vol.63**, 1989, pp.612-615.
- 2.A.Clairon, C.Salomon, S.Guellati and W.Phillips, *EuroPhys. Lett.*, **Vol.16**, Sept.1990, pp.165-170.
- 3.A.Clairon et al., *IEEE trans Instru. Meas.*, **Vol.44**, April 1995, pp.128-131.
- 4.G.D. Rovera, G. Santarelli and A. Clairon, *IEEE Trans. Ultr. Ferr. Freq. Contr.*, **Vol.43**, May 1996.

- 5.D.J.Wineland, W.M.Itano, J.C.Berquist and F.L.Walls, in *Proc. 35<sup>th</sup> Annu. Freq. Contr. Symp.*, IEEE, 1987, pp.602-611.
- 6.D.J.Wineland et al., in Enrico Fermi Summer School on Laser manipulation of atoms and ions, (E.Arimondo and W.D.Phillips eds.) 1991, also in NIST Tech Notes 1342, pp.TN182-TN192, 1992.
- 7.G.J.Dick, in *Proc.of P.T.T.I*, Redondo Beach, 1987, pp.133-147.
- 8.S.I.Ohshima, T.Kuroso, T.Ikegami and Y.Nakadan, *Proc. 5<sup>th</sup> Symp. Freq. Stand. Metro.*, Woods Hole, MA, USA, Oct 15-19, 1995.
- 9.A.Khursheed, G.Vecchi and A. De Marchi, *IEEE Trans.Ultr.Ferr.Freq.Contr.*, **Vol.43**, March 1996, pp.201-210.
- 10.W.D.Itano, L.Lewis and D.Wineland, *Phys.Rev.A*, **Vol.45**, Fev.1982, pp.1233-1235.
- 11.E.Tiesinga et al, *Phys.Rev.A*, **Vol.45**, March 1992, pp.R2671-R2673.
- 12.K.Gibble and S.Chu, *Phys.Rev.Lett.*, **Vol.70**, pp.1771-1774, March 1993.
- 13.S.Ghezali, Ph.Laurent, G.Santarelli, S.N.Lea and A.Clairon, to be submitted to *EuroPhys. Lett.*
- 14.S.N.Lea, A.Clairon, Ch.Salomon, Ph.Laurent, B.Lounis, J.Reichel, A.Nadir and G.Santarelli, *Physica Scripta*, **Vol.T51**, June 1994, pp.78-84.

# CONTINUOUS EXTRACTION MECHANISMS FROM A Cs MAGNETO-OPTICAL TRAP

N Sagna, G Duddle, P Berthoud and P Thomann

Observatoire Cantonal, Rue de l'Observatoire 58, CH-2000 Neuchâtel

## Abstract

Atomic clocks based on a continuous beam of cold atoms are a promising alternative to fountains mainly because the low atomic density of a beam would reduce collisional shifts to a negligible level; in addition dead times can be suppressed or greatly reduced. Several mechanisms allowing to continuously extract cold atoms from an anisotropic trap will be discussed, in particular atomic drift induced by a constant magnetic field and laser beam imbalance. Experimental results of trap loading times and time-of-flight measurements will be presented to assess the potential of these methods for the production of a continuous beam of cold atoms.

## 1 INTRODUCTION

### Motivation for a continuous beam of cold atoms

Cold atoms may lead to considerable improvements in both the accuracy and short term stability of atomic frequency standards. One specific requirement of frequency standards is that cold atoms must be carried into - and out of - the microwave resonator, whose geometry does not allow on-site cooling or on-site detection of the resonance. In presently realized Cesium standards, the cooling and trapping laser beams are periodically shut off to allow launching of the cold sample in a vertical ballistic trajectory (atomic fountain). Although this method already has shown its potential for accuracy [1] other approaches are worth investigating: one alternative approach would consist in extracting atoms continuously from the trap. This would basically have two advantages over a pulsed extraction: first, assuming the same average atomic flux on the detector, the average atomic density would be reduced by a factor equal to the inverse of the duty cycle in pulsed operation, typically 100. This would reduce the collisional shift [1, 2] in the same proportion. Second, the continuous atomic beam would suppress the short-term stability limitation [3] which results from the intermittent generation of an error signal in the servo-loop that controls the local oscillator. Another way of drastically reducing both the collisional shift and the short-term stability limitation would be a quasi-continuous operation where the cold atom source is pulsed at a

frequency such that at least one cold cloud is in the microwave resonator at any given time [4]. A continuous source of cold atoms would also be an extremely valuable tool for the development of compact atomic frequency standards as well as in other fields such as atom optics, atomic interferometry, collision studies and high resolution spectroscopy.

**Light shift** The light shift is an obvious drawback of a continuous or quasi-continuous beam scheme since laser light scattered from the trap and from imperfections of the optical elements can reach the microwave cavity while atoms are being interrogated. The light shift sensitivity of the  $|F = 4, m_F = 0\rangle$  state of Cs for  $\sigma$ -polarized light tuned 2.5 natural widths (13 MHz) to the red of the  $|F = 4\rangle \rightarrow |F' = 5\rangle$  transition is 7.6 kHz  $m^2/W$ . A dark channel [5] capable of producing a useful flux of  $10^6$  atoms/s is estimated to produce a fluorescence intensity in the  $10^{-4}$  W/ $m^2$  range in a microwave cavity, placed 0.25 m away from the trap, resulting in a potential light shift of the order of  $10^{-10}$ . The  $10^5$  reduction needed to bring the bias down to  $10^{-15}$  range can be obtained by placing a velocity selector (cf. [6]) on the way between source and cavity and between cavity and detection zone; the selector will absorb most scattered photons and deflect the remaining part into a light trap, while introducing only a weak and fast modulation of the atomic beam.

At least four mechanisms can be used to extract atoms continuously from an anisotropic trap: a static magnetic field, intensity imbalance, frequency shift and dark channel structures [5]. The purpose of this paper is the study of two of these mechanisms - a static magnetic field and an intensity imbalance - both along the vertical  $z$ -axis.

## 2 MEASURING METHODS

An important feature of our magneto-optical trap is the structure for the magnetic gradient. It consists of two ensembles of four straight conductors respectively parallel to the  $z$ -axis and the  $y$ -axis producing each a quadrupole field. With this set-up, isotropic and anisotropic traps can be studied [7]. The trap is irradiated by two laser beams in the horizontal plane and four laser beams at  $45^\circ$  to the vertical  $z$ -

axis. All beams are circularly polarized according to the requirements of a magneto-optical trap. A calibrated photodiode monitors the number of atoms in the trap. An additional laser beam locked to the  $4 \rightarrow 5'$  transition analyses the falling atoms 29 cm below the trap. Two coils in Helmholtz position allow to add a constant magnetic field in the  $z$ -direction. For the intensity imbalance we change the distribution between the up- and downgoing beams. The imbalance parameter  $w$  is defined by:

$$w = \frac{I_{\text{up}} - I_{\text{down}}}{I_{\text{up}} + I_{\text{down}}} \quad (1)$$

We study the drift velocities of cold atoms in both an isotropic 3D molasses and an anisotropic 2D trap.

**Time of flight technique** To investigate the influence of a static magnetic field or an intensity imbalance between up- and downgoing beams on a 3D molasses we use the *time of flight* technique (TOF): in a first step atoms are captured in an isotropic magneto-optical trap;  $10^7 - 10^8$  atoms are charged in a loading time of about 1 s. In a second step the gradients of the trap are switched off (10 ms) and the atoms thermalize in the optical molasses for an additional 5 ms. If a static magnetic field or an intensity imbalance is present along the vertical  $z$ -axis, the atoms thermalize around a velocity  $v_d \neq 0$ . Finally, we cut off the cooling laser by a mechanical shutter. The falling time of the laser intensity has been measured to be less than  $100 \mu\text{s}$ . By measuring the TOF signal  $dN/dt$  one can calculate the initial distribution of velocity  $dN/dv_0$  along  $Oz$ :

$$\begin{aligned} \frac{dN}{dv_0} &= \frac{dN}{dt} \frac{dt}{dv_0} \\ &= \frac{1}{g} \frac{dN}{dt} \end{aligned} \quad (2)$$

In this equation we use the fact that in the range of drift velocity we consider, the initial velocity  $v_0$  of an atom is related to the arrival time  $t$  in the probe beam by the relation:

$$v_0 = g \times (t - t_0) \quad (3)$$

where  $t_0 = 242 \text{ ms}$  is the time of free fall for atoms when  $v_0 = 0$ . The arrival time changes by 1 ms for an initial velocity of 1 cm/s. We define an average drift velocity of the atoms by:

$$v_d = \frac{\int_{-\infty}^{+\infty} \frac{dN}{dv_0} v_0 dv_0}{\int_{-\infty}^{+\infty} \frac{dN}{dv_0} dv_0} \quad (4)$$

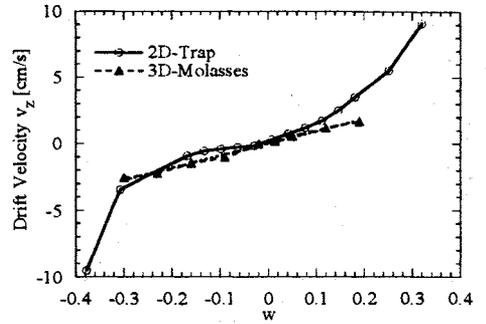


Figure 1: The drift velocity *vs.* intensity imbalance between up- and down-going beams. White circles represent the results of a 2D-trap (loading time measurement), black triangles stand for the TOF measurements in a 3D-molasses. The laser detuning and intensity are given by:  $\delta = -3\Gamma$  and  $I = 2I_s$ .

**Loading time** In the anisotropic situation only the vertical structure of the magnetic gradients is supplied, leading to cooling *and* trapping forces in the horizontal ( $xy$ ) plane and an optical molasses (without trapping forces) in the  $z$ -direction. While in a trap without extraction mechanisms the loss rate  $\Gamma$  is governed by the collision rate  $\Gamma_c$ ,  $\Gamma$  is modified if the atoms of the trap have a drift velocity  $v_d$ :

$$\Gamma = \frac{\Gamma_c}{1 - \frac{v_d}{\Gamma_c L} \left( 1 - \exp \left[ -\frac{\Gamma_c L}{v_d} \right] \right)} \quad (5)$$

where  $L$  is the length of the anisotropic cold cloud. By measuring the loading time one can deduce the average drift velocity  $v_d$ . We obtain the loading time  $\tau = 1/\Gamma$  from a fitting of the signal of the photodiode with  $(1 - \exp[-t/\tau])$ .

### 3 EXPERIMENTAL RESULTS

**Intensity imbalance** We measured the velocity of the drift induced by an intensity imbalance between up- and downgoing beams, in a 2-D trap and a 3-D optical molasses. We ensured that the static magnetic field was zero in the trapping zone. The results (see Fig.1) show a good agreement between these two velocities. Let us point out that this agreement is quite remarkable: it shows that as far as an intensity imbalance along  $Oz$  is concerned, the  $z$ -axis - free of trapping - of the 2-D trap is equivalent to that of a 3-D optical molasses. The presence of the transverse magnetic gradient does not alter the effect of an intensity imbalance along  $Oz$ . Besides, this figure shows that the drift velocity is very small ( $v_d \leq 2 \text{ cm/s}$ ) for a wide range of imbalance. Such an

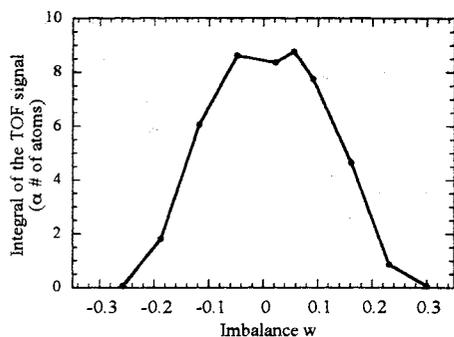


Figure 2: The integral of the TOF signal of a 3-D optical molasses *vs.* intensity imbalance between up- and down-going beams. The laser detuning and intensity are given by:  $\delta = -3\Gamma$  and  $I = 2I_s$ .

insensitivity to imbalance was known to be characteristic of sub-Doppler optical molasses [8]. Our measurements show that the transverse magnetic field in the 2-D trap does not change the drift velocity induced by an intensity imbalance. In addition to that, the drift induced in a 2-D trap exhibits a double slope that could be interpreted as Doppler and sub-Doppler behaviour. This double slope could not be observed in a 3-D molasses because of the important decrease of the TOF signals that we observed for  $|w| > 0.3$ .

To evaluate the potential in extracting atoms with an intensity imbalance, we represent in Fig. 2 the total number of atoms that cross the probe beam, when they are pushed away from the 3-D molasses. This number corresponds to the integral of the TOF signal. One can see that the total number decreases very quickly with the imbalance  $w$ . By measuring the number of trapped atoms, we found that the loss of signal is almost entirely due to the decrease of the capture process efficiency induced by the intensity imbalance. Besides, the TOF signals do not show any broadening and are only slightly distorted for the extreme values of  $w$ .

**Magnetic field** We also measured the drift velocity along the  $z$ -axis as a function of the static magnetic field. We found that the 2-D trap and the molasses show different behaviours. For the molasses, the TOF signals show that the drift velocity varies linearly with the magnetic field ( $v_d/B_z \cong 18 \text{ cm s}^{-1} \text{ G}^{-1}$ ). However, as shown in Fig. 3, the TOF signals are not only shifted along the time axis, they are also distorted and reduced in amplitude (see Fig. 4). Contrary to the case of the intensity imbalance, the loss of signal shown in Fig. 4 can not be attributed to a decrease of the capture

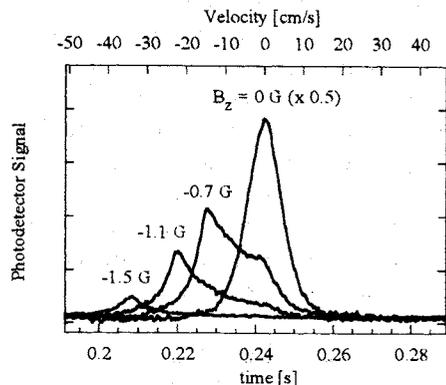


Figure 3: TOF signals for different static magnetic fields along the vertical  $z$ -axis. The signal for  $B_z = 0$  is divided by a factor 2. The arrival time of atoms having an initial velocity  $v_0 = 0$  G is 242 ms. By 1 cm/s of initial velocity the arrival time changes by 1 ms. The laser detuning and intensity are given by:  $\delta = -3\Gamma$  and  $I = 2I_s$ .

process efficiency: for  $|B_z| = 1.5$  G, the number of trapped atoms is only 30% smaller than the number at  $B_z = 0$  G. This shows that the drift velocity of the atoms also has non-vertical components. These effects already occur for very small values of the magnetic field ( $B_z \sim 0.2$  G). Although it is difficult to define a width and extract a temperature from the distorted TOF signals, the total TOF spread is too small to suggest the onset of a Doppler regime.

Contrary to the case of a 3-D molasses, the drift velocity in a 2-D trap is not linear with respect to the magnetic field. First, as shown in Fig. 5, the curve of the 2-D trap presents an asymmetry with regard to  $B_z = 0$  G. This asymmetry could be explained as follows: the atoms that are pushed towards  $+\vec{e}_z$  are not lost by the 2-D trap. Under the gravity they stay suspended at the upper edge of the cold column. Therefore the loading time method does not show any drift towards  $+\vec{e}_z$ . Secondly, for an extraction towards  $-\vec{e}_z$ , there is a threshold for the drift: for  $B_z > -0.8$  G the magnetic field induces a negligible drift on the trapped atoms and for  $B_z < -0.8$  G the atoms drift away from the trap with a velocity of  $v_d = 35 \text{ cm s}^{-1} \text{ G}^{-1}$ .

## 4 DISCUSSION

**Theoretical predictions** In order to understand qualitatively our experimental results, we developed a theoretical 1-D model using the transition  $J = 1 \rightarrow 2$  to study the effect of both intensity imbalance and longitudinal magnetic field on a  $\sigma^+ - \sigma^-$  molasses. We are aware that our experimental situation can not be fully described by this simple 1-D

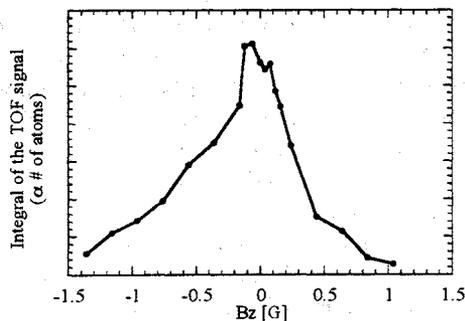


Figure 4: The integral of the TOF signal of an 3-D optical molasses vs. vertical magnetic field  $B_z$ . The laser detuning and intensity are given by:  $\delta = -3\Gamma$  and  $I = 2I_s$ .

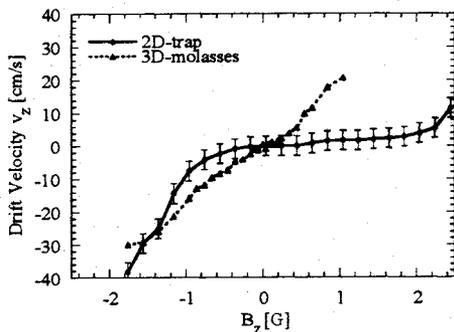


Figure 5: The drift velocity vs. the static magnetic field  $B_z$ . White circles represent the results of a 2D-trap (loading time measurement), black triangles stands for the TOF measurements in a 3D-molasses. The laser detuning and intensity are given by:  $\delta = -3\Gamma$  and  $I = 2I_s$ .

model. Our purpose, here, is to point out among all the experimental features those which can be easily understood. Using the procedure described by Dalibard *et al.* [9], we calculate the cooling force at small laser intensity for any value of the atomic velocity, while taking into account the modifications in the atomic density matrix induced simultaneously by an intensity imbalance or a magnetic field. As already predicted by Werner *et al.* [10] for an intensity imbalance and by Walhout *et al.* [11] for the magnetic field, we found that the net effect of such a perturbation is to shift the velocity to a non-zero value for which the force vanishes. This model allows to distinguish two different regimes of drift. The first one corresponds to small values of intensity imbalance ( $w \ll 1$ ) and magnetic field ( $\mu_B g_g B \ll \hbar \Gamma'$ ),

where  $g_g = 0.25$  is the Landé factor of the Cs ground state,  $\mu_B$  is the Bohr magneton and  $\Gamma'$  is the rate of optical pumping. This is the sub-Doppler regime: for our cooling parameters, the velocities of the drift induced by an intensity imbalance  $w$  and a magnetic field  $B$  are then given by:

$$\left(\frac{kv_d}{\Gamma}\right)_w^{SD} = -\frac{1}{3} \frac{\Gamma}{\delta} \frac{I}{I_s} w \quad (6)$$

$$\left(\frac{kv_d}{\Gamma}\right)_B^{SD} = \frac{g_g \mu_B B}{\hbar k} \quad (7)$$

where  $k$ ,  $\delta$ ,  $I$  are respectively the wave number, the detuning and the intensity of the cooling beams;  $I_s = 2.2 \text{ mW.cm}^{-2}$  is the intensity for which the Rabi angular frequency is equal to  $\Gamma$  for the  $|F = 4, m_g = 4\rangle \rightarrow |F' = 5, m_e = 5\rangle$  transition. Besides, for higher values of the imbalance ( $w > 0.3$ ) and the magnetic field ( $\mu_B g_g B > \hbar \Gamma'$ ) the regime becomes Doppler and the drift velocities are then given by:

$$\left(\frac{kv_d}{\Gamma}\right)_w^D = -\frac{\delta}{2\Gamma} w \quad (8)$$

$$\left(\frac{kv_d}{\Gamma}\right)_B^D = \frac{g_e \mu_B B}{\hbar k} \quad (9)$$

where  $g_e = 0.4$  is the Landé factor of the Cs excited state. Let us also point out that for intermediate values of the intensity imbalance ( $w \sim 0.3$ ) and the magnetic field ( $\mu_B g_g B \sim \hbar \Gamma'$ ), the force vanishes for three different values of the velocity. By solving the corresponding Fokker-Planck equation, one can show [10, 11], that the velocity distribution is no longer gaussian: it becomes a distribution with two peaks more or less resolved.

However, it is difficult to compare quantitatively the predictions of the model with the experimental results. Indeed our model is one-dimensional and it uses the transition  $J = 1 \rightarrow 2$ . For weak intensity imbalances ( $w \ll 1$ ), Werner *et al.* [10] calculated that the drift velocity using a  $F = 4 \rightarrow 5$  transition is eight times smaller than that given by equation (6). We use this factor to obtain the imbalance drift coefficient in the second line of Table 1 from the first line. Our model is also unable to predict the boundaries of the Doppler and sub-Doppler regime as  $w$  and  $B$  vary.

As mentioned by Steane *et al.* [12] and Werner *et al.* [10] a more complete description requires the consideration of the three-dimensional character of the cooling and that of the real atomic transition involved (i.e  $F = 4 \rightarrow 5$ ). A model that considers these two points was developed by Castin *et al.* [13]. Using the so-called Monte-Carlo wave-function technique, Y. Castin [14] calculated the effect of an intensity imbalance and a static magnetic field on a

	$(v_d)_B$	$(v_d)_w$
SubDoppler 1D-Model $J = 1 \rightarrow J' = 2$ $\delta = -3\Gamma, I = 2I_s$	29	87.4
SubDoppler 1D-Model $F = 4 \rightarrow F' = 5$ $\delta = -3\Gamma, I = 2I_s$	29	11
2D quantum Model $F = 4 \rightarrow F' = 5$ $\delta = -3\Gamma, I = I_s$	18	21
Experiment 3D $\delta = -3\Gamma, I = 2I_s$	18	9
Experiment 2D $\delta = -3\Gamma, I = 2I_s$	> -0.8 G 0 < -0.8 G 35	9 $ w  < 0.25$

Table 1: Summary of the theoretical and experimental results. Numerical values are in  $\text{cm s}^{-1} \text{G}^{-1}$  (first data column)  $\text{cm s}^{-1} \text{w}^{-1}$  (second data column)

2-D optical molasses. This fully quantum mechanical calculations were made with cooling parameters ( $\delta = -3\Gamma, I = I_s$ ) close to our experimental conditions. Its results are shown in the third line of Table 1.

**Effect of an intensity imbalance** From Fig. 1 and the last column of Table 1 we deduce that for small intensity imbalances ( $|w| < 0.25$ ) the drift along  $Oz$  induced in both an optical molasses and a 2-D trap matches the theoretical predictions of a 1-D sub-Doppler model. In addition, we verified experimentally that, as described by equation (6), the drift coefficient decreases with the saturation parameter  $s$  (same definition as in [9]) of the laser beams. The small drift coefficient ( $\sim 9 \text{ cm s}^{-1} \text{w}^{-1}$ ) that we measured is an obstacle to using this method as an extraction mechanism from a 2-D trap. Besides, for  $|w| > 0.25$  the drift velocity in the 2-D trap increases rapidly. The velocities are still too small to result from Doppler processes, but the behaviour of the curve suggests the onset of such a regime. The range of  $w$  values where the transition from sub-Doppler to Doppler regimes takes place is consistent with numerical calculations made (for  $\delta = -4\Gamma$  and  $I = I_s/\sqrt{2}$ , however) by Werner *et al.* [10] on the  $F = 4 \rightarrow 5$  Cs transition.

**Effect of the magnetic field** As shown in Table 1, the experimental drift velocity ( $18 \text{ cm s}^{-1} \text{G}^{-1}$ ) induced by a vertical magnetic field on an optical molasses is significantly lower than predicted by the

1-D sub-Doppler model [15]. This discrepancy can be explained by the effect of additional laser beams which carry other polarisation components than the  $\sigma^+ - \sigma^-$  configuration of the 1-D model. This reduces the anisotropy created by the optical pumping which is at the origin of the orientational cooling mechanism [9]. The fully quantum mechanical computation of Y. Castin [14] indeed predicts such a reduction (see the third line of Table 1).

Besides, 1-D models of magnetic induced drift predict as in the case of an intensity imbalance, a double-peaked velocity distribution in the intermediate range of magnetic field. The distorted velocity distribution emphasized by the TOF signals of Fig. 3 can not be explained in such terms: as mentioned previously, the full spread is too small to originate from Doppler processes. In addition, comparison with experimental results on the effect of a longitudinal magnetic field on a 1-D molasses [15] show that, in the range of magnetic field considered in our experiment, the drift velocity arises exclusively from sub-Doppler mechanisms. We believe that the distortion of the TOF signals is mainly due to the fact that the magnetic field forms an angle of  $45^\circ$  with the up- and down-going laser beams. With respect to the wave vectors of the two pairs of beams along the  $Oyz$  plane, the magnetic field has a non-zero transverse component. As calculated by Van der Straten *et al.* [16], in the presence of a transverse field  $B_T$  such as  $\mu_B g_g B_T \gg \hbar \Gamma'$ , the atoms are cooled towards five resonant velocities  $v_r = n \mu_B g_g B_T / 2 \hbar k$  ( $n = 0, \pm 1, \pm 2$ ) along the propagation axis of the beams. In our complex configuration, the effects of transverse magnetic fields may be expected to be as important as those of longitudinal fields. We therefore expect an important perturbation of the velocity distribution in the  $Oyz$  plane. This could also account for the rapid signal decrease (see Fig. 4): the magnetic field  $B_z$  could indeed induce a drift that has also non-vertical components. Similarly, one-dimensional models clearly can not account for the complex configuration of the 2-D trap with an axial field in which the full curve in Fig. 5 was obtained. The transverse trapping field varies between zero and approximately 0.5 G across the diameter of the elongated cloud. We believe that this transverse magnetic gradient is responsible for the discrepancy between the 2-D trap and the 3-D molasses, pointed out in Fig. 5. Indeed, the average drift coefficient in the 2-D trap is very weak ( $< 5 \text{ cm s}^{-1} \text{G}^{-1}$ ) in the low-field range ( $|B_z| < 0.8 \text{ G}$ ) where the longitudinal magnetic field is comparable to the transverse field. For higher values of the vertical magnetic field, the drift of atoms towards  $-\vec{e}_z$  appears to be much more efficient ( $\sim 35 \text{ cm s}^{-1} \text{G}^{-1}$ ). As far as the continuous extraction of atoms is concerned, our measurements indicate that the drift obtained by applying a vertical magnetic field to two

pairs of beams at  $45^\circ$  with respect to the  $z$ -axis, does not depend in such a simple way on the magnetic field as in the 1-D configuration where the applied field is parallel to the pair of laser beams. One would still need to measure the vertical velocity distribution of atoms in the 2-D trap configuration before concluding on the usefulness of such a simple method to extract atoms continuously from the trap.

## 5 CONCLUSION

Two-dimensional magneto-optical traps are being investigated as possible sources of continuous beams of cold Cs atoms. Extraction of atoms by intensity imbalance or magnetic field has been studied in detail. We demonstrated that atoms can be pushed out of the 2-D trap using these two methods. However the insensitivity of sub-Doppler molasses to intensity imbalances and the complexity of our magneto-optical configuration yield some obstacles to be overcome to get a continuous beam of cold atoms. Other extraction mechanisms such as moving molasses and dark channels are still under study.

## ACKNOWLEDGEMENT

The authors wish to thank Y. Castin for the numerical computations he made for them and C. Valentin for helpful discussions. This work was supported by the Swiss National Fund for Scientific Research, the Federal Office of Metrology and the Federal Office of Education and Science.

## References

- [1] S. Ghezali, P. Laurent, S. Lea, G. Santarelli, M. Bahoura, K. Szymaniec, and A. Clairon. A cesium fountain frequency standard: present status and future improvements. In *Proceedings of the Ninth European Frequency and Time Forum*, pages 101 - 110, March 1995.
- [2] Kurt Gibble and Steven Chu. Laser-cooled cesium frequency standard and a measurement of the frequency shift due to ultracold collisions. *Phys. Rev. Lett*, 70(12):1771 - 1774, 1993.
- [3] G. J. Dick, J. D. Prestage, C. A. Greenhall, and L. Maleki. Local oscillator induced degradation of medium-term stability in passive atomic frequency standards. In *Proceedings of the 22nd annual PTTI Meeting*, 1990.
- [4] S Ohshima *et al.* Multipulse operation of cesium atomic fountain. In *Proceedings of the Fifth Symposium on Frequency Standards and Metrology, Woods Hole, MA, USA*. World Scientific, Singapore, 1995.
- [5] W. Ketterle, K. B. Davis, M. A. Joffe, A. Martin, and D. E. Pritchard. High densities of cold atoms in a dark spontaneous-force optical trap. *Phys. Rev. Lett*, 70(15):2253 - 2256, 1993.
- [6] G Dudle, N Sagna, P Thomann, E Aucouturier, P Petit, and N Dimarcq. Generation of a continuous beam of cold cesium atoms. In *Proceedings of the Fifth Symposium on Frequency Standards and Metrology, Woods Hole, MA, USA*. World Scientific, Singapore, 1995.
- [7] N Sagna, G Dudle, and P Thomann. Atomic trapping and cooling in an anisotropic trap. In *Proceedings of the 9th European Frequency and Time Forum, Besançon, France*, pages 115 - 119, 1995.
- [8] P D Lett, W D Phillips, S L Rolston, C E Tanner, R N Watts, and C I Westbrook. Optical molasses. *J. Opt. Soc. Am. B*, 6(11):2085 - 2107, 1989.
- [9] J. Dalibard and C. Cohen-Tannoudji. Laser cooling below the doppler limit by polarization gradients: simple theoretical models. *J. Opt. Soc. Am. B*, 6(11):2023 - 2045, 1989.
- [10] J Werner, G Hillenbrand, and A Steane. Laser cooling by  $\sigma^+$  -  $\sigma^-$  circularly polarized beams of unequal intensities. *J. Phys. B*, 26:3063 - 3080, 1993.
- [11] M. Walhout, J. Dalibard, S. L. Rolston, and W. D. Phillips.  $\sigma^+$  -  $\sigma^-$  optical molasses in a longitudinal magnetic field. *J. Opt. Soc. Am. B*, 9(11):1997 - 2007, 1992.
- [12] A.M. Steane, M. Chowdhury, and C.J. Foot. Radiation force in the magneto-optical trap. *J. Opt. Soc. Am. B*, 9(12):2142-2158, 1992.
- [13] Y. Castin and K. Mølmer. Monte carlo wavefunction analysis of 3d optical molasses. *Phys. Rev. Lett*, 74:3772 - 3775, 1995.
- [14] Y. Castin Private communication
- [15] C. Valentin, M.-C. Gagné, J. Yu, and P. Pillet. One-dimension sub-doppler molasses in the presence of static magnetic field. *Europhys. Lett*, 17(2):133 - 128, 1992.
- [16] P Van der Straten, S-Q Shang, B Sheehy, H Metcalf, and G Nienhuis. Laser cooling at low intensity in a strong magnetic field. *Phys. Rev. A*, 47(5):4160 - 4175, 1993.

## THE L.H.A. 2D MAGNETO-OPTICAL TRAP: ON THE WAY TO A CONTINUOUS BEAM OF COLD ATOMS

Eve Aucouturier, Noël Dimarcq, Pierre Petit, Constance Valentin, and Stefan Weyers

Laboratoire de l'Horloge Atomique, Unité Propre de Recherche du C.N.R.S., associée à l'Université Paris-Sud,  
Université Paris-Sud, Bat 221, 91405 ORSAY Cedex, FRANCE.

### ABSTRACT

Last year we presented a proposal for the generation of a continuous beam of cold atoms as a new alternative to improve the performance of Caesium frequency standards. The starting point of our continuous beam realisation is a two-dimensional magneto-optical trap (2D MOT): the assembling of three counter-propagating circularly polarised laser beams and of two magnetic field gradients along the horizontal axes leads to a transverse cooling and trapping, whereas in the vertical z-direction the atoms are only cooled and not trapped because of the zero magnetic field in this direction. To launch the cold atoms in the z-direction it is planned to use two different techniques. We intend either to create a 1D moving molasses by a proper detuning of the frequencies of the vertical laser beams or either to apply a static magnetic field along the z-direction in order to imbalance the interaction of the atoms with the counterpropagating vertical laser beams.

In order to predict the characteristics of the beam of cold atoms, we have done a simulation of the atomic trajectories by using a two-dimensional model of Doppler cooling and magneto-optical trapping. We will also present a calculation of the velocity distribution of the atoms, which, due to the trapping and cooling process, enter the zone, where they are launched. This work might help us to understand the results of future experimental measurements on the cold beam. In the meantime, we are adding several laser beams to the two-dimensional MOT, which will be used for the analysis of the main beam characteristics: the atomic flux, its angular dispersion, and its velocity distribution.

These experiments are a burning issue for the generation of a continuous beam of cold atoms and the development of a new design for future atomic clocks.

### INTRODUCTION

The advantages of cold atoms for primary frequency standards are now well admitted (Clairon et al, (1)). At the present time, most of the research on frequency standards using cold atoms are based on pulsed operation. One ball containing about one million cold atoms is launched every second through a microwave cavity. This functioning has some disadvantages. For example, the principle of a pulsed interrogation in the microwave cavity leads to a degradation of the short-term stability due to the low duty cycle of the interrogation (Santarelli et al (2)).

The creation of a continuous beam of cold atoms might lead to a new generation of Caesium frequency standards. It is based on a simultaneous cooling and launching of the atoms. Several advantages make it very interesting for primary frequency standards. For example, it should release us from the collisional effects. Indeed, for the same number of atoms launched in one second, the atomic density is much lower with a continuous beam where the atoms are distributed all along the atomic path. One may also use the properties of the cold atoms to shorten the interrogation cavity size, while keeping the atomic clock stability to a level similar to the best Caesium thermal beam frequency standards, so as to lead to a compact and transportable frequency standard. For this, a continuous beam of Caesium atoms with an axial mean velocity between 1 and 10 m/s and axial and transversal velocity dispersions less than 10 cm/s is necessary.

We expound here the experimental set-up that has been built up at L.H.A. for trapping the atoms in an anisotropic magneto-optical trap (2D MOT). We consider launching the atoms using two methods: the moving molasses technique and the magneto-optical train. We finally present some numerical simulations we did with a very simple theoretical model based on a Monte-Carlo approach in order to predict the characteristics of the future beam of laser cooled atoms.

### THE 2-DIMENSIONAL MAGNETO-OPTICAL TRAP (2D MOT)

The 2D MOT works on the principle of a cooling and trapping of the atoms along the horizontal (Ox) and (Oy) axes, while only cooling, with no trapping occurs in the vertical (Oz) direction. Horizontal trapping results from

two pairs of circularly polarised counter-propagating laser waves and two magnetic field gradients. Vertical cooling results from two counter-propagating laser waves with no magnetic gradient along ( $Oz$ ) axis. The total magnetic field is produced in our case by two pairs of anti-Helmholtz coils. We will launch the atoms vertically, because they can easily leak from the MOT in this direction. Figure 1 shows the experimental set-up. The atomic beam must be free of any laser light, so the vertical cooling laser beam was split into two thin beams making a slight angle ( $14^\circ$ ) with the vertical axis.

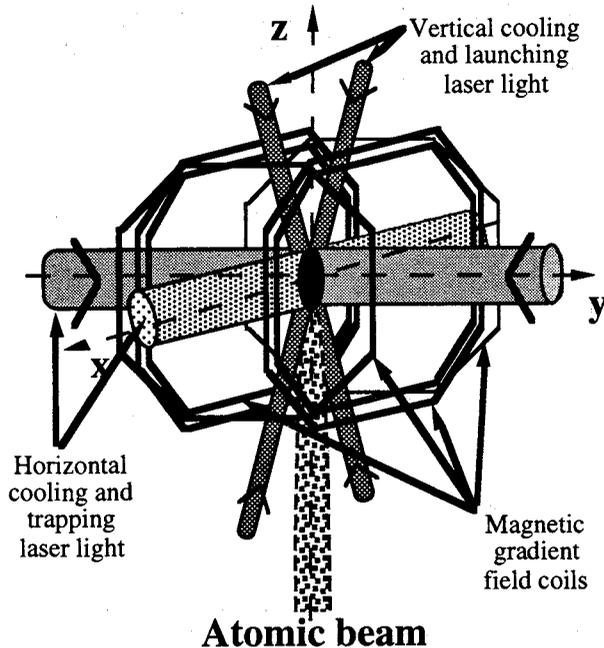


Figure 1: Experimental set-up.

### EXPERIMENTAL SET-UP FOR LAUNCHING THE ATOMS

The present set-up allows us to launch the atoms using two different methods.

The first one, called "moving molasses", consists in applying a slight positive frequency shift to the forward cooling laser beam, and the same negative frequency shift to the backward laser beam. The atoms are cooled in a moving frame, which velocity is in direct proportion to the applied frequency shift.

The second method, called "magneto-optical train", consists in applying a static magnetic field  $B$  in the vertical direction, with circularly polarised laser beams in this direction. The atomic Zeeman sublevels are split, thus modifying the resonance conditions depending on the laser light polarisation. As a result of this, the atoms are cooled in a moving frame, which velocity is in direct proportion to the magnetic field  $B$ . This extraction mechanism is also being studied by Sagna et al (3) at Observatoire de Neuchâtel (Switzerland) in a different light and magnetic

configuration, and we hope to perform worthwhile complementary experiments.

We are at the moment installing a system for the detection and analysis of the future beam of cold atoms. A pencil of resonant light located about 50 cm below the 2D MOT allows the detection of atoms falling across it through their fluorescence light. In order to measure the angular dispersion of the laser cooled atomic beam, we shall observe the light spot made by the atoms falling through the analysis laser beam with the aid of an ultra-sensitive CCD camera (see Figure 2).

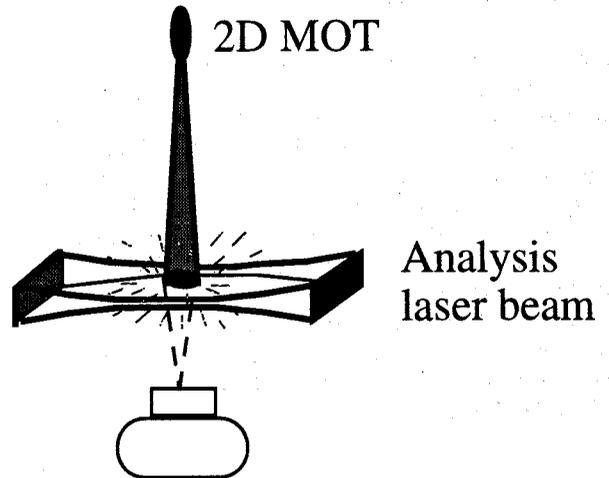


Figure 2: Measurement of the beam divergence.

The time of flight technique is the most usual method for the determination of velocity distributions in a MOT or in an optical molasses. A ball of cold atoms is created. At the instant  $t = 0$ , the laser light and if needed the magnetic gradients are switched off. The atoms fall under the gravitational force. One detects them when they cross an analysis laser beam located a few centimetres below the cooling area. The vertical velocity component of the atoms is deduced from the time of flight signal.

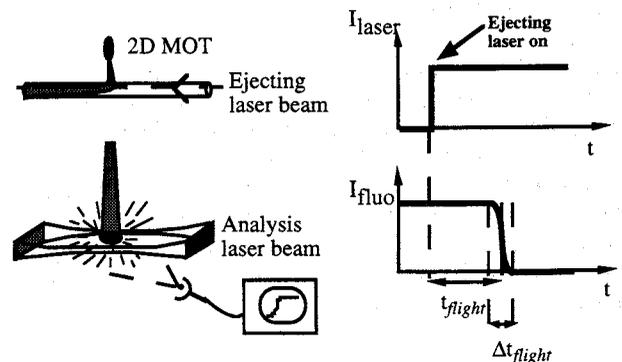


Figure 3: Measurement of the atomic axial velocity distribution. Time of flight technique.

Since our experiment shall lead to a continuous extraction of the atoms from the 2D MOT, we cannot use this method in the usual way. An addicted progressive laser wave is located a few centimetres

underneath the 2D MOT so as to switch on or off the atomic beam (see Figure 3). At an instant  $t = 0$ , this laser beam is switched on, it ejects the atoms falling across it out of their vertical path. As for a usual time of flight technique, one deduces the axial velocity distribution of the atomic beam from the time evolution of the light emitted by the atoms as they cross the analysis laser beam. The mean velocity is proportional to the time of flight  $t_{\text{flight}}$ , while the velocity dispersion is related to the characteristic duration  $\Delta t_{\text{flight}}$  of the fluorescence time evolution.

## NUMERICAL SIMULATIONS

Using a two-dimensional model of Doppler radiative cooling and magneto-optical trapping, we performed numerical simulations of the atoms interacting with laser light in the following configuration (see Figure 4). The position (res. velocity) of the atom is described by its components on the horizontal and vertical axes:  $x$  and  $z$  (res.  $v_x$  and  $v_z$ ). In an area corresponding to  $|x| < x_{\text{trap}}$  and  $|z| < z_{\text{trap}}$ , the atom undergoes a 1D MOT created by a magnetic field gradient and an adequately polarised standing wave. For  $|x| < x_{\text{launching}}$  and  $|z| < z_{\text{trap}}$ , a perpendicular 1D moving molasses performing its launching in the  $-z$  direction is superimposed on the previous 1D MOT.

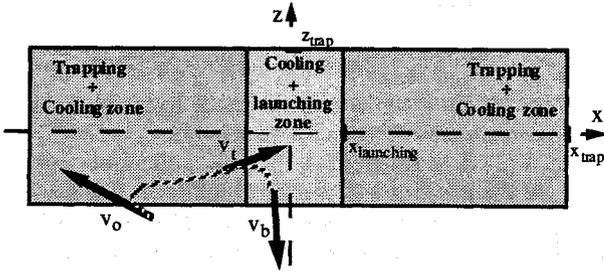


Figure 4: Scheme of the different interaction areas for the numerical simulations.

### Simulation of the atomic paths

We depict the trajectories of some atoms in this device for various initial positions and velocity directions, using a Monte-Carlo approach (Grison (4)). The photon-atom interaction is modelled as following. The photon absorption/emission cycle is described by Bloch's optical equations under weak light intensity (total saturation parameter  $s \ll 1$ ). The transition probability from the ground level to an excited level in a short-time duration  $dt$  is given by equation (1).

$$P = \frac{\Gamma}{2} \cdot s \cdot dt$$

eq. (1)

Where  $\Gamma$  is the spontaneous deexcitation rate,

and  $s$  is the total saturation parameter (equation (2)).

$$s = s_{\text{left}} + s_{\text{right}} + s_{\text{low}} + s_{\text{up}}$$

$$\text{for } |x| < x_{\text{launching}} \text{ and } |z| < z_{\text{trap}}$$

$$s = s_{\text{left}} + s_{\text{right}}$$

$$\text{for } x_{\text{launching}} < |x| < x_{\text{trap}} \text{ and } |z| < z_{\text{trap}}$$

eq. (2)

Where the saturation parameters from the left, right, low, and up light waves ( $s_{\text{left}}$ ,  $s_{\text{right}}$ ,  $s_{\text{low}}$ , and  $s_{\text{up}}$ ) are functions of the laser light intensity and frequency of each laser wave, and of the position and velocity components of the atom.

$$s_i = \frac{\frac{1}{2} \cdot \Omega_{\text{laser}}^2}{\Delta_i^2 + \frac{\Gamma^2}{4}} \quad i \equiv \text{left, right, low, up}$$

with

$$\Delta_{\text{left}} = \omega_{\text{at}} - \omega_{\text{cooling}} + k_{\text{laser}} \cdot v_x + \mu_{\text{Larmor}} \cdot x$$

$$\Delta_{\text{right}} = \omega_{\text{at}} - \omega_{\text{cooling}} - k_{\text{laser}} \cdot v_x - \mu_{\text{Larmor}} \cdot x$$

$$\Delta_{\text{low}} = \omega_{\text{at}} - \omega_{\text{cooling}} + k_{\text{laser}} \cdot v_z + d\omega$$

$$\Delta_{\text{up}} = \omega_{\text{at}} - \omega_{\text{cooling}} - k_{\text{laser}} \cdot v_z - d\omega$$

eq. (3)

Where  $\omega_{\text{at}}$  is the atomic frequency,  $\Omega_{\text{laser}}$  is the Rabi frequency of the laser wave,  $k_{\text{laser}}$  is the laser wave vector,  $\omega_{\text{cooling}}$  is the cooling frequency of the laser light,  $\mu_{\text{Larmor}}$  is a coefficient related to the Zeeman shift of the atomic levels, and  $d\omega = \frac{d\omega}{2\pi}$  is either the additional

frequency detuning of the vertical laser beams in case of the moving molasses launching technique, or the Zeeman frequency shift of the atomic levels due to the vertical magnetic field in case of the magneto-optical train launching method.

Each time the atom absorbs a photon from one of the laser beams, its velocity vector is modified by the photon recoil quantity:

$$d\vec{v} = \frac{\hbar \cdot k_{\text{laser}}}{M} \cdot \vec{u}$$

eq. (4)

Where  $\hbar$  is Planck constant,  $M$  is the atomic mass, and  $u$  is the incident photon direction. The emission process is described by a similar approach. The new atomic position and velocity are calculated after each absorption-emission cycle, until the atom gets out of the interaction area. Figure 5 shows an example of such trajectories for an initial atomic velocity value of 5 m/s.

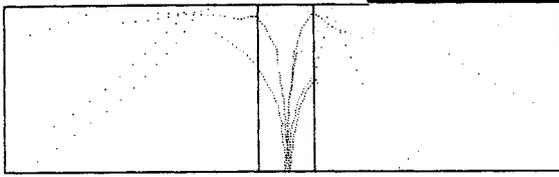


Figure 5: Typical atomic trajectories.

$$I_{\text{laser}} = 1 \text{ mW/cm}^2$$

Atoms initial velocity: 5 m/s

Additional vertical frequency detuning:  $df = \pm 1.2 \text{ MHz}$

### Simulation of the velocity distribution in the cold atoms beam

It is possible to calculate the velocity distribution of the atoms passing through a given surface of the interaction area. The code we conceived calculates the evolution of the atom in the interaction area with a model of average radiation forces:

$$\langle \vec{F} \rangle = \hbar \frac{\Gamma}{2M} \cdot ((s_{\text{left}} - s_{\text{right}}) \cdot \vec{u}_x + (s_{\text{low}} - s_{\text{up}}) \cdot \vec{u}_z) \quad \text{eq. (5)}$$

Assuming an initial thermal velocity distribution, the code follows the motion of one atom entering the interaction area with a velocity  $v_0$ . The atom is subject to the horizontal 1D MOT, until it reaches the central area (defined by  $|x| < x_{\text{launching}}$  and  $|z| < z_{\text{trap}}$ ) with a velocity  $v_1$ . It then experiences the effects of the horizontal 1D MOT and the vertical moving molasses, until it leaves the MOT, being launched downwards with a velocity  $v_b$ .

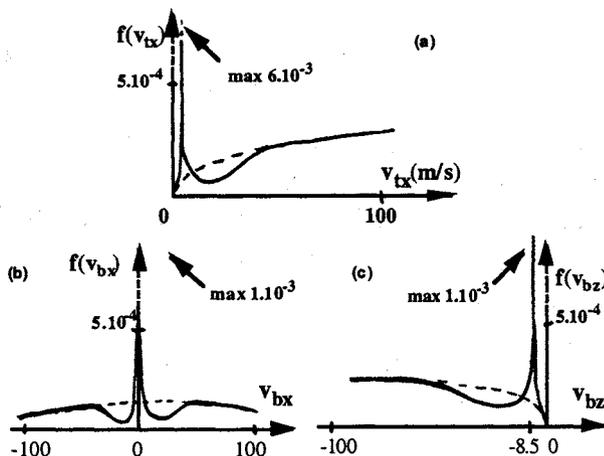


Figure 6: Velocity distributions.

- (a)  $v_{tx}$ : x component velocity of trapped atoms  
 (b)  $v_{bx}$ : transversal velocity of atoms in the beam  
 (c)  $v_{bz}$ : axial velocity of atoms in the beam  
 Dashed lines: thermal distribution ( $I_{\text{laser}} = 0$ ).

The program performs this calculation for some  $10^7$  atoms and presents the velocity distribution  $f(v_i)$  of atoms crossing the sections defined by  $x = \pm x_{\text{launching}}$  (atoms entering the central area), as well as the velocity distribution  $f(v_b)$  of atoms crossing the  $z = -z_{\text{trap}}$  section (atoms being launched downwards).

Some preliminary results of this calculation are presented Figure 6 for a laser power of  $1 \text{ mW/cm}^2$  in each beam, and a cooling frequency shift  $\omega_{\text{at}} - \omega_{\text{cooling}}$  equal to  $-3\Gamma$ . This work shall help us in predicting the angular dispersion and temperature of the cold atoms beam.

### CONCLUSION

We are studying a configuration of 2D MOT allowing the simultaneous trapping, cooling and launching of the atoms, which is now in operation at L.H.A. We intend to demonstrate experimentally the feasibility of a continuous beam of laser cooled atoms in a near future.

We performed numerical simulations so as to predict the characteristics of the future atomic beam. Nevertheless, this prediction has to be compared with experiment measurements, which may reveal sub-Doppler cooling mechanisms, that are not included in our calculation.

### ACKNOWLEDGMENTS

The authors wish to thank the Bureau National de Metrologie (Paris) and the Direction des Recherches Etudes et Techniques (Paris) for their financial support.

### REFERENCES

- (1) A. Clairon, P. Laurent, G. Santarelli, S. Ghezali, S. N. Lea, and M. Bahoura: "A CESIUM FOUNTAIN FREQUENCY STANDARD: PRELIMINARY RESULTS". 1995, *IEEE Trans. on IM*, **44**, 128-131.
- (2) G. Santarelli, P. Laurent, A. Clairon, G.J. Dick, C.A. Greenhall and C. Audoin: "THEORETICAL DESCRIPTION AND EXPERIMENTAL EVALUATION OF THE EFFECT OF LOCAL OSCILLATOR FREQUENCY NOISE ON THE STABILITY OF A PULSED ATOMIC FREQUENCY STANDARD", *10TH EFTE*, 1996, Brighton, UK.
- (3) N. Sagna, G. Dudle and P. Thomann: "CONTINUOUS EXTRACTION MECHANISMS FROM A MAGNETOOPTIC TRAP", *10TH EFTE*, 1996, Brighton, UK.
- (4) D. Grison, 1992: "ATOMES PIEGES ET REFROIDIS PAR LASER A QUELQUES MICROKELVINS", PHD Thesis, Paris, FRANCE.

# INFLUENCE OF ELECTRODE MASS-LOADING ON THE ELECTRICAL EQUIVALENT CIRCUIT OF THE TRAPPED-ENERGY AT-CUT QUARTZ RESONATORS

Jiri Zelenka

Technical University of Liberec, Faculty of Mechatronics,  
Háalkova 6, CZ-461 17 Liberec, Czech Republic.

## ABSTRACT

The static  $C_o$  and motional  $C_n$  capacitance of the electrical equivalent circuit of the AT- and BT-cut quartz resonators are computed. Another one comparison is done for infinitely large planparallel plate, for planparallel energy trapped resonator by considering the mass loading of electrodes and for the energy trapped resonator by considering the mass loading of electrodes and the spherically contoured surface of the plate. The calculated values are compared also with the measured one.

## INTRODUCTION

The exact values of the parameters of the electrical equivalent circuit of the piezoelectric resonators are very important for the design and realization of the wave filters. The parameters are functions not only of the plate and electrodes dimensions and elastic, piezoelectric and dielectric properties of the plate but they depend also on the piezoelectric stiffening, electrode mass loading, energy trapping and the flatness (curvature) of the surface of the plate in the paper. The attention will be given to the electrical equivalent circuit parameters of the thickness vibration of the AT- and BT-cut quartz plates. The computed and measured values of the motional capacitance of the infinite flat plate (by using of the relations given by Yamada and Niizeki (1)), flat energy trapped resonators (by using of the relations given by Tiersten (2)) and the contoured trapped energy resonators (by using of the relations given by Tiersten and Smithe (3))

are compared. The attention will be also given to the non-uniform distribution of the motion mentioned by Kosinski et al (4).

## QUANTITIES $C_o$ AND $C_n$

The static capacitance  $C_o$  and the motional capacitance  $C_n$  of the piezoelectric resonators vibrated in the thickness mode with the thickness of the plate  $2h_o$  and with the area of the electrodes  $A_e$  can be expressed by the relations

$$C_o = \frac{A_e k_e \epsilon_{22}}{2h_o} (1 + k_{26}^2), \quad (1)$$

$$C_n = \frac{A_e \epsilon_{22}}{2h_o} \frac{8k_{26}^2}{n^2 \pi^2} k_c, \quad (2)$$

where  $\epsilon_{22}$  is the permittivity component in the thickness direction of the plate,  $k_{26}$  is the coefficient of the electromechanical coupling

$$k_{26}^2 = \frac{e_{26}^2}{c_{66} \epsilon_{22}}, \quad (3)$$

$c_{66}$  is the elastic stiffness component,  $e_{26}$  is the piezoelectric coefficient,  $k_e$  and  $k_c$  are coefficients which depend on the shape of the piezoelectric plate and the mass loading of electrodes.

1. For the infinitely large planparallel plate the coefficients

$$k_e = 1, \quad k_c = 1. \quad (4)$$

2. For the planparallel energy trapped resonator by considering the mass loading  $\hat{R}$  of electrodes the value of the coefficient  $k_c$

follows from the relations given by Tiersten (2)

$$k_e = 1, \quad (5)$$

$$k_c = \frac{4 \sin^2(\xi_{n\nu} l) \sin^2(\nu_{n\gamma} b)}{A_e \xi_{n\nu}^2 \nu_{n\gamma}^2 L_{n\nu\gamma}}, \quad (6)$$

where the wave numbers  $\xi_{n\nu}$  and  $\nu_{n\gamma}$  in the  $x_1$  and  $x_3$  directions are given by the relations

$$\xi_{n\nu} \tan(\xi_{n\nu} l) = \left( \frac{k_n}{M_n} \Delta_n - \xi^2 \right)^{1/2}, \quad (7)$$

$$\nu_{n\gamma} \tan(\nu_{n\gamma} b) = \left( \frac{k_n}{c_{55}} \Delta_n - \nu^2 \right)^{1/2}, \quad (8)$$

$$\Delta_n = \frac{n\pi}{2h_o} \sqrt{\frac{\bar{c}_{66}}{\rho}} \left( \frac{4k_{26}^2}{n^2\pi^2} + \hat{R} \right), \quad (9)$$

$$k_n = \frac{n\pi}{h_o} \sqrt{\bar{c}_{66}\rho}, \quad (10)$$

$$M_n = c_{11} + (c_{12} + c_{66})r + 4(r\bar{c}_{66} - c_{66})(c_{22}r + c_{12}) \frac{\cot(\frac{1}{2}\kappa n\pi)}{c_{22}\kappa n\pi}, \quad (11)$$

$$\kappa = \sqrt{\frac{\bar{c}_{66}}{c_{22}}}, \quad r = \frac{c_{12} + c_{66}}{\bar{c}_{66} - c_{22}}, \quad (12)$$

$$L_{n\nu\gamma} = \left( l + \frac{\sin(2\xi_{n\nu} l)}{2\xi_{n\nu}} \right) \left( b + \frac{\sin(2\nu_{n\gamma} b)}{2\nu_{n\gamma}} \right) + \frac{\cos^2(\nu_{n\gamma} b)}{\nu_{n\gamma}^T} \left( l + \frac{\sin(2\xi_{n\nu} l)}{2\xi_{n\nu}} \right) + \frac{\cos^2(\xi_{n\nu} l)}{\xi_{n\nu}^S} \left( b + \frac{\sin(2\nu_{n\gamma} b)}{2\nu_{n\gamma}} \right) + \frac{\cos^2(\xi_{n\nu} l) \cos^2(\nu_{n\gamma} b)}{\xi_{n\nu}^S \nu_{n\gamma}^T}, \quad (13)$$

$$\xi_{n\nu}^S = \xi_{n\nu} \tan(\xi_{n\nu} l), \quad \nu_{n\gamma}^T = \nu_{n\gamma} \tan(\nu_{n\gamma} b). \quad (14)$$

The coefficients  $k_c$  computed for the AT-cut quartz plate with the square electrodes as a function of the length of square ( $2l = 2b$ ) and mass loading  $\hat{R}$  are given in Fig. 1.

3. For the energy trapped resonator by considering the mass loading  $\hat{R}$  of electrodes and the spherically contoured surface of the plate given by the radius  $R$  of the curvature the values of the coefficients  $k_e$  and  $k_c$  follows

from the relations given by Tiersten and Smithe (3)

$$k_e = 1 + \frac{l_1^2}{8Rh_o}, \quad (15)$$

$$k_c = \frac{h_o}{2A_e} \frac{g_{nmp}^2}{L_{nmp}}, \quad (16)$$

where for  $m, p = 0, 2, 4$

$$g_{nmp} = 4F_{1nm}F_{3np}, \quad (17)$$

$$L_{nmp} = \frac{\pi h_o 2^m m! 2^p p!}{\sqrt{\alpha_n \beta_n}}, \quad (18)$$

$$\alpha_n = \frac{n^2 \pi^2 \hat{c}_{66}}{8R h_o^3 M_n}, \quad (19)$$

$$\beta_n = \frac{n^2 \pi^2 \hat{c}_{66}}{8R h_o^3 c_{55}}, \quad (20)$$

$$\hat{c}_{66} = \bar{c}_{66} \left( 1 - \frac{8k_{26}^2}{n^2 \pi^2} - 2\hat{R} \right), \quad (21)$$

$M_n$ ,  $\kappa$  and  $r$  are given by the relations (11) and (12) and

$$F_{1n0} = \sqrt{\frac{\pi}{2\alpha_n}} \operatorname{erf} \left[ l_1 \sqrt{\frac{\alpha_n}{2}} \right],$$

$$F_{3n0} = \sqrt{\frac{\pi}{2\beta_n}} \operatorname{erf} \left[ l_3 \sqrt{\frac{\beta_n}{2}} \right],$$

$$F_{1n2} = 2F_{1n0} - 4l_1 \exp\left(-\frac{\alpha_n l_1^2}{2}\right),$$

$$F_{3n2} = 2F_{3n0} - 4l_3 \exp\left(-\frac{\beta_n l_3^2}{2}\right). \quad (22)$$

## RESULTS

The static and motional capacitances were computed for all three considered types of resonators. The coefficients  $k_e$  and  $k_c$  were computed for the resonators with the square electrodes. The coefficients can be also used for the circular electrodes with diameter  $2l$  if the area of electrodes  $A_e$  in equations (1) and (2) are multiplied by the correction factor  $\pi/4$ .

The computed static capacitances  $C_o$  and motional capacitances  $C_1$  are given in TABLE 1. The diameters of circular electrodes were 8 mm. The thicknesses of the

blanks  $2h_0$  were 0.1658 mm for AT-cut and 0.2554 mm for BT-cut resonators.

It is clear from the TABLE 1 that the influence of the mass-loading of electrodes on the static and motional capacitances is much smaller (especially for the large electrodes) than the influence of the curvature of the plate surface.

The measured static and motional capacitances of the AT- and BT-cut are given in TABLE 2. The blanks had a shape of the square  $15 \times 15$  mm and the thicknesses 0.1658 mm and 0.2554 mm. The Ag electrodes had the diameter 8 mm and the thickness corresponds to the mass loading  $\hat{R}$  0.02305 for AT-cut and 0.01578 for BT-cut. The difference between the measured and computed motional capacitance for resonators of type II is smaller than 7 % for AT-cut and 3 % for BT-cut. The measured static capacitances are greater than the computed one as the measured static capacitances include also the capacitances of the resonator holder.

#### ACKNOWLEDGMENT

The work was supported by the Grant Agency of Czech Republic under Grant No. 102/941/1571.

#### REFERENCES

1. Yamada T and Niizeki N, 1970, Admittance of piezoelectric plates vibrating under the perpendicular field excitation, Proc. IEEE **58**, 941-942.
2. Tiersten H F, 1976, Analysis of trapped-energy resonators operating in overtones of coupled thickness shear and thickness twist, J. Acoust. Soc. Amer. **59**, 879-888.
3. Tiersten H F and Smythe R C, 1979, An analysis of contoured crystal resonators operating in overtones of coupled thickness shear and thickness twist, J. Acoust. Soc. Amer. **65**, 1455-1460.

4. Kosinski J A et al, 1994, Inclusion of non-uniform distribution of motion effects in the transmission-line analogs of the piezoelectric plate resonator: Theory and experiment, 1994 IEEE Inter. Freq. Control Symp., 229-235.

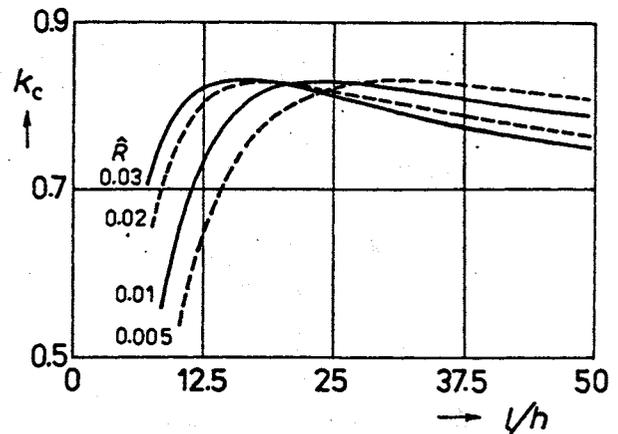


Fig. 1 The dependence of the coefficient  $k_c$  on the mass loading  $\hat{R}$  and the ratio of the half lengths  $l$  of electrodes and the half thickness  $h_0$  of the plate.

TABLE 1 - Calculated static  $C_0$  and motional  $C_n$  capacitances of AT- and BT- quartz resonators. The circular electrodes have the diameter 8 mm ( $A_e = 50,265 \text{ mm}^2$ ).

Cut	Resonators type I		Resonators type II			Resonators type III				
	$C_0$ [pF]	$C_n$ [fF]	$\hat{R}$	$C_0$ [pF]	$C_n$ [fF]	$\hat{R}$	$R$ [mm]	$C_0$ [pF]	$C_n$ [fF]	
AT	12.18	76.32	0.0050	12.18	61.76	0.0050	20000	12.197	71.03	
			0.0100	12.18	60.32	0.0100	20000	12.197	71.16	
			0.0200	12.18	58.52	0.0200	20000	12.197	71.41	
			0.0230	12.18	58.15	0.0231	20000	12.197	71.48	
						0.0050	5000	12.241	47.13	
						0.0100	5000	12.241	47.33	
						0.0200	5000	12.241	47.72	
						0.0230	5000	12.241	47.84	
	BT	7.95	20.58	0.0050	7.95	16.53	0.0050	10000	7.957	13.48
				0.0100	7.95	16.14	0.0100	10000	7.957	13.48
0.0158				7.95	15.82	0.0158	10000	7.957	13.49	

TABLE 2 - Measured static  $C_0$  and motional  $C_n$  capacitances of AT- and BT-cut quartz planparallel square resonators. The circular Ag electrodes have the diameter 8 mm.

Cut	Resonant frequency [kHz]	$\hat{R}$	$C_0$ [pF]	$C_n$ [fF]
AT	9769.8	0.02305	15.77	62.15
BT	9771.8	0.01578	9.88	16.28

## DOUBLE-MODE EXCITATION OF QUARTZ CRYSTALS WITHOUT DIPS OF ACTIVITY ON THE C AND B MODES

I. Abramson

Omsk Research Instrumentation Institute, P.O. Box 7357,  
644020 Omsk, Russia

Doubly-rotated cut quartz resonators excited on the main C-mode and the thermosensitive B-mode simultaneously look as the most promising way for essential improvement of the basic performances of the MCXOs and the temperature sensors. Nevertheless the double-mode resonators have not found wide application till now due to unacceptable fluctuations and even "dips" of activity suffered by the B-mode within a wide temperature range.

Present work is aimed at elucidation of physical origins of the phenomena as well as at search for effective method of designing the well-operating the double-mode resonators.

Careful exploration of the frequency spectrum of the fundamental mode and the 3d overtone SC-cut crystals while ambient temperature changes allowed to identify spurious modes interacting with the B-mode. Furthermore, a dependence of the interaction intensity on the "energy trapping" in the crystal plate has been revealed and explained.

On the investigations the SC-cut 10 MHz fundamental mode and 3d overtone crystals have been developed. The C-mode activity showed no noticeable fluctuations within  $(-40 + 70)^{\circ}\text{C}$  temperature range, while the B-mode activity deviated within 30 per cent for the fundamental mode and within 10 per cent for the 3d overtone crystals.

Although present work was concentrated on the SC-cut crystals developed method can be applied to other doubly-rotated cut crystals which may occur preferable in use due to more convenient the modes activity ratio or the frequency versus temperature form.

# SC-CUT RESONATOR OPERATING IN ANHARMONIC MODES WITH B-MODE REDUCTION

R. BOURQUIN, B. DULMET, J.J. BOY

ECOLE NATIONALE SUPERIEURE DE MECANIQUE ET DES MICROTÉCHNIQUES  
Laboratoire de Chronométrie, Electronique et Piézoélectricité  
26 Chemin de l'Épitaphe  
25030 BESANCON CEDEX FRANCE

## ABSTRACT

The design of a SC-cut resonator for which the C-mode has a lower resistance than the B-mode (typical ratio 1.5) is presented. This result is obtained by using an antisymmetric anharmonic mode with a particular geometry and connection of the electrodes. Experimental results (motional resistance, thermal behaviour) are given for a third overtone, 10 MHz, resonator.

## I- INTRODUCTION

In usual SC-cut resonators, the B-mode exhibits a lower motional resistance than the C-mode for which the resonator is designed. This fact comes from low acoustic attenuation and high electromechanical coupling factor of the B-mode when the driving electric field is normal to the plate of the resonator.

Until now, reduction of the electrical response of the B-mode has only been obtained by using lateral field excitation (Smythe [1], Masiukiewicz and al. [2]). This way suffers the disadvantage of increasing the motional resistance of the C-mode by an order of magnitude.

We propose an alternative way to obtain SC-cut resonators for which the B-mode resistance is 50 percent higher than this of the C-mode. This ratio is sufficient to insure that an oscillator will be able to oscillate at the right frequency without the help of any selective circuit. Contoured quartz resonators have many modes of vibration other than usual mode. These other modes, namely anharmonic modes, present different distributions of the mechanical displacement in the plane of the resonator, with nodal lines (Spencer [3], Dulmet and Bourquin [4]), the orientation of which are depending on the mode family (A, B or C) and on the overtone number. They exhibit the same Q-factor than the usual mode if the trapping of the vibration in the resonator is sufficient. They can be excited by using a proper design of the geometry of the electrodes. If the resonator operates on the third overtone, it is possible to find anharmonic of C-mode for which the motional resistance will be the lowest.

## II - MODES OF VIBRATION IN DOUBLY ROTATED CONTOURED QUARTZ RESONATORS

In a doubly rotated quartz resonator, the mechanical displacement associated with each mode (A, B or C) has three components  $A_k$  which are the components of the eigenvectors of the Christoffel's tensor  $\bar{\Gamma}_{jk}$  (Auld [5]).

$$\left(\bar{\Gamma}_{jk} - \bar{c}^{(i)} \delta_{jk}\right) A_k^{(i)} = 0 \quad (1)$$

each eigenvector  $A_k^{(i)}$  being associated with the corresponding eigenvalue, or effective elastic coefficient  $\bar{c}^{(i)}$ . The superscript (i) denotes the solution for the particular mode A, B or C. In general, all the quantities in (1) are expressed in the usual orthogonal system ( $x_1, x_2, x_3$ ) of a doubly rotated cut (IEEE Standards [6]) for which the normal to the surface of the plate is oriented along  $x_2$ . The components of the mechanical displacement and the effective elastic coefficient for the three modes of a typical SC-cut ( $\varphi = 22^\circ 45'$ ,  $\theta = 34^\circ$ ) are given in the table I.

Table I : Eigenvectors and eigenvalues of the three modes in SC-cut.

Mode	A <sub>1</sub>	A <sub>2</sub>	A <sub>3</sub>	$\bar{c}(N/m^2)$
A	0.229	0.967	0.112	120.6 10 <sup>9</sup>
B	0.189	0.068	-0.980	41.7 10 <sup>9</sup>
C	0.955	-0.245	0.167	34.6 10 <sup>9</sup>

In the case of a bounded plate, the mechanical displacement can be expressed as :

$$u_k^{(i)} = u_{nmp}^{(i)}(x_1, x_3) \sin \frac{n\pi x_2}{2h} A_k^{(i)} \cdot e^{j\omega_{nmp} t} \quad (2)$$

where  $2h$  is the total thickness of the plate and  $n$  the overtone number. The indices  $m, p$  label the different mode shapes in the plane of the plate and  $\omega_{nmp}$  is the corresponding resonant frequency. We now consider a plano-convex resonator with a thickness at its center equal to  $2h_0$  and a radius of curvature of the spherical face  $R$ . If the trapping of the vibration is sufficiently strong so that the lateral boundaries have no effect on the vibration, the

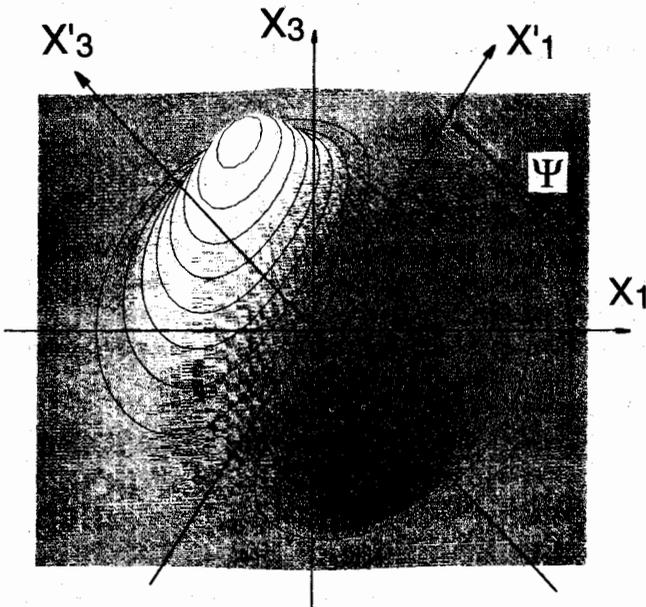
mode shape  $u_{nmp}$  in the plane of the resonator, is given by Stevens and Tiersten [7] :

$$u_{nmp}^{(i)}(x_1', x_3') = H_m\left(\sqrt{\alpha_n^{(i)}} x_1'\right) H_p\left(\sqrt{\beta_n^{(i)}} x_3'\right) e^{-\frac{\alpha_n^{(i)} x_1'^2 + \beta_n^{(i)} x_3'^2}{2}} \quad (3)$$

$n=1,3,5,\dots$

$m, p = 0,1,2,3,\dots$

$H_m, H_p$  are Hermite polynomials.  $(x_1', x_3')$  is a second coordinate system obtained by a rotation of  $\psi_n^{(i)}$  about the  $x_2$  axis. This rotation is function of the mode family and of the overtone number  $n$ .



**Fig. 1: Anharmonic mode in SC-cut (C-mode, 3rd Overtone,  $m = 0, p = 1$ )**

The figure 1 shows the in-plane distribution of the vibration for the C-301 mode ( $m=0, p=1$ ) of the C-family, third overtone. The usual mode, or metrological mode, corresponds to  $m = p = 0$ . When, at least, one indice  $m$  or  $p$  is non-zero, the mode is called anharmonic. The constants  $\alpha_n$  and  $\beta_n$  govern the mode shape. They are related to the dimensions of the resonator and to the constant of dispersion  $M_n'$  and  $P_n'$  expressed in the  $(x_1', x_3')$  frame :

$$\alpha_n^{(i)2} = \frac{n^2 \pi^2 c^{-(i)}}{8Rh_0^3 M_n'^{(i)}} ; \quad \beta_n^{(i)2} = \frac{n^2 \pi^2 c^{-(i)}}{8Rh_0^3 P_n'^{(i)}} \quad (4)$$

Stevens and Tiersten [7] have given an analytical expressions for the constants  $M_n'$  and  $P_n'$ . They depend, of course, of the mode-family.

The corresponding resonant frequency of the mode is :

$$\omega_{nmp}^{(i)} = \frac{n\pi}{2h_0} \sqrt{\frac{c^{-(i)}}{\rho}} \times \quad (5)$$

$$\left[ 1 + \sqrt{\frac{2h_0}{n^2 \pi^2 R c^{-(i)}}} \left[ (2m+1) \sqrt{M_n'^{(i)}} (2p+1) \sqrt{P_n'^{(i)}} \right] \right]^{1/2}$$

The motional resistance  $R_1 = \frac{1}{QC_1\omega}$  of a resonator having two pairs of electrodes is obtained from its motional capacitance (Lewis [8], Stevens and Tiersten [7]) :

$$R_1 = \frac{n^2 \pi^2 h_0 c^{-(i)} L_{nmp}}{4\omega_{nmp}^{(i)} (e_{26}^{(i)})^2 Q^{(i)} (J_{nmp}^{(1)} \pm J_{nmp}^{(2)})^2} \quad (6)$$

where

$$J_{nmp}^{(j)} = \iint_{S_j} u_{nmp}^{(i)}(x_1, x_3) dx_1 dx_3 \quad (7)$$

and

$$L_{nmp} = \pi 2^m m! 2^p p! (\alpha_n^{(i)} \beta_n^{(i)})^{-1/2} \quad (8)$$

The (+) sign in (6) stands when the same driving voltage  $V$  is applied on both pairs of electrodes of area  $S_1$  and  $S_2$ , the (-) sign when opposite voltages  $+V$  and  $-V$  are applied on  $S_1$  and  $S_2$ .  $e_{26}^{(i)}$  take into account the direction of vibration of the mode.

$$e_{26}^{(i)} = e_{26} A_1^{(i)} + e_{22} A_2^{(i)} + e_{24} A_3^{(i)} \quad (9)$$

$Q^{(i)}$  denotes the Q factor of the mode.

### III- CHOICE OF THE MODE

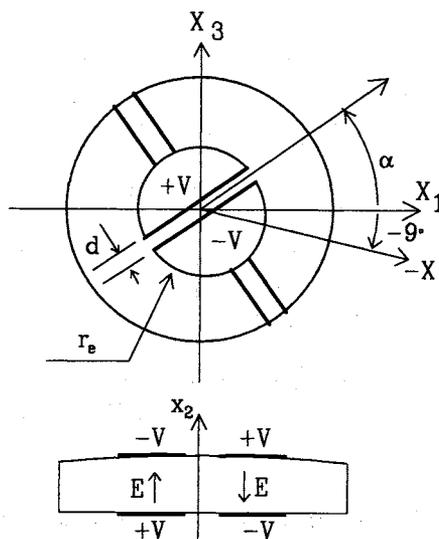
Table II gives the values of  $M_n'$  and  $P_n'$  from the fundamental mode up to the seventh overtone calculated from Steven and Tiersten's formulas [7] and the corresponding angle of orientation of the modal figure  $\psi_n^{(i)}$ .

Low values of  $M_n'$  and  $P_n'$  indicates a strong trapping. To be able to increase the motional resistance of the B-mode it is necessary to choose an overtone for which the mode-shapes are as different as possible.

**Table II : Dispersion constants (in  $10^9 \text{ N/m}^2$ ) and orientation angle of mode-shape for SC-Cut**

	Overtone n	$M'_n$	$P'_n$	$\psi'_n$
B-mode	1	61.78	130.41	$4^\circ$
	3	67.49	28.35	$32^\circ$
	5	51.27	71.62	$-37.5^\circ$
	7	66.08	48.68	$37^\circ$
C-mode	1	96.72	56.92	$6^\circ$
	3	53.14	63.29	$21^\circ$
	5	71.02	78.16	$28.5^\circ$
	7	57.68	110.74	$12^\circ$

The third overtone of the B-mode presents a higher trapping in the  $x_3$ -direction than in the  $x_1$ -direction because  $P'_3$  is the half of  $M'_3$ . For the same overtone, the trapping of the C-mode is quite the same in both directions. Furthermore, the angle of orientation of the modal figures of the two modes differs by a value of about ten degrees. So, it is possible to drive preferentially the C-301 mode by using electrodes configuration such as presented on figure 2. The electrodes are connected together to obtain the electric fields of opposite signs under each parts of them. The orientation  $\alpha$  and the dimensions of the electrodes must be determined to obtain the best ratio between the motional resistances of the B and C-modes. The x-axis indicated on figure 2 is in fact defined as the normal to the intersection between the plane of the plate and the X-face of the crystal bar from which the plate is made.

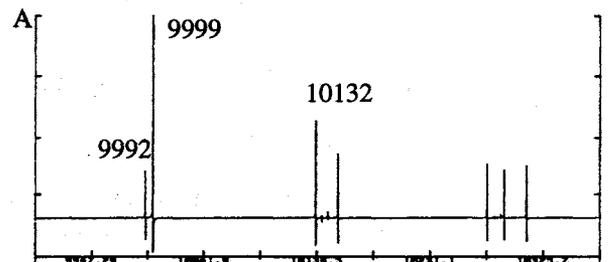


**Fig. 2 : Electrodes configuration**

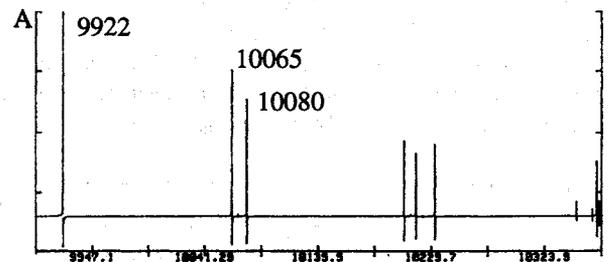
#### IV EXPERIMENTAL RESULTS

The fabricated resonators are BVA-type, SC-cut 3rd overtone. Their thickness is adjusted so that the C-301 mode has a frequency of 10 MHz. The frequency of the usual C-300 mode is then 9.922 Mhz. The active part of the resonator has a diameter of 10.2 mm. The radius of curvator of 200 mm is slightly smaller than this of a classic resonator to keep quite the same trapping.

The figure 3 shows the resonant frequencies when the electrodes are connected to drive the anti-symmetric modes. For this record, they were orientated in the -y direction ( $\alpha = 81^\circ$ ). Obviously, the C-300 usual mode, like the other symmetric modes, does not appear. When both half electrodes of a face are connected together, the symmetric modes are excited. We then observe the usual resonances for this resonator (fig. 4).



**Fig. 3 : Resonant frequencies of anti-symmetric modes**



**Fig. 4 : Resonant frequencies of symmetric modes**

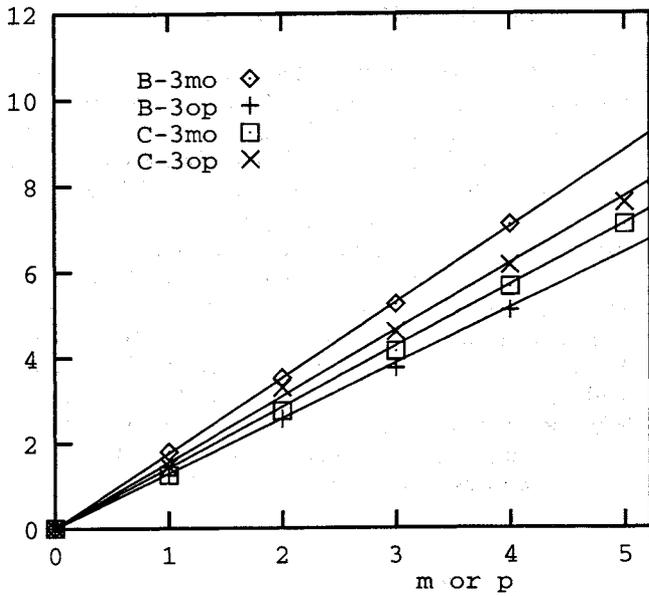
These data enable us to check the theoretical values of  $M'_3$  and  $P'_3$  for the B and C-modes. A plot of  $f_{3mp}^2$  is shown on the figure 5 for m or p equal to zero. It is in accordance with (5). In the case of the C-mode, the following values are deduced :

$$M'_3 = 57.10^9, P'_3 = 67.10^9.$$

These values are in good agreement with those of table II. For the B-mode the values are :

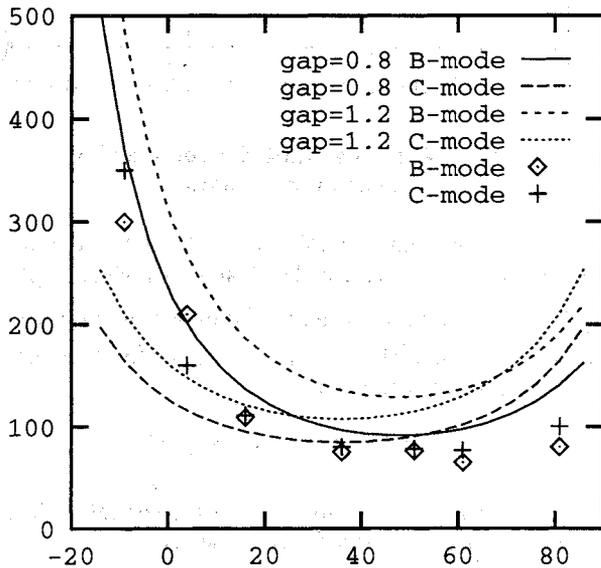
$$M'_3 = 74.10^9, P'_3 = 38.10^9$$

which are slightly different than computed values.



**Fig. 5 : Plot of  $f_{3mp}^2 - f_{300}^2$  (in MHz<sup>2</sup>) when m or p = 0**

A preceding experimental study of mode patterns in SC-cut (Dulmet and Bourquin [4]) has given the values of 36° and 48° for the orientation of the nodal lines for B and C-modes respectively. In the case of the 3-rd overtone, the motional resistances  $R_1$  have been computed with these values from (6) for the B-301 and C-301 modes. The diameter of the electrodes (fig. 2) was fixed at 6 mm and the gap between the half-electrodes at 0.8 and 1.2 mm. The used Q-factors are  $1.2 \cdot 10^6$  for the C-mode and  $1.4 \cdot 10^6$  for the B-mode. The results are plotted on the figure 6, together with experimental values obtained with a resonator having a gap of 0.8 mm.



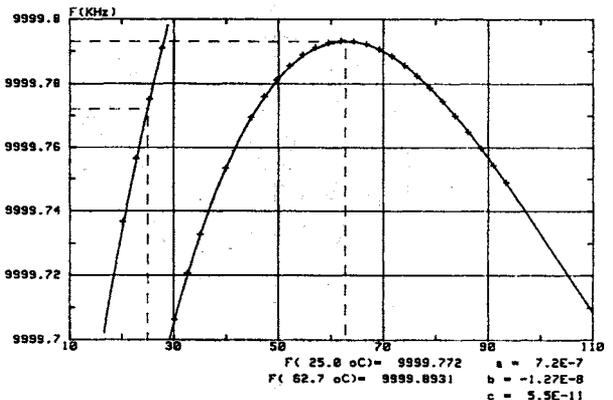
**Fig. 6 :  $R_1$  (in  $\Omega$ ) versus orientation  $\alpha$  (in degrees) of the electrodes. Experimental data are for a gap of 0.8 mm**

After several experiments we have found that  $\alpha = 36^\circ$  provides with a good compromise between optimisation of C-mode resistance and increase of B-mode one. In this case, the table III summarizes the measured motional resistances. For a gap equal to 1.2 mm the resonators have a typical resistance of 95 Ohm for the C-mode and 150 Ohm for the B-mode. This ratio is sufficient to insure a good use in an oscillator. We may note that the same resonator with conventional electrodes (diameter 4.3 mm) has resistances of 97 Ohm for the C-mode and 75 Ohm for the B-mode. An increase of the resistance ratio can be obtained by using a larger gap at the expense of an increase of the C-mode resistance.

**Table III : Motional resistances (in  $\Omega$ ) of resonators for  $\alpha = 36^\circ$**

Resonator N°	C-mode $f=9.998$ MHz	B-mode $f=10.900$ MHz	Gap (mm)
1	80	75	0.8
2	100	110	
3	85	90	
3	92	152	1.2
4	100	145	

The frequency vs. temperature curve of such a resonator operating on the C-301 mode is depicted on the figure 7. It is, of course, similar to the corresponding curve of the C-300 mode. The turnover point is located at 63°C. The same resonator, with electrodes connected to drive the C-300 mode, has then a turnover point at 71°C. This lowering of the turnover point is in accordance with the thermal behaviour predicted from the expression (5) of the plano-convexe resonator frequencies.



**Fig. 7 : Frequency vs. temperature curve for resonator operating in C-301 mode.**

## REFERENCES

- 1 **Smythe R.C.**, 1987, "An approximate expression for the motional capacitance of a lateral field resonator".  
Proc. 41st. Ann. Symp. Freq. Cont., p 311-313.
- 2 **Masiukiewicz A, Gniewinska B, Szulc W, Weiss K**, 1994, " Study of SC-cut lateral field excited resonators features", Proc. 9th EFTF, p 24-28.
- 3 **Spencer W**, 1963, "X-Ray diffraction study of acoustic mode patterns in cristalline quartz". App. Phys. Lett. 2, 7, p 133-135.
- 4 **Dulmet B, Bourquin R**, 1991, "3D analysis of thickness vibrations in doubly rotated contoured quartz plates" Proc. IEEE 1991 Ultrason. Symp., p 463-468.
- 5 **Auld B.A.**, 1990, "Acoustic fields and waves in solids" R.E. Krieger Publishing Company, Malabar Fl, USA.
- 6 **IEEE Std 176**, 1978, "Standard on piezoelectricity", IEEE New-York, USA.
- 7 **Stevens D.S., Tiersten H.F.**, 1986, "An analysis of doubly rotated quartz resonators utilizing essentially thickness modes with transverse variation". J.A.S.A. 79, 6, p 1811-1826.
- 8 **Lewis J.A.**, 1961, "The effect of driving electrode shape on the piezoelectric properties of piezoelectric crystals", B.S.T.J, Sept. 1961, p 1259-1280.

## PHASE COHERENT FREQUENCY MEASUREMENT OF VISIBLE RADIATION

H. Schnatz, B. Lipphardt, J. Helmcke, F. Riehle, G. Zinner

Physikalisch - Technische Bundesanstalt ( PTB )

Bundesallee 100

D-38116 Braunschweig

### ABSTRACT

Recently the Comité International des Poids et Mesures ( CIPM ) has recommended the frequency of the Ca- intercombination line  $^3P_1-^1S_0$  for the realization of the Metre to be  $\nu = 455\,986\,240,5$  MHz with a standard uncertainty of 200 kHz ( $\Delta\nu/\nu = 4,5 \cdot 10^{-10}$ ) [1]. Up to now this frequency has been determined by an interferometric wavelength intercomparison [2,3]. To take full advantage of its potential as a precise optical frequency / wavelength standard, we have measured this frequency with respect to the frequency of the primary standard of time and frequency, the Cs- clock. We have used the existing frequency chain at PTB to the infrared [4] and extended it to a Ca stabilized dye laser.

We have performed the first phase coherent frequency measurement of visible radiation referenced to the Cs clock and determined the absolute frequency of this transition with a relative standard uncertainty of  $9 \cdot 10^{-13}$ . The mean of the high and low- frequency recoil components of the Ca atoms stored in a magneto- optical trap is 455 986 240 493. 95 kHz with an estimated standard uncertainty of 420 Hz.

The Ca stabilized laser now represents the optical frequency standard with the lowest uncertainty being more than an order of magnitude smaller than that of the iodine stabilized HeNe laser at  $\lambda = 633$  nm. The current status and prospects will be reported at the conference.

### References

- [1] T. Quinn, *Metrologia*, **30**, 523-541,(1994)
- [2] G. Bönsch, A. Nicolaus, U.Brand, *PTB Mitteilungen* **99** 329-334, (1989)
- [3] J. Ishikawa, N. Ito and A. Morinaga, *Jpn. J. Appl. Phys.* **33**, 1652-1654, (1994)
- [4] G. Kramer, B. Lipphardt, C. O. Weiss, *Proceeding of the 1992 IEEE Frequency Control Symposium*, pp. 39-43, Hershey, IEEE Catalog No.92CH3083-3

## FM NOISE REDUCTION LIMITS OF A LASER DIODE FREQUENCY STABILIZED ON AN ATOMIC VAPOR

M. Bahoura, Ph. Laurent, G. Santarelli, S. Ghezali and A. Clairon  
BNM/Laboratoire Primaire du Temps et des Fréquences,  
Observatoire de Paris, 61, Av. de l'Observatoire, 75014 Paris, France  
E-mail: Messaoud.Bahoura@obspm.fr

N. Dimarcq  
Laboratoire de l'Horloge Atomique, Unité Propre du CNRS associée à l'Université de Paris Sud, Bat. 221,  
Université de Paris Sud, 91405 Orsay Cédex, France

### ABSTRACT

We have studied the properties of a noisy laser diode light tuned near the  $D_2$  line of Cesium (Cs) resonance transition after passing through an atomic-cesium vapour cell.

We have found that the transmitted beam acquires excess noise. This noise depends on the laser linewidth, laser detuning, intensity, and Fourier frequency.

### INTRODUCTION

Advances in diode lasers from the visible to the near infrared have opened the way to important developments in frequency standards and atomic spectroscopy. The frequency performance of such optical standards critically depends on the quality of the optical reference signal which may be an absorption signal or a laser induced fluorescence signal.

The laser frequency noise plays an important role in an optical frequency standard. It contributes to the system's short term instability directly, and it also degrades the long term stability since it may broaden the apparent linewidth in the atomic reference transition and adds excess noise to the detection signal.

Recent experiments [1,2] have found that saturated absorption is a real process for the generation of excess noise. By modelling the atomic vapour as a saturable absorber, it has been shown that a strongly saturating beam acquires less excess noise than a weaker beam for the same distance of propagation. Using a phase-diffusion model for nonsaturating laser field, McIntyre et al. [2] have shown the utility of diode-laser noise in spectroscopy. A forward four wave mixing theory [3] attributes the excess noise to the amplification of the vacuum side modes of the laser. In a study of laser induced fluorescence emitted by an atomic beam, it has been demonstrated [4] that the frequency noise of the diode laser degrades the signal-to-noise ratio of the signal with which we detect the atomic resonance and consequently the frequency stability of the locked laser.

In this paper, a semiconductor-laser noise properties of a beam propagating through a resonant medium is described by a phase-diffusion model. The main purpose

of this work is to show how the modulation frequency used in FM spectroscopy for frequency stabilisation can be adequately chosen since the excess noise extends as far as Doppler absorption profile.

Our experiment is concerned mainly with absorption in Doppler-broadened transition.

We have observed that the frequency noise inherent in a beam from a diode laser is converted into amplitude fluctuations after it passes through an atomic vapour. The measurements are performed in terms of laser linewidth, intensity, detuning, and Fourier frequency.

### THEORY

Diode lasers are inherently noisy devices. In particular, both the phase and amplitude can fluctuate and consequently there are many different statistical models to describe laser fields. In laser diodes, the amplitude fluctuations are relatively small above threshold [5]. It has been demonstrated [5] that the dominant noise mechanism in semiconductor lasers is the phase-diffusion noise caused by spontaneous emission. We have adopted the phase-diffusion model to describe the laser diode field. This model supposes that the magnitude of the field is constant and that the phase can change by any value with continuous random walk.

For our theoretical model we assume that the incident laser field has a constant electric field amplitude  $E_0$ , and a mean angular frequency  $\omega$

$$E_i(t) = \frac{1}{2} \left[ E_0 e^{-i[\omega t - kz + \varphi(t)]} + c.c \right] \quad (1)$$

where  $\varphi(t)$  is the phase, which is assumed stochastically varying, and obeys the Langevin equation

$$\begin{aligned} \frac{d\varphi(t)}{dt} &= \omega(t) \\ \frac{d\omega(t)}{dt} &= -\beta\omega(t) + f(t) \end{aligned} \quad (2)$$

where  $\omega(t)$  is the stochastic part of the laser frequency and  $f(t)$  is a random Gaussian force with correlation function

$$\langle f(t)f(t') \rangle = 2\gamma\beta^2\delta(t-t') \quad (3)$$

The laser frequency is a Gaussian random variable with the correlation function [6]

$$\langle \omega(t)\omega(t') \rangle = \gamma\beta e^{-\beta|t-t'|} \quad (4)$$

where  $\gamma\beta$  is the variance of the frequency fluctuations, and  $\beta^{-1}$  is the correlation time of the frequency fluctuations. The spectrum of the laser light with phase fluctuations is given by the Fourier transform of the correlation function

$$\langle E_i(t)E_i^*(t') \rangle = E_0^2 \exp\left[-\gamma\left[(t-t') + \frac{e^{-\beta(t-t')} - 1}{\beta}\right]\right] \quad (5)$$

The phase-diffusion model corresponds to the case of  $\beta \rightarrow \infty$  such that the frequency fluctuations become  $\delta$  correlated with a spectral density of  $2\gamma$ , giving the laser Lorentzian spectral profile.

We consider the interaction of a phase fluctuating field with an atomic vapour, which is assumed to be a two-level system (Fig. 1)

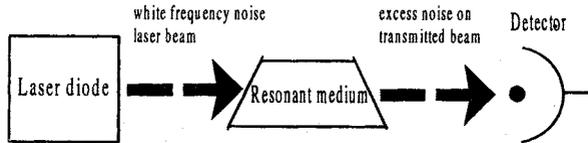


Fig. 1. Noise properties of a beam after passing through an atomic vapour

The incident field induces a macroscopic polarisation in the resonant medium

$$P(z, t) = \text{Re}\left[\Phi(z, t)e^{-i(\omega t - kz + \varphi(t))}\right] \quad (6)$$

where  $\Phi(z, t)$  is slowly varying. The induced polarisation re-radiates an em field which, in the slowly-varying field-envelope approximation, obeys the Maxwell-Bloch propagation equation

$$\left(\frac{\partial}{\partial z} + \frac{1}{c}\frac{\partial}{\partial t}\right)\mathcal{E}_r(z, t) = \frac{2\pi i\omega}{c}\Phi(z, t) \quad (7)$$

Within the resonant medium, the field  $E$  can be written in the form

$$E(z, t) = E_i(z, t) + E_r(z, t) \quad (8)$$

where

$$E_r(z, t) = \text{Re}(\mathcal{E}_r(z, t)e^{-i(\omega t - kz + \varphi(t))}) \quad (9)$$

is the field re-radiated by the induced polarisation.

According to Eq. (8), at zero we have  $E_r(0, t) = 0$ , so Eq. (9) can be expressed in integral form as

$$E_r(z, t) = \frac{2\pi i\omega}{c} \int_0^z \Phi\left(z', t + \frac{(z'-z)}{c}\right) dz' \quad (10)$$

Since  $P$  depends on  $E_r$  via the density matrix equations, the laser field after traversing the cell of length  $L$  and for small absorption is

$$E(t) = E_i(t) - \frac{\Gamma}{2} L \left\langle \int_{-\infty}^t E_i(t') \exp\left[-\left(i\Delta + \frac{\Gamma}{2}\right)(t-t')\right] dt' \right\rangle \quad (11)$$

where  $\alpha$  is the absorption coefficient,  $\Gamma$  is the natural linewidth (FWHM),  $\Delta = \omega - \omega_0 - \vec{k} \cdot \vec{v}$  is the detuning from resonance for an atom with velocity  $\vec{v}$ , and the brackets represent an average over the velocity distribution.

The incident field interferes with the emitted field to yield an heterodyne beat signal

$$I(t) = E(t)E^*(t) \quad (12)$$

The power spectrum of this heterodyne signal is obtained by a Fourier transformation of the two-time correlation function

$$I_D(\omega) = \text{Re} \int_0^{\infty} \langle I(t)I(t+\tau) \rangle e^{i\omega\tau} d\tau \quad (13)$$

The fourth-order correlation functions of Eq. (13) are well-known for the phase-diffusion fields.

Theoretical prediction we present further has been extended to the saturation case.

## EXPERIMENT

The experimental set-up used to measure the noise of the transmitted beam is shown in (Fig. 2).

The experiment is performed with two commercial laser diodes (LD), a single-mode GaAlAs diode (Hitachi HL8314E) and a DBR diode (SDL 5712H1), both emitting at 852 nm. The laser linewidth is 20 MHz for the Hitachi diode and from 3 MHz to 10 MHz for the DBR diode, varied by changing its injection current. These linewidths are measured by observation of the beat note between the output of these lasers and a narrow-linewidth (100 kHz) extended cavity laser (ECL).

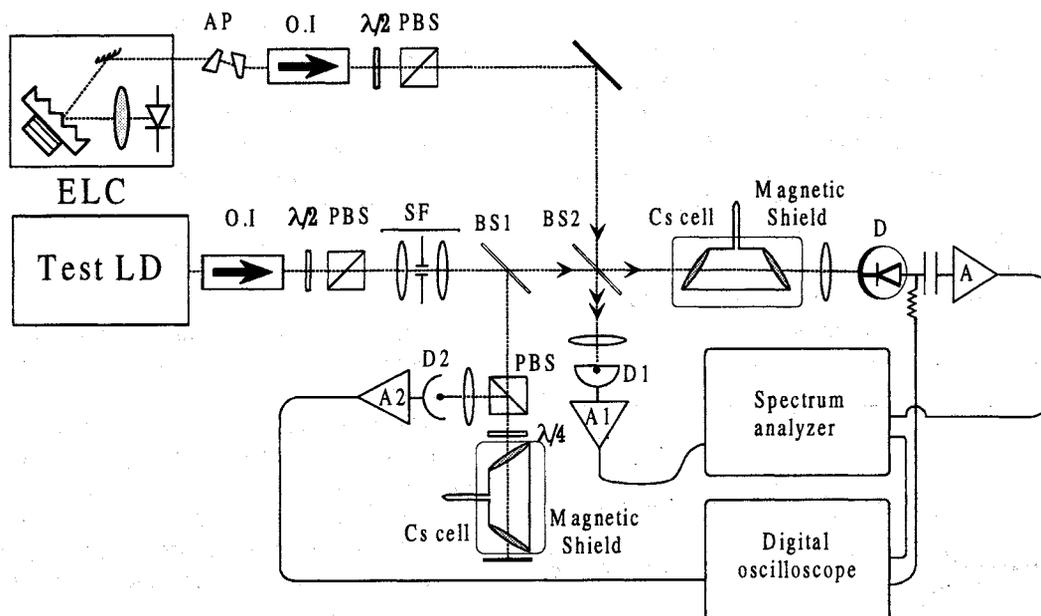


Fig.2. Experimental setup used to measure the signal-to-noise (S/N) of the transmitted laser beam. LD: laser diode, ELC: extended cavity laser, AP: anamorphic prisms, OI: optical isolator, PBS: polariser beamsplitter, SF: spatial filtering, BS: beam splitter, D: detector, A: amplifier.

The entire laser mount is placed inside two foam-lined containers which provides thermal and acoustical isolation. The output of the LD is passed through a Faraday isolator with 40 dB isolation to prevent unwanted light returning to the diode. The laser beam is linearly polarised and its intensity is controlled with a half-wave plate and a polariser beam splitter (PBS). In order to have the same Gaussian beam shape for the two laser diodes the laser beams are spatially filtered by the same set-up (SF).

A small amount (10 %) of the light emitted from the laser is reflected by a beam splitter (BS1) in order to produce saturated absorption for frequency calibration. The remaining light is transmitted through a 3 cm Cs absorption cell at room temperature. The cesium cell is surrounded by a magnetic shield in order to avoid frequency instabilities due to the fluctuations of the magnetic field.

The output of the cell is focused onto a fast (dc to 15 GHz) silicon photodiode (model Ford 4502). The DC part of the detector output is sent to a digital oscilloscope to measure the absorption signal, while the AC part, after amplification, is sent to a spectrum analyzer (HP 70004A) to record noise spectra.

Two complementary measurements are performed:

- 1) recording the noise of the transmitted beam as a function of laser detuning from resonance.
- 2) recording the noise at a fixed laser detuning.

The noise spectra  $I_D(\omega)$  are measured in terms of laser linewidth  $\gamma$ , laser intensity, detuning  $\Delta$ , and Fourier frequency  $\omega$ .

Measurements of the first type are presented in (Fig. 3), which shows the noise spectral density at two different laser linewidths under the same conditions of

absorption signal and intensity in terms of laser frequency detunings versus Fourier frequency.

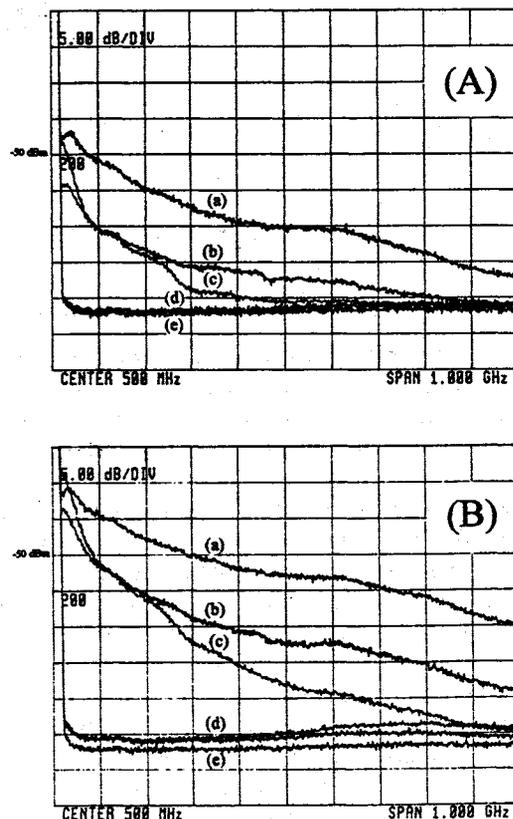


Fig. 3. Noise spectrum in Doppler absorption versus Fourier frequency for various laser frequency detunings: (a) to the blue slope side of resonance, (b) to resonance, (c) to the red slope side of resonance, and (d) off resonance. (e) is the detection noise. Laser linewidths are (A) 3 MHz, and (B) 10 MHz. For both spectra the laser intensity is  $100 \text{ mW/cm}^2$ .

The two spectra have the same behaviour but we notice that when the laser linewidth is increased, the noise increases. On the same spectrum, noises of the various detunings fall off differently, can cross each other, and depend on the Fourier frequency. The important observation is that the excess noise extends as far as Doppler absorption linewidth (500MHz).

Measurements of the second type are presented in (Fig. 4), which shows the excess noise intensity at two different Fourier frequencies under the same conditions of laser linewidth and intensity when the laser frequency is scanned over the atomic resonance.

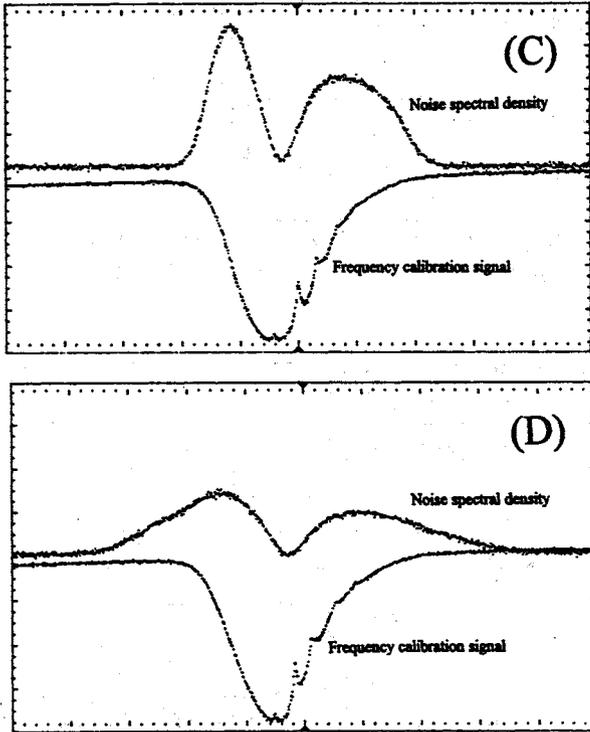


Fig. 4. Spectral density noise in Doppler absorption at (C) 20 MHz and (D) 500 MHz Fourier frequency for laser linewidth of 3 MHz and laser intensity of 1 mW/cm<sup>2</sup> versus laser detuning.

We observe that the noise exhibits a double peaked profile and it is maximum when the laser frequency is tuned to the side of the absorption profile, and is minimum when the laser is tuned to resonance. Similar spectra are predicted [7] for the case of FM spectroscopy when only the AC power of the transmitted signal, at the modulation frequency, is measured.

For high frequencies, the noisy laser spectrum can be thought of as laser field modulated by white frequency noise with continuum sidebands rather a single sinusoidal frequency in FM spectroscopy. So we have to consider the contribution of different components at different Fourier frequencies for the transmitted signal.

The asymmetry on the double peaked resonances is believed to be due to the optical pumping existing on the red side of resonance, diminishing the noise.

It is noteworthy that at resonance the noise do not dip to exact zero in contrast with pure FM spectroscopy.

This can be attributed to the correlated amplitude and phase noise in laser diode.

To compare our experimental results with the theoretical model, we have studied the dependencies of the signal-to-noise (S/N) of the resonance detection in terms of laser linewidth, intensity, and detuning versus Fourier frequency.

The signal (S) is taken as the absorption signal. The noise (N) is defined as the root square of the spectral density.

Theoretical and experimental results are each scaled to unity at their respective maximum values for comparison.

The variations of (S/N) as a function of Fourier frequency are depicted in (Fig. 5.A). for various values of the laser linewidth at a constant intensity and detuning. Theoretical predictions, with the same experimental parameters, are in good agreement with experimental results. We can notice that the (S/N) varies as  $\gamma^{-1}$  ( $\gamma$  is the laser linewidth).

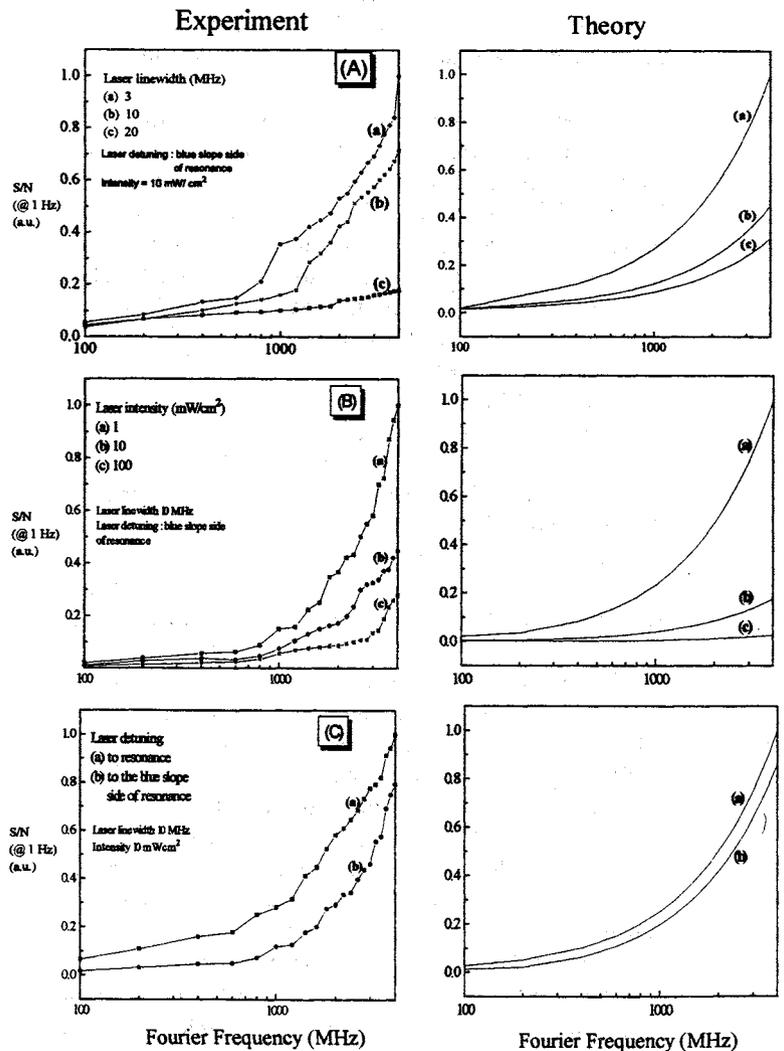


Fig. 5. Experimental measurements and theoretical predictions of the signal-to-noise of the transmitted laser beam in terms of (A) laser linewidth, (B) laser intensity, and (C) laser detuning.

(Fig. 5.B). presents the (S/N) dependence in terms of laser intensity at a fixed laser linewidth and detuning versus Fourier frequency. We notice that the (S/N) decreases when the laser intensity increases. This can be understood as with increasing laser intensity the signal saturates while the noise increases. A slight discrepancy with theoretical predictions is due to the limitations of theoretical model. These limitations come from the contribution of the amplitude noise, even small, on the noise spectra. Also from the fact that we have assumed Cs atom as a two-level system while optical pumping may modified the observed spectra.

The dependence of the (S/N) in terms of laser detuning versus Fourier frequency when the laser linewidth and intensity are constant is shown on (Fig. 5.C). The (S/N) is greater at resonance than at the side of resonance. In (Fig. 4.D), we can see a broad background noise spectral density. This noise was found to extend over the Doppler absorption profile even where there is no absorption. Then a simple interpretation that the laser noise is demodulated by the atomic resonance is no longer valid.

From these results we can see that at typical experimental conditions using DBR laser diode ( $\gamma_{\text{DBR}} \ll \Gamma$ ) and for saturation intensity, the excess noise is not negligible.

One do not expect to improve the signal-to-noise by taking a long vapour cell. This is because both the signal and the noise are increased when a long vapour cell is used.

We have used saturated absorption to obtain a narrow resonance feature. A typical probe spectral density noise is shown on (Fig. 6)

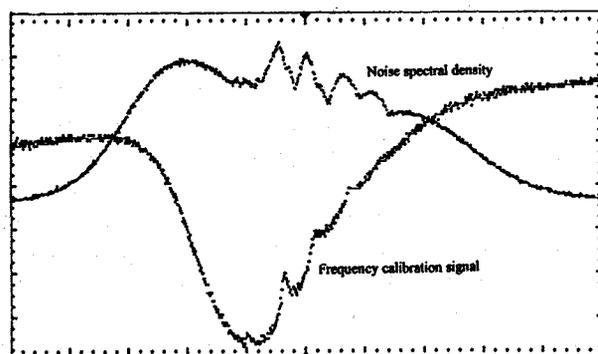


Fig. 6. Noise spectral density in saturated absorption. The lower trace is standard saturated-absorption. The upper trace is probe spectral density noise at Fourier frequency of 50 MHz.

We notice that even with sub-Doppler resonances the excess noise exhibits a double peaked profiles around these resonances on Doppler absorption background. This may compromise the laser frequency locking on these absorption profiles.

## CONCLUSION

In summary, a significant amount of excess noise is acquired by a laser diode beam after passing through atomic-cesium vapor. The nature of this noise depends, besides technical noise of the laser, on the dynamics of the resonant medium. Since the excess noise is smaller at high Fourier frequencies, the modulation frequency of FM spectroscopy for laser stabilisation, has to be greater than Doppler linewidth.

## ACKNOWLEDGMENTS

This work is supported by the Bureau National de Metrologie (BNM). The authors want to thank the Centre National d'Etudes Spatiales (CNES) via the Pharao project, and the Conseil Général de la Région Ile de France through the Sesame programme. Finally we wish to thank S. N. Lea and D. Touahri for helpful discussion in the preparation of this manuscript.

## REFERENCES

1. Varada V. Iruvanti, R. W. Boyd, 1995, "Noise properties of a beam propagating through an atomic vapour", *Opt. Comm.*, **118**, 46-50.
2. D.H. McIntyre, C.E. Fairchild, J. Cooper, and R. Walser, 1993, "Diode-laser noise spectroscopy of rubidium", *Opt. Lett.*, **18**, 1816-1818.
3. W.V. Davis, M. Kauranen, E.M. Nagasako, R. J. Gehr, A.L. Gaeta, and R.W. Boyd, 1995, "Excess noise acquired by a laser beam after propagating through an atomic-potassium vapour", *Phys. Rev. A*, **51**, 4152-4159.
4. N. Dimarcq, V. Giordano, P. Cerez, 1994, "Statistical properties of laser-induced fluorescence signals", *Appl. Phys. B*, **59**, 135-145.
5. C.H. Henry, 1982, "Theory of the linewidth of semiconductor lasers", *IEEE J. Quant. Electron.*, **QE-18**, 259-264.
6. S.N. Dixit, P. Zoller, and P. Lambropoulos, 1980, "Non-Lorentzian laser line shapes and the reversed peak asymmetry in double optical resonance", *Phys.Rev. A*, **21**, 1289-1296.
7. G.C. Bjørklund, M.D. Levenson, W. Lenth, and C. Ortiz, 1983, "Frequency Modulation (FM) Spectroscopy", *Appl. Phys.*, **B32**, 145-152.

## FREQUENCY MEASUREMENTS ON HI AND THE Kr LINE AT 1.548 $\mu\text{m}$

D. J. E. Knight<sup>a</sup>

*DK Research, 110 Strawberry Vale, Twickenham, Middx., TW1 4SH, UK*  
and

M. Zucco<sup>b</sup>, K. I. Pharaoh and D. A. Humphreys

*National Physical Laboratory, Teddington, Middx., TW11 0LW, UK*

<sup>a</sup> present address: Dipartimento di Elettronica, Politecnico di Torino, Corso Duca degli Abruzzi 24, I-10129 Torino, Italy.

<sup>b</sup> when this work was carried out, M Zucco was a guest in NPL from the Istituto di Metrologia "G Colonnetti", Torino, Italy.

### ABSTRACT

Spectral lines in the 1.55  $\mu\text{m}$  region are of interest as frequency standards for optical communication because they are in the gain region of the erbium-fibre amplifier. Using diode lasers locked to the respective spectral lines, heterodyne frequency measurements have been made between the CO R(20) line at 1.56  $\mu\text{m}$  and the HI P(3) line, for which  $f_{\text{CO}} - f_{\text{HI}}$  was 9277.0(3.6) MHz, and between the Kr optogalvanic line at 1.548  $\mu\text{m}$  and the HI R(0) 5/2-5/2 component, for which  $f_{\text{Kr}} - f_{\text{HI}}$  was 25014.5(6.4) MHz. The HI R(0) line was split by nuclear quadrupole effects into two components, the lower frequency one containing only the 5/2-5/2 transition. The latter was used for measurement, and the splitting between the two components was 438(19) MHz. A result for the Kr line frequency obtained via lambda-meter ratios against the CO R(20) line was 193 686 109 (45) MHz. The value used here for the CO R(20) locked-laser frequency was 192 068 038(31) MHz, and was derived from earlier measurements against the CO R(21) line<sup>2</sup> and from additional lambda-meter measurements.

### 1. INTRODUCTION

Spectral lines near 1.55  $\mu\text{m}$  are of interest as potential frequency standards for optical fibre communication systems. The spectrum of HI in the 1.5- $\mu\text{m}$  region has been surveyed with a diode-laser spectrometer and the centre frequencies measured by lambda meter<sup>1</sup>. Two of the HI lines are close to lines respectively in CO and Kr, of which the absolute frequencies have been measured at the National Physical Laboratory (NPL)<sup>2,3</sup>. This paper reports heterodyne frequency measurements between the NPL-measured lines and the two nearby HI lines, made with a diode laser spectrometer at NPL. The objectives were (a) to measure more precisely the frequencies of the coinciding HI lines, and (b) to improve frequency

measurement across the 1.5  $\mu\text{m}$  optical-fibre communication band, from the already-measured CO R(21) line at 1.560500  $\mu\text{m}$ <sup>2,4</sup> to the Kr optogalvanic line at 1.548  $\mu\text{m}$ , which is the subject of research<sup>5,6</sup> and of a submission to the ITU<sup>7</sup>. Both these frequencies fall within the amplification band of the Er-fibre amplifier, which at present is a key element of coherent communication systems. HI has a strong, well-spaced and regular spectrum in the 1.5  $\mu\text{m}$  region<sup>8</sup>, comprising the lines from the 3-0 roto-vibrational overtone bands<sup>1,9</sup>. There is apparently a practical difficulty in that HI dissociates at low pressures in gas cells over a period of weeks to months<sup>10</sup>.

### 2. APPARATUS

The apparatus for these experiments was similar to that developed in the initial diode-laser spectrometer<sup>4</sup>, and the CO-locked laser system has been described<sup>2</sup>. There are two independent stabilised diode laser systems, which can be connected together to a photodiode for heterodyne mixing, or, individually, to a vacuum lambda meter for wavelength measurement. The connections are in single-mode fibre, except for free-space beams through the gas cells. Figure 1 shows schematically the connections set up for the first measurement. Heterodyne measurements were made using an InGaAs/InP photodiode of bandwidth about 20 GHz, and extended-cavity diode lasers, one being locked to an NPL-measured spectral line and the other to the HI line.

The lasers were locked to the lines by a first-derivative method similar to that described before<sup>2</sup>. A new NPL-designed lock box with 500 Hz modulation frequency was used to lock to CO. The earlier system with 87 Hz modulation was used to lock to HI.

In the CO system, the laser mode tuning range without mode-hop was only a few linewidths, as a result of a failure of 1 of 3 piezoelectric transducers (PZTs) on the

grating. Careful mechanical and electrical adjustment was needed to compensate for this. Near the mode hops, which were typically of 20 or 100 nm, the intensity of the laser varied more strongly. In the HI system, the laser used had a flat power output with frequency, but background slope occurred in early use of the HI cell, possibly from effects in the non-optically-flat ends of the cell. This was much reduced by using a lens in front of the detector. To obtain adequate absorption the cell was finally set up for 6 passes before reaching the detector, the length being 10 cm and the pressure 30 torr.

The krypton-stabilized laser system was located 90 m away from the frequency mixing laboratory, and was connected via a single-mode fibre about 100 m long. An extended-cavity oscillator without amplifier was used to interrogate a commercial sealed hollow-cathode dc discharge tube, focussing the beam through the hollow cathode and anode on the cell optical axis. The tube was filled with pure krypton at 4.5 torr, and was operated at a low current, about 20  $\mu$ A, to optimise the optogalvanic signal. The laser was locked to the Kr transition via the optogalvanic signal by a first-derivative system. The lock loop did not at the time of the HI measurements have an integrator, so that the error signal was monitored to determine the oscillator offset from the peak. Good temperature control in the room ensured adequate stability for the measurements. The FWHM of the Kr optogalvanic signal was estimated from line plots to be about 650(100) MHz.

### 3. METHOD

The laboratory frequency measurements were made in a period of about three weeks. Initial measurements were made on the difference frequency between the CO 3-0 R(20) line and the HI 3-0 P(3) line, both of which were single and reasonably symmetrical. The scheme of the apparatus is shown in Figure 1. However, the HI 3-0 R(0) line near the Kr frequency proved to be split into two components. This was similar to that found by Strow for the 1-0 P(1) line at 4.5  $\mu$ m<sup>9</sup>, which he attributed to nuclear hyperfine quadrupole structure. The weaker component was a symmetrical line, identified

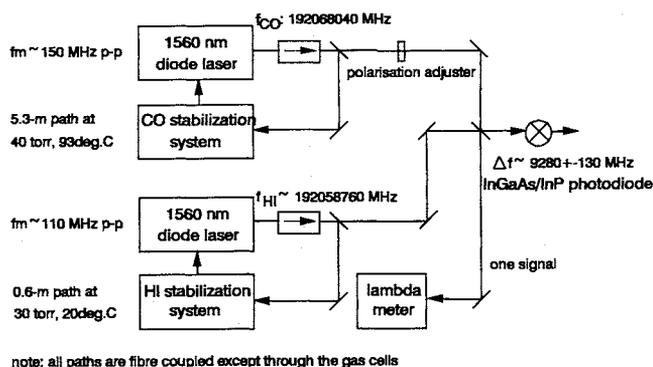


Figure 1: Scheme of CO R(20) - HI P(3) heterodyne frequency measurement

using Strow's work as containing only the 5/2-5/2 transition, and the stronger component was asymmetrical and contained the other two unresolved transitions. The weak single line was used for all measurements, apart from measurements of the splitting itself, although its dip was only about 1.7%, compared with 2.3% for the stronger component and with 6.3% for the P(3) line.

To check for background slope & other offsets, first-derivative, as well as dip, graphical plots were taken for the HI lines during measurement. Dip plots were taken for the CO line, and signal plots and lock monitor records were taken for the Kr system. Fabry-Perot fringes, with about 40 MHz period, were encountered, particularly on the first-derivative plots. These arose from reflections in the fibre system, and were minimised as far as possible by cleaning the fibre ends at the connectors.

The heterodyne beats were measured on a spectrum analyser, in the presence of up to 260-MHz peak-peak fm on the signal, by the "max-hold" method developed for the CO R(21) line frequency measurement<sup>2</sup>. The edge frequencies of the approximately square-edged beat scan were recorded after 10s, and again after 45s, the mean centre comprising 1 result. System checks were made at about every 10 readings. FET preamplifiers were used to overcome the noise figure of the spectrum analyser. A model HP 8566B (0-22 GHz) was used at 9 GHz, and a 7000-series HP system was used at 25 GHz. The spectrum analyser was typically set up with a 500 MHz span, 1-MHz rf bandwidth, 0.1-MHz video bandwidth and a 20-ms scan time, accumulating the scans under "max hold". The signal was displayed at 10 dB/division, and had a typical signal/noise ratio of 14-20 dB. If fallen the beat s/n could usually be restored by rotating the polarisation of one optical input.

Lambda-meter measurements were made with the Burleigh vacuum lambda meter, which has been investigated in previous experiments<sup>2,4</sup>. Repeated readings of the last digit in the picometre place were averaged to interpolate the relative frequency to about 30 MHz. Measurements in turn on two 1.55-1.56  $\mu$ m lines were used to obtain a ratio between the two frequencies, which then allowed one frequency to be measured from the other, amounting to a difference-frequency measurement in which the systematic offsets largely cancel. This cancellation should be more complete, the closer the frequencies, as in the difference between the peaks of the split HI R(0) line. The internal 633-nm He-Ne laser reference was thus used only as a short-term transfer standard.

### 4. RESULTS

For the frequency difference between the CO R(20) line and the HI P(3) line, see Figure 1, data was taken over

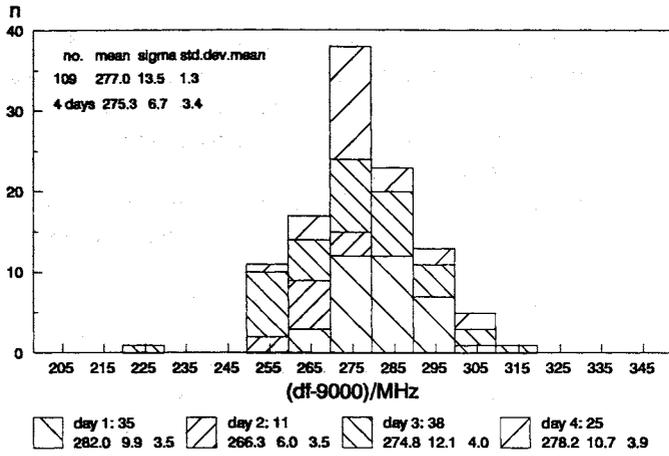


Figure 2: CO - HI difference frequency measurements at 192.06 THz

four days. The results, corrected for background-slope offsets, measured from plots taken during the runs, are shown as a composite histogram in Figure 2. Five measurements were rejected from a low group on day 3, as being outside the set after correction, and these corresponded to irregular behaviour of the CO laser before and after the low group of measurements. The corrected histogram, is approximately Gaussian, and it is valid to adopt the standard error of the mean of the 109 results, 1.3 MHz, as the statistical uncertainty, a reduction of a factor 8 on the uncorrected histogram. The corrections on two of the days were of the order of 20 MHz. The effect of correction on the final mean result was a shift of about 10 MHz.

The measurements of the splitting of the HI R(0) line at 1.548  $\mu\text{m}$  are shown in Table 1, together with the statistical part of the uncertainties. For these, the same laser was locked to each component of the line in turn, and differential measurements were made, either by lambda meter, or by heterodyning against the laser locked to the Kr line, see Figure 3 and below. Offsets

TABLE 1- HI R(0)  $f_{\text{min}}-f_{5/2-5/2}$  at 1.548  $\mu\text{m}$ , 193.661 THz: corrections and results

All entries in MHz; uncertainties are statistical components; HI R(0) splitting of peaks.

Day	no.data	type	uncorrected mean	laser differential offsets			corrected mean
				HI b/slope	Kr drift	total	
a	b	c	d	e	f	g=c+f	h=c-g
1	3	$\lambda^*$	494(34)	+6(13)	n/a	+6(13)	488(36)
2	2	$\lambda$	431(17)	+1.4(2)	n/a	+1.4(2)	430(17)
3	1	$f^\#$	425.2(9.8)	-1(1)	-1.5(1)	-2.5(2)	427.7(10)
3	1	f	446.3(8.3)	-1(2)	+3(2)	+2(3)	444.3(9)
3	2	$\lambda$	444(18)	-1(2)	n/a	-1(2)	445(18)
mean of 4 uncorrected means, rejecting the first n=4†						437(9)	
weighted mean of corrected means						n=5	438.3(5.9)

Notes

\* : about 40 readings of  $\lambda$ -meter final digit, in pm position, averaged on each peak;  
 # : groups of 6 heterodyne measurements on each peak;  
 † : 6-pass, with detector lens on HI cell, first result was 3-pass.

attributable to the laser and lock loop thus tended to cancel. The measurements were spread over three days, and comprise 3 sets taken by the lambda-meter method, and 2 sets taken by the heterodyne method. The HI differential background slope correction was taken from graph plots across the HI absorption dips. The Kr line offset, for the heterodyne results, was taken from records of the lock-loop feedback signal. The first result has a larger uncertainty and was taken whilst setting up. Using uncorrected results, the best value is obtained from the unweighted mean of the 4 later measurements, 437(9) MHz. Note that the lambda-meter and frequency methods agree. After correcting all results, the best combination is the weighted mean of all 5 results: 438.3(5.9) MHz. The agreement of the two means shows that the net effect of correction is small, except to reduce the uncertainty.

The 25 GHz difference frequency between the HI R(0) 5/2-5/2 component and the Kr line at 1.548  $\mu\text{m}$  was measured with the apparatus shown in Figure 3. Data was taken on 3 days, and the typical signal/noise ratio was 14dB. The corrected results are shown as a histogram in Figure 4. Each day's data is corrected for background slope offset of the HI locked laser, and for lock-loop offset of the krypton-locked laser. The total

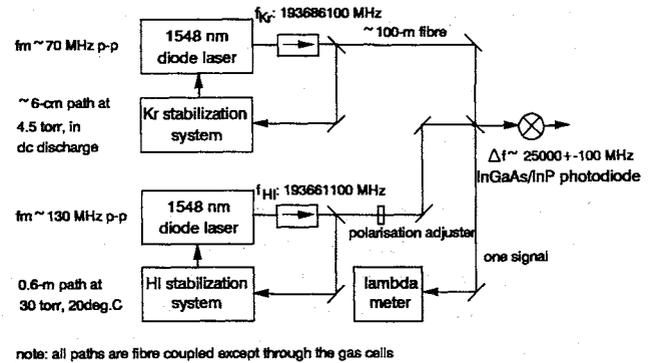


Figure 3: Scheme of Kr (o.g.) - HI R(0) heterodyne frequency measurement

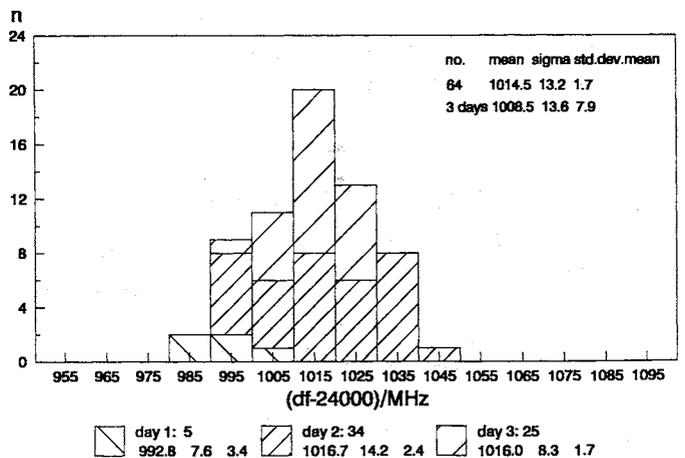


Figure 4: Kr - HI difference frequency measurement at 193.6 THz

correction is of the order of 10 MHz, half that of the  $f_{\text{CO}}-f_{\text{HI}}$  beat, but it brings the 3 daily means into coincidence and provides a near-Gaussian histogram shown in Figure 4. The mean and standard error of the mean of the 64 result histogram, 25014.5(1.7) MHz, provide the result and statistical uncertainty respectively.

### Uncertainty budgets

The statistical and estimated uncertainties for the various measurements are gathered into Table 2. The statistical, Type A, components are already discussed. The estimated, Type B, components arise broadly from the overall contributions of corrections applied already, before compiling the days' results into the histograms and table. The estimated components are 1-sigma equivalent values (uncertainty  $\pm 30\%$ ) so that the uncertainties can be combined as root sum of squares. Further contributions are given for the lambda-meter measurements, which are described below.

TABLE 2- frequency measurement uncertainty contributions

Meas.t	Type of uncertainty	contribution	amount: 1 $\sigma$ / $\pm$ MHz
$f_{\text{CO}}-f_{\text{HI}}$	A, statistical	A: corrected histogram, n=109	1.3
	B, estimated	B1: b/ground slope, pressure	3
		B2: electronic offsets, shifts	1.5
		Total B:	3.4
<b>Total (A<sup>2</sup> + B<sup>2</sup>)<sup>1/2</sup>:</b>			<b>3.6</b>
$f_{\text{HI}}-f_{\text{Kr}}$	A, statistical	A: corrected data, 5 meas.ts	5.9
	B, estimated	B1: fibre Fabry-Perot fringes	5
		B2: main peak asym.: 5%FWHM	17
		Total B:	17.7
<b>Total (A<sup>2</sup> + B<sup>2</sup>)<sup>1/2</sup>:</b>			<b>18.7</b>
$f_{\text{Kr}}-f_{\text{HI}}$	A, statistical	A: corrected histogram, n=64	1.7
	B, estimated	B1: laser offset & b/slope, press.	5.0
		B2: HI fibre F-P fringes	3.5
		Total B:	6.1
<b>Total (A<sup>2</sup> + B<sup>2</sup>)<sup>1/2</sup>:</b>			<b>6.4</b>
$\lambda$ meter(typ.)	A, statistical	A1: 2-meas.t statistics	18
	B, estimated	B1: b/slope & F-Perot fringes	14
		B2: differential $\lambda$ error	25
		Total B:	29
<b>Total (A<sup>2</sup> + B<sup>2</sup>)<sup>1/2</sup>:</b>			<b>34</b>

For the HI R(0) line splitting, apart from an allowance of 5 MHz for Fabry-Perot fringe effects that could have affected all results, a large allowance is for the "main" R(0) component asymmetry, which at 5% of FWHM amounts to 17 MHz. If the lineshape can be modelled the asymmetry allowance is not required, and the measurement to the peak of the "main" component has a total uncertainty of  $\pm 7.7$  MHz.

TABLE 3- summary of absolute frequencies and links

Measurements using locked lasers, heterodyne beats and  $\lambda$  meter

line	$\lambda/\mu\text{m}$	link type	value	from (laser)	f/MHz
Kr (o.g.)	1.547826	f, $\lambda$	<b>25015(6) MHz</b>	HI R(0) 5/2-5/2	193 686 109 (45)
HI R0main	1.548023	f, $\lambda$	<b>438(19) MHz</b>	HI R(0) 5/2-5/2	193 661 532 (50)
HI 5/2-5/2	1.548027	$\lambda$	<b>1.00829427(21)#</b>	CO R(20)	193 661 102 (59)
		$\lambda$	<b>1.00834289(9)#</b>	HI P(3)	193 661 086 (47)
		$\lambda$	<b>mean<math>\dagger</math></b>		<b>193 661 094 (45)*</b>
CO R(21)	1.560501	f	<b>f chain</b>	Rb d/f	192 113 068 (4)
CO R(20)	1.560867	$\lambda$	2 sets	CO R(21)	<b>192 068 038 (31)*</b>
HI P(3)	1.560941	f, $\lambda$	<b>9277(4) MHz</b>	CO R(20)	192 058 761 (31)

### Notes

- The entries are in order of absolute frequency, the highest at the top.
- The uncertainties of the measurements are explained in Table 2. Uncertainties of the steps to the derived absolute frequencies are combined as the root sum of squares.
- The spectral line parameters for these frequency measurements are collected together in Table 4.

#: Statistical uncertainty component, of ratio of 2 pairs of measurements. The estimated component of 29 MHz, (lambda meter, Table 2) and the CO R(20) uncertainty, 31 MHz, are summed with it to obtain the uncertainty at the right.

\*: Main  $\lambda$ -meter measurements from which the derived absolute frequencies in the table are taken.

$\dagger$ : Arithmetic mean, with statistical component of uncertainty usually appropriate to the weighted mean. The weighted mean is not used because the time and route separation of the two results averages over several types of systematic error, and weighting by the statistical uncertainty would defeat this. The same procedure was used to combine the two measurements of the CO R(20) line.

For the lambda-meter measurements the numbers entered are appropriate to differential measurement of frequency between the 193.7-THz Kr and the 192.1-THz CO frequency regions, taking readings first of one locked laser and then of the other, within a few minutes, under stable laboratory conditions. The statistical component, for two such measurements, is similar to the rounding uncertainty in interpolating the 7th decimal place from repeated readings of the 6 decimal place display, given the typical number of readings taken, and is thus a lower limit. The first estimated component of uncertainty, B1, refers to the typical net laser offset from line centre of the two locked lasers, from background-slope or lock-loop offsets or from the effect of Fabry-Perot fringes. Corrections for these effects were not applied. The remaining component, B2, is to allow for wavelength-dependent changes to the wavefront and pointing of the 1.5- $\mu\text{m}$  beam in the lambda meter as a result of (i) refractive index changes in the launching optics, and (ii) diffractive changes when different wavelengths pass through the fixed apertures.

### Absolute frequencies

The best absolute frequencies from the above heterodyne measurements, combined with lambda-meter measurements, all based on the CO R(21) frequency

measurement<sup>2</sup>, are shown in Table 3. The gas parameters for which the measurements apply, e.g. pressure and temperature, are gathered into Table 4. The uncertainties of the absolute frequencies, at the right hand side of Table 3, include both statistical and estimated components, and are largely made up of lambda-meter net contributions in the measurement chain.

TABLE 4- spectral line parameters

as appropriate to the frequency measurements in Table 3

line	dip /%	FWHM /MHz	fm p-p /MHz	path /m	P /torr	T /°C	notes
Kr (o.g.)	100*	650(100)	70	0.06	4.5	high dedisch. 20µA	
HI R(0)main	2.3(1)	320(30)	130	0.57(6 pass)+	30	20	asymmetric
HI 5/2-5/2	1.7(2)	240(60)	130	0.57(6 pass)	30	20	single
CO R(21)#	7.4(4)	585(40)	150	5.3 (2 pass)	40	94	symmetric
CO R(20)	9.3(8)	520(40)	150	5.3 (2 pass)	40	92	symmetric
HI P(3)	6.5(2)	290(60)	110	0.57(6 pass)	30	20	single

#### Notes

1. The uncertainties of the dip and FWHM are statistical, except for Kr, which includes an estimated component.

The uncertainty of the fm p-p value and of the gas pressure and temperature is about 10%.

\*: This optogalvanic signal is a Gaussian, Doppler-limited, peak, with little background signal; the path in the cell is given, but the part generating the signal is that through the hollow cathode where the discharge ionisation takes place.

†: 10-cm overall gas cell: the estimated single gas path is 95(2) mm.

#: data from reference [2].

The CO R(20) line was measured by lambda meter from the CO R(21) line<sup>2</sup>. A value<sup>2</sup> appropriate to 40 torr pressure, 192 068 043 (21,30) MHz, was combined with measurements taken 4 months later during this work, comprising 5 and 6 measurements taken on the respective lines, which had a mean result of 192 068 032 (9,30) MHz. These are combined to give the frequency 192 068 038 (8.3,30) MHz in Table 3. The uncertainties are shown here with separated (statistical, estimated) components. The R(20) measurement lies 1-sigma from the value given by the CO spectroscopic constants<sup>11</sup>: 192 068 079.85(0.4) MHz, -10(2.5) MHz at 40 torr, although our R(21) and R(22) results<sup>2</sup>, the latter obtained by lambda meter, are much closer. The reason is not known.

The frequency of the natural-isotope Kr optogalvanic line frequency obtained in this work, Table 3, is  $f_{Kr} = 193\ 686\ 109\ (45)\ \text{MHz}$ . This may be combined with the value obtained independently against the CO R(21) line by 4-laser mixing in a MIM diode<sup>2</sup>, 193 686 166 (60) MHz, to obtain a more precise value:-

$$f_{Kr} = 193\ 686\ 138\ (40)\ \text{MHz}.$$

## 5. Discussion and conclusion

The absolute frequencies of the HI lines are improved by an order of magnitude over the previous work<sup>1</sup>, and the splitting of the 3-0 R(0) line, compared with the splitting observed by Strow for the 1-0 P(1) line<sup>9</sup>, gives new information<sup>12</sup> on the vibrational dependence of the nuclear quadrupole constant of HI. This is because the P(1) line has the splitting of the ground state and the R(0) line has that for the upper,  $v=3$ , state, the splitting decreasing rapidly as J increases. The HI lines may be useful as laboratory standards, but may not be useful in communication systems because of decay over weeks of the absorption in the gas cells, thought to result from dissociation of the molecule. The uncertainty of the Kr frequency is reduced to  $\pm 40\ \text{MHz}$ .

## ACKNOWLEDGEMENTS

The authors thank F. Bertinotto for initiating the collaboration and for the loan of HI gas cells, and thank A. G. Maki for advice on the spectra of CO and HI. This work was carried out under a programme of the National Measurement System Policy Unit of the United Kingdom Department of Trade and Industry. DK thanks the Dipartimento di Elettronica of the Politecnico di Torino for facilities to prepare this paper.

## REFERENCES

1. F. Bertinotto, P. Gambini, R. Lano and M. Puleo, "Stabilisation of the emission frequency of 1.54  $\mu\text{m}$  DFB laser diodes to hydrogen iodide." *IEEE Photon. Technol. Lett.*, 4(4), 472-474 April (1993).
2. D. J. E. Knight, K. I. Pharaoh and D. A. Humphreys, "Absolute frequency measurement of the 3-0 R(21) transition of CO at 1.5605  $\mu\text{m}$  for optical communication standards." Proc. EFTF 95, Besançon, France, March 8-10, 1995.
3. D. J. E. Knight, K. I. Pharaoh and M. Zucco, "Frequency mixing to the 4th order in metal-insulator-metal diodes at 193 THz, 1.55  $\mu\text{m}$ ". Proc. Fifth Symposium on Frequency Standards and Metrology, Woods Hole, Massachusetts, USA, 15-19 October 1995.
4. D. J. E. Knight, K. I. Pharaoh, G. P. Barwood, D. A. Humphreys, C. J. Hodges, M. Lawrence and K. Cameron, "Frequency standards at 1.523  $\mu\text{m}$  based on <sup>20</sup>Ne and at 1.56  $\mu\text{m}$  derived from the 0.780  $\mu\text{m}$  Rb frequency standard." In Y. C. Chung, Ed., *Frequency-stabilized lasers and their applications*, Boston, Massachusetts, 16-18 November, 1992, Proc. SPIE, 1837, 115-123 (1993).

5. Y. C. Chung, R. M. Derosier, H. M. Presby, C. A. Burrus, Y. Akai and N. Masuda, "A 1.5  $\mu\text{m}$  package frequency-locked with a novel miniature discharge lamp." *IEEE Photon. Technol. Lett.*, **3**(9), 841-844 September (1991).
6. U. H. P. Fischer and C. von Helmolt, "Saturation and isotopic shift of the Kr 84 excited-state transition at 1547.825 nm." *IEEE Photon. Technol. Lett.*, **7**(1) 65-67 January (1995).
7. International Telecommunications Union, Telecommunication Standardization Sector, Study Group 15, contribution 45: "Wavelength division multiplexing in terrestrial line systems using OFAs." July 1993. Note: an alternative submission from Japan suggested a wavelength/frequency near 1.552  $\mu\text{m}$  (H Iwaoka, private communication, January 1995). The outcome is believed to be a recommended anchor point at 193.1 THz, 1.5525  $\mu\text{m}$ , allowing use of alternative reference lines including Kr (M Kihara, private communication, March 1996).
8. J. H. Park, L. S. Rothman, C. P. Rinsland, H. M. Pickett, D. J. Richardson and J. S. Namkung, "Atlas of absorption lines from 0 to 17900  $\text{cm}^{-1}$ ." *NASA Reference Publication* 1188. Washington D. C.: NASA, September 1987, 197pp..
9. L. L. Strow, "Observation of nuclear quadrupole hyperfine structure in the infrared spectrum of hydrogen iodide using a tunable-diode laser." *Optics Lett.*, **5**(4), 166-168 April (1980).
10. F. Bertinotto, private communication, July 1994.
11. A. G. Maki, "Infrared and far-infrared absorption frequency standards", in D. R. Lide and H. P. R. Frederikse, Eds., *CRC Handbook of Chemistry and Physics*, 75th Edition, CRC Press, Boca-Raton 1994, pp **10-319 - 10-323**.
12. A. G. Maki, private communication, 27 February 1995.

# HIGH STABILITY LASER FREQUENCY REFERENCES AT 1.5 $\mu\text{m}$ USING MOLECULAR LINES : STATE OF THE ART AND PROSPECTS.

M. de LABACHELERIE\*, K. NAKAGAWA\*\*, Y. AWAJI, M. KOUROGI and M. OHTSU

\* LPMO/CNRS, Université de Franche-Comté, 32 av. de l'Observatoire, 25044 Besançon Cedex, France

\*\* Tokyo Institute of Polytechnics, 1583 Iiyama, Atsugi 243-032 - Japan  
Tokyo Institute of Technology, Nagatsuta, Midori-ku, Yokohama 227 - Japan

## INTRODUCTION

In the field of high resolution laser frequency metrology, the ability to measure a laser frequency with a high accuracy mainly depends on the existence of a laser frequency reference with a better accuracy, close to the unknown frequency. The reason is that the most precise measurement techniques - based on the frequency counting of the beat between the unknown and the reference frequency - are limited to a frequency difference of several Gigahertz by the available electronic counting equipment. For this reason, it is necessary to generate laser waves with a very accurate frequency, as close as possible to the frequency whose measurement is needed. It is therefore easy to measure with a  $\approx 100$  kHz accuracy laser frequencies that are located in the vicinity of well-known calibrated references such as He-Ne laser<sup>1</sup>, or Rb stabilized laser diodes<sup>2</sup>. From these basic references, it is possible to generate, by non-linear optical frequency conversion<sup>3</sup>, some other references with almost the same accuracy. After that, it becomes possible to measure precisely the frequencies that are located in a narrow frequency band centered on the "converted" reference, which is usually quite far from the initial one. Recently, a considerable breakthrough has been made with the use of frequency comb generators<sup>4</sup> which allow to generate, from a stable reference, hundreds of related reference frequencies with almost the same accuracy, that extends over several terahertz around the initial reference : these devices allow to extend by more than two orders of magnitude the frequency band on which a precise comparison with the reference can be made.

In the optical fiber communication bands (around 1.3  $\mu\text{m}$  or 1.5  $\mu\text{m}$ ), the development of an accurate frequency metrology is particularly useful to perform frequency measurements that could be needed for Wavelength Dense Multiplexing applications<sup>5</sup>. However, for a long time, the most precise frequency measurement tools in this wavelength range were wavemeters which had an accuracy of only several Megahertz. The reason was that most of the available frequency references existing in these bands were not accurate enough to provide a better measurement accuracy. Moreover, the small number of potentially accurate frequency markers would not have allowed to

measure frequencies in the whole optical communication bands.

Our recent development<sup>6</sup> of ultra-narrow molecular frequency references at 1.5  $\mu\text{m}$  provides a new method to measure accurately optical frequencies on a wide frequency band. The purpose of the present paper is to review the work done on this subject, to add some technical details, and to give our opinion on the possible improvements and applications of this work.

## NARROW 1.5 $\mu\text{m}$ MOLECULAR LINES

### Detection of Saturated-Absorption Lines

The main advantage of molecular absorption lines to generate stable frequency references is that, due to their ro-vibrational structures, these lines are split in multiple narrow components which cover a large wavelength band. Therefore, a unique molecule can provide multiple frequency references spread over a large band. In the case of the 1.5  $\mu\text{m}$  region, several molecules such as acetylene<sup>7</sup> in the [1510-1550 nm] band, hydrogen cyanide<sup>8</sup> in the [1520-1565 nm] band or even methane<sup>9</sup> ( $\text{CH}_4$ ) around 1.6  $\mu\text{m}$  can provide about 60 lines covering roughly a 30 nm wavelength band for each isotope. In order to detect these lines, the main problem was their rather low absorption coefficient which comes from a low transition dipole moment (typically  $\mu \sim 0.01$  Debye for X-H stretching vibrations). Therefore, for a 10 cm long absorption cells, rather high pressures ( $> 10$  Torr) were necessary to obtain a significant absorption signal with a single-pass detection. With such a high pressure, the saturation intensity of the transitions was always too high<sup>10,11</sup> to be able to detect narrow saturated-absorption lines with usual low-power laser diodes : an optical power increase of about  $10^4$  was necessary to obtain narrow saturated-absorption lines. The linewidth of such frequency references was thus always Doppler-limited to about 500 MHz, giving a rather poor accuracy. The important problem to obtain narrow saturated-absorption lines was thus to be able to decrease the saturation intensity, mainly by reducing the gas pressure, while keeping a significant Doppler-



where  $c$  the speed of light in vacuum, and  $L$  the FP cavity length. Then, we lock the frequency of a laser  $L_1$  to the  $q$ -th FP resonance which is perturbed by the dispersion  $\delta n$  associated to the absorption line ( $n = 1 + \delta n$ ), and another laser  $L_2$  to the following FP resonance which is not perturbed by the line ( $n = 1$ ). The beat frequency  $f_B$  between both lasers should be equal to  $f_B = \nu_{q+1} - \nu_q = c/2L + \nu_q \delta n(\nu_q)$ . When the FP resonances are swept by  $\Delta \nu_q$  using a PZT driven variation of  $L$ , the beat frequency follows :

$$\Delta f_B(\nu_q) = (f_B/\nu_q)\Delta \nu_q + \nu_q \delta n(\nu_q) \quad (2)$$

with a 20 cm long FP cavity, the first term leads to a beat frequency variation of only 4 Hz for a 1 MHz sweep of  $\nu_q$  which will be neglected compared to the effect of  $\delta n$ . Therefore, any variation of  $f_B$  is directly linked to a refractive index variation  $\delta n$  of the gas. On fig.4, the beat frequency variations are shown when the FP resonance  $\nu_q$  is swept across the line P(12) of acetylene with a gas pressure of 30 mTorr.

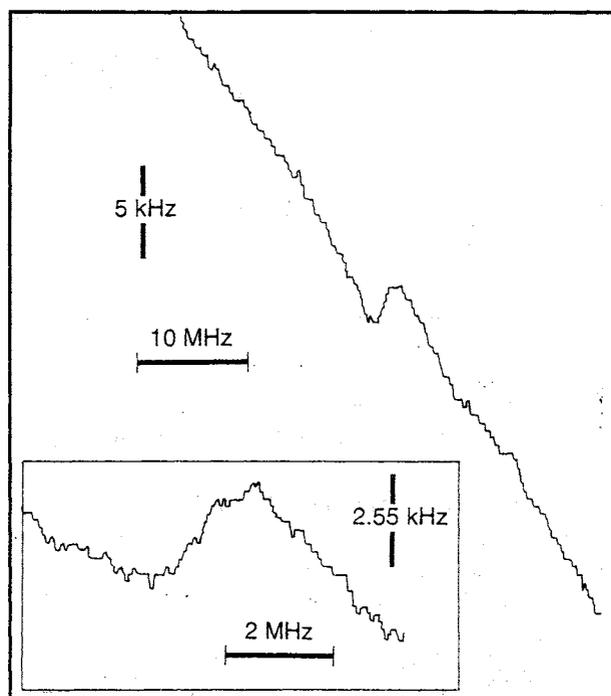


Fig. 4 : Beat frequency measurement versus FP resonance frequency showing the saturated-dispersion.

The slope (800 Hz/MHz) can be attributed to the linear dispersion of the line, and the small feature which represents a 2.5 kHz peak-to-peak variation comes from saturated-dispersion, and gives a measured peak-to-peak variation of  $\Delta n \approx 1.3 \cdot 10^{-11}$ .

The interpretation of these results is straightforward with the help of Kramers-Krönig relations from which the peak-to-peak refractive index change  $\Delta n$  is roughly given by :

$$\Delta n \approx \lambda/4\pi \cdot \Delta \kappa \quad (3)$$

where is  $\Delta \kappa$  the gas maximum absorption. Although these relations are not strictly valid in the case of non-

linear effects such as saturated-absorption, our experimental result is remarkably consistent with (3) : absorption measurements gave directly  $\Delta \kappa \approx 11 \cdot 10^{-5} \text{ m}^{-1}$  which exactly corresponds to  $\Delta n \approx 1.3 \cdot 10^{-11}$  from (3).

## 1.5 $\mu\text{m}$ FREQUENCY METROLOGY

### Accurate Frequency References at 1.5 $\mu\text{m}$

In order to lock the laser frequency to the saturated-absorption line, we have used the simplest locking procedure : PZT#2 was slightly dithered at 1 kHz. Since this frequency is lower than the first loop's bandwidth, the laser frequency still follows the FP resonance and the FP transmission will not be amplitude-modulated at 1kHz except when the laser FM is converted to AM on the slope of the absorption line. The AM signal, measured with a lock-in amplifier, is thus proportional to the absorption line derivative which nulls out at the top of the line. This signal is integrated and fed back to the PZT#2 driver in order to lock the laser frequency to the top of the line. The obtained frequency stability has been measured by the counting of the frequency beat between identically stabilized lasers : it reaches  $10^{-12}$  for an integration time of about 1s, depending on the servo adjustments, and the reproducibility is on the order of  $\pm 10$  kHz. Such a result shows that this laser wave is a frequency marker at 200 THz whose frequency can be defined with a reproducibility of only  $\pm 10$  kHz.

### Atlas of 1.5 $\mu\text{m}$ Frequency References

After having shown that these saturated-absorption lines are enough reproducible to be used as frequency references, it was necessary to calibrate their frequencies to be able to use them in frequency comparisons. This work has been done in two steps :

- First, the frequency separation between these lines has been measured with a precision much better than 10 kHz, using a frequency comb generator<sup>13</sup>.

- After that, we have measured the absolute frequency of the HCN line P(27), whose second harmonic could be easily linked to the two-photon Rb frequency standard at 778 nm<sup>14</sup>.

Presently, we have obtained<sup>15</sup> an atlas of 90 acetylene frequency references (see fig. 5), covering the band [1.52–1.55  $\mu\text{m}$ ], whose frequencies are defined with an accuracy of  $10^{-9}$ .

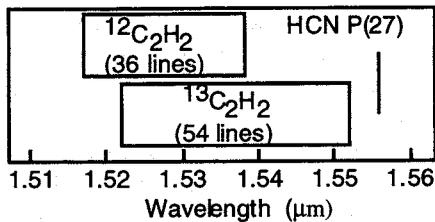


Fig. 5 : Frequency reference atlas at 1.5  $\mu\text{m}$ . The frequency spacing between lines is on the order of 60~80 GHz.

## IMPROVEMENTS

### Technical Improvements

**Gas cell.** In all our experiments, we have used a glass cell with epoxy-glued mirrors and a PZT which was also glued with epoxy. These cells suffer from residual outgassings or micro-leakages which results in a significant pressure increase after a few days. For longer times, saturated-absorption signals are broadened and lowered. A routine use of such frequency standards will require an improvement of the cell manufacturing technology. For instance, by using a well-manufactured cell with Brewster windows, with external FP mirrors, the problem could be solved at the expense of a somewhat lower finesse, and probably a higher sensitivity to acoustic noise. Another possibility is to use a similar construction with an external PZT which slightly changes the cavity length by pressing on the end mirror. In that case, the FP tube should be specially designed in order to allow for some longitudinal elasticity.

**Linewidth.** The laser stability and reproducibility are directly linked to the saturated-absorption linewidth. Therefore, an obvious improvement would be to design a FP cavity with a much wider spatial mode in order to lower the transit-time broadening. In that case, it is then necessary to compensate for the decrease of the intra-cavity optical power density to keep the same saturation conditions. Longer cavities with higher mirror radius of curvature and reflectivities are recommended, however, the cavity linewidth will decrease, which imposes a better servo electronics for the first loop.

**Size reduction of the cell.** As was already demonstrated<sup>16</sup>, it is possible to use a much shorter FP cavity with a higher finesse to obtain saturated-absorption. A 25 mm cavity with a finesse  $\mathcal{F} > 10,000$  filled with  $\text{C}_2\text{H}_2$  provided a 2.8 MHz wide absorption line with a large 140 mV signal corresponding to 4% of total transmitted power. In a cell with such a small volume, the increase of the background pressure was of

primary importance therefore, an optically contacted cell construction with an external PZT would be a quite interesting solution. With such a cell, one observed problem is the sensitivity to the chosen gas pressure : if it is too low, the absorption is over-saturated and becomes weak, very distorted, and difficult to locate. We should emphasise that we did not carry on the experiments on compact cells only because of these minor technical problems, but we are convinced that these compact references would work remarkably well with only a few technical improvements.

**Locking method.** The main improvement in the locking method would be to get rid of the residual AM at 1 kHz that we always observe : when the laser frequency is far from the absorption line, the lock-in amplifier's output is not exactly zero in most of the cases : this means that a residual AM at 1 kHz is generated at the same time as the FM and interpreted by the detection system as an offset absorption signal. This offset depends on many parameters such as alignments, frequency difference of the free-running and the FP-locked laser...and can have several origins :

- Cavity effects in the optical path which converts the laser FM to AM. This effect can be lowered by a careful design of the optical system.
- Since PZT#1 cannot follow efficiently a 1 kHz modulation, the first servo loop will correct the laser frequency to keep the laser locked to the FP mainly through a current control, which always generates AM. This effect can be lowered by using an external EO modulator to modulate the laser frequency, or any mean to frequency modulate the laser at high speed without residual AM.

Another possible detection method<sup>17</sup> consists in detecting directly the saturated-dispersion profile by using another set of high-frequency FM sidebands.

## MULTIFREQUENCY STANDARD

Since we have established a lot of well-calibrated frequency references at 1.5  $\mu\text{m}$  and a tunable laser, we can imagine a "Multifrequency standard" made of this laser whose frequency would be able to lock automatically with a very good accuracy to a predefined frequency reference contained in an absorption cell. However, the resonant method used in our experiments would always require to lock the laser to the FP which complicates significantly the automatic locking procedure. We propose a possible way to overcome this problem.

**Non-resonant detection method.** Another way to obtain increase the gas absorption length would be to use a non-resonant multiple pass cell<sup>17</sup>. However, this improvement alone is not enough to obtain the

saturated-absorption in good conditions : one must also obtain a significant optical power increase. Such an optical power increase can be obtained by using an optical amplifier which allows to reach the 100 mW region. With a typical 20 cm long multipass cell and a typical fiber amplifier, it should be possible to obtain the same conditions as in our experiments. In that case, it would be possible to detect narrow molecular lines without having to worry about frequency-locking the laser to a FP cavity. From the point of view of accuracy, improvements can be expected because the problems related to complex interactions in a system using a double servo loop would be absent. From the practical point of view, it would be possible to sweep freely the laser frequency over long ranges and lock it easily to successive lines.

**ECSL tuning.** A very widely tunable source would be necessary to make such a multifrequency standard. Since we have constructed for all experiments external-cavity lasers with a mode-hop suppression arrangement<sup>18</sup>, we believe that such lasers would be suitable, therefore we have investigated their continuously tuning capabilities. The laser rough frequency tuning was usually obtained with a manually driven screw and the fine tuning with a PZT inserted between the screw and the grating (see fig.1). We have replaced the manual screw by a motorized one, thus allowing an electrically driven tuning over long ranges. By measuring the output wavelength with a wavemeter (straight lines on fig. 6), it was possible to follow the laser frequency and to see clearly the mode-hops (labeled by two dots on fig.6). At the same time, the laser output optical power was recorded : the laser power fluctuates with a period corresponding to the solitary laser diode's intermode spacing, because the laser diode acts as an internal FP etalon. In this experiment, no mode hops are observed for a wavelength sweep of about 15 nm, but the power is gradually decreasing because the lasing mode is tuned faster than the grating resonant wavelength and gradually shifts from the top to the bottom of the grating selection curve.

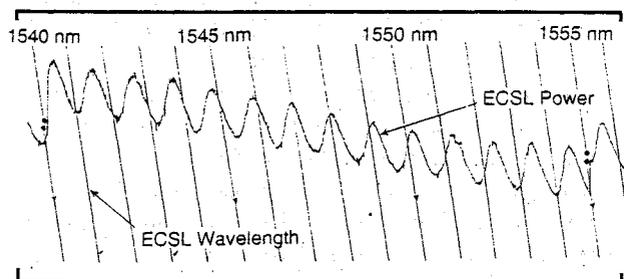


Fig. 6. Frequency and output power of the ECSL during the frequency tuning showing a 15 nm tuning free of mode hops.

A sudden, mode hop can be seen at the beginning and the end of the scan, and corresponds to a sudden increase of the power because the following mode is closer than the previous one to the top of the grating selection curve. This behaviour can be attributed to the large coupling between amplitude and refractive index which is observed in laser diodes. We can conclude that such a laser can reach without mode hops about 25 lines of acetylene, and even if there is a mode hop, the associated power jump can be easily detected to be able to take in account. Therefore, from the above considerations, we think that a convenient "multifrequency standard" could be developed in straightforward manner using our ECSL.

## CONCLUSION

In this paper, we have summarized the main results that we have obtained about 1.5  $\mu\text{m}$  frequency references using molecular lines. We have reminded how to detect their saturated-absorption and shown new experimental results about their saturated-dispersion spectra. Their great interest for high accuracy optical frequency metrology was pointed out and illustrated. We also tried to review the main problems that still remain to be solved. Meanwhile, we have proposed hints to make a compact system, and also to make a convenient "multifrequency standard" which could precisely and automatically lock to many reference frequencies.

## REFERENCES

1. O. Acef, J.J. Zondy, M. Abed, D.G. Rovera, A.H. Gerard, A. Clairon, P. Laurent, Y. Millerioux, and P. Juncar, *Optics Comm.*, **97**, 29 (1993)
2. Y. Millerioux, D. Touahri, L. Hilico, A. Clairon, R. Felder, F. Birabein, B. de Beauvoir, *Optics Comm.* **108**, 15 May 1994, pp. 91-96.
3. M. Poulin, C. Latrassé, M. Têtu, M. Breton, *Optics Lett.* **19**(16), pp. 1183-1185, (1994).
4. M. Kourogi, K. Nakagawa, M. Ohtsu, *IEEE Journ. of Quantum Electron.*, vol. **29**, pp.2693-2701, (1993).
5. D. J. E. Knight, P. S. Hansell, H. C. Leeson, G. Duxbury, J. Meldau, and M. Lawrence, *Proc. Soc. Photo-Opt. Instrum. Eng.*, vol. **1837**, pp.106-114, (1992).
6. M. de Labacherie, K. Nakagawa, M. Ohtsu, *Opt. Lett.*, vol. **19**, pp. 840-842, (1994).

7. S. Kinugawa and H. Sasada, Jap. Journ. of Appl. Phys., vol. 29, pp. 611-612, (1990).
8. H. Sasada and K. Yamada, Appl. Optics, Vol.29, pp. 3535-3547, (1990)
9. K. Suzumura, H. Sasada, Jap. Journ. of Appl. Phys, 34, Part.2, No. 12A, L1620-L1621, Dec 1995.
10. S. Sudo, Y. Sakai, H. Yasaka, T. Ikegami, IEEE Photonics Technology Letters **10**, Oct. 1989, pp. 281-284
11. M. Labachellerie, C. Latrasse, K. Diomandé, P. Kemssu, P. Cerez, IEEE Trans. Instrum. Meas. **40**, 185 (1991).
12. M. de Labachellerie, K. Nakagawa, M. Kouroggi, Y. Awaji, M. Ohtsu, Proc. Soc. Photo-Opt. Instrum. Eng., vol. **2378**, San Jose, pp. 44-51, (Feb. 1995).
13. K. Nakagawa, M. de Labachellerie, Y. Awaji, M. Kouroggi, T. Enami, M. Ohtsu, Opt. Lett. , vol. **20**(4), pp. 410-412, (1995).
14. Y. Awaji, K. Nakagawa, M. de Labachellerie, M. Ohtsu, Opt. Lett. , vol. **20**(19), pp. 2024-2026, (1995).
15. K. Nakagawa, M. de Labachellerie, Y. Awaji, M. Kouroggi, "Accurate optical frequency atlas of 1.5  $\mu\text{m}$  band of acetylene", Submitted for publication.
16. M. de Labachellerie, K. Nakagawa, M. Ohtsu, Proc. IQEC'94, Anaheim CA, May 1994, Paper QTuL.
17. Long-Sheng Ma, P. Dubé, P. Jungner, Jun Ye, John L. Hall, CLEO/QELS'95, May. 1995.
18. G. Duxbury et Al., OSA Ann. Meeting, Dallas, (1994).
18. P. McNicholl, H. Metcalf, Appl. Opt., vol. 24, 2757-2761, (1985).

# FREQUENCY STABILITY OF $1 \times 10^{-13}$ IN A COMPENSATED SAPPHIRE OSCILLATOR OPERATING ABOVE 77 K

G. John Dick, David G. Santiago, and Rabi T. Wang  
 Jet Propulsion Laboratory  
 California Institute of Technology  
 Pasadena, CA 91109-8099, U.S.A.

## Abstract\*

We report on tests of a compensated sapphire oscillator (CSO) which shows frequency-stable operation at temperatures above 77 K[1]. The frequency stability for this oscillator shows an apparent flicker floor of  $7.5 \times 10^{-14}$  for measuring times between 3 and 10 seconds, and stability better than  $2 \times 10^{-13}$  for all measuring times between 1 and 100 seconds. These values are approximately the same as for the very best available quartz oscillators. Previously, high stability in sapphire oscillators had only been obtained with liquid helium cooling. Recent improvements include a more careful analysis of the ac frequency-lock "Pound" circuitry that now enables the oscillator to reliably attain a stability 6 million times better than its fractional resonator linewidth. Measurements to date have been made with a resonator quality factor  $Q \approx 2 \times 10^6$ . Frequency stability of  $2 \times 10^{-14}$  is projected for a resonator  $Q$  of  $10^7$ , a value about one third of the intrinsic sapphire  $Q$  at this temperature.

## Introduction

Newly developed atomic and ionic frequency standards are presently limited in performance by available local oscillators. Sequentially-interrogated passive standards, which include mercury ion traps and cesium fountains, rely on an ancillary local oscillator (L.O.) which is periodically corrected by the atomic interrogation process[2,3]. In order for the standards to achieve their potential performance, a local oscillator with stability of a few times  $10^{-14}$  is required.

Up to now, L.O. requirements for passive frequency sources such as Cesium and Rubidium standards were easily met by available quartz oscillators. The general characteristics of quartz oscillators are an excellent match to L.O. requirements: They are relatively inexpensive and show their best stability for measuring times approximately equal to the required interrogation times. However, even the best "super-quartz" oscillators with stability of approximately  $1 \times 10^{-13}$  do not meet L.O. requirements for the new standards.

Active hydrogen masers and superconducting or sapphire oscillators cooled by liquid helium[4-6] do achieve the desired performance, but are roughly as expensive as the standards themselves. While such a combined standard may be very attractive when the ultimate in performance is needed, the expense is prohibitive for most applications.

A sapphire oscillator cooled by liquid nitrogen ( $\text{LN}_2$ ) could be a simpler and less expensive solution. The available quality factors ( $Q$ 's) for whispering gallery sapphire resonators at temperatures above the 77K boiling temperature of  $\text{LN}_2$  are in fact high enough to allow the required performance. However, thermally induced variations of the dielectric constant are not frozen out at 77K as they are at LHe temperatures, and prevent high stability from being attained.

We have developed a compensated sapphire resonator that reduces the effects of thermal fluctuations. This resonator incorporates a mechanical compensation process driven by the difference in expansion coefficients for the component materials (copper and sapphire). Previously reported stability of  $2-4 \times 10^{-13}$ [7] for the compensated sapphire oscillator (CSO) based on this resonator is now substantially improved, achieving a flicker floor of  $7.5 \times 10^{-14}$ .

The presently observed  $Q \approx 2 \times 10^6$  is very much lower than the intrinsic value of 30 million for sapphire at 77K, and is also below the value of 20 million we observed for other, uncompensated, modes in the same resonator. A redesign is presently underway to reduce surface contamination of the tuning gap, where resonant electric fields are large. We project a stability of  $5 \times 10^{-14}$  or better with the improved design, and calculate a noise-limited frequency stability of  $1-2 \times 10^{-14}$  for a resonator with  $Q = 10^7$ . Based on this achieved and projected performance, the CSO approach promises to meet new passive standard L.O. requirements in a compact and inexpensive cryogenic package.

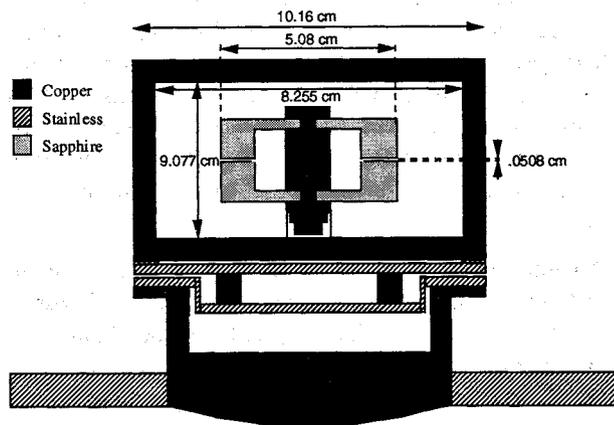
## Methodology

A detailed description and analysis of the operation of the compensated sapphire resonator has been given elsewhere[1,7,8]. This approach was anticipated by Tsarapkin, et al[9] in a room-temperature resonator with low phase noise. Our previous work analyzes a tuneable resonator constructed with a gap between two sapphire parts.

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

The analysis shows that, for the  $WGH_{n11}$  mode family, the sensitivity of resonator frequency to gap spacing is sufficient to compensate the inherent thermal frequency variation in the sapphire resonator at temperatures above 77K if the parts are separated by a material such as copper, which has a coefficient of expansion somewhat greater than that of sapphire. However, the sapphire must be made substantially reentrant, so that the effective length of the copper spacer can be larger than the gap separating the sapphire parts. When these conditions are met, the difference between thermal expansion coefficients of copper and sapphire adjusts the gap between two sapphire parts and cancels frequency variation due to thermal expansion in the sapphire and, more importantly, that due to temperature-induced variation in sapphire's dielectric constant.

The sapphire-copper composite structure is shown in Figure 1. Increasing temperature, which would tend to *decrease* resonant frequency, causes the length of the central copper post to increase, thus separating the sapphire elements, increasing the gap and thereby *raising* the resonant frequency. At a certain operating temperature these effects completely cancel, and therefore compensate the resonator frequency against the effects of temperature variation. In our tests, the  $WGH_{811}$  mode at 7.23 GHz is excited, and shows a frequency turn-over temperature of 87 K in agreement with finite element calculations[10].



**Figure 1.** Compensated sapphire resonator with frequency turn-over temperature of  $\sim 87$ K. Expansion of the copper center post with increasing temperature increases the gap spacing between the sapphire elements, counteracting an increase in dielectric constant in the sapphire. Stainless steel thermal isolation assembly reduces the effect of  $LN_2$  temperature fluctuations.

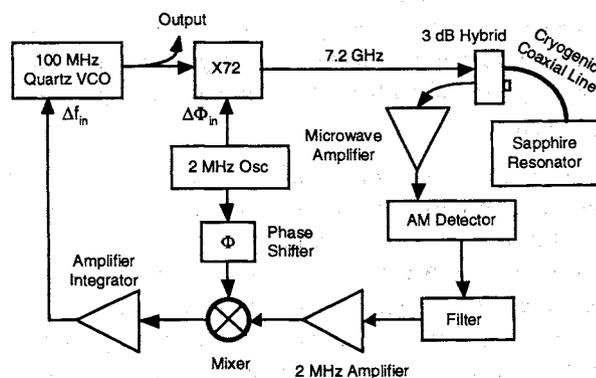
Thermal integrity of the sapphire-copper-sapphire resonator part is crucial to its frequency stability. The copper and sapphire elements are bonded using pure indium solder and an evaporated gold coating on the sapphire joint surface. This, together with the very high thermal conductivity of both sapphire and copper at  $LN_2$  temperatures, enables a low thermal time constant. The much longer time constants for the sapphire-can mounting and the can-nitrogen bath attach-

ment allow excellent short-term temperature control of the copper/sapphire resonating element and very low thermal gradients.

The internal thermal time constants for the composite resonator are  $< 5$  seconds, allowing effective operation of the compensation mechanism. Thermal time constants of 300 seconds and 1500 seconds isolate the sapphire element from the can, and nitrogen bath, respectively.

The design of the can thermal isolation, as shown by the stainless steel parts identified in Figure 1, is reentrant to minimally effect the resonator's placement in the cryostat. The original bottom plate with the copper center that sits in the  $LN_2$  bath is spaced approximately 8 mm from the copper can, but the thermal path length is approximately 6.5 cm. The thermal isolation stage is composed of a stainless steel 'deep dish' in which a copper cylinder is attached. On top of the copper cylinder is a stainless steel plate which only makes contact to the copper can with a  $\sim 0.5$  cm width ring at its outer radius. The copper cylinder has thermistors and a heater element which allow the temperature of the stage to be controlled.

The relatively conventional frequency lock circuitry is shown in Figure 2. A Pound circuit locks the 100 MHz quartz VCO to the sapphire resonator. Earlier versions used a 50-200 kHz modulation frequency injected into the VCO input. However, sufficient loop gain could not be attained without instabilities to effectively eliminate VCO frequency fluctuations. Therefore we modified the  $\times 72$  multiplier to allow injection of a higher 2 MHz modulation frequency into its L-band internal power oscillator. The allowed increase in loop gain in the frequency lock circuitry greatly improved short-term stability performance.



**Figure 2.** Pound (frequency lock) circuit with 2 MHz modulation frequency. Not shown is frequency offset circuitry associated with the  $\times 72$  multiplier which derives the exact resonator frequency of 7.226 GHz.

Because flicker noise in the rf system components is a limiting factor in the system performance we employ the lowest noise components available. Additionally, we generally design for the shortest microwave path lengths possible.

## Experimental

A number of significant sources of frequency instability were uncovered during the development process. Vibration sensitivity is a concern because of the multiple element structure of the resonator. No quantitative testing has been performed to determine vibration sensitivity, but during initial tests the system was subjected to deliberate mechanical impulses. Several mechanical resonances were observed in the range  $1 \text{ kHz} < f < 10 \text{ kHz}$  with ringing times of a few tenths of a second. Such resonances are unlikely to degrade frequency stability performance. In a frequency-locked condition, the apparent sensitivity of the resonator cryostat to applied vibration and shaking was less than that of the associated microwave components or of the 100 MHz crystal quartz oscillator.

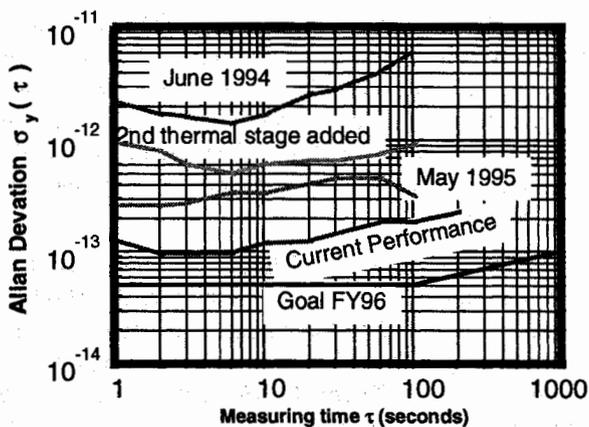


Figure 3. Allan Deviation of CSO frequency stability showing performance at various stages of development (see text).

Early performance (June 1994 in Fig. 3) was found to be limited by thermal stability of the resonator containment can. The copper can cavity which houses the resonator had been well anchored to the  $\text{LN}_2$  bath. Changes in room temperature and pressure as well as the  $\text{LN}_2$  level affect the temperature of the liquid nitrogen bath.

First tests with a 2nd thermal stage showed greatly improved long-term performance. With the resonator operating at its turnover temperature of 87K, the new isolation stage was a few degrees above the  $\text{LN}_2$  temperature. The stability of the current sources used to drive the heaters was found to be poor, so they were replaced with more stable diode-laser current supplies for subsequent experiments.

Increasing the loop gain proved to be the most significant factor in achieving the May 1995 performance indicated in Figure 3, but several other factors contribute to the improved stability. The resonator is the only component of the system contained in the cryostat. The external microwave components are exposed to the environment of our open laboratory. These mixers, hybrids, amplifiers, connectors, and cable lengths are

sensitive to temperature fluctuations and in combination greatly contribute to the instability of the system. Medium and long term performance were improved by thermally insulating the microwave components with foam.

An important source of instability is temperature fluctuations and gradients on the relatively long coaxial line which feeds the resonator in the cryostat. This cable is cooled to  $\text{LN}_2$  temperature at the resonator and held at room temperature at the cryostat's input, thus making it sensitive to both  $\text{LN}_2$  and room temperature fluctuations. This instability was reduced by better isolating the line from the cryostat wall and maintaining a more stable  $\text{LN}_2$  surface temperature and level. This improvement also contributed to the May 1995 stability in Figure 3.

By monitoring various temperatures in the system, we found that the sapphire temperature followed the outer can as well as showing its own temperature fluctuations. This indicated a thermal 'leak' in the resonator. The thermal path was a combination of a small vacuum leak and thermal radiation. Sealing the leak and adding radiation shielding significantly improved short term stability as seen in the curve labelled *current performance*.

## Analysis

Several additional improvements have been identified as necessary to achieve the desired oscillator stability. The resonator's thermal environment is currently free running and not actively controlled. Addition of feedback control electronics to the heater elements of the system with better than milliKelvin temperature resolution is required for the desired ultra-stable performance.

An important sensitivity of the flicker floor to the adjustment of the Phase Shifter identified in Figure 2 helped to identify a "false signal" in the Pound loop and a solution to the problem that this represents. It was found that the phase shifter was adjusted to be even a few degrees from the peak of the response curve at the mixer output the performance was substantially degraded. Furthermore, ultimate performance was found to be very sensitive to the length of the Cryogenic Coaxial Line. These effects have been identified as due to a 2 MHz signal at the output of the AM Detector which is phase shifted by exactly 90 degrees from the true Pound signal, and which is periodically dependent on the length of the coaxial line.

Analysis shows that such a false signal can be expected due to transmission line mismatches, and experiment shows that adjusting the line length to minimize the false signal gives the best possible stability. While to first order the Pound methodology eliminates any dependence of oscillation frequency on coaxial line length, the false signal represents a breakdown of the

method, being a signal that is not immune from changes in line length. The exact 90 degree phase shift is due to the very high Q of the resonant mode. The results indicate that caution must be observed with regard to possible sapphire modes with moderate to high Q's which have frequencies near the modulation sidebands. This problem is compounded by the fact that very many lower Q modes in the whispering gallery resonator are coupled to the output as strongly as is the desired mode.

Features of the achieved stability are a flicker floor of  $7.5 \times 10^{-14}$  and a large frequency drift of approximately  $1.5 \times 10^{-8}$ /day. A significant improvement in stability is expected with increased resonator Q. The resonator Q for the compensated mode is  $\approx 2$  million, while other modes in the same resonator show Q's up to 20 million. We believe the low compensated Q is due to poor surface cleanliness of the sapphire elements, most likely in the tuning gap where rf electric fields are large. An improved sapphire polish may allow us to achieve resonator Q's of 8 million or more. Based on our experimental results this Q would make possible a stability of  $2 \times 10^{-14}$ . The large drift rate is likely due to relaxation of the soft metal bond in the composite sapphire/copper resonator element. This is being addressed by an improved fabrication technique.

#### Conclusion

We have demonstrated a new ultra-stable oscillator capability which promises to enable improvements on the best quartz technology in a small and inexpensive cryogenic package. Projected performance is well matched to the requirements of new passive atomic standards. With a 20x performance improvement over the past 18 months, continuing improvements can be expected. Present stability ranges from  $7.5 \times 10^{-14}$  to  $2 \times 10^{-13}$  for measuring times between 1 and 100 seconds. We project a stability of  $5 \times 10^{-14}$  or better in FY96 with a resonator designed for improved Q.

#### References

- [1] G.J. Dick, D.G. Santiago, and R.T. Wang, 1995, "Temperature-Compensated Sapphire Resonator for Ultra-Stable Oscillator Capability at Temperatures Above 77 K," *IEEE Trans. UFFC*, **42** 812-819
- [2] J.D. Prestage, R.L. Tjoelker, G.J. Dick, and L. Maleki, 1994, "Recent Stability Comparisons with the JPL Linear Trapped Ion Frequency Standards," *Proc. 1994 IEEE International Frequency Control Symposium*, 739-743
- [3] A. Clairon, Ph. Laurent, G. Santarelli, S. Ghezali, S.N. Lea and M. Rahoura, 1995, "A Cesium Fountain Frequency Standard: Preliminary Results," *IEEE Trans. Instrum. Meas.*, **44**, 128-131
- [4] J.D. Prestage, R.L. Tjoelker, R.T. Wang, G.J. Dick and L. Maleki, 1993, "Hg<sup>+</sup> Trapped Ion Standard with the Superconducting Cavity Maser Oscillator as L.O.," *IEEE Trans. Instrum. Meas.*, **42** 200-205
- [5] A.N. Luiten, A.G. Mann, N.J. McDonald, and D.G. Blair, 1995, "Latest Results of the U.W.A. Cryogenic Sapphire Oscillator," *Proc. 1995 IEEE International Frequency Control Symposium*, 433-437
- [6] E.N. Ivanov, M.E. Tobar and R.A. Woode, 1995, "Advanced Phase Noise Suppression Technique for Next Generation of Ultra Low-Noise Microwave Oscillators," *Proc. 1995 IEEE International Frequency Control Symposium*, 314-320
- [7] D.G. Santiago, R.T. Wang and G.J. Dick, 1995, "Improved Performance of a Temperature-Compensated LN<sub>2</sub> Cooled Sapphire Oscillator," *Proc. 1995 IEEE International Frequency Control Symposium*, 397-400
- [8] G.J. Dick, D.G. Santiago and R.T. Wang, 1994, "Temperature-Compensated Microwave Sapphire Oscillator at Temperatures Above 80 Kelvin," *Proc. 1994 IEEE International Frequency Control Symposium*, 421-432
- [9] S.L. Abramov, E.N. Ivanov, and D.P. Tsarapkin, 1988, "Low Noise Microwave Oscillator with a Temperature Stabilized Disk Dielectric Resonator," *Radio-technika*, no. 11, 81-83, (in Russian)
- [10] A. Revilla, R.A. Osegueda, J.H. Pierluissi, G.J. Villalva and G.J. Dick, 1994, "2D Finite element Program to Compute Resonant Frequencies and 3D Visualization of Electromagnetic Modes in Cylindrical Resonators," *Proc. 1994 IEEE International Frequency Control Symposium*, 493-499

## PERFORMANCE OF A CRYOGENIC SAPPHIRE DIELECTRIC RESONATOR.

C.D. Langham and J.C. Gallop.

National Physical Laboratory, Teddington, Middx. TW11 0LW, UK.

### ABSTRACT

In this paper we present details of the design and operation of a cryogenic Sapphire Dielectric Resonator. This includes controlling the resonator's temperature to better than  $90 \mu\text{K}$  and the effect of paramagnetic impurities on the predicted performance. We will also discuss room temperature observations on the efficiency of the novel microwave choke arrangement. Finally we present some preliminary stability measurements for our Mark I system.

### INTRODUCTION

Sapphire Dielectric Resonators (SDR) at cryogenic temperatures have been used to produce oscillators of exceptional short term stability, for example Luiten et al (1). At the NPL we are developing an SDR based on a sapphire rod 140 mm long and 20 mm diameter, Langham and Gallop (2,3). Included in our design is a novel microwave choke arrangement whose aim is to axially confine the resonant field to the central region of the sapphire rod. This resonator is enclosed in an evacuated superconducting shield and cooled to 4.2 K by immersion in liquid helium.

### THE SAPPHIRE DIELECTRIC RESONATOR

The single crystal sapphire rod is circular in cross section, with a diameter of 20 mm, and is 140 mm long. At a distance of 10 mm from each end there is a region of the rod that has been turned down to produce a step of 10 mm diameter and 10 mm length. It is intended that these stepped down regions will act as microwave chokes, reducing the EM field strengths at the ends of the rod and thus the losses associated with the metal end-walls.

The rod is mounted coaxially inside a lead shield of inner diameter 48 mm, and is held in

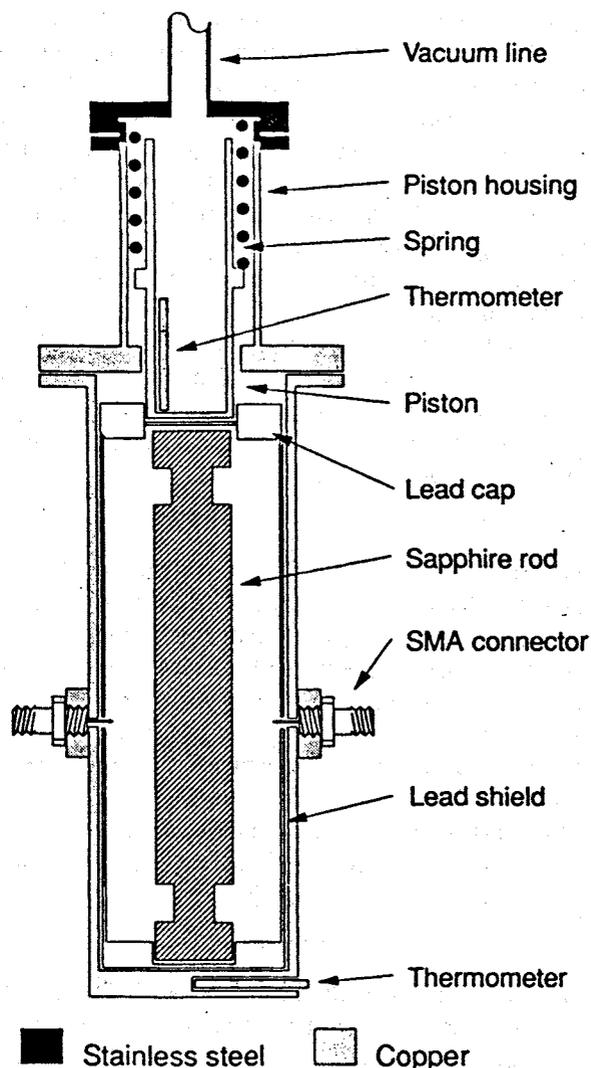


Figure 1: Cross-sectional view of the SDR

place by a spring-loaded piston. The walls and base of the shield have been machined out of a single piece of lead, and a lead cap is fitted over the top of the shield. This cap has a hole through which the piston passes, there is a lead spacer between the sapphire and piston. The resonator assembly is fitted into a copper vacuum can, see figure 1. Microwave power is fed into and out of the vacuum can by hermetically sealed SMA connectors in its side walls. Coupling to the resonator is then achieved by

probes which protrude through the lead shield by a few millimetres.

In order to provide a high degree of temperature control this resonator structure is mounted into a second vacuum jacket. Four thin film resistive heaters are mounted around the SDR can, while cooling is provided by a number of copper thermal links which connect from the flange of the can to the outer vacuum jacket. Thermometry is provided by a combination of Rhodium-Iron and Carbon glass resistance thermometers. The entire structure is cooled by immersion in liquid helium.

### THE MICROWAVE CHOKES

To investigate the efficiency of the microwave chokes, the sapphire rod was supported horizontally between a pair of copper 'book-ends'. This consists of two copper plates, each 120 mm square, mounted vertically and facing one another. To each copper plate (on their facing sides) is fastened a sheet of PTFE, 3 mm thick and again 120 mm square. In the centre of each PTFE sheet is a 20 mm diameter hole to hold the sapphire rod which then can be held horizontally above the bench.

When mounted in this structure the sapphire rod can be moved between a pair of probes which couple into and out of the evanescent microwave field of the mode under investigation. Figure 2 shows the evanescent field strength as a function of position along rod for modes with differing longitudinal mode number  $n$ . For low  $n$ -order modes (e.g.  $n=3$  &  $n=4$  of figure 2) the resonant field can clearly be seen to be confined to the region between the chokes. However, as the number of longitudinal nodes increases, the inter-nodal spacing approaches the dimensions of the choke structure, and we observe considerable field amplitude at the ends of the rod (e.g.  $n=10$  of figure 2). This behaviour is as we predicted, and the lower field strengths at the ends of the rod for  $n < 10$  modes will reduce the microwave losses associated with the shield's end-walls.

### TEMPERATURE CONTROL

In order to maintain the SDR at its optimum operating temperature of approximately 4.5 K two independent methods of temperature control are employed. The bath itself is

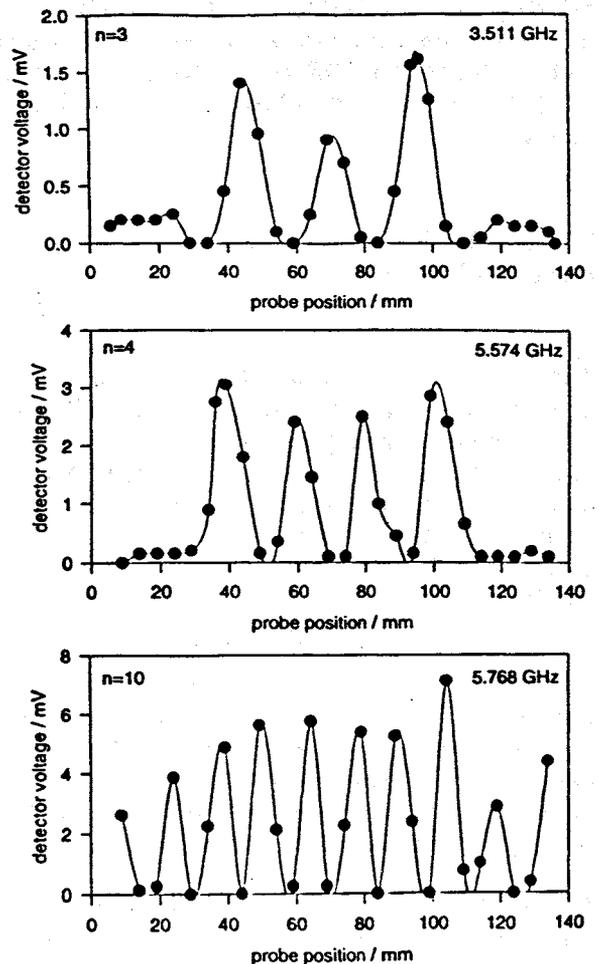


Figure 2: Field strength as a function of position along the sapphire rod

temperature controlled by means of a pump working at constant speed (which can be manually adjusted using a needle valve) and a heater immersed in the liquid helium. Power dissipated in the heater increases the helium boil-off rate and thus the vapour pressure rises. By controlling this 'excess' boil-off it is possible to servo the bath temperature to a carbon glass sensor to a stability of  $\sim 0.3$  mK. Note that this technique, by measuring the temperature in the boiling helium, is not subject to the temperature drift that would otherwise result from the time varying hydrostatic column of the liquid helium bath when it is the vapour pressure above the boiling liquid which is stabilised (by a manostat for example).

A second layer of temperature control is provided using the heaters mentioned in the description of the SDR above. A carbon glass sensor with a resistance of  $\sim 2$  k $\Omega$  and a sensitivity,  $dR/dT \sim 10^3$   $\Omega/K$  at 4.2 K is read using an a.c. resistance bridge. This system allows a

measurement resolution of  $10^{-5}$  K, which when used in conjunction with a commercial PID controller allows the SDR temperature to be stabilised to better than  $90 \mu\text{K}$  over a period of hours.

### EVALUATION OF RESONATOR PERFORMANCE

The dielectric resonator possesses many resonant modes in the frequency range 4-26 GHz. These have varying  $Q$ -values, dependent on various factors such as coupling strength to the enclosing, vacuum can, the properties of the sapphire and any impurities it may contain. Thus we can write the following, expression for  $Q$ ;

$$Q^{-1} = \frac{R_S}{\Gamma} + \tan \delta + \chi'' \quad (1)$$

where  $R_S$  is the surface resistance of the superconducting shield,  $\Gamma$  is the geometric factor of the mode,  $\tan \delta$  is the loss tangent of sapphire, and  $\chi''$  is the fractional energy loss induced by paramagnetic impurities. We selected a few of the higher  $Q$  modes for characterisation.

Essential to the characterisation of an oscillator is a study of the temperature dependencies of the resonant element. Observations of the mode line widths for the SDR, made as it warms from liquid helium temperature to the superconducting transition of the lead shield ( $T_c = 7.2 \text{ K}$ ) are shown in figure 3. This demonstrates a marked dependence on the properties of the lead shield. Just below  $T_c$  the width is

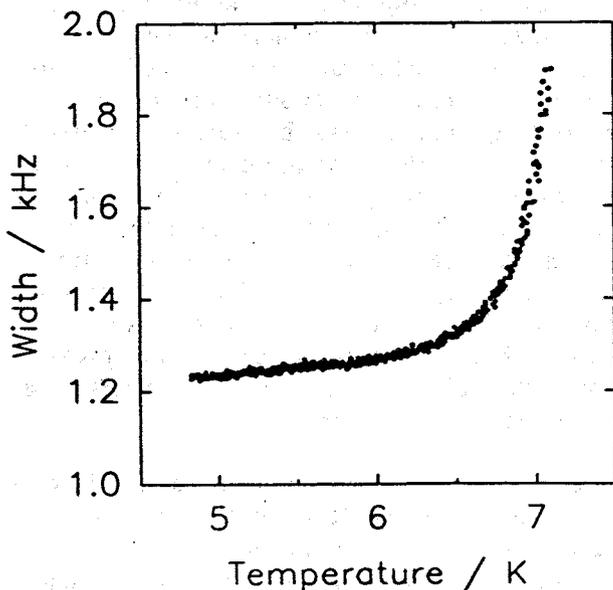


Figure 3: Width of resonance at 10.889 GHz as a function of temperature.

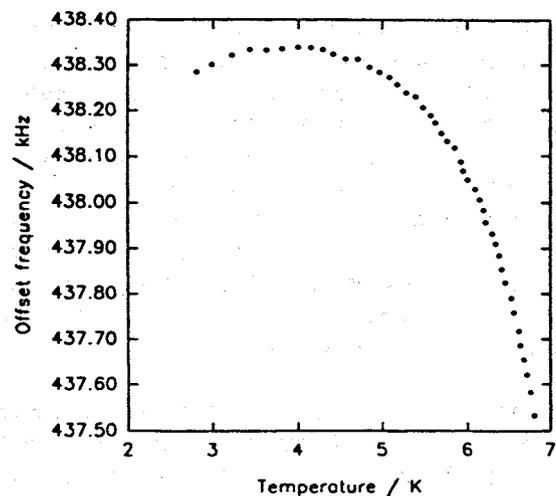
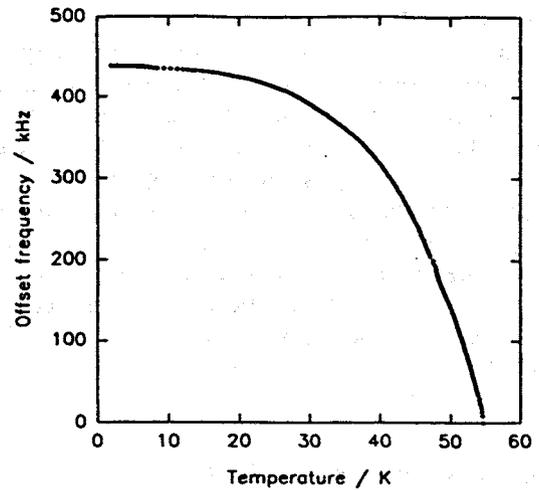


Figure 4: The upper graph shows the temperature dependence of a 10.889 GHz mode. The lower graph is showing the turning point of frequency for this mode.

falling rapidly with reducing temperature, but does not achieve a constant level which would be predicted by a simple two-fluid model of the surface resistance plus a constant residual resistance,  $R_{\text{res}}$ . This, we believe, is due either to a temperature dependent  $R_{\text{res}}$  resulting from impurities or inhomogeneities in the superconducting shield, or arises from surface contamination of the sapphire rod.

At temperatures below the normal boiling point of helium the mode frequency is observed to fall, typically to 150 Hz below the maximum at 2.5 K (see figure 4). This behaviour has been observed in both our rods and is consistent with reports from Jones et al (4). It can be attributed to the conflicting temperature coefficients caused by the surface reactance of the

superconducting shield and paramagnetic ion impurities in the sapphire. Thus;

$$\frac{1}{f} \frac{\partial f}{\partial T} = \frac{1}{2} \frac{\partial}{\partial T} \chi_0 - \frac{1}{2T} \frac{\partial}{\partial T} \chi_s(T) \quad (2)$$

where  $\chi_s$  is the static susceptibility of the paramagnetic impurities and  $\chi_s(T)$  is the surface reactance of the shield. We have estimated the concentrations of paramagnetic impurities (mainly chromium ions) in the sapphire rod to be approximately 0.3 ppm. From this result, and knowledge of the superconducting properties of the shield, the temperature dependence can be modelled. This shows a frequency maximum between 4 and 5 K depending upon the mode. Furthermore, it is predicted from this model that if the temperature is stabilised to 10 microkelvin of the maximum then frequency stabilities of 1 nHz can, in principle, be achieved. This compares to a frequency stability of  $\pm 0.4$  mHz at 4.2 K for the same degree of temperature control.

## FREQUENCY STABILITY MEASUREMENTS

We have achieved  $Q$  values as high as  $6 \times 10^7$  in our Mk I system at 4.2 K and a frequency of 17.3 GHz, using a higher order mode of the resonator shown schematically in figure 1. To demonstrate the potential of this dielectric resonator as a frequency standard a synthesised microwave source (Wiltron 6753A) has been stabilised to it by means of a 300 Hz frequency modulation, detected by a lockin amplifier referenced by this modulation frequency. The

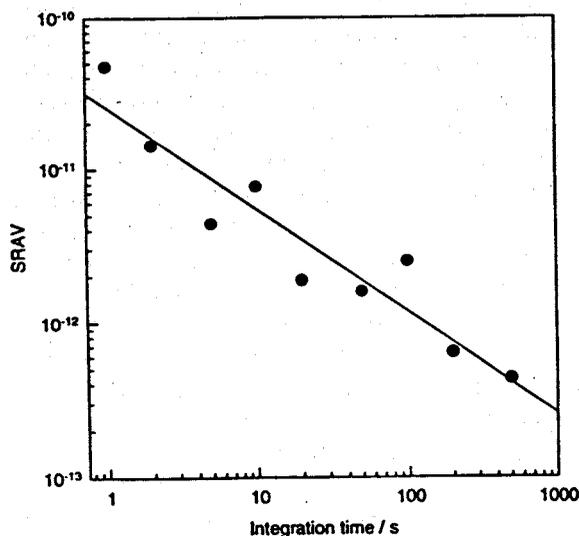


Figure 6: Allan variance (SRAV) of the stabilised oscillator as a function of sampling time

error signal is fed back to the voltage control input of a r.f. signal generator (Marconi 2031) at 20 MHz which forms part of the low frequency synthesis loop. The frequency of the stabilised source has been down converted and the IF counted by a Stanford Research SR620 counter. This counter is referenced to the 10 MHz signal from a Hydrogen maser, situated in a building some 50m from the cryogenic resonator. Already, using this remote intercomparison, Allan variance values as low as  $4 \times 10^{-13}$  at an averaging time of 500 s have been achieved (see figure 5). A second sapphire loaded resonator has been constructed and a high  $Q$  Nb cavity resonator is being evaluated, both of which can be similarly stabilised.

## CONCLUSIONS

The stability reported here is not the ultimate expected from the sapphire loaded superconducting resonator. Improvements in  $Q$  value and temperature control should provide several orders of magnitude increase in stability. The remote intercomparison method may already provide a dominant limiting factor and it will be necessary to intercompare a number of similar superconducting oscillators to explore the ultimate stability, which is expected to be better than 1 in  $10^{16}$ .

## REFERENCES

1. Luiten A.N., Mann A.G. and Blair D.G., 1995, "Latest Results of the U.W.A. Cryogenic Sapphire Oscillator". 1995 IEEE International Frequency Control Symposium
2. Langham C.D. and Gallop J.C., 1993, "Development of a High Stability Cryogenic Sapphire Dielectric Resonator". IEEE Trans. Instr. & Meas., **42** (2), 96-98
3. Langham C.D. and Gallop J.C., 1995, "Performance of Dielectric loaded Superconducting Resonators as Frequency Standards", Inst. Phys. Conf. Ser. No 148, Applied Superconductivity, Edinburgh, 1199-1202
4. Jones S.K., Blair D.G. and Buckingham M.J., 1988, "Effect of Paramagnetic Impurities on Frequency of Sapphire-Loaded Superconducting cavity resonators" Electron. Lett., **24** 346-347

## DIELECTRIC LOADED HTS RESONATORS AS FREQUENCY STANDARDS AND LOW PHASE NOISE OSCILLATORS

J C Gallop<sup>1</sup>, C D Langham<sup>1</sup> & Farhat Abbas<sup>1,2</sup>

<sup>1</sup>National Physical Laboratory, Teddington, Middx. TW11 0LW, UK

<sup>2</sup>Dept. of Elec. Eng. & Electron., UMIST, Manchester M60 1QD, UK

### INTRODUCTION

NPL has considerable experience in research into, and development of, high Q dielectric resonators based on sapphire and conventional superconductors at liquid helium temperatures (1). Stable microwave loop oscillators have been built, based on high Q resonators and cryogenic microwave functional elements such as low noise semiconductor amplifiers, phase shifters and filters operating at liquid helium temperatures and more recently at 70 K. As a result of the recent availability of reliable low power cryocoolers it has become possible to design compact dielectric loaded resonators making use of ultra-low microwave loss properties of thin films of high temperature superconductors (HTS) such as  $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$ . In this paper we address some of the design issues associated with using such high Q resonators as the reference element for frequency standards and low phase noise oscillators.

### ADVANTAGES OF CAVITY RESONATORS OVER ATOMIC FREQUENCY STANDARDS.

The most obvious advantage for frequency standards requiring good short term stability arises from the much better signal to noise ratio that a cavity resonator can offer over that achieved with an atomic standard. This improvement arises from the fact that photons are essentially non-interacting, unlike atoms, so that the photon density in a resonator can be raised to a level of  $10^{27}/\text{m}^3$ , almost comparable with solid densities, whereas in an atomic standard the density must be *many* orders of magnitude lower to eliminate the effects of atomic interactions which give rise to pressure shifts, spin exchange & dipole-dipole coupling. The result of this improved signal to noise ratio is that, by using passive resonator standards rather than atomic standards, much better Allan variance data is achievable at short averaging times ( $1\text{ms} < \tau < 10^3 \text{ s}$ ).

### SAPPHIRE DIELECTRIC LOADED SUPERCONDUCTING RESONATORS

At present the best demonstrated clock performance for averaging times to  $10^3 \text{ s}$  is that of a high - Q Nb superconducting cavity resonator (2) which has provided an Allan variance  $\sigma(\tau)$  of  $3 \times 10^{-16}$  for an averaging time  $\tau$  of up to  $10^3 \text{ s}$ . The main limitations on performance of this system for greater  $\tau$  values came from sensitivity to minute tilt angle changes and effects of small temperature changes. Both limitations may be removed by use of dielectric loaded helium-cooled superconducting resonators. This gives the promise that dielectric resonators may be capable of providing the best clocks for averaging times  $\tau$  from ms to  $\sim 10^5 \text{ s}$  or longer

In these dielectric loaded resonators the high Q cavity contains a sapphire single crystal structure in which most of the electromagnetic field energy is confined. The temperature dependent surface reactance of the superconducting walls plays a much smaller role in determining the resonant frequency than for the cavity resonator and thus higher stability may be expected. At low temperatures sapphire has a much higher Young's modulus than metals and thus the structure is more rigid and stable. Cryogenic temperatures also dramatically slow ageing processes (not only in dead humans). Temperature control to  $1 \mu\text{K}$  is readily achieved, unlike at room temperature, and the 3rd Law of Thermodynamics ensures that thermal expansion coefficients are much reduced. (Again sapphire performs much better than metals in this respect: the linear thermal expansion coefficient  $\alpha = 5 \times 10^{-12}/\text{K}$  at 2 K, 20 times less than for Nb at the same temperature and  $4 \times 10^6$  lower than for Cu at room temperature). The temperature stabilisation requirements are further relaxed by reducing the fractional frequency shift with  $T$  (1,3). Thus, to first order,  $(1/f)df/dT$  can be made equal to zero at a convenient operating temperature selected by control of paramagnetic atom doping levels, without compromising the high Q achievable as a result of the very low losses in sapphire at cryogenic

temperatures ( $\tan\delta = 2 \times 10^{-11}$  at 4 K).

## DIELECTRIC LOADED RESONATORS USING HIGH TEMPERATURE SUPERCONDUCTORS

In practice the use of liquid helium as a coolant in a non-laboratory environment does not present the major problems which are sometimes perceived. However for many applications questions of availability of liquid cryogen clearly make its use more problematic. (An extreme example is presented by the potential use of cryogenic resonator frequency standards in space applications). The development of reliable low power closed cycle coolers producing  $\sim 0.5$  W of refrigeration at 60 K from  $\sim 10$  W of input power, coupled with the recent production of HTS thin films with very low surface resistance (4) in the same temperature range ( $R_s(60\text{ K}) \sim 100 \mu\Omega$  at 10 GHz for YBCO) means that very compact low power dielectric resonators should be realisable with a total volume of not more than  $10^{-3} \text{ m}^3$ . Increasing the operating temperature from 4 K to 60 K does increase the dielectric loss in ultra pure sapphire to a value of  $\tan\delta = 10^{-7}$  but this is still low enough to not seriously compromise performance of the stable oscillators (5). At NPL we are currently working on a temperature compensation scheme to allow the resonator eigenfrequency to have a turning point in its variation with temperature at around 60 K. We have developed low-noise microwave amplifiers in the range from 4 GHz to 11 GHz which operate from the lowest temperatures to above 80 K. HTS band pass filters have been designed and evaluated in collaboration with a number of other groups and both semiconducting and superconducting phase shifters are being investigated. Below we discuss some of the design issues involved in these resonators with particular emphasis on achieving a resonant frequency which is to a high degree independent of temperature changes.

### Geometry Factor Considerations

To calculate the effect of the various contributions to both the loaded  $Q$  value  $Q_l$  and the temperature dependence of the resonator frequency it is necessary to assign a geometry factor  $\Gamma$  to each element of the resonator. Thus  $Q_l$  can be written in the following form:

$$\frac{1}{Q_l} = \frac{1}{Q_d} + \frac{1}{Q_c} + \frac{1}{Q_e}$$

where  $1/Q_d$  is proportional to the microwave power dissipation in the dielectric,  $1/Q_c$  is the loss in the superconductor and  $1/Q_e$  represents the loss in all other components such as radiated power, and dissipation in the metal (typically copper) container. For material parameters dielectric loss tangent  $\tan\delta$  and superconductor surface impedance  $Z_s$ , appropriate geometry factors  $\Gamma_d$  and  $\Gamma_c$  are introduced which are integrals of the field distribution over the volume of the dielectric and the surface of the superconductor respectively (see for example ref.4). The geometry factor representing the additional external losses can only be specified for a precise design including losses in the normal metal of the resonator housing and coupler losses. It is realistic to assume that these losses may be made negligibly small when compared with the previous two contributions. Thus

$$\frac{1}{Q_l} = \frac{\tan(\delta)}{\Gamma_d} + \frac{R_s}{\Gamma_c}$$

### Surface Impedance of YBCO

The electromagnetic loss of a superconductor at a finite frequency  $\omega/2\pi$  is not zero although the losses can be very small. For conventional superconductors, with a well defined isotropic energy gap  $\Delta(T)$ , for frequencies  $\omega/2\pi$  well below the gap frequency  $2\Delta/h$  the surface losses (characterised by the real part of a surface impedance  $R_s(T) = \text{Re}(Z_s(T))$ ) vary exponentially with inverse temperature - reflecting the density of quasiparticle excitations (normal electrons) in the superconductor. The reactive part of the surface impedance shows a similar temperature variation from a large limit  $X_s(0)$ :

$$R_s = \frac{A\omega^2}{T} e^{-\Delta/kT}$$

$$\Delta X_s(T) \propto e^{-\Delta/kT}$$

Both of the above parameters have an important role in determining the performance of frequency standards and low phase noise oscillators based on

superconducting resonators. The first determines the superconductor contribution  $Q_s$  to the overall quality factor  $Q_t$  of the resonator, this is more significant in discussing the performance of low phase noise oscillators. The latter parameter  $\Delta X_s$  determines the sensitivity to temperature variations of the resonant frequency and has more significance for the long term stability of the reference, that is of the device treated as a frequency standard or clock.

The simple behaviour summarised in the above equations would not be expected for a superconductor which exhibits a superconducting order parameter with d-wave symmetry. In this case both  $R_s(T)$  and  $\Delta X_s(T)$  are expected to vary as  $T^p$  where  $p$  is in the range 1-2. There is a growing body of evidence that the YBCO order parameter has such symmetry (6) although we, along with others, have shown that for some samples a good fit to the above equations may be obtained with a value of  $\Delta \sim 0.6 \times kT_c$ . The question of the order parameter symmetry is an important one for some HTS microwave resonator applications since both the ultimate  $Q$  and frequency stability achievable will be influenced by the outcome of this question. At this stage the ultimate  $Q$  value achievable at around 60 K does not seem to be seriously compromised by the probable existence of d-wave symmetry in YBCO (although at much lower temperatures this conclusion may no longer be valid).

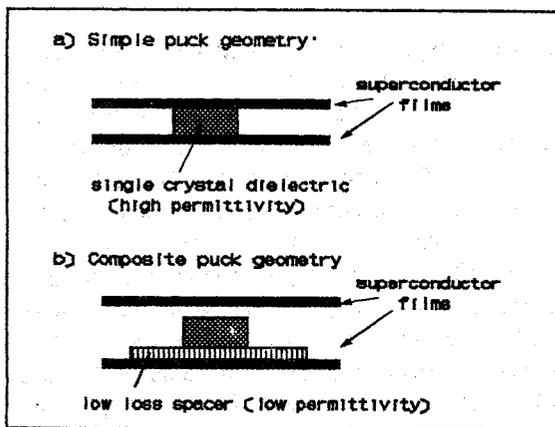


Figure 1 Schematic of dielectric puck resonator geometries a) simple b) compound

Fig.1a shows schematically the simplest design for a HTS shielded dielectric resonator. Note that the dielectric 'puck' (a short cylindrical element) is sandwiched between two planar thin or thick films of a suitable HTS material (for microwave applications this tends to be either  $\text{YBa}_2\text{Cu}_3\text{O}_{7.5}$  or  $\text{TlBaCaCuO}$  since these materials have demonstrated lower microwave surface resistance than any of the other wide range of HTS materials). Using this

design  $Q$  values as high as  $10^7$  at 5.6 GHz and  $T = 60$  K have been achieved by Shen et al. (7).

## Single Crystal Dielectric Materials

Sapphire is a widely used dielectric material at liquid helium temperatures. It has a uniaxial anisotropic permittivity in the range 9-11.5 and its thermal expansion coefficient, Young's modulus, high thermal conductivity and very low dielectric loss coefficient (when carefully grown and treated) make it ideal (3). At higher temperatures  $T$  such that  $10 \text{ K} < T < 80 \text{ K}$  the properties are less ideal, though still superior to anything that can be attained at ambient temperature. We have measured the temperature dependence of the relative permittivity of sapphire in the  $ab$  plane and also consider below the utility of other low loss dielectrics.

## TEMPERATURE DEPENDENCE OF THE RESONANT FREQUENCY OF A DIELECTRIC RESONATOR

The temperature dependence of the resonator frequency (usually the  $\text{TE}_{011}$  mode is used in these miniature dielectric resonators) arises from three main contributions: the variation of the superconducting penetration depth  $\lambda(T)$ , the linear thermal expansion of the dielectric puck  $\alpha(T)$  and the temperature dependence of the puck permittivity  $\epsilon_r(T)$ . We have made measurements of  $\lambda(T)$  recently (Clark et al. (8)) and shown that for reduced temperatures  $t = T/T_c \leq 0.2$  the fractional variation of the penetration depth is less than 0.05%/K for YBCO. The contribution that this would make to the frequency shift of the resonator depends on the geometry factor for the superconducting surfaces. Increasing the volume to surface ratio will reduce the importance of the superconductor contribution.

At higher temperatures such that  $t > 0.5$  the superconductor contribution begins to dominate over the other temperature coefficients. This is the situation which would arise with an HTS shielded dielectric puck resonator cooled in a single stage Stirling cycle refrigerator to around 60 K. In such a case it is more practical to use a compound puck geometry, as indicated schematically in fig. 1b. In this scheme a low permittivity, low loss dielectric, such as single crystal quartz, is inserted as a spacer between the high permittivity puck and the planar HTS film. An identical spacer may be used between the upper HTS film and the puck or this may be left

as a vacuum gap, whichever is more practical for the particular application envisaged. In each case the discontinuity in permittivity at the dielectric-dielectric interface acts as a dielectric mirror, having the effect of reducing the electromagnetic field strength at the superconductor surface compared with that within the puck. This of course reduces the influence of the temperature variation of the superconductor penetration depth on the resonator frequency by reducing the effective geometry factor for the superconductor  $\Gamma_c$ .

By using 2mm thick quartz spacers it has proved possible to reduce the effect of the superconductor penetration depth temperature dependence below that arising from the temperature dependence of the sapphire permittivity which has been found to have a temperature dependent component which, between 15 K and 80 K, varies as

$$\Delta\epsilon_r(T) = 2.5 \times 10^{-11} T^4$$

for several samples of single crystal sapphire from different sources.

### Design Compensation of Temperature Dependence

Below are listed (in approximate order of importance) some contributions to the temperature variation of the resonant frequency of a puck resonator, together with the sign of  $df/dT$ .

Contribution:	Parameter	Sign
HTS penetration depth	$\lambda(T)$	-ve
Dielectric permittivity	$\epsilon(T)$	-ve
Thermal expansion coefficient	$\alpha(T)$	-ve
Paramagnetic susceptibility	$\chi(T)$	+ve

In order to produce a frequency versus temperature curve which has a turning point at a convenient operating temperature (such as 50 K) it is necessary to have two dominant terms with opposite signs for  $df/dT$ . We propose here that such a turning point may be produced in a sapphire puck at high

temperature by deliberately doping the material with a paramagnetic impurity (for example chromium) in order to increase the magnitude of the paramagnetic susceptibility contribution. Calculations of the effect of the above temperature coefficients for a simple puck resonator suggest that  $f(T)$  for the  $TE_{011}$  mode at around 9.4GHz should have the following form for a  $Cr^{3+}$  doping level of 40 ppm:

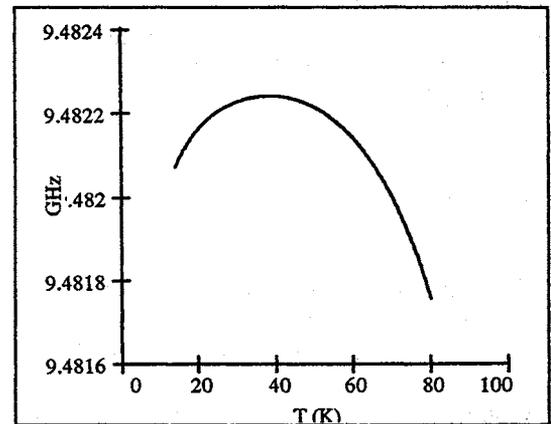


Figure 2 Predicted  $TE_{011}$  mode frequency versus temperature for 40ppm  $Cr^{3+}$  doping.

This level should be low enough to not seriously affect the loss tangent of the sapphire at the temperature of the turning point. If a resonator is operated at this turning point temperature great frequency stability should be attainable with only moderate temperature stabilisation.

The composite resonator design shown schematically in fig. 1b may also allow suitable temperature compensation. First the presence of low permittivity regions above and below the dielectric puck act as 'dielectric mirrors' in tending to confine the electromagnetic energy within the puck, thereby increasing the effective geometry factor for the superconductor plates and thus the influence of  $\lambda(T)$  on  $f(T)$ . Furthermore there are low loss dielectrics which have the opposite sign for  $\epsilon(T)$  compared with most materials such as sapphire (for example  $SrTiO_3$  or  $TiO_2$  (9)) so that use of two suitable complementary dielectrics in a composite resonator may allow temperature compensation in a useful temperature range without the need for doping.

### CONCLUSIONS

Work carried out at NPL and elsewhere has already demonstrated the great potential of helium cooled dielectric loaded superconducting resonators as clocks with the highest stability (parts in  $10^{15}$  for

averaging times extending from 1 s to  $10^4$  s). More recent developments, (including space qualified coolers with tens of thousands of hours of demonstrated life testing) suggest that similar performance may be demonstrated in a compact, light-weight system. Either system could provide excellent performance, exceeding that of some atomic standards in, for example, Ground Segment applications demanding the precision necessary to exploit the full potential of Global Navigation Satellite Systems. The light-weight version may also provide radical new capabilities for the Space Segment of such a system

9. Klein N, Tellmann N, Zuccaro C, Swiatek P & Schulz H, 1995, 'YBCO shielded high permittivity dielectric resonators for oscillators and filters', *Applied Superconductivity* 1995 2 743-8

## REFERENCES

1. C D Langham & J C Gallop, 1993, 'Development of high stability cryogenic sapphire dielectric resonator', *IEEE Trans. Instrum. Meas.* 42 96-8
2. S R Stein & J P Turneure, 1978, *Proc. AIP Conf. No. 44 (Charlottesville)* 192
3. M E Taber & D G Blair, 1994, 'Phase noise analysis of the sapphire loaded superconducting niobium cavity oscillator', *IEEE Trans. MTT* 42 344-7
4. J C Gallop & W J Radcliffe, 1994, 'Comparison of microwave measurements in HTS thin films by calorimetric and perturbation methods', *Physica C* 235 3369-70
5. F Abbas, L E Davis & J C Gallop, 1994, 'Ultra-high Q resonators for low-noise microwave signal generation using sapphire buffer layers and superconducting thin films', *Supercond. Sci. Technol.* 7 495-501
6. Kirtley J R, Tsuei C C, Sun J Z, Chi C C, Lock See YJ, Gupta A, Rupp M & Ketchen M B, 1995, 'Symmetry of order parameter in the high  $T_c$  superconductor YBCO', *Nature* 373 225-8
7. Shen Z Y, Wilker C, Pang P, Holstein W L, Face D & Kountz D J, 1992, 'High  $T_c$  superconductor-sapphire microwave resonator with extremely high Q values up to 90K', *IEEE Trans. MTT* 40 2424-32
8. Clark J, H, Donaldson G B, Gallop J C & Bowman R M, 1995, 'Temperature dependence of surface impedance of pulsed-laser deposited YBCO thin films' *IEEE Trans. Appl. Supercond.* 5 2555-8

## ENHANCED TRANPOSED GAIN MICROWAVE OSCILLATORS

M.A. Page-Jones and J.K.A. Everard  
 Department of Electronics  
 University of York  
 Heslington, York, YO1-5DD, England

### ABSTRACT

This paper describes the design of two low noise microwave oscillators operating at 7.6 GHz. These oscillators use room temperature Sapphire resonators operating in the  $TE_{01\delta}$  mode which demonstrate unloaded Q's of 44,000. Silicon transposed gain amplifiers are used to produce near-flicker-free gain. Oscillators with flicker noise corners around 1 KHz and a phase noise of  $-143\text{dBc/Hz}$  at 10 KHz offset and  $-121\text{dBc/Hz}$  @ 1KHz are demonstrated. This is within 3 dB of the theoretically predicted minimum noise. This performance is 12 dB better than in oscillators presented at the 1995 IEEE MTT conference and 1995 FCS conference<sup>1,2</sup>. This improvement is achieved by using high level mixers.

### INTRODUCTION

Microwave Oscillators are usually built using GaAs active components. These GaAs devices demonstrate high flicker noise corners which greatly degrade oscillator noise performance close to carrier. Silicon transistors have low flicker noise corners but also have low gain at microwave/mm wave frequencies. It is however fairly easy to fabricate silicon Schottky barrier mixer diodes operating at microwave frequencies.

This paper describes the design of two low noise 7.6 GHz oscillators with flicker noise corners of around 1KHz and a noise performance better than  $-143\text{ dBc/Hz}$  at 10 KHz offset and  $-121\text{dBc/Hz}$  @ 1KHz. This noise is within 3 dB of the theoretical minimum for a Sapphire dielectric resonator oscillator with an unloaded  $Q_0$  of 44,000 and a power available at the output of 1 mW. The noise changes at a rate of  $1/\Delta f^2$  down to around 1KHz. These oscillators use the transposed gain of a low frequency silicon amplifier and demonstrate a suppression greater than 58 dB of the reference LO phase noise.

### TRANPOSED GAIN

Transposed gain is achieved using two mixers, a low frequency amplifier and a local oscillator as shown in Figure 1. This upconverts the LF gain to microwave frequencies on both sides of the reference LO. The frequency of the local oscillator (which can be noisy) is therefore used at a frequency close to that required for the gain. The noise from the local oscillator is significantly suppressed ( $>58\text{ dB}$ ) in an oscillator using transposed

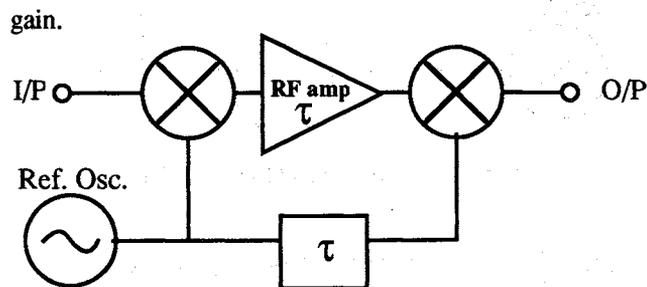


Figure 1. Transposed Gain Amplifier (TGA)

### NOISE CONVERSION

It is important to ensure that the noise performance of the LO does not degrade either the AM or PM noise of a signal passing through the Transposed Gain Amplifier (TGA).

Phase noise on the local oscillator can cause an output phase modulation because of the group delay of the amplifier. This is corrected by introducing a delay line with the same delay (and dispersion) characteristics of the amplifier in between the LO drive of the two mixers. In these experiments the delay of the low frequency (inverting) amplifier is around 1.3 ns. The delay of the amplifier and mixers matches the delay line to less than 20 ps.

AM on the reference oscillator has two effects on the output of the TGA. The first is to vary the mixer conversion losses, effectively transferring AM to the output. The second effect of LO drive variation is to alter the phase shift of the TGA. This is due to the mixer impedance changing with drive level. This is less serious with a broadband mixer. The LO should therefore have low AM by using an oscillator which incorporates 'hard limiting'.

### OSCILLATORS

With the addition of a resonator a TGA can be made into a transposed gain oscillator (TGO) as shown in Figure 2. Because of the cancellation of FM and AM described above, such an oscillator is shown to have much lower phase noise than the original local oscillator.

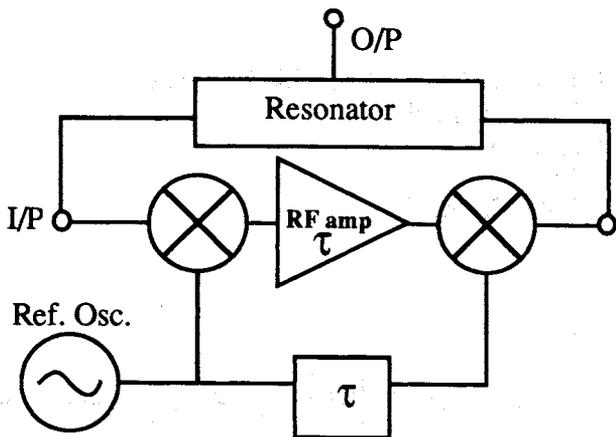


Figure 2, Transposed Gain Oscillator.

The LF amplifier consists of two bipolar transistors in parallel as shown in Figure 3.

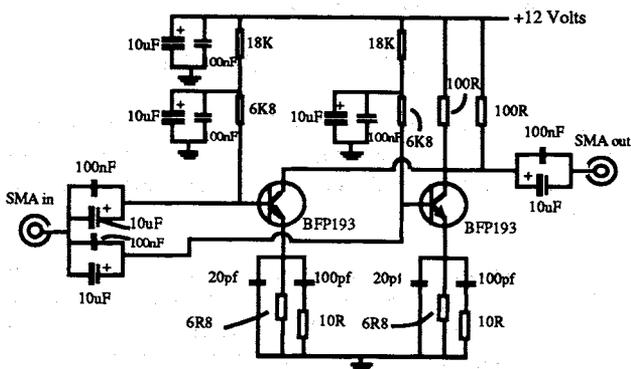


Figure 3 10KHz to 200 MHz Amplifier

A flicker noise measurement system has been built (by this group<sup>3</sup>) for measuring the cross correlation between baseband gate and drain noise and the AM and PM components that these transpose onto a carrier. This is achieved by measuring all the noises simultaneously. These measurements enable prediction of the equivalent circuit noise sources of the FET and the cross correlation coefficients between them. The final system has a residual noise floor of -180 dBc for offsets greater than 3 KHz.

The measured flicker noise corners of the LF amplifier at baseband (Figure 3) and mixers were 30KHz and less than 3 KHz respectively. The flicker noise corner of the oscillator was around 1 KHz. It should be noted that although the baseband and transposed corners are related, their ratio depends on the transposition (modulation) mechanism. In these oscillators the amplifier is not non-linear because the limiting occurs in the output mixer. The IF port of the output mixer is driven with +18dBm and the Local Port with +20dBm. The output power of the second mixer is then around 10dBm.

**Sapphire Dielectric Resonators**

The oscillators use room temperature Sapphire resonators operating in the fundamental  $TE_{01\delta}$  mode. These resonators are operated substantially as a two port with a lightly coupled third port to provide a filtered output. The unloaded Q (including the effect of the O/P port) is measured to be 44,000.

The mounting method is depicted in figure 4. The sapphire dielectric 'puck' is mounted on a hollow quartz support, connected to a height adjusting screw in the end face of the box.

The end of the quartz rod is castellated to reduce the contact area with the sapphire, as this improves the resonator Q factor. The sapphire is attached using an ultraviolet curing glue, type NOA61 manufactured by Norland. This glue is sold as an optical adhesive, and its losses are also low at microwave frequencies. The metal container is turned from three pieces of oxygen free copper, and the RG402 feeders to the coupling loops are held by compression fittings threaded into the sidewalls. The length of the quartz rod and its adjustment screw were set to give the highest possible Q. The highest Q was obtained with the puck vertically centred in the box. The sapphire pucks were 1/2" diameter 1/4" thick disks, manufactured as cryostat windows by Union Carbide. The optic axis coincides with the cylinder axis (a 'zero degree' cut), and the boules are grown by the Czochralski method. The highest Q sapphire resonators reported have used the heat exchange growth method, but this seems to be most significant at cryogenic temperatures as described by Flory and Tobar<sup>4</sup>.

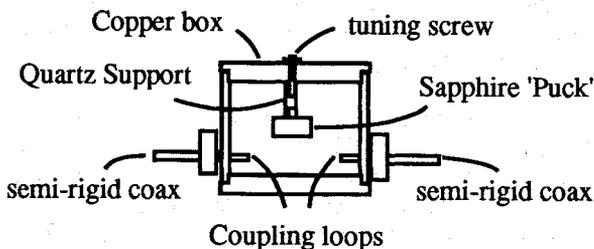


Figure 4

**Temperature Dependence of Resonators**

The resonator temperature dependence is the sum of contributions due to the expansion of the container and supports, and the changes in the dielectric constants of the sapphire itself. The effect of expansion of the box (equivalent to around 1ppm per degree) is masked by the dielectric changes in sapphire. Shelby and Fontana<sup>5</sup> have shown that the dielectric constant of sapphire changes by between 30ppm and 80ppm at room temperature, depending on the orientation of the electric field with respect to the crystal axes.

A test resonator was operated in an environmental chamber over the range from -20 to +75 °C, while monitoring Q factor and resonant frequency. Fitting a

straight line to this data gave a temperature coefficient of around -40 ppm per °C.

In a practical oscillator design the resonator temperature would have to be controlled. The ratio of thermal conductivity to specific heat is such that thermal changes take several seconds to travel through the resonator. We have demonstrated thermal control to better than 1mK with a single stage system. If better performance is required a two stage system could be implemented.

### EXPERIMENTAL MEASUREMENTS

To evaluate the transposed gain method two Transposed Gain Oscillators have been built at 7.6 GHz operating 13 MHz apart. The phase noise of the beat signal was then measured using the phase detector method (in conjunction with the low noise, HP8662A, signal source) as shown in Figure 5.

To obtain minimum noise performance using Thermal/additive noise the insertion loss of the resonator is set to 6 dB<sup>1-3, 6-10</sup>, producing a loaded Q of 22,000. The insertion loss of the mixers is around 7 dB therefore the open loop gain of the silicon amplifier should exceed 20 dB. The total noise figure of the TGA which incorporated the two mixers and the low frequency amplifier was around 9.5 dB. The single sideband output power available from the O/P mixer was around 10 dBm. The local oscillator used for each oscillator was an HP 8672A synthesised signal generator. A

GaAs based oscillator using similar dielectric resonators could also have been used. The noise performance is shown and tabulated in Figure 4. These measurements demonstrate a noise performance better than -143 dBc/Hz at 10 KHz offset (assuming that both oscillators are similar and uncorrelated, -140dBc/Hz -3dB) and a flicker noise corner of around 1 KHz. This noise is within 3 dB of the theoretical minimum for a Sapphire DR oscillator with an unloaded  $Q_0$  of 44,000 and an available amplifier power of 10 milliwatt. The equation used to describe the noise performance  $L_{fm}$  under optimum operating conditions for thermal noise is <sup>8</sup>

$$L_{fm} = (2FkT/Q_0^2P)(f_0/\Delta F)^2$$

where F (9.5dB) is the noise figure under the correct operating conditions, k is Boltzmanns constant,  $Q_0$  (44,000) is the unloaded Q, P (10dBm) is the power available at the output of the TGA,  $f_0$  (7.6GHz) is the centre frequency and  $\Delta f$  is the offset frequency. This equation incorporates the optimum operating condition that the ratio of  $Q_L/Q_0 = 1/2$  and therefore the gain of the amplifier is 6 dB.

The typical noise performance of the HP8672A signal generator at 10 KHz offset is -85 dBc/Hz. The transposed gain oscillator is therefore suppressing this noise by more than 58 dB.

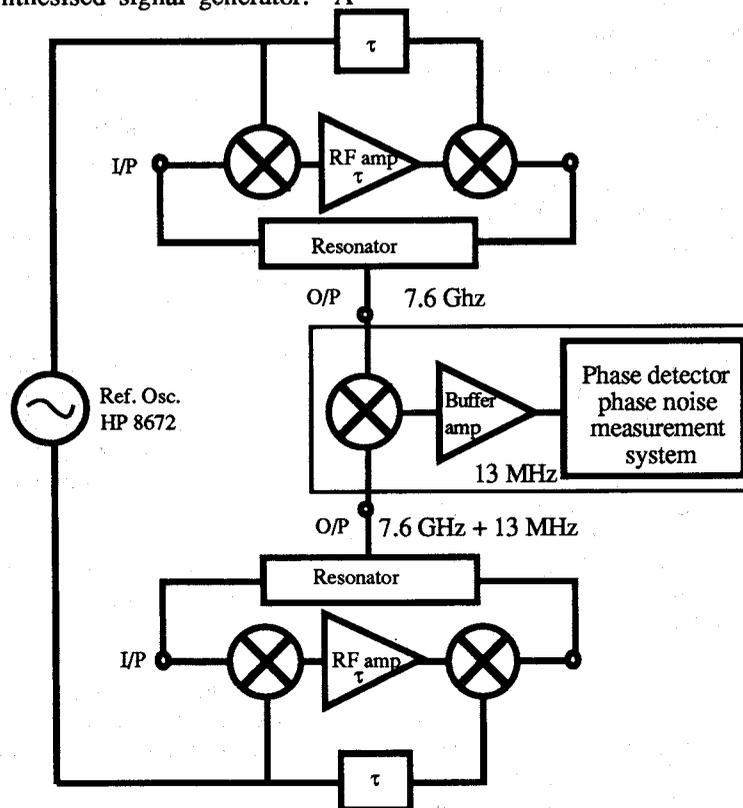


Figure 5. Experimental Oscillators and Phase Noise Measurement System

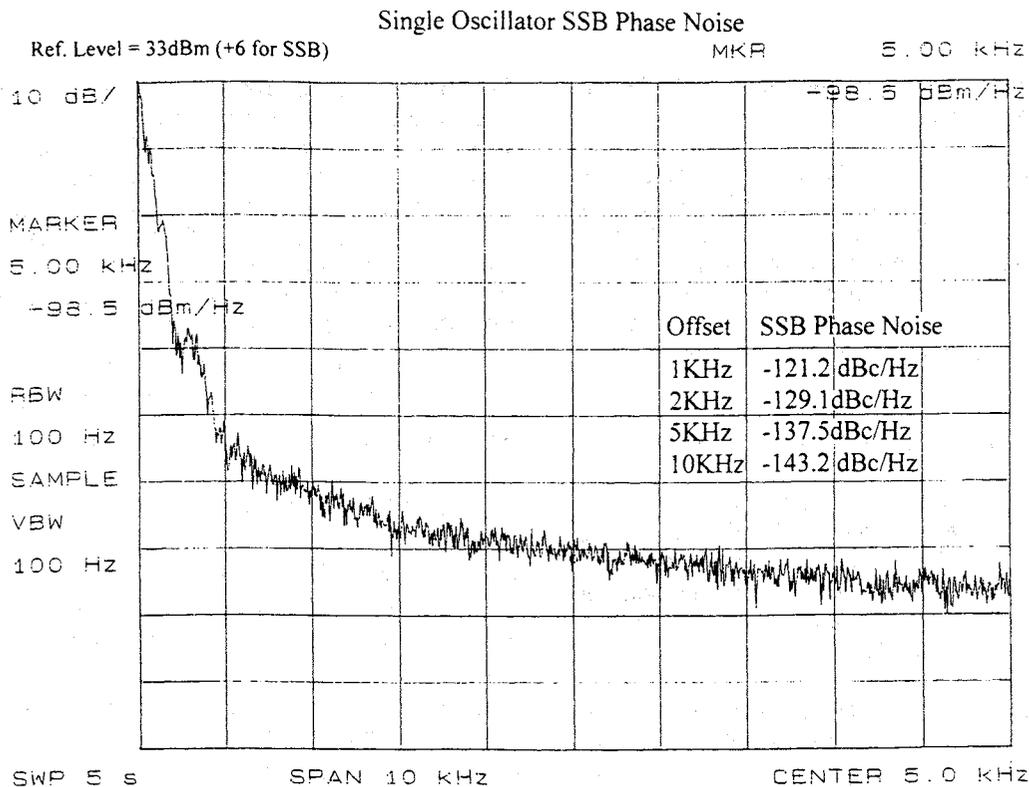


Figure 6 Phase Noise of the Beat Signal of Two Transposed Gain Oscillators at 7.6 GHz

#### FUTURE WORK

Higher unloaded Q, Higher power operation and zero IF operation are currently under investigation. It is expected that a further noise improvement of >30 dB can be achieved using these methods.

#### REFERENCES

1. J.K.A. Everard and M.A. Page-Jones, "Ultra Low Noise Microwave Oscillators with Low Residual Flicker Noise", IEEE International Microwave Symposium, Orlando May 16th -20th 1995, pp. 693-696.
2. J.K.A. Everard and M.A. Page Jones "Transposed Gain Microwave Oscillators with Low Residual Flicker Noise, Frequency Control Symposium, SanFrancisco June 1995.
3. P.A. Dallas and J.K.A. Everard. "Measurement of the Cross Correlation between Baseband and Transposed Flicker Noises in a GaAs MESFET". IEEE Microwave Theory and Techniques Conference, Dallas USA, May 1990, pp. 1261-1264.
4. C.A. Flory and R.C. Taber "Microwave oscillators incorporating cryogenic Sapphire Dielectric Resonators" IEEE Proceedings International symposium on frequency control 1993
5. R. Shelby and J. Fontanella and C. Andeen, "The Low Temperature Electrical Properties of some Anisotropic Crystal", J.Phys. Chem. Solids, Vol. 41, pp.69-74
6. T.E. Parker, " Current Developments in SAW Oscillator Stability". 31st Annual Frequency Control Symposium 1977, pp.359-364.
7. J.K.A. Everard. "Low Noise Oscillators" IEEE Microwave Theory and Techniques Conference, Albuquerque New Mexico, 1-5th June 1992, pp.1077-1080.
8. Editors D. Haigh and J.K.A. Everard. "Gallium Arsenide Technology and its Impact on Circuits and Systems". Published by Peter Peregrinus Ltd., August 31st 1989. Chapter 8 (J.K.A. Everard) on Low noise oscillators, pp. 237-280.
9. J.K.A. Everard "Low Noise Power Efficient Oscillators: Theory and Design" IEE Proceedings Part G, Vol. 133, No.4, August 1986, pp. 172-180.
10. T.E. Parker, " Characteristics and sources of phase noise in stable oscillators". 41st Annual frequency control symposium 1987.pp.99-110.

## GPS TIME TRANSFER USING CARRIER-PHASE AND P-CODE MEASUREMENTS

G rard Petit, Philippe Moussay and Claudine Thomas

Bureau International des Poids et Mesures  
Pavillon de Breteuil, F-92312 S vres Cedex, France

### ABSTRACT

GPS time transfer is usually carried out with C/A code measurements using the common-view technique. An uncertainty of a few nanoseconds can be obtained for an integration time of about 1000 seconds when ionospheric measurements and precise orbits are used. This limits the capability of frequency comparison to the level of 1 to 3 parts in  $10^{14}$  for an integration time of 1 day.

Recently new frequency standards have been developed that state an accuracy at or below  $1 \times 10^{-14}$ . For existing time transfer techniques it is a challenge to compare their frequency over a sufficiently short interval. Previous studies have already shown the potential of carrier-phase and P-code measurements from geodetic GPS receivers for the comparison of internal receiver references with an improved performance.

Several experiments on time and frequency transfer based on phase and P-code measurements are related in this paper:

- \* a zero baseline comparison at the BIPM using data from two AOA TTR4-P receivers connected to a HP 5071A clock, and
- \* short baseline comparisons (of several hundreds of kilometres) using data from TTR4-P units in operation at the BIPM and from AOA Turbo Rogue receivers in operation at the ORB, Brussels, Belgium, and the IFAG, Wettzell, Germany, both connected to a hydrogen maser.

Preliminary results show encouraging results for frequency transfer carried out with Turbo Rogue phase measurements. However, noise as high as 250 ps on 30 s phase measurements, and thus larger than expected, superimposed on unexplained excursions up to 0,5 ns for 1-2 hours periods, has been detected in results from experiments involving the TTR4-P unit belonging to the BIPM. Collaboration with the maker is already on the way for deeper investigation of this point.

### INTRODUCTION

The quality of time metrology depends on the performance of atomic clocks and on the means of time and frequency comparison. For the last five years, the wide-spread use of the GPS common-view technique has led to major improvements in precision and

accuracy, making it possible to compare the best standards in the world at their full level of performance: for integration times of 1 to 3 days, the frequency differences between atomic clocks are measured to about one part in  $10^{14}$ . However, recent developments in atomic frequency standards suggest that this performance may no longer be sufficient. The new caesium fountain FO1 built at the BNM-LPTF, Paris, France, shows a short-term white frequency noise characterised by an Allan deviation  $\sigma_y(\tau=1 \text{ s}) = 2 \times 10^{-13}$  and a type B uncertainty of  $3 \times 10^{-15}$  ( $1 \sigma$ ) [1]. Frequency standards based on linear ion traps also show an excellent short-term stability: their white frequency noise is detected at a level of  $\sigma_y(\tau=1 \text{ s}) = 6 \times 10^{-14}$  [2]. To compare the frequencies of such high-performance standards would call GPS common-view results to be averaged over 10 days of continuous operation, a time duration presently unrealistic for these devices. Previous studies have already shown the potential of using carrier-phase and P-code measurements from geodetic GPS receivers for clock comparison [3, 4, 5, 6]. This paper describes preliminary results from experimental work which compares atomic clocks using phase and P-code measurements from AOA TTR4-P and Turbo Rogue units respectively connected to a HP 5071A clock at the BIPM and hydrogen masers at the ORB, Brussels, Belgium, and the IFAG, Wettzell, Germany.

### GPS CODE AND PHASE MEASUREMENTS

The frequency of the clocks on board GPS satellites is 10,23 MHz, and all other frequencies transmitted to ground are derived from this. Each satellite transmits an L1 signal at a frequency of 1575,42 MHz ( $10,23 \text{ MHz} \times 154$ ) and an L2 signal at a frequency of 1227,6 MHz ( $10,23 \text{ MHz} \times 120$ ). Signal L1 is modulated with the P-code (Precise code) and the C/A-code (Clear/Acquisition code) while signal L2 is modulated with the P-code only. Both codes are of pseudorandom noise (PRN) type. The respective chip-rates of the C/A- and P- codes are 1,023 MHz and 10,23 MHz, which results in an improved precision when using P-code measurements.

The C/A- and P- codes give access to pseudorange (or apparent propagation time) measurements which are determined by measuring the difference between two identical codes:

- \* One is generated by the receiver and is synchronized by its internal reference (a 5 MHz

output from the receiver quartz, controlled by a 5 MHz signal issued from a local clock - HP 5071A or hydrogen maser - kept in the laboratory).

\* The other is contained in the signal emitted by the satellite and received by the receiver; it is synchronized by the satellite clock.

Most GPS time receivers are one channel C/A code devices operating on the L1 frequency and connected to the 1 pps output of the local clock. They deliver C1 pseudoranges (L1 C/A code measurements) from which are deduced time comparisons, processed over 13 minute periods, between the local clock and GPS time, the time scale broadcast by the satellites. These receivers are programmed according to the International common-view schedule issued by the BIPM which suggests, at most, 40 daily tracks. The results, referred to in the following as 'classical common views', are used for international time links. The different sources of uncertainty which affect time comparison between two clocks on the Earth through the common-view method are described elsewhere [7]. Assuming that GPS antenna coordinates are known with an uncertainty of a few centimetres ( $1 \sigma$ ), that strict common-view (synchronized within 1 s) are used to overcome Selective Availability effects and that precise satellite ephemerides are available, the accuracy limitation (uncertainties given for  $1 \sigma$  and for 13 minute tracks) then comes from:

- \* estimation of the ionospheric delay, using a model (a few ns for short-distance links),
- \* C/A code measurements noise (1 ns),
- \* multipath propagation (2 ns),
- \* estimation of the tropospheric delay (0,5 ns using temperature, humidity and pressure measurements [8]), and
- \* differential receiver calibration (1-2 ns).

In the following we neglect differential or absolute calibration of the devices involved in order to focus on frequency transfer accuracy (or time transfer precision).

The ionosphere being a dispersive medium, the related noise can easily be reduced using an appropriate linear combination of dual-frequency measurements (referred to as ionosphere-free linear combination, the coefficients of which are respectively 2,54 and -1,54 for the L1 and L2 measurements). In addition, decoding the P-code makes it possible to reduce the noise in code measurements to about 1 ns for 30 s measurements. The TTR4-P receiver made by Allen-Osborne Associates (AOA) combines these two advantages. In addition, this model based on the design of the Turbo Rogue geodetic receiver, tracks up to eight satellites simultaneously with a basic measurement interval which may be set at 1s to 3600 s, two features which considerably increase the number of available measurements. In the following, P1 and P2 pseudoranges (for L1 and L2 P-code measurements) are taken every 30 s even when Anti-Spoofing is active

(with some degradation of the precision) thanks to the original P-codeless mode developed by AOA.

The Turbo Rogue and TTR4-P models also provide phase measurements on the two frequencies L1 and L2 and thus their ionosphere-free linear combination. These measurements are provided by beating the L1 and L2 frequencies with locally synthesized frequencies and have a precision of about 1/1000 of the period, i.e. of the order of 1 ps. This helps to reduce the measurement noise, but phase measurements show an ambiguity of an integer number of cycles (the L1 and L2 periods are respectively 0,63 ns and 0,81 ns), an ambiguity which is not completely solved in the results given in the following.

Phase measurements are also much less sensitive to multipaths effects [9], already reduced by mounting the antenna on a choke ring plane. To reduce uncertainties in the estimates of the tropospheric delay would call for a mapping of the troposphere deduced from global processing of phase data taken by a number of GPS receivers spread world-wide.

TTR4-P models are adapted to time measurements thanks to an interface between a local 1 pps signal and the 5 MHz internal reference of the receiver. It is thus possible to extract the time differences between the local clock and the satellite clock (REFSV) or the local clock and GPS time (REFGPS) with a sample interval of 30 s. The quantities REFSV and REFGPS are computed by the receiver from the ionosphere-free linear combination of P1 and P2 pseudoranges and can be used to apply the classical common view method.

## EXPERIMENTAL SET-UP

This section describes several experiments using P-code and phase measurements carried out in 1995 at the BIPM and between the BIPM, the ORB, Brussels, Belgium and the IFAG, Wettzell, Germany.

### At the ORB and the IFAG

The ORB and the IFAG have at their disposal Turbo Rogue receivers for which the internal reference is derived from the 5 MHz provided by local hydrogen masers with very good short-term stability. The data used for this paper cover two days of L1 and L2 phase measurements taken within the framework of the International GPS Service for Geodynamics (IGS).

### At the BIPM

A TTR4-P receiver (serial number 109) has been in operation at the BIPM since several years. We had also at our disposal a TTR4-P receiver (serial number 120) belonging to the NPL, Teddington, United Kingdom, for a three month period in 1995.

One of the most important problems encountered with early TTR4-P models is poor operation of the antenna. This was solved by an antenna upgrade provided by AOA for unit 109. Unfortunately the antenna shipped with the unit 120 does not make it possible to carry out measurements with sufficient gain. We thus operate both units in a common-antenna mode, the local HP 5071A clock provides a common 1 pps input and a common 5 MHz internal reference to both receivers (zero-baseline configuration).

A problem encountered with the TTR4-P model is frequent resetting of the station coordinates. This occurs even in 'Clock' mode for which, theoretically, the coordinates are held fixed. It should also be noted that the internal reference has frequent resets so it is not possible to have continuous data for more than one day on average. A last drawback is that the output of data blocks on the serial port of the device does not operate correctly at sample intervals up to 30 s. It is thus necessary to purchase a Flashcard which allows storage of the complete measurement lines (including C1, P1 and P2 pseudoranges and L1 and L2 phases) for any sample rate.

#### ZERO BASELINE EXPERIMENT

The first check was to observe individual satellites from their appearance above the horizon to complete disappearance. This led to the detection of an anomaly in the 30 second data REFGPS of three satellites PRN 12, 2 and 21. This is shown in Fig. 1 for PRN 12, which is not affected by SA, and should thus present a smooth profile during the 6 hours of continuous tracking. Instead, structures with amplitudes as large as 40 ns are observed and cannot be assigned to poor measurements at low elevations, for which errors would be random, or to coordinates, ephemerides or troposphere errors which would be very slowly varying. The origin of this anomaly is probably an erroneous computation of the distance between the satellite and the receiver. It follows that a selection of satellites must be used to obtain a safe classical common-view time transfer using TTR4-P units. It should also be noted that the time difference between the local clock (1 pps input) and the internal reference of the receiver is not output in the data block and is thus unknown, a drawback to linking phase measurements to quantities such as REFGPS and REFSV.

A second test is to compute the differential quantity  $[REFSV_{109} - REFSV_{120}]$  obtained from observation of the same satellite at the same moment by the TTR4-P units 109 and 120 in the zero baseline configuration. Individual 30 s points obtained over a 12 hour period from all visible satellites are shown in Fig. 2. The values obtained are specified to 1 ns. This is another disadvantage of the TTR4-P model: it outputs data with too low a resolution (1 ns) so it is not possible to apply the Technical Directives for standardization of GPS

time receiver software [10]. In Fig. 2, the standard deviation of one individual point is about 4,1 ns, which corresponds to subnanosecond noise for the classical 13 minute track length of a single satellite. In the chosen configuration, most of the sources of uncertainty are cancelled, but not the measurement noise from the two receivers: P pseudoranges and links between the 1 pps input and the internal reference. The level of noise which is observed here is thus larger than expected from the P pseudorange measurements themselves (3-4 ns for individual measurements instead of the expected 1-2 ns).

Time transfer using GPS phase measurements has also been carried out with the zero baseline configuration. The results of phase differences, obtained with a sample interval of 30 seconds from the two TTR4-P units, are shown in Fig. 3 for all satellites visible over one day (the cycle ambiguities are not solved in an absolute way). The observed short-term noise is about 150-200 ps which is much larger than expected from phase measurements. In addition, structures are apparent, up to 0,5 ns for 1-2 hour periods. Though not explained, these may be due to excursions of the internal reference in one TTR4-P unit. The results of phase measurements in the zero baseline configuration are thus disappointing.

#### SHORT BASELINE EXPERIMENT

A comparison of time transfer results between the HP 5071 A clock at the BIPM and the hydrogen maser at the ORB is shown in Fig. 4 for a one day period and for all visible satellites: the sample time is 30 s, values are deduced from ionosphere-free combinations of dual-frequency P-code and phase measurements (the TTR4-P receiver used at the BIPM is unit No 109). The standard deviation of individual points is 2,5 ns for P pseudoranges, which is the expected level for the case of a non-zero baseline, and 250 ps for phases, which seems to be too high.

To clarify the situation with phase measurements, a more detailed analysis was conducted satellite by satellite, both BIPM TTR4-P units being involved. An example for PRN 21 (see Fig. 5) shows noise on individual 30 s points of about 250 ps for the unit 109 and 50 ps for the unit 120. There are also structures with amplitude of several hundreds of picoseconds for averaging times of several tens of minutes. The intrinsic noise of the HP 5071A clock in operation at the BIPM explains parts of the observed noise: it presents white frequency noise for averaging times  $\tau$  such that  $10 \text{ s} \leq \tau \leq 10^4 \text{ s}$  at a level of about  $\sigma_y(\tau) = 2,7 \times 10^{-12}$  for  $\tau = 10 \text{ s}$ . This integrates into a random walk phase noise of about 45 ps over  $\tau = 30 \text{ s}$  and to 270 ps over  $\tau = 1000 \text{ s}$ . The phase measurement noise is thus dominated by the clock noise in the short baseline experiment between the BIPM and the ORB when TTR4-P unit No 120, belonging to the NPL, is

used. However, it is predominant when measurements are taken with TTR4-P unit No 109, belonging to the BIPM. This suggests that this receiver operates poorly when used for phase measurements.

The same analysis is shown in Fig. 6 for the time link between the ORB and the IFAG, where local clocks are hydrogen masers. The short-term term noise (about 10 ps) indicates the high performance of the Turbo Rogue receivers for phase measurements. Taking into account the excellent short-term stability of the local hydrogen masers, the observed structures, of amplitudes several tens of picoseconds, probably comes from poor estimation of the tropospheric delays.

## CONCLUSIONS

Several experiments in time and frequency transfer using dual-frequency P-code and L phase measurements have been conducted using AOA TTR4-P units in operation at the BIPM. These show that, even if the precision obtained with P pseudoranges is nearly correct, the level of noise obtained with short-term phase measurements is higher than expected, at least for one TTR4-P unit. In addition, structures of amplitude 0,5 ns for averaging times of 1-2 hours, probably due to instabilities of the internal reference of this particular unit, limit frequency transfer to several parts in  $10^{15}$  over one day. It should also be noted that the TTR4-P model is not well adapted to accurate classical time transfer via common views for at least two reasons: resets in the station coordinates and anomalies in the computation of the distance satellite - ground station.

A test using two Turbo Rogue receivers connected to hydrogen masers at the ORB and the IFAG shows that phase measurements can provide time transfer with a precision of order 10 ps over averaging times of 30 s. In the longer term, a limitation to frequency transfer at the level of a few parts in  $10^{16}$  over one day is due to poor estimations of the tropospheric delays. Tropospheric mapping deduced from a global treatment of phase measurements obtained with a number of GPS receivers spread world-wide could help to improve frequency transfer to about 1 part in  $10^{16}$ , a performance which would then probably be limited by hardware problems.

## Acknowledgements

The authors are grateful Dr J.P. Chambers from the NPL, Teddington, UK, for the loan to the BIPM of a TTR4-P unit.

## References

1. Clairon A., Ghezali S., Santarelli G., Laurent P., Lea S.N., Bahoura M., Simon E., Weyers S. and Szymaniec K., 1995, "Preliminary accuracy evaluation of a cesium fountain frequency standard", Proc. Symp. on Frequency Standards and Metrology, in press
2. Maleki L., 1996, personal communication
3. Dunn C.E., Lichten S.M., Jefferson D.C. and Border J.S., 1991, "Sub-nanosecond Clock Synchronization and Precision Deep Space Tracking", Proc. 23rd PTTI, 89-99
4. Dunn C.E., Jefferson D.C., Lichten S.M., Thomas J.B., Vigue Y. and Young L.E., 1993, "Time and Position Accuracy Using Codeless GPS", Proc. 25th PTTI, 169-179
5. Douglas R.J. and Popelar J., 1994, "PTTI Applications at the Limits of GPS", Proc 26th PTTI, 141-152
6. Baeriswyl P., Schildknecht T., Utzinger J. and Beutler G., 1995, "Frequency and Time Transfer with Geodetic GPS Receivers: First Results", Proc. 9th EFTF, 46-51
7. Lewandowski W. and Thomas C., 1991, "GPS time Transfer", Proc. IEEE, 79, 7, 991-1000
8. Lewandowski W., Grudler P. and Baumont F., 1994, "Study of tropospheric correction for GPS common-view time transfer between the BIPM and the OCA", Proc. 8th EFTF, 981-997
9. Allan D.W. and Thomas C., 1994, "Technical Directives for Standardization of GPS Time Receiver Software", Metrologia, 31, 69-79
10. Jaldehag K., 1995, "Space Geodesy Techniques: An Experimental and Theoretical Study of Antenna Related Error Sources", ed. Chalmers University of Technology, Göteborg, Sweden, 276, 23-27

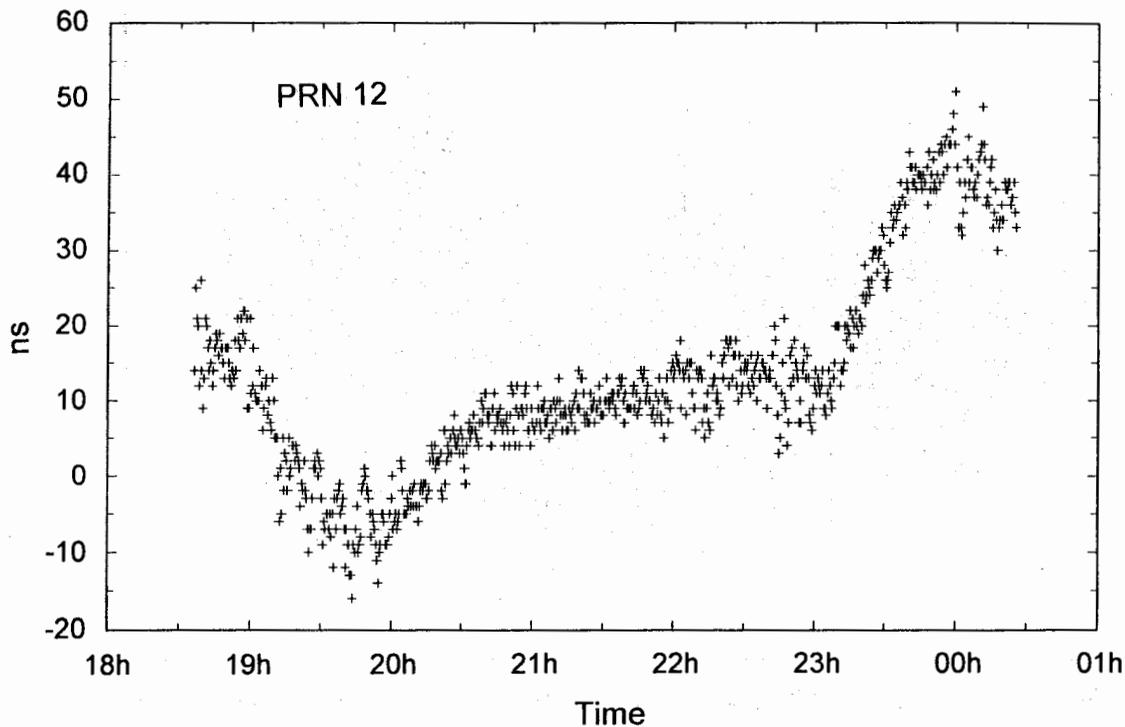


Figure 1. Time differences REFGPS between the local clock and GPS time as delivered at a 30 s sample rate from observation of PRN 12 by a TTR4-P receiver (serial number 109) in operation at the BIPM.

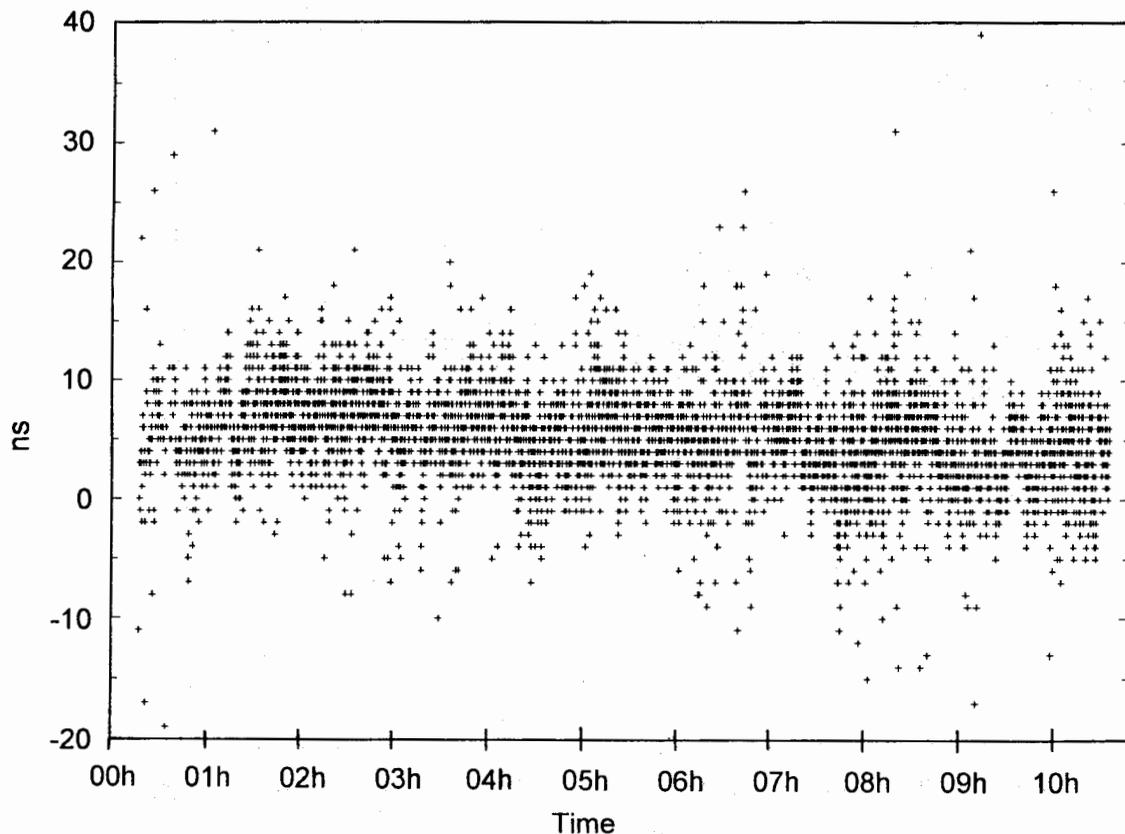
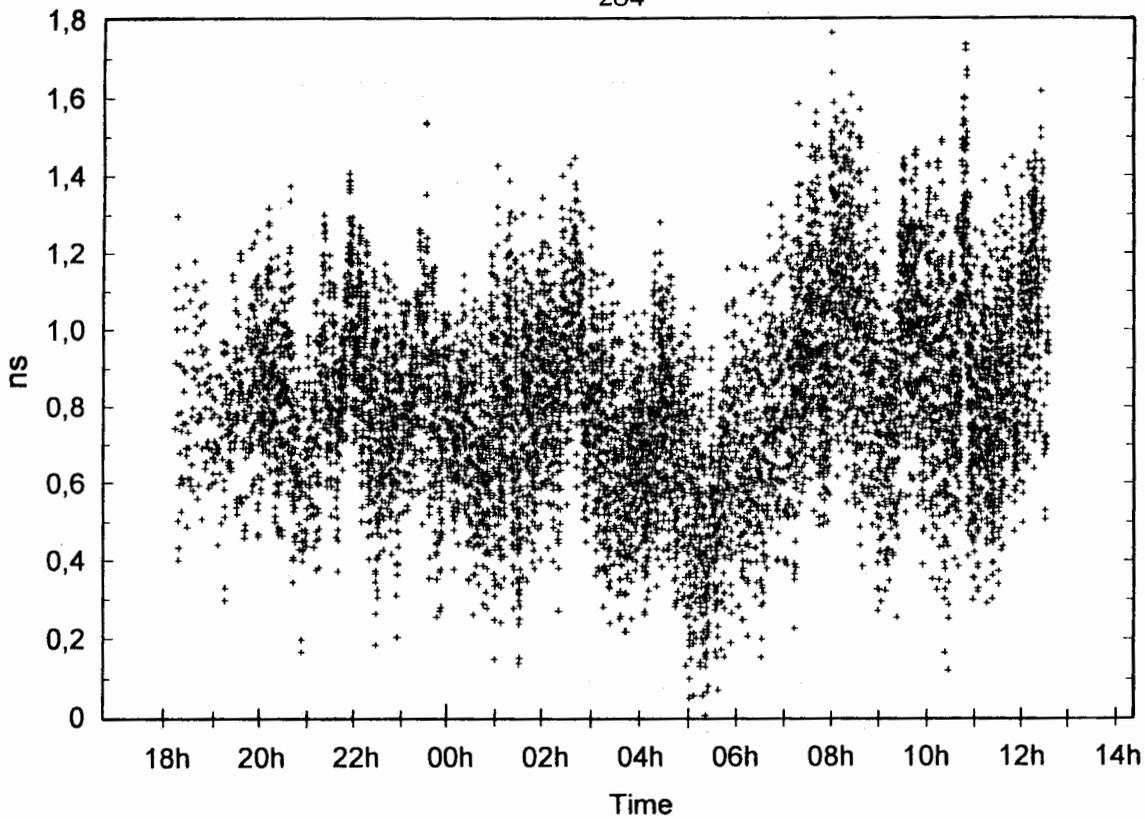
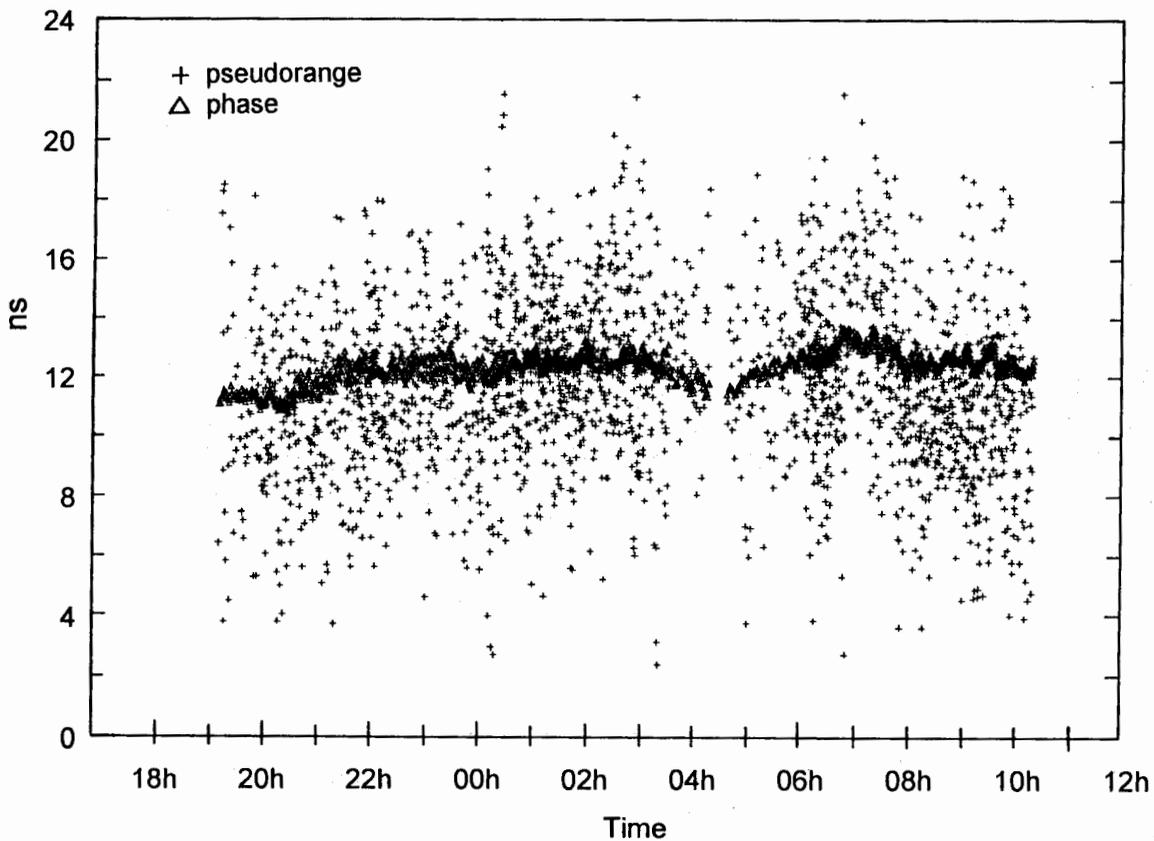


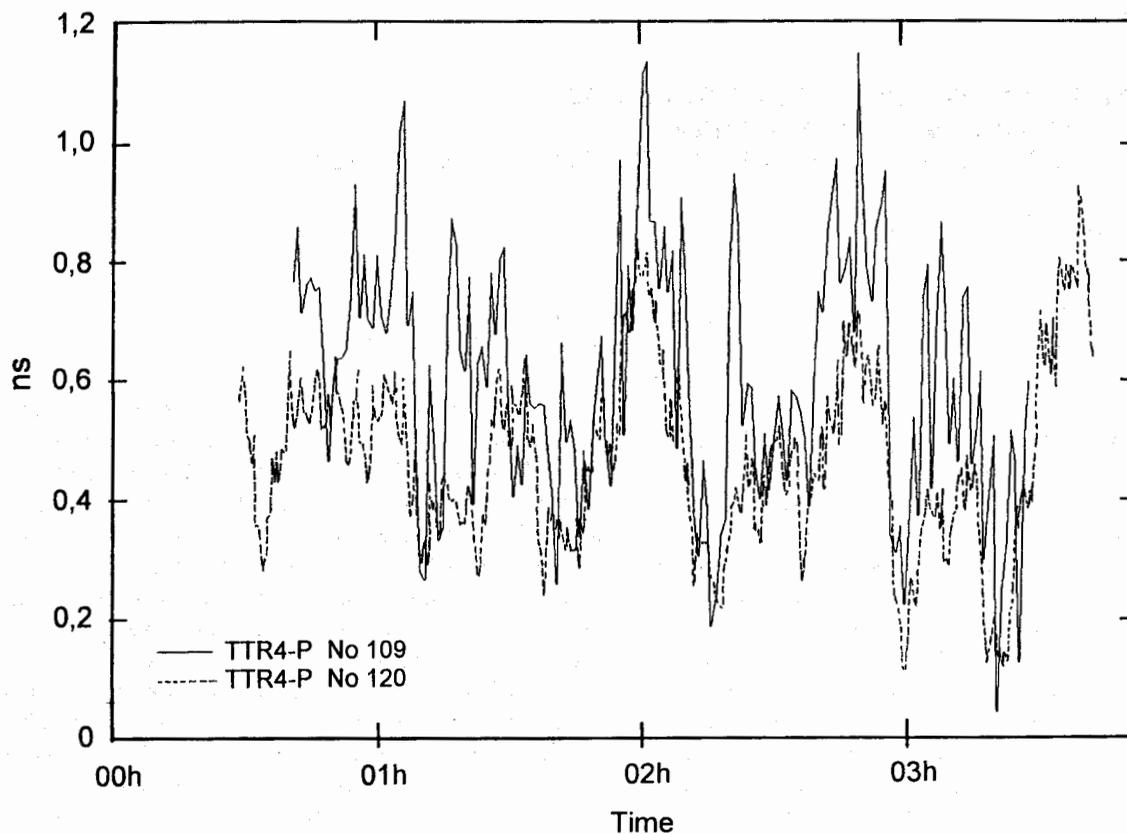
Figure 2. Zero-baseline REFGPS differences obtained from observation of all visible satellites at a 30 s sample rate. The involved GPS receivers are two different TTR4-P units (serial numbers 109 and 120) at the BIPM.



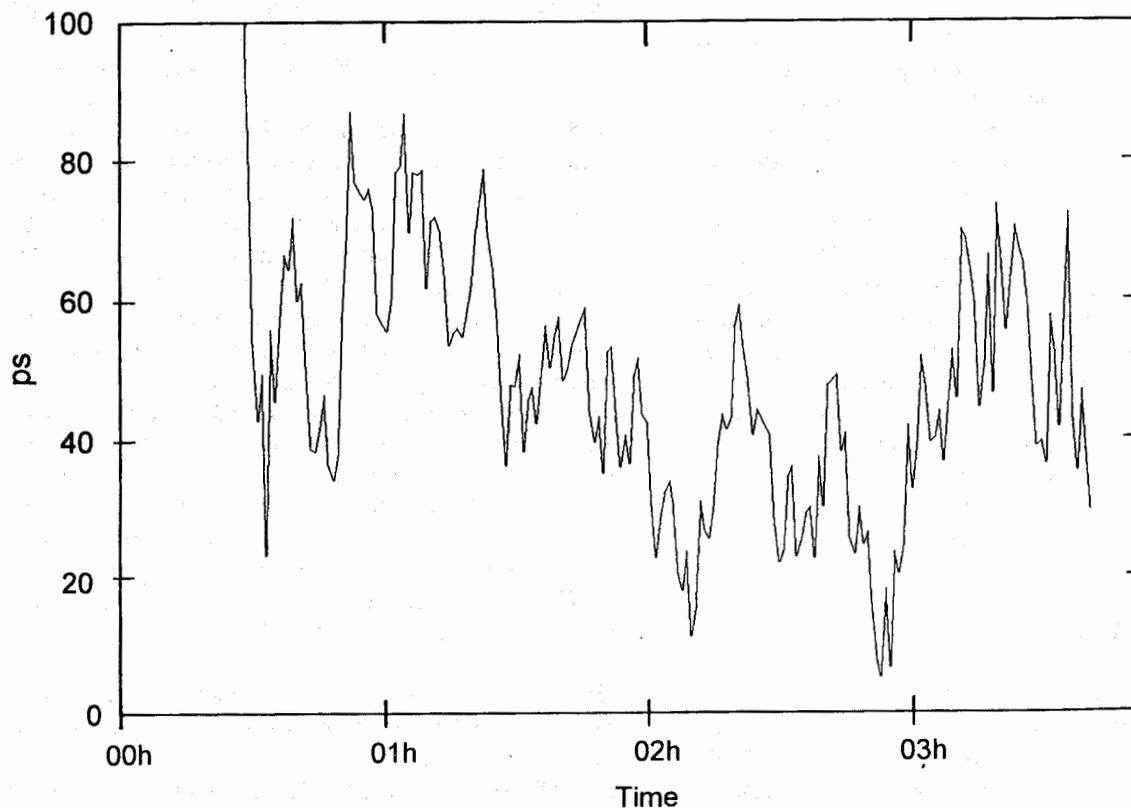
**Figure 3.** Zero-baseline phase differences obtained from observation of all visible satellites at a 30 s sample rate. The involved GPS receivers are two different TTR4-P units (serial numbers 109 and 120) at the BIPM.



**Figure 4.** Pseudorange and phase differences between the ORB hydrogen maser and the BIPM HP 5071A clock obtained from observation of all visible satellites at a 30 s sample rate. The involved GPS receivers are a Turbo Rogue unit at the ORB and a TTR4-P unit (serial number 109) at the BIPM.



**Figure 5.** Phase differences between the ORB hydrogen maser and the BIPM HP 5071A clock obtained from observation of PRN 21 at a 30 s sample rate. The involved GPS receivers are a Turbo Rogue unit at the ORB and two different TTR4-P units (serial numbers 109 and 120) at the BIPM.



**Figure 6.** Phase differences between the ORB and the IFAG hydrogen masers obtained from observation of PRN 21 at a 30 s sample rate. The involved GPS receivers are Turbo Rogue units.

## TIMING OPTIONS AND FREQUENCY CONTROL CONCEPTS FOR MEETING USER REQUIREMENTS WITH THE SECOND-GENERATION GLOBAL NAVIGATION SATELLITE SYSTEM (GNSS2)

<sup>1</sup>Clifford C. HODGE and <sup>2</sup>John A. DAVIS

<sup>1</sup>Basic Metrology Section, Centre for Quantum Metrology

<sup>2</sup>Time & Frequency Services, Centre for Electromagnetic and Time Metrology

National Physical Laboratory, Queens Road, Teddington, Middlesex, United Kingdom TW11 0LW

### ABSTRACT

This paper begins with a brief insight into the Sixties' conception of the global positioning systems for the United States (GPS) and the former Union of Soviet Socialist Republics (GLONASS). These satellite systems were developed with the requirements of the military in mind. But, the Nineties have brought increasing pressure upon the GPS controllers to release these systems into civilian control. We predict that as this happens, new key market areas will emerge, though the users will require not just a global **positioning** service, but a global **navigation** service. We identify some key market areas for the early part of the 21st Century. Frequency-band allocation issues will play a key role in defining the system performance for meeting the new user requirements: a complex system using L-band through to J-band for the ground-to-space-segment (and space-to-ground-segment) will be shown to have distinct advantages over existing L-band systems, whilst also allowing for cost recovery mechanisms. The use of K-band frequencies and above, through to the optical, are being actively researched, and may be the preferred link for inter-satellite communication. A brief summary of clock developments for the next twenty years is given, and likely predictions where clocks could be used in the ground- and space-segment are made. A space-segment payload clock frequency in the microwave, rather than acoustic region, is called for, if anticipated technological developments in the 21st Century are to be realised. But, quartz resonators are not dismissed, having a role to play in microwave clock synchronisation. The need for a truly global timing system is highlighted with an example from typical UTC data: but, this initiative will require close international collaboration to succeed. The relative benefits of time-transfer techniques are evaluated, with one-way, common-view and two-way time-transfer having differing roles to play. Two-way time-transfer to geostationary satellites is likely to play a significant role in underpinning the system performance. Synchronisation of clocks in the space- and ground-segment are described in the framework of the next-generation of Global Navigation Satellite Systems (GNSS2).

### 1) INTRODUCTION

The tearing down of the Berlin Wall in 1990 heralded the start of a new chapter in History: the Cold War, which had raged on since the Sixties, was finally at an end. In the intervening thirty-year period huge advances were made in space exploration, which had led into spin-off application

areas such as global digital telecommunications and global satellite positioning systems. During the Cold War period, the United States (US) developed the Global Positioning System (GPS), whilst the former Union of Soviet Socialist Republics engineered the Global Navigation Satellite System (GLONASS) with comparable capability (Thomas A. HERRING, 1996). Both systems were developed, and have been maintained under, military control. The GPS network currently consists of 24 satellites, the last of these was launched and has been in operation since 13th December 1995. It might have been sooner had the Challenger disaster not occurred in 1987, which halted the US Shuttle Programme for three years, and reduced the Space Shuttle launch rate from two per month to two per year. Similarly, the final satellite in the GLONASS constellation was launched, and became fully operational, in January 1996.

In the years that followed the cessation of the Cold War, the military markets have diminished: but, civilian applications look set to fill that void. This is being partly driven by cost, as the civil market becomes enticed into using GPS by the receiver manufacturers: for instance, U.S.\$ 350 now buys the user a receiver with nominal 30 metre precision anywhere on the Earth, with 1 second update capability.

GPS is a military system, and this is apparent from the intentional degradation in the form of Selective Availability (S/A), that is added to the satellites' ephemerides, thereby limiting the accuracy obtainable at the user's receiver. Of course, the US can switch S/A on, or off, at any time (as evidenced during the 1992 Gulf War). This has been a matter of deep concern to European civilian users considering whether to meet future sole-means navigation requirements using GPS for aviation. But, despite the pressure to remove S/A mounting, and efforts to reduce its effect having been published (D. W. ALLAN, 1995), rumours that S/A will be removed completely in the near future have failed to materialise. This has prompted Europe to investigate the feasibility of working towards a wholly civil European Global Navigation Satellite System (GNSS1), which could be followed by a next-generation (potentially stand-alone) system, GNSS2.

Growth in the global mobile communications market is set to explode in the next twenty years. By being involved in satellite communication activities now, Europe will hopefully be able to take advantage of the future commercial opportunities that arise. In 1996, a tri-partite consortium

consisting of ESA, EUROCONTROL and the European Commission will be participating in the exploitation of four INMARSAT-3 satellites to augment the existing GPS and GLONASS satellite constellations. They will bring increased accuracy, integrity, availability and continuity of service to the GPS and GLONASS user community. Likewise, governments are grasping the opportunities being made available: for example, the United Kingdom's Heritage Secretary recently announced that the Government would be signing up to a package of eighteen digital channels to be broadcast in the year 2005 from six geostationary satellites.

Frequency standards and time-transfer technology have routinely been developed in National Standards Laboratories. It takes industry up to ten years to develop this research for the commercial market. Over the years, the fundamental physical constants (such as time, length and mass) have been invariably traceable to the measurement of a stable frequency. During the last half-Century, metrologists have seen a factor of ten improvement in clock performance per decade: and, this progress appears to be showing no signs of slowing down. So, whereas Global Navigation Satellite Systems experts talked of 10 metre accuracy in the 1970's, the advances made with clock and time-transfer technology by the 1990's seems to suggest 10 cm accuracy in real-time being a real possibility with next-generation GNSS in the next decade (C. C. HODGE, 1996). The only barrier to meeting these objectives would appear to be political, rather than technological (Walter BLANCHARD, 1995), as issues of cost recovery and control become of paramount importance. In summary, the ability to measure time and frequency with increasing accuracy has meant that the 1970's satellite systems technology, which concentrated on global **position** has, in the 1990's, switched emphasis to global **navigation**.

Not wanting to fall behind in GNSS developments, Europe has instigated a study into the feasibility of a second-generation system (GNSS2). It has the potential to be a totally independent navigation satellite system under civilian control. But, meeting cost recovery issues will depend on the uptake of the civilian market to the notion of GNSS2. Therefore, it is crucial that GNSS2 fully meets the requirements of the future generation of users. We provide some key market areas that will find the availability of a real-time, high accuracy, GNSS attractive for navigation purposes.

## 2) USER COMMUNITY REQUIREMENTS

### 2.1.1 Civil Aviation Authorities

The transfer of GPS (and GLONASS) into civilian control is being driven by the civil aviation authorities, who wish to see GPS used for sole-means navigation during the phases of high-precision approach. Although the positional **accuracy** specifications are not too demanding (and recent trials using GPS appear to indicate that the required accuracy is achievable using Local Area Differential (LAD)

techniques, even for CAT III precision approach (P. KLASS, 1995)), it is the specification for **integrity**, stipulated by the civilian airport authorities, which is currently **unattainable** using the existing global positioning systems: so much so, that GNSS1 (GPS plus GLONASS plus INMARSAT-3) will only be able to fulfil near-CAT I precision approach (see TABLE 1). GNSS1 broadcasts ionospheric corrections by using a combination of C-band and L-band links in differential mode. But, because these corrections are transmitted to the user over a wide area, major problems could occur when there is unplanned failure in the up-link or down-link, affecting users over a wide area: the whole of Europe, for instance.

However, the driver for GNSS is there. The cost of maintaining existing Microwave Landing Systems (MLS) by international airport authorities is currently under scrutiny (The Role of GPS in European Civil Aviation, 1995), on the pretence that were GPS to provide sole means navigation capability, fulfilling civil aviation's CAT III integrity requirements, it may be a cheaper option to maintaining MLS capability. Still, even if the specification for CAT III integrity were attainable, the civilian airline community would be unlikely to accept having to be the sole financiers of a sole-means GNSS. With only 2 % of the potential user market, they might be asked to provide 50 % of the costs! So, there need to be other market areas for cost recovery.

Precision approach	Horizontal	Vertical	Integrity
Non-precision	33 metres	10 metres	T.B.D.
CAT I	19 metres	5.4 metres	$3.3 \times 10^{-7}$
CAT II	8 metres	2.5 metres	$3.3 \times 10^{-8}$
CAT III	4.5 metres	0.8 metres	$3.3 \times 10^{-9}$

TABLE 1: Civil aviation authority requirements for precision and non-precision approach phases of flight

### 2.1.2 Rail container monitoring and tracking

The potential market for rail container monitoring and tracking is vast. GNSS2 would stand to benefit the mass transport system by providing a service which reduced the dependence for expensive signalling systems. Rail transportation systems require special attention, in that the direction of carriage motion is predictable, but failure to classify whether the carriage is on the up-line or the down-line can have disastrous consequences. Any global satellite broadcast signal will also be blocked as the carriage passes through a tunnel. Dead-reckoning for interpolation, combined with laser diode sensing may offer a solution, but much research needs to be initiated. The rail transportation community would require 0.3 metres horizontal accuracy, with an update time between 0.1 seconds and 1 second, using on-board navigation systems capability for interpolation and integrity monitoring. Integrity figures for

rail transportation will depend on the number of nodes at a rail junction. So, the level of effort required to achieve the highest integrity would be directly related to the level of nodal activity. Initial trials for rail container monitoring using GPS experienced receiver drifts, and failed to evaluate wheel slippage (Chris de la FUENTE, 1993). But, this is a market sector of significant importance, and it should be pursued.

### 2.1.3 Understanding of earthquake shock-wave propagation

The understanding of earthquake shock-wave propagation is in its infancy stage, but remains an active research topic in the Western Continental United States and the Pacific Basin area of Japan and South-East Asia. New seismological events seem to bring unpredictable changes to theory, as the experts attempt to predict future events and the structure of fault lines (C. FROHLICH, 1995). GPS is already being used to determine the after-effects of earthquakes using Local Area Differential (LAD) techniques, using a "before-and-after" analysis. But, only through better understanding can scientists determine whether the underlying mechanism for these events is predictability or probability. Nominal accuracies of 0.03 metres, with 0.1 second to 1 second update may be required, which is likely to be difficult to achieve over a regional area.

### 2.1.4 Intelligent Vehicle ((Automated) Highway Systems (IVHS/AHS)

Undoubtedly the largest market (20 million users in Europe alone), it is therefore not surprising that the Departments of Transport in both Europe and the United States are keenly following developments in GNSS2. IVHS, or "in-car navigation" facilities, are already being incorporated by the vehicle manufacturers for the more expensive series of vehicles, and IVHS will continue to penetrate into the lower end of the market. But, with traffic levels on major highways throughout Europe continuing to grow, and demand exceeding supply, governments are investigating candidate road management schemes, from road tolls through to Automated Highway Systems (AHS). AHS (J.-P. BARBOUX, 1994) may prove to be too difficult to achieve, with accuracies in the 0.1 metres to 0.01 metres range, with update times of 0.1 seconds to 0.01 seconds. However, were GNSS2 able to meet AHS specifications, it would bring a good return on the investment.

### 2.2 When to meet User Requirements with GNSS2

Existing GPS and GLONASS systems currently fulfil user requirements down to approximately 10 metres without LAD and dual-frequency receiver techniques. In developing GNSS2, Europe should concentrate on meeting the needs of those users with more demanding positional and navigational requirements below the 10 metre threshold: competing on both cost and quality of service.

GPS (and GLONASS) seem to be adequate for non-precision approach for the civil aviation community,

whilst GNSS1 will aim to meet near-CAT-I requirements. But, it is the CAT II and CAT III accuracy and integrity that GNSS2 needs to concentrate upon. LAD techniques using up to three pseudolites in the vicinity of the CAT II/CAT III approach (radius of 1 km) with up to eight satellites in view, may satisfy airport authorities' specifications. GNSS2 is likely to comprise of a number of local services to international airports, which together would form a regional European service. This calls for a GNSS with spot beams for the civilian community. By comparison, rail monitoring only requires the highest integrity where rail lines meet, particularly in urban areas, or at termini, which makes the use of pseudolites and LAD at selected areas desirable.

But, because of the unpredictability of an earthquake over a vast area, and the user requirements on accuracy and update time, economies of scale may make the implementation of GNSS2 for real-time earthquake prediction a difficult market to penetrate. Similarly, AHS would require high accuracy and integrity over a wide area, with all the associated problems of dead-reckoning and signal attenuation through obstacles, particularly in urban areas. But, this remains a key cost recovery area and the solutions will need to be aspired to, if GNSS2 is to meet its goal.

## 3) FREQUENCY-BAND ALLOCATION

GPS broadcasts on two frequencies in the L-band ( $L_1$ : 1.575 GHz and  $L_2$ : 1.227 GHz), with GLONASS operating on L-band at around 1.6 GHz. Radio communications agencies are under increasing pressure to make more frequency-band space available. With the mobile communications industry also broadcasting in the L-band, the competition for frequency-band allocation in the L-band region (0.390 GHz - 2.5 GHz) is under intense pressure. But, an indication of how the International Telecommunications Union (ITU) are able to maintain frequency-band allocation integrity is demonstrated by the GLONASS system, which had, until recently, been transmitting in a radio astronomy L-band. The USSR has since been told to revise its transmission frequencies: satellites on diametrically opposite sides of the Earth now share frequencies, thereby preserving the band for the radio astronomy community.

Calls for backward compatibility of GNSS2 with GPS and GLONASS could be difficult to satisfy. Recent European Space Agency (ESA) studies indicate that to launch GNSS2 with L-band capability would cause unacceptable interference with existing GPS and GLONASS systems. But, another reason why GNSS2 should not become entirely dependent on L-band arises from interference due to ground-based communications systems, which has the potential to limit accuracy, and more critically, integrity performance for users requiring the highest performance. And, finally, L-band transmission suffers from ionospheric delays, which are inversely proportional to the square of the frequency, and are typically 10 nanoseconds for GPS in a

one-way time-transfer mode. Reliance on L-band requires modelling to remove both static and dynamic ionospheric corrections, which are then broadcast to the user, which inevitably impacts on system complexity and integrity performance.

S-band (2.5 GHz - 4.0 GHz) is largely assigned for research purposes, and ionospheric effects are still significant in the user error budget.

C-band (4.1 GHz - 7.0 GHz) has been largely maintained for telecommunications and research purposes. It forms a key service for both the INMARSAT-3 broadcast and, more importantly, for the civilian Microwave Landing System in the band 5.000 GHz - 5.250 GHz. At the World Administrative Radio Conference for the Mobile Services in 1987 (International Telecommunication Union, 1994), the protection of this vital aeronautical radionavigation system was expressed as being "paramount". The civil aviation community's reluctance to utilise GNSS has been in ensuring that there is backward compatibility between MLS and GNSS. As a result, the civil aviation community have recently approved the notion of a Multi-Mode Receiver (MMR), which integrates the hardware for Inertial Landing Systems (ILS), MLS and GNSS (primarily GPS). But, if civil CAT III precision approach requirements are to be met, GNSS2 transmission at MLS frequencies in the C-band would automatically ensure that backward compatibility of instrumentation occurred, without reliance on GPS and GLONASS developments.

Combined L-band and C-band (two-frequency, even multi-frequency) transmission would allow the ionospheric delay (and ionospheric delay delays) to be determined to high precision, and eliminated with respect to other error sources through broadcasting to the user in a local area.

X-band (7.0 GHz - 11.5 GHz) is also attractive, and has been much researched during the Cold War period by the military. This is, in part, due to the higher frequency and bandwidth available when compared with the L-band. But, the primary advantage of X-band is that it is sufficiently high in frequency to make the ionospheric contribution to the user range error (URE) budget sufficiently small, whilst being high enough in frequency such that rain fade due to tropospheric absorption is not a major problem. X-band requires antennae of typical diameter 2 metres, whilst being sufficiently directional to enable regional coverage to be obtained. During the Cold War period, much progress was made on the effects of multipath evaluation (Yaakov BAR-SHALOM, 1994 and Richard D. J. VAN NEE, 1992), which cannot be modelled easily, but must be determined directly by the user. But, despite the advantages of using X-band, frequency-band allocation for radionavigation, fixed- and broadcasting-satellite services will need to be considered through the official channels, in discussion with the International Telecommunications Union (ITU).

J-band (11.5 GHz - 18.0 GHz) does experience some rain fade attenuation, which seems to contradict its being of use for GNSS2 in times of heavy rain-fall. But, this frequency-band provides good directionality, and would seem to be the ideal candidate for broadcasting to geostationary satellites in real-time. The National Physical Laboratory (NPL), in partnership with five other European laboratories, have routinely participated in two-way satellite time- and frequency-transfers using an INTELSAT satellite at 307 °E (J. A. DAVIS, 1996). Although rain fade effects using two-way time-transfer are seen in the United Kingdom, signal outages are rarely experienced, and could be eliminated with higher power levels from the geostationary satellite. This idea would also build on ESA's experience gained during GNSS1 development using INMARSAT-3, with the potential for backward compatibility. By transmitting at spot frequencies in the J-band, it may also be possible to reconstruct the Kramers-Kronig relationship for tropospheric absorption, and evaluate the concentration of water in the path of the satellite and the user in real-time.

The use of K-band (18.0 GHz - 33.0 GHz) for space-to-ground- (and ground-to-space-) broadcasting is precluded by rain fade from the troposphere. However, K-band and higher frequency transmission (V-band, W-band, etc.) would be well suited for inter-satellite links between a network of geostationary satellites, which would formulate the basis of a global time- and frequency-dissemination service to the ground-segment. These ideas will be developed in the time-transfer section, later in this paper.

Laser ranging via inter-satellite links should also not be dismissed, where studies have suggested that 5 picoseconds timing accuracy is feasible (G. APPLEBY, 1994).

Two methods for cost recovery appear to have surfaced. The first involves the use of "smart cards", whereby the signal transmitted to the user is encrypted. But, the concept of a civilian system being encrypted has met with some resistance over who would control the system, and there is also opposition from the maritime community.

Ideas to transmit at a number of frequencies may add to the complexity of GNSS2, but there is justification for doing so. Where this may be to an advantage would be in offering different levels of service, dependent upon: the complexity of the receiver for obtaining certain frequencies (single L-band, single C-band and dual X-band, for instance) and the carrier coding rates used (10 MHz, 30 MHz or 100 MHz services on offer). By levying a payment on the complexity of the receiver, cost recovery mechanisms could be established, without the need for data encryption, which also adds to the complexity of the receiver, the transmission data-rate, and the complexity of the payload in the space-segment.

NPL are therefore proposing that future studies concentrate

upon:

- the use of L-band, and C-band (or X-band, or J-band) for the real-time determination of local ionospheric delay delays;
- the investigation of multi-path effects using multi-frequency C-band and X-band directional-array antennae receivers;
- the use of J-band for two-way time-transfer from the ground-segment to geostationary satellites for real-time satellite ephemeris determination, forming the basis for inter-satellite links to a global network of satellites;
- the use of multi-frequency J-band for the real-time determination of local tropospheric delay delays;
- the use of frequencies at K-band and above (including the optical region) for inter-satellite links utilising a geostationary satellite network (via two-way time-transfers) with (pseudo-ranging or) two-way time-transfer to orbiting satellites in the GNSS2 constellation.

#### 4) STABLE CLOCKS

FIGURE 1 provides a typical overview of Allan variance figures for state-of-the-art clock performances in 1995:

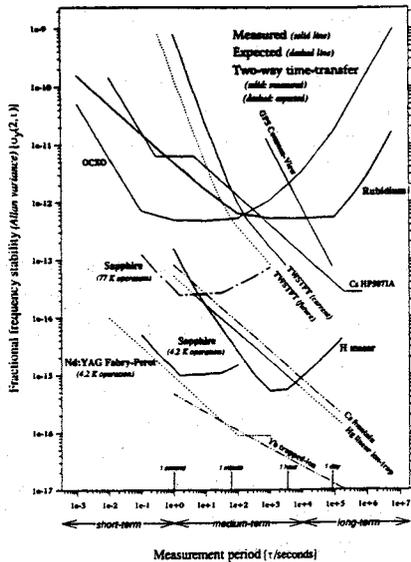


FIGURE 1: Allan variance versus measurement period for space atomic clocks

Over the years, frequency standard development has seen a steady progression towards higher stability, with the development of atomic clocks operating at higher frequencies (see TABLE 2). This trend looks set to continue. Of perhaps the greatest significance is the lack of a clear link between optical and microwave frequencies, which has precluded the realisation of stable frequency standards operating in the terahertz region (L. MALEKI, 1993). These frequency standards can be broken down into short-term ( $10^1$  seconds -  $10^2$  seconds), medium-term ( $10^2$  seconds -  $10^5$  seconds) and long-term ( $10^5$  seconds -  $10^7$  seconds) stability.

Clock	Resonant frequency
Quartz (S)	0.005 GHz
H-maser (M)	1.420 GHz
Rubidium (M)	6.835 GHz
Caesium (M-L)	9.192 GHz
Sapphire (S-M)	8 GHz - 17 GHz
Rutile (S-M)	8 GHz - 17 GHz
Mercury linear-ion trap (M-L)	40.51 GHz
Optical parametric oscillators (S-M)	190 THz
Cavity-stabilised Nd:YAG (S-M)	282 THz

TABLE 2: Resonant frequency of sample clocks, with stability regions indicated (S = short-term; M = medium-term; L = long-term)

Ovenised quartz resonators have been used as the short-term "workhorse" local oscillator for GPS, with a stability of  $10^{-12}$  for averaging times between 1 second and 100 seconds. The resonance is determined by the dimensions of the material (through the mass and stiffness). Quartz has the advantage that it is relatively cheap, low-power and light-weight. But, generally, the resonant frequency does not exceed 10 MHz in a commercial device, which means that medium- and long-term stability frequency standards with resonant frequencies in the microwave need to be mixed down in frequency, at the expense of bandwidth.

With the availability of low-power, light-weight space-qualified cryocoolers (J. C. GALLOP, 1996), cryogenic sapphire dielectric resonators operating at liquid nitrogen temperatures offer a realistic alternative as a short-term stability local oscillator, operating nominally in the X-band. Sapphire dielectric resonators (D. G. BLAIR, 1994) have many advantages over atomic frequency standards because the theoretical limit to the signal-to-noise is limited by radiation pressure from the microwave photons, whilst the signal-to-noise for an atomic frequency standard experiences a trade-off between atomic flux and pressure broadening effects. Also, the lifetime of cryogenic dielectric resonators is limited by the lifetime of the cryocooler (in excess of eight years), whereas atomic clocks are limited by the lifetime of the molecular supply (usually five years). Cryogenic sapphire dielectric resonators are also more tolerant to external magnetic field fluctuations, with the size of the cryogenic resonator system being comparable to commercial rubidium atomic frequency standards. The dimensions of the dielectric resonator can be reduced further by a factor of four by choosing rutile (N. KLEIN, 1995) instead of sapphire, with negligible loss in stability, allowing ultra-miniature cryocoolers to be used, with minimal power requirements. In summary, cryogenic dielectric resonators have distinct advantages over atomic frequency standards, and are attractive for the space-segment.

Rubidium atomic clock technology has progressed to the point where  $10^{-13}$  stability may be realised between  $10^2$  seconds and  $10^4$  seconds, through miniaturisation of the bulb, and by using stabilised laser to pump the rubidium vapour (G. MILETI, 1995), making rubidium a relatively cheap and compact medium-term stability clock for the user. But, without space-qualified single-mode lasers for pumping, the order-of-magnitude improvement in stability cannot be reaped for the space-segment applications.

In 1967, the second became defined as the number of cycles in a hyperfine transition of the caesium atom. Although the caesium atomic clock is larger than its rubidium counterpart, the attainable stability of  $10^{-13}$  is an important improvement over rubidium in the long-term. With six-beam laser-stabilised pumping and generation of a molasses (K. GIBBLE, 1992), a caesium fountain has been realised with stability of  $3 \times 10^{-15}$  in the long-term, with the potential for approaching  $10^{-16}$  in a micro-gravity environment (Philippe LAURENT, 1996). To realise the potential of this caesium fountain would require a time-transfer capability of 30 picoseconds with GNSS2, which has yet to be proposed. But, to reap the full benefits of the caesium fountain, space-qualified single-mode lasers will need to be developed.

The active cavity room-temperature hydrogen maser has been able to achieve  $2 \times 10^{-15}$  stability over  $10^3$  seconds -  $10^4$  seconds, and the caesium fountain standard may prove to be more stable than the hydrogen maser in the medium-term. But, the hydrogen maser has distinct advantages over caesium and rubidium because it relies on a discharge to generate the atomic beam, precluding the need for space-based lasers. The potential to develop a hydrogen maser-in-space, with a cryogenic dielectric resonator as the local oscillator, has potential for the next-generation of clocks in the space-segment. Attempts to design a cryogenic hydrogen maser system with better medium-term stability than the room-temperature version appear to be in difficulty, with problems in maintaining a uniform bulb coating over the long-term.

Progress on the mercury linear-ion trap system (LITS) has been steady and promising, with stability comparable to the caesium fountain already demonstrated (L. MALEKI, 1995). The United States are intending to launch a mercury LITS using an ovenised quartz crystal as the local oscillator.

Close attention also needs to be paid to the development of new optical frequency standards. Nd:YAG lasers stabilised to cryogenic Fabry-Pérot cavities have the potential of better than  $10^{-16}$  stability for averaging times up to  $10^4$  seconds (S. SCHILLER, 1995), and have the potential to be the next-generation of frequency standards, if the gap between optical and microwave frequencies can be bridged (M. NOTCUTT, 1996). Optical parametric oscillators may also offer a means of bridging that gap (N. C. WONG, 1995).

FIGURE 1 shows that existing two-way time-transfer requires better stability than that attainable with rubidium. But, rubidium clocks do have a role to play for the user, where cheap, light-weight operation is required. Quartz may also have a role to play in syntonising dielectric resonators, by mixing a fraction of their frequency from a cryogenic dielectric resonator.

Research effort is actively geared towards the production of space-borne miniature clocks: cryogenic dielectric resonators; atomic rubidium; atomic caesium fountain; room-temperature active hydrogen maser; and, the mercury linear-ion trap.

It should be mentioned that a space-segment GNSS2 will place emphasis on short- and medium-term stability clocks for **navigation** systems (for the space-segment and pseudolites), rather than on medium- and long-term stability clocks for **positioning** systems (using the 12-hour orbit global satellite systems). Non-reliance on semiconductor laser devices in the harsh environment of space favours a space-segment dependent upon the cryogenic dielectric resonator and active hydrogen maser technology. Maintenance of the time scale using long-term stability clocks is likely to be the responsibility of national timing centres throughout the world, partly because of the time to launch and deploy clocks in space, and partly because national timing centres usually research and develop the best frequency standards before industry realises their potential. Predictions are that the caesium fountain will be the future long-term "workhorse" for ground-segment global time-keeping.

Continent	Clocks for UTC evaluation
United States	69
Western Europe (Austria, Belgium, France, Germany, Italy The Netherlands, Spain, Switzerland, Turkey United Kingdom)	75
Australasia (Australia, China, Hong Kong, Japan Korea, New Zealand)	42
Eastern Europe (Czech Republic, Hungary, Poland, Russia)	16
Continental America (excluding USA) (Argentina, Canada)	9
Continental Africa (Israel)	1

TABLE 3: Atomic clocks considered in the calculation of International Atomic Time (TAI) for the 60-day interval commencing 23rd April 1995

However, a study of the main contributors to International Atomic Time (TAI) for the week commencing 23rd April 1995, a typical date (see TABLE 3), indicates

that the majority of clocks are located in the northern hemisphere in the United States and Western Europe, with a dispersion over Australasia. But, for any truly global time- and frequency-dissemination service, this imbalance would have to be addressed.

## 5) TIME- AND FREQUENCY-TRANSFER OPTIONS IN THE GNSS2 SYSTEM

### 5.1) Introduction

Time- and frequency-transfer will play a key role in the next-generation GNSS, the most significant internal application being the synchronisation of ground-segment and satellite clocks. Inter-satellite links would allow direct synchronisation of satellite clocks. GNSS2 should be designed to satisfy next-generation user requirements for both time and frequency dissemination, and enable time-transfers between primary timing laboratories in real-time to become common-place. There will need to be significant advantages in using the GNSS2 system for these applications when compared with the use of present-day GPS, if cost recovery is to be achieved.

Magnitude of Error	One-Way	Two-Way	Ranging
Satellite Position	> 6 ns*	0	0
Earth-Station Position	300 ps	0	300 ps*
Instrumentation Delay Changes	200 ps	200 ps*	200 ps*
Ionosphere	0	0	0
Troposphere	200 ps	0	400 ps*
Satellite Motion	0	0	0
Relativistic Effects	0	0	0
Absolute Delays	500 ps	500 ps*	500 ps*
Measurement Noise	20 ps	20 ps	20 ps
Total Errors	> 6 ns	≈ 600 ps	≈ 800 ps

\* indicates the major error source(s) requiring future research to mitigate their effects

0 indicates that the magnitude of the error source is (through modelling of the static component) negligible compared to other factors

TABLE 4: Typical time-transfer and ranging error budgets

### 5.2) Synchronisation of Ground Segment and Satellite Clocks

The synchronisation of ground-segment and satellite clocks within the GNSS2 system offers a major challenge to the time and frequency community. Currently, navigation signals from the present-day GPS are used to perform this synchronisation. But, the accuracy of this one-way time-transfer method is principally limited by the instantaneous determination of the satellite ephemeris (see TABLE 4). To improve the clock synchronisation, either the systematic errors limiting the one-way method must be

reduced, or a two-way satellite time- and frequency-transfer (TWSTFT) method needs to be implemented. The methods selected for clock synchronisation will be linked to the method chosen for satellite ephemeris determination, through the selection of common instrumentation. We propose that TWSTFT methods should be combined with microwave ranging of the satellite for real-time satellite ephemeris determination.

The TWSTFT method offers several clear advantages for the synchronisation of earth station and satellite clocks: the precise determination of the earth station coordinates and the satellite ephemeris are not required; the effects on the time-transfer of ionospheric delays are reduced to only a small fraction of the effect resulting from a one-way link; tropospheric delays cancel. The use of satellite simulators and calibration loops may enable the instrumentation delays to be calibrated during operation. Low measurement noise may be achieved with both the two-way and one-way methods, through the use of high frequency transmissions and directional receiving antennae, which favours J-band transmissions. Multipath interference effects could also be ameliorated. A minor disadvantage of the TWSTFT method is that an additional payload must be flown on the satellite, although there are examples today where experiments have been performed using on-board instrumentation (S. BEDRICH, 1995). Clock synchronisation significantly better than a nanosecond is ultimately achievable using the TWSTFT method. But, substantial work will be required to achieve this accuracy, particularly in the calibration of the instrumentation delays.

If one-way time-transfer were to be used for satellite clock synchronisation, then a substantial improvement must be achieved in satellite ephemeris determination (David W. ALLAN, 1996). One possible method is through the use of inter-satellite links. Higher transmission frequencies would significantly reduce ionospheric delay errors, and modelling could reduce the tropospheric delay error (D. KIRCHNER, 1994). Overall, clock synchronisation with an accuracy of one nanosecond may be possible. But, it should be stressed that the accuracy attainable using the two-way method will ultimately be higher.

### 5.3) Satellite Ephemeris Determination

The determination of the satellite ephemeris results in the largest contribution to the navigation error in the existing GPS system. GPS determines the satellite ephemeris by making pseudo-range measurements using the navigation signal. This method requires both small propagation delay errors and accurate clock synchronisation at the receiving station. The disadvantage of this technique is that the geometrical arrangement used to determine the satellite's position is particularly poor (FIGURE 2). This is due to the relatively short baselines in comparison with the distance to the satellite. Microwave ranging (FIGURE 3) dramatically improves the accuracy with which the range component of the satellite's ephemeris may be determined, and recent

results suggest that the range component of the satellite's ephemeris may be determined with an error of only a few centimetres. When taking geometrical considerations into account, this result suggests that the satellite ephemeris may be determined to an uncertainty well below a metre.

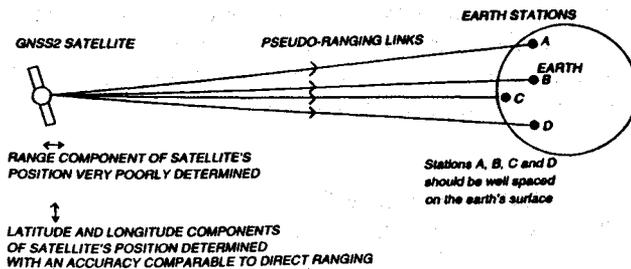


FIGURE 2: Pseudo-ranging of GNSS2 satellites

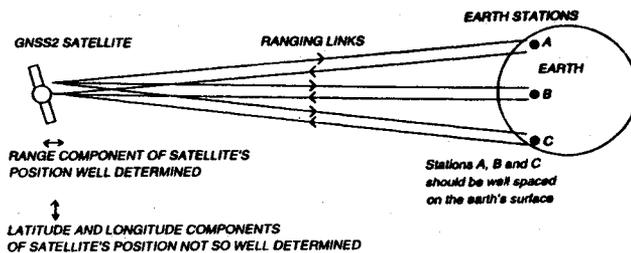


FIGURE 3: Microwave ranging of GNSS2 satellites

#### 5.4) Inter-Satellite Links

The use of inter-satellite links has several advantages that will be of value to GNSS2. The absence of atmospheric absorption enables a wide variety of transmission frequencies to be considered for these links. Inter-satellite links can improve the geometrical arrangement used to determine the satellite ephemeris, allowing high integrity to be realised, and making direct synchronisation between satellite clocks possible. The absence of ionospheric and tropospheric delays is particularly advantageous, and both one-way and two-way links are possible. Of particular interest may be links between geostationary satellites, or inclined geostationary satellites at the extremes of their orbit. With these satellites having a relatively fixed position, it may be possible to transmit and receive signals using small directional antennae. Low noise two-way links may be established, both for clock synchronisation and for improving the geometry of the ranging links used for ephemeris determination (FIGURE 4). Another interesting proposal is to transmit one-way links from geostationary satellites to satellites in 12-hour orbit (FIGURE 5). If the ephemeris of the geostationary satellites are well determined, then pseudo-ranging techniques should also enable the satellite ephemeris for the constellation in 12-hour orbit to be determined to high accuracy.

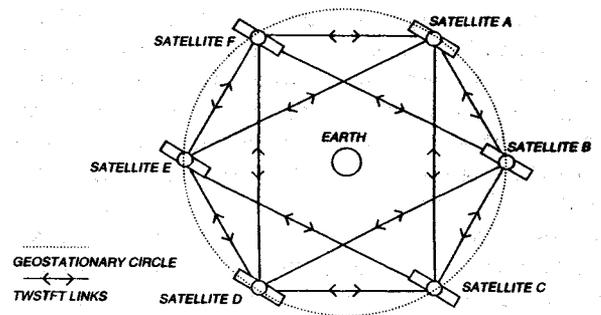


FIGURE 4: Two-way inter-satellite links between geostationary satellites

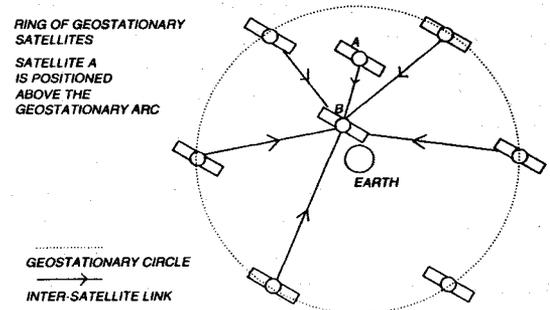


FIGURE 5: One-way links between geostationary satellites and satellites in 12-hour orbit

#### 5.5) GNSS2 Navigation Signals

In the existing GPS system the largest sources of error in the navigation signals are due to inaccurate satellite ephemeris determination and clock synchronisation. From the above discussions, it is clear that GNSS2 performance could be improved significantly over present-day GPS. Techniques to reduce the remaining sources of error in the GNSS2 navigation signal, such as ionospheric delay and tropospheric delay and multipath effects, have already been addressed through the use of multi-channel, multi-frequency receivers with directional antennae, and the use of higher multi-frequency broadcasting. Urban areas will pose particular attention to detail. Possible future studies have also been highlighted. The measurement noise and delay stability of the receiver may limit the performance of the navigation system, although carrier-phase detection with code rates up to 100 MHz may be used to limit the measurement noise. There is also work in progress to improve the delay stability of the receiver instrumentation, with special attention on alleviating the temperature-dependence effects on instrumentation, perhaps through miniaturisation of receivers and temperature-stabilisation.

To obtain the maximum performance from the GNSS2 navigation system, the choice of satellite constellation must be optimised. To determine the position of a GNSS2 receiver, four or more satellites must be observed

simultaneously. The geometrical orientation of the satellites, the dilution of precision (DOP), is a major factor in the user's navigation accuracy. The options are numerous for GNSS2: both in the choice of the number of satellites; and in the possible orbit types. Twelve-hour intermediate circular (ICO), geostationary (GEO), inclined geostationary (IGSO), medium-elliptical (MEO), highly-elliptical (HEO) (and even low-earth (LEO)) orbits are all under consideration. With an optimised choice of satellite constellation, a navigation accuracy below 1 metre ( $1\sigma$ ) should be realised, before the implementation of LAD techniques.

#### 5.6) Time- and Frequency-Transfer, and Dissemination, Services

GNSS2 offers several advantages over the existing GPS for both time-transfer and time-dissemination. The inclusion of geostationary satellites in the GNSS2 constellation enables the continuous reception of signals from a single satellite. Operation at high transmission frequencies would enable reception using small directional antennae resulting in high carrier-to-noise (C/No) values and low measurement noise. GNSS2 will not require Selective Availability degradation, and lower systematic errors will give GNSS2 a distinct advantage for both one-way and common-view time-transfer, over the existing GPS system. An option worth considering is the operation of a regional, near-real-time, common-view service. With many primary timing laboratories already possessing up-linking facilities, the possibility exists to transmit GNSS2 common-view data via geostationary GNSS2 satellites, so making real-time common-view time-transfer available to GNSS2 users.

#### 5.7) Time and Frequency in the Ground-Segment

To obtain the best synchronisation between a GNSS2 time scale and UTC, it would be sensible to reference GNSS2 time (UTC(GNSS2)) to atomic clocks at primary timing laboratories. As many primary timing laboratories already possess up-linking capabilities, there would be the possibility of up-linking signals from several different laboratories worldwide in real-time. A key issue will be the re-synchronisation time for correcting drifts in the satellite clocks, which will depend on accuracy and integrity requirements of the user, and the stability of the on-board clocks in the space-segment. This could typically range from 60 seconds to 3,600 seconds. (Clock drifts as low as 3 picoseconds are achievable, with the clock drift for the combined cryogenic dielectric resonator and active hydrogen maser  $< 0.1$  ps per second/ $\tau^{-1/2}$  between  $\tau = 1$  second and 1,000 seconds).

#### CONCLUSIONS

This paper has described the first Phase of an on-going study of work investigating the feasibility of a wholly-civil European second-generation Global Navigation Satellite System. The issues of time and frequency underpin the whole system, and this has been highlighted through the options available for frequency-band allocation, clock

development, and time-transfer techniques for global navigation and global time- and frequency-dissemination. The NPL propose that further studies on issues such as local real-time ionospheric and tropospheric delay evaluation, two-way time-transfer to geostationary satellites for real-time satellite ephemeris determination, and multipath effects and methods to ameliorate them through multi-frequency receivers and directional antennae, would be beneficial towards developing the next-generation of global navigation satellite systems.

#### ACKNOWLEDGEMENTS

This work was co-funded under ESA Contract 11420/95/NL/US for GNSS2 Mission Analysis and the relevant National Measurement System Programmes for Time and Frequency, and Basic Metrology. The authors thank Dr David W. ROBINSON and Dr Stuart POLLITT for their continuing support of this work, and to the ESA GNSS2 Mission Analysis Team for providing NPL with the opportunity to participate in this exciting project.

#### REFERENCES

- D. W. ALLAN, R. GIFFARD, L. S. CUTLER and J. A. KUSTERS, "Improved time and frequency transfer opportunities via GPS", Proc. 9th EFTF, pp 71-74, Besançon, France, 8-10 March 1995
- David W. ALLAN, Neil ASHBY, "Navigation and timing accuracy at the 30 centimeter and sub-nanosecond level", GNSS2 Workshop, National Physical Laboratory, 4 March 1996 (paper to be published)
- G. APPLEBY, P. GIBBS, "Precise optical observations of geostationary satellites from Hertsmenceux", Proceedings of the Workshop: Accurate orbit determination and observations of high-earth satellites for geodynamics, 17th-19th October 1994, Walferdange, Luxembourg
- Jean-Pierre BARBOUX, "Practical real time kinematic applications of GPS", Proc. 3rd DSNS, Paper 36, 18-22 April 1994
- Yaakov BAR-SHALOM, Anil KUMAR, William D. BLAIR, Gordon W. GROVES, "Tracking low elevation targets in the presence of multipath propagation", IEEE Trans. Aero. Electron. Systems, pp 973-979, July 1994
- S. BEDRICH and A. HÄRTING, "Monitoring of the PRARE clock on-board Meteor-3: operations and performance analysis", Proc. 9th EFTF, pp 34-39, Besançon, France, 8-10 March 1995

D. G. BLAIR, M. E. COSTA, A. N. LUITEN, A. G. MANN and M. E. TOBAR, "Cryogenic sapphire dielectric resonator microwave oscillator with improved stability", Proc. 8th EFTF, 1994

Walter BLANCHARD, "The Implementation of GNSS", IEE, Savoy Place, London, 16 November 1995

C. FROHLICH, "Shaken to the core", New Scientist, pp 42-46, 9 December 1995

Chris de la FUENTE, K. PRESTON, "The use of GPS on British Rail", Proc. ION GPS-93, pp 1631-1636, 1993

J. C. GALLOP and C. D. LANGHAM, "Dielectric loaded HTS resonators as frequency standards and low phase noise oscillators", Proc. 10th EFTF, Session 9, Brighton, UK, 5-7 March 1996

K. GIBBLE and S. CHU, "Future slow-atom frequency standards", Metrologia, pp 201-212, 1992

Thomas A. HERRING, "The Global Positioning System", Sci. Amer., pp 32-38, February 1996

C. C. HODGE and J. A. DAVIS, Task 2 and Task 3 Reports for GNSS2 Mission Analysis, "Stable clocks and time-transfer", April 1996

International Telecommunication Union, "Radio Regulations", Volumes 1, 2, 3, ISBN 92-61-04141-8, 1994

D. KIRCHNER, C. LENTZ, "Tropospheric corrections to GPS measurements using locally measured meteorological parameters compared with general tropospheric corrections", Proc. 25th PTTI, pp 231-248, November 1994

Philip J. KLASS, "GPS to be tested for Cat.-3 landings", Aviation Week & Space Technology, pp 62-63, 14th February 1995

N. KLEIN, C. ZUCCARO, U. DÄHNE, H. SCHULZ, N. TELLMANN, R. KUTZNER, A. G. ZAITSEV, R. WÖRDENWEBER, "Dielectric properties of rutile and its use in high temperature superconductors", J. Appl. Phys., pp 6683-6866, November 1995

Philippe LAURENT, E. SIMON, G. SANTARELLI, A. CLAIRON, P. LEMONDE, C. SALOMON, N. DIMARCQ, P. PETIT, C. AUDOIN, F. Jamin CHANGEART, F. GONZALES, "A space frequency standard using cold atoms: the PHARAO project", GNSS2 Workshop, National Physical Laboratory, 4 March 1996

Lute MALEKI, "Frequency standards from government research laboratories over the next 25 years", Proc. 25th PTTI, pp 549-560, November 1993

Lute MALEKI, "Trapped ion standards: current status and future prospects", Proc. 9th EFTF, pp 1-8, Besançon, France, March 1995

Gaetano MILETI and Pierre THOMANN, "Study of the S/N performance of passive atomic clocks using a laser pumped vapour", Proc. 9th EFTF, pp 271-276, Besançon, France, 8-10 March 1995

M. NOTTCUTT, C. T. TAYLOR, A. G. MANN, R. GUMMER and D. G. BLAIR, "Cryogenic system for a sapphire Fabry-Pérot optical frequency standard", Cryogenics, pp 13-16, 1996

S. SCHILLER, S. SEEL, R. STORZ and J. MLYNEK, "Experimental studies of cryogenic optical resonators", Proc. SPIE, 2378, pp 138-146, 1995

Richard D. J. VAN NEE, "Multipath effects on GPS code phase measurements", Navig. J. Inst. Navig., pp 177-190, Summer 1992

N. C. WONG, "Optical-to-microwave frequency chain utilising a two-laser-based optical parametric oscillator network", Appl. Phys. B, pp 143-149, 1995

"The Role of GPS in European Civil Aviation", IEE, Savoy Place, London, 2nd March 1995

## A GPS PARALLEL COMMON-VIEW EXPERIMENT

Jan Čermák

Institute of Radio Engineering and Electronics, Academy of Sciences of the Czech Republic

### ABSTRACT

In May-June 1995 an experiment was carried out at the Institute of Radio Engineering and Electronics (IREE), Prague, in collaboration with the Physikalisch-Technische Bundesanstalt (PTB), Braunschweig, and the Istituto Elettrotecnico Nazionale (IEN), Turin to assess the performance of the GPS common-view time transfer to IREE.

At each laboratory two GPS receivers were in parallel operation. While the first receivers were scheduled according to the recommendation of Bureau International des Poids et Mesures (BIPM), the second receivers followed a special schedule with equally-spaced tracks and with satellites at high elevations. As a comparative measure of the white-noise PM present in the common-views, the time deviation  $\sigma_x(\tau_0)$  was used. The best noise performance with  $\sigma_x(\tau_0) = 1.5$  ns for the sampling interval of 1 hour was shown in the case of the UTC(TP)-UTC(IEN) common-view difference measured by the second-receiver pair. A favourable impact of higher satellite elevations has been observed. Daily variations with a variable amplitude up to several nanoseconds were found in all common-view pairs. As a first step in seeking the main source of these variations, antenna temperature sensitivity of the two IREE's receivers has been measured. All filtering throughout the analysis was realised by using Vondrák smoothing.

### INTRODUCTION

The time laboratory of IREE is responsible for the Czech National Time and Frequency Group Standard whose performance has been described by Buzek et al (1) and Buzek (2). IREE is also generating the UTC(TP) time scale (TP stands for *Tempus Pragense*).

As with most of the time laboratories nowadays, IREE is making use of the common-view GPS time transfer to ensure the access to BIPM and other time and frequency centres. Since the GPS reception was implemented in IREE, routine comparisons of UTC(TP) are being made against UTC(PTB) and UTC(IEN) which are very helpful in evaluating the metrological parameters of the Czech Standard.

Though the behaviour of generic GPS in common-view is known, in practice the performance of a concrete link between two laboratories may differ depending on local conditions. Thus two laboratories may establish a special common-view link and may attempt to optimise it (e.g. through hardware and software calibration, choice of tracking schedule and satellites, corrections for temperature sensitivity etc.).

Having this as motivation, we asked our colleagues of PTB and IEN to take part in an experiment which would provide redundant data to better assess the current common-view links to IREE. The idea was to employ the spare GPS receivers of IREE, PTB and IEN with a special schedule different from that of BIPM. Thus the second receivers of each laboratory were scheduled for equally-spaced tracks so that the common stability measures could be applied. In addition, the schedule was chosen so as to allow the satellites to be tracked at only high elevations.

Some interesting results obtained in the post-process analysis of all common-view pairs are described in this paper.

### BACKGROUND

#### GPS in Parallel Common-View

The time transfer from place A to place B via the GPS in common-view is based on simultaneous observations of the same satellite. Thus the strict common-view technique requires at least one pair of receivers that would have an identical tracking schedule (i.e. the same start of track, length of track and satellite to be tracked).

If there is a spare receiver at one site it can follow either the same schedule as a backup receiver, which is usually the case, or it can be scheduled in a different manner making a pair with another equally-scheduled receiver. In this way the two observers can establish two (or more) GPS common-view links to provide redundant time transfer between each other. For these multi-scheduled common-views we use here the term "parallel" common-views.

Let us first introduce several basic definitions concerning GPS in common-view. The GPS time receivers make it possible to measure the difference

between local time and GPS time. Hence the measured time difference is a function of time. If there are more receivers at site A, we can write for the time difference obtained from receiver k at time t

$${}^k\Delta_A(t) = x_A - {}^k g_A(t) \tag{1}$$

where  $x_A$  is the local time reference (used for all receivers at the place) and  ${}^k g_A$  is the GPS time as interpreted by receiver k. We can write a simple model relation

$${}^k g_A(t) = g + S_A(t) + {}^k R_A(t) \tag{2}$$

where g is the system GPS time as generated by the GPS master-clock and the time varying functions represent the fluctuations in the transfer of GPS time to site A. By  $S_A(t)$  we denote the fluctuations arising outside the receiver (in the satellite and on the path to A) and  ${}^k R_A(t)$  represents the fluctuations inside the receiver k. We assume that time  ${}^k g_A$  has been corrected for the systematic errors that is we are interested in the relative time transfer (or frequency transfer) only.

In practice, the difference (1) measured by the time receiver at time t is not a kind of one-shot value but a result of certain statistical data treatment within the receiver related to time t. The time receivers usually provide difference (1) along with the slope (frequency deviation) and the residual standard deviation resulting from the ultimate linear fit made over the whole track.

When substituting  $x_A$  and  ${}^k g_A$  in (1) as epochs then for  $x_A < {}^k g_A$  (time  $x_A$  is ahead of time  ${}^k g_A$ ), the difference  ${}^k \Delta_A$  becomes negative. To avoid confusion we should point out that if the difference  ${}^k \Delta_A$  is understood as a result of a counter measurement (pulses  $x_A$  and  ${}^k g_A$  triggering Start and Stop inputs, respectively), it takes by convention a positive sign when  $x_A$  arrives before  ${}^k g_A$ .

Since relations (1) and (2) hold also for the equally-scheduled k-th receiver at site B, we can make a common-view difference

$${}^k \Delta_{A,B}(t) = {}^k \Delta_A(t) - {}^k \Delta_B(t). \tag{3}$$

Substituting from (1) and (2) in (3) gives

$${}^k \Delta_{A,B}(t) = x_A - x_B + S_B(t) - S_A(t) + R_B(t) - {}^k R_A(t). \tag{4}$$

Because  $x_A \approx x_B$  and the time functions in the above equations are slowly varying, time t may be represented by either  $x_A$  or  $x_B$ .

Equation (4) indicates that for an ideal time transfer equalities  $S_A(t) = S_B(t)$  and  $R_A(t) = {}^k R_B(t)$  should hold at any time. Obviously, as B approaches A (by shortening the base-line) then  $S_B - S_A$  approaches zero (paths A and B become

identical). The residual difference  ${}^k R_A - {}^k R_B$  depends on local receiving conditions and therefore is of main concern if any improvements in the common-view time transfer should be made.

### Stability Measure

In order to assess the phase (time) stability of the GPS common-views one needs an adequate stability measure. Out of the variances commonly used in the time domain, the best suited measure in our case is the time variance,  $\sigma_x^2(\tau)$  or TVAR, which is defined, e.g. Allan (3), as

$$\sigma_x^2(\tau) = \frac{1}{6} \langle (\Delta_2 \bar{x})^2 \rangle, \tag{5}$$

where  $\Delta_2$  is the operator of the second difference and  $\bar{x}$  represents the average over  $\tau = n\tau_0$  (n is integer and  $\tau_0$  is the basic sampling interval). The symbol  $\langle \rangle$  indicates time average over an infinite time.

An estimate of  $\sigma_x^2(\tau)$  for a finite data length with N samples is

$$\sigma_x^2(\tau) = \frac{1}{6n^2M} \sum_{k=1}^M \left[ \sum_{i=k}^{n+k-1} (x_{i+2n} - 2x_{i+n} + x_i) \right]^2 \tag{6}$$

where  $M = N - 3n + 1$  and i, k are integers.

The plot of  $\sigma_x(\tau)$  makes it possible to clearly distinguish between the specific types of noise assumed to be present in the GPS common-views. The  $\sigma_x(\tau)$  plot decays as  $\tau^{-1/2}$  for white-noise PM (-1/2 slope on log-log scale) and increases as  $\tau^{1/2}$  for white noise FM (+1/2 slope on the log-log scale).

### Vondrák Smoothing

A digital filter well suited to filtering the GPS common-views in the time domain is Vondrák smoothing which allows unequally spaced arguments to be treated. The principle of smoothing is based on minimising the sum

$$\frac{1}{N-3} \sum_{i=1}^N (\xi_i - x_i)^2 + \frac{\tau_s^6}{t_B - t_A} \int_{t_A}^{t_B} \left[ \frac{d^3 \xi(t)}{dt^3} \right]^2 dt = \min \tag{7}$$

where

- $x_i$  measured value at  $t_i$ ,
- $\xi_i$  smoothed value at  $t_i$ ,
- $\xi(t)$  smoothed function
- N number of measurements,
- $t_A$  beginning of smoothing interval,
- $t_B$  end of smoothing interval,
- $\tau_s$  smoothing time constant.

The algorithm approximating relation (7) is described in detail by Vondrák (4) and (5). Huang and Zhou (6) have derived a response function in the frequency domain as

$$\frac{\xi(\omega)}{\chi(\omega)} = \frac{1}{1 + (\omega\tau_s)^6} \quad (8)$$

which shows a good selectivity of the algorithm when utilised as a filter. For the finite data length, equation (8) holds best in the middle of the smoothing interval.

## EXPERIMENT SPECIFICATIONS

The parallel common-views started 25 April 1995 (MJD 49832) and finished 19 June 1995 at PTB and 29 June at IREE and IEN.

The base-line between IREE and PTB is 368 km, TP and IEN 768 km, and finally PTB and IEN 835 km.

As local time references, UTC time scales were used that is UTC(TP) at IREE, UTC(PTB) at PTB and UTC(IEN) at IEN.

The standard track length of 13 minutes was used for both the BIPM and experimental schedules.

The BIPM schedule provided on average 25 daily common-view tracks of  $(TP-PTB)_1$ , 36 tracks of  $(TP-IEN)_1$  and 23 tracks of  $(PTB-IEN)_1$ . Average satellite elevations were about  $45^\circ$ . To make the symbols more legible when treating these concrete common-view pairs, we will use an alternative notation for differences (3), e.g.  $(TP-PTB)_k$  instead of  $^k\Delta_{TP, PTB}$ .

Since the satellites for the experimental schedule were chosen to be at the highest elevations possible ( $70^\circ$  on average), a compromise choice of the minimum spacing between tracks was approximately one hour. Thus the experimental schedule provided 24 hourly-spaced tracks a day with 14 different satellites to be tracked (some of them had to be tracked twice). The interval between the tracks immediately before and after midnight was shorter by 4 minutes due to receiver approximation of sidereal time.

The receivers employed in the experiment were as follows (numbers in parentheses are serial numbers):

Allen Osborne TTR-6 (260) and (412) at IREE, TTR-5 (156) and Rockwell Collins at PTB, and lastly NBS (31) and TTR-5 (196) at IEN. Each first receiver of the laboratory was that of BIPM schedule.

The measurement protocols produced by the receivers gave 0.1 ns resolution in difference (1) for the IREE and PTB receivers, and 1 ns for the IEN receivers. Slope resolution was 1 ps/s for all receivers.

## EVALUATION

Before post-processing the data, inconsistent outliers had been removed and the IEN record was corrected for a step of +600 ns made by IEN on MJD 49840.4. All common-view calculations were related to the middle of the tracking interval. Figure 1 shows the records of relative common-view differences against UTC(TP) obtained from the rough data (systematic drift removed).

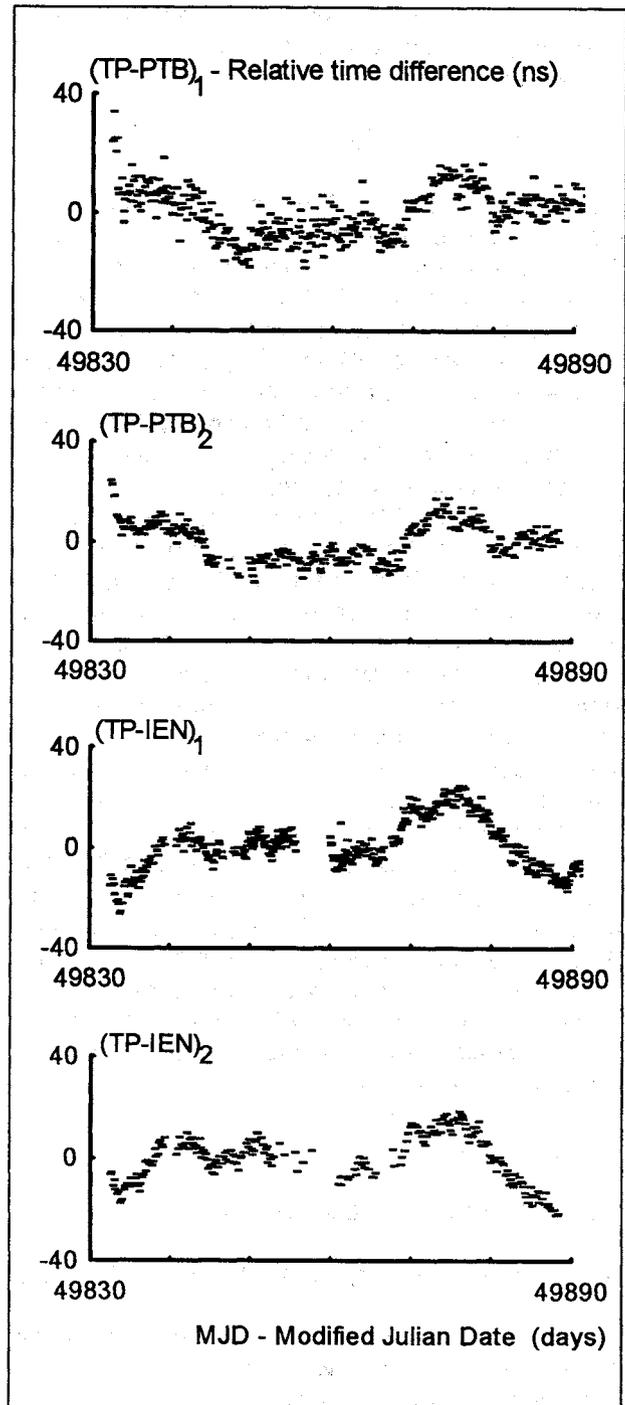


Figure 1: Common-views against UTC(TP) calculated from rough data

Though the measurements were interrupted several times due to problems with the GPS reception in one or another laboratory, the continuous segments provided data sufficiently robust for analysis.

**Stability Results**

According to (5), the time deviation  $\sigma_x(\tau)$  is defined for equally-spaced samples. Thus it can be strictly applied only to the common-view differences obtained from the experimental schedule ( $k=2$ ). Figure 2 shows the plots of  $\sigma_x(\tau)$  as calculated for common-view differences (TP-PTB)<sub>2</sub>, (TP-IEN)<sub>2</sub> and (PTB-IEN)<sub>2</sub>.

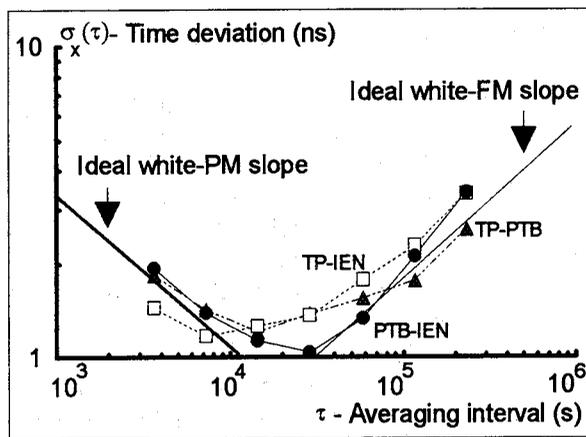


Figure 2:  $\sigma_x(\tau)$  for common-views obtained from the experimental schedule

As expected, two types of noise can be identified in the plots in Figure 2. For short averaging intervals, the prevailing noise with  $-1/2$  slope is the white noise PM due to GPS. For long averaging intervals, the  $+1/2$  slope indicates the white noise FM due to clocks involved in the common-views. For a given sampling interval  $\tau_0$ , function  $\sigma_x(\tau)$  takes its minimum for an interval which we refer to as  $\tau_c$ . In the above common-views,  $\tau_c$  falls in the range of  $1 \times 10^4$  to  $2 \times 10^4$  s for  $\tau_0 = 1$  hour. In order to give  $\sigma_x(\tau)$  a comparative meaning for white noise PM, we assume that the effective noise bandwidth is identical in all receivers.

It is apparent that  $\sigma_x(\tau_0)$  can be taken as a single-value measure to compare the phase noise of different common-views. However, one must be concerned with the condition  $\tau_0 \ll \tau_c$  which should be fulfilled to avoid contamination with the FM noise of clocks. We should note that  $\sigma_x(\tau_0)$  is equal to the classical standard deviation of the pure white-noise PM.

The estimate of  $\sigma_x(\tau)$  can also be applied to non-equidistant data (now having the basic sampling interval variable,  $\tau_{oi}$ , in contrast to constant  $\tau_{oi}$ ). If

$\tau_{oi}$  complies with the condition  $\max(\tau_{oi}) \ll \tau_c$  then  $\sigma_x(\tau_{oi})$  is again equal to classical standard deviation of white noise PM.

As an alternative measure, the classical standard deviation can be used if calculated from the common-views high-pass filtered by means of Vondrák smoothing with an appropriate smoothing time constant  $\tau_s \approx \tau_c$ . In this way the white noise FM prevailing for  $\tau > \tau_c$  is attenuated and the standard deviation becomes independent of data length. The result of high-pass filtering (with  $\tau_s = 10^4$  s) applied to the common-views of the experimental schedule can be seen in Figure 3 (see the difference against the plots in Figure 2). Obviously, for sake of comparison, the value of  $\tau_s$  should be the same for all common-views.

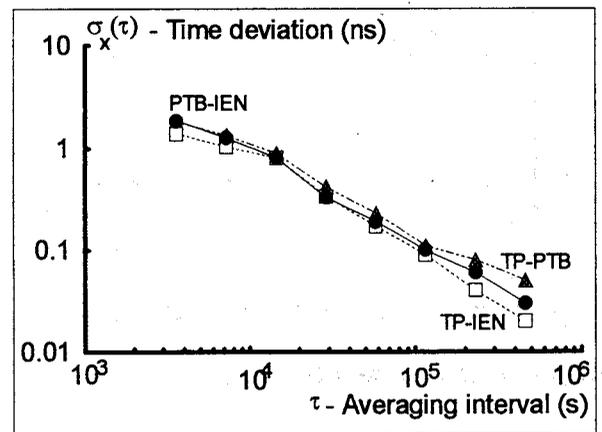


Figure 3:  $\sigma_x(\tau)$  for high-pass filtered common-views

The phase instability results for all common-views obtained during the experiment are summarised in Table 1. The quantities  $\sigma_x(\tau_0)$  and  $\sigma_x(\tau_{oi})$  are the above described time deviations and  $\sigma$  represents the classical standard deviation calculated from the high-pass filtered common views (with  $\tau_s = 10^4$  s).

**TABLE 1 - Common-view time deviations (ns)**

	Schedule			
	BIPM k=1		Experimental k=2	
	$\sigma_x(\tau_{oi})$	$\sigma$	$\sigma_x(\tau_0)$	$\sigma$
(TP-PTB) <sub>k</sub>	3.9	3.6	1.8	1.8
(TP-IEN) <sub>k</sub>	1.5	1.6	1.5	1.4
(PTB-IEN) <sub>k</sub>	4.1	3.6	1.9	1.7

Comparing the deviations in Table 1 for both schedules one can see noticeable differences between the common-view pairs in which PTB's first receiver was involved and the other pairs. In addition, close examination of the  $(\text{TP-PTB})_1$  record has revealed sudden changes in noise strength as if there were two modes of operation. This effect is illustrated in Figure 4 for high-pass filtered common-views. But even for the less noisy mode (see area marked by the arrow), the deviation  $\sigma_x(\tau_{oi})$  yields around 2.5 ns which is still considerably more than in the case of other common-view pairs.

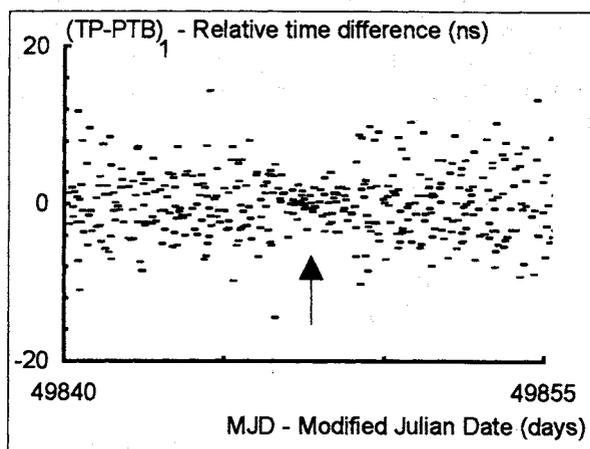


Figure 4: Changes in  $(\text{TP-PTB})_1$  noise strength

### Influence of Satellite Elevation

It is generally assumed that high elevations have a favourable effect on the phase noise of GPS in common-view. In association with our experiment one can be interested to what extent this holds for averaging intervals of one hour or longer.

Comparing the time deviations in Table 1 for both the BIPM and experimental schedules (the latter had on average higher elevations) one can really see better noise figures for the experimental schedule. However, regarding to what has been said about the common-views with the first receiver of PTB, this comparison is of no value. So only the  $(\text{TP-IEN})_1$  and  $(\text{TP-IEN})_2$  common-views remain to be compared but there is no significant difference between them.

A difference of about 0.5 ns in favour of high elevations has been found by comparing two data sets selected out of the  $(\text{TP-IEN})_1$  record according to elevations (with the decision elevation of  $40^\circ$ ).

### Daily variations

Close inspection of all common-view records has revealed daily variations up to several nanoseconds occurring irregularly in all common-view pairs. Actually, this is not unexpected since there are several potential sources of daily variations in

GPS in common-view. As an example, an interval with diurnal variations observed in  $(\text{TP-IEN})_1$  is displayed in Figure 5.

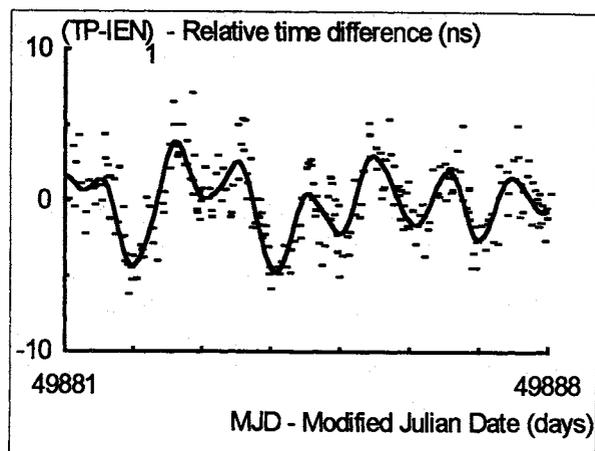


Figure 5: An example of diurnal variations

It is apparent that one should start with seeking a correlation with the environmental changes. However, this would require a good record of the environmental data from all places involved in the measurement. Unfortunately, during the experiment we had no such data available.

A possible approach to attenuate the daily variations would be to simply pass the common-views through a band-stop filter with 1/day centre frequency. However, if the potential source (as temperature sensitivity) produced seasonal components, these would remain interfering.

### Antenna Temperature Sensitivity

The main source of the daily variations may be the temperature sensitivity of the receiving system. The outdoor elements, which comprise the cables and the antenna with down-converter (in NBS-type receivers), are particularly critical as they are exposed to large temperature changes with daily periodicity. The suspicion is supported by the fact that daily variations have also been observed in the common-views between the two IREE receivers.

The problem arises in how to measure this effect without a GPS-signal simulator.

The temperature sensitivity makes the GPS time a function of the temperature change,  $v_k$ . Assuming that the dependence is linear within the temperature range, equation (2) can be modified in the form (omitting index A)

$${}^k g(t, v_k) = g + S(t) + {}^k R(t) + a_k v_k(t). \quad (9)$$

where  $a_k$  is the temperature coefficient which we are looking for. For  $a_k > 0$ , the GPS time gets delayed as temperature increases.

By changing the temperature and measuring the changes induced in difference  $x - k_g$ , one could theoretically obtain  $a_k$ . In practice, however, this is impossible because of the large noise due to selective availability which entirely masks the temperature effect. Perhaps using the Block I satellites might be of some help but for a longer measurement needed to collect enough data to filter out the phase noise, the white noise FM would be of trouble anyway. Therefore we have to make use of the common-view technique again. Considering that  $A \equiv B$  we may write  $x_A = x_B$  and  $S_A = S_B$ . Substituting (9) for receivers 1 and 2 in (4) yields the changes in common-view differences (indexes indicating sites are omitted)

$$\Delta(t, v_1, v_2) = {}^2R(t) - {}^1R(t) + a_2 v_2(t) - a_1 v_1(t). \quad (10)$$

Based on the above relation, two IREE's TTR-6 antennas have been tested with both receivers operated with the same BIPM schedule. To induce observable changes in the common-view difference (10), large steps in temperature  $v_1(t)$  were periodically made while keeping temperature  $v_2(t)$  relatively constant. Then, alternatively, the steps were made in temperature  $v_2(t)$  while keeping  $v_1(t)$  nearly constant. During the measurement hourly samples of both  $v_1(t)$  and  $v_2(t)$  were taken. The evaluation was made by using the standard least-square estimation technique. The coefficients were found to be  $a_1 = 0.1 \text{ ns}^\circ\text{C}$  and  $a_2 = -0.11 \text{ s}^\circ\text{C}$  with 10% confidence.

## CONCLUSION

Phase noise performance of the GPS common-view links from PTB and IEN to IREE has been found good ( $< 2 \text{ ns}$ ) except for  $(\text{TP-PTB})_1$  obtained from the first receivers. Confronting  $(\text{TP-PTB})_1$  with other links leads to a conclusion that the problem is associated with the first receiver of PTB. The  $(\text{TP-IEN})_2$  common-views showed the best performance giving  $\sigma_x(\tau_0) = 1.5 \text{ ns}$ . Comparing this figure with  $1.2 \text{ ns}$  obtained from the common-views made directly between the two IREE receivers, it can be concluded that in practice the base-line up to  $800 \text{ km}$  does not affect the relative time transfer.

Influence of satellite elevation on the common-view phase noise has been observed. One could conclude that it is desirable to avoid very low elevations ( $< 30^\circ$ ) but it is not necessary to insist on very high elevations.

The daily variations encountered in all common-view pairs remain a problem which has yet to be investigated. The antenna temperature coefficients of the order of  $0.1 \text{ ns}^\circ\text{C}$  found for the IREE receivers may account for part of the effect but not entirely. It is clear that if the problem should be

solved the detailed environmental data from both common-view sites would be needed.

The fact that the temperature coefficients have been found with the opposite sign differs from what would be anticipated. Though the estimation confidence is acceptable, one still may have some doubts whether the linear model applies to large temperature steps needed to induce observable changes in phase.

The experiment described in this paper should be understood as the first step in our effort to assess the current performance of the GPS common-view time transfer to the Czech National Time and Frequency Standard. We hope that the better knowledge of this performance will allow its further improvement.

## ACKNOWLEDGEMENTS

The author would like to thank A. Bauch, P. Hetzel and T. Polewka of PTB as well as F. Cordara and V. Pettiti of IEN for helping to realise the experiment. He is also grateful to O. Buzek and B. Čemusová of IREE for their support and helpful discussions and to J. Vondrák of the Astronomical Institute, Czech Academy of Sciences, for the smoothing program.

## REFERENCES

1. Buzek O, Čermák J and Čemusová B, 1995, "Czech National Time and Frequency Group Standard", 9th EFTF, 400-403
2. Buzek O, 1996, "Temperature sensitivity of a fibre optic timing link", 10th EFTF
3. Allan D W, 1991, "Time and frequency metrology: Current status and future considerations", 5th EFTF, 1-9
4. Vondrák J, 1969, "A contribution to the problem of smoothing observational data", Bull Astron Inst Czech, 20, 349-355
5. Vondrák J, 1977, "Problem of smoothing observational data II", Bull Astron Inst Czech, 28, 84-89
6. Huang Kun-Yi, Zhou Xiong, 1981, "On essentiality of the Whittaker-Vondrak method as a filter, and estimation of standard deviation and correlation for digital filter", Acta Astron Sinica, 22, 120-128

## USE OF DIRECT BROADCASTING SATELLITE SIGNALS FOR TIME TRANSFER AND ORBITOGRAPHY

Ph. Guillemot\*, J. Folliard\*, M. Brunet\*, F. Meyer\*\*, M. Vincent\*\*, M. Granveaud\*\*\*, P. Grudler \*\*\*\*

\* CNES Toulouse, \*\* Observatoire de Besançon, \*\*\*LPTF Observatoire de Paris,  
\*\*\*\* Observatoire de la Cote d'Azur - France

### ABSTRACT

Several time transfer experiments have been performed in Europe, to compare clocks in the 10 nanoseconds range, using TV satellites such as ASTRA, BRITSAT - Davis et al (1) - or TDF 2 - Meyer et al (2, 3).

More recently, the Observatoire de Besançon (OB) and the CNES Toulouse (CT), with the support of the LPTF-Observatoire de Paris (OP) and the Observatoire de la Côte d'Azur (OC), have conducted a new 6 week experiment using the French telecommunication satellite TELECOM 2A, and four reception stations each equipped with a 80 cm diameter antenna, caesium clock and GPS time receiver. Each raw data was corrected of the motion of the satellite and others delays. Time transfer results were then compared to GPS results, the accuracy is limited by the delay uncertainty of each station equipment.

Using the same measurements (TV and GPS), the position of TELECOM 2A spacecraft was calculated and compared with classical technique (Ranging and Turn-Around).

This paper describes techniques, stations, uncertainty budget, time transfer results and presents a promising technique for geostationary orbit determination. As a conclusion, we propose the use of the Telecom 2 satellite for a near operational, automatique and independent to GPS time transfer system.

### INTRODUCTION

The principle of time transfers through direct TV satellites has been described in (Refs. 1, 2). Time differences between two clocks are obtained by timing the arrival of a given pulse at two different location and by taking into account the co-ordinates of both receiving antennas and satellite.

Since 1990, we have been using four stations observing the TDF2 satellite, receiving D2-MAC standard television signal and timing the arrival of the line 624

(Refs. 2, 3). Each station is equipped with a caesium clock and linked to the others by GPS, allowing the calibration of the results.

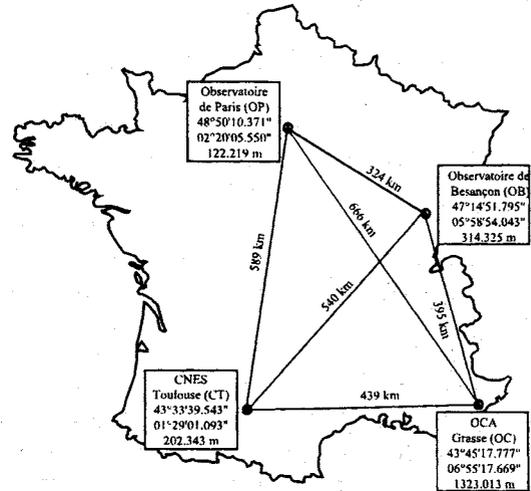


Figure 1 : The 4 stations in use

Considering two stations, the time diagram is presented on figure 2, where :

- Sat : Satellite  $t_{sat}$  : Emission date
- St<sub>i</sub> : Ground station i  $H_i$  : St<sub>i</sub> time reference
- $t_{Ri}$  : St<sub>i</sub> receiving date  $\Delta t_i$  : St<sub>i</sub> measurement
- $\tau_i$  : Satellite-St<sub>i</sub> propagation delay
- $RST_i$  : St<sub>i</sub> internal propagation delay (calibration)

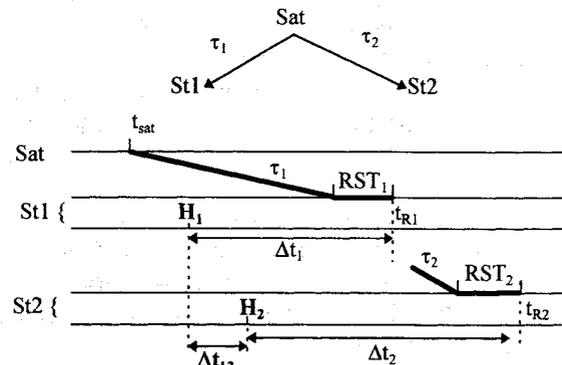


Figure 2 : Time diagram

In order to obtain the time difference  $\Delta t_{12} = H_2 - H_1$ , we have :

$$t_{Ri} = t_{sat} + \tau_i + RST_i = H_i + \Delta t_i$$

So the relationship between the time difference of the two concerned clocks, the measured time intervals and the propagation delays can be written as follows :

$$\Delta t_{12} = H_2 - H_1 = (\Delta t_1 - \Delta t_2) - (\tau_1 - \tau_2) - (RST_1 - RST_2) \quad (1)$$

In this relationship, the propagation delay term can be decomposed into :

$$\tau_i = \frac{D_{Sat-Sti}}{c} + \tau_{iono} + \tau_{tropo} + \tau_{Sagnac} \quad (2)$$

Ionospheric, Tropospheric and Sagnac effects can be easily computed. The distance between the satellite and the station can be estimated by knowing the position of the station antenna and the theoretical position of the satellite on the geostationary orbit. However, to reach an uncertainty of less than 50 nanoseconds, it has been proved that we must take into account the satellite periodic motion in latitude, longitude and radius (Ref. 1 and figure 3 shows time differences without taking into account the satellite motion). Two methods have been explored :

- explicitly determining the satellite position by using data from three stations, as shown in F. Meyer et al (1, 4), in that case, the time differences between the 3 stations must be obtained by another way,
- computing the satellite position with classical technique and models using station keeping measurements provided by the Satellite Control Centre (SCC).

Only this second way is discussed in this paper

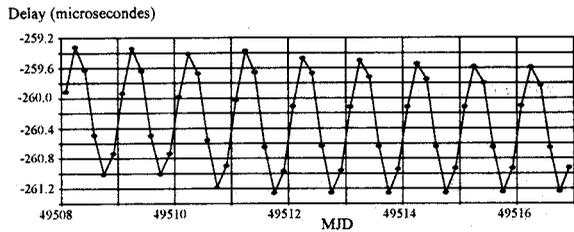


Figure 3 : Time difference OP-CT with Telecom 2 without satellite motion cancellation

To improve this method, we first used TDF2 satellite. Because of some lack of accuracy on the computed position of the satellite, we decided in May 1994 to follow on the experiment for 6 weeks with the Telecom 2A satellite for which accurate ephemeris were available. The satellite position is evaluated by a Kalman filter from station keeping measurements (Ranging and Turn-Around with 3 stations from the CNES 2 GHz network : Aussagel - France, Kourou - French Guyana, Hartebeeshtoeck - South Africa), dynamic models of moon and solar perturbation, solar pressure, earth potential and nutation.

During the experiment,  $\Delta t_i$  measurements between the 1 pss and the television signal are made every 4 hours,

each session, one point per minute and 16 points per session, lasts 15 minutes. At the end of the session, data are averaged to product the  $\Delta t_i$  value, using a least squares method, and stored. Files are sent to the Observatoire de Besançon to be pre-processed (correction of Sagnac, Ionospheric, Tropospheric and station delays). After the 6 week campaign, data are sent back to CNES, and time differences are computed. As a calibration of the method, results are compared with GPS time difference.

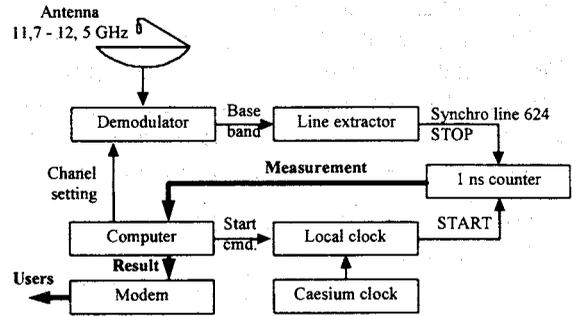


Figure 4 : Station synopsis

### UNCERTAINTY BUDGET

In order to evaluate the uncertainty of our method, we propose here an evaluation of the error budget. Each term is connected to (eq.1) or (eq.2).

#### Measurement precision

Elementary measurement showed an amplitude noise of 8 nanoseconds (Figure 7), which can be reduced by averaging.

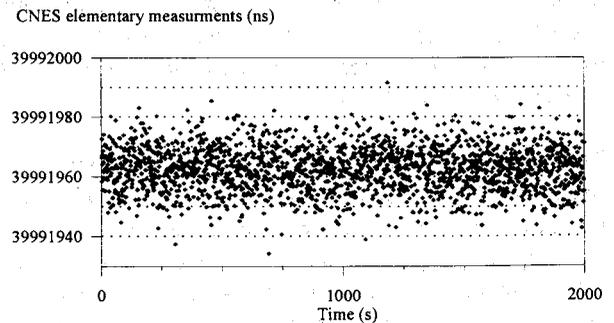


Figure 7 : elementary measurement

As we take 16 points per session we have, for each result  $\Delta t_i$  (eq.1), a precision of :

$$\sigma_{mes} = \frac{8}{\sqrt{16}} = 2 \text{ ns}$$

However, we must mention that the noise observed on measurement is quite more important than the noise foreseen by the link budget. The excess noise probably comes from the receiver, which is made with a commercial tuner.

### Receiver calibration

The main part of  $RST_1 - RST_2$  in (eq.1) comes from station delay between antenna and the counter. This delay can be obtained in two ways :

- a differential calibration (figure 5) using two co-located stations which gives the delay difference between the two stations, this method has been applied at the Observatoire de Besançon before sending stations to Toulouse, Grasse or Paris ;

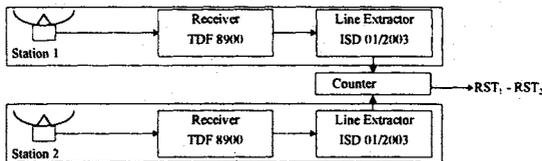


Figure 5 : Differential calibration

- a direct calibration (figure 6) using a satellite simulator in order to obtain  $RST_i$  for each station; this system has been developed in CNES.

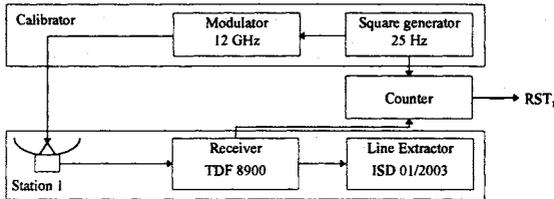


Figure 6 : Direct calibration

The comparison of the two methods shows some incoherence that may be due to some imperfections on the CNES equipment : the pulse generator does not produce the same signal as the satellite and the line extractor is not calibrated in this way. So we decided to use only differential measurements, which are sufficient (we need  $RST_1 - RST_2$  in eq.1).

The delays have been measured when stations were receiving TDF2 satellite. We have not made new measurements after having modified stations for Telecom 2 (new antenna, new tuner setting). The measurement accuracy of  $RST_1 - RST_2$  difference is estimated by the Observatoire de Besançon at about 8 nanoseconds.

### Receiver to local time link

Each receiver has a local clock which is linked to the laboratory caesium time. We estimate that we can measure the delay between local and main time with an uncertainty of :

$$\sigma_{\text{clk}} = 1 \text{ ns}$$

### Antenna position

The antenna position operates in the distance computation between satellite and station ( $D_{\text{Sat-Sti}}$  in eq. 2). They are known with an uncertainty of either about one meter or 3 nanoseconds.

### Satellite position

The satellite orbit is determined from Ranging and Turn-Around measurements. These measurements are computed with specific software and extrapolated at the time transfer session dates. Due to both noise measurement and differential models used, we can estimate that the satellite position, for  $D_{\text{Sat-Sti}}$  (eq. 2), is given with a uncertainty of :

$$\sigma_{\text{sat}} = 5 \text{ ns}$$

### Sagnac effect

Electromagnetic waves suffer a relativistic effect in a referential in rotation : the Sagnac effect. This delay, which is function of both position and velocity of transmitter and receiver, can be easily computed with a precision better than one nanosecond. However, as we suppose that the satellite is fixed, we have a differential effect on the time transfer ( $\tau_{\text{Sagnac1}} - \tau_{\text{Sagnac2}}$ ) :

$$\sigma_{\text{Sagnac}} = 1 \text{ ns}$$

### Ionospheric delay

Ionospheric delay ( $\tau_{\text{iono}}$  in eq. 2) comes from signal travel through ionising environment. It is low at the signal frequency (12 Ghz) ; differential effects ( $\tau_{\text{iono1}} - \tau_{\text{iono2}}$ ) is under one nanosecond because of very short distances between station :

$$\sigma_{\text{iono}} = 0.2 \text{ ns}$$

### Tropospheric Delay

Tropospheric delay  $\tau_{\text{tropo}}$  (eq. 3) results from the atmospheric bending of the satellite signal. This delay can be decomposed in a dry part and a wet part, which depend on both local weather and satellite elevation. For our stations (30° elevation), differential delay ( $\tau_{\text{tropo1}} - \tau_{\text{tropo2}}$ ) is about 1 or 2 nanoseconds, 5 nanoseconds in a worst case, which will degrade the time transfer of :

$$\sigma_{\text{tropo}} = 3 \text{ ns}$$

**Synthesis**

Table 1 gives the total uncertainty budget, which is valid only for stations placed in France (elevation of about 30 degrees). For lower elevations, ionospheric and tropospheric delays and errors become more important.

A uncertainty better than 10 nanoseconds is possible, if we are able to reduce the receiver calibration part : It seems conceivable to obtain 2 nanosecond after some new calibrations, which gives a total uncertainty of 7.3 nanoseconds at 1  $\sigma$ .

	Precision	Accuracy
Station measurement	2 ns	-
Receiver to local time link	-	1 ns
Receiver calibration	-	8 ns
Antenna position	-	3 ns
Satellite position	-	5 ns
Sagnac effect	-	1 ns
Ionospheric delay	-	0.2 ns
Tropospheric delay	-	3 ns
<b>Total uncertainty at 1 <math>\sigma</math></b>	<b>10.6 ns</b>	

Table 1 : Total error budget

**GPS comparison**

In order to calibrate the method, we have compared our results with GPS measurements.

$$\varepsilon = \Delta t_{ij-TC2} - \Delta t_{ij-GPS} \quad (3)$$

The total error budget on  $\varepsilon$  comes not only from the Telecom 2 link, but also from GPS link for which accuracy is about 4 nanoseconds

**TIME TRANSFER RESULTS**

To improve the Telecom 2 time transfer, we realised a 6 week campaign in May and June 1994. Data were collected in Toulouse (CT), Besançon (OB), Paris (OP) and Grasse (OC). After being pre-processed in Besançon (Ionospheric, Tropospheric, Sagnac and station delay cancellation), time differences were computed at CNES thanks to the satellite position of the satellite control centre.

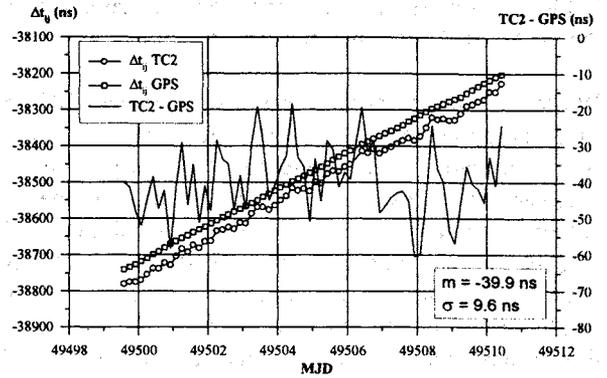


Figure 8 : (CT - OB) link

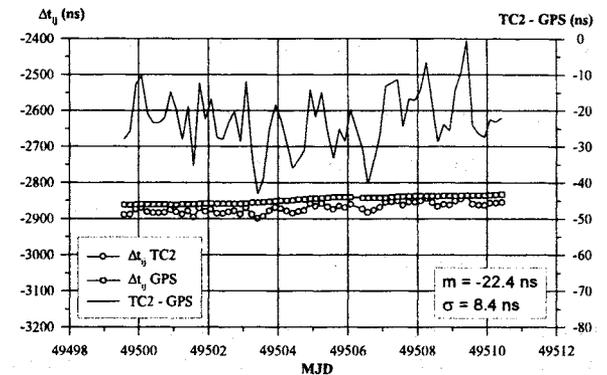


Figure 9 : (OB - OP) link

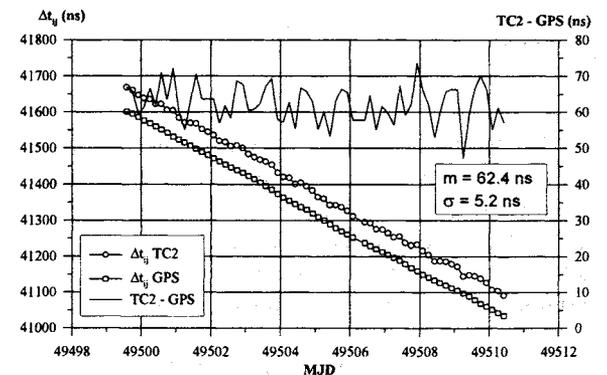


Figure 10 : (OP - CT) link

Figures 8 to 10 show results for 3 links. We made a comparison with smoothed common-view GPS data, to confirm our method accuracy.

Link	Mean	Standard deviation
CT - OB	-39.9 ns	9.6 ns
OP - CT	62.4 ns	5.2 ns
OB - OP	-22.8 ns	8.4 ns
CT - OC	-24.7 ns	8.5 ns
OB - OC	14.5 ns	6.2 ns
OP - OC	37.0 ns	7.8 ns

Table 2 : TC2 - GPS time transfer comparison

Table 2 summarizes means and standard deviations on the TC2 - GPS difference for all links.

We observe an important mean error between GPS and TC2 measurements. This difference probably comes from constant biases, finding their origin in the receivers : Our results were worse before we checked antenna position in each station and local time link, receivers are the only system of the station which have not been controlled since we have been using Telecom 2 satellite. Assuming that the observed mean errors are due to these biases, measured standard deviation are in accordance with the error budget, lower than 10 nanoseconds.

As a validation, we can evaluate the closing equation :

$$\begin{aligned} (CT - OB) + (OB - OP) + (OP - CT) &= -0.3 \text{ ns} \\ (CT - OB) + (OB - OC) + (OC - CT) &= -0.7 \text{ ns} \\ (CT - OC) + (OC - OP) + (OP - CT) &= 0.7 \text{ ns} \\ (OP - OB) + (OB - OC) + (OC - OP) &= 0.3 \text{ ns} \end{aligned}$$

Results lower than one nanosecond prove the reliability of the method and the existence of constant undetermined biases during the 6 week campaign.

### ORBIT DETERMINATION

We decided to evaluate an original orbit determination method using time measurements, a short base (all stations located in France instead of located in France, South Africa and French Guyana) and very simple stations, using 80 centimetres commercial. This is permitted by eq. 1 which can be re-arranged in :

$$(\tau_1 - \tau_2) = (\Delta t_1 - \Delta t_2) - \Delta t_{12} - (RST_1 - RST_2) \quad (3)$$

The knowledge of  $\Delta t_{12}$  by another way, in our case GPS, permits to evaluate the pseudo-range  $\tau_1 - \tau_2$  for a given link :  $\tau_1 - \tau_2$  determined a hyperboloid on which the satellite is. Three hyperboloids, and so three links, are necessary to know the satellite position.

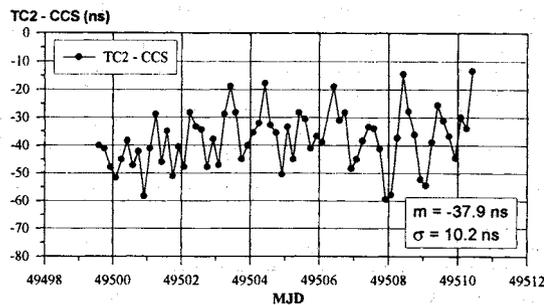


Figure 11 : TC2 - CCS pseudo-range for CT-OB

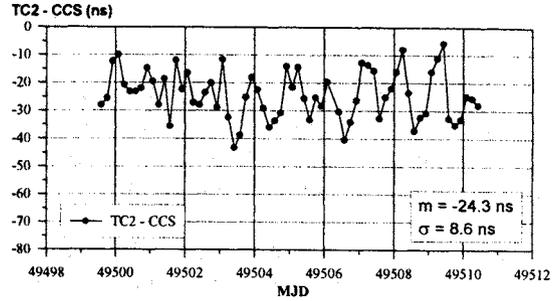


Figure 12 : TC2 - CCS pseudo-range for OB-OP

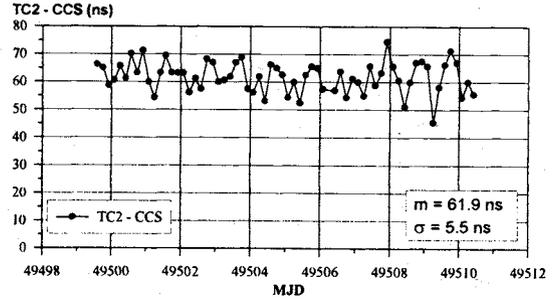


Figure 13 : TC2 - CCS pseudo-range for OP-CT

We first compared the pseudo-ranges computed from time differences measurements with those given by the satellite control centre for each link (Figure 11 to 13). We observed the same biases (Table 3, measured biases) and standard deviations as for the comparison with GPS, but no periodic component due to the satellite motion : That proves a good coherence between the two methods.

	Ct-Ob	Op-Ct	Ob-Op	Op-Oc	Ct-Oc	Ob-Oc
M	-37.9	61.9	-24.3	39.9	-22.4	15.7
F	-41	63	-26	42	-17	18

Table 3: Biases in nanoseconds measured (M) or filtered (F)

Then, pseudo-range data have been introduced in a Kalman filter to determine the orbit. Biases on each link have been identified by the filter in order to obtain the best orbit (Table 3).

	One way time measurement method	Ranging / Turn Around methods
$\sigma_A$	2.0 m	5.0 m
$\sigma_{Ex}$	$5.0 \cdot 10^{-7}$	$1.0 \cdot 10^{-7}$
$\sigma_{Ey}$	$5.0 \cdot 10^{-7}$	$2.0 \cdot 10^{-7}$
$\sigma_{Ix}$	$0.5 \cdot 10^{-4} \text{ d}^\circ$	$2.0 \cdot 10^{-4} \text{ d}^\circ$
$\sigma_{Iy}$	$0.5 \cdot 10^{-4} \text{ d}^\circ$	$2.0 \cdot 10^{-4} \text{ d}^\circ$
$\sigma_{I\Omega}$	$1.0 \cdot 10^{-4} \text{ d}^\circ$	$2.0 \cdot 10^{-4} \text{ d}^\circ$

Table 4 : Orbit parameters precision

The orbit determination consists in evaluating its semi major axis (A), eccentricity (Ex, Ey), inclination (Ix, Iy) and mean longitude. The standard deviation on each parameter is compared with those obtained with classical methods (Table 4). Results show that the time measurement method is compatible with station keeping exigencies.

## FUTURE DEVELOPMENTS

The results obtained during this 6 week campaign demonstrate all the potentiality of the method, as well for time transfer as for orbit determination :

- For time transfer, we are able to propose a low cost alternative or redundancy to GPS ;
- For orbit determination, we prove that the station keeping could be possible with a short base and very low cost station, if it is able to know the time difference between the local time scales, by GPS or another method.

However, we have to improve the system accuracy by reducing biases, which seem incompatible with the error budget. For the next campaign, planned in 1996, the CNES will explore the following ways :

- Making new calibration of stations, in order to reduce biases, with a new calibrator developed at the CNES and based on a television signal generator (Figure 14);
- Developing a new professional receiver, with a lower noise and stable delays, in order to assure delay stability and a better noise measurement, compatible with the link budget ;

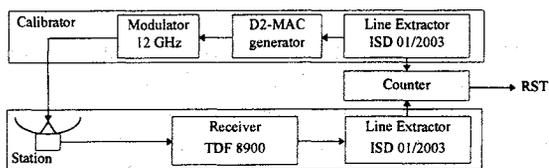


Figure 14 : CNES new calibrator

- Automating data transfers between stations, using a computer network, to propose a full real time system able to give time transfer results next to each measurement session without human work.

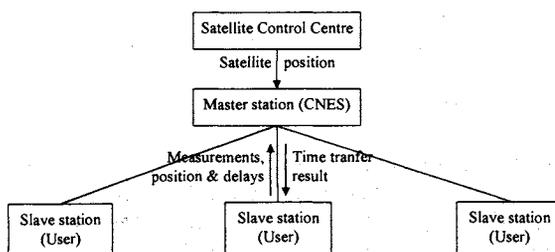


Figure 15 : Station network

Then, we will be able to propose an near operational time transfer method, independent of GPS, using geostationary direct broadcasting Telecom 2 satellites for which television and station keeping services will be provide until 2005; or an orbit determination method using time difference measurements.

## ACKNOWLEDGEMENTS

The authors wish to thank J.F Dutrey (CNES), B. Mougine (Observatoire de Besançon) and R. Helmlinger (LPTF) for their helpful work of collecting and filing data, P. Debay, C. Fremeaux and M. Moury ( TC2 Satellite Control Centre - CNES) for their co-operation and the TC2 station keeping measurements, J.P. Prost (Dassault Aviation) for his calculation on orbit determination and A. Vincent (Observatoire de Besançon) for the modification of the receivers.

## REFERENCES

1. J.A. Davis, P.R. Pearce, P. Menary, M. Hutchins, M. Sagin, 1991, " Time transfer by passive use of direct broadcasting satellite signals ", *EFTF*, 5, 74-81
2. F. Meyer, M. Brunet , M. Granveaud ,F. Vernotte, M. Vincent, 1991, " Time transfer using a geostationary direct TV satellite ", *EFTF*, 5, 82-88
3. F. Meyer, F. Vernotte, M. Vincent, M. Granveaud, B. Laporte, 1993, " Improved time transfer using a geostationary direct TV satellite ", *EFTF*, 7, 129-134
4. F. Meyer, M. Granveaud, B. Laporte, F. Vernotte, M. Vincent, 1992, " Synchronisation d'horloges atomiques par TDF1 et TDF2 : Résultats et perspectives ", *Bulletin du BNM*, 90, 15-22
5. Ph. Guillemot, F. Meyer, J.P. Prost, M. Brunet, M. Vincent, J. Foliard, 1995, "Synchronisation d'horloges atomiques par TDF2 & Télécom2", *CNES, CR-95-8-CT/AE/TTL/TF*

STUDIES OF GaPO<sub>4</sub> CRYSTALS AND RESONATORS

A.ZARKA\*, B.CAPELLE\*\*+J.DETAINT\*\*, D.PALMIER\*\*\*,  
E.PHILIPPOT\*\*\* & O. V. ZVEREVA\*\*\*\*

\* Laboratoire de Minéralogie-Cristallographie. Université P.et M Curie. 4.Place Jussieu. 75252 Paris-Cedex 05.France.

\*\* C.N.E.T. Laboratoire de bagueux, 196 Av. H. Ravera. 92225 Bagneux. France.

\*\*\* Laboratoire de Physicochimie des Materiaux Solides . U.R.A. D0407 C.N.R.S. Université de Montpellier II. France.

\*\*\*\* Institute of Crystallography, 117333. Leninsky pr. 59. Moscow. Russia.

+ L.U.R.E. Batiment 209D.Centre Universitaire Paris-Sud. 91405.Orsay-Cedex.

## ABSTRACT

The crystalline perfection of GaPO<sub>4</sub> samples grown in different conditions have been studied using the X-ray topographic technique and the synchrotron radiation delivered at L.U.R.E. (Orsay, France). The results concerning the quality and the piezoelectric features of the obtained samples are given.

## INTRODUCTION

The gallium phosphate (GaPO<sub>4</sub>) is, as berlinite, a close crystallographic analog of quartz with the advantage that it has no  $\alpha$ - $\beta$  phase transition. It was shown that this material has very attractive properties for the applications to piezoelectric devices, due to its large coupling coefficients and to the existence of compensated cuts (1,2). Furthermore, it can be used up to extremely high temperatures (900°C). As this material is obtained by the hydrothermal growth method, seeds are necessary. To obtain rapidly much larger crystals, epitaxial growth on berlinite plates can be performed. With seeds cut in the X zone of such crystals a new growth in phosphoric acid gives large crystals. Using the synchrotron radiation delivered by the DCI storage ring at L.U.R.E. (Orsay, France), we have studied by the X-ray topography technique the crystalline perfection of new GaPO<sub>4</sub> samples. For this study we have used mainly Laue traverse topography .

## PROPERTIES OF MATERIALS (3)

The most important properties related these piezoelectric materials which belong to the class of symmetry 32 are summarized in the table 1 with those of quartz and berlinite.

Quartz and its analogue berlinite have phase transitions below their melting point in contrary to the gallium phosphate . They display non negligible solubilities in both alkaline and acid aqueous solutions at elevated temperature and high pressure. Consequently the natural choice for their growth is the hydrothermal method. Moreover the quartz and berlinite crystals have a comparable low density whereas gallium phosphate has higher one (more than twice that of quartz).

GaPO<sub>4</sub> CRYSTAL GROWTH

Some interesting results can be emphasized from many growth experiments carried out at  $T_c \leq 240^\circ\text{C}$  in different acid media investigated (H<sub>3</sub>PO<sub>4</sub>, H<sub>2</sub>SO<sub>4</sub> and HCl) and their mixtures:

TABLE 1- Comparison of some physical properties.

	Quartz	AlPO <sub>4</sub>	GaPO <sub>4</sub>
coupling coeff.*k %	8.5	11.0	>16.0
surtension coeff.* Q	3.10 <sup>6</sup>	10 <sup>6</sup>	>5.10 <sup>4</sup>
Transition $\alpha \rightleftharpoons \beta$ °C	573	586	no

\* AT cut

- only pure phosphoric acid (or its mixtures with a low proportion of other acids), leads to significant  $V_z$  growth rate. The  $V_x/V_z$  ratio decreases to a value close to 1.0, while the H<sub>3</sub>PO<sub>4</sub> acid concentration increases to 20 M.l<sup>-1</sup>.

- in H<sub>2</sub>SO<sub>4</sub> medium (or its mixtures with a low proportion of other acids), the  $V_x/V_z$  ratio is always very high ( $V_z$  very weak). So, this solvent cannot be used to large crystal growth. Nevertheless, the high solubility and the low dissolution kinetics of GaPO<sub>4</sub> in this solvent authorise higher growth temperatures and, so, should lead to crystals with lower OH impurity content.

- in HCl medium (or its mixtures with a low proportion of other acids), the  $V_x/V_z$  ratio is rather good. Unfortunately, both growth rates,  $V_x$  and  $V_z$  are too slow (generally less than 0.05 mm/d.face).

A few conclusions can be extrapolated from all the growth experiments in the investigated range of retrograde solubility for  $T_c \leq 240^\circ\text{C}$ . The sulphuric acid is theoretically the best solvent for GaPO<sub>4</sub> crystal growth ( $T_c \geq 200^\circ\text{C}$  conducting to crystals with low OH content, but the too strong growth rate anisotropy does not enable an industrial development. For this purpose, *the concentrated phosphoric acid seems to be the most suitable* ( figure 1.).

## Seed Lengthening

Excepted for the most concentrated pure phosphoric acid, any of the other acid media investigated have presented  $V_y$  growth rate  $> 0.01$  mm/d.face. On the other hand, if plates of good enough dimensions can be synthesized by the epitaxial growth of GaPO<sub>4</sub> on large berlinite seeds (4), their very weak lengthening along

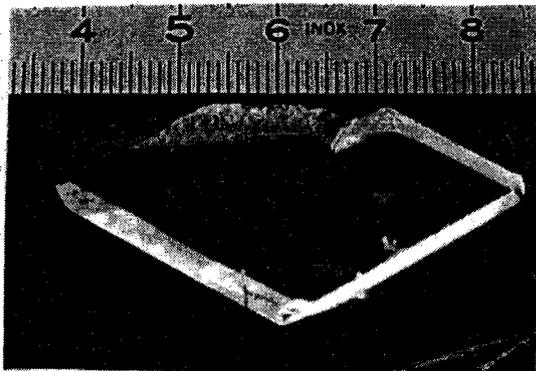


Figure.1:Photograph of a sample obtained with  $H_3PO_4$  solution.

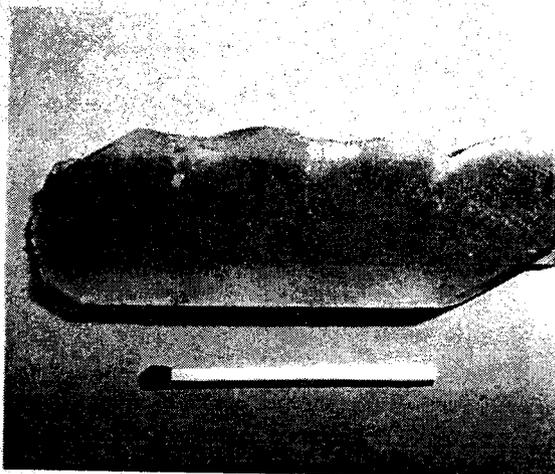


Figure.2:Photograph of a crystal obtained by Seed lengthening.

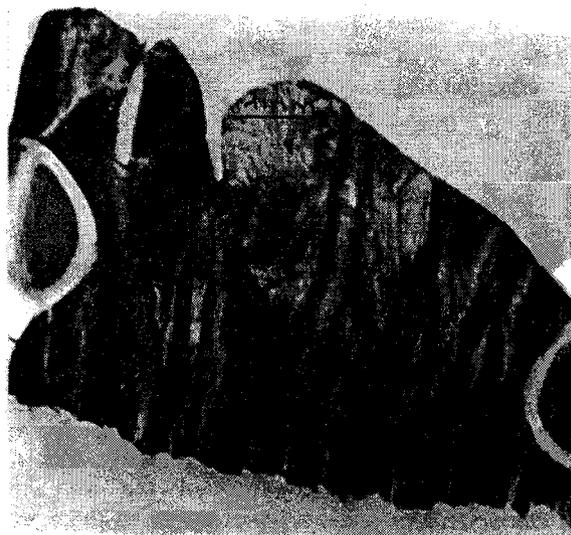


Figure.3:X-ray topograph of a slice obtained from sulfuric acid. Dendritic aspect.

the Y direction cannot be improved without a purely manual method.

Indeed, we have undertaken the growth of large crystals directly from several small Z-seeds spliced on a quartz plate. Crystal growth experiments have been carried out through a modified horizontal composite gradient method (HTG)  $\pm 10^\circ$  oscillating angle around an horizontal axis perpendicular to the glass vessel, with growth conditions;  $T_c = 150^\circ C$  and  $15 M.l^{-1}$  phosphoric acid concentration.

The resulting crystals show a very flat surface without any obvious defects at the interface of the small seeds (figure 2).

Nevertheless, after the first crystal growth, X-ray topographs revealed an important dislocation density close to the seed junction. These defects can be easily and drastically decreased using the "crossed" crystal growth method (5,6), i.e., successive growths along both X and Y directions.

### X-RAY TOPOGRAPHY

Several plates of gallium phosphate crystals have been studied by X-ray transmission topography using the white beam delivered by the D.C.I. storage ring at LURE (Orsay, France). Some examples are given in the following. In fig.3 and fig.4 are given two typical topographs of Y cut slices of two crystals grown by VTG method in  $H_2SO_4, 230^\circ C$  (7,8,9). In fig.3 we can observe the great amount of dislocations which are grouped in bundles and the dendritic aspect of the growth. In Fig.4. the dislocations are also very numerous and they start from the seed in the two parts of the grown crystal. The circular white contrast is due to the fixation of the slice.

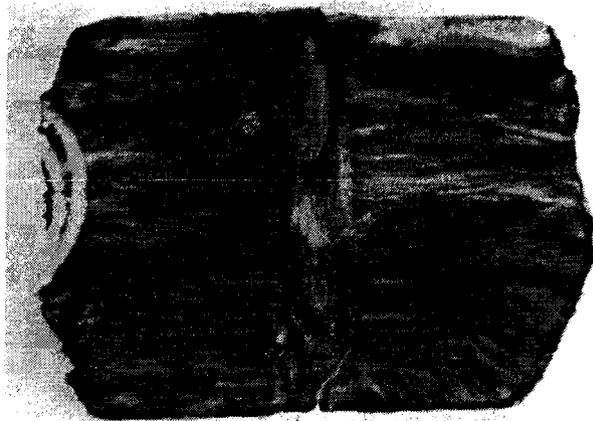


Figure.4:X-ray topograph of a slice obtained from sulfuric acid. Observe the dendritic aspect and the numerous dislocations.

In Fig.5 and Fig.6 are given two other topographs corresponding to crystals grown by the SHT method in  $H_3PO_4$  (4). The two samples correspond to Y cut slices. It can be observed that in the two cases the density of the dislocations is not so high, and that a slight disymmetry is observed in the two grown parts (above and below the seed in the Fig.5). In Fig 6 a stop in the growth has induced the creation of several bundles of dislocations. As in the previous cases (Fig. 3 and Fig.4) the samples presented in Fig. 5 and Fig.6 do not present any twin.

In Fig.7, 8 and 9 are presented the topographs of other crystals grown by O. V. Zvereva et al. (10). The quality of the crystal presented in Fig.7 is quite good and the sample presents very few dislocations. In the following topograph (Fig. 8) the sample is of a very good quality but it shows some growth bands in different parts.

In Fig 9 is shown a topograph of a resonator. In the centre it can be observed a black contrast due to the vibration. The fringes on the side are due to coupled modes.

To summarize this part, it seems that the quality of several crystals obtained by the different techniques of growth is quite good. The density of the dislocation is generally not so high and depends of the quality of the seeds and of the growth conditions.

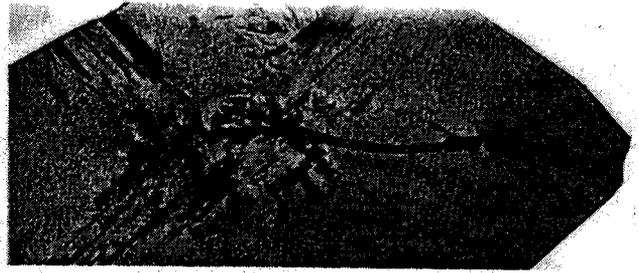


Figure.7:X-ray topograph of another slice. The quality is good and there are very few dislocations.



Figure.8:X-ray topograph of another slice. The quality is good and there are very few dislocations. Some growth bands appear in different parts of the crystal.

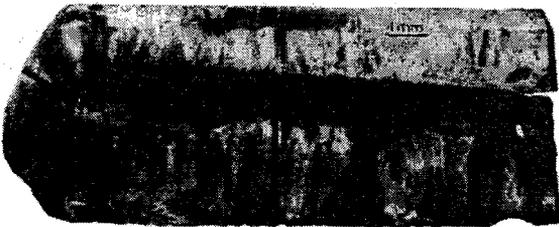


Figure.5:X-ray topograph of a slice obtained from phosphoric acid. In the two grown parts the dislocations are grouped in bundles.

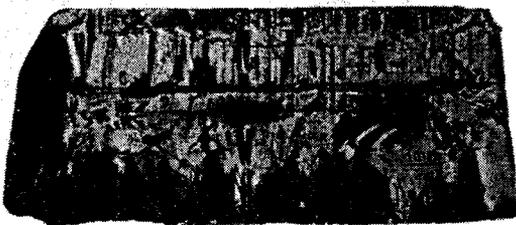


Figure.6:X-ray topograph of a slice obtained from phosphoric acid. The quality of the crystal is quite good.

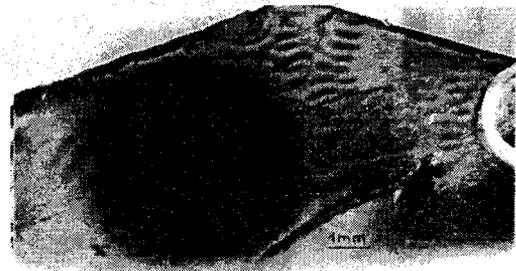


Figure 9: X-ray topograph of a resonator under vibration.

#### PIEZOELECTRIC CHARACTERIZATIONS

These characterizations have been done on Y rotated cut slices from different crystals grown in sulfuric acid.

To correlate the piezoelectric properties and the growth conditions we have measured the characteristics of resonators between  $-30^{\circ}C$  and  $150^{\circ}C$ . From these measures we have extracted the piezoelectric characteristics and we have compared them with the same properties calculated with the constants measured by Krempf and al (3).

In table 2 we have compared the values of the surtension, the coupling and temperature coefficients. The coupling coefficient was taken from the formula of the theory of unidimensional slices. The value noted  $K^{**}$  was obtained by adjusting the antiresonance frequency from the stray capacitance effect. We observe that the surtension is generally about 20000 instead of the values obtained with the crystals grown at lower temperature. In these last cases these values were comprise between 10000 and 15000. The coupling coefficient is quite equal to the calculated values and the temperature coefficients are quite always lower than the calculated ones.

The temperature coefficient of the slices with positive rotation angle (Y+14) is given in table 2.

For the Y-13 cut which is near of the temperature compensated curve we observe a cubic behaviour with a zero of the temperature coefficient between 130°C and 160°C (this can be very useful for application at high temperatures). The measurements obtained with orientations between 13°02' and 13°25' show that the compensated cut at room temperature has its orientation shifted by 1.5° to 2° in comparison with the low temperature crystals.

TABLE 2-Coupling coefficients and thermal properties of some Y rotated Cuts.

Coupling Coefficients and Thermal Properties of some Y rotated cuts of GaPO<sub>4</sub>

	Y-13°02'	Y CUT	Y+14°09'	Y+15°	Y+21°
Q factor	21300	8340#	26200	20500	19800
k (%)	15.39#	18.56	16.72	16.68#	16.15
k <sup>**</sup> (%)	16.08	19.52	19.33	17.33#	17.45
TCF <sub>exp</sub> (10 <sup>-6</sup> /°C)	+3.92	19.9	40.5	42.08	46.5
k <sub>comp</sub> (%)	16.39	19.10	19.17	19.10	18.18
TCF <sub>comp</sub> (10 <sup>-6</sup> /°C)	0.980	35.7	56.8	57.2	59.6
TCF <sub>a<sub>comp</sub></sub> (10 <sup>-6</sup> /°C)	+3.37	33.4	50.2	50.5	52.2

\*\* corrected from stray capacitance  
# larger values were measured with other samples  
TCF are measured or computed for 25°C

TABLE 3-Piezoelectric properties of some Y rotated Cuts.

Piezoelectric Properties of some Y rotated Cuts

	Y-13°02'	Y CUT	Y+14°09'	Y+15°	Y+21°
Fr (Hz)	7894200.	8751762.	9485655.	9169240.	5039750.
Fa (Hz)	7971620.	8874059.	9594238.	9273918.	5094200.
2h (μm)	159.4μ	143.5μ	144.5μ	153.0μ	285.0μ
2Re(mm)	1.500	1.500	1.15	2.00	4.12*3.24
Cp+Co (pF)	0.727	0.865	1.210	2.198	0.271
Fa <sup>**</sup> (Hz)	7978753.	8891285.	9634088.	9283709.	5103518.
R̄ (%)	1.9026	.9947	1.053	.6621	1.1356
Δ (%)	1.5008	1.156	1.483	.7683	0.7178
F <sub>s</sub> (1+R̄-Δ)	8005057.	8859716.	9592618.	9264067.	5115499.
C̄ <sub>exp</sub> (10 <sup>9</sup> N.m <sup>-2</sup> )	23.25	23.08	27.21	28.69	30.35
V <sub>exp</sub> (ms <sup>-1</sup> )	2552.0	2542.7	2760.8	2834.8	2915.8
C̄ <sub>exp</sub> <sup>**</sup> (10 <sup>9</sup> N.m <sup>-2</sup> )	23.29	23.17	27.44	28.75	30.46
V <sub>exp</sub> <sup>**</sup> (ms <sup>-1</sup> )	2554.3	2547.6	2772.3	2837.8	2921.1
C̄ <sub>comp</sub> (10 <sup>9</sup> N.m <sup>-2</sup> )	22.03	23.23	26.16	26.31	28.88
V <sub>comp</sub> (ms <sup>-1</sup> )	2484.2	2551.2	2706.9	2715.2	2794.6

\*\* corrected from stray capacitance Cp

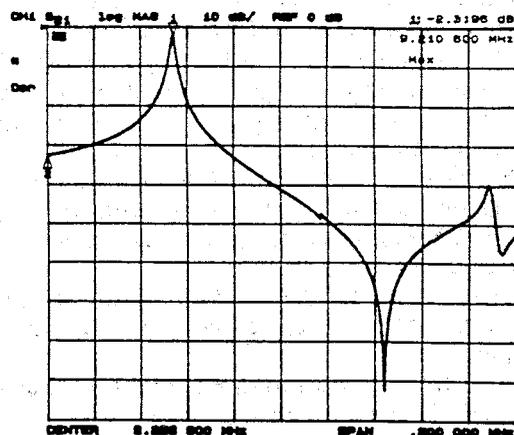


Figure 10: Typical response of a resonator GaPO<sub>4</sub> grown in H<sub>2</sub>SO<sub>4</sub> at 230°C.

In table 3 are shown some results concerning one resonator of each cut (Y-13, Y+14...) and we have given some values of stiffened elastic constants and phase velocities. This was done by correcting the mass loading and the electrodes dimensions so as to obtain the antiresonance frequency of the infinite non-electroded slice from the data of the resonators with finite dimensioned electrodes. For that, we have used a tridimensional model of resonator which has permit the correction of the electrode dimension noted Δ in table 3. The total correction R+R̄-Δ was applied to the antiresonance frequency Fa directly measured or to Fa<sup>\*\*</sup> the antiresonance frequency corrected of the stray capacitance effect. The effective constant values (stiffened values) and those of the velocities which were obtained are compared to the data obtained with the constants of Kreml and al. (3). It can be observed that the values of the constants are generally greater than those calculated. It is the same for the values of the velocities. It can be noted that the direct extraction from the antiresonance frequency measured on the resonators give also good results. This seems very surprising. For the same reason the corrections of the parasitic capacitances are weak. The figure 10 shows a typical response of a resonator GaPO<sub>4</sub> (H<sub>2</sub>SO<sub>4</sub>, 230°C). This curve was obtained at 125°C and shows an excellent response and that these devices can be used in oscillators or filters. We have observed that in these crystals, the acoustic losses are independant of the temperature. This seems to prove that the impurities amount and particularly those of OH is greatly reduced.

CONCLUSION

Study of crystalline perfection of GaPO<sub>4</sub> has been done in relation with the crystal growth. Different parameters of the growth have been studied. In relation with that, X-ray topography was used either to examine the quality of the samples or to observe the vibration modes of the corresponding resonators. Crystal growth results are discussed from many growth experiments carried out at Tc < 240°C in different acid media (H<sub>3</sub>PO<sub>4</sub>, H<sub>2</sub>SO<sub>4</sub> and HCl and their mixtures) with different growth techniques. In the present states of our investigations, pure phosphoric acid or its mixtures with low proportions of other acids can be

considered for an industrial development. For the crystals obtained from  $H_2SO_4$  solutions at low temperatures (200°C-230°C) the X-ray topographs have shown few defects. At higher temperatures the growth presents a dendritic aspect (good quality can be obtained by reducing the growth rate).

Concerning the crystal growth in  $H_3PO_4$  it appears that the quality is better when the temperature is higher (160°C-300°C). In this particular case, some crystals have been also examined.

Resonators using six Y rotated-cut and doubly rotated one were made from the different growth technique samples.

The frequency responses and the thermal behaviour were measured for the fundamental mode and some other overtones. It was observed a slight angular shift of the main compensated cut with an increase of the growth temperature and a confirmation of the interest of  $GaPO_4$  crystals (large coupling coefficient and outstanding thermal stability).

In conclusion, these studies have shown that an excellent quality is already obtained and this permits to make good resonators for filters or VCXO with large frequencies and can lead to good high stability oscillators.

## REFERENCES

- (1) Krempl P.W., Stadler J., Wallnöfer W., Ellmeyer W. and Sclic R., 1991, Present state of  $GaPO_4$  reseach. Proc. 5th European Frequency and Time Forum. Besançon, 143-147
- (2) Philippot E., Ibanez A., Goiffon A., Cochez A., Zarka A., Capelle B., Schwartzel J. and Detaint J., 1993, A quartz-like material: gallium phosphate ( $GaPO_4$ ) crystal growth and characterization, J. of Cryst. Growth 130, 195-208.
- (3) Krempl P.W., Krispel F., Wallöfer W. and Leprecht G., 1995,  $GaPO_4$ . A critical review of material data. Proc. 9th European Frequency and Time Forum. Besançon, 66-70.
- (4) Cochez M., Foulon J.D., Ibanez A., Goiffon A., Philippot E., Capelle B., Zarka A., Schwartzel J. and Detaint J., 1994, Crystal growth and characterization of a quartz-like material :  $GaPO_4$ . J. de Physique IV. Colloque C2 Vol.4, 83-194.
- (5) Zarka A., Detaint J., Schwartzel J., Toudic Y., Capelle B., Zheng Y.L., Philippot E., Buisson X. and Arnaud R., 1991. French Patent 91.14007.
- (6) Zarka A., Liu Lin and Buisson X., 1981, Influence de la localisation sectorielle du germe sur la qualité cristalline du quartz de synthèse, J. of Cryst. Growth 54, 394-398.
- (7) Detaint J., Schwartzel J., Joly C., Toudic Y., Capelle B., Zarka A., Zheng Y.L., Toudic Y., Philippot E., 1992., Proc. 6th European Frequency and Time Forum. Noorwijk 383-388.

(8) Capelle B., Zarka A., Schwartzel J. and Detaint J., Ibanez A., Philippot E. and Denis J.P., 1994, study of gallium phosphate and langasite crystals and resonators by X-ray topography. Proc. 1994 IEEE International frequency control symposium, 48-57.

(9) Philippot E., Ibanez A., Goiffon A., Capelle B., Zarka A., Schwartzel J. and Detaint J., 1992, Comparison of energy trapping resonators and filters using the AT cut of quartz, berlinite and gallium phosphate. Proc. 6th European Frequency and Time Forum. Noorwijk, 223-227.

(10) Zvereva O.V., Mininon Yu. M, and Demianets L.N., 1994, Hydrothermal growth of OH-free  $AlPO_4$  and  $GaPO_4$  crystals. The way of twin reducing. J. de Physique IV. Colloque C2 Vol.4, 19-24.

## TOLERABLE ORIENTATION ERRORS IN DETERMINATION OF THE MATERIAL CONSTANTS OF CRYSTALS

Carl K Hruska

Piezoelectricity Research Laboratory, York University  
North York, Ontario, M3J 1P3 Canada

Orientation errors in the crystal units used to study crystal properties affect the accuracy of the calculated material constants. Tolerable orientation errors satisfy two conditions: 1) they do not invalidate the calculation of errors due to experiment; and 2) they do not substantially add to their values. Their values depend on the experiment and the number of crystal units used and can be determined by simulations. In five cases analyzed in detail, they have been found to be in the range between 9' and 8".

### 1. INTRODUCTION

In order to assess the quality of published values of material constants of crystals and the significance of a disagreement among them in case of conflicts, one needs reliable values of their standard errors. The least-squares method alone will not produce such errors because it ignores some factors adding to their values. One of them is the misorientation of the crystal units used in experiments to determine the material constants. As some orientation errors are unavoidable, two questions need to be answered:

- 1) what is the effect of orientation errors on the accuracy of material constants, and
- 2) what are the tolerable orientation errors.

As the problem is difficult to deal with analytically, numerical answers are sought by means of the Monte Carlo method using simulations described in Hruska (1). Five real-life examples made for  $\alpha$ -quartz provide illustrative results and sufficient data to formulate several general conclusions.

The numerical values of all material constants and their errors appearing in this article are stated for room temperature, right-hand quartz and the frame of reference according to IEEE Standard of 1978 (2).

### 2. TERMINOLOGY AND SYMBOLS

1) Material constants of crystals are determined by means experiments

conducted on suitably chosen crystal units (rods, plates, cubes etc.). Their orientation is given by the IEEE rotational symbol  $xzlw(\psi/\phi/\theta)$  (2). Each unit is generally described by three rotational angles  $\psi_k$ ,  $\phi_k$ , and  $\theta_k$ ,  $k = 1, 2, \dots, t$ , where  $t$  is the number of crystal units used in a given experiment.

Experimentally determined values of the rotational angles are their 'best' available values. As such they are denoted  $\psi_k^B$ ,  $\phi_k^B$ , and  $\theta_k^B$ . Because of measurement errors, they are only nominal values differing from the 'true' rotational angles which remain unknown. For each angle, the difference between its true and its nominal value represents an orientation error. To make this study tractable, it is assumed that the errors of all 3t rotational angles are random and normally distributed with a common standard deviation  $\sigma^M$ . The superscript M indicates that  $\sigma^M$  characterizes errors caused by 'misorientation'.

Using these assumptions the nominal values of the rotational angles can be independently randomized by means of

$$\alpha^R = \alpha^B + N(0,1) \cdot \sigma^M, \quad (1)$$

where  $\alpha$  stands for any of  $\psi_k$ ,  $\phi_k$  and  $\theta_k$ ,  $k = 1, 2, \dots, t$ .  $N(0,1)$  is the standard normal variate whose value is selected at random. Symbols  $\psi_k^R$ ,  $\phi_k^R$  and  $\theta_k^R$  will represent the randomized values of the rotational angles.

2) Experiments to yield the material

constants consist of taking readings of a suitable experimental quantity  $E$  defined on the crystal units. Each unit is used in one or more ways to provide one or more different values of  $E$ . In all, a set of  $m$  experimental values  $E_i$  is obtained,  $i = 1, 2, \dots, m$ . It is assumed that errors in  $E_i$  are random errors. In this paper these errors are referred to as experimental errors.

3) Unless referring to a specific material constant from literature, material constants in this paper are generally denoted  $x_j$ ,  $j = 1, 2, \dots, n$ , where  $n$  is the number of different material constants calculated from the  $m$  experimental values  $E_i$ . Typically  $m > n$ , and the values of  $x_j$  are obtained from an overdetermined linear system by the least-squares fit (Draper and Smith (3)). The difference  $m - n$  is referred to as the number of degrees of freedom of the linear system.

The rotational angles  $\psi_k$ ,  $\phi_k$ , and  $\theta_k$  are necessary participants in the calculation as they are needed to obtain the matrix of the linear system. If the best values  $\psi_k^B$ ,  $\phi_k^B$ , and  $\theta_k^B$  are used there, the results are regarded as the best obtainable values of the material constants and denoted  $x_j^B$ . If the randomized values  $\psi_k^R$ ,  $\phi_k^R$ , and  $\theta_k^R$  are used, the obtained values of  $x_j$  are denoted  $x_j^R$ .

The least-squares process produces also the standard errors of  $x_j$ .

Corresponding to  $x_j^B$  and  $x_j^R$ , they are denoted  $\vartheta^E(x_j^B)$  and  $\vartheta^E(x_j^R)$ ,

respectively. As will be explained later, these errors are related to the experimental errors. The superscript  $E$  is used to indicate this relationship.

### 3. ERRORS OF MATERIAL CONSTANTS DUE TO EXPERIMENTAL ERRORS

These errors are the above mentioned standard errors  $\vartheta^E(x_j^B)$  obtained by the least-squares process. It is generally accepted that, for each calculated

value  $x_j^B$ , the pair of points  $x_j^B - \vartheta^E(x_j^B)$  and  $x_j^B + \vartheta^E(x_j^B)$  defines a  $\approx 68\%$  confidence interval for the true value of the sought material constant  $x_j$ .

This interpretation deserves some comments:

1) The least-squares process is designed to work on the assumption that the only source of uncertainty in the linear system solved is the experimental errors. Correspondingly,  $\vartheta^E(x_j^B)$  reflect only these experimental errors.

2) The interpretation of  $\vartheta^E(x_j^B)$  in terms of a 68% confidence interval is valid only if the distribution of the experimental errors is normal or if the number of degrees of freedom  $n - m$  of the linear system is sufficiently large. The latter can always be arranged and a potential problem can thus be avoided.

3) To obtain the correct values of  $\vartheta^E(x_j^B)$ , it is required that the true rotational angles be used in their calculation. However, under practical conditions, only their nominal values are available. This generally distorts  $\vartheta^E(x_j^B)$  (Table 1) so that they no longer accurately reflect the effect of the experimental errors. This condition must be rectified.

Standard errors  $\vartheta^E(x_j^R)$  are continuous functions of  $\sigma^M$  and, for  $\sigma^M = 0$ ,

$$\vartheta^E(x_j^R) = \vartheta^E(x_j^B), \quad j = 1, 2, \dots, n. \quad (2)$$

For this reason it is possible to find  $\sigma_1^M$  for any choice of  $k_1 > 0$  and  $0 \leq p < 1$ , such that the probability

$$P(|\vartheta^E(x_j^R) - \vartheta^E(x_j^B)| < k_1 \cdot \vartheta^E(x_j^B)) = p, \quad j = 1, 2, \dots, n, \quad (3)$$

holds for all

$$\sigma^M \leq \sigma_1^M. \quad (4)$$

A choice of  $p$  close to 1 and  $k_1$  sufficiently small provide a practical certainty that the difference between

$\varphi^E(x_j^B)$  and all possible values of  $\varphi^E(x_j^R)$  is arbitrarily small. This is likely to include also  $\varphi^E(x_j^R)$

calculated for the combination of the true values of the rotational angles. As a result,  $\varphi^E(x_j^B)$  become a

sufficiently accurate representation of the errors due to experiment, no matter what the true values of the rotational angles may be. As such, condition (3) defines the first limitation placed on the tolerable size of orientation errors.

#### 4. ERRORS OF MATERIAL CONSTANTS DUE TO MISORIENTATION

The values of material constants obtained by the least-squares process generally depend of the values of the rotational angles used in the calculation (Table 1). The fact that, instead of their true values, only the nominal rotational angles  $\psi_k^B$ ,  $\phi_k^B$ , and  $\theta_k^B$  are available to calculate  $x_j^B$  is a source of errors in  $x_j$  - errors due to misorientation. These errors are not included in  $\varphi^E(x_j^B)$  and must be assessed separately.

Using Eq. (1) to obtain randomized values  $\psi_k^R$ ,  $\phi_k^R$ , and  $\theta_k^R$ ,  $k = 1, 2, \dots, t$ , and using these values in turn in the least-squares process produces a distribution of randomized values  $x_j^R$ ,  $j = 1, 2, \dots, n$ . For each  $j$ , the standard deviation of this distribution has been adopted (1) as the error in  $x_j$  due to misorientation. This standard deviation can be accurately estimated as the standard deviations of a large spectrum of  $x_j^R$  which can be easily

generated. The estimate is denoted  $\varphi^M(x_j^R)$ ; the superscript M underlines

that this quantity represents the error in  $x_j$  due to 'misorientation'.

The distribution of  $x_j^R$  has been investigated under various conditions. For values of  $\sigma^M$  which are not excessively large (to be discussed below), the following common results have been obtained:

1) The distribution of  $x_j^R$  is normal with mean  $x_j^B$  (Fig.1).

2) For each  $x_j$ , the value of  $\varphi^M(x_j^R)$  is proportional to  $\sigma^M$  so that

$$\varphi^M(x_j^R) = s(x_j) \cdot \sigma^M, \quad (5)$$

where  $s(x_j)$  is introduced to represent the sensitivity of a material constant  $x_j$  to angular errors (Fig.1).

3) For any material constant,  $s(x_j)$  decreases with the increasing number of crystal units used; consequently, the adverse effect of inaccurate orientation can be reduced by employing a larger number of crystal units (Table 2).

4) There exists a strong positive correlation between  $s(x_j)$  and  $\varphi^E(x_j^B)$ .

The correlation coefficient values found in the course of this work are 0.99, 0.98, 0.92, 0.97 and 1.00.

Consequently, in a given experiment, the sensitivity of  $x_j$  to the orientation errors is very nearly proportional to its error  $\varphi^E(x_j^B)$  due to experiment.

5) The contributions towards  $\varphi^M(x_j^R)$

from the first, second and third rotational angles  $\psi_k$ ,  $\phi_k$ , and  $\theta_k$  individually were obtained by randomizing all  $t$  values of one rotational angle, while keeping the rest equal to their nominal value. The contributions were denoted  $\varphi^\psi(x_j^R)$ ,  $\varphi^\phi(x_j^R)$  and  $\varphi^\theta(x_j^R)$ . For each of these standard errors the qualitative features 1) - 4) remain in effect.

For any  $\sigma^M$ , the three standard errors combine

$$\varphi^M(x_j^R) = \sqrt{\{[\varphi^\psi(x_j^R)]^2 + [\varphi^\phi(x_j^R)]^2 + [\varphi^\theta(x_j^R)]^2\}} \quad (6)$$

as is appropriate given that the three rotational angles should act as three independent sources of errors.

The comparison of  $\varphi^\psi(x_j^R)$ ,  $\varphi^\phi(x_j^R)$  and  $\varphi^\theta(x_j^R)$  does not generally point to any particular rotational angle as a

dominant source of error. However, in individual experiments, a hierarchy of one kind or another has been detected. Moreover,  $\vartheta^{\theta}(x_j^R) = 0$ ,  $j = 1, 2, \dots, n$ ,

for material constants studied by the resonance method using thickness modes of plates of infinite lateral dimensions.

## 5. TOTAL ERRORS OF MATERIAL CONSTANTS

The total standard errors of material constants are denoted  $\vartheta^T(x_j)$ . They result from the combined effect of errors due to experiment and misorientation. For  $\sigma^M \leq \sigma_1^M$ , one can write

$$\vartheta^T(x_j) \cong \left\{ [\vartheta^E(x_j^B)]^2 + [\vartheta^M(x_j^R)]^2 \right\}^{1/2}. \quad (7)$$

This is so because Eq. (4) guarantees that  $\vartheta^E(x_j^R)$  and  $\vartheta^M(x_j^R)$  are standard deviations of two nearly independent distributions. Eqs. (7) and (5) can be combined

$$\vartheta^T(x_j) \cong \left\{ [\vartheta^E(x_j^B)]^2 + [s(x_j) \cdot \sigma^M]^2 \right\}^{1/2} \quad (8)$$

and  $\vartheta^T(x_j)$  is obtained as a function of  $\sigma^M$ . The dependence of  $\vartheta^T(x_j)$  on  $\sigma^M$  is illustrated graphically in Fig. 3.

Eq. (8) shows the contribution of the orientation errors to the overall errors. Trying to determine  $\sigma^M$  such that the contribution is not too large, it must hold

$$(\vartheta^T(x_j) - \vartheta^E(x_j^B)) / \vartheta^E(x_j^B) \leq k_2, \quad (9)$$

where  $k_2$  is a chosen small number  $> 0$ .

Combining (8) and (9), we find Eq. (9) satisfied for all

$$\sigma^M \leq \sigma_2^M, \quad (10)$$

where

$$\sigma_2^M = \{ [(1+k_2)^2 - 1] \}^{1/2} \cdot \vartheta^E(x_j^B) / s(x_j). \quad (11)$$

The validity of Eqs. (7) and (8) has been tested numerically and under a variety of conditions. They were found to produce reliable results to within less than 5 percent of the actual values of  $\vartheta^T(x_j)$  determined by

simulations. With this reservation the interval  $x_j^B - \vartheta^T(x_j)$  and  $x_j^B + \vartheta^T(x_j)$ , where  $\vartheta^T(x_j)$  is defined by (7) or (8) represent  $\approx 68\%$  confidence intervals for the true value of the sought material constants  $x_j$ . This is because both  $\vartheta^E(x_j^B)$  and  $\vartheta^M(x_j^R)$  are standard deviations of two nearly independent normal distributions.

The results in this section depend on the results of Section 4 which in turn hold for values of  $\sigma^M$  which are not 'excessively large'. This limit is not fixed and depends on the numerical properties of each investigated case. For the experiments analyzed in this paper (Section 7) this limit was never reached.

## 6. TOLERABLE ORIENTATION ERRORS

Defining the tolerable orientation errors as such that they 1) do not invalidate the calculation of the errors  $\vartheta^E(x_j^B)$  due to experiment, and 2) do not substantially add to their values, the tolerable orientation error  $\sigma^M$  must satisfy the following condition

$$\sigma^M \leq \sigma_{\text{TOL}}^M = \min(\sigma_1^M, \sigma_2^M). \quad (12)$$

## 7. ANALYZED EXPERIMENTS AND NUMERICAL ILLUSTRATIONS

Three experiments, referred to as Experiment 1, 2 and 3, have been analyzed in detail to provide a basis for generalizations made in Section 4 as well as numerical illustrations needed throughout the entire paper. The original purpose of the experiments was to determine the third-order electro-mechanical constants of  $\alpha$ -quartz. Their choice for this analysis was influenced by the availability of all data needed here.

Experiment 1 consisted of 50 observations (Hruska (4)) taken on 50 rods vibrating in length to determine 5 material constants (Hruska (5)).

Experiment 2 consisted of 134 observations taken on 56 plates vibrating in thickness (4). Originally, the experiment was used to determine 10

material constants (Hruska and Brendel (6)). Its true potential is to determine 11 material constants. It is for these 11 constants that the illustrations in this paper are made.

Experiment 3 consisted of 24 observations (Kittinger et al (7)) of the transit time of acoustic pulses taken on cube-shaped crystal units of two different orientations to determine 17 material constants (Hruska and Ng (8)). As details concerning the actual number of physically different units of the two orientations have not been given in (7), the calculations made here proceeded as if only two physically different crystal units were used. As a result, only six different values of orientation angles,  $\psi_1, \phi_1,$

$\theta_1, \psi_2, \phi_2, \theta_2,$  have been treated as possible sources of errors.

Experiments 1 and 2 were analyzed also for a reduced number of crystal units, 24 in each case. Each of these units produced one useful observation. The units were selected to minimize experimental errors in the calculated material constants. This was done in order to show that, in contrast with this, the errors due to misorientation exhibit the opposite trend and also to obtain to most stringent demand on the tolerable orientation errors.

The results of the analysis of the 5 cases are shown in Table 2 (sensitivity  $s(x_j)$ ); Tables 3, 4 and 5 show  $\sigma_1^M$  and  $\sigma_2^M$  and  $\sigma_{TOL}^M$ , respectively. The results in Tables 3-5 are obtained for  $p = 0.99, k_1 = 0.05$  and  $k_2 = 0.01$ . This choice is arbitrary and made for illustrative purposes only. Table 6 shows the size of standard errors obtained by simulation for  $\sigma^M$  outside the domain of the tolerable orientation errors.

## 8. CONCLUSION

The evaluation of the reliability of material constants and of conflicts among them require their standard errors. The errors  $\varphi^E(x_j^B)$  traditionally obtained by the least-squares process do not generally fulfill this role.

$\varphi^E(x_j^B)$  may be seriously distorted by orientation errors. The distortion cannot be avoided completely but it can be reduced to acceptable limits by making sure that the standard error of no rotational angle exceeds  $\sigma_1^M$ .

$\varphi^E(x_j^B)$  reflect only the experimental errors. The misorientation of crystal units is an independent source of errors  $\varphi^M(x_j^R)$ . To obtain valid errors of the material constants,  $\varphi^E(x_j^B)$  and  $\varphi^M(x_j^R)$  must be combined. In order to make the contribution of the orientation errors minimal, the standard error of the rotational angles should not exceed  $\sigma_2^M$ .

If the standard errors of all rotational angles do not exceed  $\sigma_{TOL}^M$  which is the lesser of  $\sigma_1^M$  and  $\sigma_2^M$ , then  $\varphi^E(x_j^B)$  represent the total errors of the material constants reasonably well. It is in this sense that this paper introduces the 'tolerable' standard orientation error  $\sigma_{TOL}^M$ .

Examples of the  $\sigma_1^M$  and  $\sigma_2^M$  given in this paper show that the demands on the orientation accuracy can be more stringent than might be expected. For different experiments their values are different. One cannot generally say which of the two is the smaller and more restrictive. Predictably, with an increasing number of crystal units used in an experiment, both  $\sigma_1^M$  and  $\sigma_2^M$

increase and the orientation accuracy requirements become more relaxed.

The determination of  $\sigma_{TOL}^M$  is a numerical problem. It needs to be solved for each experiment individually. To be of practical use as a requirement on orientation accuracy,  $\sigma_{TOL}^M$  must be known when crystal units are being prepared. To be calculable at this stage, prior knowledge of the size of experimental errors and the approximate values of the material constants to be determined must be available. As such, the calculation of  $\sigma_{TOL}^M$  can be done only after a preliminary round of experimentation; it should be viewed as part of the

optimization process of the determination of material constants.

#### ACKNOWLEDGEMENTS

Pauline Hruska helped with the final stages of the preparation of this paper. This work was supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

#### REFERENCES

1. Hruska C K, 1994, "Sources of errors in nonlinear constants of crystals and their assessment", J.Appl.Phys. **75**, 77-83.
2. IEEE Standard on Piezoelectricity, 1978, "IEEE Std 176-1978". Available from The Institute of Electrical and Electronics Engineers, Inc., 345 East 47th Street, New York, NY 10017.
3. Draper N R and Smith H, 1966, "Applied regression analysis", J.Wiley, New York, 59-63 and 77-81.
4. Hruska C K, 1990, "Determination of the complete tensor of nonlinear piezoelectricity of quartz using the resonator method", 4th EFTF, Besancon, 295-300. Available from ENSMM, Univeriste de Besancon, France.
5. Hruska C K, 1990, "Linear combinations of the third-order elastic and piezoelectric constants", IEEE-UFFC, **37**, 329-331.
6. Hruska C K and Brendel R, 1990, "Combinations of the electroelastic and electrostrictive constants of quartz determined using thickness modes of plates", J.Appl.Phys., **67**, 1676-1679.
7. Kittinger E, Tichy J, and Friedel W, 1986, "Nonlinear piezoelectricity and electrostriction of alpha quartz", J.Appl.Phys. **60**, 1465-1471.
8. Hruska C K and Ng P, 1993, "Material nonlinearities in quartz determined by the transit-time method using dc field interactions", J.Acoust.Soc.Am., **93**, 1426-1430.

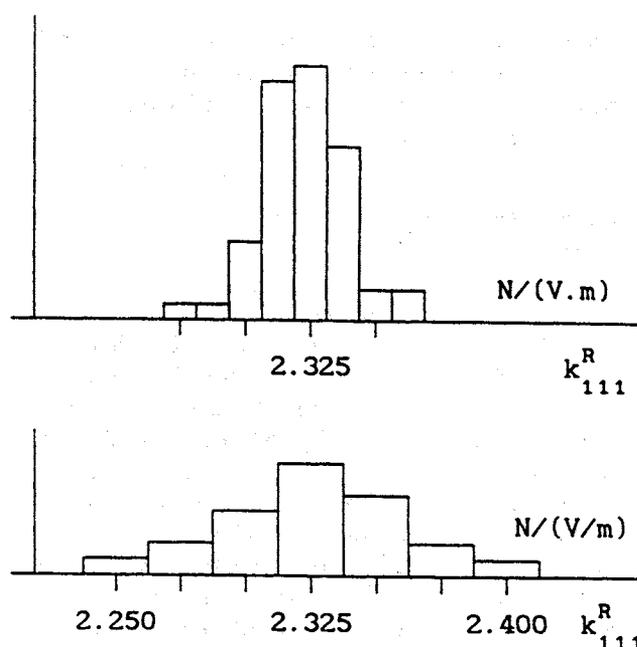


Figure 1: Typical distribution pattern of a material constant resulting from the randomization of the three rotational angles of all crystal units used for its calculation. Obtained for material constant  $k_{111}^R$  based on data by Kittinger et al (7).  $k_{111}^R$  is a linear combination of four third-order elastic and electroelastic constants according to Hruska and Ng (8). The histograms are for  $\sigma^M = 10.5'$  (top) and  $21'$ . The pattern suggest a normal distribution of  $k_{111}^R$  centered at  $k_{111}^B = 2.323$  N/(V.m) (8); standard errors  $\phi^M(k_{111}^R)$  represent errors of  $k_{111}^R$  due to misorientation. Their values 0.0166 (top) and 0.0331 N/(V.m) are proportional to  $\sigma^M$  and illustrate the linear relationship between  $\phi^M(k_{111}^R)$  and  $\sigma^M$ .

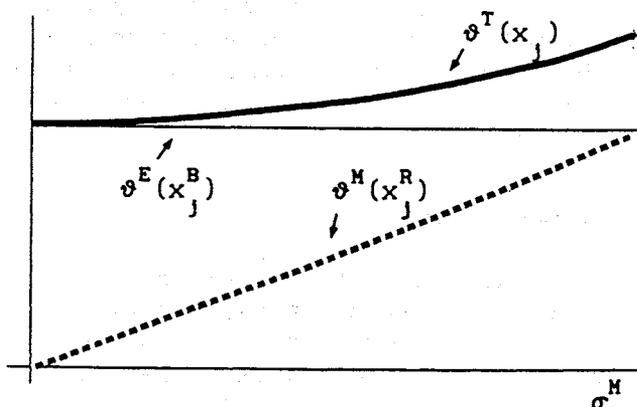


Figure 2: Total error of material

constants. For any material constant  $x_j$ , standard errors due to experiment  $\phi^E(x_j^B)$  and misorientation  $\phi^M(x_j^R)$  combine to produce the total standard error  $\phi^T(x_j)$ . For  $\sigma^M \leq \sigma_1^M$ ,  $\phi^E(x_j^B)$  is nearly constant (Section 3). For  $\sigma^M$  not excessively large,  $\phi^M(x_j^R)$  grows linearly with  $\sigma^M$  (Eq. (5) and  $\phi^T(x_j)$  depends on  $\sigma^M$  according to Eq. (8).

TABLE 1 - Effect of misorientation on the values of material constants and their standard errors.

material constant	$x_j^B \pm \phi^E(x_j^B)$	$x_j^R \pm \phi^E(x_j^R)$
$k_{124}$	$-0.24 \pm 0.08$	$-0.44 \pm 0.09$
$k_{33}$	$-9.08 \pm 7.87$	$-17.76 \pm 8.85$
$q_{111} \cdot 10^{20}$	$-0.24 \pm 0.26$	$0.22 \pm 0.30$

Values in the left column are obtained by the least-squares fit using the nominal rotational angles of 56  $\alpha$ -quartz plate units listed in Tab I, parts b-d in Hruska (4). The right column is for the nominal angles changed by  $-30'$ .  $k_{124}$ ,  $k_{33}$  and  $q_{111}$  are linear combinations of the third-order electromechanical constants of  $\alpha$ -quartz with the respective dimensions of  $N/(V.m)$ , 1, and  $F/V$ .

TABLE 2 - Average sensitivity  $s(x_j)$  for various material constants calculable from experiments 1, 2 and 3.

number of units	$s(x_j)$	experiment		
		1	2	3
full	$s(k_{ijk})$	0.001	0.002	0.001
	$s(k_{ij})$	-	0.083	0.124
	$s(q_{111})$	-	0.009	0.104
24	$s(k_{ijk})$	0.003	0.005	-
	$s(k_{ij})$	-	0.282	-
	$s(q_{111})$	-	0.021	-

For any material constant  $x_j$  and  $\sigma^M$  not excessively large,  $\phi^M(x_j^R) = s(x_j) \cdot \sigma^M$ .  $\phi^M(x_j^R)$  is the standard error of the material constant due to misorientation.  $\sigma^M$  is the common standard deviation of the random errors of all rotational angles used to calculate  $x_j$ .  $k_{ijk}$ ,  $k_{ij}$  and  $q_{111}$  are combinations of the third-order electromechanical constants of  $\alpha$ -quartz calculable from data listed in Hruska (4) and Kittinger et al (7).  $s(k_{ijk})$ ,  $s(k_{ij})$  and  $s(q_{111})$  are, respectively, in  $[N/(V.m)]/\text{angular minute}$ ,  $1/\text{angular minute}$ , and  $(F/V)/\text{angular minute}$ . The full number of crystal units used in experiments 1, 2 and 3 is, respectively, 50, 56 and 2. The increase in  $s(x_j)$  as the number of the crystal units is reduced to 24 is significant.

TABLE 3 - Values of  $\sigma_1^M$ .

number of crystal units used	experiment		
	1	2	3
full	9'	9'	2'
24	1'	8"	-

Errors in the rotational angles generally distort the calculated values of experimental errors  $\phi^E(x_j^B)$  and render them invalid. To keep the distortion within 5% bounds, the standard deviation of orientation errors  $\sigma^M$  must satisfy the condition  $\sigma^M \leq \sigma_1^M$ . The full number of crystal units used in experiments 1, 2 and 3 is, respectively, 50, 56 and 2.  $\sigma_1^M$  decreases as the number of crystal units is reduced to 24.

TABLE 4 - Values of  $\sigma_2^M$ .

number of crystal units used	experiment		
	1	2	3
full	11'	4'	4'
24	3'	1'	-

Errors in material constants due to misorientation  $\phi^M(x_j^R)$  add to the experimental errors  $\phi^E(x_j^B)$  to produce the total error  $\phi^T(x_j)$ . To keep the increase under 1% of  $\phi^E(x_j^B)$ , the standard deviation of orientation errors  $\sigma^M$  must satisfy the condition  $\sigma^M \leq \sigma_2^M$ . The full number of crystal units used in experiments 1, 2 and 3 is, respectively, 50, 56 and 2.  $\sigma_2^M$  decreases as the number of crystal units is reduced to 24.

TABLE 5 - Tolerable orientation errors.

number of crystal units used	experiment		
	1	2	3
full	9'	4'	2'
24	1'	8"	-

Tolerable orientation errors  $\sigma_{TOL}^M$  satisfy two conditions: (1) they do not invalidate the calculated errors  $\phi^E(x_j^B)$  due to experiment; and (2) they do not substantially add to the values of  $\phi^E(x_j^B)$  to produce unnecessarily large values of total errors  $\phi^T(x_j)$ . To achieve this, the standard deviation  $\sigma^M$  of all rotational angles must satisfy the condition  $\sigma^M \leq \sigma_{TOL}^M = \min(\sigma_1^M, \sigma_2^M)$ . The full number of crystal units used in experiments 1, 2 and 3 is, respectively, 50, 56 and 2.  $\sigma_{TOL}^M$  decreases as the number of crystal units is reduced to 24.

TABLE 6 - Typical errors of material constants for large orientation errors.

number of units	type of error	experiment		
		1	2	3
full	$\phi^E$	1.31%	8.56%	4.04%
	$\phi^M$	0.30%	6.83%	2.08%
	$\phi^T$	1.42%	11.9%	4.88%
24	$\phi^E$	0.63%	3.17%	-
	$\phi^M$	0.50%	17.2%	-
	$\phi^T$	1.13%	74.6%	-

Calculated for common standard orientation error  $\sigma^M = 21'$ .  $\phi^E$ ,  $\phi^M$  and  $\phi^T$  stand for average standard errors  $\phi^E(k_{ijk}^B)$ ,  $\phi^M(k_{ijk}^R)$  and  $\phi^T(k_{ijk})$ .

$k_{ijk}$  are combinations of third-order elastic and electroelastic constants of  $\alpha$ -quartz calculable from data (4,7).  $\phi^E$  are obtained by the least-squares fit,  $\phi^M$  and  $\phi^T$  by simulation. As  $\sigma^M > \sigma_1^M$ ,  $\phi^E$  no longer reliably represent the total standard errors due to experimental errors.  $\phi^M$  are the errors due to misorientation. The total standard errors  $\phi^T$  include the effect of experimental errors and misorientation. As  $\sigma^M$  is too large, Eq. (7) cannot be generally used to predict  $\phi^T$  from  $\phi^E$  and  $\phi^M$ . The full number of crystal units used in experiments 1, 2 and 3 is, respectively, 50, 56 and 2.

## DEFECTS IN QUARTZ RELATED TO INTERSTITIAL Li AND Na IONS AS REVEALED BY LUMINESCENCE

A. Halperin

**Racah Institute of Physics, The Hebrew University, Jerusalem 91904, Israel**

### INTRODUCTION

Quartz usually contains Al impurities substituted for Si in the lattice. Monovalent ions, e.g.  $\text{Li}^+$  and  $\text{Na}^+$ , provide then the charge compensation at the Al. These monovalent ions ( $\text{M}^+$ ) have been found to affect considerably the thermoluminescence (TL) of quartz, effects involving migration of  $\text{M}^+$  ions along the c-axis channels in quartz. Thus, the glow curve of Li containing quartz subjected to certain double irradiation procedures involving migration of Li ions away from the AL was shown to exhibit a very strong TL peak near 190K, as shown by Halperin (1) and Halperin et al. (2).

Radiation induced drift of  $\text{M}^+$  ions in quartz is also known to produce acoustic losses and frequency offsets in quartz resonators. (King and Sunder (3) and Martin (4)) The TL peaks related with the migration of the  $\text{M}^+$  ions give an insight into the radiation induced drift of the interstitials in quartz. It provides thus a good tool for the investigation of the related  $Q^{-1}$  losses and frequency offsets in quartz resonators. The present paper will review some earlier results on the  $\text{M}^+$  related defects in quartz and describe results of still unpublished work on the topic.

### EXPERIMENTAL

As this paper contains a review of several investigations there is no use to describe here the various experimental setups and procedures taken in the investigations. The reader is referred to the references cited below for details of the experimental work. In short, the samples were all cut from synthetic quartz crystals and mounted in cryostats fitted to the specific work. An Air Product Displex double stage refrigeration system was used at 10-370K, and a liquid nitrogen cooled metal cryostat served in the range 80 - 700K. The

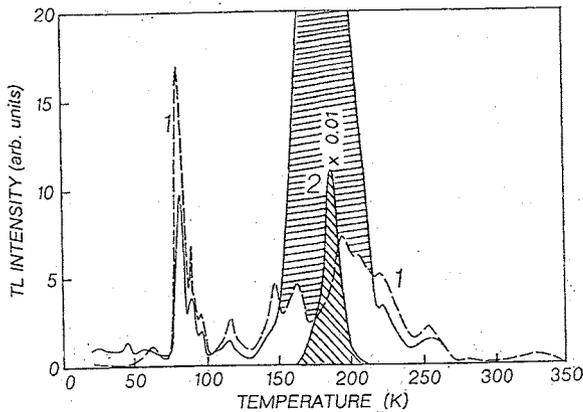
excitation was by x-rays. The emitted light in TL measurements was focussed on the photocathode of a suitable photomultiplier, amplified and recorded on a chart recorder. A monochromator was inserted in the optical path between the sample and the photomultiplier for spectral measurements. A heat absorbing glass was used to reduce the black body radiation from the sample holder when reaching temperatures above 600K. The warming rate in the TL measurements was generally  $10^\circ/\text{min}$ .

### $\text{M}^+$ ( $\text{Li}^+$ , $\text{Na}^+$ ) RELATED TL IN QUARTZ

#### TL Below 350 K In Li Containing Quartz

In conventional TL measurements the sample is excited by a single irradiation at the desired temperature, after which it is warmed up (at a constant heating rate) when the TL is emitted. Figure 1, Curve 1 shows such a so-called glow curve for a sample from a synthetic electronic grade quartz block (Sawyer), cut from the +X-growth zone of a Y-plate cut from a crystal block. It contained about 30 ppm Al per Si atom. Prior to the TL measurement it has been x-irradiated for 5 min at 10 K. The sample was then subjected to a double irradiation procedure consisting of one 5 min x-irradiation at 10K followed by warming to 250K and cooling back to 10K, when a second 5 min x-irradiation took place. The TL was then recorded during warming. Such a procedure of two irradiations with warming to a given intermediate temperature between the two irradiations will be referred to below as the Intermediate Warming procedure, or the IW procedure. Curve 2 in Figure 1 shows the resulting glow curve. Most of the TL peaks are now somewhat weaker compared to Curve 1. Unexpected was the extremely strong TL peak

near 190K. It was absent in Curve 1 and exceeds now all the other peaks in intensity by two orders of magnitude. This 190K peak can also be obtained by performing the 1st irradiation at 250K and the second at 10K. This Two Temperature Irradiation procedure will be referred to as the TTI procedure. The 190K TL peak can be formed (see Figure 4B below) by a first irradiation at temperatures over the whole 150 - 300K range by the TTI procedure. In the IW procedure the same temperature range applies for the intermediate warming temperature (1,2). The second irradiation temperature can be performed at any temperature up to about 180K(1).

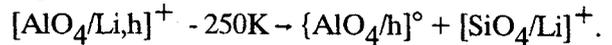


**Figure 1** - 1-Glow curve for quartz (5 min x-irradiation at 10K). 2 - after a double irradiation (10K, 4 min. each), with intermediate warming to 250K.

The appearance of the strong 190K peak was fully explained in a correlated ESR and TL study by Jani et al.(5), Halperin et al.(6) and (2). The electron trap related with the 190K peak was found to be the  $[\text{SiO}_4/\text{Li}]^+$  defect formed by the migration of the Li ion away from the Al. The scheme of the steps involved in the excitation and emission of the 190K peak is summed up below (for the IW procedure):

1. Positive holes provided during the 1st irradiation get captured at the Al-Li center:  
 $[\text{AlO}_4/\text{Li}]^0 + \text{h}^+ \rightarrow [\text{AlO}_4/\text{Li,h}]^+$
2. At 250K the now loosely bound  $\text{Li}^+$  ions at the Al-Li-h centers leave and migrate along the

c-axis channels where they get trapped near regular Si in the lattice:



3. The 2nd low temperature irradiation provides electrons and part of them get trapped at the Si-Li Traps:  $(\text{SiO}_4/\text{Li})^+ + \text{e}^- \rightarrow [\text{SiO}_4/\text{Li}]^0$ .

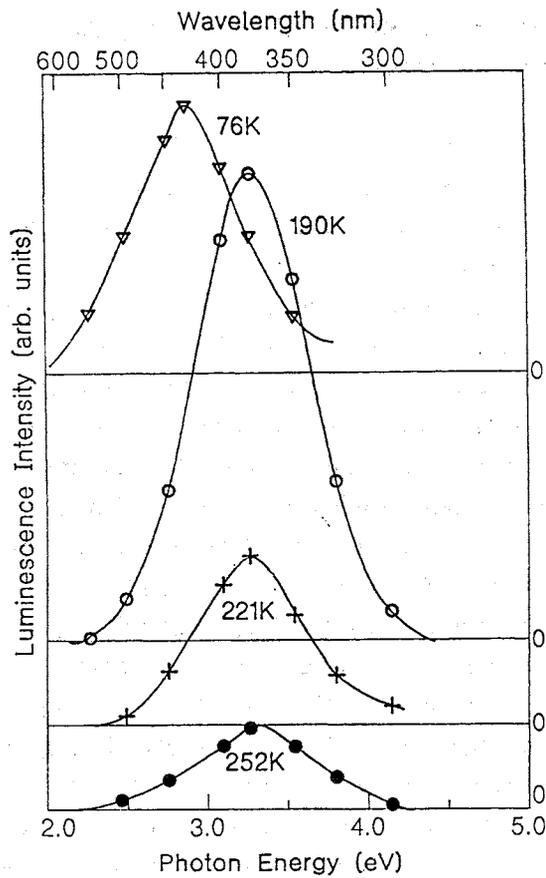
4. Subsequent warming of the sample releases the electrons from the Si-Li traps with the emission of the 190K TL peak by radiative recombination at the luminescence centers.

The observation of the ESR spectrum of the  $[\text{SiO}_4/\text{Li}]^0$  centers and the correlation between the ESR spectrum and the 190K TL peak provide strong support to the above scheme of the formation of the 190K TL peak (2,5,6).

In the TTI procedure the first irradiation takes place at a temperature high enough for the release of the Li ions from the Al-Li-h centers so that 1 and 2 combine into one step.

The 190K TL peak gives an emission band at 380nm (3.25 eV). It was shown by Katz and Halperin (7) that practically all the peaks in the range 120-260K emit the 380 nm band. This implies that all these TL peaks involve the release of electrons from various electron traps and their recombination at the same luminescence centers. Emission spectra for a few TL peaks are shown in Figure 2. Most peaks below 115K emit at longer wavelengths (7). For example, the 76K peak is shown in Figure 2 to emit at 430nm.

Sweeping of quartz in an electric field along the c-axis at temperatures of about 500°C replaces the monovalent ions in the crystal by other monovalent ions present at the negative electrode. Sweeping in a hydrogen containing atmosphere with clean electrodes replaces Li and Na ions by hydrogen. Hydrogen sweeping was found to eliminate completely all the TL peaks emitting the 380nm band, including the 190K peak (2). This suggests that like the 190K peak the other electron traps related to the above TL peaks contain Li. It should be noted that the later electron traps do not need any excitation for their formation in contrast with the 190K TL related  $[\text{SiO}_4/\text{Li}]^+$  electron trap which is formed by the irradiation.

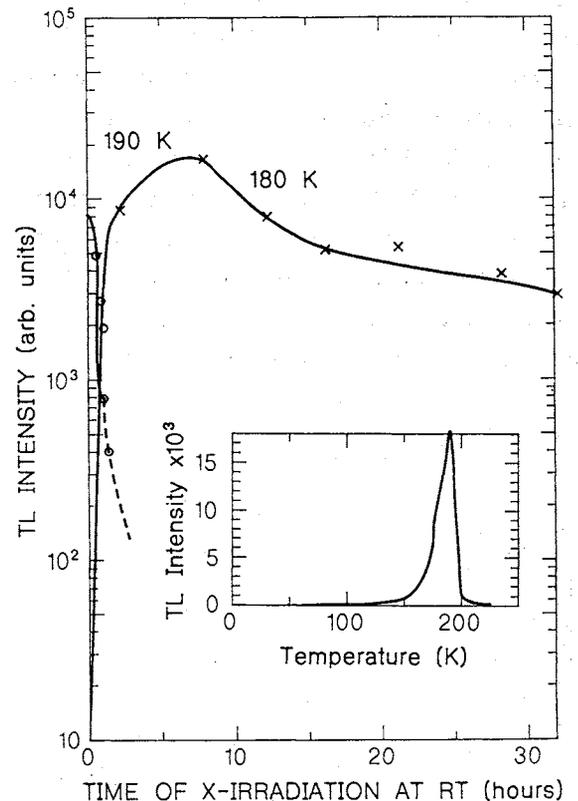


**Figure 2.** - Emission spectra for a few TL peaks of quartz

#### The radiation produced 180K TL peak.

Halperin and Katz (8) have shown that prolonged irradiation near room temperature changes the glow curves of quartz crystals. The following description will be limited to the replacement of the 190K peak by one at about 180K. Figure 3 gives the intensities of the 190 and 180K peaks as function of the accumulated preirradiation time at room temperature. The 190K peak is seen to drop by more than an order of magnitude after 1 hour of preirradiation. The 180K peak then emerges and rises fast with radiation time. It reaches a maximum after 7 hours of preirradiation after which its intensity decreases with irradiation time. The inset in Figure 3 gives the glow curve obtained after 7 hours of preirradiation at RT. It shows practically only the 180K peak. This 180K peak is not formed in hydrogen swept samples and emits the 380nm band. Its optimal intensity is almost the same as that of the original 190K peak. The related trap must therefore contain a Li ion at

some defect site in the crystal. The equal intensities of the 180 and 190K peaks suggest that the defect related with the 180K TL is also a Li ion at a regular Si but at a site with a higher binding energy to the Si. Irradiation at temperatures in the range 250-400K seems to transfer the Li ions from the sites related with the 190K TL peak to the more stable sites related with the 180K peak. Warming of the preirradiated sample to above 450K eliminates the 180K TL peak when the 190K peak is restored (8).

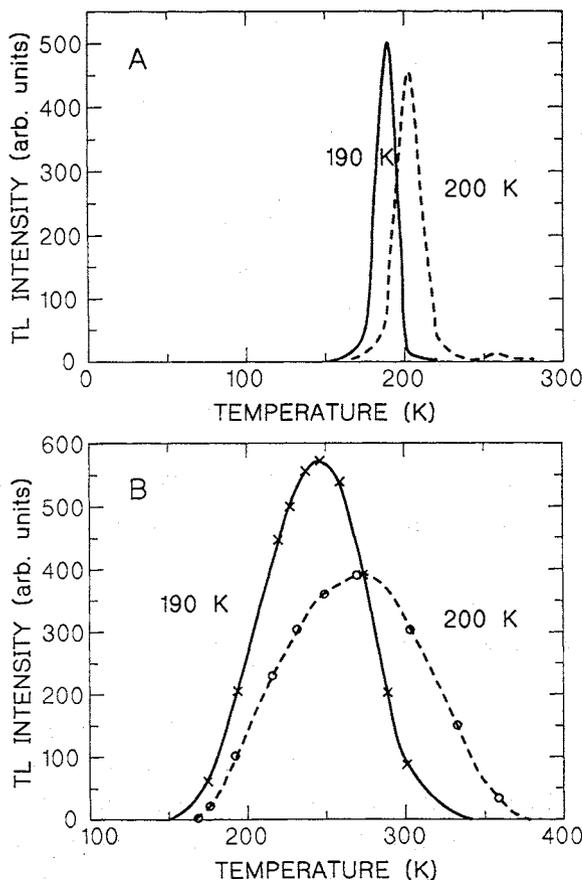


**Figure 3.** - The decay of the 100K TL peak and the growth of the 180K peak as a function of the preirradiation time.

#### The $[\text{SiO}_4/\text{Na}]$ Related TL Peak.

Some samples from the same electronic grade quartz block of the Li containing samples were subjected to electric sweeping in which the Li was completely replaced by Na. The conventional glow curves obtained for the Na-swept samples were not much different from those obtained with the Li containing ones. In the following only the Na related TL peak analogous to the 190K peak will

be dealt with. The peak appeared at 202K and will be referred to as the 200K peak. It was formed by the same double irradiation procedures described for Li above. The optimal intensities obtained for the 200K peak were almost the same as those of the 190K peak as shown in Figure 4A where both glow curves were obtained by the IW procedure. Figure 4B gives the formation curves for the 190 and 200K peaks, obtained by the IW procedure. In this case the abscissa gives the temperature reached in the intermediate warming between the two low temperature irradiations. The 200K Na peak emitted the 380nm just as the 190K Li peak. It was also eliminated by hydrogen sweeping. These characteristics support strongly the model in which the  $[\text{SiO}_4\text{Na}]^+$  center is the electron trap related to the 200K peak. Unfortunately attempts to reveal the ESR spectrum of the  $[\text{SiO}_4\text{Na}]^0$  centers were not successful until now. For further details of the characteristics of the 200K TL peak the reader is referred to Halperin and Katz (9,10).



**Figure 4.** - A. The 190K Li related TL peak and the 200K Na related one. B. Formation curves for the 190 and 200K TL peaks.

### Single Irradiation Excitation of the $[\text{SiO}_4/\text{M}]$ Related TL Peaks

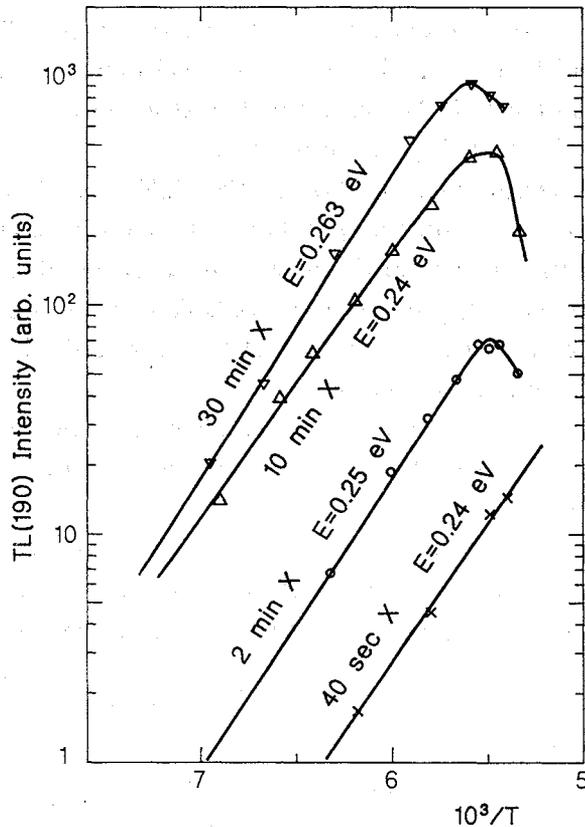
The formation of the  $[\text{SiO}_4/\text{M}]^0$  centers was described above. In early publications (1,2) it has been claimed that a double irradiation procedure was necessary for the formation of these TL peaks. This categorical statement proved not to be accurate. Halperin (11) and Halperin and Sucof (12) have shown that though at lower intensities compared with those obtained by a double irradiation and at a limited range of temperatures a single irradiation excitation is possible. This can be achieved at a temperature at which the  $\text{M}^+$  ions can leave the Al-M-h centers and up to a temperature at which the TL peak is still relatively stable. For the 190K peak the lower temperature limit is above 150K (Figure 4B) and the upper limit is about 180K. For the Na related peak the temperature range will be about 170-190K.

The single irradiation excitation of these TL peaks provides (11,12) a convenient method for the determination of the thermal binding energies of the  $\text{M}^+$  ions to the Al-M-h centers. Figure 5 gives a few curves obtained, each with a different excitation dose (irradiation time) for the 190K TL. The intensities are plotted as a function of the reciprocal excitation temperatures. The slopes of the curves give an activation energy of 0.25 eV compared to 0.27 eV obtained from the Li ion mobility by Hughes (13). The binding energy for the 200K Na peak was found to be 0.31 eV (12) which compares very well with the 0.3 eV reported by King and Sunder (14).

### The Release of the M Ions From the Al Centers

The scheme for the production of the 190K TL peak by the IW procedure given above shows that during the first low temperature irradiation the  $[\text{AlO}_4/\text{Li}]^0$  centers trap holes thus forming  $[\text{AlO}_4/\text{Li,h}]^+$  centers. The Li ions remain still at the Al sites (binding energy 0.25 eV) until the sample is warmed up to above 150K. The thermal energy then overcomes the 0.25 eV binding energy when the Li is released from the Al site. In the TTI procedure the first irradiation takes place

above 150K when the Li ion is released immediately after the trapping of the hole. For Na containing samples the process is the same except that the higher binding energy of the Na to the Al-Na-h center (0.31 eV) results in a formation curve shifted to higher temperatures compared to that for Li (Figure 4B).



**Figure 5.** - Single excitation curves of the 190K TL peak as a function of  $10^3/T$ , obtained for various excitation times. Slopes give the binding energy of the Li at the Al centers.

The above model differs from those reported in literature. Hitt and Martin (15) adopted a model very common in literature by which ionizing irradiation above 200K removes the alkali ions from the Al sites. Trapping of the hole occurs by this model after the removal of the alkali ion. Hughes (13) realized that irradiation at about 200K can not overcome the very strong Coulomb binding energy between the Al and the alkali ion. He suggested therefore that holes created during the irradiation diffuse into the Al-M centers where they get deeply trapped and neutralize the site, when the M ion is free to drift away.

The present work shows clearly that both the above models are wrong. Holes get trapped even at the lowest temperatures when the hole mobility is negligibly small. The release of the M ions by warming is of course purely thermal. The 0.27 eV energy obtained for Li in (13) is therefore in fact the thermal binding energy as obtained in the present work and not hole mobility as claimed in (13). The method for the measurement of the binding energies offered in the present work is very simple and more accurate compared to the indirect method of "hole mobility" given in (13).

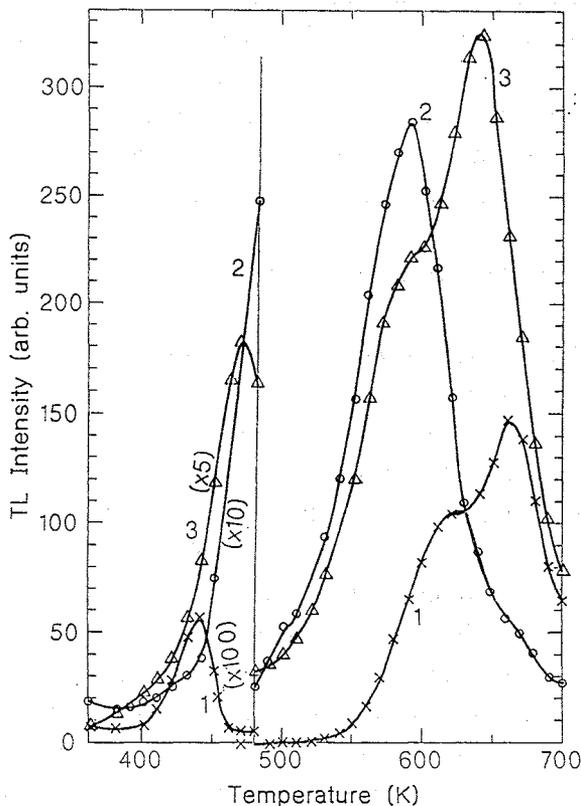
### The 350-700K TL In Quartz

This paragraph will concentrate on the role of the M ions in the traps related with the TL peaks in the range 450-700K (the deep traps). Jani et al (16) have observed in high quality synthetic quartz intense TL peaks in the range 180-300°C. Other investigators, for example Bernhardt (17), report on TL peaks in the range 600-700K and even above 700K. A close relation has been reported to exist between the low temperature TL peaks and those at high temperatures (17). Of interest is the observation that all the deep-trap-related TL peaks do not appear in hydrogen swept samples. This leads to the conclusion that they are alkali associated electron traps (16).

The replacement of the 190K Li associated TL peak at room temperature by a peak at 180K (see above) and the restoration of the 190K peak by warming take place in the temperature range 400-600K (8). This points to a relation between the shallow traps related with the 180 and 190K TL peaks and the deep traps.

The TL peaks at high temperatures change with the dose of the x-irradiation. Figure 6 shows three glow curves obtained for the same sample (+x-zone, electronic grade). Curve 1 was obtained after 1 min of x-irradiation by x-rays at 350K. It shows the main peak at 660K. After 32min irradiation at 350K (Curve 2) only traces of the 660K peak are seen and a peak at about 590K is dominating. Curve 3 gives the glow obtained after 5 hours of irradiation at 350K. The 590K peak is now weaker and is being replaced by one near 640K.

The TL peaks below 480K shown in Figure 6 are much weaker and are given on more sensitive ordinate scales. A TL peak at 440K appears at low doses (1 min irradiation, Curve 1). In Curve 3 (5 hours irradiation) it is seen to be replaced by a peak near 460K.



**Figure 6.** - Glow curves for a sample after 1, 32 and 300 min x-irradiation at 350K (curves 1, 2 and 3 respectively).

The total TL intensity in the range 500-700K, (the area under the glow curve in this temperature range) was found to remain nearly constant throughout the range of 3-300 min irradiation times. More than that, the TL peaks at 660, 590 and 640K reached each under its optimal irradiation dose nearly the same intensity. This suggests equal concentrations of the electron traps related to each of the three peaks. At short irradiation times most of the available alkali ions released from the Al sites get trapped at a given site in the quartz lattice. At higher doses the alkali ions are gradually transferred to other sites in the same defect, which results in the replacement of one of the above TL peaks by another. This is

similar to the replacement of the 190K peak by the 180K peak shown above. The exact nature of the deep electron traps is yet unknown. Work is going on now with the hope that it will lead to an identification of the related defects.

## CONCLUSIONS

1. The procedures for the formation of the 190 and 200K Li and Na related TL peaks involve the migration of alkali ions from the Al and formation of the  $[\text{SiO}_4/\text{M}]^+$  electron traps.
2. During preirradiation at 250-400K, alkali ions get trapped more deeply at other sites, presumably at the same Si in the quartz lattice. Some alkali ions get trapped at still more stable sites, forming the deep electron traps related with the high temperature TL peaks.
3. In all the above cases the alkali ions migrate back to the Al sites during the emission of the related TL.
4. Irradiation itself can not release the M ions from the  $[\text{AlO}_4/\text{M}]^0$  centers. The release takes place only after the formation of  $[\text{AlO}_4/\text{M},\text{h}]^+$  centers by the trapping of a hole when the M will leave the Al center at temperatures near or above 200K.
5. Research on the radiation induced  $Q^{-1}$ -losses and frequency offsets of quartz resonators in correlation with the above TL peaks is believed to lead to a better understanding of the effects of radiation on the resonators which may help in the elimination of the unwanted effects.

## REFERENCES

1. Halperin A. and Katz S., 1984, "X-ray transferred TL in quartz crystals", *J. Lumin.* **31-32**, 129-132.
2. Halperin A., Jani M.G. and Halliburton, L.E., 1986, "Correlated ESR and TSL study of the  $[\text{SiO}_4/\text{Li}]$  center in quartz", *Phys. Rev.* **B34**, 5702-07.
3. King J.S. and Sunder H.H., 1975, "Transient changes in quartz resonators following exposure to pulse ionization", *Radiation Effects* **26**, 203-212
4. Martin J.J., 1990, "Radiation induced frequency

- offsets and acoustic loss in AT cut quartz crystals", J. Appl. Phys. **68**, 5095-5104.
5. Jani M.G., Halliburton L.E. and Halperin A., Observation of a simple lithium associated electron trap in crystalline SiO<sub>2</sub>, Phys. Rev. Lett. **56**, 1392-94.
  6. Halperin A., Jani M.G. and Halliburton L.E., The defect involved in the emission of the TL peaks at 190K in quartz", Cryst. Lett. Def. and Amorph. Mat. **16**, 7-12.
  7. Katz S. and Halperin A., 1988, "The low temperature phosphorescence and TL of quartz crystals", J. Lumin. **39**, 137-143.
  8. Halperin A. and Katz S., 1988, "Irradiation effects on the low temperature TL of quartz crystals", J. Phys. Chem. Solids **49**, 577-583.
  9. Halperin A. and Katz S., 1987, "The TL related with the [SiO<sub>4</sub>/Na]<sup>o</sup> center in quartz", Sol. State Commun. **63**, 697-99.
  10. Halperin A. and Katz S., 1988, "The TL of Li and Na swept quartz crystals", Proc. of 42nd Freq. Control Symp. 184-88.
  11. Halperin A., 1989, "Single irradiation excitation of the TL related to the [SiO<sub>4</sub>/Li]<sup>o</sup> center in quartz", J. Phys. Chem. Sol. **52**, 1039-40
  12. Halperin A. and Sucov E.W., 1991, "Single irradiation excitation of the TL related to the [SiO<sub>4</sub>/Li] center in quartz", J. Phys. Chem. sol. **52**, 1039-40.
  13. Hughes R.C., 1975, Electronic and ionic charge carriers in irradiated single crystals and fused quartz", Rad. Effects **26**, 225-235.
  14. King, J.C. and Sunder, H.H., "Transient charges in quartz resonators following exposure to pulse ionization", Rad. Effects **26**, 203-212.
  15. Hitt K.B. and Martin J.J., 1983, "Radiation-induced mobility of lithium and sodium in alpha-quartz", J. App. Phys. **54**, 5030-31, and references therein.
  16. Jani M.G., Halliburton L.E. and Kohnke E.E., 1983, "Point defects in crystalline SiO<sub>2</sub>: thermally stimulated luminescence above room temperature", J. App. Phys. **54**, 6321-28.
  17. Bernhardt H.J., 1982, "A contribution to high temperature TL investigation in quartz crystals", Phys. Stat. Sol. (a) **74**, 159-168, also Bernhardt H.J., 1984, "Investigations on low temperature TL centers in quartz", Crystal Research & Technol. **19**, 133-1143.

# COMPOSITE RESONATOR SPECTROSCOPY OF QUARTZ FERROBIELASTIC TWINNING UNDER UNIAXIAL STRESS

P. L. GUZZO, R.J. BESSON

Ecole Nationale Supérieure de Mécanique et des Microtechniques  
26 Chemin de l'Épitaphe 25030 Besançon cedex - France

G. D. MANSFELD

Institute of Radioengineering and Electronics, Russian Academy of Sciences  
11, Mokhovaya Str. 103907 Moscow - Russia

## Abstract

In the present work, the acoustic waves in X-cut quartz crystals subjected to uniaxial stress were used as a tool for investigation the ferrobielastic twinning phenomenon. The propagation characteristics of media was measured using the previously developed composite resonator spectroscopy. Al-ZnO-Al layers were deposited on one of the X-faces of quartz samples forming composite resonators structures. It was observed that the position of peaks for longitudinal waves increased with the increase of the compressive stress. The ferrobielastic switching was characterized by an abrupt decrease of frequency of peaks. Subsequent increase of pressure resulted in the smaller slope of frequency vs. stress dependence than before twinning. The quantitative and qualitative explanation of the observed data and the evaluation of the sound velocity are given.

## 1. Introduction

Under high uniaxial pressure crystalline lattice of solids may be distorted or rearranged and therefore proper material constants be noticeably changed. It is known that the X-axis of quartz crystals can even be reversed by applying uniaxial stress at favorable directions and this is known as ferrobielastic twinning.

In the light of the free energy formulation, the ferrobielastic twinning is classified as a secondary ferroic phenomenon [1]. The ferrobielastic twinning is induced when a suitable stress is applied along a crystallographic direction that favors the switching from the existing domain to the alternate domain [1,2]. The symmetry operation relating quartz ferrobielastic domains (the two Dauphiné twin states) is a 180° rotation about [001]. Consequently, the signs of the elastic compliance  $s_{14}$ , the piezoelectric coefficient  $d_{11}$  and the piezo-optical coefficients  $\pi_{14}$  and  $\pi_{41}$  are reversed between twin states [2]. Besides strain-stress relationships, these changes on tensor properties provide additional means for detection of stress induced twinning. The propagation of the twin walls was optically observed with the sample positioned between crossed polarizers [1, 3-5]. The

piezoelectric polarization charge along [100] was measured as a function of the applied stress [6, 7] and the acoustic emission generated during the ferrobielastic twinning was recorded [6].

In the present work, we describe a new method to investigate the ferrobielastic twinning of quartz using the composite resonator spectroscopy. The propagation of acoustic waves was provided by the composite resonator structure [8] having X-cut quartz crystals as substrate. The resonant peculiarities of composite resonator structure contain informations about acoustic wave velocity and attenuation [9]. The position of resonant peaks in the frequency range of hundreds MHz as a function of the compressive stress was measured before and after twinning. Our experimental data of sound velocity were compared to those given by the formula proposed by Thruston and Brugger [10]. The changes of resonance peaks positions in the frequency domain contain informations about lattice transformation and sound velocity at different stress levels.

## 2. Evaluation of sound velocity

From the theory of finite deformations certain functions of sound velocity can be derived and expressed in terms of the second and third-order elastic constants. An expression for the isothermal stress derivatives of  $\rho_0 w^2$  for any homogeneous stress system was previously derived [10]. In the case of an uniaxial stress we have :

$$\left(\rho_0 w^2\right)'_{P=0} = -(\mathbf{N} \cdot \mathbf{M})^2 - 2c_0 F - G \quad (1)$$

where  $w$  is so called "natural" wave velocity referred to undeformed dimensions of the sample and  $\rho_0$  is the density at zero stress.  $\mathbf{M}$  and  $\mathbf{N}$  are unit vectors along the directions of stress and the propagation of acoustic waves in the absence of a static stress, respectively.  $c_0$  is related to the actual sound velocity ( $V$ ) at zero stress and is given by :

$$c_0 = \left( \rho_0 v^2 \right)_{P=0} = c_{prqs}^s N_p N_q U_r U_s \quad (2)$$

where  $c_{prqs}^s$  are the second order adiabatic stiffness coefficients and  $U$  is the unit vector along the direction of the particle displacement referred to undeformed dimensions. The summation over all repeated indices is implied. The parameter  $F$  is the stress derivative of the Lagrangian strains and is associated to the geometrical change of the specimen due to the uniaxial stress.  $F$  is defined by the following expression :

$$F = s_{abrs}^T M_a M_b U_r U_s \quad (3)$$

where  $s_{abrs}^T$  are the second-order isothermal compliance coefficients. The parameter  $G$  is associated to the dependence of elastic stiffness with the uniaxial stress and is defined as :

$$G = s_{abuv}^T C_{uvprqs} M_a M_b N_p N_q U_r U_s \quad (4)$$

where  $C_{uvprqs}$ , the third-order stiffness coefficients, are the isothermal strain derivatives of the adiabatic second order stiffness coefficients. Since all quantities entering into equation (1) refer to the unstressed state, neither the deformation nor changes in crystal symmetry under stress have to be considered.

We are interested in the variation of the sound velocity in relation to the uniaxial stress. The dependence of the relative change of velocity with the uniaxial stress is obtained by solving equation (1). Considering the first and the second terms of the expansion  $\sqrt{1+a} = 1 + \frac{1}{2}a - \dots$  ( $a = \text{constant}$ ), the relative change of natural velocity ( $\Delta w/w_0$ ) can be written as follows:

$$\frac{\Delta w}{w_0} \cong \frac{1}{2c_0} \left[ -(N \cdot M)^2 - 2c_0 F - G \right] P \quad (5)$$

where  $P$  represents an uniaxial stress load per unit of natural undeformed area in the direction prescribed by  $M$ . The term  $M \cdot N$  vanishes when the propagation direction is perpendicular to the direction of uniaxial stress.

The equation (5) evaluates the dependence of  $\Delta w/w_0$  with  $P$  taking into account the elastic coefficients of the crystal. On the other hand, it is known that the frequency changes is connected with the sound velocity changes in the media. The experimental method that will be described in next section is based on the measurement of the frequency change as a function of  $P$ . The changes of samples dimensions are also needed to investigate the relative change of the actual velocity ( $\Delta v/v_0$ ).

For a compressive stress  $P$  it is possible to connect the lateral dimension along the  $X$ -axis ( $x_p$ ) and the actual sound velocity ( $v_p$ ) by the following relations :

$$x_p = x_0(1 + \epsilon_x) \quad (6)$$

and

$$v_p = 2x_p \delta f_p \quad (7)$$

where  $x_0$  and  $x_p$  is the path length at the initial condition (zero stress) and that after application of  $P$ , respectively.  $\delta f_p$  is the value of the frequency difference between subsequent resonant peaks under  $P$ .  $\epsilon_x$  is the relative strain along the  $X$ -axis. It is possible to write  $\delta f_p$  as a function of  $\delta f_0$ , the difference between positions of the same peaks at zero stress, as follows :

$$\delta f_p = \delta f_0 + \Delta f \quad (8)$$

where  $\Delta f$  represents the change in the difference between peaks position due to  $P$ . Substituting (7) into (8) and taking into account that  $v_0 = 2x_0 \delta f_0$  is the sound velocity at zero stress we have :

$$\Delta f = \frac{1}{2} \left( \frac{v_p}{x_p} - \frac{v_0}{x_0} \right) \quad (9)$$

The substitution of equation (6) in the equation (9) will give the dependence of  $\Delta v/v_0$  as a function of  $\Delta f/\delta f_0$  and  $\epsilon_x$  :

$$\frac{\Delta v}{v_0} = \frac{\Delta f}{\delta f_0} (1 + \epsilon_x) + \epsilon_x \quad (10)$$

In the present work, we did not measure  $\epsilon_x$ . However, it is possible to determine  $\epsilon_x$  using the strain-stress relation. Using the matrix notation,  $\epsilon_x$  is determined as follows :

$$\epsilon_x = s_{ij}^T T_j' \quad (11)$$

where  $s_{ij}^T$  are the isothermal elastic compliance coefficients and  $T_j'$  are the components of the stress tensor. In the rotated set of axis  $XY'Z'$ , in which  $M$  coincides with  $Z'$ ,  $T_j'$  may be written as follows:

$$T_j' = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -P \end{pmatrix} \quad (12)$$

where  $P$  is the compressive stress along  $Z'$  direction. Making the transformation of coordinates from the system  $XY'Z'$  to the cristal system  $XYZ$  by using the direction cosines that relates both frames, we obtain the following expression for  $\epsilon_x$  :

$$\epsilon_x = -P(\sin^2 \theta s_{12}^T + \cos^2 \theta s_{13}^T - \sin \cos \theta s_{14}^T) \quad (13)$$

where  $\theta$  is the rotation angle between  $XYZ$  and  $XY'Z'$ . Using the equations (10) and (13) it is possible to evaluate  $\Delta\nu/\nu_0$  before and after twinning by using the relative change of frequency as a function of  $P$ .

In the present work, we use the elastic compliances determined by Bechmann [11] and the third-order elastic coefficients are those determined by Thurston et al. [12]. The third-order elastic stiffness coefficients that change sign between twin states are  $C_{114}$ ,  $C_{124}$ ,  $C_{134}$  and  $C_{444}$ . In accordance with the IRE Standard [13],  $s_{14}$  is positive and  $C_{114}$  is negative for the existing twin state.

### 3. Experimental Procedure

The apparatus used to apply almost uniaxial compressive stress was detailed in a previous work [14]. Here we just describe the method used to measure the axial strain. Two pairs of inductive gauge heads were positioned at  $180^\circ$  in relation to the axis of the applied force. Each pair of gauge heads measured the differential displacement between the platens in which the sample was set. The axial strain at each side of sample was evaluated from the differential displacement values.

The composite resonator structure [8] was built using as substrate synthetic quartz crystals of X-cut rotated of  $\theta = -45^\circ$  and  $\theta = -60^\circ$ . The faces of the specimens were ground and polished optically. The error of parallelism between faces was near  $40''$ . The dimensions along  $X$ ,  $Y'$  and  $Z'$  were 2, 5 and 5 mm, for samples rotated of  $-45^\circ$ , and 2, 10 and 4 mm, for samples rotated of  $-60^\circ$ . Longitudinal waves were excited along the  $X$ -axis of the specimen. For this, a ZnO layer of  $9 \mu\text{m}$  thick was deposited on one of the  $X$ -faces between Ti-Al and Al electrodes of 0.18 and  $0.15 \mu\text{m}$  thick, respectively.

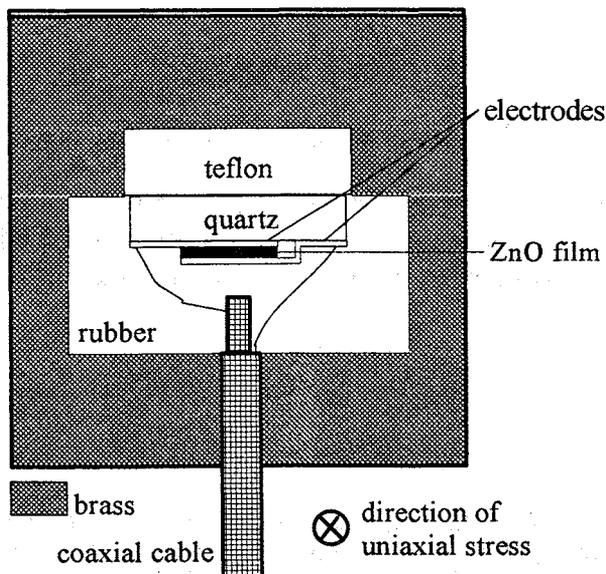


Figure 1 - Schematic representation of experimental arrangement (top view).

The composite resonator was set in a holder as it is shown in Figure 1. Their lateral faces were isolated in order to reduce the effect of parasitic charges. The frequency signal was supplied by a 4396A Hewlett Packard network spectrum analyser.

The basic quantity of measurement was the frequency difference between subsequent resonant peaks ( $\delta f_p$ ). The characteristic pattern of resonant peaks in the frequency-phase domain is illustrated in Figure 2. Several peaks near 300 or 400 MHz were followed as the compressive stress was varied. The average value of  $\delta f_p$  was evaluated in order to minimize the error in measurements of peak positions. The experimental error due to the measurement of minimums or maximums of resonant peaks is of the order of 13 ppm.

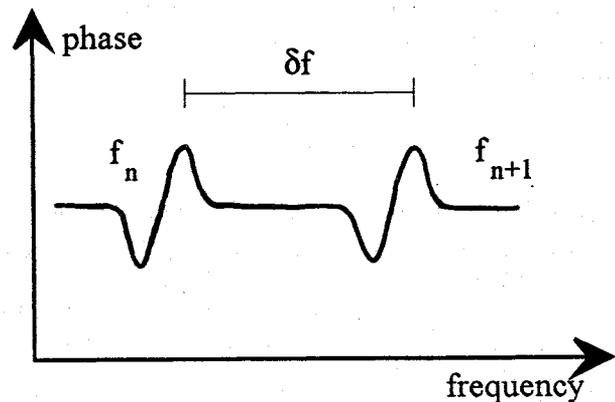


Figure 2 - Schematic phase-frequency resonant spectrum of a composite resonator used in compressive experiments.

### 4. Results and Discussion

Figure 3 shows the dependence of the axial strain measured by each pair of gauge heads in relation to the compressive stress ( $P$ ) for a sample rotated of  $-60^\circ$ . The discrepancy observed in the values of measured strains ( $\epsilon_z$ ) can be associated to non homogeneous compression due to non parallelism of faces.

The relative position of resonant peaks ( $\Delta f / \delta f_0$ ) are plotted in Figure 4 as a function of the applied stress ( $P$ ). Both  $\epsilon_z$  and  $\Delta f / \delta f_0$  increase almost linearly with the increase of  $P$ . Near 400 MPa one observes an abrupt increase of  $\epsilon_z$ , simultaneously with an abrupt decrease of  $\Delta f / \delta f_0$ . By still increasing the pressure, it is observed that  $\Delta f / \delta f_0$  increases in a smaller slope than before the jump while  $\epsilon_z$  increases in a bigger slope. With the subsequent decrease of pressure, both  $\Delta f / \delta f_0$  and  $\epsilon_z$  monotonously decrease to zero keeping almost the same slope which was set after the jump.

As it was previously reported [4, 5, 14] the ferroelastic twinning is usually followed by an abrupt jump in strain-stress relationships. Thus, the jump observed in Figure 4

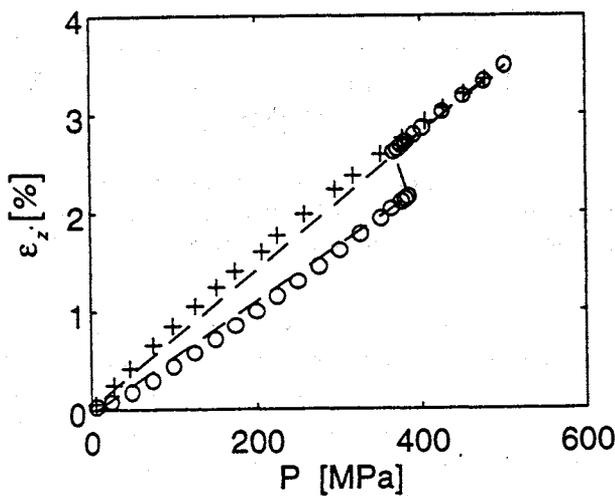
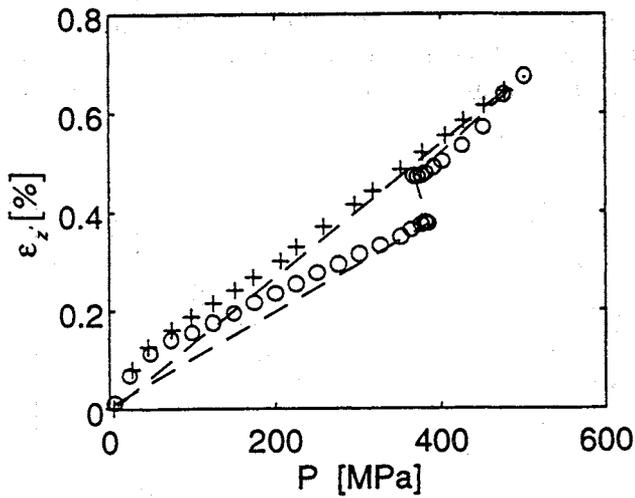


Figure 3 - Strain-stress dependences for a X-cut sample rotated of  $-60^\circ$  (o increase of stress, + decrease of stress).

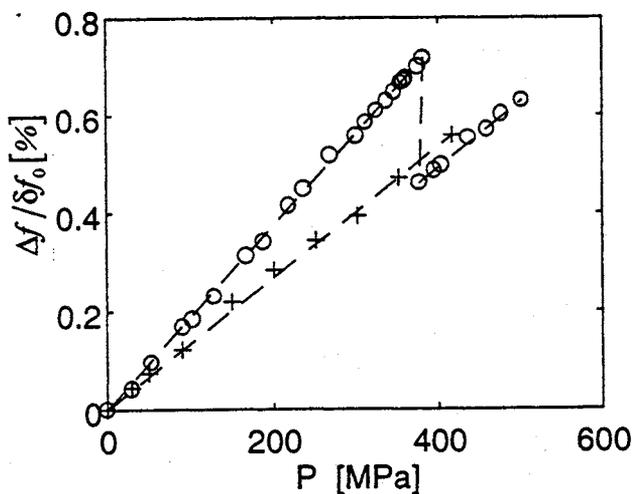


Figure 4 - Relationship between the relative change of frequency and stress for a X-cut sample rotated of  $-60^\circ$  (o increase of stress, + decrease of stress).

are also connected with the transition from the existing twin state (domain I) to the alternate twin state (domain II). As the jump in  $\varepsilon_z$  vs.  $P$  dependence was detected by the both pair of gauge heads, we suggest that the twinning took place in the whole sample. It was also observed that  $\varepsilon_z$  vs.  $P$  and  $\Delta f/\delta f_0$  vs.  $P$  slopes did not change at any stress level after the transition point. This fact submits the irreversibility of the lattice after putting off the pressure. It is in accordance with previous strain-stress experiments [4,5,14].

The evaluation of  $\Delta v/v_0$  for the sample rotated of  $-60^\circ$  is shown in Figure 5. The linear dependence of  $\Delta v/v_0$  vs.  $P$  shows the same slope before and after the transition point. From the equation (13) and the experimental measurements of lateral strains [4], it is noticed that  $\varepsilon_x$  changes sign after twinning.  $\varepsilon_x$  is negative for domain I and positive for domain II. The equation (13) shows that the absolute value of  $\varepsilon_x$  is higher for domain II than for domain I under the same stress. Thus, during the increase of the pressure, when the crystalline lattice belongs to domain I, the length along the X-axis decreases and the influence of  $\varepsilon_x$  on  $\Delta v/v_0$  is not significant. If the pressure is still increased after the transition point, the specimen expands along the X-axis and the influence of  $\varepsilon_x$  on  $\Delta v/v_0$  is much more important. In this experiment, the effect connected with the decrease of  $\Delta f/\delta f_0$  was fully compensated by the expansion of the wave path length in X-direction and thus the actual sound velocity was not affected by twinning.

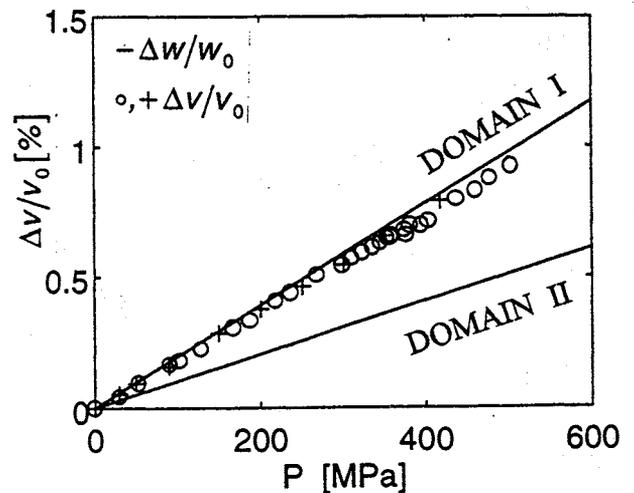


Figure 5 - Relationships between sound velocities and stress for a X-cut sample rotated of  $-60^\circ$  evaluated before and after twinning (o increase of stress, + decrease of stress).

The relationship between the natural velocity and  $P$  is also plotted in Figure 5.  $\Delta w/w_0$  was evaluated for both twin states taking into account both sets of elastic

coefficients. During the increase of stress,  $\Delta w/w_0$  vs.  $P$  dependence has a bigger but not so different slope compared to that of  $\Delta v/v_0$  vs.  $P$  dependence. It is explained by the small decrease of the wave path length when the pressure is increased before the transition point. After twinning, the change of  $\Delta w/w_0$  vs.  $P$  slope is connected with changes in second and third order elastic coefficients while  $\Delta v/v_0$  vs.  $P$  slope does not change due to the compensation of  $\Delta f/\delta f_0$  by  $\epsilon_x$ . From figures 4 and 5 it is observed that  $\Delta f/\delta f_0$  has the same behaviour of  $\Delta w/w_0$  under stress. This fact suggests that  $\Delta f/\delta f_0$  is associated with changes in elastic coefficients.

Quite another situation takes place when the transition in the whole sample is not finished. This may be illustrated by the data obtained from the sample rotated of  $-45^\circ$ . Figure 6 shows the axial strain measured by each pair of gauge heads as a function of  $P$ . Figure 7 shows the relationship between  $\Delta f/\delta f_0$  and  $P$ . The abrupt jump in strain-stress relationship was observed only in one side

of the sample near 500 MPa. Simultaneously, an abrupt jump of about 0.7% was noticed in  $\Delta f/\delta f_0$  vs.  $P$  dependence. After this point, the pressure was released and both  $\epsilon_z$  and  $\Delta f/\delta f_0$  kept the same slope which was set after twinning.

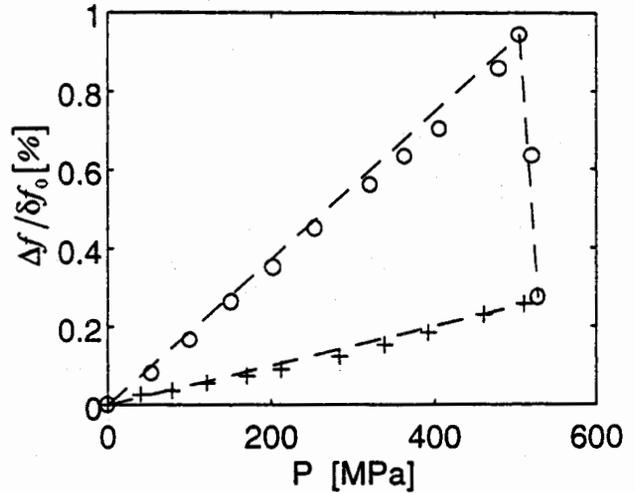


Figure 7 - Relationship between relative change of frequency and stress for a X-cut sample rotated of  $-45^\circ$  (o increase of stress, + decrease of stress).

The disagreement between axial strains showed in Figure 6 compelled us to increase the pressure again. The dependences  $\epsilon_z$  vs.  $P$  and  $\Delta f/\delta f_0$  vs.  $P$  for the second load cycle are shown in Figures 8 and 9, respectively. It appears that the slope of  $\Delta f/f_0$  vs.  $P$  dependence is similar to that set during the increase of pressure in the previous cycle. This observation supports the assumption that crystal parts oriented with the symmetry of domain II switched-back to the symmetry of domain I when the pressure was released in the first cycle.

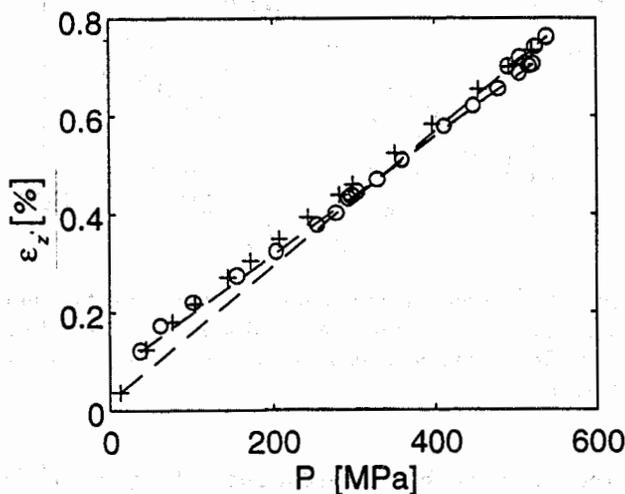
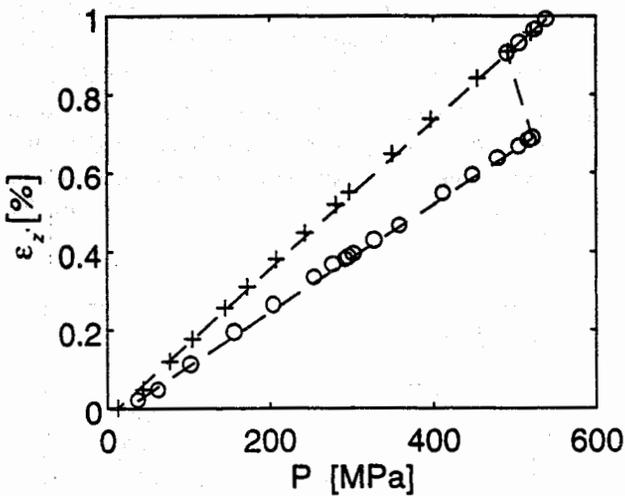


Figure 6 - Strain-stress dependences for a X-cut sample rotated of  $-45^\circ$  (o increase of stress, + decrease of stress).

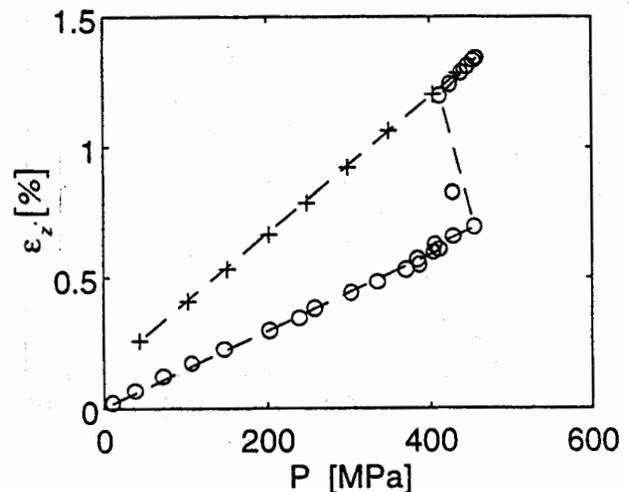


Figure 8 - Second strain-stress dependence for a X-cut sample rotated of  $-45^\circ$ . Here  $\epsilon_z$  is the average of both measurements. (o increase of stress, + decrease of stress).

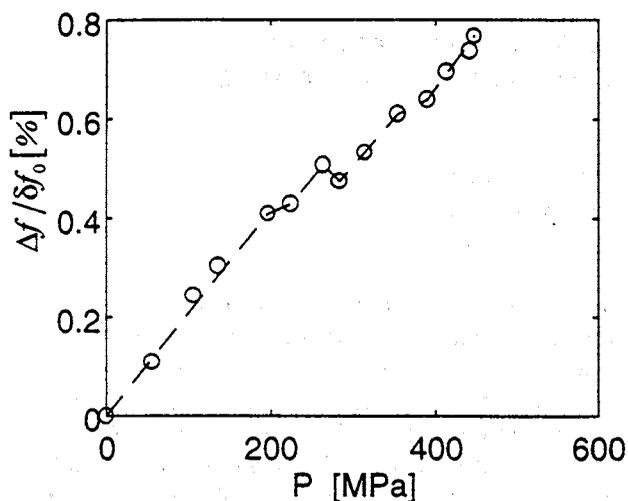


Figure 9 - Second relationship between relative change of frequency and stress for a X-cut sample rotated of  $-45^\circ$  (o increase of stress, + decrease of stress).

Our previously observations of domains patterns [5] showed that the switch-back to the original twin state occurs at very low stress levels when the twinning is not entirely overcome and without change in strain-stress slope.

The relationship between  $\Delta f / \delta f_0$  and  $P$  for the second strain-stress cycle shows several interruptions. It can be qualitatively explained if we consider the existence of strained clusters of  $\text{SiO}_4$  tetrahedra that could favorize the nucleation of the alternate twin state. These interruptions are probably associated to partial transitions that can take place when the pressure is increased again. It suggests that the dependence of  $\Delta f / \delta f_0$  vs.  $P$  is influenced by the existence of both twin states in the crystal. The abrupt jump observed in  $\epsilon_z$  vs.  $P$  dependence near 450 MPa was observed in both sides of the sample. It suggests that the twinning occurred in the whole sample. Unfortunately, at the same time, the sample was partially broken and the electric contact with the composite resonator was lost. In Figure 8, the residual strain observed at zero stress is due to sample's damages. The sound velocity was not evaluated for the sample rotated of  $-45^\circ$  because neither equation (5) nor equation (10) can be applied when both twin states exist simultaneously. This problem must be considered in the future together with the evaluation of sound attenuation.

## 5. Summary

The results of the present work can be summarized as follows :

- (1) Composite resonator spectroscopy can be used to investigate the ferrobielastic twinning in quartz.
- (2) The linear relationship between the relative change of frequency and the compressive stress is strongly modified by the ferrobielastic twinning.

(3) The relative change of frequency and the natural velocity show similar behaviour in ferrobielastic twinning.

(4) The linear relationship between actual sound velocity and compressive stress is not influenced when the ferrobielastic twinning occurs in the whole sample.

## Acknowledgements

One of the authors (PLG) is grateful for scholarship provided by CNPq (Brazil).

## References

- [1] Aizu K, 1973, "Second-order ferroic state shifts", *J. Phys. Soc. Japan*, **34**, 121-128.
- [2] Newnham R.E. and Cross L.E., 1974, "Symmetry of secondary ferroics II", *Mat. Res. Bull.*, **9**, 1021-1032.
- [3] Anderson T.L., Newnham R.E. and Cross L.E., 1977, "Coercive stress for ferrobielastic twinning in quartz", *31st Ann. Symp. on Frequency Control*, 171-177.
- [4] Markgraaff J, 1986, "Elastic behavior of quartz during stress induced Dauphiné twinning", *Phys. Chem. Minerals*, **13**, 102-112.
- [5] Guzzo P.L. and Boy J.J., 1995, "Influence of crystallographic orientation on ferrobielastic twinning in quartz", *9th European Frequency and Time Forum*, 438-442.
- [6] Cline T.W., Laughner J.W., Newnham R.E. and Cross L.E., 1978, "Electrical and acoustic emission during ferrobielastic twinning in quartz", *32nd Ann. Symp. on Frequency Control*, 43-49.
- [7] Bertagnoli E, Kittinger E and Tichy J, 1979, "Ferrobielastic hysteresis in  $\alpha$ -quartz", *J. Appl. Phys.*, **50**, 6267-6271.
- [8] Mansfeld G.D., Freik A.D. and Krutov B, 1993, "A high overtone BAW resonator as a tool for acoustic characterization of thin films and layers", *7th European Frequency and Time Forum*, 261-266.
- [9] Mansfeld GD, 1994, "BAW composite resonator spectroscopy", *IEEE Ultrasonics Symposium*, 655-658.
- [10] Thurston R.N. and Brugger K, 1964, "Third-order elastic constants and the velocity of small amplitude elastic waves in homogeneously stressed media", *Phys. Rev.*, **133**, A1604-A1610.
- [11] Bechmann R, 1958, "Elastic and piezoelectric constants of  $\alpha$ -quartz", *Phys. Rev.*, **110**, 1060-1061.
- [12] Thurston R.N., McSkimin H.J. and Andreatch Jr P, 1966, "Third-order elastic coefficients of quartz", *J. Appl. Phys.*, **37**, 267-275.
- [13] "Standard on piezoelectric crystals", 1949, *Proc. I.R.E.*, **37**, 1378-1395.
- [14] Guzzo P.L., Boy J.J., Dulmet B and Maitre P, 1994, "Influence of the sample geometry on twinning generation in quartz : experimental measurements and finite elements analysis", *8th European Frequency and Time Forum*, 802-810.

## ON THE FOUNDATIONS OF CONTINUUM PHYSICS INCLUDING ELECTROMAGNETIC INTERACTIONS AND INTERFACIAL PROPERTIES

N. DAHER

*Laboratoire de Physique et Métrologie des Oscillateurs du CNRS,  
Associé à l'Université de Franche-Comté-Besançon  
32, Avenue de l'Observatoire, 25044 Besançon Cedex - France*

### Abstract:

This work concerns the development of a general nonlinear theory including discontinuities and interfacial properties in the framework of continuum physics. After recalling the basic ideas on which the different theories are founded, difficulties that occur at interfaces are placed in evidence and discussed. A solution is proposed through a new reformulation where the idea of invariance under different transformations plays a major role. Finally, some features relative to the construction of an approach including complex structures and particularly deformable ferroelectric and piezoelectric semiconductors are outlined.

### Introduction :

Nonlinear continuum theories [1-7] including electromagnetomechanical interactions reversible and irreversible processes such as diffusion and conduction phenomena and interfacial properties have been developed through either vectorial or energy formulations. In a first step, the basic ideas and concepts at the origin of these approaches are recalled, then the advantages and drawbacks of each formulation are placed in evidence, particularly in the study of complex structures where different continua interact nonlinearly with each other. In a second step, it is shown that although the two approaches have been recently extended to account for interfacial properties, the obtained results are partial, sometimes inconsistent and may lead to certain ambiguities. This is not due to the fact that the methods are not applied properly but rather to intrinsic difficulties intimately related to the basic postulates. These aspects are developed and a solution is proposed where a new energy method emerges. The latter constitutes a unified mode of expression based on the use of three invariance principles associated with (i) mechanical and coupled effects (invariance under rigid body motion) (ii) electromagnetism (gauge invariance) and (iii) passage from a  $n$  dimensional space to a  $n-1$  dimensional embedded subspace (scale-change invariance). Finally, the interest of such an approach is placed in evidence by its application to complex structures such as nonlinear deformable ferroelectric semiconductors including interfacial properties, where none of the conventional formulations apply.

The present theory requires the establishment of some mathematical tools that allow the generalization of

Green-Gauss and transport theorems to singular surfaces and lines in addition to a number of physical principles. All these local aspects were recently developed in a number of papers and in different contexts (some of the references are given along the paper). The development of these ideas in mathematical form being impossible in a brief space only the important features and results are outlined.

### Two basic approaches in physics : the vectorial and scalar formulations

The history of physics can be analyzed in terms of oscillations between two viewpoints. One of them, favoured by Newton in the XVIIth century, takes a vectorial concept (force) as fundamental and the other, favoured by Leibniz at the same period, takes a scalar concept (energy) as the basic entity. Later developments showed that although the two methods take their origin in two opposite philosophies, they are equivalent in their predictions as to classical mechanics and electromagnetism and hence scientifically indistinguishable. However, since both methods are neither philosophically nor logically identical in so far as the basic hypotheses are concerned, it is natural to question their validity and to establish their potentialities when trying to move from simple mechanical and electromagnetic situations into the unknown.

### Advantages and drawbacks in Quantum and Classical field theories

Let us first have a general view on physics as a whole by making a distinction between the quantum and the classical field theories. In quantum field theory [8] the concept of force in the Newtonian sense is absent while the concept of energy is very useful particularly through the so-called Lagrangian approach. One of its main interests is its accommodation to symmetry properties and invariance principles. As to the classical field theory, the vectorial approach is not only present but it supersedes the energy method. The main reason is that unlike the microreversibility principle that governs the microcosm, irreversibility is essential in the macrocosm. A "force" does not derive from a scalar in general. Consequently, although the Lagrangian approach is a precious tool to quantum field theory it is not sufficiently wide to include certain phenomena encountered in macroscopic physics.

Since the early sixties, the vectorial approach has been developed on an axiomatic basis where continuum mechanics and thermodynamics have merged into a single science before including electromagnetism at a later stage [7]. The combination of the laws of mechanics, thermodynamics and electromagnetism in a nonlinear framework is a necessity in a number of practical and realistic applications. This is particularly the case in the problem we are interested in concerning electromagnetoelastic and acoustooptic effects as well as thermal and electronic conduction, diffusion and transport phenomena in semiconductor devices.

### Difficulties associated with the vectorial approach in complex structures

Although the vectorial approach is applicable in principle to complex structures one needs to postulate and construct several interacting and interpenetrating continua that complement the mechanical well-known lattice continuum. In deformable semiconductors, for example, these are the polarizable, magnetizable, impurity, electronic and hole continua [3, 4, 5]. The account for all these phenomena requires the construction of a great number of laws. More precisely, the law of conservation of momentum and moment of momentum have to be established for each continuum. In proceeding through this vectorial formulation, not only the number of postulates rapidly increases but the construction of the postulates leads to a certain degree of arbitrariness.

### Introduction of a hybrid approach: Partially scalar, Partially vectorial

The lack of a general energy method that applies simultaneously to mechanics, electromagnetism and irreversible thermodynamics led some authors to use a hybrid approach established in the last decades [3-6]. The basic idea was to benefit from the advantages of both approaches. In dealing with Maxwell equations, the vectorial approach allows to obtain the bulk as well as the jump relations across discontinuity surfaces, by the use of a generalized version of Stoke's and Green-Gauss theorems. As to the electromagnetomechanical equations, they were modeled by an energy approach of a mechanical origin, that generalizes Lagrangians and which is well adapted to the use of invariance under a rigid body motion, needed for obtaining different coupled effects in a consistent manner. This energy method takes its origin in a scalar mechanical principle (d'Alembert's principle) practically unknown by physicists while it is used in mechanics and was particularly developed by french theoretical mechanicians [3-6, 9, 10].

### Back to the origin of the three different approaches and possible extension

Historically, d'Alembert's principle [11, 12] emerged through the attempts to reconcile the different views defended by Newton and Leibniz concerning the basic concepts related to the physical world. More precisely, in order to place in evidence the link between the inertial Newtonian force and the leibnizian "vis viva" (kinetic energy), d'Alembert adopted Leibniz's idea according to which the Newtonian measure of motion by the time rate of change of momentum is to be replaced by the following statement [12]: "The change of the kinetic energy is equal to the work done by the force". D'Alembert put this idea in a formal manner extending to dynamics the principle of virtual work that goes back to Aristotle but that was recognized much later by Bernouilli as a general principle of statics. Mathematically, this principle is expressed in the form of an orthogonality relation:

$$\langle \mathbf{F}^{\text{eff}}, \mathbf{V}^* \rangle = 0, \forall \mathbf{V}^*, \Rightarrow \mathbf{F}^{\text{eff}} = 0 \quad (1)$$

where the dynamical field  $\mathbf{F}^{\text{eff}}$  is the dual of a kinematical field  $\mathbf{V}^*$ , the star indicates the virtual character of the velocity field  $\mathbf{V}$  and  $\mathbf{F}^{\text{eff}}$  includes two contributions: gravitational force plus the negative product of mass times acceleration which is considered as a force created by the motion. Although apparently nothing is gained since one recovers Newton's law, it is precisely this apparent triviality which makes of this reformulation an ingenious invention that was greatly distorted and misunderstood in the past.

In order to place in evidence the generality as well as the unifying character of this principle, let us start with the simplest example of a continuous medium. Here, tensorial elements enter into the picture since scalars and vectors are not sufficient to account for some physical entities such as the notion of deformation (symmetric part of the displacement gradient noted  $u_{(ij)}$ ) and stress tensor  $\sigma_{ij}$ . In a purely mechanical but deformable body, the orthogonality relation (1) globally holds except that one has to construct a scalar which includes the velocity field  $\mathbf{V}$  as well as its first gradient. In order to provide a direct link with the Lagrangian approach, one may replace the velocity  $\mathbf{V}$  by an infinitesimal displacement  $\delta \mathbf{u}$ . Thus in a first gradient approximation one is led to  $\mathbf{F}_i^{\text{eff}} \delta u_i - \sigma_{ij} \delta u_{i,j}$ , where  $\sigma_{ij}$  does not necessarily derive from a scalar as in the Lagrangian approach ( $\sigma_{ij} = \partial L / \partial u_{i,j}$ ). In addition, one distinguishes between external and internal forces (here matter-matter interaction represented by the stress tensor). Internal forces for which one ultimately needs to construct constitutive relations have to be formulated in such a manner that they satisfy the invariance of the response of a material body under a superimposed rigid

body motion. This invariance requirement, places restriction on the stress tensor. Keeping this in mind, we can now appreciate, the generality and the unifying character of d'Alembert's principle. Firstly, on assuming that eq. (1) holds true for any virtual motion  $\mathbf{V}^*$  one recovers the fundamental law of Newtonian mechanics (consequence of conservation of momentum). Secondly, the use of the above-mentioned invariance property leads to the symmetry of the stress tensor (consequence of conservation of moment of momentum) and thirdly, on letting the virtual vector  $\mathbf{V}^*$  coincide with the actual velocity  $\mathbf{V}$ , one obtains the conservation of mechanical energy. We have thus three basic informations condensed in one scalar statement. Let us note at this point that this procedure maybe extended to complex structures by introducing generalized kinematical fields [3-6]. In the study of electromagnetomechanical interactions including heat and electronic diffusion and conduction phenomena, the generalized kinematical fields will be the mechanical velocity, the material time derivative of the polarization and magnetization and the velocity of the different carriers (impurities, electrons, holes).

By recalling the basic elements and principles on which the vectoriel, energy and hybrid approaches are based as well as the advantages and drawbacks of each formulation, we have paved the way for the examination of the different approaches when applied to interfacial science. Before moving to interfacial considerations, let us note that d'Alembert's approach may be used to derive Maxwell equations. This fact is very useful when dealing with interfaces as it will be shown in the forthcoming developments. In proceeding in this manner one does not only obtain new results but one also gains a certain unity when a comparison is performed with mechanics. Indeed, in the same manner as the symmetry of the stress tensor is a consequence of rotational invariance, the skewsymmetry of the magnetic tensor will be a consequence of gauge invariance. In addition, one may show that in a Lorentzian four dimensional space one recovers the whole set of Maxwell equations through one postulate subject to one constraint (gauge invariance) [13].

#### Position of the Physical problem concerning interfacial science

In the forthcoming developments, I shall raise a number of questions that deserve a special attention since they concern the basic structure of continuum physics. Indeed, although the fundamental postulates seem to apply to simple situations with no ambiguity using either the vectoriel approach or the scalar one, serious difficulties occur when dealing with interfacial properties.

#### a) Vectorial approach: impossibility of lowering spatial dimensions

In the study of discontinuities and interfaces where the classical jump and boundary conditions are not sufficient anymore some of the pillars on which the vectorial structure is based collapse. This is the case of Ampère's law and the like (line integrals). To see this clearly let us recall that the introduction of surface parameters in addition to well-known volume ones, requires a lowering of the space dimension by one order ( $n \rightarrow n-1$ ). Thus for Gauss's law for example one sets

$$\int_{\partial D - \Sigma} \mathbf{D} \cdot \mathbf{n} \, da + \int_{\partial \Sigma} \hat{\mathbf{D}} \cdot \mathbf{b} \, dl = \int_{D - \Sigma} q_f \, dv + \int_{\Sigma} \hat{q}_f \, da \quad (2)$$

as to Ampere's law, even in magnetostatics, the lowering procedure breaks off

$$\oint_{\partial S - \gamma} \mathbf{H} \cdot \mathbf{t} \, dl + (?) = \int_{S - \gamma} \mathbf{J} \cdot \mathbf{n} \, da + \int_{\gamma} \hat{\mathbf{J}} \cdot \mathbf{b} \, dl \quad (3)$$

$\mathbf{D}$ ,  $q_f$ ,  $\mathbf{H}$ ,  $\mathbf{J}$  and  $\hat{\mathbf{D}}$ ,  $\hat{q}_f$ ,  $\hat{\mathbf{H}}$ ,  $\hat{\mathbf{J}}$  denote respectively the electric displacement, the free charge, the magnetic field, the electric current and their interfacial counterparts. The three dimensional domain  $D$  is a simply connected, open material region of  $E^3$  with regular boundary  $\partial D$  having a unit oriented normal  $\mathbf{n}$ .  $\Sigma$  is a discontinuity surface contained in  $D$ . The two dimensional domain  $S$  is an open material surface in  $E^3$  enclosed by a closed curve  $\partial S$  having a unit binormal  $\mathbf{b}$  ( $\mathbf{b} = \mathbf{t} \times \mathbf{n}$ )  $\gamma$  is a discontinuity curve on  $S$ .  $dv$ ,  $da$  and  $dl$  denote respectively volume, surface and line elements.

#### b) Scalar approach : Impossibility of the multiplication of distributions

One of the basic differences between the vectorial and scalar approaches is that the basic fields in an energy like formulation are postulated in the bulk while these fields are postulated at boundaries (closed surfaces or lines) as in eqs. (2) and (3), when the vectorial approach is used. Consequently, no problem occurs with the energy formulation when lowering the dimension, but there is another problem in relation with the scalar or energy approach. To see this let us recall that (i) an energy is a scalar product, (ii) the usually continuous functions must be replaced by discontinuous ones (distributions) since the bulk medium is now cut by a discontinuity surface and (iii) the product of two distributions is not a well defined mathematical operation. The three above mentioned propositions clearly show the difficulty when distributions are to be used simultaneously with scalar products. Notice,

however, that unlike the vectorial approach, the energy method does not degenerate but the operation of multiplication is ill defined, when the distribution theory is used. This means that if a solution is available, then it must avoid the use of distribution theory and one should look for it in the framework of the energy approach where not only the dimension can be lowered from volumes to singular surfaces but also to singular lines. Notice that in this last case, the vectorial approach degenerates in electromagnetism and in mechanics as well, since there exists no closed boundary for a singular line. In brief, one may say that the trouble comes from the gradients of the volume functions when the latter are in close contact with the interfacial region [5, 10]. Mathematically one writes:

$$\int_D W^T dv = \int_{D-\Sigma} W(\nabla_j a_i) dv + \int_{\Sigma} \left\{ ? + \hat{W}(\hat{\nabla}_j \hat{a}_i) \right\} da \quad (4)$$

where  $W$  and  $\hat{W}$  denote respectively the volume and surface energies whose expressions can be given with no ambiguity while  $?$  indicates the energy that couples the volume to the interface and that needs to be determined.

#### Towards a new approach based on invariance principles

Eq. (3) shows that the degeneracy occurs with surface coefficients, while the impossibility of distribution products concerns the discontinuity of the bulk field at either side of the interface. One may take benefit from this opportunity and combine the two methods to obtain a consistent result. However, it must be emphasized that such a procedure is neither practical since it is feasible only in simple situations nor desirable since one loses the elegance and the unity of the basic postulates. It is simply a first step, of a heuristic value, that leads to the discovery of a correct result, but the method is neither self-consistent nor self-contained. This lack of rationality and unity requires further investigations in order to avoid having one foot in the vectorial approach and the other in the scalar formulation and to obtain independent, intrinsic postulates by cutting all links with the vectorial approach. This second phase, essential for further generalizations, is obtained by imposing internal constraints on the scalar approach in the form of invariance principles. More precisely, one starts by postulating a first gradient theory in the bulk and at the interface as follows [13]:

$$\begin{aligned} \delta W &= H_i \delta A_i + H_{ij} \nabla_j \delta A_i, \\ \delta \hat{W} &= \hat{H}_i \delta \hat{A}_i + \hat{H}_{ij} \hat{\nabla}_j \delta \hat{A}_i \end{aligned} \quad (5)$$

where a caret (^) is used to indicate surface parameters and operators. Then, one imposes gauge invariance

(invariance under the addition of the gradient of any scalar field) leading to the following restrictions:

$$\begin{aligned} H_i &= 0, H_{ij} + H_{ji} = 0, \hat{H}_i = 0, \\ \hat{H}_{ij} + \hat{H}_{ji} &= \hat{H}_{ik} n_k n_j + \hat{H}_{jk} n_k n_i \end{aligned} \quad (6)$$

Notice here that unlike the volume tensor, the surface tensor is not skewsymmetric in the general case. This fact is very important since it clearly shows that as long as one persists in treating electromagnetism in terms of vectors and pseudovectors as is usually the case in engineering sciences [14], one is bound to deal only with particular solutions implicitly assuming  $\hat{H}_{ik} n_k = 0$  for example. Indeed, in this particular case  $\hat{H}_{ij}$  becomes skewsymmetric and only then, it can be identified to a pseudovector. There is a second invariance principle needed for the full determination of eq. (4) which we call "scale-change invariance" that we do not reproduce here for lack of space [13, 15]. Let us note however that unlike gauge invariance which characterizes Maxwellian electromagnetism, scale-change invariance is of a mathematical nature. It is a substitute for the impossibility of distribution products. Its form remains invariant when dealing with mechanics, electromagnetism or electromagnetomechanical interactions. One would like to say it is of a geometrical nature but there is a risk of ambiguity here since gauge invariance may also be interpreted as a curvature in a certain geometrical space [16, 17].

#### Application to ferroelectric and piezoelectric semiconductors

Now that we have established all the necessary ingredients needed for the construction of complex structures including interfacial properties, by use of a unified mode of expression, let us give a plan showing how the use of invariance principles associated with an energy formulation and using Green-Gauss and transport theorems [19] allows the construction of a consistent theory of deformable, ferroelectric and piezoelectric semiconductors.

*Electromagnetism* : As pointed out in the last section, gauge and scale-change invariance lead to volume and surface differential equations generalizing thus Maxwell equations [13, 15].

*Electromagnetomechanics* : Here, the invariance under rigid body motion and again the scale-change invariance play a major role in showing the way the different continua couple together and lead to a local balance equation for each continuum [18].

*Thermodynamics* : The above mentioned equations are combined with the first and second principles of thermodynamics, leading to the so-called Clausius-Duhem inequality, essential for the determination of reversible and irreversible constitutive relations [4, 5].

### Conclusion

Although the attention has been focused on the methods and on the basic principles, the origin of this work lies in practical problems associated with the study of volume and surface phenomena such as wave propagation problems. As is usual in scientific investigations one looks for new ideas, methods and principles only when the old ones fail and cease to satisfy. Here, the difficulties encountered through the study of interfacial properties in semiconductors were the chief motive for these general investigations. Finally, as pointed out in the introduction, my concern here was to give a qualitative and general understanding of this subject matter through an overall picture. This obviously required to work with one eye on the past another on the future.

### References

- [1] D.F. Nelson, *Electric, Optic and Acoustic Interactions in Dielectrics* (Wiley, New York, 1979).
- [2] H.G. de Lorenzi and H.F. Tiersten, *J. of Math. Physics*, 16 (4), april 1975
- [3] G.A. Maugin, *Acta Mechanica* 35,1-70 (1980)
- [4] N. Daher and G.A. Maugin, *Int. J.Eng. Sci.* 25,1093-1129 (1987)
- [5] N. Daher, "Principe des puissances virtuelles étendu aux discontinuités et interfaces, application à l'acousto-électronique". Doctoral Thesis, Besançon (1987)
- [6] G.A. Maugin, J. Pouget, *J. Acoust. Soc. Amer.* 68, 575-582 (1980)
- [7] A.C. Eringen, G.A. Maugin, "Electrodynamics of Continua I, Foundations and Solid Media" Springer Verlag (1989)
- [8] C. Itzykson and J.B. Zuber "Quantum field theory", Mc GRAW HILL (1980)
- [9] P. Germain, *J. de Mécanique* 12, 235-274 (1973)
- [10] N. Daher and G.A. Maugin, *Acta Mechanica* 60, 217-240 (1986)
- [11] J. d'Alembert, *Traité de dynamique*, Paris (1743)
- [12] C. Lanczos "The variational principles of mechanics", University of Toronto Press
- [13] N. Daher, *Synthetic Metals* 67 (1-3), 287-291 (1994)
- [14] A.M. Albano, D. Bedeaux and J. Vlieger, *Physica* 102 A, 105-119 (1980)
- [15] N. Daher, EFTF 94, Technical University Munich, Germany, (1994)
- [16] B. Felsger "Geometry, particles and fields", Odense University press (1981)
- [17] M. Nakahara "Geometry, Topology and Physics", Institute of Physics Publishing Bristol and Philadelphia (1990)
- [18] N. Daher, *Synthetic Metals*, vol. Syn. 076/1-3, pp. 327-330 (1995)
- [19] N. Daher, IUTAM Symposium on nonlinear instability and transition in 3D boundary layers, Manchester (1995)

# AN ADVANCED ULTRA-SENSITIVE DETECTOR FOR THE MEASUREMENT OF AM AND PM NOISE IN MICROWAVE COMPONENTS

R.A. WOODE, E.N. IVANOV, M.E. TOBAR

The Department of Physics, The University of Western Australia,  
Nedlands, Perth 6907 WA, Australia

**Abstract** - The phase and amplitude noise in many microwave components have yet to be understood. This is because conventional measurement systems are not sensitive enough. We have been able to measure the phase and amplitude noise in components such as isolators, power limiters and voltage controlled phase shifters and attenuators. This was possible due to the advent of the ultra-sensitive Ivanov-Tobar-Woode (ITW) measurement system (Patent Pending). For example, the phase noise in an isolator is shown to be flicker of phase noise with a level of -184 dBc/Hz at 1 kHz offset. The noise floor of the measuring system was also flicker of phase below 1 kHz, and flat at a level of -193 dBc/Hz above 1 kHz.

## I. INTRODUCTION

The latest generation of low noise microwave oscillators have such low noise, that the noise level of individual components such as isolators and varactor phase shifters needs to be known [1,2]. The phase noise of these components has rarely been measured because the noise floor of a typical measuring system is usually worse than the devices under test. A standard X-Band double balance mixer phase bridge has a noise floor of approximately -150 dBc/Hz at 1 kHz offset frequency.

The noise floor of a simple phase bridge may be lowered using a technique such as cross correlation developed by the National Institute of Measurements and Technology (NIST) [3]. This signal processing

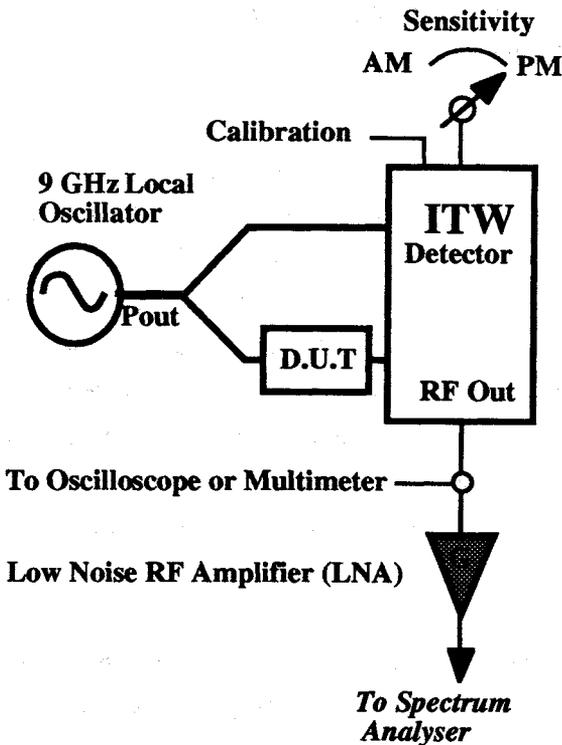
technique reduces the noise floor by  $N^{-0.5}$ , where  $N$  is the number of averages. This technique has been used to improve the noise floor by approximately -15 dB, achieving -165 dBc/Hz at 1 kHz offset [3]. The drawbacks are; the amount of time needed to reduce the noise floor; and that two measurement systems are needed for the correlation. We have built a new measurement system with a noise floor of -193 dBc/Hz above 1 kHz, and with the ability of measuring the noise in real time.

We report on noise measurements of microwave isolators, power limiters, voltage controlled, phase shifters and attenuators. The measurement of noise in these components was only possible due to the advent of the new ITW noise measurement system. For example, we measured a level of -184 dBc/Hz at a kHz offset in an isolator. This is substantially lower than the level of Isolator phase noise calculated by N.N. Kolachevsky [4] and D.P. Tsarapkin [5]. Note that even the NIST cross correlation technique is inadequate for measuring the noise in isolators.

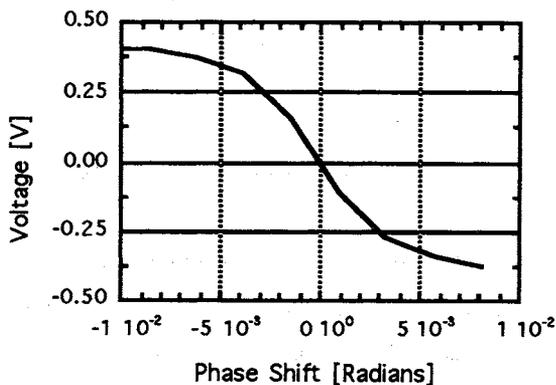
## II. THE ITW MEASURING SYSTEM

The ITW detector can be made either AM or PM sensitive with quadrature settings ( $0^\circ$  or  $90^\circ$ ) with a internal phase shifter as shown in figure 1. Thus the tuning procedure is similar to a simple phase bridge which incorporates a double balance mixer. The ITW detector is calibrated by introducing a known phase modulation (PM) and measuring the phase to voltage

conversion ratio. A typical S shaped detector output curve is shown in figure 2, with the maximum sensitivity of 110 V/Radian at 0°.



**Figure 1.** The ITW detector combined with a device under test (D.U.T). The sensitivity of the detector in V/Radian is measured by introducing a known phase modulation and measuring the output on the oscilloscope (C.R.O). The system noise floor is measured by replacing the DUT with a piece of plain cable.



**Figure 2.** The detector output voltage as a function of phase shift for the ITW measurement system. The conversion ratio from phase to voltage [V/Radian] may be calculated from the slope of the output voltage. The maximum slope is at the zero crossing and is 110 V/Radian. The conversion ratio was also confirmed by

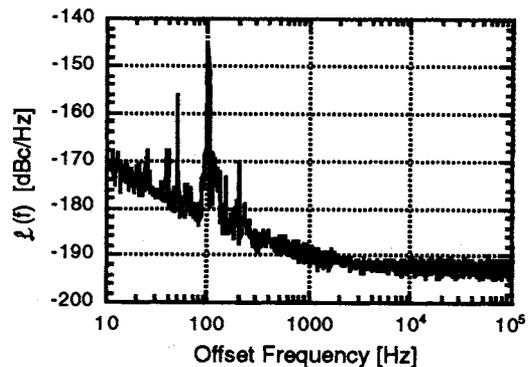
introducing a phase modulation into the measurement port while monitoring the detector output. The conversion ratio is dependent on the input power level to the detector.

The device under test adds its insertion loss to the measurement. This can vary with each individual component which increases or decreases the sensitivity of the detector. The peak sensitivity of the detector measured was approximately 160 V/Radian which resulted in a measured noise floor of -193 dBc/Hz, above a kHz. At frequencies below a 1 kHz offset, the noise floor followed the flicker law given by;

$$\mathfrak{L}(f)_{SSB}^{N.F} = -163 - 10 \log(f) \quad (1)$$

where  $\mathfrak{L}(f)$  is the single sideband noise floor in [dBc/Hz] and  $f$  is the offset frequency.

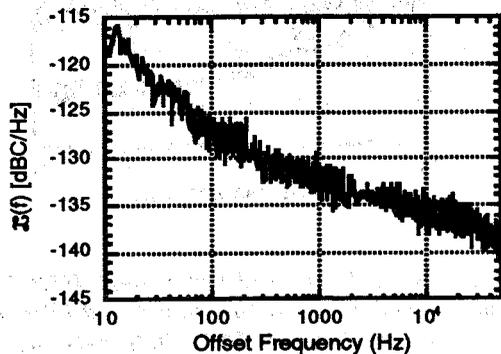
The detector noise floor in figure 3 showed some extra noise at low offset frequencies, which was due to external vibration and pickup of spurious electromagnetic radiation.



**Figure 3.** Measurement noise floor of the ITW ultra-sensitive phase detector with +20 dBm incident power. There is some excess noise at low offset frequencies which is due to the vibration sensitivity. There is also some evidence of electromagnetic interference.

A microwave amplifier with known phase noise was tested. The amplifier was configured with attenuators both before and after which maintained an acceptable input power. Figure 4. shows the phase noise of the microwave amplifier measured using the ITW detector

as  $-133$  dBc/Hz at 1 kHz offset. This level of phase noise is fairly typical and was within a dB of the noise measured by a conventional mixer system.



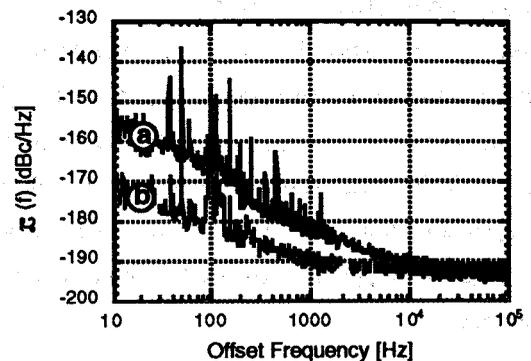
**Figure 4.** The phase noise of a microwave amplifier measured at 9 GHz using the ITW detector. This measurement was confirmed when comparing the phase noise with the measured noise using a conventional system.

### III PHASE NOISE RESULTS

The phase noise in a variety of active and passive microwave components were successfully measured at 9 GHz using the ITW detector.

#### Ferrite Isolators.

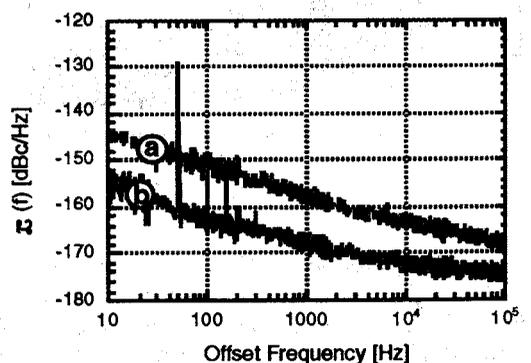
The phase noise in ferrite isolators has never before been reliably measured. However there has been some conjecture about the exact levels of noise, and the processes involved [4,5]. Here we present the first reliable phase noise measurements of an isolator. This was possible because of the ITW measurement system as shown in figure 1. The ITW detector was configured to be phase sensitive with a maximum power of +20 dBm incident on the isolator under test. The phase noise of the isolator was measured as  $-184$  dBc/Hz at 1 kHz offset and is shown in figure 5. The phase noise measurements were repeated with multiple isolators from different manufacturers. In all cases the noise measured was at least 10 dB above the noise floor and in the range of  $-183$  to  $-186$  dBc/Hz at 1 kHz offset.



**Figure 5.** Phase noise in a single microwave isolator at 9 GHz and +20 dBm of incident power. Curve a; The phase noise of a microwave isolator. Curve b; The measuring system noise floor. Below a few kHz the phase noise of the isolator was well above the noise floor.

#### Microwave power Limiters.

The second device tested was a microwave power limiter. The limiter operated in the limiting regime when more than +7 dBm was passed through it. The measured phase noise of the microwave limiter is shown in figure 6. The phase noise of the limiter was  $-158$  and  $-168$  dBc/Hz at 1 kHz offset for the limiting and non limiting regimes respectively.

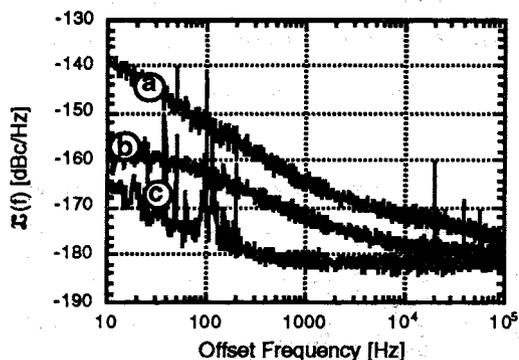


**Figure 6.** The phase noise of a microwave power limiter measured in the limiting and non limiting regimes. Curve a; The phase noise measured in the limiting regime with +10 dBm input power. Curve b; The phase noise in the non limiting regime, with an input power of +4 dBm. The noise floor for each measurement was well below the levels measured here.

### Voltage Controlled Devices.

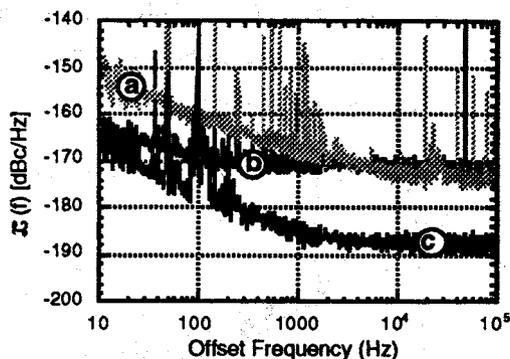
Microwave voltage controlled devices are used in many applications in which their phase noise is critical [1,2]. A selection of voltage controlled varactor diode phase shifters (VCPs) and voltage controlled pin diode attenuators were tested (VCAs). Figure 7. shows the phase noise measured in two VCPs from different manufacturers.

The phase noise level measured varied from -165 to -172 dBc/Hz for VCP 1 and 2 respectively. Both had a similar dynamic range and were biased with batteries at +3 volts. The VCP phase noise was found not to be significantly dependent on the bias voltage. Changing the batteries for a noisy bias supply saw the general level of phase noise remain unchanged, however there was a significant increase in electromagnetic spikes. The phase noise shown in figure 7 was determined to be intrinsic to the VCPs and not due to the conversion of bias voltage noise to phase noise.



**Figure 7.** The phase noise of two voltage controlled phase shifters from different manufacturers measured at 9 GHz with +10 dBm of input power. Both VCPs were biased with batteries at +3 volts. Curve a; The Phase noise of VCP 1. Curve b; The phase noise of VCP 2. Curve c; The measuring system noise floor.

The phase noise measured in two voltage controlled attenuators is shown in figure 8. The attenuators were both biased so that the sensitivity of ITW detector was the same in both cases.

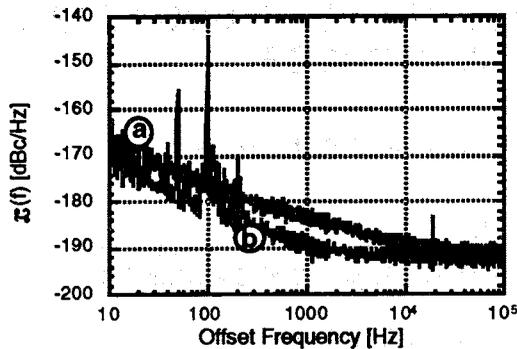


**Figure 8.** The phase noise of two voltage controlled attenuators from different manufacturers, at 9 GHz and with an input power of +20 dBm. Both were biased with voltage supplied by a battery. Curve a; The phase noise measured for VCA 1. Curve b; The phase noise of VCA 2. Curve c; The measuring system noise floor.

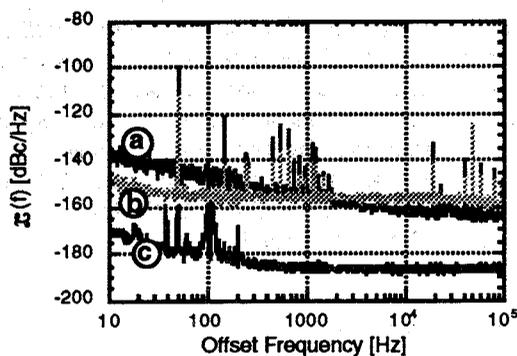
### IV AMPLITUDE NOISE RESULTS

It is also important to characterise the amplitude noise of the voltage controlled components as it can be as critical as the phase noise for some applications.

The amplitude noise of VCP 2 shown in figure 9 was very low at -184 dBc/Hz at 1 kHz offset, which was more than 10 dB lower than its measured phase noise. The amplitude noise measured for VCA 1 and 2 is shown in figure 10. The attenuators were biased so that the sensitivity of the measuring system was the same as the previous phase noise measurements. The attenuators amplitude noise was -157 dBc/Hz at 1 kHz offset, which is more than 10 dB higher than the phase noise.



**Figure 9** Amplitude noise of VCP 2. These results were measured at the same levels of bias and power as the phase noise plot for VCP 2 in figure 7.



**Figure 10.** The Amplitude noise of the voltage controlled attenuators VCA 1 and 2, measured at 9 GHz with an input power of +20 dBm, and bias voltage supplied by battery. Curve a; The amplitude noise measured for VCA 1. Curve b; The amplitude noise of VCA 2. Curve c; The measuring system noise floor.

## V CONCLUSION

We have measured the phase noise and in some cases amplitude noise of microwave isolators, power limiters, and voltage controlled, phase shifters and attenuators. These measurements were made possible with the advent of the ultra-sensitive ITW detector with a measurement noise floor of -193 dBc/Hz at offset frequencies greater than 1 kHz. The flicker phase noise of a ferrite microwave isolator was measured to be -184 dBc/Hz at 1 kHz offset frequency which was much lower than the level calculated in [4,5]. The knowledge of the noise of these components is very

important for understanding the noise in advanced low noise microwave circuits.

## VI ACKNOWLEDGEMENTS

This project was supported by the University of Western Australia, and the Australian Research Council through their Collaborative Research and APRA-I schemes.

## REFERENCES

1. E.N. Ivanov, M.E. Tobar, R.A. Woode, 1995, "Advanced phase noise suppression technique for the next generation of ultra-low noise microwave oscillators," Proc. of 1995 IEEE Freq. Control Symp., pp 315 to 320.
2. R.A. Woode, M.E. Tobar, E.N. Ivanov, 1995, "An ultra-low noise microwave oscillator based on a high-Q liquid nitrogen cooled sapphire resonator," , " Proc. of 1995 IEEE Freq. Control Symp., pp 420 to 426.
3. F.G. Ascarrunz, E.S. Ferre and F.L. Walls, 1993, "Investigations of AM and PM noise in X-Band devices," Proc. of 1993 IEEE Freq. Control Symp., pp 303 to 311.
4. N.N. Kolachevsky, 1985, Moscow, Nayaka (In Russian), Fluctuation phenomena in ferromagnetic materials. ,
5. D.P. Tsarapkin, 1994, "Low phase noise sapphire disk dielectric resonator oscillator with combined stabilization," Proc. of 1994 IEEE Freq. Control Symp., pp 451 to 458.

# SPECTRAL ANALYSIS OF IRREGULARLY SPACED TIMING DATA : COMPARISON OF SEVERAL METHODS

**F. Vernotte\*, G. Zalamansky\*, M. McHugh\* and E. Lantz\*\***

**\*Observatoire de Besançon**

41 bis, avenue de l'Observatoire - B.P. 1615 - 25010 BESANÇON Cedex - FRANCE  
☎ : (33) 81.66.69.22 - Fax : (33) 81.66.69.44 - E-mail : FRANCOIS@OBS-BESANCON.FR

**\*\*Laboratoire d'Optique P. M. Duffieux, U. A. CNRS 214**

Faculté des Sciences et des Techniques - 25030 BESANÇON Cedex -FRANCE

**Abstract** - In some applications of time metrology, the measurement cannot be regularly spaced in time. This problem occurs for instance in the case of time transfer using GPS satellites in common view, since it is necessary to wait to see the same satellite from the two clocks being compared. This problem may also be observed in millisecond pulsar timing measurements since pulsars cannot be observed at regularly spaced dates. However, it is of importance to measure the time stability in these cases.

We developed several methods of spectral analysis of irregularly spaced data based upon the multi-variance method. The first class of methods involves different types of interpolation of the missing data, as described in Vernotte et al (1). The second class is an adaptation, in the multi-variance framework, of the lowest mode estimator introduced by Deeter and Boynton (2). The details of these methods are described in this paper. In order to compare their efficiency, we applied all of these methods over different types of simulated signals. The results are discussed in this paper.

## INTRODUCTION

The time and frequency team of the Observatoire de Besançon is involved in the spectral analysis of time sequences for several years. In this connection, we developed the multivariate method (4). However, such a method need regularly spaced timing data. Of course, this is not the case for millisecond pulsar timing sequences.

The first idea we got consists in interpolating the missing data in order to get a sequence of irregularly spaced data. Unfortunately, such a method may add erroneous information. Among four types of interpolation method, we finally keep only one of them. Following the works of Groth, Deeter and Boynton (2) developed a new method for the estimation of the power spectrum of pulsar timing data : the lowest mode estimator. This approach may be considered as an extension of the structure function concept, introduced by Lindsey and Chie, to the case of irregularly spaced timing data, since the lowest mode estimator may be optimized to one type of noise.

We have shown in a previous paper (3) the importance of the choice of the spectral density estimator (variance or structure function), particularly in the case of the multi-variance method (4). For instance, a very low frequency noise ( $f^3$  or  $f^4$  frequency noise which may occur in pulsar timing) could be confused with a linear frequency drift. In this case, a set of variances containing variances insensitive to linear frequency drift and converging for very low frequency noises associated with variances sensitive to linear drift must be chosen in order to determine both the low frequency levels and the linear drift coefficient.

## SPECTRAL ANALYSIS

The Power Spectral Density (PSD) of time error of an oscillator may be modeled as a sum of power laws :

$$S_x(f) = k_0.f^0 + k_{-1}.f^{-1} + k_{-2}.f^{-2} + k_{-3}.f^{-3} + k_{-4}.f^{-4}$$

Moreover, for the millisecond pulsar study, a  $f^{-5}$  phase noise should be added.

On the other hand, since the instantaneous normalized frequency deviation  $y(t)$  is the derivative of the time error  $x(t)$ , its PSD is given by :

$$S_y(f) = 4 \pi^2 f^2 S_x(f)$$

Then :

$$S_y(f) = \sum_{\alpha = -3}^{+2} h_{\alpha} f^{\alpha}$$

Estimating the PSD of  $y(t)$  consists then in measuring the noise levels  $h_{\alpha}$ .

The most obvious way to find the PSD is by taking the square of the Fourier transform (or the numerical Fast Fourier Transform) of the time series. Although straightforward, this is not always the best approach nor is it always a possibility. Variances are calculated using  $r^{\text{th}}$  order difference operators, one example being the two sample variance or Allan variance:

$$\sigma_y^2(\tau) = \frac{1}{2} \langle [y(t+\tau) - y(t)]^2 \rangle$$

Variances in general have a smoother response than the Fourier transform. However, a single variance is not well suited to measure all possible power laws. The multi-variance method was developed to better deal

with a wide range of power law noise spectra (4). This method uses different variances and gives the most probable set of power law noise levels (most probable in the sense of the least-squares). One severe drawback for the study of pulsar timing data is the fact that none of these methods can be used on non-equispaced data.

### INTERPOLATION OF THE MISSING DATA

In order to perform spectral analysis of non-equispaced data, we first reconstruct a sequence of equally spaced data by interpolating the missing data. The goal of the reconstruction is to get  $N$  equally time-spaced data from  $M$  irregularly spaced data without losing information or adding information.

From  $M$  data we measure  $\tau_0$ , the smallest interval between 2 consecutive timing data. This smallest interval  $\tau_0$  is the basic interval between all consecutive data of the reconstructed sequence.  $N$ , the new number of data, is then equal to the duration of the sequence divided by  $\tau_0$ . Actually, in order to easily perform Fast Fourier Transform over the reconstructed data, we choose as  $N$  the first power of two greater than the duration of the sequence divided by  $\tau_0$ . Let us define  $t_0$  the date of the first sample of the sequence and  $t_{M-1}$  the date of the last sample of the sequence,  $N$  is given by the relationship :

$$N = 2^{\text{int}\left(\log_2 \frac{t_{M-1} - t_0}{\tau_0}\right) + 1}$$

Generally, the available data are time error  $x(t)$  measurements between the oscillators and a reference oscillator. However, the time stability is mainly studied from the instantaneous normalised frequency deviation samples  $\overline{y_k}$ , obtained from the  $x(t)$  data by the relationship :

$$\overline{y_k} = \frac{x(t_k + \tau) - x(t_k)}{\tau}$$

Reconstructing equally spaced data from the  $x(t)$  data or from the  $\overline{y_k}$  samples yields different ways of reconstruction.

#### Reconstruction by linear interpolation of the $x(t)$ data

This first method keeps the same  $\overline{y_k}$  samples as in the original irregularly spaced sequence. The only difference this method yields, is the division of each initial  $\overline{y_k}$  sample into several  $t_0$ -long samples with the same value. Thus, the added information is the constancy of the frequency deviation during the initial samples.

#### Reconstruction of the $x(t)$ data by cubic spline functions

Obviously, the real frequency deviation  $y(t)$  is not constant over the time interval of each initial  $\overline{y_k}$  samples. In order to avoid this hypothesis of constant samples within each initial sample, it is possible to fit the  $x(t)$  data with cubic spline functions. The new  $\tau_0$ -long samples vary smoothly while preserving the same average over the initial samples. The added information is then an hypothesis of continuity (and derivability) of the  $\overline{y_k}$  samples, due to the continuous variation (derivability of second order) of  $x(t)$ .

Although the  $x(t)$  samples are strongly correlated for the low frequency noises, the hypothesis of continuous variation of the  $x(t)$  samples is completely wrong in the case of a white noise ! Since the types of frequency noises can vary from  $f^{-3}$  (only in the case of millisecond pulsars (5)) to  $f^{+2}$ , i. e. from  $f^{-5}$  to  $f^0$  phase noises, this method may be justified only for correlated  $x(t)$  data, but not in the case of a white phase noise ( $f^{+2}$  frequency noise).

#### Reconstruction by linear interpolation of the $\overline{y_k}$ samples

On the other hand, it is possible to reconstruct directly the  $\overline{y_k}$  samples by linear interpolation. The new  $\overline{y_k}$  sequence is then continuous but not derivable. The  $x(t)$  sequence is obtained by the relationship :

$$x(t_k + \tau_0) = x(t_k) + \tau_0 \cdot \overline{y_k}$$

In this case, the  $x(t)$  function is only derivable once. However, the hypothesis of continuity of the  $\overline{y_k}$  samples is wrong in the case of a white frequency noise ( $f^2$  phase noise) or higher frequency noise. This method may only be applied to low frequency noises.

#### Reconstruction of the $\overline{y_k}$ samples by cubic spline functions.

Theoretically, this method could only be justified for very low frequency noises ( $f^{-3}$  frequency noise) ; nevertheless we decided to observe the behaviour of such a method for all the types of noises in order to confirm our theoretical considerations.

We demonstrated (1) that the reconstruction by linear interpolation of the  $x(t)$  data gives the better results. In the following we will only use this reconstruction method.

## LOWEST MODE ESTIMATION

An other way for estimating the PSD of a sequence of irregularly spaced timing data consists in formulating a general method of analysis. Let  $x(t)$  represent the time series or noise process and  $\{h_1(t), \dots, h_j(t), \dots, h_n(t)\}$  be a set of  $n$  sampling functions. We can now define :

$$P(f_j) = \left[ \int_{-\infty}^{+\infty} h_j(t) x(t) dt \right]^2$$

where  $P(f_j)$  is an estimator of  $S_x(f)$ , the PSD of  $x(t)$ . This continuous function formula can easily be adapted to the case of discrete data. We have  $m$  discrete data obtained at times  $\{t_1, \dots, t_k, \dots, t_m\}$

$$h_j(t) = \sum_{k=1}^m h_{jk}(t_k) \delta(t-t_k)$$

denoting  $h_{jk} = h_j(t_k)$ , the estimation is given by :

$$P(f_j) = \left[ \sum_{k=1}^m h_{jk} x_k \right]^2$$

It can be shown that the expectation of  $P(f_j)$  for a large number of processes may be written as :

$$\langle P(f_j) \rangle = \int_{-\infty}^{+\infty} |H_j(f)|^2 S_x(f) df$$

where  $H_j(f)$  is the Fourier transform of the sampling function  $h_j(t)$ . If for example we choose  $e^{i2\pi f_j t}$  as  $h_j(t)$ , we see that  $H_j(f)$  is a delta function and  $P(f_j)$  is identified with  $S_x(f)$  at each  $f = f_j$  just as we expect. For a variance calculation  $h_j(t)$  is an  $r^{\text{th}}$  order difference function, and  $S_x(f)$  is related to  $P(f_j)$  via the transfer function  $H_j(f)$ . We may choose anything we like for  $h_j(t)$ , and as we shall see, a polynomial can give useful results. Regardless of our choice for  $h_j(t)$ ,  $P(f)$  is a filtered form of  $S_x(f)$ , related by the transfer function  $|H_j(f)|^2$ .

We can now state that the sampling function  $h_j(t)$  gives an unbiased estimate of  $S_x(f)$  if  $H_j(f)$  is centered on some nominal frequency  $f$  and has a narrow bandwidth. For a finite length data set,  $H_j(f)$  will not be a single peak but will have sidebands. A sinusoidal  $h_j(t)$  will no longer give a delta function and instead yields a sinc function. This does not pose too great a problem if  $S_x(f)$  is flat, i. e. a white noise spectrum. But if  $S_x(f)$  rises too steeply for low frequencies, i. e. a red noise spectrum, the sidebands in fact dominate the product of  $|H_j(f)|^2 \cdot S_x(f)$  and the integral diverges.

## The moment condition

To ensure the convergence of the estimator the moment condition is imposed. We want  $H_j(f)$  to satisfy

$$\int_{-\infty}^{+\infty} |H_j(f)|^2 f^{-\alpha} df \quad \text{finite}$$

where  $f^{-\alpha}$  is the red noise part of the power spectrum. If  $h_j(t)$  satisfies the following moment condition :

$$\int_{-\infty}^{+\infty} h_j(t) t^q dt = 0 \quad \text{for } q < n$$

it can be shown that the estimator will converge for  $\alpha < 2n+1$ . Thus for  $f^5$  noise ( $f^3$  frequency noise) the moment condition must be verified for  $n = 2$ . For discrete data the moment condition is

$$\sum_{k=1}^N h_{jk} t_k^q = 0 \quad \text{for } q \leq 2$$

Moreover, this condition implies  $h_j(t)$  is non-sensitive to quadratic phase drift, i. e. linear frequency drift. In fact we shall see that it is equivalent to having removed a linear and quadratic term from the time series.

## The orthogonality conditions

The choice of an estimator verifying the moment condition ensures that it will be unbiased on a red noise, since it will reject the (unavoidable) drifts. If we want it to be insensitive to the measurement white noise, and be mostly sensitive to the suspected red one, we are led to the orthogonality condition and to the lowest mode estimator (2).

A set  $\{h_1(t), \dots, h_N(t)\}$  of estimators applied over a measurement noise (assumed as white phase noise) has a covariance matrix, of which the elements may be defined as :

$$\langle h_i | h_j \rangle = \sum_{k=1}^m h_{ik} h_{jk} \sigma_k^2$$

where  $\sigma_k^2$  is the variance of the observational noise at the time  $t_k$  and  $m$  the number of time measurements. This is a kind of scalar product. Uncorrelated estimators must have a diagonal covariance matrix and are consequently "orthogonal".

On a red noise of known spectrum  $S_x(f) = k_{-2r} f^{2r}$ , the estimation is given by :

$$P(f_j) = \int_{-\infty}^{+\infty} |H_j(f)|^2 k_{-2r} f^{2r} df \quad (1)$$

Moreover, denoting  $h_j^{(-r)}(t)$ , the  $r^{\text{th}}$  integral of  $h_j(t)$ , its Fourier transform is given by :

$$h_j(t) = \frac{d^r}{dt^r} h_j^{(-r)}(t) \Leftrightarrow H_j^{(-r)}(f) = \frac{H_j(f)}{f^r} \quad (2)$$

Then, from (1) and (2), it becomes :

$$P(f_j) = k_{-2r} \int_{-\infty}^{+\infty} |H_j^{(-r)}(f)|^2 df$$

Thus, the covariances of the estimations for a  $f^{2r}$  noise defines a new  $r$ -scalar product :

$$\langle h_i^{(-r)} | h_j^{(-r)} \rangle$$

The independance of the estimations for a  $f^{2r}$  noise leads us to a new diagonal  $r$ -covariance matrix.

The choice of sets of estimators, orthogonal on both measurement noise and red noise has the advantage that we can estimate the variance of any sum of them, as the sum of individual variances (assuming that uncorrelation implies independence, even if opposite examples are well known for non-gaussian distributions).

Deeter and Boynton suggest a basis of  $m$  orthogonal estimators for measurement noise  $h'_i(t)$ , using  $m$  orthonormal polynomials  $p_i(t)$  of ascending degrees  $i$ , defined iteratively by :

$$\sum_{k=1}^m p_i(t_k) p_j(t_k) \sigma_k^2 = \delta_{ik}$$

and

$$h'_i(t) = \sum_{k=1}^m p_i(t) \sigma_k^2 \delta(t - t_k)$$

$\{h'_1(t), \dots, h'_m(t)\}$  are orthogonal on measurement noise even for irregularly sampled data. It may be demonstrated that each polynomial  $p_i(t)$  verifies the moment condition up to an order equal to  $i$ , the degree of this polynomial. Thus, removing the first  $n+1$  polynomials ensures that they respect the moment condition up to order  $n$ .

From  $\{h'_{n+1}(t), \dots, h'_m(t)\}$  we derive an  $r$ -orthogonal set of estimators by rotating the  $r$ -covariance matrix to its principal axes, obtaining  $\{h_{n+1}(t), \dots, h_m(t)\}$  :

$$h_j(t) = \lambda_j^{-1/2} \sum_i u_{ji} h'_i(t)$$

$u_{ji}$  the eigen-vectors components  
with  $\lambda_j$  the eigen values of the  $r$ -covariance matrix

Thus :

$$\langle h_i^{(-r)} | h_j^{(-r)} \rangle = \delta_{ij}$$

### Practical calculation of the lowest mode estimator

We want to study  $f^3$  and  $f^4$  (if it exists) frequency noise in pulsar timing, and we assume that no "redder" noise is present. We have 165 residual times of the monitoring of PSR 1937+21, obtained at Arecibo

Observatory (courtesy Joe Taylor). Our estimators must then verify the moment condition up to order 2 in order to converge up to  $f^4$  frequency noise ( $f^6$  phase noise) :

$$\sum_{k=1}^{165} h_{ik} t_k^q = 0 \quad \text{for } 0 \leq q \leq 2$$

On the other hand, the a priori knowledge of the spectrum of the residuals leads us to optimize the estimator for  $f^2$  frequency noise by diagonalizing the 2<sup>nd</sup> order covariance matrix.

The computation of all 165  $p_i(t_k)$  requires a large number of significant digits, but once it is done we easily get the 162 doubly orthogonal estimators  $\{h_0, \dots, h_{161}\}$ , and we normalize them so that  $\langle h_i^{(-2)} | h_j^{(-2)} \rangle = \delta_{ij}$ .

It may be demonstrated that the polynomials  $h_j(t)$  estimates the PSD of the derivative of  $y(t)$ , for the frequency  $f_j$ .

### DISCUSSION

In order to compare the performances of these methods, we generated simulated signals of 8192 regularly spaced data. We kept only 165 data, spaced as in a real millisecond pulsar sequence.

For each choice of noise level, 100 realizations of the same signal were generated. In the following, we will call *reference* the measurements obtained with the multivariance method over the whole sequence of 8192 equally spaced data. Of course, the two methods (linear interpolation and lowest mode estimator) are applied over the sequences of 165 irregularly spaced data. Thus, it is possible to see the effects due to the missing data for both methods.

### Pure noises

Table I shows the results obtained for signals composed of only one type of noise ( $f^3$ ,  $f^2$  or  $f^{+2}$  frequency noise). For  $f^3$  frequency noise (see figure 1), the lowest mode estimator (LME) yields a lower dispersion than the linear interpolation (LI). This dispersion is closer to the reference dispersion than to the LI one. The LME is actually optimized for very low frequency noises.

Noise	Method	Reference	Linear interpol.	Lowest mode
$f^3$		17 %	45 %	27 %
$f^2$		4.2%	21 %	27 %
$f^{+2}$		0.6%	18 %	21 %

Table I : Dispersion of the noise level measurements for pure noises. These dispersions have been averaged over 100 realizations of the same noise. The reference is obtained for sequences of 8192 regularly spaced data

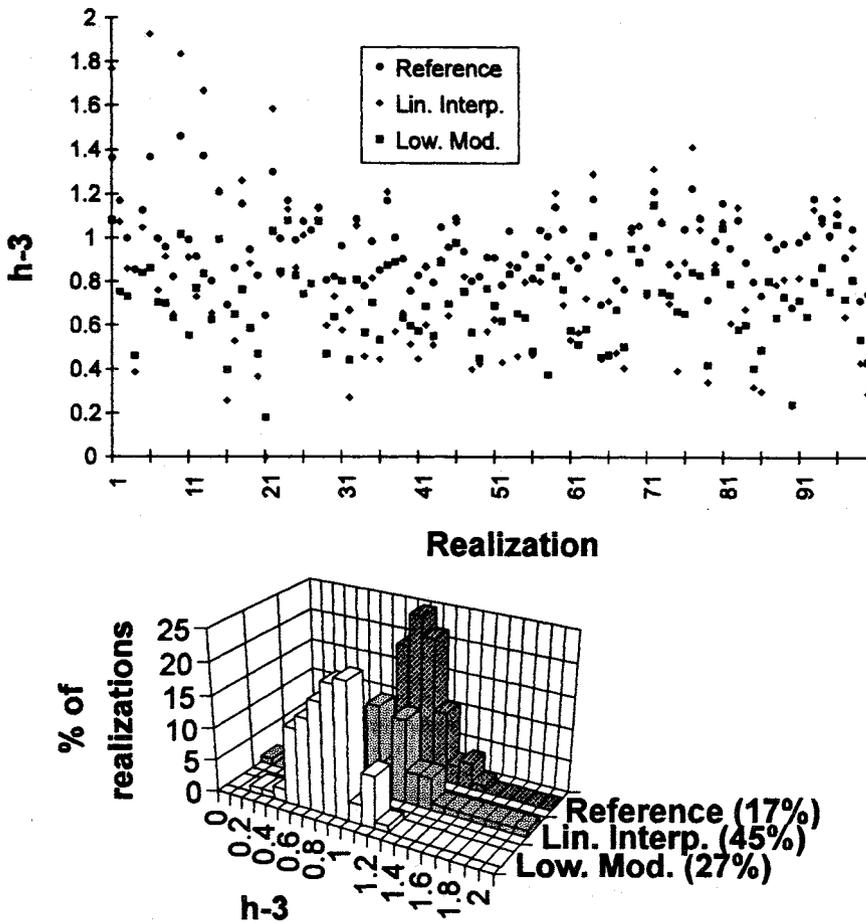


Figure 1 : Comparison of the results obtained over 100 realizations of pure  $f^3$  frequency noise with 8192 regularly spaced data (reference), and 165 irregularly spaced data with the linear interpolation and the lowest mode estimator.

For  $f^2$  and  $f^{+2}$  frequency noises, the LI gives better results than the LME. Obviously, the higher frequency noise we study, the more important the effects of the missing data are.

**Mixed noises**

Table II shows the results obtained for signals composed of 2 or 3 types of noises. The differences between the LME and the LI are not very significant. However this tables seems to confirm that the LME is a bit better for the low frequency noises.

**Dynamics**

In order to go further, we studied the dynamics of the methods. The dynamics may be defined as the range in which a noise level can vary without

swamping the other ones or being swamped by the other ones.

Thus, for a signal composed of  $f^2$  and  $f^{+2}$  frequency noise, we varied the ratio  $h_{-2}/h_{+2}$  (see figure 2).

In the frequency range we used, the two types of noise are about of the same order of magnitude for a ratio  $h_{-2}/h_{+2} \approx 5 \cdot 10^{-9}$ . In this case, the dispersion of the LI (35%) is lower than the one of the LME (40%). However, the dispersion of the LI increases more quickly than the one of the LME when the ratio  $h_{-2}/h_{+2}$  changes.

For instance, if we only want a rough estimate (i. e. the order of magnitude) of the noise levels (e. g. a uncertainty of 100%), the range in which the ratio

$h_{-2}/h_{+2}$  can vary is wider for the LME ( $2 \cdot 10^{-11} \rightarrow 3 \cdot 10^{-7}$ , dynamics  $\approx 40$ dB) than for the LI ( $10^{-10} \rightarrow 1.5 \cdot 10^{-7}$ , dynamics  $\approx 30$ dB).

Thus, the LI may provide more accurate measurements of the noise levels, but the dynamics of the LME is greater.

**Separability**

An other useful criterion is the number of null measurements : for instance, since the  $f^2$  frequency noise exists in the signal, if a method yields a null measurement, it means that this method confused this noise with another one ( $f^1$  or  $f^3$  frequency noise). So, this criterion informs us about the separability of this method.

Figure 3 shows that, if we are interested in a 20% of null measurement, i. e. 1 chance over 5 to confuse a

Signal	Noise	Method	Reference	Linear interpol.	Lowest mode
$h_{-2}f^2+h_{+2}f^{+2}$ $h_{-2}=2E-10 ; h_{+2}=1$	h-2		53 %	69 %	59 %
	h+2		0.5 %	24 %	24 %
$h_{-2}f^2+h_{+2}f^{+2}$ $h_{-2}=4E-9 ; h_{+2}=1$	h-2		22 %	36 %	41 %
	h+2		0.8 %	35 %	39 %
$h_{-3}f^3+h_{-2}f^2+h_{+2}f^{+2}$ $h_{-3}=1E-10$ $h_{-2}=5E-8 ; h_{+2}=1$	h-3		51 %	55 %	53 %
	h-2		52 %	100 %	126 %
	h+2		0.6 %	74 %	73 %

Table II : Dispersion of the noise level measurements for mixed noises. These dispersions have been averaged over 100 realizations of the same noise. The reference is obtained for sequences of 8192 regularly spaced data.

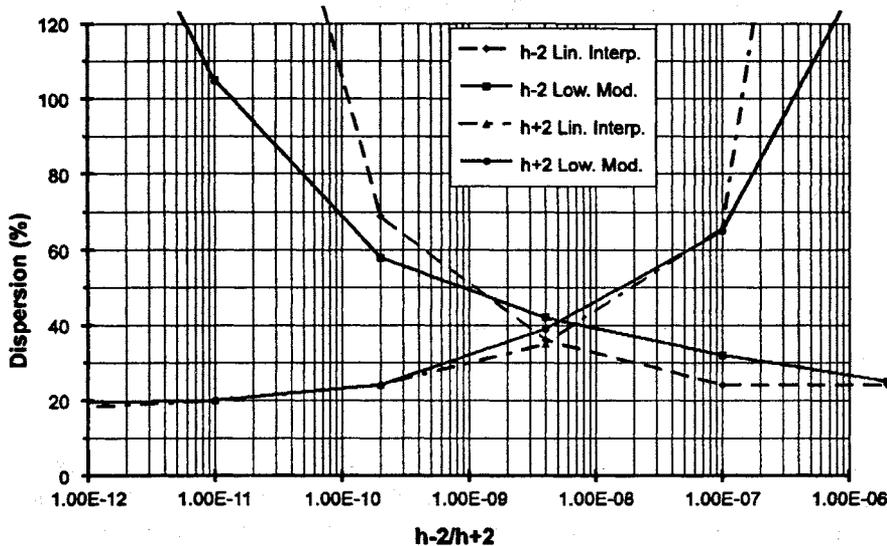


Figure 2 : Dynamics. Dispersion of the measurements versus the ratio  $h_2/h_{+2}$

noise type with another one, the range is wider for the LME (40dB) than for the LI (30dB).

## CONCLUSION

The first conclusion of this study is that these two methods shows performances similar. The linear interpolation may be more accurate, especially for high frequency noises, but its dynamics for rough estimates (30 dB) is smaller than the dynamics of the lowest mode estimator (40 dB). Moreover, the separability of the lowest mode estimator seems also to be higher than the the one of the linear interpolation.

Recently, using the lowest mode estimator we set an upper limit (at 95% of confidence) to the  $f^3$  frequency noise level of the millisecond pulsar PSR 1855+09 (5).

This noise level is too low to be directly measured, thus, in this case the lowest mode estimator must be used because of its wider dynamics.

Concerning the prospects of this work, we will improve this method by using it in the multi-variance framework. Thus, we will define several lowest mode estimators and use them over the same signal.

First, these estimators may be optimized for different types of noises (the exponent must be even).

Secondly, these estimators may be defined with different moment conditions.

Thus, it is possible to build up to 8 different estimators. Using all these estimators in a multi-estimator system yields the

advantage of the multi-variance method (4) : a wider frequency range, a better accuracy for the noise level measurement and an estimation of the uncertainties about each noise level measurement.

With such a system it should be possible to find a more stringent limit on the  $f^3$  frequency noise level of millisecond pulsar and, perhaps, a lower limit.

## REFERENCE

1. Vernotte F, Zalamansky G and Lantz E, 1993, "Noise and drift analysis of non-equally spaced timing data", *PTTI 93*, 379-388.
2. Deeter JE and Boynton PE, 1982, "Techniques for the estimation of red power spectra. I. Context and methodology", *Ap J*, **261**, 337-350.
3. Vernotte F, Lantz E, McHugh M, Zalamansky G, 1995, "Cut-off frequencies and noise power law model of spectral density : adaptation of the multi-variance method using the structure function approach", *9th EFTE*, 373-376.
4. Vernotte F, Lantz E, Gros Lambert J, Gagnepain JJ, 1993, "Oscillator noise analysis : multi-variance measurement", *IEEE Trans. on Instrum. and Meas.*, **IM-42.2**, 342-350.
5. McHugh M, Zalamansky G, Vernotte F, Lantz E, 1996, "Pulsar timing and the upper limits on a gravitational wave background : a Bayesian approach.", submitted to *Phys. Rev. D*.

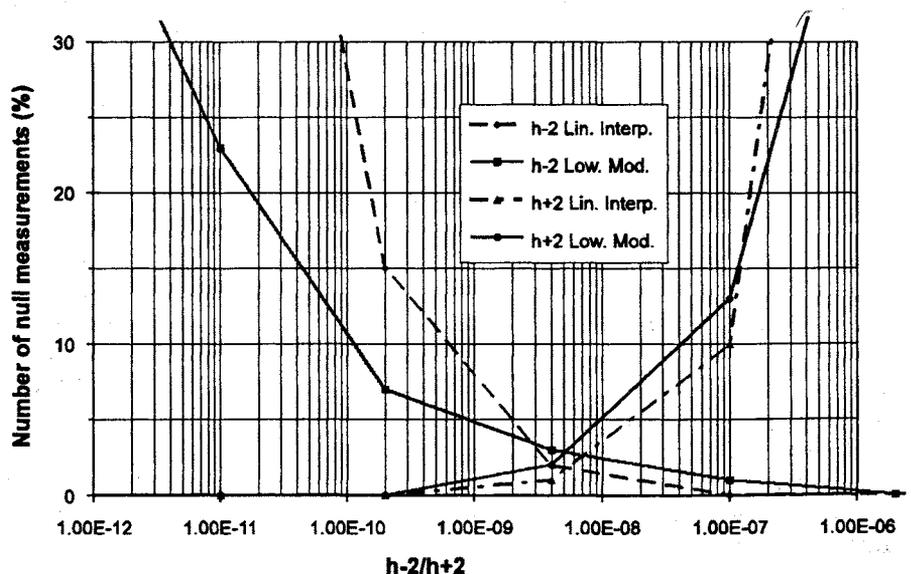


Figure 3 : Separability. Number of null measurements versus the ratio  $h_2/h_{+2}$

## A QUANTUM CHAOLOGIST'S VIEW OF QUARTZ RESONATORS

Mark Oxborrow

Department of Physics, Theoretical Physics,  
1 Keble Road, Oxford OX1 3NP, England

and

Center for Chaos and Turbulence Studies (CATS), Niels Bohr Institute,  
Blegdamsvej 17, DK-2100, Copenhagen Ø, Denmark

Physicists have, over the last two decades, gained many non-trivial insights into the behaviour of quantum systems whose corresponding classical dynamics are chaotic; the study of such systems is called "quantum chaology". The main function of this paper is (i) to tell the reader about the **existence** of this knowledge and (ii) to explain **how** it relates to the discipline of frequency control. Resonators are analogous to closed quantum systems, in the sense that the trajectory of a high-frequency wavepacket bouncing between the bounding surfaces of a resonator is similar to the trajectory of a quantum wave-particle bouncing inside the walls of a confining potential. Through this analogy, the same universal laws that control the behaviour of closed quantum systems should be applicable to the sorts of resonators that are used for frequency-control applications.

Any finite resonator has a discrete set of normal modes. Quantum chaology places certain restrictions both on the way in which the frequencies of these normal modes are distributed in frequency space and on the way in which these frequencies change as the body is perturbed (by, say, a change in its operating temperature). These restrictions motivate the use of "chaotic" resonators, which are inherently immune to activity dips and sudden shifts in frequency.

An experiment is described, which verifies the predictions of quantum chaology for a matchbox-sized block of crystalline quartz.

### QUANTUM CHAOLGY

Before going any further, it is perhaps worth stressing that quantum chaology has absolutely nothing to do with the sort of chaos that is induced through non-linearities. This paper concerns the behaviour of linear resonators. Since quantum chaology is a rather obscure subject, a brief introduction to it is given forthwith.

"Quantum chaology" is the study of quantum sys-

tems, which, in the classical limit, exhibit chaotic dynamics. For the purposes of this definition, a "quantum system" is any finite system that can be described by a wave equation; the words "classical limit" denote the limit in which the characteristic wavelengths of the solutions to this wave equation are much smaller than the size of the system; "chaotic dynamics" implies that the evolution of the system is critically dependent on the initial conditions. Systems that exhibit quantum chaology include:

- 1) The helium atom;
- 2) A hydrogen atom in a strong magnetic field;
- 3) An irregularly shaped quantum dot;
- 4) An irregularly shaped microwave-cavity resonator;
- 5) An irregularly shaped quartz resonator.

It is perhaps worth stating here that most ideal shapes, and almost all real ones, are irregular. The articles by Berry[1] and Keating[2] explain the philosophy of quantum chaology in more detail. Many of the quantitative results in quantum chaology stem from Random Matrix Theory[3, 4, 5, 6] and Periodic Orbit Theory[7, 8, 9, 10].

A finite object made from a continuous, elastic, loss-free medium has the potential to oscillate in any one of a countably infinite set of normal modes, provided the object is subject to loss-free boundary conditions. Despite this incircumventable fact, most high-Q (almost loss-free) resonators are designed to oscillate at a single frequency. Any resonances that lie, in frequency space, near to the "desired" one are deemed to be "unwanted"[11, 12, 13], since their presence can give rise to activity dips and/or sudden frequency shifts. As regards the design of reliable frequency standards, it is perhaps as important to understand the nature and origin of unwanted modes as it is to understand the behaviour of desired ones.

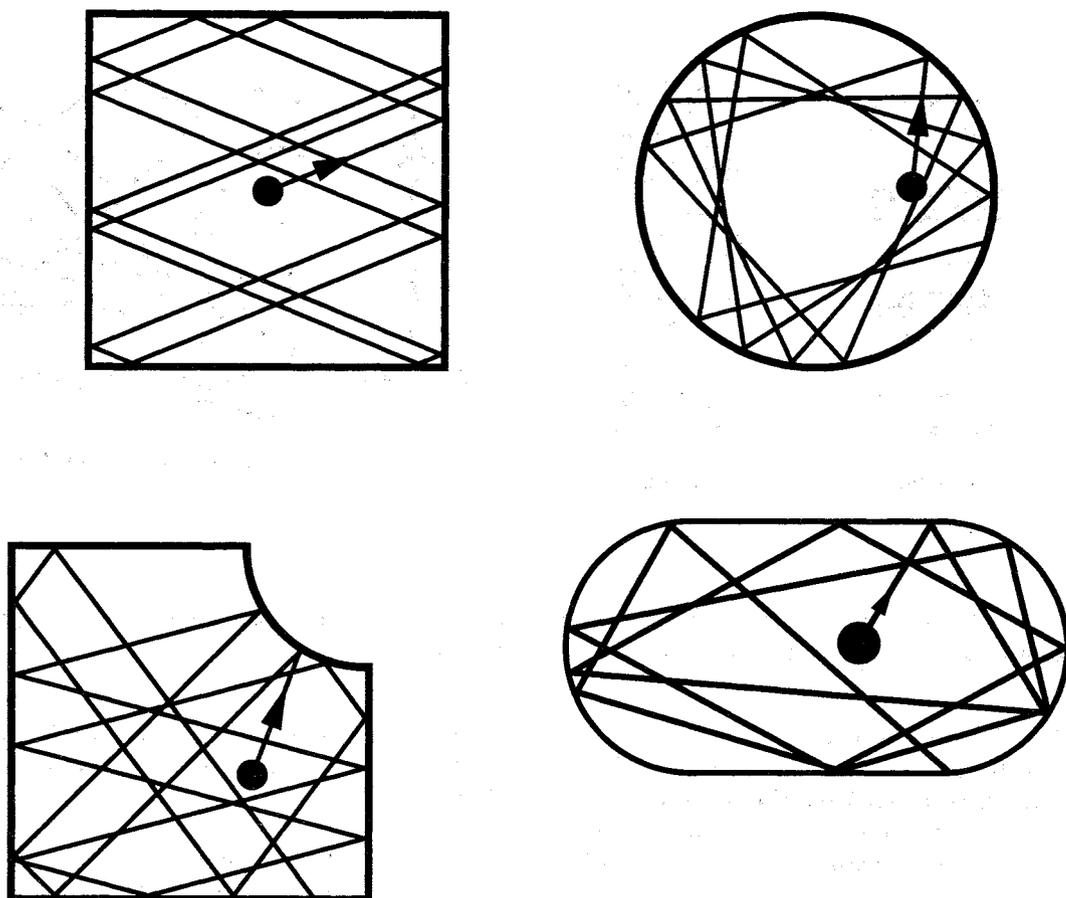


Figure 1: From top left clockwise: the square billiard; the circle billiard; the “Bunimovich” billiard; the “Sinai” billiard.

The modelling of real, high-Q, bulk-wave, elastomechanical resonators is fraught with problems. Experimental uncertainties often limit the accuracy to which a resonator can be specified, as regards its shape, orientation, and the properties of the material from which it is made. Furthermore, there exist no truly effective/efficient algorithms for calculating the frequencies at which an arbitrarily shaped/oriented three-dimensional object resonates. The list of solvable (integrable) systems includes the isotropic sphere[14] the infinite isotropic cylinder[15, 16] but little else[17, 18]. Though there exist a plethora of numerical methods for calculating the frequencies at which solid elastic bodies resonate[19, 20, 21, 22, 23, 24], the performance/convergence/efficiency of these methods is poor for (generic) systems, that cannot be represented as a solvable system plus a small perturbation.

Quantum chaology concerns the way in which resonances are distributed in frequency space; the list of frequencies at which an oscillator resonates is regarded as a data set, which exhibits certain statistical properties. Even though it is often difficult, if not impossible, to predict the frequencies of in-

dividual resonances, the form of certain statistical measures, defined as sums over all of the resonance frequencies, can be predicted exactly. The frequencies at which an object resonates can be regarded as “levels” in frequency space; quantum chaology deals with the statistics of these levels. Though the idea of “level statistics” might sound a bit devious and contrived, it leads to some powerful results; these are presented in the next section.

## PREDICTIONS

This section describes those aspects of quantum chaology that are relevant to the design of resonators. To date, most of the relevant theoretical work has been done on ideal, two-dimensional, scalar, isotropic, hard-walled, quantum billiards. These “toy” resonators are both easy to visualize and amenable to computer-based calculations.

The behaviour of the (single) quantum wave-particle inside such a system is governed by Shrödinger’s equation:

$$-\frac{\hbar^2}{2m}\nabla^2\Psi(\mathbf{r},t) + V(\mathbf{r})\Psi(\mathbf{r},t) = -i\hbar\frac{\partial\Psi(\mathbf{r},t)}{\partial t}, \quad (1)$$

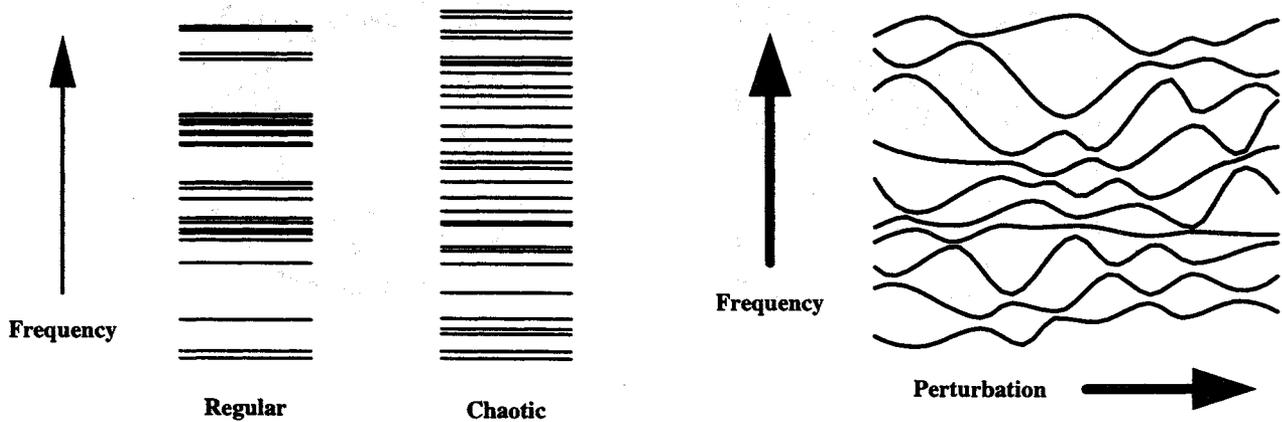


Figure 2: From left to right: comparison between the resonance spectra of a regular and a chaotic system; level dynamics of a resonance spectrum in response to a perturbation.

where  $\mathbf{r}$  is the two-dimensional position vector,  $t$  is the time,  $m$  is the mass of the wave-particle,  $\nabla^2$  denotes the two-dimensional Laplacian operator, and  $\Psi(\mathbf{r}, t)$  is the wavefunction of the quantum wave-particle.  $V(\mathbf{r})$  is the potential of the wave-particle at position  $\mathbf{r}$ ; it is zero everywhere inside the billiard and arises abruptly to infinity at the walls of the billiard. As a result, the wavefunction  $\Psi(\mathbf{r}, t)$  vanishes at the walls (“Dirichlet” boundary conditions). The billiard is thus defined solely by the shape of its walls. Such a system has an infinite discrete set of standing-wave solutions (eigenfunctions)  $\Psi_n(\mathbf{r}, t) = \Psi_n(\mathbf{r})e^{iE_n t/\hbar}$ , where  $E_n$  is the eigenvalue of the  $n$ -th eigenfunction.

In the classical ( $\hbar \rightarrow 0$ ) limit, the particle inside the billiard will travel in straight lines at a constant speed, and bounce specularly off the walls. In the case of elastodynamical systems, the classical limit corresponds to the situation where the motion of the system’s medium can be represented as the propagation of high-frequency, short-wavelength wave-packets, where the trajectories of these packets obey the laws of geometrical acoustics[18, 25]. Waves of this sort are often referred to as “acoustic beams”. Such conditions are realised in (bulk-wave) acoustic delay lines[26, 27, 28, 29], which once had important military/commercial applications.

### Universal Level Statistics

For certain simple shapes of boundary, the trajectory of a particle inside a billiard is **regular**—as is the case for the top two billiards in Fig. 1; these are known as integrable (or regular) billiards. For other, more complicated shapes of boundary, the particle’s

trajectory is **chaotic**—as is the case for the bottom two billiards in Fig. 1; these are known as chaotic billiards.

One might ask whether the regularity/chaoticity of the classical trajectories affects the way in which a billiard’s eigenvalues (resonance frequencies)  $E_n$  are distributed in energy space (frequency space). The answer to this question is quite remarkable: **all** integrable billiards exhibit the **same** sort of level statistics, known as Poissonian statistics; **all** chaotic billiards obey the **same** sort of level statistics, known as GOE statistics[5]. Thus, the level statistics of a given system does not depend on the precise details of the system; rather it depends only on whether or not the system’s classical trajectories are chaotic. This amazing result is known as the **universality** of level statistics.

### Level Repulsion

The left half of Fig. 2 compares a typical series of eigenvalues (resonance frequencies) for a regular (integrable) system with a typical series of eigenvalues for a chaotic system. Note that the spacing between the eigenvalues of the regular system is relatively more erratic, where the separation between consecutive eigenvalues can be both extremely large and extremely small. In regular systems, the desired modes often lie perilously close to unwanted ones. In contrast, the eigenvalues of the chaotic system are spread much more evenly in frequency space. This property is due to **level-repulsion**; this effect is the hall-mark of “chaotic” resonators. In any system whose classical trajectories are chaotic, level repulsion prevents unwanted modes from contaminating

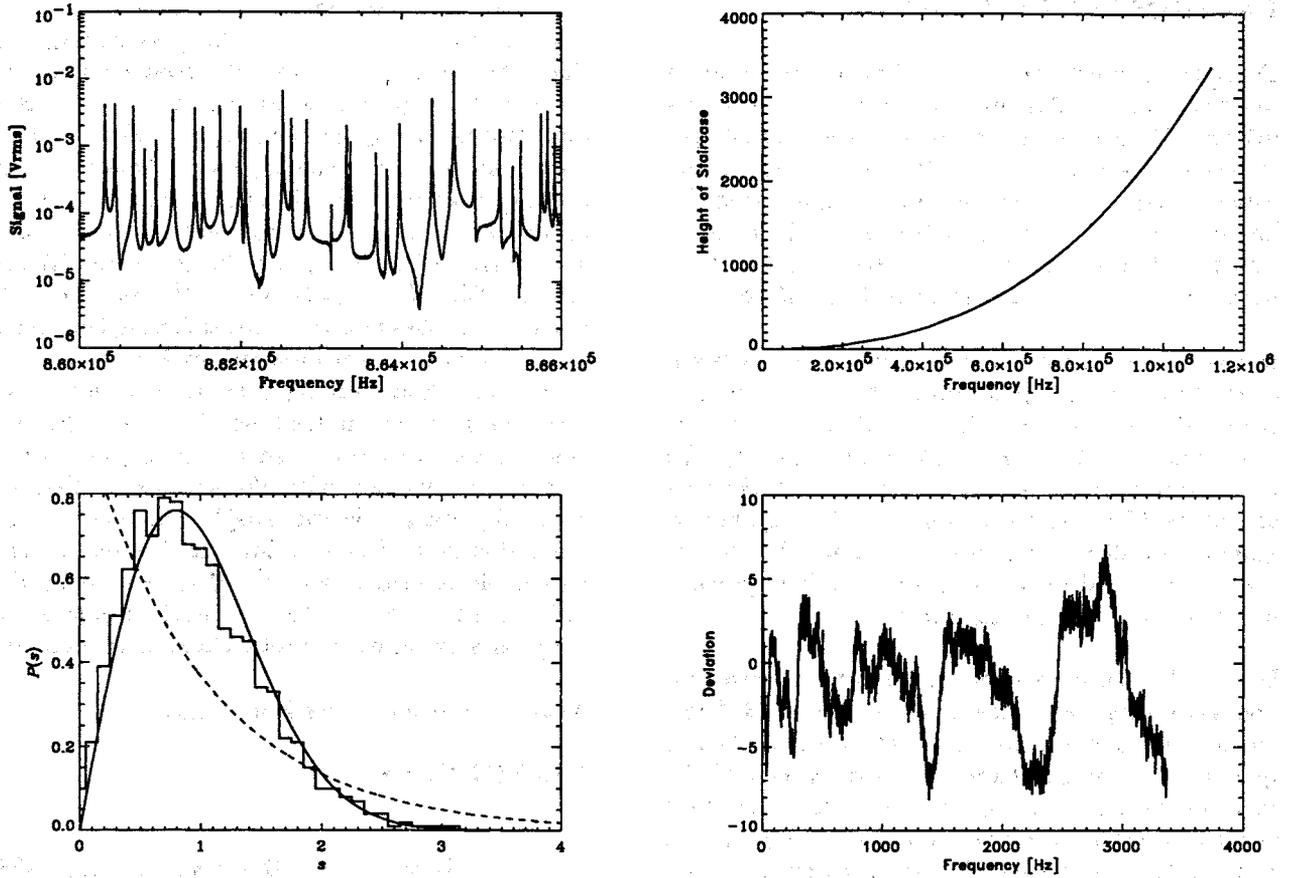


Figure 3: From top left clockwise: a segment of the raw spectrum; the staircase function; the deviation of the staircase function from a polynomial best-fit curve, the unfolded nearest-neighbour spacing distribution.

the desired one(s).

Though not understood in terms of quantum chaos, level repulsion has been used in the quartz industry for many decades: plate resonators are often bevelled/contoured on one or both sides. The act of contouring a plate does not just facilitate energy trapping; it also makes the resonator's classical trajectories more chaotic. Through the effect of level repulsion, this ensures that the area in frequency space around the desired mode is devoid of (unwanted) resonances.

### Level Dynamics

The properties of any real resonator vary with time (ageing), temperature, magnetic field, and other variables. A change in any one of these parameters perturbs the resonator, and thus alters the frequencies at which it resonates. The variation of these frequencies (levels) as a function of the perturbation constitutes a type of dynamics; this is depicted in the right half of Fig. 2. Here, it is assumed that the resonator's classical trajectories are chaotic; as a result the levels obey a **no-crossing** law. The properties of the level dynamics can be described in terms of

certain statistical measures (e.g. the autocorrelation function of the rate of change of frequency with respect to the perturbation). Though these measures have only been calculated for abstract theoretical models[30, 31, 32, 33, 34, 35, 36], it is conjectured that they are universal: **every** resonator whose classical trajectories are chaotic exhibits the **same** sort of level dynamics, **regardless** of the sort of perturbation that is applied. This hypothesis has been partially verified in the case of a hydrogen atom in variable magnetic field[37]; to date, no equivalent experimental work has been done for elastodynamical systems.

### Wavefunctions and Energy Trapping

In physics, the concept of energy trapping is known as **localization**. Standing waves  $\Psi_n(\mathbf{r})$  corresponding to resonators whose classical trajectories are chaotic have universal properties and exhibit certain characteristic features ("scars"). It is possible to localize eigenfunctions through chaotic scattering and/or disorder. Various papers[38, 39, 40] discuss these issues in detail.

## EXPERIMENT EVIDENCE

Do all these unbelievably general predictions work in practice? In two dimensions, for standing-wave solutions, Maxwell's equations and the Schrödinger equation both reduce to the scalar Helmholtz equation[41]. This correspondence implies that the law of universal spectral statistics should hold for suitably shaped microwave cavities –as has been born out by various experiments[41, 42, 43, 44].

In contrast, no direct correspondence exists between elastodynamics and quantum mechanics, other than in trivial (one-dimensional) cases. Given the present, non-existent state of relevant theory [45, 46], the question as to whether, and to what extent, the laws of universal level statistics apply to elastodynamical systems is still an open one. The main objective of the experimental work described below is to go some way towards resolving this matter.

Using the technique of **resonant ultrasound spectroscopy**[49], I measured the lowest 3364 resonances of a matchbox-sized block of crystalline quartz. The block's shape was (approximately) the three-dimensional analogue of the Sinai billiard that is depicted in the bottom left quadrant of Fig. 1, i.e. a rectangular parallelepiped with a spherical octant removed from one of its corners. The experimental set-up was similar to the instruments described in the articles by Migliori et al[50] and Maynard[49], but differed in some important technical details[51].

I have naively compared the experimental spectrum to the predictions of Random Matrix Theory, for the so-called GOE (Gaussian Orthogonal Ensemble)[4, 5, 6]. These predictions correspond to those of Periodic Orbit Theory, in the case of a fully chaotic, time-reversible dynamics. Several papers have undertaken similar analyses[47, 45, 46, 48], and the reader is referred to them for the mathematical details.

The top left quadrant of Fig. 3 shows a small segment of the raw experimental data; the entire spectrum was judged to contain 3364 resonances.

The top right quadrant of Fig. 3 shows the "staircase function" of the experimental spectrum, whose value at a given frequency  $f$  is equal to the the number of resonances with frequencies lower than  $f$ . A third-order polynomial in frequency ( $a + b f + c f^2 + d f^3$ ) was fitted (least squares) to this staircase function. The bottom right quadrant of Fig. 3 shows the deviation of the experimental staircase function from this best-fit polynomial; the curve reveals certain periodic oscillations, which are related to the Periodic Orbit Theory[46].

The list of resonance frequencies was then "unfolded"[6] using the best-fit polynomial, such that the average density of resonances in the "stretched" frequency space was every unity. Random Matrix Theory predicts that the distribution of the distances  $s$  (in frequency space) between neighbouring resonances should obey a universal (GOE) curve, which can be approximated, to a reasonable degree of accuracy, by the so-called "Wigner surmise"[6]:  $P(s) = (\pi/2) s e^{-\pi s^2/4}$ . Note that this function vanishes at  $s = 0$ , thus enforcing level repulsion. The bottom left quadrant of Fig. 3 shows the nearest-neighbour spacing distribution (a histogram, solid lines) for the unfolded list of resonance frequencies, together with the Wigner surmise (solid curve), and, for comparison with the spectral statistics of classically integrable systems[6, 10],  $P(s) = e^{-s}$  (dashed curve). The experimental distribution is in reasonable agreement with the Wigner surmise, thus indicating that the "3-D Sinai" block of quartz is an irregular system, with chaotic classical trajectories.

More experimental work is in progress.

## REFERENCES

- [1] M.V. Berry, "The Bakerian Lecture, 1987: Quantum chaology", *Proc. Roy. Soc. Lond. A*, **413**, 183-198
- [2] J. Keating, 1993, *The quantum mechanics of chaotic systems*, in "The Nature of Chaos", ed. T. Mullin, pp. 282-310, Oxford University Press
- [3] C.E. Porter, 1965, "Statistical theories of spectra: Fluctuations", Academic Press, New York
- [4] M.L. Mehta, 1991, "Random Matrices", 2nd edn., Academic Press, Boston
- [5] O. Bohigas, M.-J. Giannoni, and C. Schmit, 1984, "Characterization of Chaotic Quantum Spectra and Universality of Level Fluctuation Laws", *Phys. Rev. Lett.*, **52**, 1-4
- [6] O. Bohigas and M.-J. Giannoni, 1984, *Chaotic Motion and Random Matrix Theories*, in "Chaotic Motion and Random Matrix Theory", Lecture Notes in Physics Vol. 209, p. 1, Springer-Verlag, Berlin
- [7] M.C. Gutzwiller, 1971, *J. Math. Phys.*, **12**, 343-58
- [8] M.C. Gutzwiller, 1990, "Chaos in Classical and Quantum Mechanics", Springer-Verlag, New York
- [9] M.C. Gutzwiller, 1991 *The semi-classical quantization of chaotic Hamiltonian systems*, "Les

- Houches Lecture Series Vol. 52", eds. M.J. Gianoni, A. Voros, and J. Zinn-Justin, pp. 201-49, North-Holland, Amsterdam
- [10] M.V. Berry, 1985, "Semiclassical theory of spectral rigidity", Proc. R. Soc. Lond. A, **400**, 229-251
- [11] M. Onoe, 1964, "Analysis and Detection of Unwanted Modes in Quartz Crystal Units for Oscillators", J. Inst. Elec. Commun. Engrs. (Japan), **47**, pp. 82-93
- [12] J.G. Miller and D.I. Bolef, 1970, "Sampled-cw Study of Inhomogeneous Ultrasonic Responses in Solids", J. Appl. Phys., **41**, 2282-93
- [13] H. Bahadur and R. Parshad, "A New Method for Determination of the Frequency Spectrum of a Quartz Crystal and Its Applications", IEEE Trans. Sonics and Ultrason., **SU-25**, pp. 24-9
- [14] H. Lamb, 1882, London Math. Soc. Proc., **13**
- [15] L. Pochhammer, 1876, J.f. Math. (Crelle), **Bd. 81.**, 324
- [16] C. Chree, 1889, "The Equations of an Isotropic Elastic Solid in Polar and Cylindrical Coordinates, their Solution and Application", Cambridge Phil. Soc. Trans., **14**, 250-369
- [17] A.E.H. Love, 1944, "A Treatise on the Mathematical Theory of Elasticity", Dover, New York
- [18] K.F. Graff, 1991, "Wave Motion in Elastic Solids", Dover, New York
- [19] H. Ekstein and T. Schiffman, 1956, "Free Vibrations of Isotropic Cubes and Nearly Cubic Parallelepipeds", J. Appl. Phys., **27**, 405-412
- [20] H.H. Demarest, Jr., "Cube-Resonance Method to Determine the Elastic Constants of Solids", J. Acoust. Soc. Am., **49**, 768-75
- [21] I. Ohno, 1976, "Free Vibration of a Rectangular Parallelepiped Crystal and Its Application to Determination of Elastic Constants of Orthorhombic Crystals", J. Phys. Earth, **24**, 355-79
- [22] I. Ohno, S. Yamamoto, O. Anderson, and J. Noda, 1986, "Determination of Elastic Constants of Trigonal Crystals by the Rectangular Parallelepiped Method", J. Phys. Chem. Solids, **47**, 1103-8
- [23] R.C. Peach, 1988, "Rigorous Two-Dimensional Equations for the Analysis of Contoured Crystal Resonators", Proceedings of the 42nd Annual Frequency Control Symposium, 38-44
- [24] W.M. Visscher, A. Migliori, T.M. Bell, and R.A. Reinert, "On the normal modes of free vibration of inhomogeneous and anisotropic elastic objects", J. Acoust. Soc. Am., **90**, 2154-60
- [25] L. Couchman, E. Ott, and T.M. Antonsen, Jr., 1992, "Quantum chaos in systems with ray spitting", Phys. Rev. A, **46**, 6193-210
- [26] L.G. Merkulov and L.A. Yakovlev, 1962, "Propagation and Reflection of Ultrasonic Beams in Crystals", Soviet Physics - Acoustics, **8**, 152-155
- [27] V.J. Hammond, 1962, "Quartz Delay Lines ... the state of the Art", British Communications and Electronics, February 1962, pp. 104-10
- [28] J.S. Palfreeman, 1965, "Acoustic Delay Lines—A Survey of Types and Uses", Ultrasonics, January-March 1965, pp. 1-8
- [29] R.W. Gibson, "Solid Ultrasonic Delay Lines", Ultrasonics, April-June 1965, pp. 49-61
- [30] B.D. Simons and B.L. Altshuler, 1993, "Universal Velocity Correlations in Disordered and Chaotic Systems", Phys. Rev. Lett., **70**, 4063
- [31] B.D. Simons and B.L. Altshuler, 1993, "Universalities in the spectra of disordered and chaotic systems", Phys. Rev. B, **48**, 5422
- [32] M. Faas, B.D. Simons, X. Zotos, and B.L. Altshuler, 1993, "Universality in the spectra of strongly correlated systems", Phys. Rev. B, **48**, 5439
- [33] J. Zakrzewski and D. Delande, 1993, "Parametric motion of energy levels in quantum chaotic systems. I. Curvature distributions", Phys. Rev. E, **47**, 1650
- [34] J. Zakrzewski and D. Delande, 1993, "Parametric motion of energy levels in quantum chaotic systems, II. Avoided-crossing distributions", Phys. Rev. E, **47**, 1665
- [35] D. Delande and J. Zakrzewski, 1993, "Statistical Approach to Parametric Motion of Quantum Levels", J. Phys. Soc. Jpn., **63 Suppl. A**, 101
- [36] J. Zakrzewski, 1995, "On the Universal correlations in disordered and chaotic systems", Z. Phys. B, **98**, 273
- [37] B.D. Simons, A. Hashimoto, M. Courtney, D. Kleppner, and B.L. Altshuler, 1993, "New Class of Universal Correlations in the Spectra of Hydrogen in a Magnetic Field", Phys. Rev. Lett., **71**, 2899
- [38] S. Sridhar and A. Kudrolli, 1994, "Experiments on Not Hearing the Shape of Drums", Phys. Rev. Lett., **72**, 2175

- [39] H. Weidenmüller, 1994, "Why different drums can sound the same", Physics World, August 1994, 22
- [40] V.N. Prigodin, 1995, "Spatial Structure of Chaotic Wave Functions", Phys. Rev. Lett., 74, 1566
- [41] S. Sridhar, D.O. Hogenboom, and B.A. Willemssen, 1992, "Microwave Experiments on Chaotic Billiards", J. Stat. Phys., 68, 239-58
- [42] H.-D. Gräf, H.L. Harney, H. Lengers, C.H. Lewenkopf, C. Rangacharyulu, A. Richter, P. Schardt, and H.A. Weidenmüller, 1992, "Distribution of Eigenmodes in a Superconducting Stadium Billiard with Chaotic Dynamics", Phys. Rev. Lett., 69, 1296-9
- [43] P. So, S.M. Anlage, E. Ott, and R.N. Oerter, 1995, "Wave Chaos Experiments with and without Time Reversal Symmetry: GUE and GOE Statistics", Phys. Rev. Lett., 74, 2662-5
- [44] U. Stoffregen, J. Stein, H.-J. Stöckmann, M. Kuś, and F. Haake, 1995, "Microwave Billiards with Broken Time Reversal Symmetry", Phys. Rev. Lett., 74, 2666-9
- [45] O. Bohigas, O. Legrand, C. Schmitt, and D. Sornette, 1991, "Comments on spectral statistics in elastodynamics", J. Acoust. Soc. Am., 89, 1456-8
- [46] D. Delande, D. Sornette, and R. Weaver, 1994, "A reanalysis of experimental high-frequency spectra using periodic orbit theory", J. Acoust. Soc. Am., 96, 1873-80
- [47] R.L. Weaver, 1989, "Spectral statistics in elastodynamics", J. Acoust. Soc. Am., 85, 1005-13
- [48] C. Ellegaard, T. Guhr, K. Lindemann, H.Q. Lorensen, J. Nygård, and M. Oxborrow, 1995, "Spectral Statistics of Acoustic Resonances in Aluminium Blocks", Phys. Rev. Lett., 75, 1546-9
- [49] J. Maynard, 1996, "Resonant Ultrasound Spectroscopy", Physics Today, January 1996, 26
- [50] A. Migliori, J.L. Sarao, W.M. Visscher, T.M. Bell, M. Lei, Z. Fisk, and R.G. Leisure, 1993, "Resonant ultrasound spectroscopic techniques for measurement of the elastic moduli of solids", Physica B, 183, 1-24
- [51] M. Oxborrow and Clive Ellegaard, "Quantum Chaology in Quartz", Proceedings of the 3rd Experimental Chaos Conference, Edinburgh 1995, to appear.

## Low Dose Proton Radiation Sensitivity of Quartz Resonators

P.E. Cash, D.A. Emmons - Frequency and Time Systems

W. Stapor - U.S. Naval Research Laboratories

### Introduction

The effect of low dose irradiation of 3rd overtone SC (Stress Compensated) quartz crystals is not a completely understood phenomenon. Earlier work of Norton<sup>1,4</sup> and Suter<sup>2,3</sup> proposed the following concepts:

- (1) The low dose radiation sensitivity of quartz crystal resonators is manufacturing lot dependent. This effect contrasts to high dose rate sensitivity, which is dependent on impurities within the crystal structure.
- (2) The low dose radiation sensitivity of quartz crystals cannot be predicted by examination of high dose data.
- (3) The largest change in frequency per Rad occurs at low dose rates.
- (4) There is evidence that the influence of gamma ray and proton irradiation is nearly identical.

Norton and Suter's studies, and those which preceded them, were performed on quartz crystals both AT-cut conventional adhered electrode units, and BVA SC-cut units, using 1.25 MeV photons from Cobalt 60 sources and protons from a cyclotron source.

Frequency and Time Systems (FTS) performed quartz resonator proton radiation susceptibility testing under contract from Lockheed Martin Corporation (LMC) with the support of the U.S. Naval Research Laboratories at the University of California at Davis Crocker Nuclear Laboratory using 65 MeV protons. The testing environment was selected to best simulate the Low Earth Orbit conditions in which a space borne Global Positioning System Receiver (GPSR) would operate. Periodic exposures of 0.07 Rad(Si)/min., for 10 minute durations, corresponding to  $10^3$  to  $10^4$  particles/cm<sup>2</sup>sec were used to approximate the orbital environment. The following report presents the results from the testing which was completed on September 26, 1995. Distinguishing features of this testing were the number of devices tested (13) and methods used to monitor and control the environment. Discussion will include, for comparison, test data taken by FTS on 4.0 MHz 3<sup>rd</sup> Overtone SC-Cut crystals. These crystals are used in the FTS9500 oscillator currently being tested for LMC on the EOS (Earth Observing Satellite) program under Cobalt 60 gamma irradiation.

### Test Conditions

The test vehicles for the quartz resonators were FTS1130 oscillators. The FTS1130 is a double ovenized high performance oscillator which exhibits temperature coefficients of less than  $5 \times 10^{-12}/^\circ\text{C}$ . The crystals tested were 3<sup>rd</sup> overtone SC-Cut crystals in a HC-40 style holder manufactured by Bliley Electric Company. The crystals were made from synthetic swept quartz and were preconditioned by irradiating the devices to a level of 23.3K (dosimetry average) Rad using a Cobalt-60 source. The crystals were of two frequencies 10.24 MHz (6 units) and 10.95 MHz (7 units). These resonators were spares from the production lots of space flight oscillators. Dosimetry at the cyclotron facility utilized LiF thermoluminescent diodes (TLD) on top of the crystal enclosure, on top of the inner enclosure and on the outside of the test bed oscillator. The cyclotron beam pattern was nearly uniform over an 8 cm wide diameter. The scintillation counter detector (NaI and photomultiplier) was placed in the beam path in a position symmetrical to the area of irradiation on the crystal enclosure. Dosimetry for each exposure was accurate to better than 10%. A scintillation counter measured the energy spectrum and total integrated beam charge for each exposure. 65 MeV proton energy was chosen as a realistic indicator for effects in the 40-100 MeV range representing the broad peak for solar protons passing through a typical 1 gm/cm<sup>2</sup> of aluminum. The quartz blank of the resonator ( $\approx 0.3$  gm/cm<sup>2</sup>) absorbs approximately 2-3% of the energy of each proton. The test bed oscillator shielded the crystal by 1 gm/cm<sup>2</sup> of equivalent aluminum. The effective shielding degraded the proton spectrum peak to 57 MeV which resulted in a 15% increase of dose at the crystal relative to the observed scintillator detector dose. The incident doses were adjusted accordingly. The frequencies of two oscillators were measured concurrently. The oscillator in the path of the protons was the device under test, and the second out of beam device served as the control unit. The control oscillator showed no effects during testing, demonstrating the validity of the results from the device under test. This verified that the stray magnetic fields and other environmental effects had a negligible effect. The oscillator under test was attached to a temperature controlled aluminum plate held to within  $\pm 0.2^\circ\text{C}$ . This temperature, as well as humidity and barometric pressure were monitored and

recorded during the testing.

The frequency of the oscillators were measured by mixing the output frequency against a synthesizer with a 100 Hz offset from the nominal oscillator frequency. The synthesizer was referenced to an FTS Cesium Standard to provide absolute frequency accuracy. The output of the mixer was measured using HP53132A counters. The fractional frequency, temperature, humidity, barometric pressure, event markers and time interval were recorded on a computer for subsequent analysis. The fractional frequency was also recorded on a strip chart recorder, for the purpose of making continuous measurements and to back up the computer recorded data.

The sequence of irradiation was consistent throughout the testing. The test bed oscillators were positioned to place the crystal in the path of the protons with the quartz blank perpendicular (+ Z-axis) to the beam. Ten minute irradiation periods at rates of 0.07 rad (Si)/min ( $8 \times 10^3$  part./cm<sup>2</sup>-sec) and 0.15 rad (Si)/min ( $1.8 \times 10^3$  part./cm<sup>2</sup>-sec) were followed by 20 minute recovery periods. The majority of the crystals experienced 4 cycles at both radiation levels.

Three important additional tests were performed. The first was measurement of recovery response following 9 hour and 36 hour time periods after the initial cycles of radiation. The second measured directional sensitivities of two crystals in the  $\pm Z$ ,  $\pm Y$  and X direction. The third test measured the effects of high levels of radiation exposure, 300K Rad and 1M Rad on two oscillators.

#### Definitions

The following characteristics are defined and analyzed:

- y =  $\Delta f/f$  Change in frequency versus a reference frequency.
- y=0 The arbitrary baseline or starting point
- $\Delta y$  Change in offset with respect to a previous y
- dy/dt Time rate of change of y

#### Measurement Precision and Noise

The intrinsic frequency fluctuations of the cesium standard used as the primary frequency reference show Allan deviation,  $\sigma_y(\tau)$ , equal to  $3 \times 10^{-11}/\tau^{0.5}$  within the observation time period. During the observation period of 10 seconds, the reference will contribute  $1 \times 10^{-11}$  of noise. The observed noise in

the measured frequency of each oscillator under test is a combination of device under test and reference noise. The Allan deviation of the oscillator is  $\approx 1 \times 10^{-12}$  over the measurement period of 10 seconds.

#### Test Results

The summary of test results is shown in Tables I and II. These tables show the frequency variation after each irradiation cycle, the total change during testing, and the mean variation in the frequency normalized on a per rad basis for the 0.07 rad/min and 0.15 rad/min exposure levels. Note that the sequential exposures show, in general, a trend of diminishing sensitivity within each sequence. Tables I and II include the maximum frequency variation measured and environmental variations (pressure, humidity and temperature.) Tables I and II include a brief statement concerning the relative sensitivity of the crystal and observations which distinguish the performance of the crystal. The relative sensitivities were labeled as either low, moderate or high. Figures 1-3 were selected from 13 charts produced from the measured data. These figures show the fractional frequency and dy/dt.

Figures 4 and 5 show total dose effects on two oscillators, SN1490 and SN1489, at accumulated levels of 300K Rad and 1M Rad, respectively.

Table III presents susceptibility of ten 4 MHz 3rd overtone SC-Cut crystals used for the EOS program.

#### Analysis

The analysis of the data is divided into key characteristics;  $\Delta y$  sensitivity, recovery characteristics, drift after testing, directional sensitivity, total dose frequency variations.

#### $\Delta y$ Sensitivity

Generally, the accumulation of 0.7 rad and 1.5 rad doses of proton radiation over ten minute intervals causes positive excursions of 3 to  $5 \times 10^{-11}$  per rad. This rate is specific to each crystal and can vary from least sensitive to most sensitive by a factor of 50. Two notable exceptions showed negative  $\Delta y$  for initial doses. Sensitivity to successive doses of radiation at a given rate usually diminishes as dose accumulates. This conditioning effect appears to be permanent, over hours, based upon the data for SN1488, in which irradiation of the crystal was repeated after 9 hour and 36 hour time periods. Obviously a larger sample set of long term conditioning would be required fully to support this statement. The testing showed two crystals, SN1479

and SN1487, with sensitivities of  $<1 \times 10^{-11}$  and  $2 \times 10^{-11}$ . These devices were the least sensitive, in comparison, SN1488 had initial sensitivities of  $6 \times 10^{-10}$  per rad.

#### Recovery Responses

The 10 minute on and 20 minute off irradiation testing revealed a wide range of frequency drift recovery responses, from nearly zero to complete return to the pre-irradiation levels. One possible theory is that some recovery always takes place, however the rate of each crystal is unique and is dependent on the rate of irradiation. Data from SN1488 strongly supports this concept. Notice that the post irradiation frequency drift after exposures of 0.7 rad is nearly zero. After exposure to 1.5 rad, the post irradiation drift contains a negative slope which flattens out after a time period of 4 to 5 hours. The variation in drift slope is indicated in serial numbers 1481, 1482, 1484 and 1485. Within the context of "conditioning" of the sensitivity effects, Figure 2 shows an interesting result for the most sensitive of the crystals: SN1488. Following a total of 10 exposures, the oscillator was left unperturbed for 9 hours. It was then subjected to a low dose rate exposure of 0.7 rad and the response was  $12.9 \times 10^{-11}$  / rad; i.e. less than for any previous 0.7 rad exposures. The oscillator was then left for 36 hours and exposed again to 0.7 rad and 1.5 rad in succession. The responses were  $8.6 \times 10^{-11}$ /rad and  $5.3 \times 10^{-11}$ /rad, both less than for any of the original 10 exposures.

#### Post Irradiation Drift

In cases where we could establish pre- and post irradiation drift rates, the drift rates appeared to show little long term change after the transient effects were recorded. By way of contrast, Norton<sup>4</sup> has shown one example of low dose data for a BVA crystal in which the post irradiation drift was  $-1.5 \times 10^{-11}$  per hour compared to nearly zero before exposure.

#### Directional Sensitivities

Testing included measurement of the sensitivities with respect to the orientation of the crystal to the proton beam. This test was required to justify the methodology of tests made normal to the crystal surface. The sensitivities were measured in both normal (Z) and three parallel directions ( $\pm Y$  and  $+X$ .) The sixth orientation ( $-X$ ) was obscured by the oscillator electronics and was not measured. In general, the sensitivities are similar in the five directions. This supports the idea that the testing which performed is valid for the typical space (spherical) exposure.

#### Total Dose Sensitivities

The total dose frequency variations of SN1490( 300K Rad ) and SN1489( 1M Rad) of  $-5.4 \times 10^{-9}$  and  $-32.4 \times 10^{-9}$ , respectively, are important since they clearly show the non-linear effect of the total vs. low dose radiation. The sensitivity of the two oscillators are  $2 \times 10^{-14}$ /rad and  $3 \times 10^{-14}$ /rad. This is three orders of magnitude lower than the low dose rate measured sensitivities on the same resonators.

#### 4 MHZ Resonator Test Results

Table III shows summary results from testing of 4 MHZ SC-Cut quartz crystals( conventional, adhered-electrode, Bliley Electric Company) at low dose rates of Co-60 gamma irradiation. Here also there is a wide scatter of results, but the general trend seen is positive  $\Delta y$  for single low doses, and are very similar in magnitude to the present proton results.

#### Conclusions

The results of the testing of 4 MHZ, 10.24 MHZ and 10.95 MHZ 3rd overtone SC-Cut quartz crystals confirms the earlier work of Norton<sup>1,4</sup> and Suter<sup>2,3</sup> on 5<sup>th</sup> overtone AT-Cut, SC-Cut crystals. The testing showed that low dose rate radiation sensitivity is manufacturing lot dependent. Frequency variations due to low dose rate proton and gamma ray exposures varied between  $<1 \times 10^{-11}$  and  $4 \times 10^{-10}$  per Rad. The majority of the crystals exhibited radiation sensitivities of  $3-5 \times 10^{-11}$  per Rad. The hypothesis that a non-linear relationship exists between low dose rate and high total dose rate was confirmed. Testing indicated that the majority of crystals exhibited a conditioning effect during successive exposures to low dose radiation.

The effect on performance of these types of sensitivities is highly system dependent. However, the capability exists to screen low sensitivity resonators for use in applications where this characteristic is important. In fact, crystal screening was employed for the EOS program.

In comparison to previous data taken on low dose sensitivities of quartz resonators, the 4 MHZ, 10.24 MHZ and 10.95 MHZ crystals have similar responses.

In conclusion, the sensitivities of 10 MHZ 3<sup>rd</sup> Overtone SC Cut quartz resonators are most likely manufacturing lot dependent, and vary within the lot. Furthermore, the effect of repeated exposure to low dose proton irradiation appears to systematically reduce the sensitivity of the crystals.

**Acknowledgments:**

We would like to thank Dr. Carlos Castaneda and the staff of the Crocker Nuclear Laboratory at UC- Davis for their support and cooperation, Dr. Joseph Suter of The Johns Hopkins University Applied Physics Laboratory for review and comments to this manuscript and for help from staff members of The Aerospace Corporation. Measurements of the EOS crystals were aided by David Swant of LMC. Finally, this study would have been impossible to complete without the contribution of Jon Holland of FTS and Patrick McDonald of NRL.

**References:**

1. J.R. Norton, J.M. Cloeren, and J.J. Suter et al., The Johns Hopkins University Applied Physics Laboratory, "Results from Gamma Ray and Proton Beam Radiation Testing of Quartz Resonators." IEEE, Transactions on Nuclear Science, NS-31, p 1230 (1984).
2. J.J. Suter and R.H. Maurer, "Low and High Dose Photon Irradiation of Quartz Crystal Resonators.", The Johns Hopkins University Applied Physics Laboratory, Proceedings from 17th Annual PTTI Conference
3. J.J. Suter et al., The Johns Hopkins University Applied Physics Laboratory and B. Schlueter et al., Oscilloquartz, SA, "The Susceptibility of BVA-SC Resonators to Photon and Ionization Effects."
4. J.R. Norton, "Performance of Ultra stable Quartz Oscillators Using BVA Resonators", Proc. 8th European Frequency and Time Forum, pp. 457-465 (1994)

Table I  
Analysis of Low Dose Frequency Variations( 1e-11)

10.95 MHz Crystals

Osc. SN	Crystal SN	0.07 Rad/min. Exposures	0.15 Rad/min. Exposures	Total Change	Mean dy per rad 0.7 rad exp.	Mean dy per rad 1.5 rad exp.	Maximum Change	Pressure Change Torr.	Temp. Change deg. C	Humidity Change %	Relative Sensitivity	
1478	9450-10	5,4,3,6	7,5,6	36	6.4	4.0	7	0.8	0.1	0	Moderate	
1479	E-07	0,0,0,0	2,0,0,0	1.5	0.0	0.3	1.5	1.6	0.1	2	Least	
1480	9512-04	3,5,2,2	2,3,3,1	21	4.3	1.5	5	0.9	0.2	3	Moderate	Drift rate after irradiation not established( short observation time), although it is different than precursor drift rate.
1481	9512-02	2,4,3,2	5,5,5,5	31	3.9	3.3	5	0.4	0.1	1	Moderate	Final drift rate less than pre-irradiation rate and 0.7 rad recovery periods.
1482	E-12	5,5,6,6	8,7,8,6	51	7.9	4.8	8	0.4	0	1	Moderate	Initial drift rate is low and established.
1483	9450-07	-20,-18,-18,0	-15,-2,5,4,3	-63	-19.6	-0.7	5	1.8	0.3	7	High	Negative y for 0.7 rad dose reaching zero on fourth event, positive going transients for 1.5 rad doses.
1484	9450-11	5,5,3,2	6,6,5,3	35	5.4	3.3	6	1.6	0.1	3	Moderate	Diminishing dy changes vs radiation during sequence of exposures, constant drift throughout test.

Table II  
Analysis of Low Dose Frequency Variations( 1e-11)

10.24 MHz Crystals

Osc. SN	Crystal SN	0.07 Rad/min. Exposures	0.15 Rad/min. Exposures	Total Change	Mean dy per rad 0.7 rad exp.	Mean dy per rad 1.5 rad exp.	Maximum Change	Pressure Change Torr.	Temp. Change deg. C	Humidity Change %	Relative Sensitivity	
1485	9451-08	6,5,3,3	8,8,8,9	50	6.1	5.9	9	1.1	0.2	1	Moderate	Initial drift high, recovery during intermediate periods changes slope for 0.15 rad rate, final drift rate not established.
1486	9450-02	-3,-3,-3,-2	-4,-4,-5,-6	-27.5	-3.6	-3.1	-2	1.7	0.1	6	Moderate	Drift rate before and after irradiation is approximately 3e-11/hr
1487	9509-09	1,2,2,2	2,2,2,2	13	2.1	1.3	2	0.7	0.1	5	Low	Drift rate of 2e-11/hr was unaffected by radiation.
1488 (1)	9450-01	40,36,33,29,27	47,36,15,10	260	47.1	19.2	47	3	n/a	5	Highest	Annealing of sensitivities observed during testing and confirmed after 9 hour recovery period and retest 36 hours later.
1489	E1-12	3,2,2,1	2,0	10	2.9	0.7	3	1.7	0.1	2	lowest	Drift <1e-11/hr unaffected by radiation.
1490	9509-14	4,3,2,2	5,4,4,2	26	3.9	2.7	5	1.9	0.1	3	Moderate	Significant transient effects, drift is 3e-11/hr

Table III  
Summary of Results for 4 MHz Crystals on EOS Program

Units are 1e-11

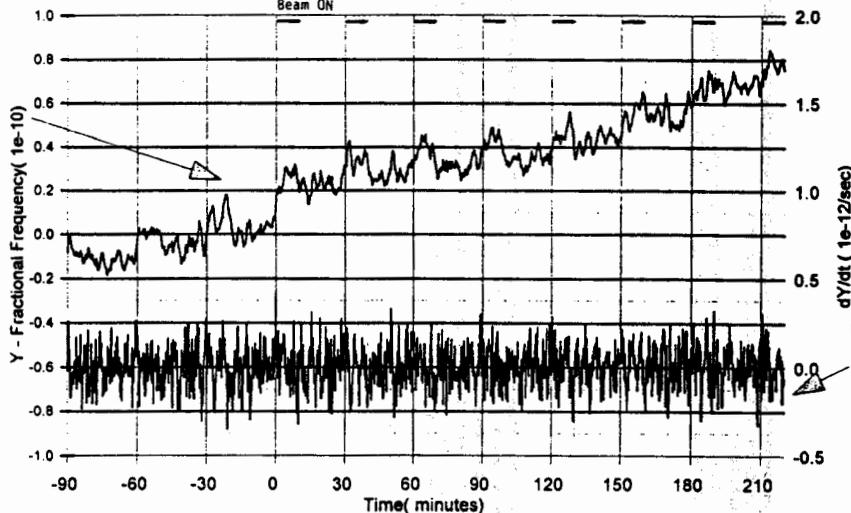
Test Bed	Xtal No.	Pre-Rad Drift removed from Data Set. E-11 per day	Max. Dev. During 1h Exposure	Deviation at end of Exposure	Max. Dev. During Recovery Overshoot	Dev. 12 hours later	"per-rad" effects	
							prompt	longer-term
717	10	-3.7	3.6	3.6	-5	-3	* 1	-1
718	11	3	2.6	0	-8	-8 or less	0.62	-2
719	9	5	4.0	4	-3	-2	1	-0.5
720	15	10	1.8	1.8	-18 (min)	-15	* 0.6	-4.7
720	repeat 15	-0.5	1.1	0.1	-11	0 to -6	0.3	-2 or less
721	12	-2	11.0	11	1.7	2	2.6	0.5
722	17	5.5	6.2	6.2	-17	-17(est.)	1.5	-4
723	18	0	5.0	4	-7	-2	1.2	5
724	23	2	+5; -10	-10	-22	-10	-2.4	(-2.4 to -5.2)
725	25	-3.8	6.3	4.8	-5	2	1.5	0.5
726	26	-1.5	2.6	-3.8	-9	7	0.6	1.7

\* NOTES: Xtl's 10, 15 = 3.4 rads; all others, 4.2 rads [ref. D.Swant memo 3 Jan 1965]

Low Dose Radiation Testing SN1487

10 Minute exposures of 4 x 0.07 rad/min and 4 x 0.15 rad/min with 20 minute recoveries  
 10.240000 MHz Crystal 9451-08 Y: 2 minute average dY/dt: 2.5 minute average  
 Beam ON

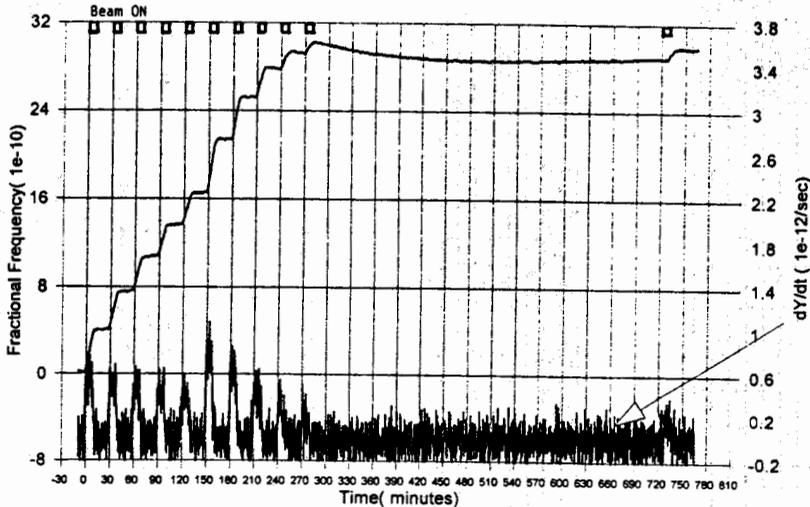
Figure One



Low Dose Radiation Testing SN1488

10 Minute exposures of 5 x 0.07 rad/min, 5 x 0.15 rad/min and 1 x 0.07 rad/min with 20 minute and 7 hour recoveries  
 10.240000 MHz Crystal 9450-01 Y: 2 minute average dY/dt: 2 minute average

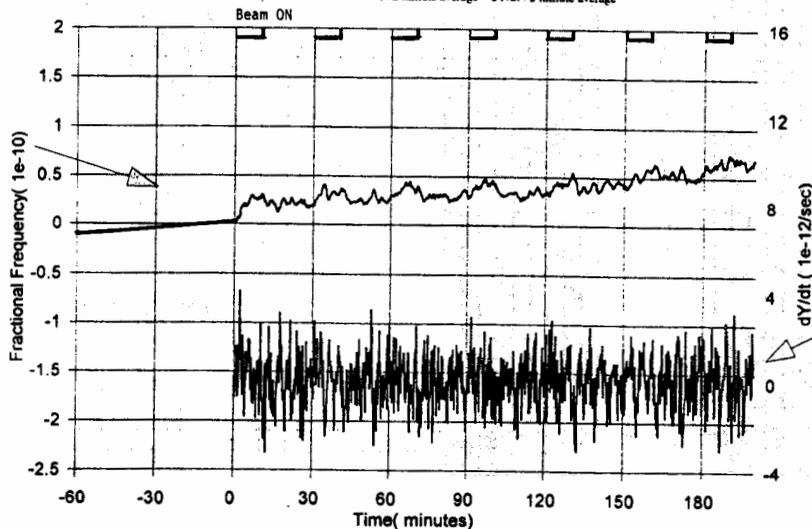
Figure Two



Low Dose Radiation Testing SN1489

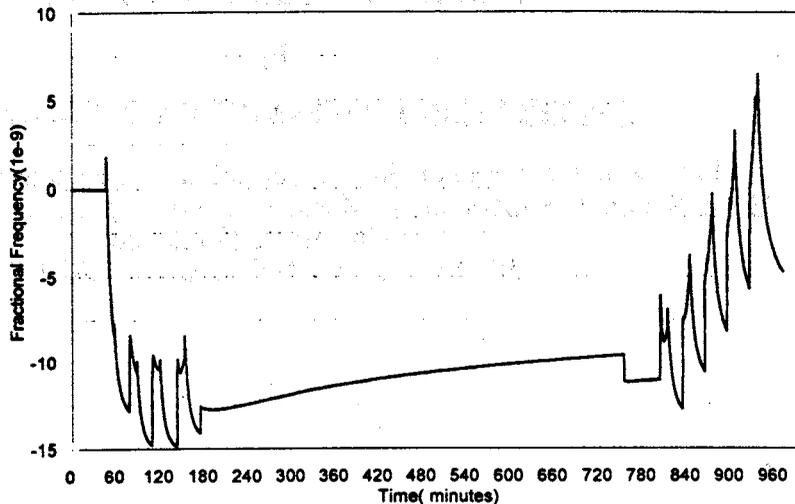
10 Minute exposures of 4 x 0.07 rad/min and 2 x 0.15 rad/min with 20 minute recoveries  
 10.240000 MHz Crystal E1-12 Y: 2 minute average dY/dt: 2 minute average

Figure Three



SN1490 300 KRad Total Dose Sensitivity  
Exposures of 3x20K, 6 x 40K

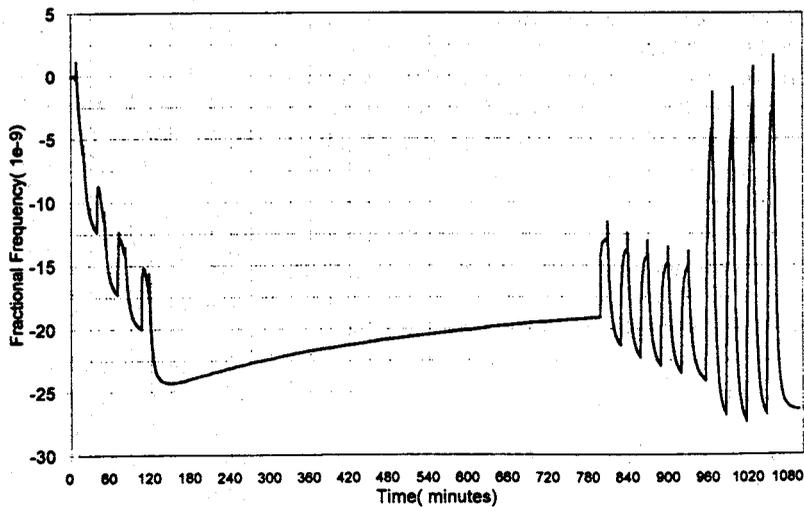
Figure  
Four



Sensitivity: 2 e-14/ rad  
Crystal: 9509-14

SN1489 1 M Rad Total Dose Sensitivity  
Exposures of 3x20K, 6x40K, 4x160K

Figure  
Five



Sensitivity: 3 e-14/ rad  
Crystal SN:E1-12

# CONTRIBUTION OF A 3D NUMERICAL SIMULATION TO THE DESIGN OF MICRO-MACHINED QUARTZ MICROSTRUCTURES

by

**C.R. TELLIER, T. MESSAOUDI and T.G. LEBLOIS**

Laboratoire de Chronométrie, Electronique et Piézoélectricité (LCEP)  
Ecole Nationale Supérieure de Mécanique et des Microtechniques (ENSM)  
26 Chemin de l'Épitaphe  
25030 BESANÇON CEDEX - FRANCE

## ABSTRACT

This paper is an attempt to prove the interest of a 3D simulation for micromachining of quartz. For this purpose new results related to doubly rotated plates are discussed. The 3D simulation based on the tensorial model of dissolution constitutes a convenient computer aided design for masks patterns. But care must be taken that for quartz crystal the major problem remains the accurate adjustment of the dissolution slowness surface.

property of this model which is used in the simulation results from this behavior : the propagation vector  $\mathbf{P}$  which at any etching duration describes the displacement of a given surface element can be calculated from the equation for the slowness surface. By comparing all the displacements  $\mathbf{P}$  of elements potentially present at the starting surface it becomes possible to reconstruct numerically the final etching shape.

## THEORETICAL BASIS OF SIMULATION

### The tensorial model of dissolution

The graphical and numerical simulation is derived in the framework of the tensorial model for the anisotropic dissolution previously proposed by C.R. TELLIER and co-workers [1]. In this model at a surface element of orientation  $(\varphi, \theta)$  we associate a reciprocal dissolution slowness vector  $\mathbf{L}(\varphi, \theta)$ , whose magnitude,  $L$ , is equal to the dissolution rate,  $V_N$  and whose direction coincides with that of the unit inward normal,  $\mathbf{n}$ , to the surface element.

It is obvious that when the angles of cut  $(\varphi, \theta)$  vary the vector  $\mathbf{L}$  generates in space a dissolution slowness surface whose equation can be conveniently written as

$$L(\varphi, \theta) = L(n_1, n_2, n_3) = D_0 + D_i n_i + D_{ij} n_i n_j + D_{ijk} n_i n_j n_k + \dots$$

where  $D_0, D_i, D_{ij}, \dots$  are components of dissolution tensors and  $n_1, n_2, n_3$  are the cartesian components of the unit normal  $\mathbf{n}$ . The number of dissolution constants is reduced by the crystal symmetry to give for the class 32 a general analytical expression involving the angles of cut  $(\varphi, \theta)$ . For a dissolution process exclusively governed by orientation surface elements move along straight lines within the crystal during etching. The main

### Applications to micromachining

In micromachining etched structures are obtained by starting with an inert mask. So depending on the shape of the initial mask we are concerned with 2D or with 3D etching and consequently we work either with a polar diagram of  $L$  lying in a particular cross-section of orientation  $(\varphi_0, \theta_0)$  or with a large part of the slowness surface. For example dissolution profiles for grooves of infinite length can be derived from a polar graph whereas the simulation of etching shapes for membranes or mesa micromachined with an initial circular mask involve the entire slowness surface. So when we start with a square mask (figure 1a) we can distinguish regions on membranes and mesa which clearly are matter for a 2D etching such as the square sides and for a 3D etching such as corners.

At this point it is of interest to recall the criteria [3] which are commonly used to predict approximate etching shapes. Consider two planes  $P_1$  and  $P_2$  forming an intersection of angle  $\delta$  so that the criteria can be stated as follows :

1. A concave intersection is stable (its angle  $\delta$  remains unchanged) provided there is no plane between  $P_1$  and  $P_2$  with a larger dissolution slowness.
2. Conversely a convex intersection remains stable when there is no plane between them with a smaller dissolution slowness.

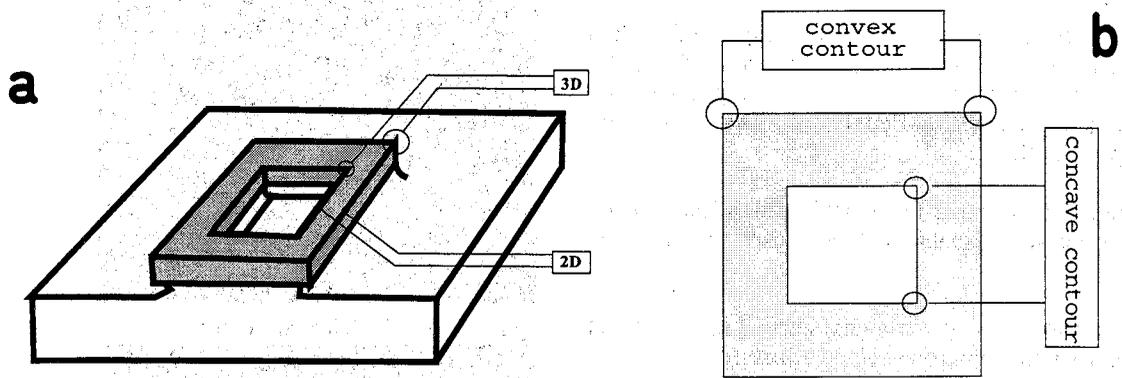


Figure 1 : a : 2D and 3D etchings for hole and mesa structures, b : convex top contour of mesa.

It is clear that stable intersections are connected with converging displacements whereas diverging trajectories give rise to curved intersections. Applying these criteria we infer that maxima maximora of  $L$  determine primarily the final etching shape of a starting "circular" membrane. For a membrane micromachined with a square mask maxima maximora of  $L$  give the final shape of corners whereas the local etching under the mask at the square side is correlated to the maxima in polar graphs of  $L$  provided the inert mask remains perfectly rigid during the etching process. Thus depending on the orientation of the cross-sectional plane these maxima are not necessarily connected with protuberances in the slowness surface.

The analysis of limiting shapes for mesa like structures seems to be more complicated because depending on the viewing angle the structure appears to be composed of successive either concave or convex intersections. Effectively if we put attention to dissolution profiles under the mask we work with successive concave intersections but as soon as we look to the upper contour of mesa we are concerned with convex shapes (figure 1b). Consequently in terms of previous criteria the final shape results from a competition between minima minimora and maxima maximora of  $L$ . In the numerical simulation we have thus to decide which type of extrema plays the most important role because in the simulation we provide numerical tests to distinguish diverging and converging trajectories which depend on the nature of intersections [2]. To answer to this question it seems reasonable to undertake a prior analysis of experimental mesa shape.

## EXPERIMENTALS

### Experimental details

Several singly-rotated plates (500  $\mu\text{m}$  thick) with angles of cut ( $\varphi = 0^\circ$ ,  $\theta = \pm \theta_0$ ) and doubly-rotated plates with angles ( $\varphi = 90^\circ$ ,  $\theta = \theta_0$ ) denoted,  $X \pm \theta_0$  and  $Y \pm \theta_0$  plates respectively were cut in the same quartz blank. Prior to etching they were mechanically lapped

and then optically polished. To etch structures in a concentrated  $\text{NH}_4\text{F.HF}$  solution maintained at a constant temperature of  $60^\circ\text{C}$  we used conventional photolithographic process passing through the fabrication of a mask chemically inert in etchant. For this purpose Chromium-Gold films were thermally evaporated. To induce relaxation of internal stresses due to structural defects frozen in the film during the evaporation as-deposited films suffered an "in-situ" "natural" restructuration at ambient temperature. Then the annealing process was continued with several isothermal agings performed at increasing temperatures. Figure 2 shows a magnification of a Cr-Au mask after a prolonged etching indicating that films present only some threading "dislocations" which extend to the free surface of films. Such films allow us to provide etchings whose duration can reach 24 hours. It is in particular the case for the cut  $X+30^\circ$  which was immersed during 24 hours in the etching to obtain a depth of etch of about 250  $\mu\text{m}$  on each face. For the  $Y+\theta_0$  cuts we adopted an etching duration of about 7 hours because the two faces etch differently so that we have to perform two successive etchings to characterize completely the anisotropic behavior.

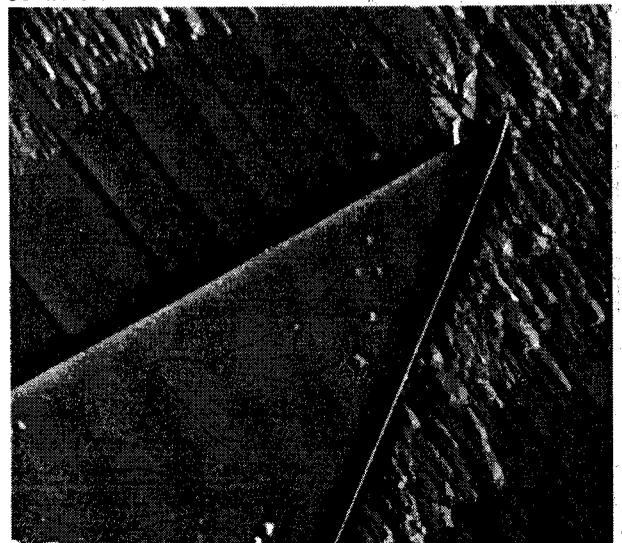


Figure 2 : The Cr-Au mask after prolonged etchings

### Micromachined structures

Figures 3 to 6 give some SEM images of etched membranes and mesa obtained with mask of circular shape. These images indicate that all etching shapes are, as expected, orientation-dependent. The anisotropy affects principally :

(i) the number of planar facets (PF) which limit etched structures and whose extent and inclination vary with the cut,

(ii) the development of somewhat curved regions (CR) which partly bound micromachined structures,

(iii) the formation of dissolution profile with acute angle in the so-called sharp region (SR) characterized by the presence of a sharp edge which penetrates deeply within the crystal under the mask.

Figures 3 to 6 reveal also other features which merit some comments :

(i) For several cuts we observe the presence of jagged sidewalls which in most cases are constituted by grooves aligned along preferential directions. Hence one can expect to recognize the mean orientation of such sidewalls because dissolution figures such as pits and hillocks are directly connected with the cut. In contrast curved regions must be free on any dissolution texture provided the absence of structural defects on the sidewalls.

(ii) The bottom surface of structure is also covered by dissolution figures initiated by defects due to the mechanical lapping. However, for the  $Y+\theta_0$  cuts the geometrical features of these dissolution figures are only slightly modified for  $\theta_0$  increasing from  $30^\circ$  to  $60^\circ$ . In this condition for these cuts it remains difficult to distinguish the presence of a shoulder slightly disoriented with respect to the reference surface.

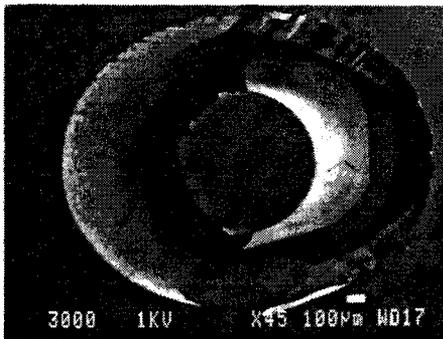


Figure 3 : X+30 etched structure

But the most important experimental fact which emerges from the SEM micrographs is that the bevelling planes and curved shoulder related to membrane resemble to those connected with mesa. This typical behavior is conveniently illustrated by the SEM image of deeply micromachined X+30 mesa and membrane. It should be also pointed out that sharp edges seem to penetrate much

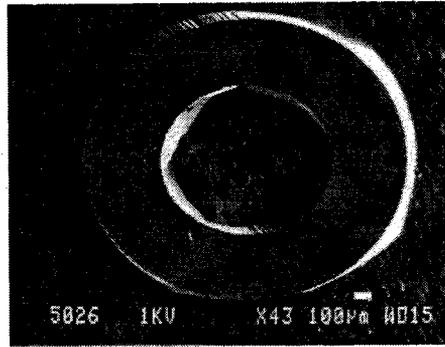


Figure 4 : Y+70 etched structure

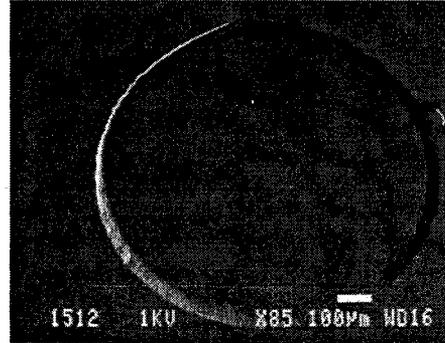


Figure 5 : Y+15 diaphragm

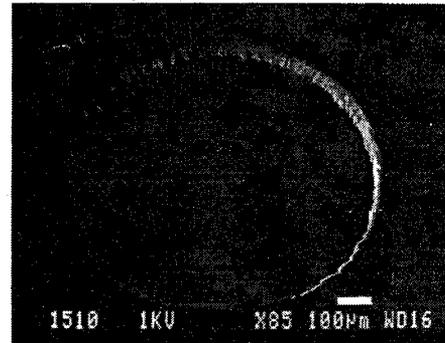


Figure 6 - Y+15 mesa

more within the crystal for mesa (convex undercutting) than for membrane (concave undercutting).. Thus we infer that for mesa as for membranes the final etching shape is primarily connected with the maxima in the dissolution slowness. This conclusion deviates from previous explanations which argue for large convex undercutting induced by minima in L.

### THEORETICAL 3D SHAPES

#### Theoretical procedures

The dissolution constant are evaluated from experiments using a procedure proposed in a previous paper [3]. In practice we analyze essentially two types of 2D etching shapes :

1. Out-of-roundness profiles associated with the etching cross-sectional shape of initially circular thick quartz plates. Such analysis gives informations on the shape of

polar diagram for  $L$  lying in the same sections but does not offer a great accuracy on the amplitude of successive extrema in  $L$ .

2. Profilometry traces whose geometrical features depend on the nature and on the number of extrema for  $L$  present in a particular part of the polar diagram. Here again we cannot extract directly the true amplitude of extrema from data. For this reason the procedure for the evaluation of dissolution constants remains iterative and requires several verification steps in which we compare predicted and experimental 2D shapes. Following this procedure we have estimated 41 dissolution constants to generate a dissolution slowness surface from which 3D shapes can be numerically derived. The 3D simulation is performed by making tests of some nature for membranes. Tests of different nature are successively made for mesa assuming that final etching shape is also composed of slow etch surface elements. Actually the graphical simulation gives two types of representations (see figure 7 for example) : a constant level contour shape which allows us to depict easily regions where planar facets or curved sidewalls develop and cross-sectional profiles corresponding to various intersecting lines which make an angle  $\psi_i$  with the rotated axis  $x''_1$ .

### Predicted shapes

Figures 7 to 10 give some predicted shapes for structures micromachined on X+30, Y+15 and Y+70 plates. Let us first put attention to the deeply etched X+30 plate. The adequation between experimental and theoretical shape seems very good. The agreement concerns the formation of planar facets and curved region as well as the development of sharp edges along specific directions. Moreover predicted value and experimental value for the ratio of the maximum underetch (which occurs in the so-called sharp region) to the depth of etch are in close accord.

Owing to the general agreement observed for various X+  $\theta_0$  cuts. Surprisingly a careful comparison of figures 8 to 10 with figures 4 to 6 reveal some deviations which affect the extent of curved regions. In contrast inclinations of planar facets show very small departure. Moreover in theory the Y+15 cut dissolves more slowly than the Y+70 cut but the predicted ratio for the two etch rates departs from the experimental ratio by about 30%. These departures suggest that we have only to adjust a part of the Y-polar diagram of  $L$ .

### Practical aspects of simulation

In this section we put attention to the possibility of evaluating technical problems such as lateral underetch and convex undercutting generally encountered for in micromachining. As seen in the above section constant-level contour diagrams give conveniently evidence for the formation of curved regions or planar facets with large extent which give rise to non-symmetrical mechanical structures. Dissolution profiles furnish rapidly qualitative informations on the geometrical features of built-in structure such as presence of a curved

shoulder or of a sharp edge. More important than this qualitative approach is the numerical evaluation of some parameters characterizing the geometry of micromachined structures. The lateral underetch,  $u_L$ , under the mask is one of parameters generally retained to discuss the faisability of structures. For a given cut we have to follow the evolution of  $u_L$  with the angle  $\psi_i$  related to cross-sectional line  $L_i$ . Figure 11 gives changes in polar plots  $u_L(\psi_i)$  with angle of cut  $\theta_0$  for several Y+  $\theta_0$  cuts. Clearly we observe a continuous evolution in the  $u_L(\psi_i)$  plot which with increasing  $\theta_0$  reproduces progressively the polar plot  $L(\psi)$  corresponding to the Z-cut. This behavior may be interpreted in terms of the great influence of maxima-maximora in  $L$  which precisely lie on the Z-polar plot. Sharp edges penetrate more deeply within the crystal for mesa than for membranes. Thus we need to follow the evolution of such convex undercutting with several cuts. Figure 12 gives the changes in the largest underetch  $u_s$  (independently to the depth at which marked undercutting occurs) with angle  $\psi_i$  for a X+30 mesa. We observe that the convex undercutting is, along specific lines  $L_i$ , three times more larger than  $u_L$ . A numerical evaluation of  $u_s$  is thus necessary to design new structures micromachined on non standard cuts.

### CONCLUSION

Micromachining was performed on various singly and doubly rotated quartz plates. Etching shapes for mesa and holes are analysed and compared with 3D shapes derived from a numerical simulation which starts with the equation for the dissolution slowness surface. Clearly etching results in unsuitable geometrical features such as blocking facets and sharp edges. The 3D simulation allows us to evaluate these features provided an accurate experimental determination of dissolution constants was achieved.

### REFERENCES

1. TELLIER C.R., VATERKOWSKI J.L., 1989, "A tensorial representation of the dissolution slowness : Application to etched singly rotated quartz plates", *J.Mater. Sci.*, **24**, 1077-1088.
2. LEBLOIS T, TELLIER C.R., 1992, "Determination of the dissolution slowness surface by study of etched shapes. Part I", *J. Phys. III*, **2**, 1259-1286.
3. TELLIER C.R., BRAHIM-BOUNAB A., 1994, "Anisotropic etching of silicon crystals in KOH solution. Part I : Experimental etched shapes and determination of the dissolution slowness surface", *J. Mater. Sci.*, **29**, 5953-5971.

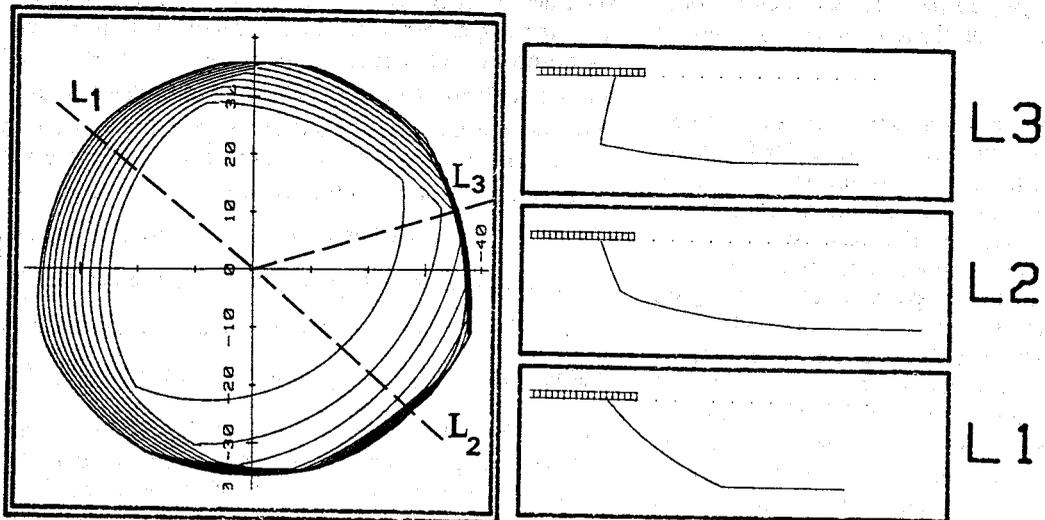


Figure 7 : Predicted constant-level contour diagram and dissolution profiles for a X+30 hole.

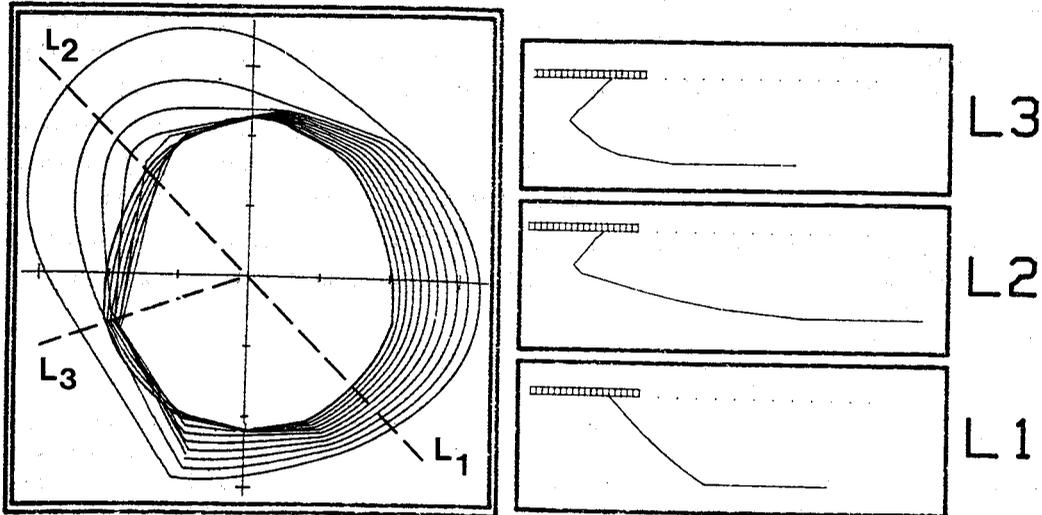


Figure 8 : Predicted constant-level contour diagram and dissolution profiles for a X+30 mesa.

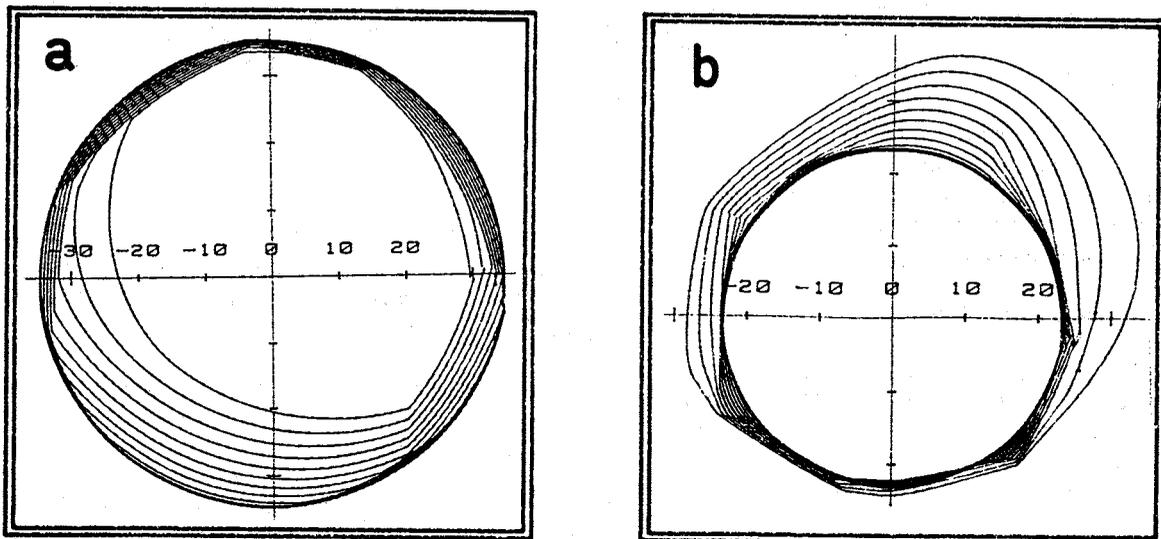


Figure 9 : Predicted shapes for Y+70 diaphragm (a) and mesa (b).

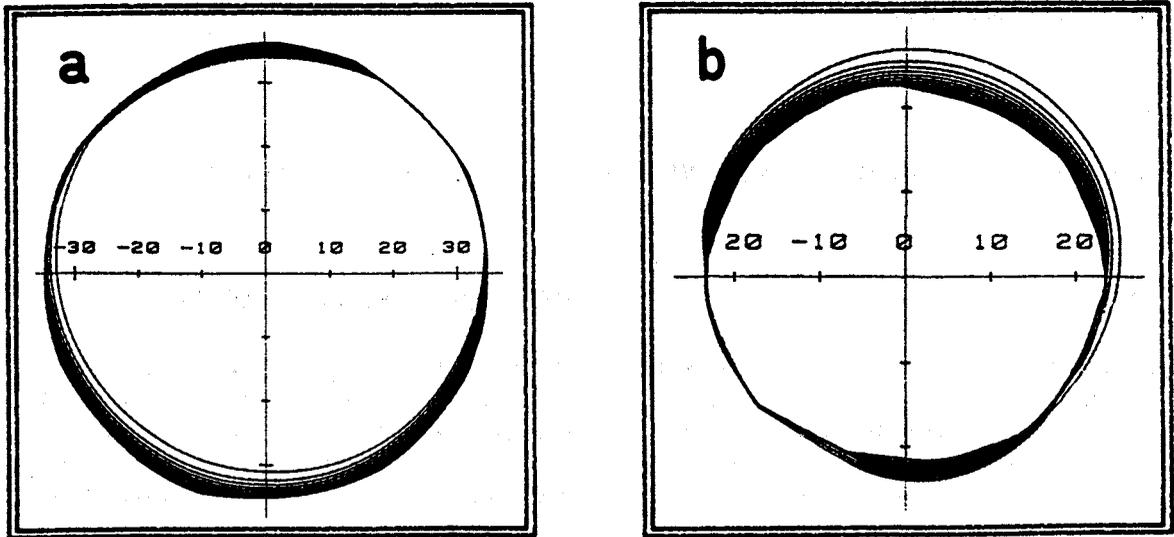


Figure 10 : Predicted shapes for Y+15 diaphragm (a) and mesa (b).

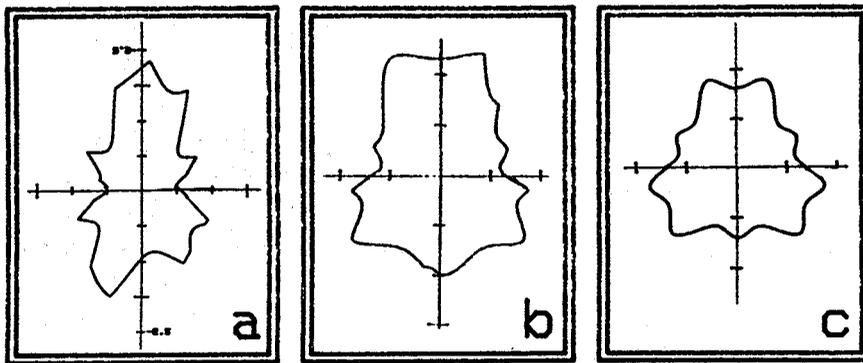


Figure 11 : Polar plots of the lateral underetch  $u_L$ . a, b, c are for Y+15, Y+40 and Y+70 membranes respectively.

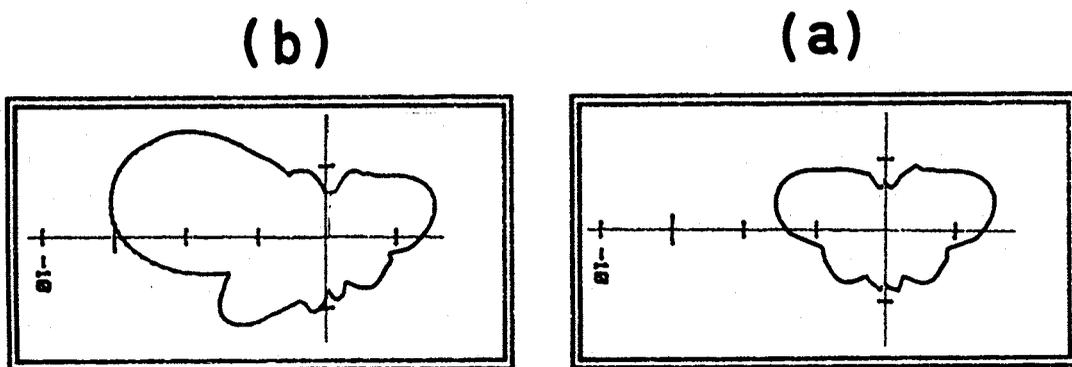


Figure 12 : Comparison of polar plots of (a) the lateral underetch,  $u_L$ , and (b) sharp underetch  $u_S$  : the case of X+30 mesa.

# THE INFLUENCE OF PcCu LAYER CRYSTALLINE STRUCTURE ON THE PARAMETERS OF SAW GAS SENSORS

J. Hechner, T. Wróbel

Institute of Electronic Materials Technology, Wólczyńska 133, 01-919 Warsaw, Poland

## ABSTRACT

This paper presents some problems concerned with SAW NO<sub>2</sub> sensors employing copper phthalocyanine (PcCu) as the chemical interface. We have investigated the influence of crystalline structure of PcCu coating on the sensor sensitivity of generated frequency.

The experiments were carried out for SAW delay lines on STX quartz and 128°YX lithium niobate.

It was found that for the investigated range of parameters the β-PcCu layer provides several times better sensor sensitivity and temperature stability of SAW delay line generation frequency than α-PcCu layer of the same mass. This is valid for both used substrates.

## INTRODUCTION

SAW gas sensors have been the subject of intensive research over the last years [ 1-4 ].

We have been working on nitrogen dioxide SAW sensors. The monitoring of NO<sub>2</sub> concentration is of increasing interest [ 5, 6 ] because of the importance of the gas in environmental pollution as a consequence of its emission from power plants, industry and traffic. SAW sensors are made by applying a thin chemically sensitive film onto the surface of the SAW delay line. In general the detection principle of the SAW gas sensor relies on the change of the SAW velocity upon absorption of the reactant by the sensor film. This velocity change can be brought about by three effects: the change in the mass density of the film, the change in its elastic constants or the change in its electrical conductivity.

This paper presents some problems concerned with SAW NO<sub>2</sub> sensors employing copper phthalocyanine (PcCu) as the chemical interface of SAW delay lines on STX-quartz and 128°YX lithium niobate.

The influence of piezoelectric substrate material, crystalline structure and morphology of the PcCu coatings on the sensors sensitivity, on the response time and on temperature stability of generated frequency is discussed.

## EXPERIMENTAL PROCEDURE

We used polished single crystal STX-quartz and 128° YX LiNbO<sub>3</sub> wafers produced at ITME. Dual-delay lines on both these substrates have been designed.

The SAW sensing element consists of two identical delay lines. One of the lines is covered with PcCu as the chemical coating, while the other is used as a reference in order to compensate for the variation of such parameters as temperature, pressure, etc. (Fig.1.).

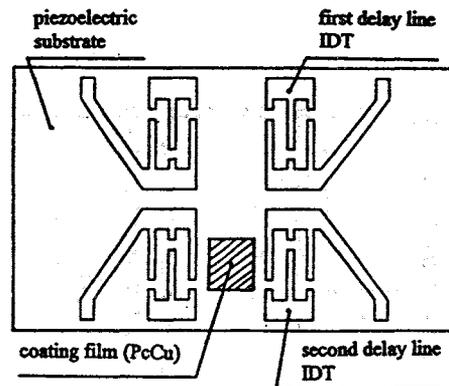


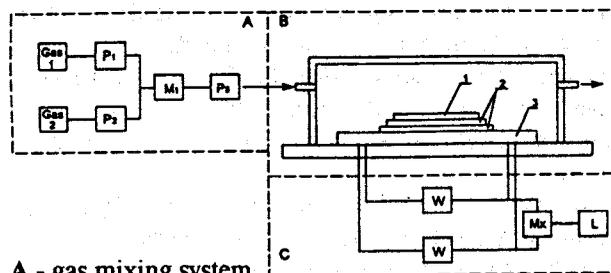
Fig.1. Schematic diagram of SAW dual-delay line

The geometrical and electrical characteristics of the delay lines were chosen to provide low insertion loss of about 17dB before and 20dB after coating with PcCu and appropriately large surface area to be coated. The parameters of delay lines on both substrates are shown in Table 1.

Table 1. The parameters of delay lines.

Parameters	LiNbO <sub>3</sub>	Quartz
center frequency	70.1 MHz	71.7 MHz
insertion loss	17 dB	17 dB
number of transducers	2	2
number of electrodes	41	201
transducer period	55.6 μm	44 μm
electrode width	7 μm	11 μm
acoustic aperture	2080 μm	6082 μm
spacing between IDTs	4000 μm	6000 μm
total chip area	15.0 × 10.2mm	24.2 × 18.2mm

Fig.2. illustrates the experimental measurement system. It contains three modules: the gas-mixing system (A), SAW device (B) and electronics circuits (C).



**A - gas mixing system**

P<sub>1</sub>, P<sub>2</sub>, P<sub>3</sub> - mass flowmeters  
M - mixer of gases

**B - SAW device**

1 - piezoelectric substrates  
2 - heater  
3 - flatpack

**C - electronics circuits**

W - amplifier  
M<sub>x</sub> - frequency mixer  
L - frequency counter

Fig.2. Experimental measurement system.

SAW delay lines were fabricated using standard lift-off process. Aluminium was used as a material of transducer electrodes.

Thin films of PcCu were prepared in a two stage process. At the first stage the input material was purified by vacuum sublimation and at the second one vacuum deposited on piezoelectric substrates at a pressure of the order of  $10^{-6}$  Tr.

The source temperature was 400°C, the distance between the source and the substrate was 40÷100mm. The substrate temperature was varied between 100°C and 400° C.

The structure of the films was determined by X-ray diffraction measurements, while their morphology was examined using a scanning electron microscope. Film thickness was monitored in situ with a quartz crystal oscillator and later ex situ by an optical interferometer. The piezoelectric substrates were bonded to alumina ceramic plates by means of the special silicone adhesive, which proved to be an adequate material for the suppression of bulk modes and unwanted SAW edge reflections.

The sensor was heated by a cermet thick film resistor deposited on the back side of the ceramic plate. The SAW device was mounted in the flatpack and introduced into the exposure chamber, containing the gas inlet and outlet. The surface of the SAW delay-lines was aligned in parallel with the gas stream line within the chamber.

The electronic system is connected to the delay lines by coaxial cables. This system includes amplifiers, mixer and frequency counter.

Oscillation is obtained using a wide band amplifier for each SAW delay line. The frequencies from these oscillators are mixed and the frequency difference is measured.

Nitrogen dioxide (NO<sub>2</sub>) with nitrogen (N<sub>2</sub>) was used as the test gas. The mass flowmeters were used to control the input gas flow and concentration.

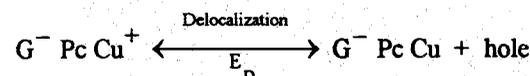
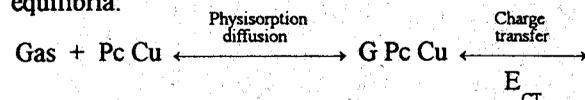
## RESULTS

### The Sensitive Film

The interface material was copper phthalocyanine, which is an organic p-type semiconductor [ 7 ].

PcCu is not an intrinsic semiconductor. However, the chemisorption of NO<sub>2</sub> causes the creation of a charge-transfer complexes with PcCu molecules by accepting electrons from PcCu, which results in increasing of the concentration of holes.

These processes may be represented by the following equilibria:



where G is the molecule of gas interacting with the PcCu cycles of the film, G<sup>-</sup> the ionized gas molecule, E<sub>CT</sub> -the energy of charge transfer, and E<sub>D</sub> the energy required to free the charge, i.e., to create a mobile hole.

In our experiments we have obtained monocrystalline films of α-PcCu and β-PcCu structure, and also two phase α and β films. These types of PcCu film morphology are illustrated in Figs. 3 - 5.

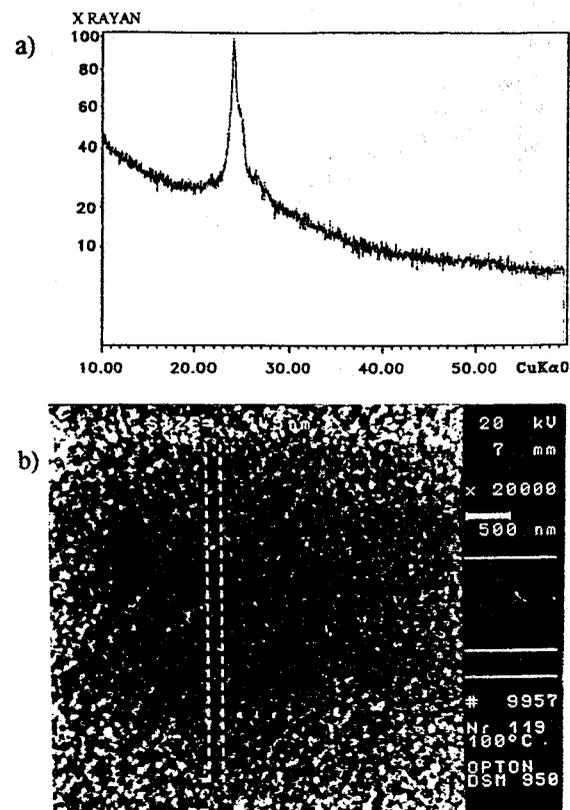


Fig.3. α - PcCu diffraction spectrum and morphology  
a - diffraction spectrum (CuKα radiation)  
b - film morphology

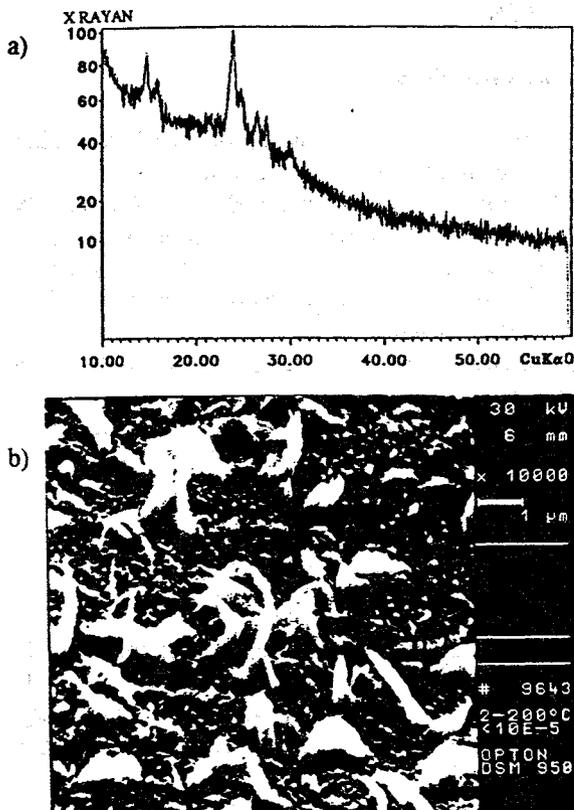


Fig. 4.  $\alpha$  and  $\beta$  - PcCu diffraction spectrum and morphology  
 a - diffraction spectrum (CuK $\alpha$  radiation)  
 b - film morphology

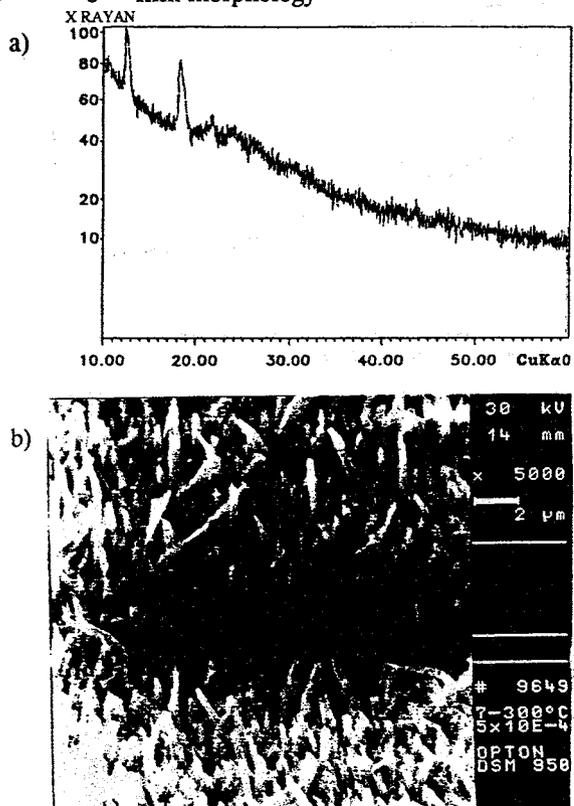


Fig. 5.  $\beta$  - PcCu diffraction spectrum and morphology  
 a - diffraction spectrum (CuK $\alpha$  radiation)  
 b - film morphology

The  $\alpha$ -PcCu film depicted in Fig.3. was deposited at substrate temperature of about 100°C. Microscopic examination has shown that the film is homogeneous and formed by small particles ( $d \approx 100$ nm). X-ray diffraction pattern exhibits a single distinguished peak at  $2\theta = 24^\circ$ .

Two-phase  $\alpha$  and  $\beta$  film shown in Fig.4. was evaporated at substrate temperature of 200°C.

The  $\beta$ -PcCu layer illustrated in Fig.5. was deposited at substrate temperature of 300°C. The layer is also homogeneous but formed by needles ( $L \geq 1\mu$ m). Two distinct peaks at  $2\theta = 12.5$  and  $2\theta = 18.5$  respectively are visible in the X-ray diffraction pattern.

We varied the coating layer thickness from 0.1 $\mu$ m to 1.5 $\mu$ m and its density from 0.17g/cm<sup>3</sup> to 1.6g/cm<sup>3</sup>.

In this paper we present as an example the results concerning only well ordered monocrystalline films of  $\alpha$  and  $\beta$ -PcCu.

### Temperature Stability of Generated Frequency

The temperature stability of SAW delay line generated frequency was investigated in the temperature range from 20°C to 120°C.

The frequency of a single oscillator and the differential frequency were continuously measured during the heating process. The measurement was stopped approximately 1 hour after temperature stabilization. Typical plots are shown in Figs. 6 - 8.

The frequency change of coated and uncoated delay lines on quartz substrate during heating up to the three different temperatures is illustrated in Fig.6.

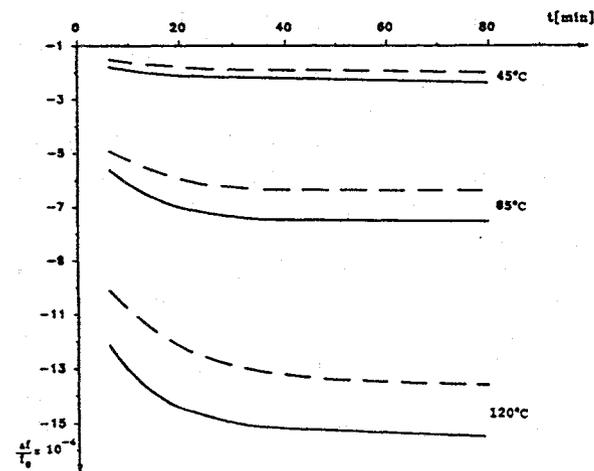


Fig.6. The relative frequency shift of PcCu-coated and uncoated delay lines on quartz substrate during heating up to three different temperatures

- $\Delta f$  - frequency shift
- $f_0$  - frequency of generation at room temperature
- - uncoated delay line
- - PcCu - delay line coating  $\alpha$  - PcCu, mass -  $3.2 \times 10^{-6}$  g

The frequency versus time curves for the uncoated and PcCu-coated delay lines almost coincide for temperatures below 45°C.

The discrepancy between the curves grows with the increase of temperature. The shape of the curves indicates that the differential frequency increases with the temperature.

Heating causes frequency decrease of a single oscillator. This is true regardless of the PcCu morphology and type of the substrate, however the value of frequency change depends on the structure of the layer.

Figure 7. illustrates the relative frequency shift of PcCu-coated delay line on quartz substrate during heating up to 85°C.

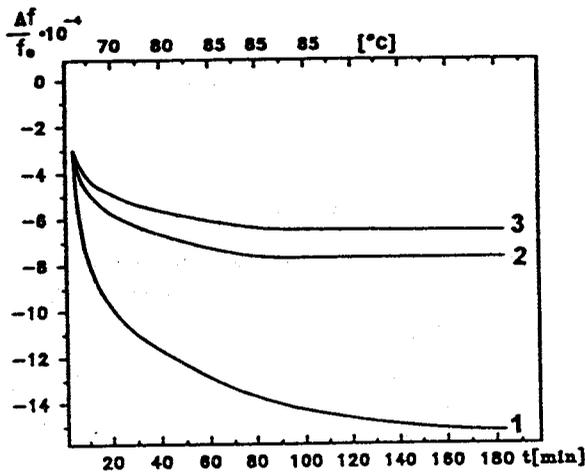


Fig. 7. The relative frequency shift of PcCu-coated delay line on quartz substrate during heating up to 85°C.

$\Delta f$  - frequency shift

$f_0$  - frequency of generation at room temperature

1 -  $\alpha$  - PcCu, coating mass -  $4.8 \times 10^{-6}$  g

2 -  $\alpha$  - PcCu, coating mass -  $3.2 \times 10^{-6}$  g

3 -  $\beta$  - PcCu, coating mass -  $4.8 \times 10^{-6}$  g

From the graph it follows that the value of frequency change with temperature is greater for the delay line covered with  $\alpha$ -PcCu film, than for the line employing the  $\beta$ -PcCu layer of the same mass.

The similar effect is observed in the case of devices using lithium niobate as a substrate, as can be seen in Fig. 8.

Of course the loading mass and the type of piezoelectric substrate affect the relation between the frequency shift and the temperature [ 8, 9 ].

Our measurement results show that for the investigated range of parameters the  $\beta$ -PcCu layer provides several times better temperature stability of SAW delay line generation frequency (hence also the stability of differential frequency) than  $\alpha$ -PcCu layer of the same mass.

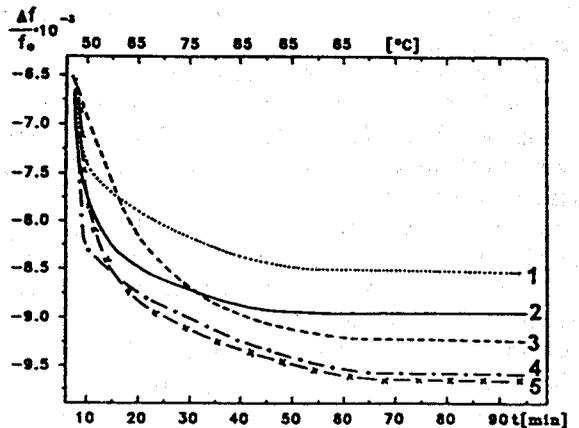


Fig. 8. The relative frequency shift of PcCu-coated delay line on  $\text{LiNbO}_3$  substrate during heating up to 85°C.

$\Delta f$  - frequency shift

$f_0$  - frequency of generation at room temperature

1 -  $\beta$  - PcCu, coating mass -  $5 \times 10^{-7}$  g

2 - uncoated delay line

3 -  $\alpha$  - PcCu, coating mass -  $1 \times 10^{-7}$  g

4 -  $\alpha$  - PcCu, coating mass -  $3 \times 10^{-7}$  g

5 -  $\alpha$  - PcCu, coating mass -  $5 \times 10^{-7}$  g

The difference in behavior between the  $\alpha$  and  $\beta$ -PcCu coated delay lines can be explained by the difference in the value of stress imposed by the two coatings on the substrate surface [ 10 ]. Small round grains of  $\alpha$ -PcCu coating form densely packed film. In contrast to  $\alpha$ -structure the film of  $\beta$ -PcCu is not continuous. For this reason for  $\beta$ -PcCu the influence of elastic properties of the film and its thermal expansion on the stress is much weaker than for  $\alpha$ -PcCu.

At constant temperature the differential frequency drift is practically independent of the PcCu film parameters (see Table 2).

Table 2. The differential frequency drift.

Temperature [°C]	Differential frequency drift	
	Delay line on quartz	Delay line on lithium niobate
20	$\leq 1$ Hz / min	$\leq 2$ Hz / min
45	$\leq 2$ Hz / min	$\leq 6$ Hz / min
120	$\leq 3$ Hz / min	$\leq 7$ Hz / min

### Sensor Sensitivity and Response Time

The sensor sensitivity was determined as the value of differential frequency shift during the film exposition on the reaction with  $\text{NO}_2$ , till saturation took place.

The relation between the sensor parameters and the structure of PcCu films will be presented for  $\alpha$  and  $\beta$ -PcCu layers of the same mass deposited on lithium niobate substrates.

The frequency change as a function of time was measured for  $\text{NO}_2$  concentration in the carrier gas equal to 10ppm. The flow rate was 1l/min. These measurements were taken at room temperature. Typical plots are shown in the Figs.9-11.

It can be seen that the response time and the sensitivity strongly depend on the film structure and its morphology, but the largest sensitivities do not correspond to shortest response times.

The response time of SAW device designed on lithium niobate employing  $\alpha$ -PcCu (Fig.9), was shorter than 5min., whereas that of  $\beta$ -PcCu was approximately 50min. (Fig.10). On the other hand the measured sensitivity of  $\beta$ -PcCu film was several times higher than the sensitivity of  $\alpha$ -PcCu film.

The response rate of SAW sensors is determined by the physical and chemical processes taking place at the chemical interface. The physisorption, the pore diffusion, the charge creating process and its transfer can be the factors which limit the response time.

Contrary to the  $\beta$ -PcCu film the  $\alpha$ -PcCu layer has revealed low saturation levels due to densely packed film and small working surface. So in the case of  $\alpha$ -PcCu the response was rapid and in the case of  $\beta$ -PcCu it was slow.

To illustrate the influence of the type of piezoelectric substrate on the sensor parameters we compared the results of experiments carried out for  $\beta$ -PcCu layers deposited on quartz and lithium niobate in the same conditions (Figs 10, 11).

In the case of SAW device based on quartz the frequency change is significantly smaller than in the case of lithium niobate with the similar response time.

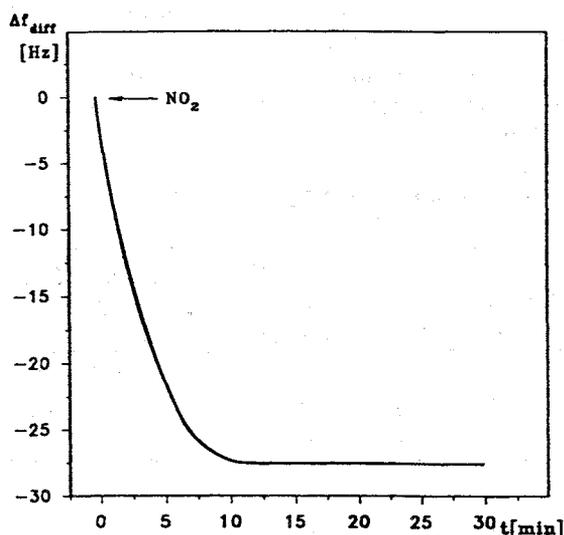


Fig. 9. The differential frequency change versus time for  $\alpha$ -PcCu covered SAW device on lithium niobate (coating mass -  $5 \times 10^{-7}$  g)

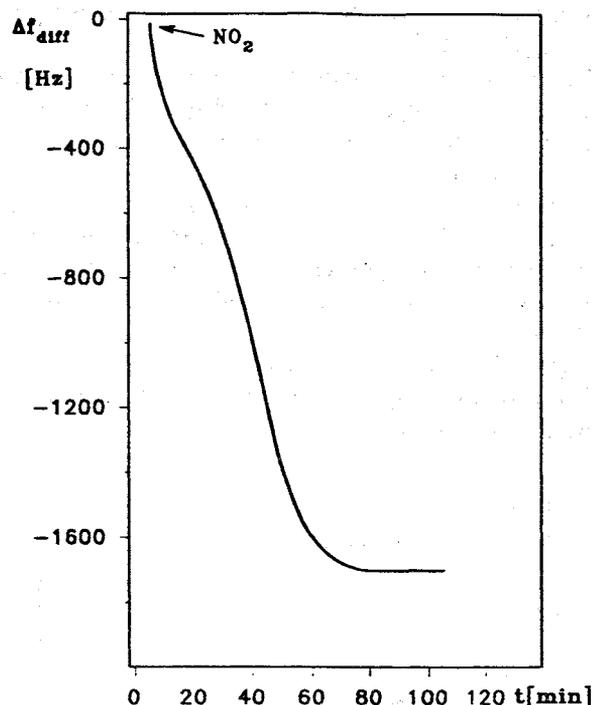


Fig. 10. The differential frequency change versus time for  $\beta$ -PcCu SAW device on lithium niobate (coating mass -  $5 \times 10^{-7}$  g)

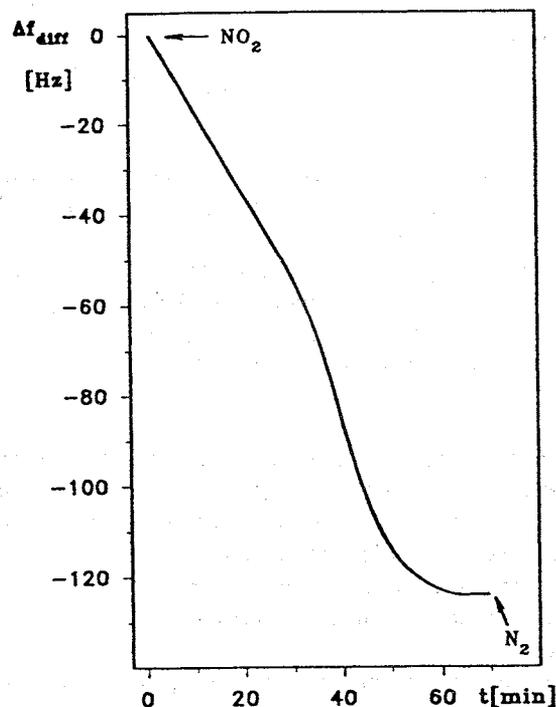


Fig. 11. The differential frequency change versus time for  $\beta$ -PcCu SAW device on quartz (coating mass -  $3.2 \times 10^{-6}$  g)

The SAW sensor fractional frequency shift due to conductivity change is proportional to electromechanical coupling coefficient, which is a measure of the piezoelectric strength of the substrate.

The electromechanical coupling for  $\text{LiNbO}_3$  is about 40 times higher than for quartz.

The fractional frequency shift due to mass loading is proportional to the mechanical coupling constant which is several times lower for lithium niobate than for quartz. The larger the electromechanical coupling the greater the influence of layer's conductivity change on sensor's response. The larger the mechanical coupling the greater the effect of change in loading.

For our experimental situation both mass loading and conductivity changes may cause a sensor response. In the case of low concentration of  $\text{NO}_2$ , the effect of change in mass was insignificant in comparison with the effect of conductivity change (therefore the sensitivity of SAW device on quartz was much smaller than on  $\text{LiNbO}_3$ ), but if the concentration is greater than about 100ppm conductivity change will cease to be a function of  $\text{NO}_2$  concentration.

### CONCLUSION

The results of this work confirm that SAW delay line used as a reference element to compensate for the ambient temperature variations is not a universal solution and should be implemented with great caution. Our measurements show that for the investigated range of parameters the  $\beta$ -PcCu layer provides several times better sensor sensitivity and temperature stability of SAW delay line generation frequency than  $\alpha$ -PcCu layer of the same mass.

It seems interesting to try to optimize  $\text{NO}_2$  detection by simultaneously using both lithium niobate and quartz sensors in the same detection system, utilizing  $\beta$ -PcCu as the chemically sensitive layer. This should make it possible to widen the measurement range of gas concentration and to increase the sensor sensitivity as well as selectivity.

### Acknowledgements

Special acknowledgements are due to Waldemar Soluch and Piotr Nagłowski for the interest in this work and valuable assistance.

### REFERENCES

- [1] Arn D. et al, 1992, "Surface acoustic wave gas sensors: developments in the chemical industry". Sensors and Actuators, B. 8 27-31.
- [2] von Shickfus M. et al, 1994, "Improving the SAW gas sensor: device, electronics and sensor layer". Sensors and Actuators, B. 18 -19, 443 - 447.
- [3] Ricco A.J., 1994, "SAW Chemical Sensors", The Electrochemical Society Interface, Winter 38-44.
- [4] Falconer R.S., 1995, "A versatile SAW-based sensor system for investigating gas-sensitive-coatings". Sensors and Actuators, B. 24-25 54-57.
- [5] Zhou R. et al, 1993, " $\text{NO}_x$  sensitivity of monomeric and polymeric-N-macrocylic compounds". Sensors and Actuators, B. 15-16 312-316.
- [6] Shu Ji Qin and Bott B., 1991 "The sensitivity to  $\text{NO}_2$  of sandwich devices based on lead phthalocyanine and copper phthalocyanine". Sensors and Actuators, B. 3 255-260.
- [7] Passard M. et al, 1994, "Gas sensitivity of phthalocyanine thin films". Sensors and Actuators, B. 18-19, 489-492.
- [8] Liron Z. et al, "Temperature effect and chemical response of surface acoustic wave (SAW) single-delay-line chemosensors". Sensors and Actuators, B.12, 115 - 122.
- [9] Mrwa A. et al, 1995, "Response of lead phthalocyanine to high  $\text{NO}_2$  concentration". Sensors and Actuators, B.24-25, 596-599.
- [10] Bartley D.L. and Dominquez, 1990, "Elastic effects of polymer coatings on surface acoustic waves". Anal. Chem., 62 1649 - 1656.

## MEASUREMENT OF MECHANICAL QUANTITIES USING QUARTZ SENSORS

R Grossmann, J Michel, T. Sachs, E. Schrüfer

Chair for Electrical Measurement, Technical University of Munich

D - 80 290 Munich, Germany

### ABSTRACT

Quartz crystals supply a resonant frequency highly stable in time. However, temperature, strain, etc. can affect this frequency. Usually a quartz crystal is part of an oscillator and housed in order to avoid any influence of the environment. On the other hand by clamping it in the most sensitive direction it may work as a sensor for strain and thus for any mechanical quantity.

So far quartz sensors have no advantage to conventional wire strain gauge. The important difference is that quartz crystals can mechanically store energy. Once they are provided with a certain amount of energy they can operate for a while without any active part, i.e. without power supply or oscillators. Provided with this capability and considering the high resonant frequency quartz sensors are ideally suited for wireless measurements. This allows applications on inaccessible or moving objects, e.g. torque measurement on rotating shafts or tyre pressure measurement while driving.

There are three kinds of quartz sensors to be presented: SAW (surface acoustic wave) delay lines, SAW resona-

tors and BAW (bulk acoustic wave) resonators. The measurement principle is the same for all of them. A transceiver sends RF pulses over an antenna in order to stimulate the sensor that produces a pulse response. There is enough energy to retransmit this signal for some time. In the case of the sensor being a resonator, the signal consists of a damped harmonic oscillation with the sensor's resonant frequency, whereas a SAW delay line produces several echo pulses after defined transit intervals. In either case the information on the measured quantity (strain) is contained in the signal. Considering known geometrical and material properties any interesting mechanical quantity is determinable. The advantage of this system is based on the passive operation of the sensor which only has to be coupled with an antenna but works without direct power supply or other devices. So very small and very reliable arrangements are possible. The achieved resolution comes to 0.5% with measurement rates of up to  $10^4$  per second. If two sensors are used, an effective temperature compensation can be ensured.

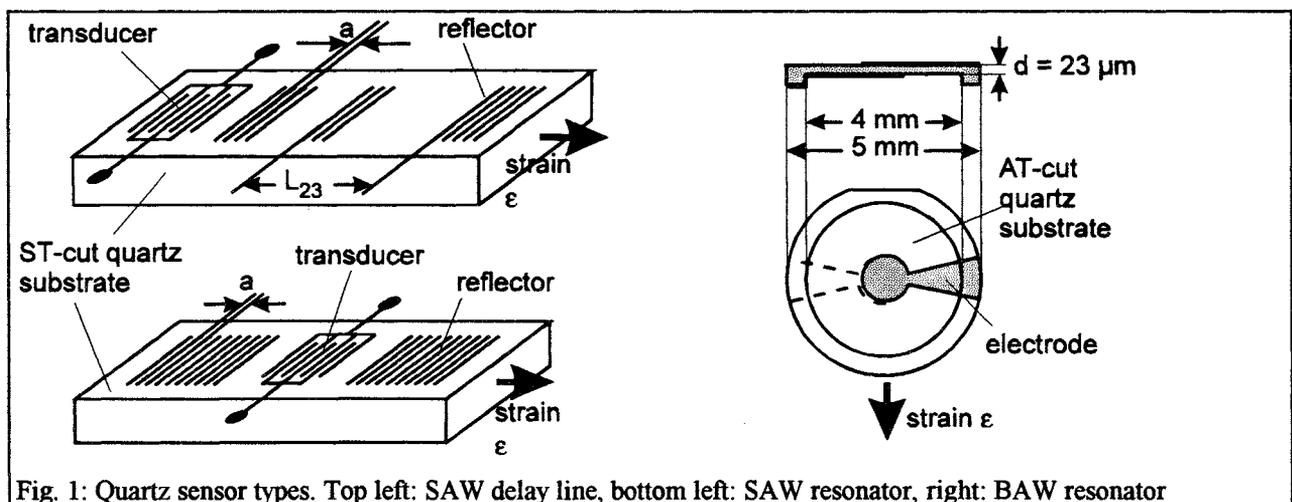


Fig. 1: Quartz sensor types. Top left: SAW delay line, bottom left: SAW resonator, right: BAW resonator

## SENSORS

### SAW Delay Lines

Generally, a SAW device consists of a strip of piezoelectric material, in this case ST-cut quartz (fig. 1). On its surface an interdigital structure works as transducer between electrical and mechanical fields. The distance  $a$  between the digits determines the system frequency  $f_0$ . A delay line furthermore contains reflectors (fields of aluminium strips). An RF pulse with the system frequency at the transducer produces a surface acoustic wave train. It propagates like an earth quake wave along the delay line and is partly reflected at each reflector. After certain intervals the echo pulses arrive at the transducer where they effect an electrical signal. The transit times  $T_{ik}$  depend on the distances  $L_{ik}$  between the corresponding reflectors and the propagation velocity  $v$ . The phase difference between two echoes follows from

$$\varphi_{ik} = 2\pi f_0 \cdot T_{ik} = 2\pi f_0 \cdot \frac{2L_{ik}}{v} \quad (1)$$

A strain  $\varepsilon$  in  $x$ -direction leads to a change of geometry and - influencing material properties - propagation velocity. This results in a change of the phase differences:

$$\frac{\Delta\varphi_{ik}}{\varphi_{ik}} = \frac{\Delta L_{ik}}{L_{ik}} - \frac{\Delta v}{v} \approx 1.24\varepsilon \quad (2)$$

However, the sensitivity in directions orthogonal to propagation is negligible.

As the surface wave only runs on its top, a SAW device may be glued with its bottom side onto the object the strain of which is to be determined. The glue transfers up to 99% of the strain from the object to the sensor,

### SAW resonator

Beside the delay line, there is also a SAW resonator. It is shaped, too, as a small beam of ST-cut quartz. How-

ever, its transducer is positioned in the middle and the reflector fields on either side cover nearly the whole top. An induced acoustic wave with the system frequency causes a standing wave by constructive interference. The resonant frequency is given by the interdigital distance  $a$  and the propagation velocity  $v$ :

$$f_{res} = \frac{v}{2a} \quad (3)$$

The pulse response of a resonator is a damped harmonic oscillation. The damping originates from resonant and load resistance or an antenna. A strain  $\varepsilon$  in  $x$ -direction effects a change  $\Delta f$  of the resonant frequency:

$$\frac{\Delta f}{f_{res}} = -\frac{\Delta a}{a} + \frac{\Delta v}{v} \approx -1.24\varepsilon \quad (4)$$

Like a delay line, a SAW resonator shows no remarkable sensitivity in other directions than  $x$ . The mounting onto an object is as easy as with a delay line.

### BAW resonator

A quartz crystal may not only be stimulated to surface waves, but can perform several modes of volume oscillations (bulk acoustic waves, BAW). By far the highest resonant frequency is achieved using the thickness shear mode. For this one usually uses thin discs of AT-cut quartz with electrodes on top and bottom. The quartz may have an enforced rim for easier mounting. The resonant frequency shows the same dependence of wavelength and wave speed as a SAW resonator. Since in the case of a BAW there are no coupled modes, one can replace the wave speed with an expression containing but material properties:

$$f_{res} = \frac{v}{2d} = \frac{1}{2d} \sqrt{\frac{c_{66}}{\rho}} \quad (5)$$

Ion etching makes frequencies up to 200 MHz possible. The stiffness  $c_{66}$  is responsible for shear movements. In  $x$ -direction this stiffness is extremely sensitive to strain:

$$c_{66}(\varepsilon) = c_{66} + c_{661} \cdot \varepsilon \quad (6)$$

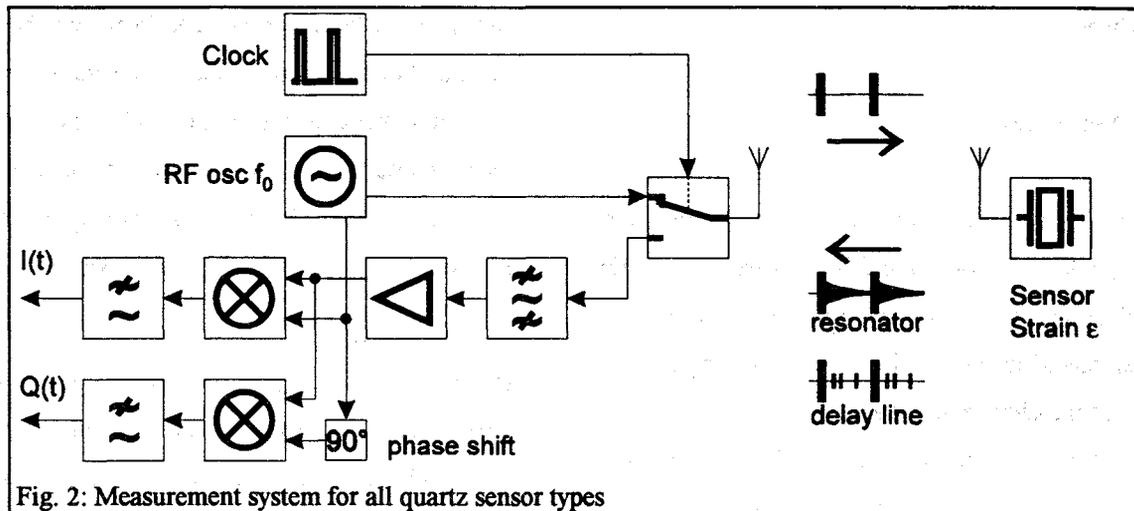


Fig. 2: Measurement system for all quartz sensor types

Neglecting other (geometrical) effects this induces the relative frequency shift

$$\frac{\Delta f}{f_{\text{res}}} = \frac{1}{2} \frac{c_{661}}{c_{66}} = -3.6\epsilon \quad (7)$$

Hence, and considering that the quartz plates are very thin (e.g. 23 μm for a resonant frequency of 72 MHz), BAW sensors are especially suited for measurements of small quantities (e.g. < 1 mN). However, since the active volume must be able to freely vibrate, either the quartz needs an enforced rim or one has to carve a hollow into the measured object or use a special mounting. Interesting are applications on a diaphragm, which amplifies forces onto its surface by transforming them into a bending momentum. The thin quartz plate does not interfere with the bending of the diaphragm. This layout fits well for pressure measurements.

## MEASUREMENT SYSTEM

Usually resonators operate as the frequency determining device of an oscillator. However, since quartz sensors are capable of storing energy and their resonant frequencies are in the RF range, a wireless operation mode suggests itself. The principle is the same for the three kinds of quartz devices (fig. 2): The sensor is attached only to an antenna. At regular intervals a transceiver emits RF pulses close to the resonant fre-

quency  $f_{\text{res}}$ , which stimulate the sensor (fig. 3). Its antenna transmits its pulse response back to the transceiver. As the bandwidth of the sensor signal is tight, it is advisable to mix it down in an IF range with the emitted frequency  $f_0$  in order to increase the accuracy of the measurement. In case a quadrature demodulator is used, one can avoid ambiguities in the frequency position of the signal. After demodulation the signals  $I(t)$  and  $Q(t)$  are digitally evaluated. For a delay line, the measured value  $\varphi_{ik}$  follows directly from the time domain:

$$\varphi_{ik} = \arctan \frac{Q(t)}{I(t)} \quad (8)$$

For resonators, the exact resonant frequency is obtained after a (complex) Fourier transform  $\text{FFT}\{I + jQ\}$ , in-

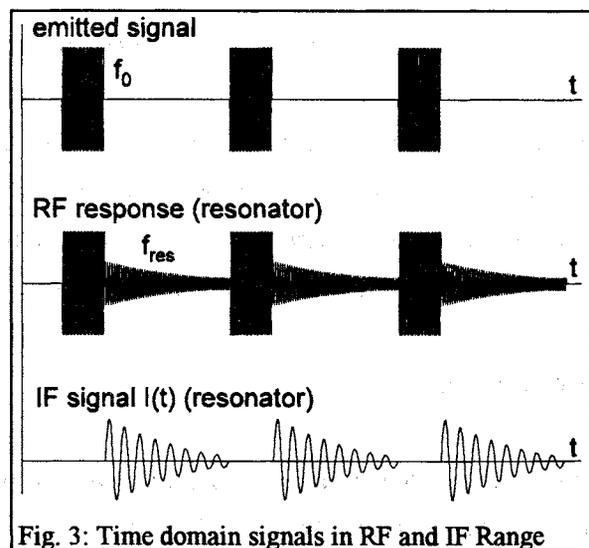
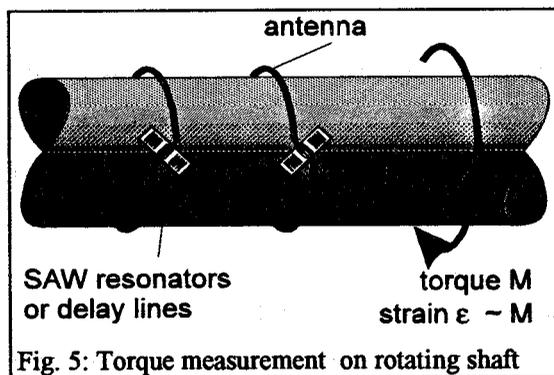
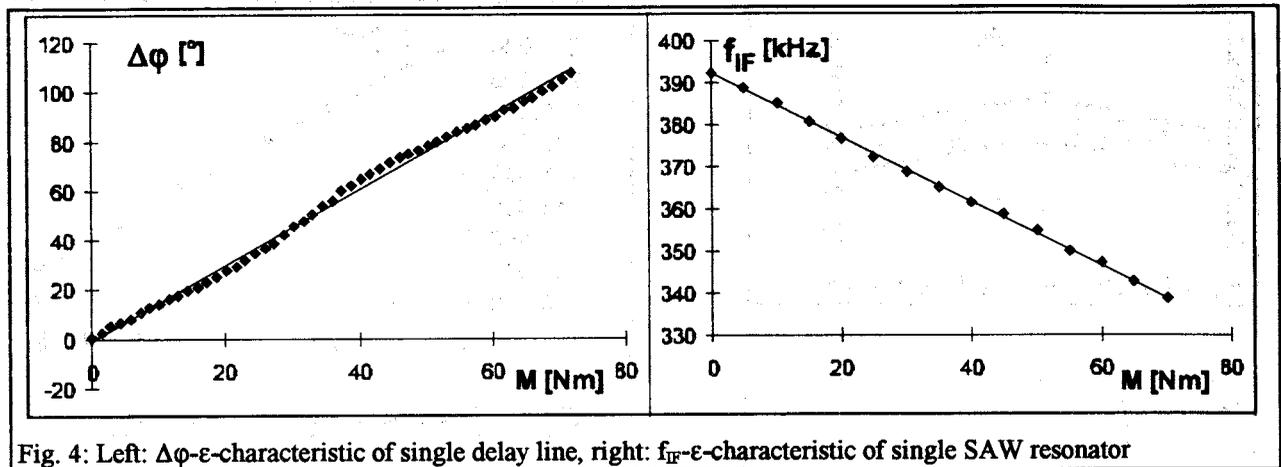


Fig. 3: Time domain signals in RF and IF Range



terpolation with second order terms and maximum search.

There are several conditions to satisfy for the measurement system to operate: The band width of the emitted RF pulse has for all possible strains to cover the sensor spectrum. However, a band width too broad results in a poor energy transfer into the quartz. Furthermore, an amplifier in the receiver branch must compensate for the insertion loss of the sensor (30 dB for resonators and 50 dB for delay line) and twice the free space attenuation (30 dB for antenna distance 30 cm).

## APPLICATIONS

### Torque measurement on a rotating shaft

Often it is important to measure the torque on a rotating shaft, e.g. for boost development, process control, stress analysis or power measurement. Torque causes a

torsion on the shaft, which results in strain proportional to the effecting torque at  $\pm 45^\circ$ -angles with respect to the axis of rotation.

For the measurement (fig. 5) two sensors of either the SAW resonator and delay line are glued onto the shaft at  $\pm 45^\circ$  angles and, for comparison, a conventional torque sensor has been added. Each SAW sensor has its own antenna wound around the shaft. The mounting of the sensors has hence hardly an effect on the measurement. Another advantage of this layout is that there is no hard wear.

The SAW resonator has a sensitivity of  $-760$  Hz/Nm (fig. 4) with variations of about 1 kHz at the resonant frequency of 433 MHz (390 kHz in IF range after demodulation). The driving range comes to 200 kHz, thus the resolution of the system is 0,5%. The power emitted by the transceiver is about 25 mW. The delay line performs not so good and needs more emitted power (1 W), but still grants an accuracy of 1%. On the other hand, the measurement rate of a delay line can be  $10^4$  per second, which is 10 times more than that of a resonator. The sensitivity of a single delay line can be determined to  $1.5^\circ$ /Nm (fig. 4).

The second sensor in both cases is needed for temperature compensation. A change of temperature effects before all a strain equal for either sensor.

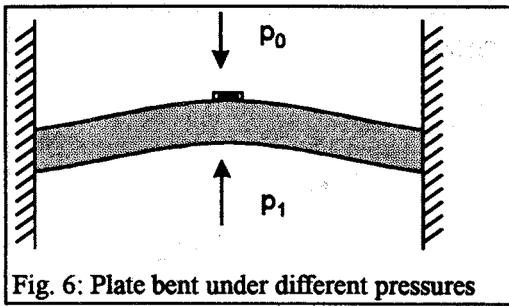


Fig. 6: Plate bent under different pressures

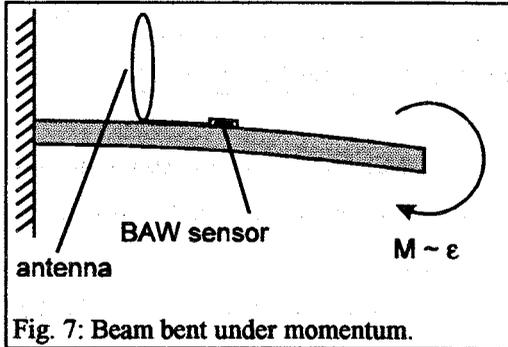
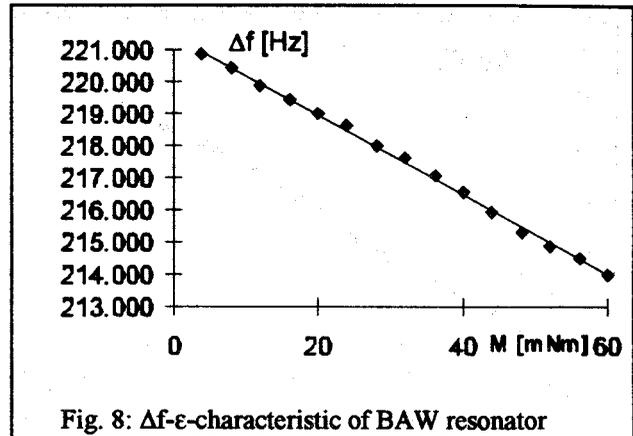


Fig. 7: Beam bent under momentum.

#### Measurement of bending momentum

The measurement of pressure in moving parts such as tyres plays an important role for safety and reliability. A layout well suited to this job is a diaphragm with a BAW sensor glued on its top (fig. 6). The different pressures on either side cause a bending and hence strain, which can be determined. In this context, the bending momentum of a plastic beam has been measured (fig. 7). The sensor was glued onto a hollow in the middle of the beam and connected to a small antenna. This situation is equivalent to bending under pressure, but easier to realise.

The measured characteristic of the BAW sensor (fig. 8) with the resonant frequency of 72 MHz ( $f_{IR} = 220$  kHz) illustrates the high sensibility to small quantities; in this case a bending moment of 1 Nm would effect a frequency shift of about 3 MHz and surely destroy the sensor. The sensibility originates partly in the larger dependence of strain and partly in the small thickness of the quartz disc.

Fig. 8:  $\Delta f$ - $\epsilon$ -characteristic of BAW resonator

#### COMPARISON

Each sensor is very small and will hardly influence a measurement. They all operate wirelessly, passively and reliably. Several sensors with slightly different system frequencies can be connected parallelly to one single antenna.

The advantage of the SAW delay line lies in its high measurement rate, but comes along with a lower efficiency, which demands more emitted power. The SAW resonator shows a good sensitivity and low power consumption. Either type of SAW sensor can easily be mounted, as the bottom face is passive and may be glued.

Ideal for measurements of bending or any small value is the BAW resonator. For mounting the quartz needs either an enforced rim or a special clamping, or the measured object must have a hollow. Though it is a little bit more difficult to handle with, it offers by far the highest sensitivity.

The temperature dependence of each sensor type is reproducible; in most cases thermal strain from the mounting dominates thermal variations of material properties. It can easily be eliminated using two sensors.

sensor type	SAW delay line	SAW resonator	BAW resonator
dimensions [LxWxH]	11 x 3 x 0.6 mm <sup>3</sup>	8 x 3 x 0.6 mm <sup>3</sup>	5 x 5 x 0.023 mm <sup>3</sup>
mounting	direct	direct	clamping / rim
system frequency	856 MHz	433,9 MHz	72 MHz
evaluated quantity	phase	frequency	frequency
evaluated domain	time	frequency (FFT)	frequency (FFT)
measurement rate	10 <sup>4</sup> /sec	10 <sup>3</sup> /sec	10 <sup>3</sup> /sec
rel. sensitivity	$\Delta\phi/\phi = 1.1 \epsilon$	$\Delta f/f = -1.1 \epsilon$	$\Delta f/f = -3.8 \epsilon$
efficiency $P_{out}/P_{in}$	10 <sup>-5</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>
instability	< 1 %	< 0.5 %	< 0.5 %

Table 1: Comparison among the three discussed sensor types

## REFERENCES

- [1] J Michel, T Sachs, E Schrifer, 1996: „Drehmomentmessung mit passiven Oberflächenwellenresonatoren“, 8. GMA/ITG-Fachtagung Sensoren und Meßsysteme, Bad Nauheim
- [2] S Datta, 1986: „Surface Acoustic Wave Devices“, Englewood Cliffs
- [3] C Reichinger, 1995: „Berechnung und Einsatz von piezoelektrischen Schwingern zur Kraftmessung“, Dissertation, München
- [4] R Grossmann, 1995: „Messung mechanischer Größen mit funkabfragbaren Quarzresonatoren“, 40. Internationales Wissenschaftliches Kolloquium, Ilmenau
- [5] J Tichý, G Gautschi, 1980: „Piezoelektrische Meßtechnik“, Springer, Berlin
- [6] Landolt-Börnstein, 1979: „Elastische, piezoelektrische, pyroelektrische, piezooptische, elektrooptische Konstanten und nichtlineare dielektrische Suszeptibilitäten von Kristallen“, vol. III/11, Springer, Berlin

## COMPARISON OF TWO-WAY SATELLITE TIME AND FREQUENCY TRANSFER AND GPS COMMON-VIEW TIME TRANSFER DURING THE INTELSAT FIELD TRIAL

J A Davis<sup>1</sup>, W Lewandowski<sup>2</sup>, J A De Young<sup>3</sup>, D Kirchner<sup>4</sup>, P Hetzel<sup>5</sup>, T Parker<sup>6</sup>, W Klepczynski<sup>3</sup>, G De Jong<sup>7</sup>, A Söring<sup>8</sup>, F Baumont<sup>9</sup>, K A Bartle<sup>1</sup>, H Ressler<sup>10</sup>, R Robnik<sup>10</sup> and L Veenstra<sup>11</sup>.

National Physical Laboratory, Queens Road, Teddington, Middlesex, UK<sup>1</sup>,  
 Bureau International des Poids et Mesures, Sevres, France<sup>2</sup>,  
 United States Naval Observatory, Washington, DC, USA<sup>3</sup>,  
 Technical University Graz, Graz, Austria<sup>4</sup>,  
 Physikalisch-Technische Bundesanstalt, Braunschweig, Germany<sup>5</sup>,  
 National Institute of Standards and Technology, Boulder, Colorado, USA<sup>6</sup>,  
 NMI, Van Swinden Laboratorium, Delft, The Netherlands<sup>7</sup>,  
 Forschungs- und Technologiezentrum, Deutsche Telekom, Darmstadt, Germany<sup>8</sup>,  
 Observatoire de la Cote d'Azur, Grasse, France<sup>9</sup>,  
 Space Research Institute, Graz, Austria<sup>10</sup>,  
 RSI, Comsat World Systems, 6560 Rock Spring Drive, Bethesda, Maryland, USA<sup>11</sup>.

### ABSTRACT

For a decade and a half GPS Common-View time transfer has greatly served the needs of primary timing laboratories for regular inter-comparisons of remote atomic clocks. However, GPS as a one-way technique has natural limits and may not meet all challenges of the comparison of the coming new generations of atomic clocks. Two-Way Satellite Time and Frequency Transfer (TWSTFT) is a promising technique which may successfully complement GPS. For two years, regular TWSTFTs have been performed between eight laboratories situated in both Europe and North America, using INTELSAT satellites. This has enabled an extensive direct comparison to be made between these two high performance time transfer methods. The performance of the TWSTFT and GPS Common-View methods are compared over a number of time transfer links. These links use a variety of time transfer hardware and atomic clocks and have baselines of substantially different lengths. The relative merits of the two time transfer systems are discussed.

### 1) INTRODUCTION

The comparison of ultra-stable atomic clocks located at distant laboratories requires the development of high precision methods of time and frequency transfer. For many years the use of common-view of GPS satellites has been the only method used regularly to compare atomic clocks contributing to International Atomic Time (TAI) (Lewandowski and Thomas (1)). The recent development of Two-Way Satellite Time and Frequency Transfer (TWSTFT) has provided a wholly independent method of time transfer that may be used both as an independent check on the GPS common-view method and ultimately as an alternative method (Kirchner (2)).

An increasing number of primary timing laboratories are commissioning TWSTFT systems. TWSTFT offers several advantages over the one-way GPS common-view method. The use of small directional antennae with the

TWSTFT system enables high C/No links to be established, resulting in low measurement noise. The use of a two-way time transfer method either eliminates or substantially reduces many of the systematic errors associated with one-way time transfer methods. These include errors associated with earth station coordinates, satellite coordinates, ionospheric and tropospheric delays.

The cost of both commissioning and operating TWSTFT instrumentation is however significantly higher than the costs of GPS receiving instrumentation. A small transmit and receive satellite earth station must be commissioned at each laboratory. Satellite time must also be purchased on a commercial geostationary satellite. The performance of a TWSTFT system is of importance when justifying the additional costs required over GPS receiving equipment.

Preliminary comparisons of regular TWSTFT and GPS common-view time transfers, made between five laboratories have been reported (Davis et al (3)). In this paper the work is extended. Comparisons are made on both European and transatlantic time transfer links. Values of  $\sigma_y$  are calculated from the TWSTFT measurements, GPS measurements and the (TWSTFT-GPS) differences. The results obtained are explained in terms of clocks used at each location, and the characteristics of the time transfer methods. During the TWSTFT campaign, the Intelsat satellite was upgraded from the VA-F13 satellite to the VII-F6 satellite. The effects of this change on the TWSTFTs are examined. At present TWSTFT is performed between all possible combinations of laboratories. Operating a reduced schedule would result in some inter-comparisons being made using intermediate stations. The characteristics of links performed via intermediate stations are compared with direct TWSTFT links. Finally discrete changes in either the TWSTFT or GPS time transfers are considered in terms of the delay changes in the corresponding instrumentation.

Station	Clock	TWSTFT Modem	GPS Receiver
TUG	HP5071A	Mitrex 2500	STI TTS-502 No 4 and NBS 03
NPL	Hydrogen Maser	Mitrex 2500	AOA TTR5A NBS type
PTB	Primary Caesium	Mitrex 2500A	Rockwell Collins NBS
NIST	Composite	Mitrex 2500	NBS Prototype No 10
USNO	Hydrogen Maser	Mitrex 2500A	STI 502B

Table 1: Atomic clocks and TWSTFT instrumentation

## 2) METHOD

Regular TWSTFT and GPS common-view measurements have been performed over a period of two years. Six European and two North American laboratories have been participating in the measurement campaign. The results presented in this paper have been obtained from inter-comparisons of data from five of the Laboratories. These were the Technical University Graz, Graz, Austria (TUG), National Physical Laboratory, Teddington, UK (NPL), Physikalisch- Technische Bundesanstalt, Braunschweig, Germany (PTB), National Institute of Standards and Technology, Boulder, Colorado, USA (NIST), United States Naval Observatory, Washington, DC, USA (USNO). The atomic clocks, GPS receivers and TWSTFT modems used at each location are summarised in Table 1. Two years cost-free satellite time was provided by INTELSAT. Two separate schedules of European and transatlantic TWSTFT measurements were each performed three times per week, starting at 10:00 UTC and 14:00 UTC or 15:00 UTC respectively. Each individual time transfer lasted for five minutes. GPS measurements were made according to the BIPM International GPS common-view schedules 22, 23, 24 and 25.

## 3) DATA ANALYSIS

The characteristics of the TWSTFT and GPS data sets are significantly different. The measurement noise of the TWSTFT data is much lower. TWSTFT data consist of spot measurements of five minutes duration made every two or three days. In contrast, GPS data is obtained from the processing of a series of up to 30 thirteen-minute measurements, spread throughout each day. Several days data may contribute to the final time transfer. Underlying both sets of measurements are the variations of the atomic clocks. These short and medium term clock variations are small in the case of the Hydrogen Maser, but significantly

larger in the case of both primary and commercial caesium clocks.

TWSTFT measurements were made on Mondays, Wednesdays, and Fridays. This measurement schedule resulted in a regular but unevenly spaced data set. The conventional formula for the calculation of  $\sigma_y$  has to be modified in order to efficiently calculate  $\sigma_y$  values. The following equation is used to calculate  $\sigma_y$  from the available data set:

$$\sigma_y(\tau) = \frac{1}{\tau\sqrt{2}} \left( \frac{1}{N} \sum [a_1(y_1 - y_2) - a_2(y_2 - y_3)] \right)^{0.5} \quad (1)$$

where  $\tau = \frac{1}{N} \sum \tau_1^{1/2} \tau_2^{1/2}$ ,  $a_1 = \sqrt{\tau_2}/\sqrt{\tau_1}$  and  $a_2 = \sqrt{\tau_1}/\sqrt{\tau_2}$ .  $y_1$ ,  $y_2$  and  $y_3$  are successive elements of the data set.  $\tau_1$  is the intervals between the occurrence of  $y_2$  and  $y_1$ . Similarly  $\tau_2$  is the intervals between the occurrence of  $y_3$  and  $y_2$ .  $N$  is the number of elements in the summation. The above approximation has several important characteristics. When  $\tau_1 = \tau_2 = \tau$  equation (1) reverts to the standard formula for  $\sigma_y$  (A Gerber and A Ballato(4)). When applied to a data set consisting only of a linear drift, a zero result is obtained, again similar characteristics to the standard  $\sigma_y$  function.

A detailed description of the processing of the GPS common-view data is given in (3) including the use of Vondrak smoothing (Vondrak (5)). Longer transatlantic GPS links were not corrected for ionospheric measurements, more accurate tropospheric modelling and post-processed precise ephemeris. This is being done and should be addressed as part of a later study.

LINK	$\sigma$ (Total Period)	Subset (MJDs)	$\sigma$ (Subset)
NIST-NPL	6.9 ns	49600-49950	3.7 ns
NIST-PTB	5.9 ns	49600-49950	2.6 ns
USNO-NPL	4.0 ns	49387-49900	2.9 ns
USNO-PTB	4.9 ns	49485-49796	3.5 ns
USNO-TUG	5.6 ns	49387-49850	3.0 ns
PTB-NPL	2.5 ns	49600-49950	1.4 ns
PTB-TUG	3.4 ns	49567-49950	2.3 ns
NPL-TUG	2.4 ns	49471-49840	2.4 ns

Table 2: Standard deviations calculated from the (TWSTFT-GPS Common-View) differences.

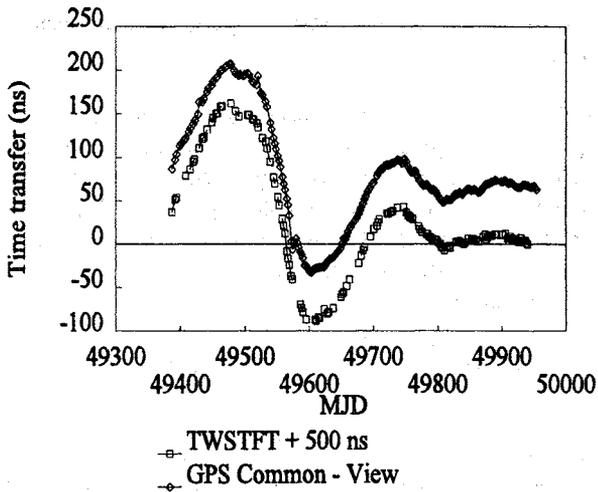


Figure 1: NIST-NPL TWSTFT and GPS common-view time transfer.

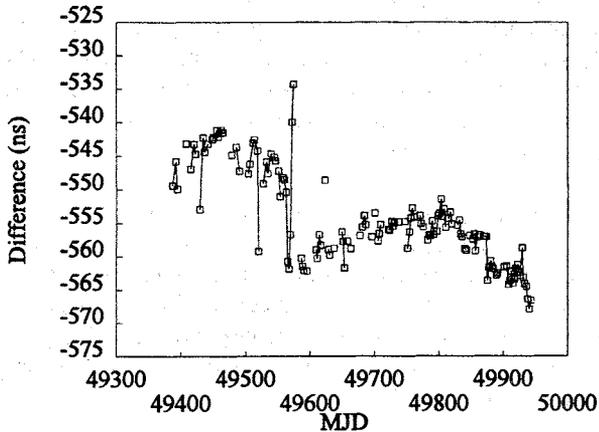


Figure 2: NIST-NPL TWSTFT - GPS common view differences.

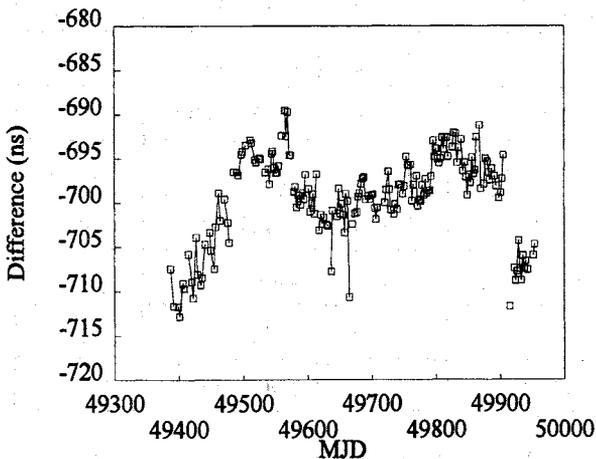


Figure 3: USNO-PTB TWSTFT-GPS common-view differences.

#### 4) RESULTS

##### 4.1) Time Transfers, Time Transfer Differences and Standard Deviations.

Curves of the (NIST-NPL) time transfer made over a two year period are shown in Figure 1. The offset between the two curves is due in part to the delay asymmetries of the TWSTFT instrumentation not being calibrated, and in part to an offset of 500 ns being added to display the two curves clearly. Curves of the (TWSTFT-GPS) differences, again with arbitrary offsets, are shown in Figures 2, 3 and 4 for the (NIST-NPL), (USNO-PTB) and (PTB-NPL) links.

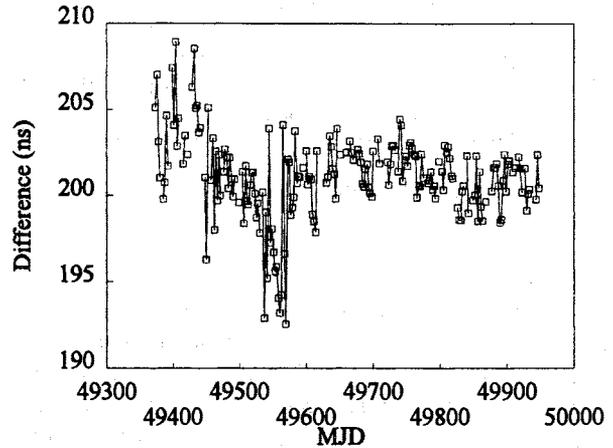


Figure 4: PTB-NPL TWSTFT - GPS common-view differences.

Several trends emerged. There is good agreement in the shape of the time transfer curves obtained using the TWSTFT and GPS common-view methods. Values of the standard deviation calculated from the (TWSTFT-GPS) differences are shown in Table 2, both for the complete data set and for a sub-section. Outlying points that deviated substantially from the mean value were removed before calculating the standard deviation using the points shown in Figures (2, 3 and 4). The NPL-TUG differences exhibited the lowest standard deviation when calculated over the whole period. In order to compare (TWSTFT-GPS) differences over shorter periods standard deviations have been calculated from subsets of the data which are free from known instrumentation changes. The results obtained were encouraging. Standard deviations of between 1.4 ns and 3.7 ns were obtained over periods ranging from 350 to over 500 days. The most stable operation occurred on the (PTB-NPL) link where a standard deviation of 1.4 ns was obtained for the (TWSTFT-GPS) differences, over a period of 350 days. Delay changes occurring in the NIST-NPL (TWSTFT-GPS) differences (Figure 2) correlated with known but uncalibrated TWSTFT instrumentation changes at NIST on MJD 49560 and MJD 49595. TWSTFT instrumentation delay changes also occurred at USNO on MJD 49485 and MJD 49842, corrections have been applied before producing Figure 3.

#### 4.2) Calculation of $\sigma_y$ values

Values of  $\sigma_y$  ( $\tau = 2.3$  days) are calculated for TWSTFT, GPS Common-View time transfers and (TWSTFT-GPS) differences. Results are shown in Table 3. In order to average out the white phase noise present within the transatlantic GPS common-view time transfers the cut off period of the Vondrak smoothing may be as high as four days. In these cases the presence of common data prevents a meaningful value for  $\sigma_y$  being calculated when  $\tau = 2.3$  days. In these cases an averaging time of 4.7 days is used. When comparing against TWSTFT two successive TWSTFT readings were averaged before calculating the  $\sigma_y$  values. Under these circumstances no additional benefit would be obtained from the direct calculation of  $\sigma_x$  as all the available data is averaged before calculating  $\sigma_y$ . The data sets contained only data that was collected on the same MJDs. When data was missing from one data set, the corresponding data was removed from the other data sets before processing. Any discrete delay steps occurring due to known instrumentation changes were removed before the  $\sigma_y$  values were calculated.

Values of  $\sigma_y$  varied considerably between the time transfers. The principle reason for this was the variation in the stability of the atomic clocks used at each location. The (NIST-NPL), and (USNO-NPL) links ( $\tau = 4.7$  days) had the lowest values of  $\sigma_y$ . These results were attributed to the use of an active Hydrogen Maser at NPL and USNO and real time composite clocks, heavily weighed in Hydrogen Masers at NIST. With these clocks the values of  $\sigma_y$  were lower for the TWSTFT time transfers when compared against GPS common-view time transfers for the (USNO-NPL) links and slightly higher for the (NIST-NPL) link. It must however be remembered that a four day cut-off period has been used for the Vondrak smoothing on the NIST-NPL GPS link, this may also have the effect of smoothing some of the clock noise. With Active Hydrogen Maser based clocks at each location, the  $\sigma_y$  values ( $\tau = 2.3$  days) for the (TWSTFT-GPS) differences were higher than the  $\sigma_y$  of the component time transfers, and were reasonably consistent with the relationship:

$$\sigma_y^2(\text{TWSTFT-GPS}) = \sigma_y^2(\text{TWSTFT}) + \sigma_y^2(\text{GPS}) \quad (2)$$

This suggests that there is no correlation between the instabilities in the TWSTFT and GPS common-view time transfers. The clock instabilities on these links are small ( $\tau = 4.7$  days) and the instabilities are due to uncorrelated delay changes occurring in the two time transfer systems.

Most of the time transfer links under study operated with either a primary caesium clock or commercial HP5071A caesium clock at at least one location. With averaging times ( $\tau$ ) of 2.3 days the principle instability contributing to the  $\sigma_y$  values was then clock noise.

For a given time transfer using caesium clocks, values of  $\sigma_y$  were in almost all examples lower for the GPS common-view measurements when compared against the

TWSTFT measurements. In most cases the difference in  $\sigma_y$  values was quite small but clearly significant. Despite the lower  $\sigma_y$  values obtained from the GPS time transfers, the conclusion should not be drawn that the GPS common-view method offers the best technique for comparison between caesium clocks. The lower  $\sigma_y$  values is most likely due to the choice of TWSTFT and GPS measurement schedules, rather than to an intrinsically higher accuracy of the GPS method (3). A significant amount of data is averaged before computing the  $\sigma_y$  values from GPS common-view data. In contrast, only a relative small amounts of data is averaged before computing  $\sigma_y$  from the TWSTFT data. This averaging of data results in a lower  $\sigma_y$  when calculated from the GPS common view method.  $\sigma_y$  values calculated from co-located clock measurements have been used to illustrate this result (3). In all examples involving caesium atomic clocks, the values of  $\sigma_y$  ( $\tau = 2.3$  days) obtained from the (TWSTFT-GPS) difference were significantly lower than the  $\sigma_y$  values obtained from the individual time transfers. This indicated that when using caesium clocks, the major contribution to the  $\sigma_y$  values is from clock noise. A significant proportion of this noise cancels in the (TWSTFT-GPS) differences, due to the partial elimination of noise from the clocks.

Link	TWSTFT $\sigma_y \times 10^{14}$	GPS $\sigma_y \times 10^{14}$	TWSTFT-GPS $\sigma_y \times 10^{14}$
	$\tau = 2.3$ days	$\tau = 2.3$ days	$\tau = 2.3$ days
NIST-PTB	1.7	0.7	1.5
USNO-PTB	1.5	1.0	1.1
USNO-TUG	2.3	2.1	1.7
PTB-NPL	1.8	0.8	1.7
NPL-TUG	2.4	2.3	1.3
PTB-TUG	2.9	2.6	1.5
	$\tau = 4.7$ days	$\tau = 4.7$ days	$\tau = 4.7$ days
NIST-NPL	0.5	0.4	0.7
USNO-NPL	0.5	0.6	0.7

Table 3: Values of  $\sigma_y$  calculated from TWSTFT and GPS time transfer and (TWSTFT-GPS) differences.

#### 4.4) Delay Steps

Hardware changes will induce discrete changes in the instrumentation delay of either the TWSTFT or GPS system. Where possible these steps are calibrated and corrected. However, there are several examples of delay steps that do not correlate with known instrumentation delay changes. Determining the laboratory where the delay step occurred is straightforward, however

determining whether the step occurred in either the GPS common-view or TWSTFT instrumentation is more difficult. Averaging data from the laboratory where the step occurred over several time transfer links may reduce other instabilities and enable the origin of the steps to be determined. An example of uncorrelated delay steps occurred on MJD 49558 and MJD 49576 observed on the PTB (TWSTFT-GPS) differences. These steps were identified as occurring within the PTB GPS receiver.

#### 4.5) Satellite Transponder Delays

The effects on a TWSTFT, of delay changes within the satellite transponders depends upon the link configuration. European TWSTFTs use the same transponder for both transmissions forming the time transfer. This results in zero transponder delay asymmetry. Transatlantic TWSTFT transmissions are made via separate transponders, resulting in non zero delay asymmetry. Any delay change in the transponder will effect the overall time transfer. The replacement of the Intelsat satellite on MJD 49908 may result in a change in transponder delays. To investigate this further the four time transfers (NIST-NPL), (NIST-PTB), (USNO-NPL) and (USNO-PTB) were all averaged over the period of the satellite change. Examining both TWSTFT and (TWSTFT-GPS) data revealed that delay changes of several nanoseconds have occurred in some of differences at the time of the satellite change. These steps were observed on links involving USNO but not NIST suggesting that their might be an alternative cause. The study did however place a limit on the possible change in the transponder delay asymmetry of a few nanoseconds.

Link Combination	SUM $\sigma_y \times 10^{14}$	DIFFERENCE $\sigma_y \times 10^{14}$
NIST-NPL USNO-PTB	1.8	1.9
NIST-PTB USNO-NPL	2.0	1.6
NIST-NPL USNO-TUG	2.3	2.3

Figure 4:  $\sigma_y$  calculated from the sum and differences of transatlantic TWSTFT links.

Delay changes in the satellite transponder may effect the stability of the transatlantic TWSTFT links. TWSTFT measurements taken from pairs of transatlantic time transfers, coupled only through the use of common transponders, were combined to form both a sum and difference. Any transponder-induced delay instability would then correlate in the sum of the time transfers but cancel in the differences. Values of  $\sigma_y$  ( $\tau=2.3$  days) are shown in Table 4. The results demonstrate that over a period of a few days there is no significant delay instability introduced by the satellite transponders.

#### 4.7) TWSTFT Links via Intermediate Stations

At present TWSTFTs are performed between every possible pair of laboratories. A significant advantage may be obtained in terms of both usage of satellite time and a reduced measurement schedule if the number of time transfers were reduced. This would however result in several "two stage" comparisons between clocks at many locations. TWSTFT have been calculated using both single links and intermediate stations. Values of  $\sigma_y$  are given in Table 6. The use of an intermediate station does not significantly degrade the  $\sigma_y$  values.

TWSTFT Time Transfer	$\sigma_y \times 10^{14}$ single link	$\sigma_y \times 10^{14}$ via intermediate station
NIST-NPL	1.0	1.0
NIST-PTB	1.5	2.0
USNO-NPL	0.9	1.1
USNO-PTB	1.6	1.7
USNO-TUG	2.3	2.3
PTB-NPL	1.6	1.6
PTB-TUG	2.9	2.9
NPL-TUG	2.5	2.5

Table 5: Values of  $\sigma_y$  calculated from TWSTFTs using both direct links and intermediate stations.

#### 5) CONCLUSIONS

The TWSTFT and GPS common view methods have been shown to be capable of providing high precision time and frequency transfers. Values of  $\sigma_y$  ( $\tau=4.7$  days) as low as  $4.3 \times 10^{-15}$  and  $5.3 \times 10^{-15}$  have been reported for GPS common view and TWSTFT respectively.  $\sigma_y$  values calculated between laboratories using caesium clocks were found to be dominated by clock noise. (TWSTFT-GPS) differences displayed a standard deviation as low as 1.4 ns when considering 350 days of data. Further work is required to obtain the optimum performance from both systems, and to calibrate the absolute delay of the TWSTFT system.

#### 6) ACKNOWLEDGEMENTS

This work has been supported under the UK National Measurement System Program for Time and Frequency. The work has also been funded by grants from the Austrian Academy of Science and the Jubilee fund of the Austrian National Bank.

**REFERENCES**

- (1) W. Lewandowski, C. Thomas, "GPS Time Transfer," in Proceedings of the IEEE, Special Issue on Time and Frequency, vol 79 no 7, pp. 991-1000, July 1991.
- (2) D. Kirchner, "Two-Way Time Transfer Via Communication Satellites," in Proceedings of the IEEE, Special Issue on Time and Frequency, vol 79, no 7, pp 983-990, July 1991.
- (3) J A Davis , W Lewandowski, J A De Young, D Kirchner, P Hetzel, G De Jong, A Söring, F Baumont, W Klepczynski, T Parker, K A Bartle, H Ressler, R Robnik and L Veenstra. "Preliminary comparison of Two-Way Satellite Time and Frequency Transfer and GPS Common View Time Transfer during the INTELSAT field trials", Proceedings Precise Time and Time Interval (PTTI 95) Applications and Planning meeting, November 1995.
- (4) E A Gerber and A Ballato, "Precision Frequency Control", Volume 2, Oscillators and Standards, pp 202, Academic Press inc 1985.
- (5) J. Vondrak, "A contribution to the problem of smoothing observational data," Bull. Astron. Inst. Czechoslovakia, vol 20, pp. 349-355, 1969.

## GLONASS COMMON-VIEW TIME TRANSFER BETWEEN NORTH AMERICA AND EUROPE AND ITS COMPARISON WITH GPS

W. Lewandowski  
Bureau International des Poids et Mesures  
Pavillon de Breteuil  
92312 Sèvres Cedex, France

J. Danaher  
3S Navigation  
23141 Plaza Pointe Drive  
Laguna Hills, California 92653, USA

W.J. Klepczynski  
United States Naval Observatory  
3450 Massachusetts Avenue NW  
Washington DC 20392-5420, USA

### ABSTRACT

GLONASS common-view time transfer between two North American Laboratories and the BIPM in France is in operation since June 1995. GLONASS time receivers manufactured by the 3S Navigation, operating in these three locations, observe according to the 1st BIPM International GLONASS Common-View Schedule since 4 January 1996. As GPS common-view time transfer between Europe and North America is accomplished with an uncertainty of several nanoseconds, it gives here an excellent reference with which to evaluate the ultimate performance of GLONASS common-view time transfer. This comparison shows a slightly lower performance for GLONASS which is mainly due to very approximate evaluation of ionospheric delay for GLONASS observations and to limited possibilities to obtain accurate GLONASS ground antenna coordinates. In addition, the principles of developing a GLONASS common-view schedule are explained and some basic features of GLONASS time receivers are reported. Data processing and possible standardization of GLONASS international common-view time transfer are also discussed.

### INTRODUCTION

Russian Global Navigation Satellite System, (GLONASS) was inaugurated in 1982. The full constellation of twenty four satellites plus one spare satellite was completed on 18 January 1996. Because the GLONASS signal is free of intentional degradation and available world-wide, it offers to the international time metrology community an

additional tool for high accuracy time transfer. Its combined use with GPS provides robustness, redundancy and reliability. The first comparison of GLONASS and GPS common-view time transfers was performed three years ago between the BIPM and the VNIIFTRI near Moscow. The baseline is about 3000 km and period of comparison extended over one month (Lewandowski et al (1)). Although the results were encouraging, this comparison was limited by the necessity to operate the GLONASS navigation receivers, which were roughly adapted for the needs of time transfer, in a manual mode. In addition, at this epoch the adapted observation mode of GLONASS satellites in common-view necessitated a weekly renewal of the schedule. Since then another mode of GLONASS observation in common-view was devised allowing fully automatic operations. GLONASS common-view time receivers were developed according to these new procedures, and the 1st BIPM International GLONASS common-view schedule was published and implemented on 4 January 1996. Its principles and a comparison with the GPS schedule is provided in this paper. The data analysis presented here covers the first month of implementation of the 1st BIPM GLONASS schedule at two laboratories situated in North America, the USNO and the 3S Navigation, and one situated in Europe, the BIPM. All three laboratories are equipped with R-100/10, the first C/A Code GLONASS receivers for automatic common-view time transfer. Two are equipped with permanently operating C/A Code GPS time receivers (at the BIPM and at the USNO). This allows a comparison of GLONASS and GPS common-view time transfers over a baseline of about 6000 km. Detailed description of the treatment of GLONASS and GPS data is provided. The difficulties linked to the use of two different coordinate reference frames by the two systems are described. The need for GLONASS time transfer standardization is also discussed.

## GPS AND GLONASS COMMON-VIEW SCHEDULES

The GPS constellation is composed of 24 satellites situated 20000 km above Earth on 6 quasi-circular orbits inclined by 55°. GPS orbits are sidereal so the satellites are observed every sidereal day at nearly the same location on the sky and the scheduled common views are repeated every 23 h 56 min. The GPS common view schedule is kept without change for about 6 months, then a new schedule is issued.

The GLONASS constellation also has 24 satellites. The satellites are located on 3 near-circular orbits at a height of 19100 km with periods of about 11 h 15 min and an inclination of 64,8°. Within each plane satellites are located in 8 almanac slots (plane 1: slots 1 to 8, plane 2: slots 9 to 16, plane 3: slots 17 to 24). Due to their height GLONASS satellites do not have, as GPS, sidereal period. That apparently creates a problem of scheduling automated GLONASS common views. But GLONASS orbits are, in certain sense, sidereal. GLONASS satellites repeat the same path across the sky every 8 sidereal days. Most importantly slots within each plane have such rotation that on successive sidereal days successive slots appear at the same location on the sky (Fig. 1). This feature of GLONASS constellation allows fully automated GLONASS common-view tracking schedule resembling the GPS one. All the track times should be decremented by 4 minutes each day. Slot number should be increased by 1 each day, within each of 3 orbital planes. There is only one difference between GLONASS and GPS schedules: with GPS we observe on successive sidereal days in the same direction of the sky the same GPS satellite and with GLONASS a different satellite.

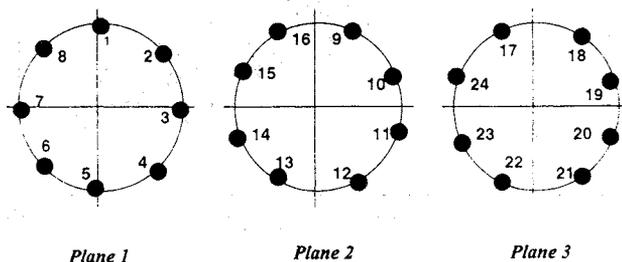


Figure 1. GLONASS slots.

1st BIPM international GLONASS common-view schedule was implemented on 4 January 1996 in several time laboratories around the world.

## DESCRIPTION OF HARDWARE

During this experiment, C/A Code GLONASS receivers, of type R-100/10, for automatic common-view time transfer, were used for the first time. Among the three sites involved in this exercise two are equipped with permanently operating C/A Code GPS time receivers: at the BIPM and at the USNO. At the USNO the GPS time receiver is a STI 502 unit and at the BIPM a SERCEL unit. These are connected to the same clocks as the GLONASS receivers. The clocks are a hydrogen maser at the USNO, a HP5071A caesium at the USNO and a commercial standard caesium at the 3S Navigation. The two GPS receivers were differentially calibrated with uncertainty of 2 ns. The GLONASS receivers were roughly compared side-by-side at 3S Navigation for several days before being shipped to the BIPM and the USNO. Their differential delays are thought to be known to within a few tens of nanoseconds. The calibrated GPS link allows in this study a better assessment of GLONASS time link.

## DESCRIPTION OF GEODETIC PARAMETERS

GPS satellites operate in the WGS 84 reference frame, but the coordinates of the GPS antennas we use, are expressed in the reference frame ITRF88 (close to WGS 84 to within decimetres, but more accurate) with an uncertainty of 10 cm for the USNO and 30 cm for the BIPM (Lewandowski et al (2), Lewandowski (3)). The uncertainties of the GPS ground antenna coordinates can have an impact on the accuracy of the BIPM-USNO GPS common-view link of no more than 1 nanosecond. GLONASS satellites operate in the SGS 90 reference frame which differs by several metres from WGS 84. No accurate relationship between SGS 90 and WGS 84 is yet known (Misra (4)). Therefore, the sole method of determining GLONASS antenna coordinates is to average a series of navigation solutions. The uncertainties of such coordinates are no better than several metres and can have an impact on the accuracy of the common-view link of a few tens of nanoseconds. The coordinates of the GLONASS antennas at the BIPM and the USNO were expressed, for this experiment, in the ITRF88 reference frame with an estimated differential uncertainty of 30 cm. This ensures coherence of the ground-antenna coordinates between the two sites. The disadvantage of this solution is the lack of coherence between

ground-antenna coordinates and satellite ephemerides. For short baselines of several hundreds kilometers this difference has a negligible impact on common-view time transfer. This impact, however, increases with the length of the baseline. The coordinates of the GLONASS antenna at the 3S Navigation were determined from GLONASS navigation solutions with an uncertainty of several meters. The accuracy and precision of GLONASS links with 3S Navigation are affected by this uncertainty.

The uncertainty of GPS broadcast ephemerides range within 10 m to 15 m. They are not affected by Selective Availability (SA). Uncertainty of GLONASS broadcast ephemerides is believed to be slightly higher and should range within 15 m to 20 m. Post-processed precise ephemerides of GPS satellites are publicly available with a delay of several days. Their uncertainty of several decimetres is continuously improving. There is no available information on possible computation of GLONASS post-processed precise ephemerides. However, there are already some geodetic civil institutions considering international efforts to compute publicly available GLONASS precise ephemerides. During this study we did not apply GPS precise ephemerides as their sole use, without applying measurements of ionosphere, would have limited impact on the quality of GPS time link.

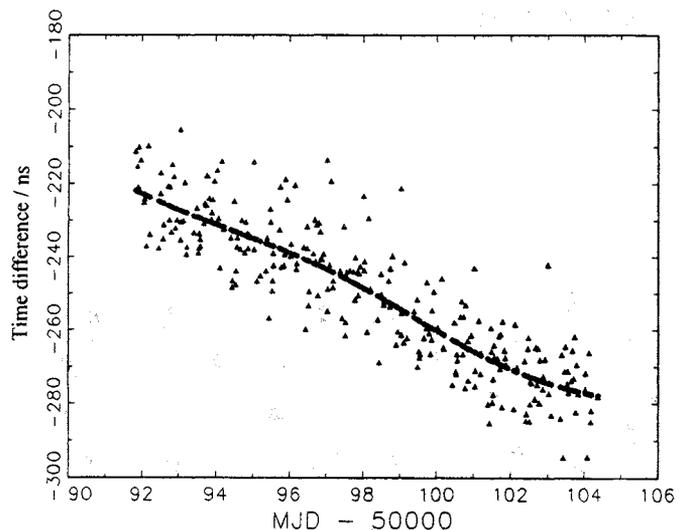
## IONOSPHERIC DELAY

The GPS time receivers correct time observations for ionospheric delay using a model based on broadcast parameters. As GLONASS navigation message does not contain ionospheric parameters, GLONASS time receivers use a model based on fixed parameters. This limits precision of GLONASS links studied in this paper. Especially as these links are long distance where low elevation angles must be considered.

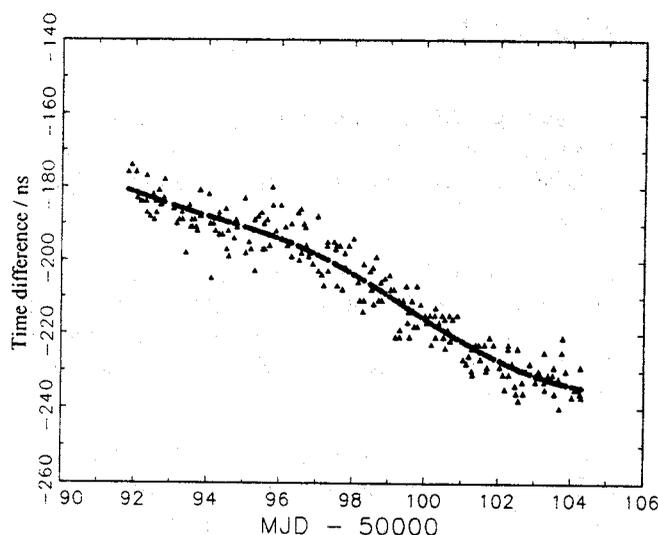
Ionospheric measurements using two GPS frequencies are performed at the BIPM and the USNO. But USNO GPS ionospheric data was not yet operationally available at the time of this study. For this reason BIPM-USNO GPS link was not here corrected for these measurements. There are already commercially available double-frequency GLONASS time receivers measuring ionospheric delay. Their use should greatly improve quality of GLONASS measurements.

## DISCUSSION OF RESULTS

In this paper we consider three time links of the baselines of respectively 3700 km, 6000 km and 8400 km. The length of baseline affects precision and accuracy of satellite common-view time transfer. The longer is the distance, the larger is the effect of uncertainty of satellite ephemeride and ionospheric delay on time transfer. The data analysis covers the first month of implementation of the 1st BIPM international GLONASS schedule which corresponds to the 26th BIPM GPS schedule. Both, GLONASS and GPS common views were computed at the midpoints of the tracks. There were about 40 GLONASS common views available daily for 3S Navigation-USNO link, 17 common views for USNO-BIPM link and 13 common views for 3S Navigation-BIPM link. GPS USNO-BIPM link had about 15 common views available daily. During this exercise GPS satellites were subject to Selective Availability, so strict common views were required (Allan et al (5)). A Vondrak smoothing (Vondrak (6)), which acts as a low-pass filter with a cut-off period of about 5 days, was performed on the raw common-view values of three GLONASS links and one GPS link (Figures 2 and 3). This cut-off period was chosen as representing, approximately, the limit between short time intervals, for which measurement noise is dominant, and longer intervals for which clock noise prevails.



**Figure 2.** BIPM clock - USNO clock by GLONASS common views raw and smoothed.



**Figure 3.** BIPM clock - USNO clock by GPS common views raw and smoothed.

The rms of residuals to the smoothing was computed for all time links. It was adopted to express the uncertainty of these links. Table 1 gives its numerical values.

Lab(i) - Lab(j)	Distance (cord) /km	Uncert. of GLO /ns	Uncert. of GPS /ns
3SNAV-USNO	3700	8	-
USNO - BIPM	6400	9	5
3SNAV-BIPM	8400	11	-

**Table 1.** GLONASS and GPS common -view links and their uncertainties.

The better performance of GPS USNO-BIPM link is easily explained by the use of very accurate ground antenna coordinates expressed in a reference frame coherent with the one used for broadcast ephemerides and better quality of the modeling of ionospheric delay. We can also observe increasing of the uncertainty of GLONASS links with increasing distance between stations. This is associated, as stated above, to the effects of the uncertainties of satellite ephemerides and ionospheric delays. The calibrated GPS common-view time link of good quality is here an excellent reference with which to evaluate the performance of GLONASS common-view time transfer. We have interpolated smoothed GLONASS

and GPS values of BIPM-USNO for the 0 h UTC of each day (Table 2).

MJD	BIPM - USNO			
	by GPS /ns	by GLO /ns	GPS -GLO /ns	GPS -GLO -59.3 /ns
50092.0	-166.3	-222.7	61.4	2.1
50093.0	-169.6	-227.1	57.5	-1.8
50094.0	-172.7	-231.1	58.4	-0.9
50095.0	-175.5	-234.8	59.3	0.0
50096.0	-178.9	-238.8	59.9	0.6
50097.0	-183.1	-243.2	60.1	0.8
50098.0	-188.3	-248.3	60.0	0.7
50099.0	-194.3	-254.1	59.8	0.5
50100.0	-200.6	-260.2	60.2	0.9
50101.0	-206.5	-265.6	59.1	-0.2
50102.0	-211.8	-270.6	58.8	-0.5
50103.0	-215.9	-274.2	58.3	-1.0
50104.0	-218.7	-276.9	58.2	-1.1

**Table 2.** BIPM clock - USNO clock smoothed and interpolated values for 0 h UTC.

There is a constant bias of -59,3 ns between two methods which is due partly to the lack of very accurate calibration of GLONASS receivers and partly to less accurate geodetic parameters and modeling of ionospheric delay for GLONASS. We report also residual to the mean with rms = 1.1 ns. These are strikingly low, most of them below 1 ns, with three departures above 1 ns at the edges, due probably to the effects of smoothing.

### NEED FOR GLONASS STANDARDIZATION

GLONASS common-view tracks are now observed and reported according to the standards developed for GPS common-view time transfer drawn up by the CCDS Group on GPS Time Transfer Standards (Allan and Thomas (7)). This satisfies most immediate needs, but must be considered as a provisional solution. There are some particular issues which should be defined specifically for GLONASS and for GLONASS and GPS used together in double system time receivers which will be soon commercially available. Use of different coordinate reference frames, modeling of ionosphere and troposphere can be cited among them. It is suggested that the upcoming 13th CCDS meeting considers transformation of the CCDS Group on GPS Time

## Transfer Standards into the CCDS Group on GPS and GLONASS Time Transfer Standards.

### CONCLUSION

- 1) A successful automated international GLONASS common-view time transfer according to the 1st BIPM GLONASS schedule has been proven by this exercise.
- 2) Performances of GLONASS time links are only slightly lower than those of GPS, which is easily explained.
- 3) More accurate determinations of GLONASS antenna coordinates in the SGS 90 reference frame and use of double-frequency ionospheric delay measurements would significantly improve GLONASS common-view time transfer.
- 4) Combined use of GLONASS and GPS provides robustness, redundancy and reliability.
- 5) Need for more rigorous standards for GLONASS common-view time transfer is stated.

### ACKNOWLEDGMENTS

The experimental part of this work was done at the BIPM thanks to the loan from Hewlett Packard of a 5071A caesium clock and from 3S Navigation of a GLONASS R100/10 time receiver. The staff of the Time Section of the BIPM is grateful to Hewlett Packard and 3S Navigation for their generosity.

### REFERENCES

1. Lewandowski W., Moussay P., Cherenkov G.T., Koshelyaevsky N.B., Pushkin S.B., 1993, "GLONASS Common-View Time Transfer", Proc. 7th EFTF, 147-152.
2. 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 Symp. on Frequency Control, 218-224.
3. Lewandowski W., 1992, "A Worldwide Unification of GPS Antenna Coordinates for High Accuracy Time Transfer", Proc. Journées spatiaux temporaires, Paris Observatory, 142-147.
4. Misra P.N. and Abbot R.I., 1994, "SGS 85 - WGS 84 transformation", Manuscripta Geodetica, 19, 300-308.
5. Allan D.W., Granveaud M., Klepczynski W.J. and Lewandowski W., 1990, "GPS time transfer with implementation of Selective Availability", Proc. 22nd PTTI meeting, 145-156.
6. Vondrak J., 1969, "A Contribution to the Problem of Smoothing Observational Data", Bull. Astron. Inst. Czechoslovakia, 20, 349-355.
7. Allan D.W. and Thomas C., 1994, "Technical Directives for Standardization of GPS Time Receiver Software", Metrologia, 31, 67-79.

## "COMMON VIEW" CLOCK SYNCHRONIZATION OF REMOTE ATOMIC CLOCKS USING GPS AND PRARE ONBOARD ERS-2

J Hahn<sup>1</sup>, S Bedrich<sup>2</sup>

<sup>1</sup> DLR, Institut für Hochfrequenztechnik, Postfach 1116, D-82230 Oberpfaffenhofen, joerg.hahn@dlr.de

<sup>2</sup> GeoForschungsZentrum Potsdam, Abt. 1, Telegrafenberg A17, D-14473 Potsdam, c/o DLR / D-PAF, Postfach 1116, D-82230 Oberpfaffenhofen, bedrich@dfd.dlr.de

### SUMMARY

The paper presents a validation test for two-way clock comparison with the help of the PRARE instrument onboard ERS-2 and with two of the PRARE system stations, which are provided with atomic clocks. A "Common View" ground-clock synchronization scenario has been performed, using the PRARE instrument and GPS time receivers connected to an atomic clock at each ground station site. By computing and comparing the results respectively the accuracy for the PRARE instrument has been estimated. The achievable synchronization accuracy for the PRARE equipment in this "Common View" mode is around 1 ns, due to the sub-optimum design of the existing PRARE system for time purposes. The experiment turns out that the PRARE system combined with a highly stable frequency source onboard a suitable satellite is qualified to be a basic element for a future GNSS II.

### 1. INTRODUCTION

When investigating navigation systems of the next generation new methods for synchronization of the ground and satellite clocks (S/C's and G/C's) have to be discussed. Besides the conservative one-way method with its known drawbacks in combining clock and ephemeris errors, two-way methods seem to be important concerning better synchronization accuracy. In Bedrich (1) a precise two-way microwave time transfer link has been described. It has been shown that the existing PRARE system can be modified for such purposes. The PRARETIME instrument has been introduced. A detailed study can be found in Hahn (2).

To test the described in these papers two-way clock comparison method, an experiment has been planned between DLR and GFZ in 1995 performing a "Common View" (CV) G/C synchronization scenario by means of PRARE onboard ERS-2 and GPS time receivers. For DLR it was the first experiment with PRARE equipment data. The main goal of this experiment was clock comparison via two-way microwave time links. Existing systems like MITREX use geostationary satellites (GEO) for this purpose. But a GEO only covers a certain area of the Earth. For global time compari-

son operation or time dissemination a polar orbiting reference has to be used. Two-way clock synchronization using an orbiting time reference has been studied.

### 2. MEASUREMENT PRINCIPLE

During the measurement campaign two different clock comparison techniques have been used simultaneously: clock comparison via

- GPS CV and
- PRARE instrument onboard ERS-2.

These methods are outlined below.

#### 2.1 GPS-"Common View"

The GPS CV is a well proved time comparison method of remote clocks using the BIPM GPS tracking schedule (presently schedule no 26). Following this schedule, all participating time receivers monitor a dedicated GPS satellite clock simultaneously.

Each GPS time receiver with its reference clock  $REF_i$  records the time difference  $(REF_i - SV_j)_m$  averaged over a certain interval (780 sec) at time  $t_m$ , where  $j$  is the pseudo range number of GPS satellite  $SV$ . Data correction accounting for ionospheric, tropospheric, relativistic and range delays are implemented in the receiver with different effort.

By exchange and subtracting the recorded files the clocks  $REF_i$  and  $REF_k$  can be synchronized respectively and the clock offset  $D_{ik}^{GPS-CV}$  at time  $t_m$  is

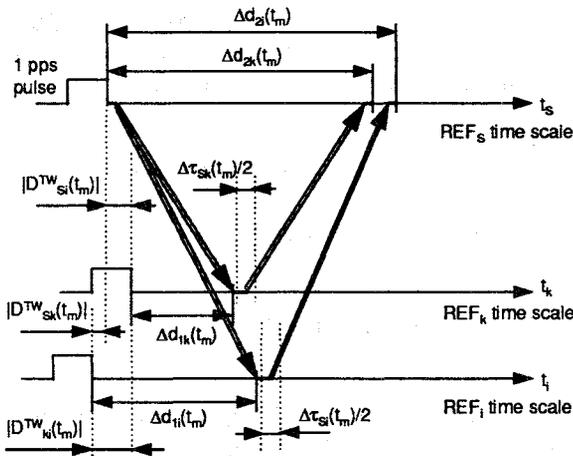
$$\begin{aligned} D_{ik}^{GPS-CV}(t_m) &= (REF_i - SV_j)_m - \\ &\quad - (REF_k - SV_j)_m \\ &= (REF_i - REF_k)_m. \end{aligned} \quad (1)$$

Repeating this procedure for any available time  $t_m, t_{m+1}, t_{m+2}, \dots$  the rate  $R$  between the clocks can be computed over a certain time interval. The accu-

racy for this comparison method is around 3.6 ns following Lewandowski (3).

## 2.2 PRARE Clock Comparison

The PRARE instrument operates in two-way mode (signal flow: space-ground-space). This allows to measure the clock offset  $D_{Sk}^{TW}(t_m)$  of the involved S/C  $REF_S$  and G/C  $REF_k$ , if the asymmetry  $\Delta\tau_{sk}(t_m)$  of the two paths is determined. This is fulfilled by measuring the signal travel time at the S/C, round-trip travel time  $\Delta d_{2k}(t_m)$ , and additionally (simultaneously) at the G/C site, one-way travel time  $\Delta d_{1k}(t_m)$  (cf. fig. 1).



**Fig. 1** Principle of two-way clock comparison (CV mode)

The clock offset can be determined with

$$D_{Sk}^{TW}(t_m) = \Delta d_{1k}(t_m) - \frac{\Delta d_{2k}(t_m) + \Delta\tau_{sk}(t_m)}{2}. \quad (2)$$

The microwave signal transmitting the frequency and time information is subject to diverse disturbances on its way from clock to clock causing asymmetries ( $\Delta\tau_{sk}$ ) between the signal paths.

One can distinguish between internal, i.e. hardware and measurement errors on both sides, and methodic, i.e. signal transmission and correlation errors. The residual accuracies of models or supplementary measurements which correct those deviations define the achievable clock synchronization accuracy.

While the microwave signal travels the ionosphere with frequency  $f_X$ , a coherent signal on a second frequency  $f_S$  with correlated reception gives the

possibility to correct very precisely for ionospheric delays (different down- and uplink frequencies in X-band). If  $\Delta d_{1k}^X(t_m)$  and  $\Delta d_{1k}^S(t_m)$  indicate the one-way travel times in X- and S-band ( $f_X$  and  $f_S$ ) respectively, the ionospheric delay time  $\Delta\tau_{IX}(t_m)$  for the X-band signal will be equal to

$$\tau_{IX}(t_m) = \frac{\Delta d_{1k}^X(t_m) - \Delta d_{1k}^S(t_m)}{f_X \left( \frac{1}{f_X^2} - \frac{1}{f_S^2} \right)}. \quad (3)$$

Tropospheric corrections have to be included by meteorological measurements and appropriate modelling; all other influences depend on the calibration efforts.

The simultaneous PRARE Doppler measurements  $f_{D2}$  (two-way in PRARE) have to be included to solve for the range asymmetry  $\Delta\tau_d$

$$\Delta\tau_d = \frac{\Delta d_{2k}}{2} \cdot \frac{f_{D2}}{f_{Down}}, \quad (4)$$

with  $f_{Down}$  being the downlink frequency.

Relativistic effects have to be considered for utmost accuracy requirements.

In the neighborhood of culmination point at time  $t_{cul}$  the asymmetry  $\Delta\tau_{sk}(t_m)$  reaches its minimum due to almost reciprocal signal paths. This time moment has to be preferred for data processing.

Exchange of the measurement results between the clocks has to be done by data transfer (for example in the message frame of the PRARE microwave links). The main advantage of two-way measurements is the fact that slant range  $r$  and clock offset  $D_{Sk}^{TW}(t_m)$  are achievable one without the other, i.e. the clock offset  $D_{Sk}^{TW}(t_m)$  is obtained without knowledge of the clocks' distance, and the slant range  $r$  is obtainable without information about the clocks' deviation. On the other hand for picosecond level of accuracy relativistic effects have to be calculated using precise orbit information.

Synchronization of remote G/C's mutually is possible by CV observation of the S/C. Fig. 1 gives a scheme for comparison of clocks  $REF_i$  and  $REF_k$ .

When subtracting the computed clock offsets  $D_{Sk}^{TW}(t_m)$  and  $D_{Si}^{TW}(t_m)$ , the clock offset  $D_{ki}^{TW}(t_m)$  between the ground clocks can be calculated.

### 3. EXPERIMENTAL SETUP

Fig. 2 presents an overview of the experimental setup with ground station sites in DLR's branch in Lichtenau near Weilheim and GFZ's branch in Oberpfaffenhofen near Munich. Each station was equipped with an appropriate PRARE station and GPS time receiver and are described in section 4.

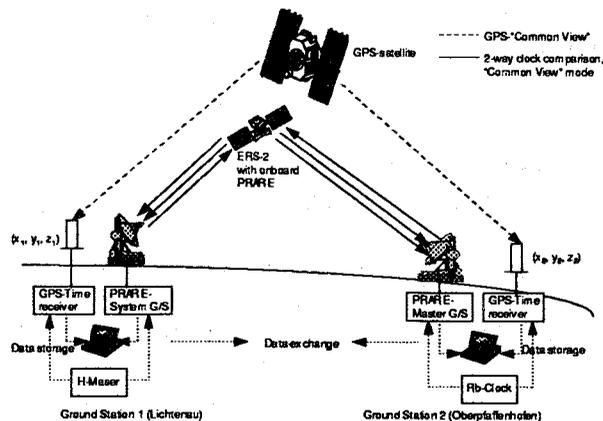


Fig. 2 Scheme of experimental setup

### 4. HARDWARE CALIBRATION

The always critical matter is careful calibration of instrumental hardware. Only by this way utmost accuracy can be reached.

The PRARE instrument performs the calibration by itself (calibration loops, etc.). Corresponding figures are available in the recorded data files. The additional one-way measurement must be calibrated by the experimenter. This concerns essentially cable delays between clock, PRARE processor and time interval counter (TIC), and electronic switches, plugs, isolation amps.

In the GPS time receiver, common receiver, antenna, and clock cable delays must be accounted for. These delays are considered in the data processing. During a 7 weeks measurement campaign in 1990 a comparison between the two involved GPS time receivers had been performed. Here both receivers had been operated at the same location with the same clock. A bias could be determined with an order of about 200 ns.

To overcome this problem in the present experiment, a new comparison was necessary and initialized. A third GPS time receiver AO TTR5 was used, operating at each ground station site with the same ground clock respectively for a duration of one week. The bias between the ground station receiver and this third receiver has been evalu-

ated. Finally the bias between the two ground receivers presently valid could be estimated with presently about 40 ns and considered in the data computation.

#### 4.1 Ground station site in Lichtenau (station 1)

The principal structure of the experimental setup in DLR's branch in Lichtenau can be obtained from fig. 3.

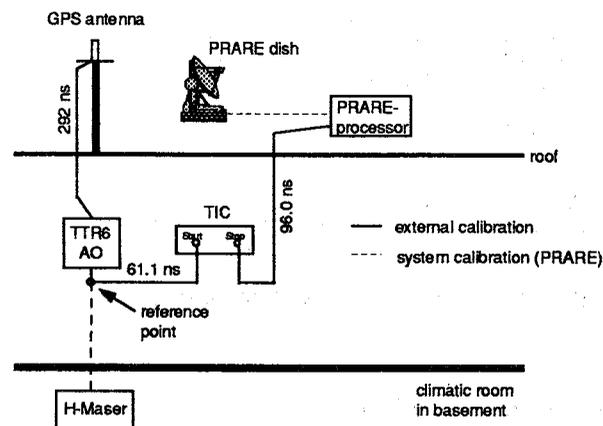


Fig. 3 Setup at DLR's ground station site

The reference clock, a H-maser CH1-75 of Russian company "KVARZ" is located in a climatic room in the basement. The GPS time receiver AO TTR6 operates in the time lab where the TIC for the one-way measurement is placed too. The PRARE processor with antenna dish at roof is connected via cable to the TIC in the time lab. The cable delays have been determined as shown in figure 4 and will be denoted as  $\Delta\tau_{start1} = 61.1$  ns,  $\Delta\tau_{stop1} = 96$  ns. The cable delay of 292 ns to the GPS antenna is considered in the TTR6 data processing.

#### 4.2 Ground station site in Oberpfaffenhofen (station 2)

The setup in GFZ's branch in Oberpfaffenhofen which corresponds particularly to the PRARE Master Station configuration can be seen in fig. 4. The reference clock is represented by a Rb-clock in phase locked redundancy by EFRATOM which is DCF-77 disciplined to follow the UTC(PTB) time scale, cf. Bedrich (4). Also a GPS time receiver AO TTR6 with a TIC for registration of one-way measurements is located in the operation room. The PRARE processor with antenna dish at roof is connected via cable to the TIC in the operation room. The cable delays have been determined

and can be obtained from fig. 4 respectively and will be denoted as  $\Delta\tau_{start2} = 5 \text{ ns}$  and  $\Delta\tau_{stop2} = 209.5 \text{ ns}$ . The cable delays of 298 ns and 5 ns to GPS antenna and clock reference point are considered in the TTR6 data processing.

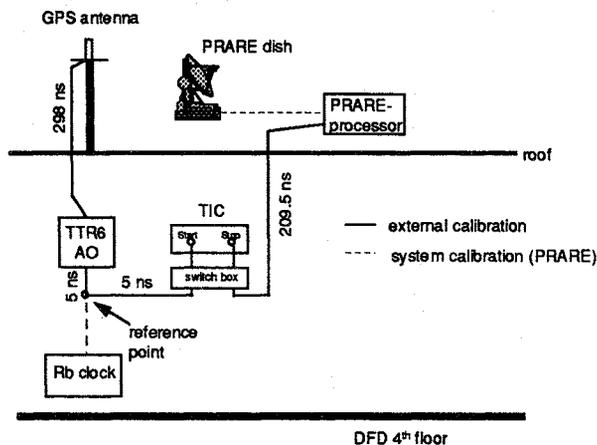


Fig. 4 Setup at GFZ's ground station site

## 5. DATA EVALUATION

All data have been recorded between December 1995 and February 1996. Below an example of the data analysis of GPS and PRARE observables is given.

### 5.1 GPS-CV

Fig. 5 presents the result of GPS clock comparison between the two clocks involved in Lichtenau (H-maser) and Oberpfaffenhofen (Rb) respectively over the whole observation duration.

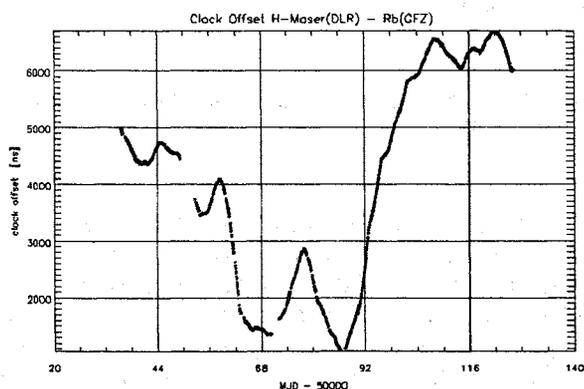


Fig. 5 Result of GPS CV between DLR's H-maser and GFZ's DCF77 disciplined Rb-clock

It has to be mentioned again that GFZ's rubidium clock is DCF77-disciplined to follow the UTC(PTB)

time scale. That can be seen in fig. 5 - each third day a correction is applied, the rate is due to this Rb-clock. The observed offset of 40 ns between the two GPS time receivers has been considered.

### 5.2 PRARE data

For clock comparison, the following data were available from the PRARE instrument data processing (the station index will not be added for simplification),

in either range and Doppler measurement files:

- date and time of event,  $t_m$ ;
- azimuth  $az$  and elevation  $el$  at ground station, [°];
- air pressure  $pr$ , [hPa], air temperature  $T$ , [K] and air humidity  $hu$ , [%] at ground station;
- tropospheric refraction correction (2-way)  $\Delta\tau_{T2}, \Delta\tau_{TC2}$ , [ps] or [cycles];
- ionospheric refraction correction (2-way)  $\Delta\tau_{I2}, \Delta\tau_{IC2}$ , [ps] or [cycles];
- satellite centre of mass correction (2-way)  $\Delta\tau_{M2}, \Delta\tau_{MC2}$ , [ps] or [cycles];
- onboard PRARE antenna phase centre correction (2-way)  $\Delta\tau_{P2}, \Delta\tau_{PC2}$ , [ps] or [cycles];
- ground station mechanical centre correction (2-way)  $\Delta\tau_{C2}, \Delta\tau_{CC2}$ , [ps] or [cycles];
- external calibration correction (2-way)  $\Delta\tau_{E2}, \Delta\tau_{EC2}$ , [ps] or [cycles];
- X-band versus S-band travel delay measurement (1-way, space to ground)  $\Delta\tau_{XS}$ , [ps];

in only range measurement files:

- two-way travel time (each second) including internal calibration of the space segment and ground station and 91 value range correction  $\Delta d_2$ , [ps];

and in only Doppler measurement files:

- measured Doppler cycles (2-way) including internal calibration of the space segment and ground station  $\Delta d_{C2}$ , [cycles];
- frequency offset relative to nominal satellite frequency 8.489 GHz  $\Delta\tau_{OF}$ .

One-way travel time measurements  $\Delta d_1$  were recorded at each ground station site with date and time of event  $t_m$  and figures in [ps].

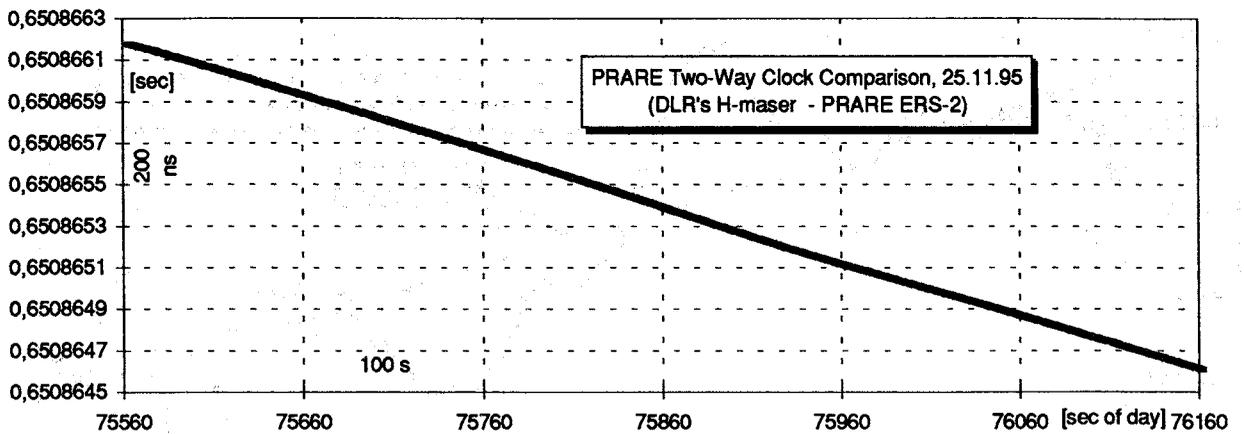


Fig. 6 Example of two-way clock comparison between DLR's H-maser and PRARE onboard clock

Generally the time offset has been determined following equation (2). The data designated at

$$t_m, \Delta\tau_{I2}, \Delta\tau_{IC2}, \Delta\tau_{P2}, \Delta\tau_{PC2}, \Delta d_1, \Delta\tau_{E2}, \Delta\tau_{EC2}, \Delta\tau_{XS}, \Delta d_2, \Delta d_{C2}, \text{ and } \Delta\tau_{OF}$$

have been applied in the present work to compute the time offset and account for asymmetries on the propagation path.

The two-way range measurement  $\Delta d_2$  has been corrected to  $\Delta d_{2c}$  following instructions from the PRARE processing team to be

$$\Delta d_{2c} = \Delta d_2 - \Delta\tau_{I2} + \Delta\tau_{P2} + \Delta\tau_{E2} + \Delta\tau_d, \quad (5)$$

where  $\Delta\tau_d$  being the correction for range asymmetries due to different up- and downlink paths and calculated as

$$\Delta\tau'_d = \Delta d_{C2} + \Delta\tau_{IC2} + \Delta\tau_{PC2} + \Delta\tau_{EC2} + \Delta\tau_{OF}, \quad (6)$$

$$\Delta\tau_d: \Delta\tau'_d \longrightarrow \text{eq. (4).}$$

The one-way signal travel time  $\Delta d_1$  was corrected to  $\Delta d_{1c}$  using

$$\Delta d_{1c} = \Delta d_1 - \Delta\tau_{IX} + \Delta\tau_{start} - \Delta\tau_{stop}, \quad (7)$$

$$\Delta\tau_{IX} [ps]: \Delta\tau_{XS} \longrightarrow \text{eq. (3).}$$

Finally the time offset between satellite  $S$  and ground station  $k$  was evaluated with

$$D_{Sk}^{TW}(t_m) = \Delta d_{1ck} - \frac{\Delta d_{2ck}}{2}, \quad (8)$$

(here the station index is added).

An example of two-way clock comparison between the H-maser reference in Lichtenau and PRARE onboard clock (ERS-2) during a 600 sec pass on 25 November 1995 presents fig. 6. Here the rate of the PRARE oscillator can be identified; it includes relativistic effects which have not been accounted for yet.

## 6. DISCUSSION OF RESULTS

In the following, the results of the differential data processing of the PRARE pass from 25 November 1995 (cf. fig. 7) are presented.

The pass had a duration of about 650 sec. Data points calculated are smoothed over 30 sec. The standard deviation from the linear regression line is given by a figure of 770 ps ( $1\sigma$ ). To compare the processed time offset with the GPS CV results, the midpoint of the regression line has been determined to be

$$D_{12}^{TW}(t_m = 75878 \text{ sec of 25 Nov 1995}) = 4485.66 \text{ ns.}$$

For the time corresponding to the above evaluated midpoint GPS CV clock data have been processed in the following manner: For the 25 November 1995 a linear regression line has been computed. The clock offset corresponding to the midpoint of PRARE pass has been evaluated by interpolation to be

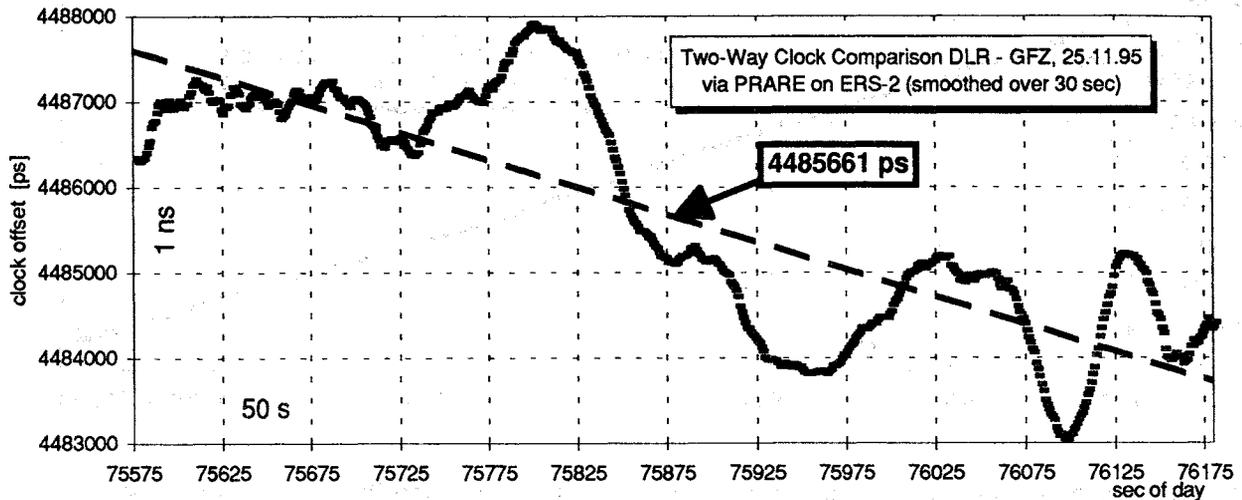


Fig.7 Result of PRARE clock comparison, DLR's H-maser - GFZ's Rb, 25.11.95

$$D_{12}^{GPS-CV}(t_m = 75878 \text{ sec of 25 Nov 1995}) = 4550.63 \text{ ns.}$$

Here the standard deviation from the linear regression line is given by a figure of 8.91 ns ( $1\sigma$ ).

With a great number of other PRARE CV data processings the fact could be verified that the precision of PRARE two-way clock comparison in present design is always around 1 ns.

Concerning accuracy, a bias discrepancy of 65 ns to GPS CV computation was computed. The reasons are:

- uncertainty of time offset between the GPS time receivers;
- the calibration measurement was not sufficient for accurate GPS calibration;
- missing one-way travel time calibration of the PRARE ground stations.

## 7. CONCLUSION

Differential two-way clock comparison (ground to ground) by means of PRARE onboard ERS-2 is presently possible with a precision of around 1 ns. A similar figure is expected for space to ground clock comparison causing from the two-way principle. The achievable accuracy which is supposed to be of the same order must be further investigated. Due to time constraints this matter could not be discussed finally (calibration problems with GPS).

The experiment emphasizes the possibility of highly accurate global clock comparison and time transfer if a two-way time and ranging system of

PRARE like type is carried onboard a polar orbiting satellite. The results point out that in conjunction with an ultrastable onboard reference this configuration could be applied for clock synchronization within future satellite navigation systems (GNSS I and II) and for benefit of the whole time community.

## ACKNOWLEDGEMENT

The authors thank Mr H Nau for many helpful discussions on the performance and evaluation of the experiment. We also would like to acknowledge the efforts of Mr R Manetsberger, Mr Ch Lehner and Dr J Furthner for their support in installation of the PRARE station in Lichtenau and data taking, and GFZ's director Prof Dr Ch Reigber for approval of experiment running.

## REFERENCES

1. Bedrich S, Hahn J, 1995, Precise Two-Way Microwave Time Transfer to Space-Borne Atomic Clocks, *Proc. 9<sup>th</sup> European Frequency and Time Forum*, Besançon, 485-491
2. Hahn J, Bedrich S, Nau H, 1995, Study on "H-Maser in Space" as Ultra-Stable Clock for Time Transfer and Time Dissemination, Final Report, DLR, Oberpfaffenhofen
3. Lewandowski W, Thomas C, 1991, GPS Time Transfer, *Proc. IEEE 79 (7)*, „Special Issue on Time and Frequency“, 991-1000
4. Bedrich S, Flechtner F, 1996, PRARE One resp. Two Years in Orbit - Performance Analysis of the Instruments' Oscillators onboard ERS-2 and Meteor-3/7, to be published at *Proc. 10<sup>th</sup> European Frequency and Time Forum*, Brighton

## IMPROVING THE NIST IONOSPHERIC MEASUREMENT SYSTEM<sup>1</sup>

Marc A. Weiss, Steven Jefferts, Laurent Gaudron<sup>2</sup>

NIST Time and Frequency Division  
325 Broadway, Boulder, CO 80302, U.S.A.

### ABSTRACT

Two problems of the NIST Ionospheric Measurement System have been addressed. A new software lock on the received satellite frequency allows the system to lock more robustly in the presence of selective availability. We obtain about 30% more measurements. Also, biases in the measurements largely come from the front end antenna system. Preliminary results indicate that for the most part the problem is neither due to multi-path interference nor phase center offsets in the antennas, as was previously thought. There is indication that some of the effect is due to an interaction between the two quadrafler helix antennas.

### INTRODUCTION

Measurements of the delay of GPS timing signals through the ionosphere are important for common-view time transfer. The longer the baseline, the more important that real measurements be used in place of an ionospheric model. The NIST Ionospheric Measurement System (NIMS) is used routinely for international time transfer for the generation of international atomic time (TAI) [1]. Measurements from various NIMS's and other ionospheric measurement systems have been shown to significantly improve international time transfer, especially during periods of maximum solar sunspots [2].

The NIMS measures the relative phase of the pseudo-random code called the P-code as received between the two frequencies L1, 1.6 GHz, and L2, 1.2 GHz, transmitted by the Global Positioning System (GPS) satellites [3]. The P-code is transmitted coherently on the two frequencies.[4] Since the group delay of the code is inversely proportional to the square of the carrier frequency, the differential arrival time is a measure of the ionospheric delay on the signals. The NIMS measures the differential arrival time of the P-code between the L1 and L2 signals using a codeless technique [1,5,6]. The system does not use the actual P-code, rather it determines the relative phase of the code using a delay-and-multiply technique. Since we do not track the pseudo-random codes per se, we must find some way to differentiate among the GPS satellite signals received, since they are all nominally on the same frequency. The differences of the received frequencies are due to Doppler shifts of the received signals resulting from the motion of the satellites. We use the frequency offsets of the satellites and the rates of

change of these offsets due to satellite motion to discriminate among satellites. The NIMS software tracks each satellite by using individual frequency-locked loops to each satellite.

The GPS signals are deliberately corrupted using a process called selective availability (SA) to deny the full accuracy of the system from users who are not authorized by the U.S. military. SA causes the received frequencies to fluctuate so rapidly that the original NIMS frequency-locked loop was unable to maintain lock consistently. We have redesigned and implemented a new loop, increasing the bandwidth. In the redesign we now implement a frequency lock on both the L1 and L2 signals instead of only the L1 signal. Since we receive the signals sequentially, dwelling 7.5 s on each, we now have lock information every 7.5 s instead of every 15 s. Unfortunately, this does not double the bandwidth. Because the frequency average is a difference of the measured phase now minus the measured 15 s ago, the frequency measurement on L1 at time T is correlated with the correction applied based on the measurement on L2 at time T-7.5s. Yet the increase in bandwidth is significant enough to provide consistent locking on GPS satellites with approximately a 30% increase in available data.

Another problem with the NIMS has been biases in measurements of the order of  $\pm 6$  ns. Measurements of the delay from the same satellite at the same time with two different receivers show such offsets, which repeat each day. Research suggests that averaging these biases over all satellites tracked in one hour by correcting the received delay for the vertical ionosphere shows an agreement with Faraday rotation measurements of the

<sup>1</sup>Contribution of the U.S. government not subject to copyright.

<sup>2</sup>Guest researcher from Technical Institute, Strasbourg University, Strasbourg, France

ionosphere at about the 3 ns level [7,8].

Our studies suggest that the biases come from the front end antenna system. While at first glance the problems seem to relate to the geometric relationship of the receiver to the satellites and multi-path interference, our study indicates that for the most part this is not so. Therefore, the problem is probably neither due to multi-path interference nor phase center offsets in the antennas, as was previously thought. There is indication that some of the effect is due to an interaction between the two quadrafler helix antennas.

#### THE NEW TRACKING LOOP

A complete description of the NIMS has been published elsewhere. We describe here three functional portions of the design relevant to our discussion. The front end mixes a signal from either L1 or L2 with a frequency midway between them. The system sequentially locks on L1 for 7.5 s and L2 for 7.5 s. This design allows both frequency channels to follow the same paths after the first mixer. This is an important feature for stability and accuracy of the NIMS.

The codeless phase measurement is based on a delay-and-multiply technique. Both the L1 and L2 signals carry a pseudo-random code called the P-code, with a chip rate of 10.23 MHz. The NIMS recovers this 10.23 MHz clock by delaying the signal by half of one chip, about 50 ns, multiplying the delayed signal by the direct one, then band-limiting the result. We now have a signal consisting of a sum of sine waves coherent with the 10.23 MHz clocks from the satellites. This signal is then mixed down so the central frequency is approximately 78.74 Hz. The Doppler offsets vary about  $\pm 25$  Hz. This signal is sampled at 250 Hz by a 8 bit A/D converter and passed into a microprocessor.

The microprocessor tracks each satellite in the 250 Hz bit stream using a frequency lock loop as well as deriving the ionospheric delay for each satellite on the L1 signal. We process the 250 Hz data for 7.5 s, obtain a phase of the P-code on the received L-band, then switch to the other L-band for the next 7.5 s. The 250 Hz data are processed as follows. At each 250th of a second, we compute, for a given satellite, a sine and cosine value based on an estimate of the satellite's P-code received frequency and rate-of-change of the frequency. We sum for the entire 7.5 s the product of the sine value and the received data, as well as summing the product of the cosine value and the data. The arctangent of the ratio of the sine product sum divided by the cosine sum gives us a phase of the P-code for the 7.5 s interval for the received L-band.

Thus for each satellite, we have a sequence of P-code phases, one every 7.5 s, alternating between L1 and L2. Let us label phases as  $\Phi^1_1, \Phi^2_1, \Phi^1_2, \Phi^2_2, \Phi^1_3, \Phi^2_3, \dots$ , where the superscript refers to the L band frequency,

and the subscript is a sequential count of the phases. The difference between neighboring phases, such as  $\Phi^2_1 - \Phi^1_1, \Phi^2_1 - \Phi^1_2$ , give us our measurements of the ionospheric delay. The difference  $\Phi^1_2 - \Phi^1_1$ , being the change in the relative L1 phase over 15 s, is a measure of the offset in our estimate of received frequency for this satellite. The change in the relative L2 phase over 15 s,  $\Phi^1_2 - \Phi^1_1$ , also measures our estimate of received frequency for this satellite.

We report here that we now use both the change in the relative L1 phase over 15 s,  $\Phi^1_2 - \Phi^1_1$ , and the relative L2 phase over 15 s,  $\Phi^2_2 - \Phi^2_1$  in the frequency-locked loop. Previously we used only the change in L1 phase over 15 s to close the frequency lock loop. This was done for two reasons. First of all, it was enough to allow us to lock consistently on satellites before the advent of SA. Secondly, the L2 change in phase is coupled with the L1 change in phase, so that using both measures does not double the information over using one of them. In addition, the signal power of L2 is specified to be 6 dB lower than L1. The L2 and L1 phase changes are coupled because we alternate measurements of phase on L1 and L2. Hence between consecutive L1 phase measurements there is an L2 phase measurement, and vice versa. If we adjust our received frequency estimate after an L1 phase measurement using  $\Phi^1_2 - \Phi^1_1$ , then the phase change  $\Phi^2_2 - \Phi^2_1$ , will be corrupted by that adjustment, since the steering will have occurred midway through the  $\Phi^2_2 - \Phi^2_1$  measurement.

We implemented a tracking loop which adjusts the L2 phase difference measurement by subtracting 1/2 of the previous phase correction which had been applied midway through the L2 measurement. So if A1 was the previous phase correction applied after the L1 measurement, we now use

$$A2 = \Phi^2_2 - \Phi^2_1 - \frac{1}{2}A1$$

as the phase applied to close the frequency lock loop after this L2 measurement. Similarly, after the next L1 7.5 s measurement, we now use

$$A3 = \Phi^1_3 - \Phi^1_2 - \frac{1}{2}A2.$$

The factor of 1/2 comes since the rate adjustment from the previous cycle occurred half-way through the current one.

The result of the new lock loop is that the receiver takes about 30% more data. As a result, the effect of SA seems no longer to interfere with the operation of the receiver. Since measurements of the ionospheric delay must be made nearly simultaneously with GPS common-view measurements in order to correct the common-view measurements, this will allow for less noise in international time transfer.

#### MEASUREMENT BIASES

If we measure ionospheric delays using two NIMS's and

difference these values, biases appear. For a given satellite we found a non-constant bias pattern that repeats with each pass, once per sidereal day. Examples of this are shown in figures 1a and 1b. Each point is the difference of midpoints to 15 minute linear fits to the 7.5 s ionospheric measurements. Note that the pattern is different for different satellites, even if they are tracked at common times.

We wish to point out that even with these biases, agreement with measures of the vertical delay through the ionosphere using Faraday rotation is approximately at the 3 ns level [4]. This level of agreement requires averaging all NIMS data taken over one hour after correcting them for the vertical delay.

Because we are using a codeless technique, we expected that multi-path interference would be a significant cause of biases. We decided to attempt to reduce these biases by physically rotating the antenna. We hoped rotation would reduce multipath corruption since measurements of the antenna pattern of one such antenna showed a field reversal with a rotation of 180 degrees of azimuth.

Figure 2 shows how the two NIMS antennas are offset from each other and surrounded by a choke ring ground plane. For simplicity we decided to rotate the entire unit and look for a reduction in the changing biases. Though this rotation was less than optimum since it meant that the antennas moved as well as rotated, we felt this should serve to demonstrate the possibility. Figure 3 illustrates the system built for rotating the entire front end antenna system. We built two rotators, one for NIMS#106 and one for #110. They both rotated coherent with the 7.5 s measurement sequence of their respective NIMS processor. The antenna with NIMS#106 rotated in 7.5 s in one direction, then reversed for the next 7.5 s. The antenna system with NIMS#110 rotated in 22.5 s in one direction, then reversed direction for the next 22.5 s. Thus in one direction we measure sequentially L1, L2, then L1 each for successive 120° intervals, then reverse and measure L2, L1, then L2 in the opposite 120° intervals. For the 15 minute linear fits, these should average appropriately.

The results of rotating the antenna are null. We see no significant change in the pattern of biases. Compare the data in figures 4a and 4b, the biases while rotating, with the data in figures 1a and 1b. The two satellites chosen for these figures display biases whose day-to-day variances are typical. The patterns themselves and the peak-to-peak changes in the biases vary among satellites. The fact that rotating the antenna produced a null result suggests that the biases may not be a function of the geometry of the received signal and any multipath interference. In particular, the orientation of the antenna systems does not seem to contribute to the biases we are measuring, or we would expect to find a significant change in the biases after rotating the antenna system continuously. From this we conclude that a potential

phase center offset between the antennas does not significantly contribute to the problem represented in figures 1a and 1b.

Two other simple experiments were done. In the first one, the antennas and front end electronics were swapped between the two rotating NIMS units. Thus the systems consisting of L1 and L2 antennas, choke ring ground plane, and front end electronics were reversed between their positions on the roof, and connected to the opposite systems of cables, rotators, and the electronics in the lab. The result of this reversal was a reversal in the sign of the biases, keeping data associated with the laboratory electronics. The values of the biases were apparently the same within the uncertainty, but with the opposite sign. This result suggests that the biases are not associated with antenna position. That is, the differential biases we measure are not associated with the differential multipath interference associated with position, but rather with biases associated with differences in the specific front end antenna systems.

A second experiment added another piece to this puzzle. The L1 antenna on one of the NIMS antenna systems was rotated by 45°. The resultant change in the biases was large. Figures 5a and 5b show the biases for the satellites of 4a and 4b after this rotation. We want to emphasize that though the peak-to-peak values in figures 5 are small, they represent the differential biases between the two systems. We have no information about the biases in either single system.

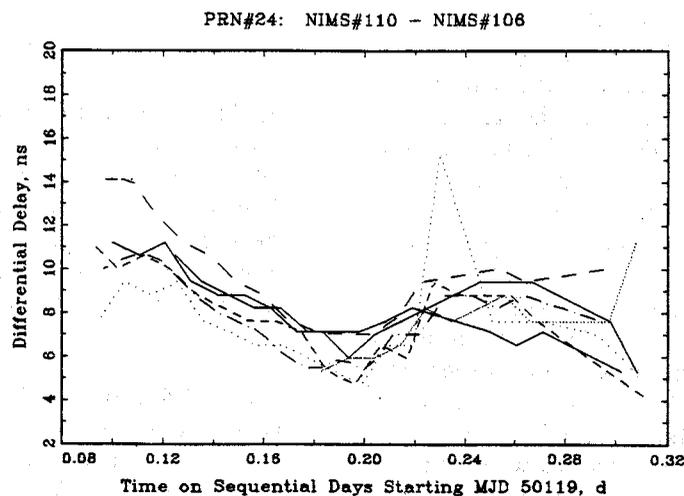
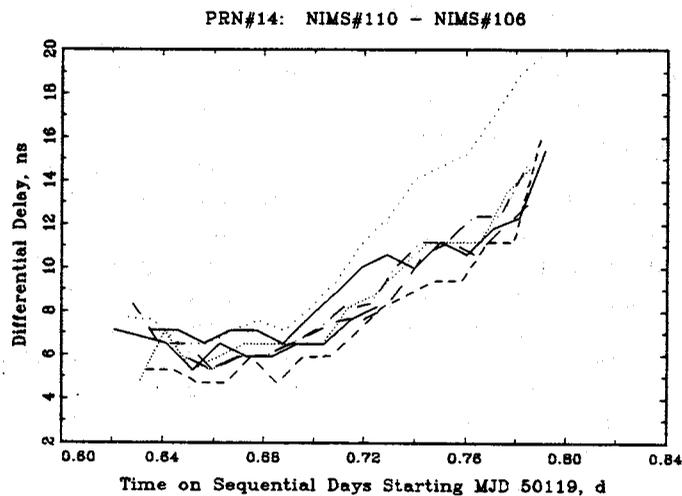
## CONCLUSIONS

We have discussed work on two problems of the NIMS. For the problem of loss of lock because of SA we developed a new frequency lock. The new loop has largely eliminated the problem caused by SA. We now reliably lock on most satellites. The problem of biases in the ionospheric measurements was not solved, but substantial progress has been made toward discovering their cause. At this time the most probable cause lies in the relationship between the two quadrifilar antennas for receiving the L1 and L2 signals. These results on biases are preliminary, and more research is necessary.

## BIBLIOGRAPHY

- [1] B. Guinot and C. Thomas, "Establishment of international time," BIPM Annual Report, Part D, pp. D1-D22, 1988.
- [2] W. Lewandowski, G. Petit, C. Thomas, "Precision and accuracy of GPS time transfer," *IEEE Trans. Instrum. Meas.*, v. 42, pp. 474-479, April, 1993.
- [3] D. Davis, M.A. Weiss, M. Vidmar, "A Codeless ionospheric calibrator for time

- transfer applications," *Proc. Institute of Navigation, Satellite Division*, 2nd International Meeting (ION GPS-89), 1989.
- [4] B. W. Parkinson, S. W. Gilbert, "NAVSTAR: Global positioning system--ten years later," *Proc. IEEE*, v. 71, pp.1177-1186, Oct. 1983.
- [5] P. MacDoren et al, U.S. Patent 4,797,677, Jan. 10, 1989.
- [6] M. Imae, M. Miranian, W. Lewandowski, and C. Thomas, "A dual-frequency codeless GPS receiver measuring ionospheric effects and its application to time comparison between Europe and USA," *Proc. 3rd European Time and Frequency Conf.*, pp.89-93, 1989.
- [7] K. Davies, "Comparison of Hourly Total Electron Contents from GPS and GOES-2," *Proc. International Beacon Satellite Symposium*, Aberystwyth, United Kingdom, 11-15 July, 1994.
- [8] D. Davis, M.A. Weiss, K. Davies, G. Petit, "Improving GPS Time Transfer Accuracy with the NIST Ionospheric Measurement System," *Proc. Institute of Navigation, Satellite Division*, 4th International Meeting (ION GPS-91), 1991.



Figures 1a and 1b: The difference of ionospheric measurements from NIMS#110 minus NIMS#106 made at the same time on the same satellites. The two units were not rotating during these measurements. Curves represent data taken on successive days, and adjusted for the approximately 4 min/d shift of the satellite ground track. We see that there is a repeated changing bias in the offset between the two systems.

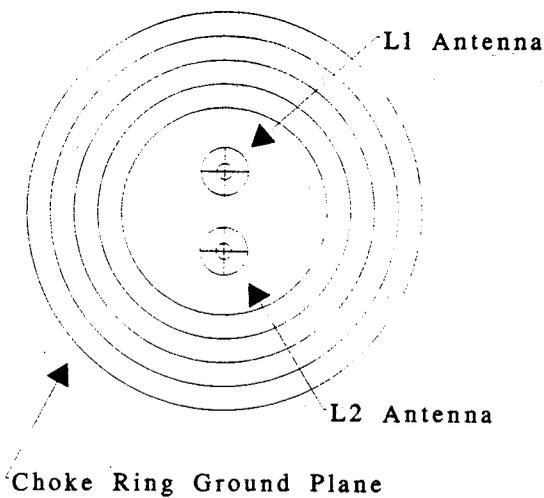


Figure 2: Shows the geometric relationship among the two antennas of the NIMS, L1 and L2, and the choke ring ground plane. The rotation system rotates the entire system around its center.

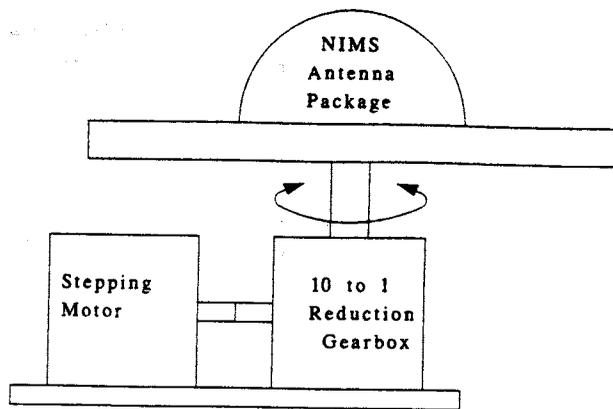
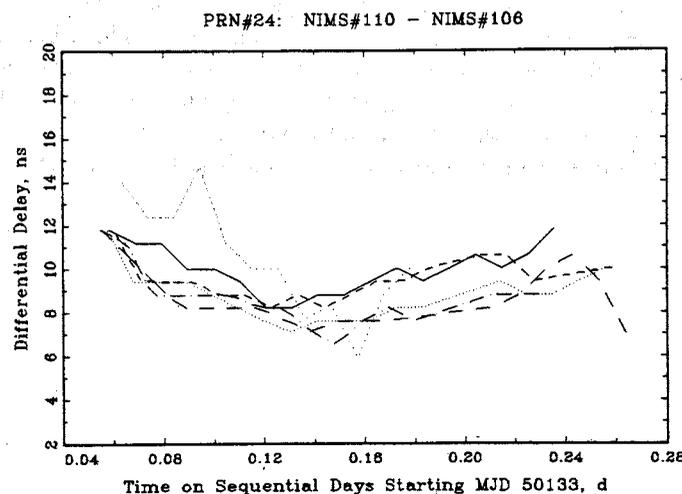
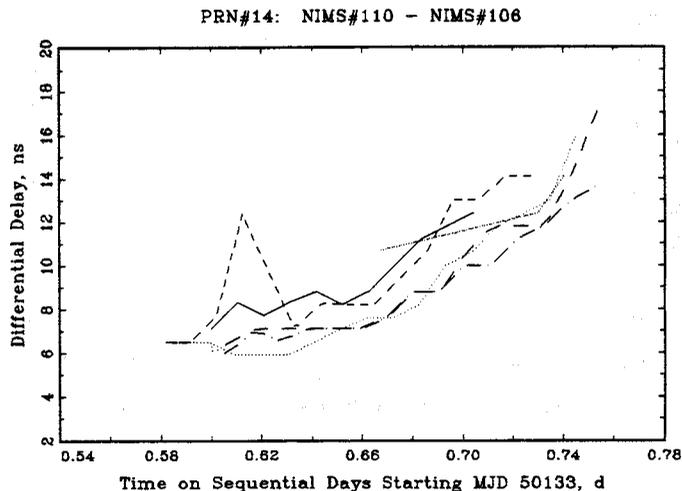
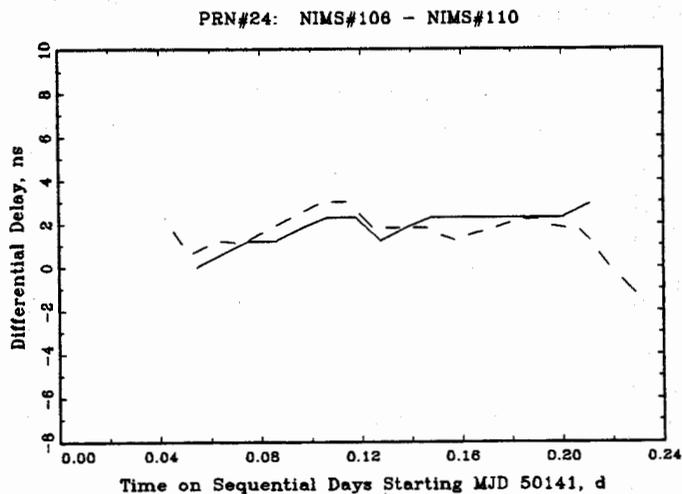
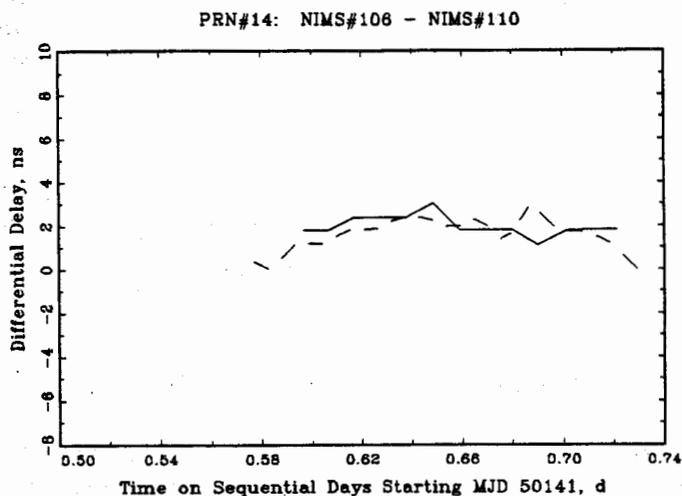


Figure 3: An illustration of the NIMS rotation system. One system, NIMS#106, rotated each direction in 7.5 s, coherent with the NIMS measurements. The other, NIMS#110 rotated each direction in 22.5 s, also coherent with the measurements.



Figures 4a and 4b: As in figures 1, the difference of ionospheric measurements from NIMS#110 minus NIMS#106 made at the same time on the same satellites. In this case the two units were rotating during these measurements. Curves represent data taken on successive days, and adjusted for the approximately 4 min/d shift of the satellite ground track. The repeated changing bias in the offset between the two systems does not seem to have changed.



Figures 5a and 5b: As in figures 1 and 4, the difference of ionospheric measurements from NIMS#110 minus NIMS#106 made at the same time on the same satellites. The antenna packages have been reversed between the units, and the L1 antenna for the unit #110 as been rotated by  $45^\circ$ . The two units were rotating during these measurements. Curves represent data taken on successive days, and adjusted for the approximately 4 min/d shift of the satellite ground track. The repeated changing bias in the offset between the two systems seems to have changed significantly. Note that a reduction in the differential bias does not necessarily imply that the bias itself has been reduced.

## SPECIFICATION OF LONG TERM RELIABILITY OF FREQUENCY CONTROL PRODUCTS

David Brown (C-MAC Quartz Crystals) and Jonathan Clarke (Solicitor with Attwater & Liell)

C-MAC Quartz Crystals Ltd and Attwater & Liell, Harlow, Essex, England

The paper is intended to highlight some technical, commercial and legal issues in the specification of frequency control products, which are of concern to manufacturers. The difficulties inherent in predicting frequency ageing are looked at as are some legal implications.

### BACKGROUND TO SPECIFICATIONS

The products of our industry are intended to define a reference frequency with a relatively high degree of precision. That frequency is then used to enable the function of some communication, navigation or other system. When such products are sold, particularly to large customers, they are normally supplied to meet a customer specification.

The specification for a frequency control product, such as a crystal oscillator, typically may include ten or twenty different statements. Some of these describe basic functional parameters such as the amount of power it is allowed to consume, its physical size and shape and the waveform of its output.

The specification will usually also limit the allowed perturbation of the frequency by external factors. Typically there will be bounds on the amount the frequency may vary when subject to a variation in temperature or supply voltage or perhaps under conditions of mechanical stress.

For any particular example of a product it is normally a relatively simple matter to discover whether its performance in these respects conforms to specification. Although we can sometimes be working close to the achievable limits of parameter measurement the measurements can be repeated or carried out by different people with an expectation of reasonable agreement. In short there is a definite and immediate answer to the question "does the performance of this item fall within the range allowed by the specification". When tested it is either a pass or fail and objective reference can be made to independent standards or experts in case of dispute.

If it is required that the product has a certificate of

conformance one can be issued based on these factual results.

### Ageing

Unfortunately there is also a tendency for the frequency of crystal products to age, for the frequency to drift slightly with time.

Very likely some aspects of the specification will attempt to set a limit on the maximum amount of this "ageing" of the frequency. From the user point of view this is reasonable because there will be some limit on the acceptable variation of frequency to ensure overall performance within the system requirements. In a typical specification over half the paper area is occupied by statements about ageing.

Now we have made a huge step from parameters which are fully verifiable at any time to predictions only verifiable at some future time, if at all. Since the frequency ageing performance varies with time it normally is not possible to exactly repeat and verify an earlier measurement.

What does a typical specification for an oscillator actually call for. First there may well be a general statement like this example for a TCXO:

*The devices shall start up and operate under any combination of supply voltage and ambient temperature specified below. The devices shall meet all of the parameters of the specification under these conditions.*

Parameter	Min	Max	Unit
Supply Voltage	+4.7	+5.2	V
Operating temperature	-20	+70	°C
Operating Life	20		years

This introduces the general expectation of a long life and as far as catastrophic failure is concerned we should use materials and design so that one might expect the device to function for 20 years. That is not a major problem in

itself, but the specification goes on to define limits on the future accuracy of the frequency.

Parameter	Max	Unit
Stability in 24 hours after an initial 72 hours operation	$\pm 0.3$	ppm
Ageing 90 days	$\pm 0.75$	ppm
Ageing first year	$\pm 1.0$	ppm
Ageing 20 years	$\pm 4.6$	ppm

So here is a requirement to meet an ageing specification, what do we do about it?

### Prediction of Ageing Performance

There has been much past study of ageing mechanisms and a comprehensive report was presented in 1991 by Vig and Meeker (1). One conclusion of that paper was that "Even though several possible ageing mechanisms are well understood, the ageing of resonators is not well understood."

Because absolute proof of the real life ageing performance of a particular product example can only be obtained in real time we cannot be absolutely certain that every product we ship will meet the specification. Even products made at the same time through the same intended process will vary somewhat in their behaviour.

Clearly we can't wait ten years measuring the ageing before shipping the product and anyway if we did the next ten years would produce a different result. In the modern environment of short lead times we would prefer not to wait at all. Clearly we need to observe the ageing for a relatively short time and deduce what it will be over life.

Now we have to rely on several assumptions, the big one being that the drift of frequency over a short period allows reliable prediction of the drift over a much longer period.

Conventionally there is a working assumption that normal ageing behaviour corresponds to a logarithmic function of time in the form:

$$f = A \ln(1 + Bt)$$

where  $f$  is the fractional frequency shift from time  $t=0$  and both  $A$  and  $B$  are temperature dependent parameters.

This relationship has been incorporated in MIL Std

MIL-O-55310 although Bower et al (2) concluded it should be regarded as unreliable.

The particular customer ageing specification shown above is also shown graphically in Fig. 1 and it can be seen that it does not anyway correspond exactly to a "MIL" curve.

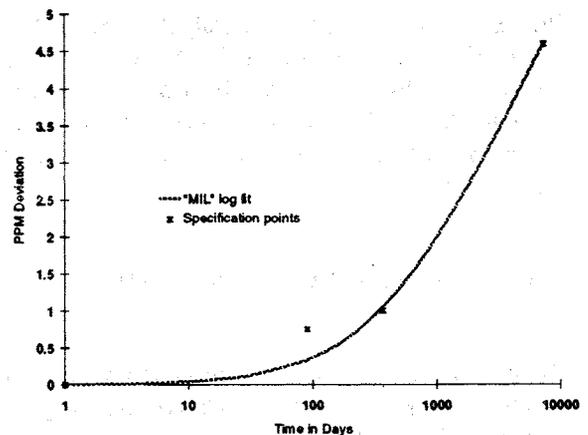


Fig 1.

### Ageing Measurements

In a precision oven controlled oscillator (OCXO) the environmental sensitivity of the device frequency is normally small enough that it is possible to make meaningful measurements of the ageing drift over periods of a few days. In such devices the frequency determining element, the crystal, is at essentially constant temperature during its operating life so such measurements can be expected to be reasonably accurate predictors of likely drift in normal operation although the accuracy will not reach 100%.

Temperature compensated oscillators can now provide a frequency sufficiently constant to replace OCXO products in many applications and TCXOs are widely used in communication systems. In a TCXO the crystal is subject to the full rigours of the environment and the frequency drift with changes in temperature etc. may well be much larger than the weekly ageing drift. Determination of the ageing rate in a short time now becomes a problem and the crystal may be subject to a wide range of conditions.

It is customary to accelerate the ageing by subjecting the device to a high temperature so that the ageing is faster and can be measured with more certainty. Unfortunately this now introduces a further variable which is the actual acceleration factor i.e. the real relationship between the

frequency drift measured in a short time under conditions of acceleration and real life. The observed frequency drift under any particular conditions is at least in part a function of the previous history of the device. In a typical product specification a wide range of conditions are possible and it is simply not reasonable to study performance under every possible combination.

Fig. 2 shows some ageing results for a group of 52 MHz fundamental resonators when aged at 125°C and -2°C after an initial burn in period at 125°C.

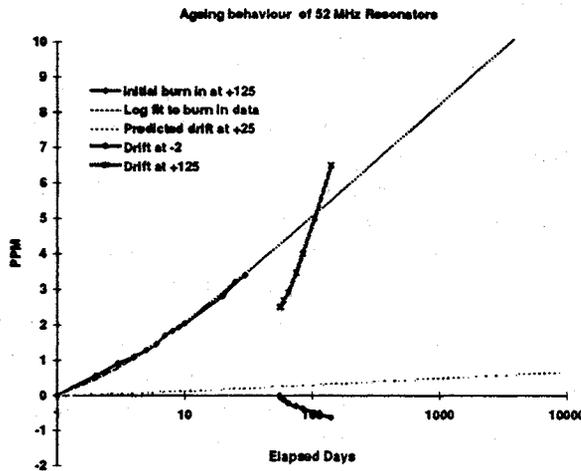


Fig. 2

It is interesting to note that operation at -2°C produces an ageing sign change compared with 125°C, rather than just a reduction in rate which might be expected. This is thought to reflect movement back towards some equilibrium state. Incidentally storage at 125°C and operation well below -2°C are within the range required by this specification. Certainly in this example prediction is not simple.

## CONSEQUENCES

I hope there is agreement that prediction of long term performance from short term information carries with it some real risk of being wrong. Usually we are proved to be conservative in our assumptions but our customers tend to use the best they have been offered as the reference point for future specifications.

Well we only give a one or two year warranty don't we? Or does the ageing specification override our warranty clause? What really would be the legal situation if some product which was supposed to stay within limits for 10 or 20 years drifted out after 8 or 9?

In practice we would no doubt aim to satisfy our customer if possible but in the case of a large problem we might have to consider the legal options.

## THE LEGAL IMPLICATIONS OF SELLING A PREDICTION

### The Commercial Problem

The practical situation which I propose to address is as follows:-

A seller manufactures a device which has to comply with a specification. That specification provides that the device will perform within specified performance parameters throughout a pre-determined period.

When he sells the goods the manufacture warrants that the goods comply with the specification. Additionally he guarantees the goods for a year. He says that if they are found not to fall within the definition and parameters contained in the specification, he will either replace them or take them back and refund the purchase price.

Two problems might arise. Firstly goods which should have continued to operate within the specified parameters throughout say a 25 year period, drift away from that parameter in say year 20.

What remedy does the buyer have bearing in mind the year's guarantee has expired?

The second problem is as follows. The situation is the same but the purchaser does his own tests on the device and concludes from those tests that, on his prediction, the goods will not operate within the defined parameters within say, 20 years time. What remedy does the purchaser have against the seller in these circumstances, if any?

In both situations can the seller protect himself against liability or avoid liability altogether?

### The Legal Background

The law relating to the sale of goods in England is an enormous topic. It was consolidated in one substantial statute about 100 years ago. The development of that statute and the refinement of the law has been undertaken by the courts and the legislators to take into account modern trading conditions ever since.

There is now one body of law which has two distinct sections. However, there are links between those two sections. The first is the law that relates to a private consumer buying from a merchant, trader, manufacturer or supplier.

The second section deals with a trader, merchant or trade purchaser (who has some expert knowledge of the goods that he is buying) purchasing from another manufacturer, trader, merchant or supplier.

**The Private Purchaser.** Thus if I am a normal "off the street" consumer when I go to buy any goods from a "professional seller" (not my next door neighbour or fellow man in the street) the law gives me certain rights and guarantees. Whether I am buying a teaspoon or television I am entitled to expect that the item being bought will be of "merchantable quality". Merchantable quality is rather unhelpfully defined as being fit for the purpose for which goods of that type are commonly bought. If I have explained the purpose for which I want the goods or the purpose is reasonably clear from its description or general usage, I can expect that the items bought will be fit for its purpose. Similarly if I have asked for the goods by reference to some sort of specific description I can expect the goods to comply with that description.

### The Exclusion Clause

The trader who is selling me the goods can affect my position, or attempt to, in two ways. Firstly, he can try and limit the rights that I have as a consumer. He might in his invoice which is presented to me with the goods try to say that he does not accept responsibility if the goods are not of merchantable quality. He will not succeed in such an attempt. The law will declare such an exclusion clause invalid. It will be ineffective and ignored.

**The Manufacturer's Guarantee.** The other possibility is that he might give me a guarantee with the goods. By and large such guarantees are of negligible value. They rarely give a purchaser more rights than he has already under law. The purchaser may feel comforted to see them set out on paper. A guarantee is normally limited in time. Invariably they are said to last for a year, three years or perhaps five years and so on. However, if you buy goods which can reasonably be expected to continue to work faultlessly for five years (if there is such a thing) and your guarantee gives you the right for replacement during only one year, the guarantee in those circumstances appears to reduce your existing legal rights. It will not stop your existing legal rights

taking precedence. In law you have a "guarantee" which is worth more than the written guarantee.

**The Commercial Contract.** The situation between traders or merchants or colleagues in business is different. The law recognises the fact that they are entitled to regulate their own business arrangements between themselves. Therefore a lot of the rights of the average domestic consumer has either do not exist in the business world, or if they do exist, can more easily be restricted by the business parties in their trade dealings between each other.

The notions of fitness for purpose and merchantable quality are still there. However, what the trader is entitled to do is to exclude his liability in respect of them and he has a greater power to do so. The law considers the respective trading positions of buyer and seller in these circumstances. It looks at the reasonableness of such an exclusion clause with more commercial realism that it would in the case of a private domestic purchaser. It looks at the respective trading strengths of the two parties. It looks as their respective powers to insure against difficulties arising from faulty goods. In short it allows them more control on regulating their respective positions in the market place.

**The First Problem.** Against that background let us consider the practical problem that has been identified.

Let us assume that a manufacturer produces some sort of component which he believes will perform true to specification for the next 25 years. He cannot run a 25 year test on the item. He has to sell it now and not in 25 years time.

Under his standard terms and conditions he says that the product is free from defect. Secondly, he says that it has built in accordance with its specification. Thirdly, he says that he guarantees it for a year.

In the specification he sets out the physical characteristics of the component. He then goes on to say that it will operate within certain performance parameters for the 25 year period. Thus he has made the "25 year condition" part of the contract.

**Limitation of Liability.** It should also be noted at this stage that elsewhere in his conditions of sale he has probably stipulated that his financial liability under the contract only goes to the extent of the cost or value of the goods sold. In other words if the items sold only cost £10.00 but as a result of its use and a deficiency in it, damage occurs which costs a million pounds to rectify, his limit is the original £10.00.

This clause is the exclusion clause. It will be subject to the scrutiny of the Court who will consider whether or not it is fair. The Court will look at the trading relationship between the parties. If it is considered to be unfair the Court will not allow the supplier/seller to rely on its terms.

**The Seller's Guarantee.** Let us go back to the seller's guarantee. He is guaranteeing the goods for a year. That is part of his normal terms and conditions for all goods that he sells.

This is an example of a guarantee which is perhaps a comfort to the purchaser but is not likely to give the seller any greater obligation that he has already under Law.

The goods fail to operate within the parameters in year twenty. The buyer might want to try two remedies in this situation. He might like to give the goods back and recover his money. He is unlikely to succeed in that respect. He has had the goods 20 years. The Court will tell him that he has had a fair crack of the whip. If he has a remedy it will be for damages from the seller. In the normal course of events the amount of damages would relate to the financial loss that the buyer has suffered as a result of the deficiency.

As can be seen in this situation the seller's guarantee is irrelevant.

**Limiting the Amount of Damages.** We must remember that when looking at the law and applying it, the Courts must take account of the fact that they are working in the real world with real people having to go about their business and earn a living and attempt to make a profit. What the court has to do is react to the problem taking into account the commercial realities.

Thus if a supplier says I have done all that I can reasonably do to test the reliability of this component over the next 25 years. I have used all available current technology. I can demonstrate that my testing of this component is sensible and reasonable bearing in mind its cost and the facilities available to me. I have made known to the purchaser the processes which I am using. I warranted the physical construction of the component and made the purchaser aware of that construction. In other words I have done all that is reasonably possible for me to do.

Let us then propose that the seller goes onto say that having done all that he feels reasonable he cannot guarantee 100% forecasting of what is going to happen in 24 years time. He therefore seeks to exclude his liability for something that he cannot test against and

which neither he nor the average purchaser could possibly foresee.

In those circumstances one can look favourably on the proposition that the court should allow the seller to make use of and take advantage of this sort of exclusion clause. It seems reasonable in all the circumstances.

**How does the court test the reasonableness of the Limitation Clause?** The Court considers whether the clause is fair and reasonable having regard to the circumstances which existed when the Contract was made.

In reaching that decision the Court is directed to consider the following:-

- (i) The strength of the bargaining positions of the two parties relative to each other including taking into account alternative means by which the purchaser's requirements could have been met, i.e. whether he could have bought the goods elsewhere.
- (ii) Whether the buyer received an inducement to agree to the clause or whether there was available to him at the time that he entered into a contract the opportunity to enter into a similar contract with another seller but without the limitation clause.
- (iii) In a case where the buyer is saying that he did not know of the existence of the limitation clause, whether he should have known bearing in mind the previous trading relationship that he had with the seller and trading situation in the business generally.
- (iv) Where the clause excludes or restricts liability if some condition is not complied with, whether it was reasonable at the time of the contract to expect that compliance with that condition would be practicable.
- (v) Whether the goods were manufactured or processed to the special order of the customer.

Where, as is usually the case, the limitation clause restricts the seller's liability to a particular sum of money, and considering whether the clause was reasonable the Court must have regard to:-

- (a) The resources which a seller could expect to be available to him for the purpose of meeting the liabilities should it arise and
- (b) How far it was open to him to cover himself by insurance.

Taking all this into account. And on the hypothesis that I set out above it does seem that in those circumstances

the limitation of liability could well be effective. As a result the seller can provide himself with some protection, but he will end up paying a reasonable amount of damages.

### The Second Problem

We also have the interesting problem of a conflict of predictions. The seller says that he has done his tests on the component and concluded that it will not vary outside the defined parameters for a period of 25 years. He has reached that conclusion as a result of accelerated tests that he has carried out over the space of 6 months. The purchaser buys the goods on that basis. During the space of 2 years he carries out testing which is "accelerated" but less accelerated than the seller's. He says that according to his tests after 10 years the performance drifts beyond the required parameters. Let us suppose for the purpose of this example that we are now 3 years from the date of the original purchase. Ten thousand of these components have been bought that fitted to the end product and shipped. The buyer now says that they are only going to work for 3 fifths of the period required. He wants to know what recourse he has got against the seller.

**Is a Prediction Proof?** Clearly the goods do not comply with the specification, or do they? The hurdle that the seller has to overcome first of all, before he can consider any remedy is that of proof. He would have to go to the court and satisfy the court that on the balance of probabilities his prediction of what is going to happen in 10 years time is reliable and sufficiently reliable for a Judge to act upon it. I would have thought that the average Judge is going to be very cynical about deciding a case on the basis of a prediction.

**What is the Remedy?** However, let us assume that the Judge does agree with the buyer. How is he going to give the buyer some remedy for the breach. What is the buyer's loss at the stage that he is in court? The difficulty is that he has no loss. The components are still working perfectly satisfactorily and as yet nobody has complained. He might have a go at persuading the Judge that he has to recall all the goods and refit the components. The cost of doing that would presumably be enormous. It is likely to be caught by the seller's exclusion clause limiting his liability. Additionally and in any event a Judge is going to take a good deal of persuading to accept that result when he has only based his decision on a prediction.

Thus in those circumstances the buyer has a fairly

substantial hurdle to overcome in proving his case. If he can prove it he is at a very practical disadvantage in trying to demonstrate what remedy he is entitled to.

**The Guarantee?** Let us then consider the situation of the component having been sold with a years guarantee. After 9 months the buyer having done his tests states that his prediction is different from that of the seller. He says that therefore the goods are not within specification. He relies on the guarantee to say that he wants to either give the goods back and get his money back, or he wants them replaced.

The strict terms of the guarantee allow him replacement goods. However we go back to the argument of whether predictive tests of the buyer are any better of those than the seller. Again ultimately that is a question of evidence before the Judge. The Judge will have to hear the expert evidence of each side and, given that both are predictions, make up his mind as to which he thinks the best. If he decides the buyer has the best expert and can prove his case he will allow him a remedy under the Sale of Goods Act or any additional remedy given under the terms of the guarantee. For the buyer to get his money back and give back the goods would be difficult to achieve. He would have to show that fundamentally what he had bought was different from that which he had contracted to buy. Almost certainly that would not be the case. It would however depend upon the precise nature of the goods sold and the purpose for which it was going to be put, and how useful or useless the component would be in carrying out that purpose given the diminished specification.

I am assuming that the guarantee does not specifically state that under its terms the buyer can require a refund of the price. If the guarantee says that the buyer is entitled to get it. The seller might not be happy with that result bearing in mind that we are talking about is a divergence in predictions.

### Does English Law apply to the Contract

In deciding what law governs a contractual situation between two parties the determination is made in one of three ways:-

1. By the parties choosing which law should be applicable,
2. By the Court inferring a selection from the circumstances of the contract or, failing either of these,
3. By the Court deciding which system of law has the closest and most real connection with the contract.

In most commercial contracts there is a clause which states which law is to apply. Often the seller puts into his terms and conditions a clause stating that the law of his country shall be applicable.

The Court will usually accept that. Similarly the Court would probably accept a situation where the parties choose the law of the country different from that of either party. English law is often selected to govern contracts and the Courts usually accept that.

In the unlikely event of there being no choice of law the Court will look at the circumstances and where appropriate reach a conclusion on whether the parties have an inferred choice of law or one can be implied. They can reach a conclusion from such things as the currency expressed in the contract, the language used, the connection with preceding transactions and so on.

If there is no place that it can be implied and none stated the Court then decides what is the closest and most real connection. It will look at such factors as the place of contracting, the place of performance of the contract, the place of residence or business of the parties and the nature and subject matter of the contract.

## REFERENCES

1. Vig J R and Meeker T R, 1991, "The Aging of Bulk Acoustic Wave Resonators, Filters and Oscillators", 45th Symposium on Frequency Control.
2. Bower D, Churchill N, Dowsett J and Salt D, 1995 "A look into the crystal ball: some thoughts on the prediction of the ageing of BAW and SAW resonators", EFTF 1995.

## SAW TWO RESONATOR FILTER USING SYNCHRONOUS COUPLING INTERDIGITAL TRANSDUCERS

Waldemar Soluch  
Institute of Electronic Materials Technology  
Wolczynska 133, Warsaw, Poland

### Abstract

Surface acoustic wave (SAW) two resonator filter with coupling interdigital transducers (IDT) placed inside reflectors, was designed, fabricated and tested. In the filter, electrodes of the coupling IDTs are in spatial synchronism with reflectors' electrodes. Compared to the case when the coupling IDTs are between reflectors, resonators are shorter, the unwanted longitudinal modes are suppressed and better rejection should be obtained in the stopband. The filter, fabricated on ST-cut quartz, had the centre frequency of 302.6 MHz, insertion loss about -14 dB, and minimum rejection in a stopband higher than 40 dB.

### 1. INTRODUCTION

Coupling networks in SAW two resonator filters are usually placed between reflectors (Fig.1). To obtain reasonable coupling in the case of low electromechanical coupling substrates, the coupling networks should be long (Rosenberg and Coldren (1)).

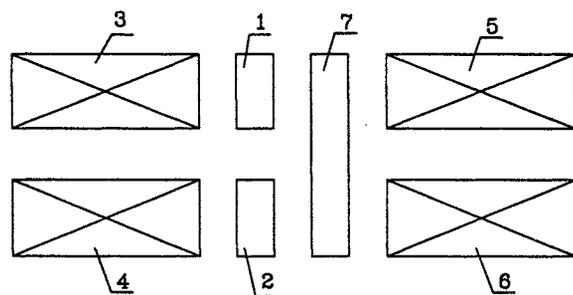


Fig. 1. SAW two resonator filter  
1, 2 - IDTs, 3, 4, 5, 6 - reflectors,  
7 - coupling network.

Therefore resonators are also long, the nearest longitudinal modes are close to the centre frequency, and rejection near the passband is low. A possible solution to this problem can be placement of the coupling network inside reflectors. For example, it was recently shown, that coupling network can be formed by connecting open reflectors' strips (Soluch (2)).

It is the purpose of this work to show, that in the case of short circuited reflectors strips, coupling IDTs can be placed inside reflectors.

### 2. FILTER DESIGN

The structure of the filter is shown in Fig.2. Each reflector consists of three parts: bus bars for input and output IDTs, coupling IDTs and the remaining part of short circuited strips. Coupling IDTs are parallelly connected by the rear parts of reflectors. Input and output IDTs are apodised, while coupling IDTs are unapodised.

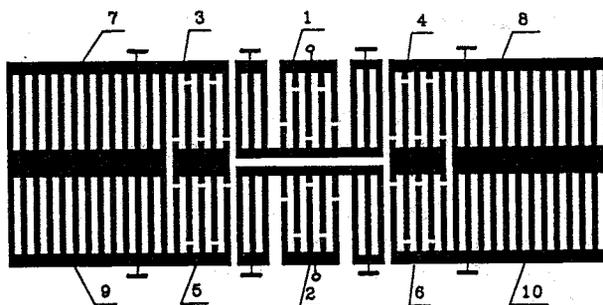


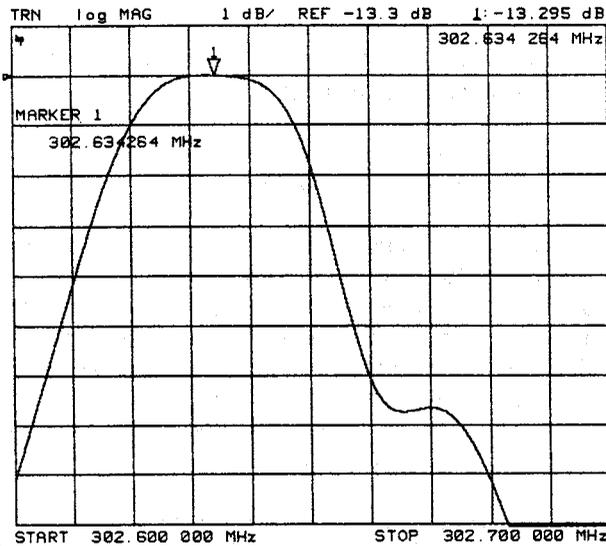
Fig. 2. SAW two resonator filter with coupling IDTs placed inside reflectors  
1, 2 - input and output IDTs,  
3, 4, 5, 6 - coupling IDTs,  
7, 8, 9, 10 - rear parts of reflectors.

Scattering matrix theory was used for the calculation of the amplitude transfer function of the filter (1). To determine the scattering coefficients of reflectors, the reflection coefficient for one aluminium strip on ST-cut quartz  $r = -0.5h/\lambda$  was used, where  $h$  is the thickness of the aluminium layer, and  $\lambda$  is the wavelength (Dundrowicz (3), Tanski and van de Vart (4)). The scattering coefficients of the apodised IDTs were calculated from their input admittances (Soluch (5)). ST-cut quartz parameters were taken from Slobodnik (6). In this calculations, the cross channel coupling coefficient was used as a parameter. It will be determined from the comparison of calculated and measured amplitude responses of the filter.

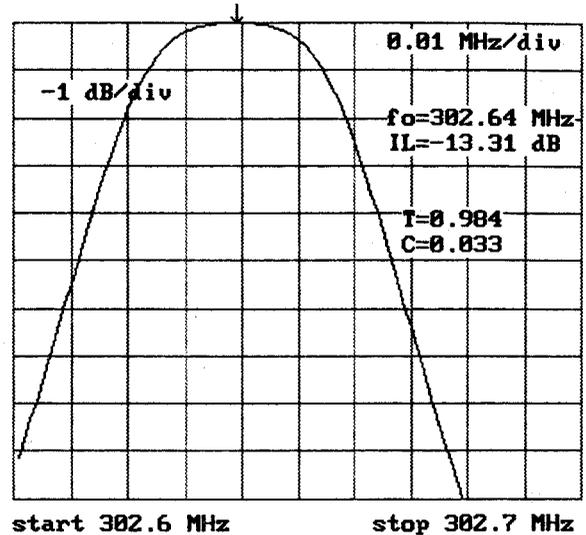
The following data of the filter were used: acoustic aperture - 1 mm, total number of each reflector strips - 600, number of electrodes of coupling IDTs - 101, number of electrodes of the input and output IDTs - 61, number of bus bar strips - 10, period of electrodes - 5.2  $\mu\text{m}$ , width of electrodes - 2.6  $\mu\text{m}$ , lateral shift of resonators - 0.3 mm, aluminium layer thickness - 0.12  $\mu\text{m}$ . The input and output IDTs were placed at the positions of maximum coupling.

### 3. MEASUREMENTS AND CALCULATIONS

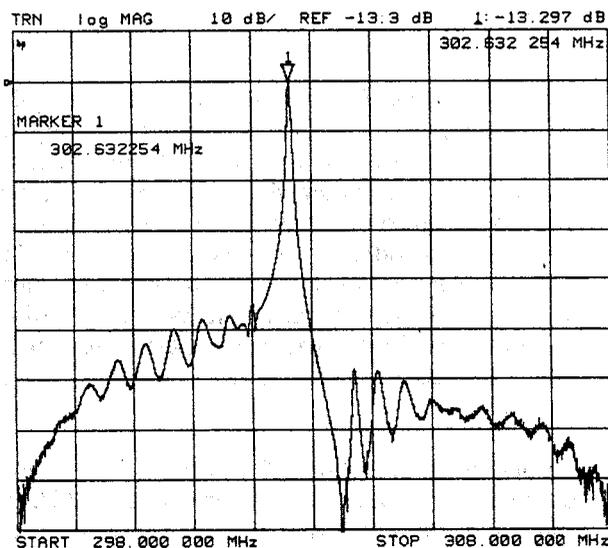
Electron beam was used for photolithographic mask fabrication. Electrodes of the filter were deposited on the ST-cut quartz substrates by the lift-off method. Filters were mounted in metallic packages. Transfer properties of the filters were measured in a 50  $\Omega$  system (HP Network Analyser 8752A). Typical



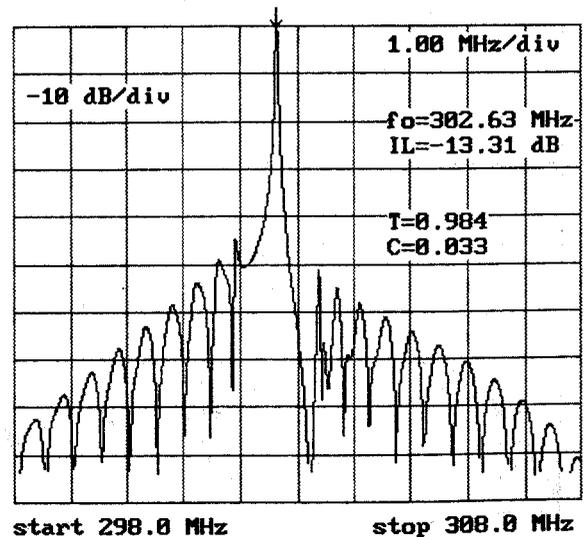
a)



a)



b)

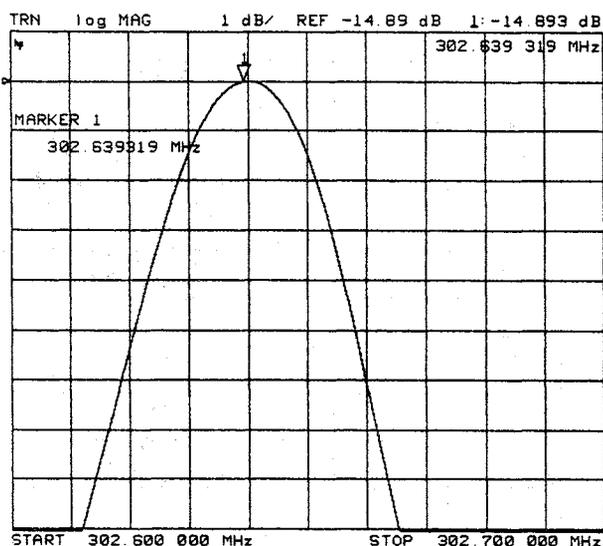


b)

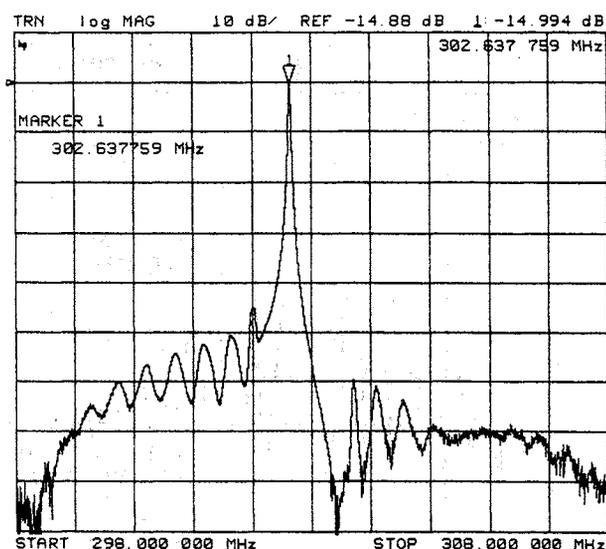
Fig.3. Measured amplitude responses  
a) narrowband response,  
b) broadband response.

Fig.4. Calculated amplitude responses  
a) narrowband response,  
b) broadband response.

amplitude responses of the filter are shown in Fig.3. Spurious transverse mode is seen above the centre frequency (Fig. 3a). Characteristic asymmetry of the transfer response (Fig.3b) is the result of the synchronous placement of the coupling IDTs. In this case, the centre frequency and insertion loss were equal to about 302.6 MHz and -13.3 dB, respectively. The stopband rejection was higher than 40 dB at frequencies below the passband, and higher than 50 dB at frequencies above the passband.



a)



b)

Fig.5. Measured amplitude responses for the case of one coupling IDTs short circuited  
 a) narrowband response,  
 b) broadband response.

To obtain an agreement between the measured and calculated results, several parameters used in calculation should be properly corrected. It was found that:

1. Velocity of the SAW in the periodically metallised areas  $v_h = 3147.46$  m/s,
2. Coefficient, which takes into account an attenuation of the SAW at the effective length of a resonator,  $T = 0.984$ ,
3. Cross channel coupling coefficient  $C = 0.033$ ,
4. Load impedance  $Z = R + j\omega L$ , where  $R = 50 \Omega$  and  $L = 0.05 \mu\text{H}$ .  $L$  represents a parasitic inductance of connecting wires.

Calculated value of the reflection coefficient  $\Gamma$  at the centre frequency was equal to 0.999.

The amplitude responses, calculated from an analytical expression (1), are shown in Fig. 4.

Reasonable agreement exists between the measured and calculated amplitude responses of the filter. In particular, the asymmetry of the responses is similar, and the stopband rejections are in good agreement. The shape of the stopband is determined mainly by the frequency responses of the coupling IDTs and reflectors. It is worthy of note, that excellent feedthrough suppression was obtained in the filter.

Fig.5 presents measured amplitude responses for the case when one of the coupling IDTs was short circuited. In this case the spurious transverse mode does not exist and the insertion loss is equal to about -14.9 dB.

#### 4. CONCLUSIONS

SAW two resonator filter with synchronous coupling IDTs was designed, fabricated and tested. As compared to the filter with coupling IDTs located between reflectors, these filters are shorter and better rejection should be obtained in the stopbands. This conclusion is especially important in the case of low electromechanical coupling substrates, like quartz. However, asymmetry of the frequency response, which is characteristic for the synchronous placement of IDTs, makes the rejection smaller at the frequencies below the passband.

Location of the coupling IDTs in the reflectors' areas can be used as a method of obtaining additional coupling in filters with coupling IDTs located between reflectors.

Compared to the filters with stopband multistrip couplers (2), application of the synchronous coupling IDTs should be more useful, since they can be easily weighted. Therefore the frequency response of the cross channel coupling and coupling to the spurious transverse modes can be varied.

When coupling IDTs are located inside reflectors, a so called QARP method can be used for SAW to bulk scattering loss suppression (Ebata (7)). In this case, no additional gaps between input and output IDTs and reflectors exist, but period of the IDTs is smaller than that of reflectors. It makes possible to use large reflection coefficient for one strip, reflectors can be shorter, which leads to farther reduction of filter length. It also means, that the aluminium layer thickness will be larger, which is important for lowering of ohmic losses at high frequencies. Smaller number of reflectors strips also means a higher yield in technology.

## 5. REFERENCES

1. Rosenberg R L, Coldren L A , 1979, "Scattering analysis and design of SAW resonator filters", IEEE Trans. on Sonics and Ultrason., SU-26, 205-230
2. Soluch W, 1994, "Application of a stopband multistrip coupler in a SAW resonator filter", IEEE Ultrason. Symp. 249-251
3. Dundrowicz C, 1976, "Reflection of surface waves from periodic discontinuities", IEEE Ultrason. Symp. 386-390
4. Tanski W J, van de Vart H, 1976, "The design of SAW resonators on quartz with emphasis on two port", IEEE Ultrason. Symp., 260-265
5. Soluch W, 1993, "Admittance matrix of a surface acoustic wave interdigital transducer", IEEE Trans. on Ultrason., Ferroelec., Freq. Contr., 40, 828-831
6. Slobodnik A J, 1976, "Surface acoustic waves and SAW materials", Proc. IEEE, 64, 581-595
7. Ebata Y, 1988, "Suppression of bulk-scattering loss in SAW resonator with quasi-constant acoustic reflection periodicity", IEEE Ultrason. Symp., 91-96

## ACKNOWLEDGEMENT

I would like to thank M. Teodorczyk and E. Lipinska for filters fabrication and to T. Wrobel for mask design and filters measurements.

This work was supported by the State Committee for Scientific Research (KBN) under grant number 8 S501 051 04.

# FREQUENCY STABILITY CONTROL OF CARRIER-ENVELOPE RESONANCE IN A SURFACE ACOUSTIC WAVE DELAY LINE OSCILLATOR.

Franck LARDET-VIEUDRIN, Gilles MARTIN, Serge DOS SANTOS, Gilles MARIANNEAU and Michel PLANAT.

*Laboratoire de Physique et Métrologie des Oscillateurs*

*associé à l'Université de Franche Comté*

32, Av. de l'Observatoire, 25044 BESANÇON CEDEX, FRANCE

## INTRODUCTION.

Low frequency noise with hyperbolic power spectral density (the so-called  $1/f$  noise) is a ubiquitous and fascinating phenomenon of nature. Despite its chief importance in solid state high technology devices its explanation still remains a challenge [1,2].  $1/f$  noise is also observed in the frequency generated by LC and quartz crystal oscillators and in atomic clocks. The state of the art in the domain of frequency fluctuations metrology results from the use of two basic tools for stability characterization: Allan variance in the time domain and Fourier spectra in the frequency domain. In the time domain  $1/f$  noise leads to a flicker floor of the Allan variance. The statistical analysis of the series of frequency deviations cannot be of any help for further progress.

In order to get round the statistical bottleneck, we recently proposed an alternative method [3,4]: the series of frequency readings was processed by means of an effective sliding filtering window and interpreted in terms of a series of local multiscale stability exponents. The discontinuities in this series were transformed into a binary coding in which remote correlations were observed. The binary data were mapped onto the steps of a devil's staircase [5] which clearly demonstrated global correlations in the time series, correlations that are not detected in the traditional statistical approach. These results were obtained in the low frequency part of the spectrum of a LC [3] and of a quartz oscillator [4].

It is shown in this paper that the frequency fluctuations are controlled to a large extent by synchronization processes and can be accounted for by a nonlinear mapping. This provides the theoretical framework for the previously observed correlations in the low frequency noise. In addition we present a simple experimental set-up and data to support our conclusion.

## EXPERIMENTAL SETUP.

The experimental arrangement is aimed at the study of what we term 'carrier-envelope' resonances

of a delay line oscillator. These resonances are revealed as steps, that is synchronization zones on a frequency-amplitude characteristic, on which the frequency stability is considerably enhanced (up to eight orders of magnitude). The frequency  $f_0$  and the root mean square (r.m.s.) amplitude  $u$  of the carrier is set from an external injection source. Thanks to an active modulator (MOD) this high frequency signal is splitted in wavepackets of which the duration  $\kappa$  is controlled in a loop built from a surface acoustic wave delay line (SAW), an envelope detector (DET) and a low frequency amplifier (A) as shown in Fig 1.

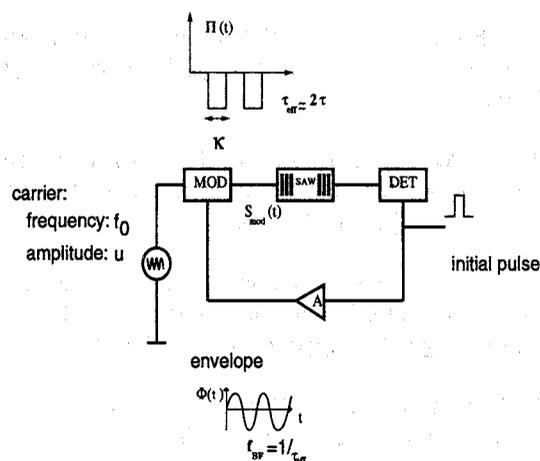


FIG. 1. Schematic of the experimental set-up allowing the study of carrier-envelope resonance. MOD: rectangular modulator; SAW: surface acoustic wave delay line, DET: linear envelope detector; A: low frequency amplifier.

The delay line consists of two transducers deposited on a quartz crystal doubly rotated cut. The interval between two adjacent fingers is  $l_0 = V_r/2f_0 \sim 15\mu\text{m}$ , with the velocity of the monochromatic surface acoustic wave equal to  $V_r \sim 3.10^3\text{ms}^{-1}$ , so that a carrier frequency  $f_0 \sim 100\text{MHz}$  has been selected. The center to center distance between input and output transducers  $l \sim 4.5\text{mm}$  leads to the delay  $\tau = l/V_r \sim 1.5\mu\text{s}$ ; this delay is observed for short pulses in the open loop configuration.

This effective delay  $\tau_{eff}$  results from the non linear propagation of wave packets in the acousto-electronic loop. Its mean value appears to remain close to  $\tau_{eff} = 2\pi/\omega_{BF} \sim 2\tau$ , i.e. twice the monochromatic delay. The resulting envelope frequency  $\omega_{BF} \sim 300\text{KHz}$  is amplitude dependant:

$$\omega_{BF} \sim \omega_{BF}^0(1 + \beta Y^2), \quad (1)$$

with a non linear parameter  $\beta \sim 0.12V^{-2}$ . In addition there are wide **synchronization steps** at integer values of the carrier-to-envelope frequency ratio  $\Omega = \omega_0/\omega_{BF}$ .

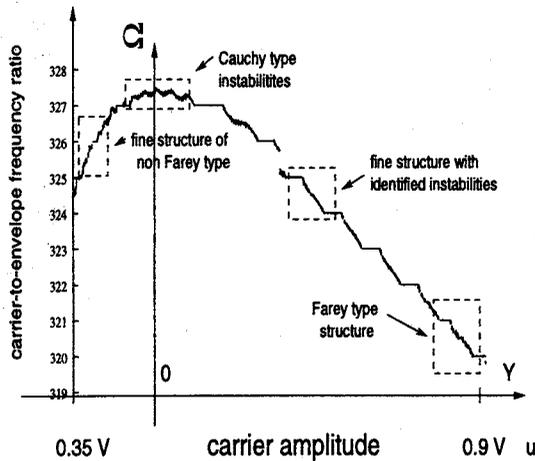


FIG. 2. Schematic of the frequency-amplitude characteristic: synchronization zones and fine structure.

They are attached to a lorentian shape characteristic  $\Omega = \Omega^0/(1 + \beta Y^2)$  with  $\Omega^0 = \omega_0/\omega_{BF}^0$  as shown in Fig 2. Finally there is a fine structure which is dependant on the zone of the characteristic: it will be described later.

### THEORETICAL APPROACH.

The propagation of wavepackets  $\Phi(t)$  follows from the equ.:

$$i \frac{d\Phi}{dt} + \omega_{BF}^0 \Phi(1 + \beta |\Phi|^2) = \sqrt{2}u\omega_0 \exp(i\omega_0 t). \quad (2)$$

It is easily observed that in the undriven case, if the wavepackets are switched off from the modulator, constant amplitude modulations with frequency  $\omega_{BF}$  satisfies the dispersion relation (1).

In the driven case slowly varying amplitude and phase modulations are obtained, that is:

$$\Phi = Y(t) \exp i(\theta(t) + \pi/2), \quad (3)$$

so that the following first order system is obtained:

$$\begin{cases} \dot{Y} = -\sqrt{2}u\omega_0 \cos(\omega_0 t - \theta(t)), \\ \dot{\theta} = \omega_{BF}^0(1 + \beta Y(t)^2) - \frac{\sqrt{2}u\omega_0}{Y(t)} \sin(\omega_0 t - \theta(t)). \end{cases} \quad (4)$$

If the wavepackets are switched on, synchronization states are described by assuming that the amplitude  $Y$  remains constant and phase varies linearly over time intervals equal to the carrier period, i.e.:

$$\begin{cases} \theta = \omega_0 t + \theta_0 & \text{and} & Y = Cte, \\ \text{when} & \frac{2\pi n}{\omega_0} < t < \frac{2\pi(n+1)}{\omega_0} & \text{and } n \text{ integer.} \end{cases} \quad (5)$$

As a result the continuous differential equation is transformed into a two dimensional mapping:

$$\begin{cases} Y_{n+1} = Y_n - c \cos \theta_n, \\ \theta_{n+1} = \theta_n + \frac{2\pi}{\Omega^0}(1 + \beta Y_{n+1}^2) - \frac{c}{Y_{n+1}} \sin \theta_n, \end{cases} \quad (6)$$

with the non linear parameter  $c = 2\pi\sqrt{2}u$  proportional to the injection amplitude  $u$ .

Fundamental fixed points of the mapping satisfy  $Y_{n+1} = Y_n$  and  $\theta_{n+1} = \theta_n + 2\pi s$  where  $s$  is integer, that is  $\cos \theta_n = 0$ ,  $\sin \theta_n = \pm 1$  and correspond to a set of amplitudes given by:

$$\beta Y^3 + Y(1 - s\Omega^0) \pm \sqrt{2}u\Omega^0 = 0. \quad (7)$$

The real solutions of equ.(7) with index  $s$  point to the fundamental steps  $327 < \Omega < 320$  on both sides of the characteristic.

The counting of an effective frequency ratio  $\nu = \omega_0/\omega_{BF}$  is performed by introducing the average phase (or winding number) as follows:

$$\nu = \lim_{n \rightarrow \infty} \frac{\theta_n - \theta_0}{2\pi n}. \quad (8)$$

It is well known [6,9] that for Arnold type maps of the form (6) the  $\nu$  versus  $\Omega^{-1}$  graph is a devil's staircase with the wider steps located at  $\frac{0}{1}$  and  $\frac{1}{1}$  and the sub-harmonics located at  $(\frac{1}{2})$ , then  $(\frac{1}{3}, \frac{2}{3})$ , ...,  $(\frac{p+p'}{q+q'})$ ,  $p$  and  $q$  integers). This so-called Farey tree was observed before in experiments involving Josephson junctions or nonlinear electronic circuits [7]. In the present case the devil's staircase structure is attached to the characteristic owing to the non linear relation (1).

## FINE STRUCTURE.

Let us point out two significant regimes in the experiment and their theoretical explanation. One regime occurs at the peak of the characteristic when  $Y = 0$  ( and  $c/Y \rightarrow \infty$  ) so that the  $(n+1)^{th}$  phase jump calculated as

$$X_{n+1} = \theta_{n+1} - \theta_n - \frac{2\pi}{\Omega^0} = \tan \theta_n. \quad (9)$$

In this case phase jumps may have arbitrarily large amplitudes. Numerical and theoretical analysis [6,9] predicts that the phase jump density obeys the Cauchy law

$$f(X) = \frac{1}{\pi} \frac{1}{1+X^2}.$$

In our set-up the envelope oscillator is indeed the most unstable in a range of amplitudes around the top of the characteristic (the relative stability is of the order  $10^{-3}$ ) and the observed density is close to the predicted one. Another regime is obtained at high amplitude  $Y$ . In such a case there are limit cycles located at fixed points of the amplitude  $Y$  and the phase mapping converts into the generic Arnold map [6,7]

$$\theta_{n+1} = \theta_n + \frac{2\pi}{\Omega} - \frac{c}{Y} \sin \theta_n. \quad (10)$$

As long as the coupling parameter  $d = c/Y \leq 1$ , the mapping (6) is invertible and resonances ordered as a Farey tree are expected as claimed before. At  $d > 1$  higher order resonances interactions should appear and the dynamics become chaotic. To identify the subharmonics we make use of the stability exponent  $\alpha_k^{(N)}$  given by:

$$\begin{cases} \alpha_k^{(N)} = \frac{\log|\Delta\nu_k^{(N)}|}{\log(N\Delta Y)}, \\ \Delta\nu_k^{(N)} = \frac{1}{N} \left( \sum_{i=N+1}^{i=2N} \nu_{k+i} - \sum_{i=1}^{i=N} \nu_{k+i} \right), \end{cases} \quad (11)$$

where  $\nu_k$  is the winding number of the  $k^{th}$  amplitude point,  $\Delta Y$  is the amplitude gap between two adjacent samples,  $N$  is the number of averaging points and  $\Delta\nu_k^{(N)}$  is the corresponding winding number deviation. The stability exponent takes its maximum close to the center of a step on the characteristic. Fine structure of Farey type is effectively observed in the upper limit of the characteristic (Fig. 3).

Resonance ratios up to the fourth branching level are identified. The best stability  $10^{-11}$  is observed on the fundamental levels and approximately one decade of frequency stability is lost at each branching.

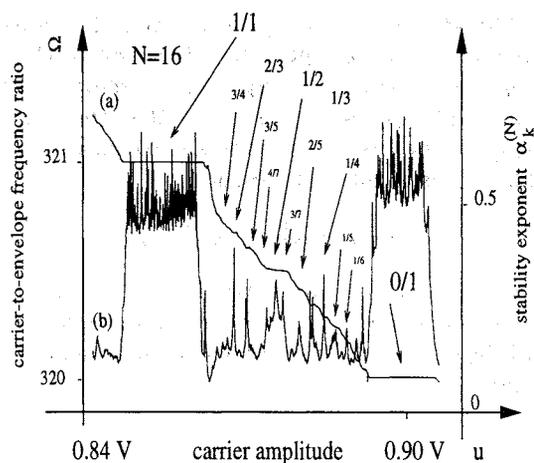


FIG. 3. The Farey type fine structure at high amplitudes. (a) Winding numbers  $\nu_k^{(N)}$ . (b) Local stability exponents  $\alpha_k^{(N)}$ ; Averaging number:  $N = 16$ .

Two others amplitude zones may be mentioned. At the right side of the top ( $Y > 0$ ) the characteristic shows amplitude zones with increasing instability. They correspond to well identified zones of winding numbers with bad rational approximations. A theoretical explanation for such a behavior is currently under investigation and will be reported in the near future. At the left side of the top ( $Y < 0$ ) the characteristic shows subharmonic zones of the non Farey type. In particular subharmonics at  $\frac{p}{2p \pm 1}$  ( $p$  integer) up to the fourth branching order have been identified. The explanation for such a dependence is still lacking.

## CONCLUSION.

We have succeeded in measuring and explaining the frequency locking between the carrier and the envelope of a delayed electronic oscillator. It is shown that a nonlinear mapping accounts for the overall features of this synchronization and that the fine structure in the frequency noise can now be understood on an analytical basis for the first time. This extends previous work concerned with the origin of the correlations in  $1/f$  frequency noise of a single frequency oscillator [3,4]. This is an important step towards the development of new kinds of stable devices: frequency dividers, analog to digital converters, neural networks and highly sensitive sensors.

- [1] D.W.Allan , *IEEE Trans. on Ultras. Ferroelect. and Freq. Control, UFFC - 34* , 647 (1987).
- [2] M.B. Weissmann, *Reviews of Modern Physics*, **60**, 537, (1988).
- [3] M. Planat, G. Marianneau, N. Ratier and F. Lardet-Vieudrin, *Appl. Phys. Lett.* **21**, 3206 (1995).
- [4] M. Planat, V. Giordano, G. Marianneau, F. Vernotte, M. Mouret, C. Eckert and J. Miehé, *IEEE Trans. on Ultras. Ferroelec. and Freq. Control*, (march 1996).
- [5] M.C. Gutzwiller and B.B. Mandelbrot, *Phys. Rev. Lett.*, **60**, 673, (1988).
- [6] M. Planat, P. Koch, *Fractals* , **1** , 727 (1993).
- [7] P. Bak, T. Bohr and M. Jensen , in "Directions in Chaos", edited by Hao Bai-lin ( World Scientific, Vol 2, 1988), p. 16.
- [8] M. Remoissenet. "Waves called solitons.", (Springer Verlag ed, 1994) , p. 106.
- [9] M. Planat. in "Noise in Physical Systems and 1/f fluctuations." ed by T. Musha, S. Sato and M. Yamamoto, (Ohmsha, Ltd, 1991), p. 489-492.

## HYBRID WIDE-BAND VCOs USING LOW-LOSS LEAKY SAW RING FILTERS

S.Doberstein, A.Martynov, V.Malyukhov

ONIIP, Maslennikov str., 231, Omsk, 644063, Russia

The transceivers for the mobile communication systems require wide-band voltage controlled oscillators (VCOs). A widely known Colpitts-type VCO with a SAW resonator provides the fractional variable frequency range of 0.38% [1]. To extend the fractional variable frequency range to 3-5% wide-band slanted interdigital transducers (IDTs) may be used in this VCO [2]. We used VCO with SAW delay line and external phase shifter in the amplifier feedback loop [3] for 146-174, 440-470 MHz mobile communication receivers. For delay line we used previously developed low-loss leaky SAW ring filter with insertion loss of 1 dB having linear phase response with the phase shift of  $\pm 180^\circ$  at 3 dB fractional bandwidth of 5% on  $49^\circ\text{YX LiNbO}_3$  [4]. Ring filters consisted of input/output IDTs placed in parallel acoustic tracks and 2 reflective multistrip couplers, which provided the SAW transmission between these tracks. The filters did not required matching networks because they provided the specified real input/output impedances in the passband by self-matching. The hybrid module with VCO contained the SAW filter, amplifier, phase shifter, buffer stages. The amplifier of VCO contained the single bipolar transistor with the cutoff frequency of 7.5 GHz. The optimization of the amplifier-phase shifter-SAW filter system was provided for achieving the maximal variable frequency range with specified maximal control voltage of 6 volts, current consumption of 5 mA and the carrier to noise ratio C/N no less than 70 dB (offset frequency from the carrier: 25 kHz, bandwidth: 3 kHz). 185, 480 MHz hybrid VCO using leaky SAW filters on  $49^\circ\text{YX LiNbO}_3$  provided variable frequency range of 5 and 13 MHz (2.7%) when the control voltage varied from 0.5 to 6 volts and C/N was 77-82 dB. The current consumption was not more than 5 mA with the source voltage of 5 volts. Hybrid SAW modules with VCO had small size and were mounted in 22x19.5x5 mm glass-and-metal packages.

## REFERENCES

- [1] H.Sato, S.Morishita. Jap.J.Appl.Phys., v.24, (1985), Suppl. 24-1, pp.118-120.
- [2] K.Yamanouchi et al. Proc. IEEE Ultrason. Symp., 1992, pp.57-60.
- [3] T.E.Parker, G.Montress. IEEE Trans. UFFC, v.35, N3, 1988, pp.342-364.
- [4] S.Dobershtein, V.Malyukhov. Proc.IEEE Freq. Contr. Symp., 1994, pp.330-336.

## FREQUENCY AND TIME DISSEMINATION USING SPREAD-SPECTRUM CODED SIGNALS TRANSMITTED VIA DIRECT-TV BROADCASTING SATELLITES

Gerhard Bethscheider\* , Jos Giannandrea\* , Wolfgang Schäfer\*\*

\*SES, Société Européenne des Satellites, Château de Betzdorf, Luxembourg

\*\*TimeTech GmbH, Nobelstr. 15, D-70569 Stuttgart, Germany

Wide-band spread spectrum-coded signals are routinely used for satellite tracking and positioning purposes. These signals are transmitted using very low power levels simultaneously to the main digital or analogue direct-TV signal without any interference. The round-trip time-delay is measured, which provides the range at precisions of better than 5 cm. The line-of-sight velocity is determined at the same time with accuracies of better than 0.1 mm/sec.

These ranging-signals have been received with independent time-receivers located in the footprint area of the satellite. The locally re-generated spread-spectrum symbol-rate stabilizes a low-phase-noise crystal oscillator, whose frequency is corrected for the doppler shift induced by the satellite's relative movement. Orbital information and measured velocity is continuously broadcast from the ranging site to the user with minimum time delay in the form of user data modulating the spread-spectrum signal.

Frequency and time-transfer accuracies achieved so far are presented. The system shows a potential frequency transfer accuracy compatible with good H-maser clocks for integration times of about one day and more.

Further enhancements to the system are investigated to provide a reliable, highly accurate frequency and time reference.

### REFERENCE

Michael Dienert, Claus Baumer: HIGH ACCURACY TWO-WAY TIME TRANSFER AND RANGING EXPERIMENTS USING SATRE MODEMS. 6th EFTF, 1993, Neuchâtel, Switzerland

ABSTRACT

## TIME TRANSFER USING TV SATELLITE TC2A

F. Meyer, M. Vincent

Observatoire de Besançon,  
41 bis Avenue de l'Observatoire  
25010 Besançon Cedex

TV signals from direct TV satellites have been used for time transfer in various experiments. The results of a previous experiment [1] with satellite TDF2 let us expect this technique to allow remote atomic clock comparisons on continental (up to 1000 km) links, at an accuracy level of 10 ns.

We show in this paper that the only obstacle to better results (5 ns or less) is the control of the equipment internal delays.

This paper presents the results of a 6 week measurement campaign using slightly modified equipment and satellite TC2A, for which accurate ephemeris were available.

A theoretical estimation of the accuracy of the pseudo-ranging method that was applied [1] is presented, and appear to be consistent with both (TDF2 and TC2A) experimental results.

Experimental assessment of the accuracy has been performed using accurate GPS common view data from the 4 stations. Biases that were observed during the TDF2 experiment [1] have been significantly reduced after precise antenna coordinate determination, and after Sagnac effect was taken into account. As no re-calibration was performed after modification of the equipment, residual biases (ranging between 26 and 62 ns) were observed in the time transfer results.

A new measurement session is currently being set up, including re-calibration of the 4 stations, with the final aim to obtain a consistency within a few ns with accurate GPS common-view data.

## Reference :

- [1] F. Meyer, M. Granveaud, M. Laporte, F. Vernotte, M. Vincent : "Time transfer using a geostationary direct TV satellite", Proceedings of the 7th EFTF Neuchâtel 16-18 Mars 1993, pp 129-134.

## ONE-WAY FREQUENCY AND TIME TRANSFER IN THE FRAME OF THE INTELSAT TWO-WAY EXPERIMENT

G. Brida, F. Cordara, V. Pettiti, A. Godone

Istituto Elettrotecnico Nazionale Galileo Ferraris  
Strada delle Cacce 91, 10135 Torino, Italy

### Abstract

During 1995, a synchronization experiment in a one-way mode was performed at IEN, using a DBS (Direct Broadcasting Satellite) receiving station, in the frame of the international two-way synchronization experiment based on INTELSAT satellites at 307° East. To stabilize the received frequency at the level required by the 2500A MITREX modem, two approaches were followed: first the Low Noise Converter oscillator was locked by injection via a standard frequency in the Ku-band and successively the satellite beacon frequency has been used as reference to lock an oscillator of the down conversion chain. The instability of the receiving station in both configuration plus the measurement system has been evaluated receiving the PRN codes from different laboratories, after having removed the effect of the satellite movement. Using the second equipment setup, one-way synchronization measurements were regularly performed since October 1995, according to the schedule for the European Laboratories, during their range sessions. Correcting the measurement results for the satellite position obtained from the range data of at least three European Laboratories, the time and frequency transfer capabilities in the one-way mode, for such equipment and stations configuration, have been evaluated.

### 1 - Introduction

The one-way synchronization technique, based on geostationary satellites, has been widely experimented, and the obtainable results have been reported in the literature [1, 2, 3].

In most cases, television frame synchronizing pulses have been used to obtain the time differences between the clocks compared. The main problem to be solved in this kind of time comparisons, is the correction of the synchronization results for the effects of the satellite movement. This has been obtained in different way by the experimenters: i) determining some orbital parameters of the satellite from range measurements, ii) determining the instantaneous position of the satellite using a second synchronization system, iii) removing the effect of the periodic satellite motion by averaging the data over one sidereal day, and the effect of the long-

term longitudinal drift by calibrating one link with another synchronization system (e.g. GPS).

These problems can be overcome using the two way-satellite synchronization technique that almost eliminates the errors contributions due to the satellite movement. This technique, using pseudo random noise (PRN) coded signals that carry the local 1PPS reference information, supplied by special modems (e.g. MITREX), has proved to allow time comparisons over very long baselines with precisions lying in the subnanosecond region and accuracy at the nanosecond level [4].

An international experiment of time and frequency transfer (TWSTFT) on INTELSAT satellites at 307°E between eight laboratories, using this technique, was regularly performed three days a week along 1994 and 1995 with both range and synchronization sessions [5]. Being available at IEN a Ku-band satellite receiving station already used in past experiments and a MITREX modem, it was decided to join this synchronization link in a one-way mode following the schedule agreed for the European area. The participating laboratories in Europe were the Forschung und Technologiezentrum, Darmstadt (FTZ), the National Physical Laboratory, Teddington (NPL), the Physikalisch - Technische Bundesanstalt, Braunschweig (PTB), the Observatoire de la Côte d'Azur, Grasse (OCA), the Technische Universität, Graz (TUG) and the NMI-Van Swinderen Laboratorium, Delft (VSL). Other laboratories in the USA, namely the National Institute for Standards and Technology, Boulder (NIST) and the U.S. Naval Observatory, Washington (USNO), took also part to this experiment. The distribution of the European earth stations is shown in Fig. 1.

The time transfer sessions were performed weekly by these laboratories on Mondays, Wednesdays and Fridays respectively and the first 5 minutes of each session were devoted to range measurements. Since October 20 up to December 22 1995, the PRN codes transmitted by the different European laboratories during these range sessions have been received - once at a time - to test the time and frequency transfer capabilities of such a measurement system, correcting the time measurement performed at IEN for the satellite position obtained from three sets of range data supplied by the other laboratories. In such a way, the need of a second synchronization system to correct the time differences for the effects of the satellite movement is eliminated.

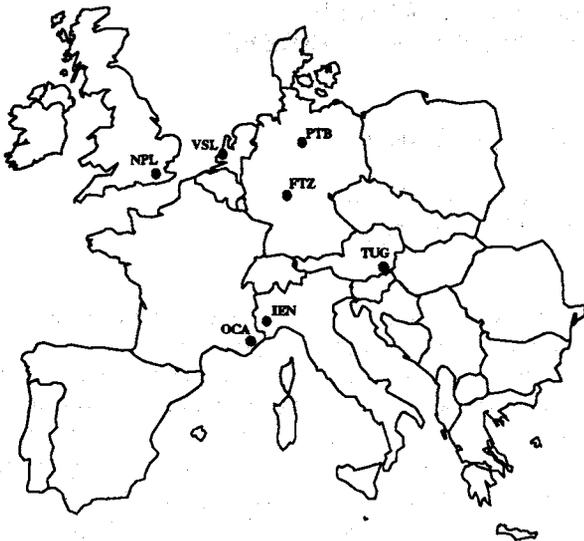


Fig. 1 - Location of the European laboratories

Measurement sessions had been previously performed in June and July 1995 to characterize the frequency instability of the TV Low Noise Converter (LNC) used, of the satellite beacon and of the carrier frequencies transmitted in Europe and in USA. This work was done in order to adapt the existing satellite station to the MITREX requirements [6].

## 2 - The receiving station

The satellite receiving station used at IEN is composed of a 3 m diameter paraboloidal dish, a commercial TV LNC, a second frequency conversion group, a MITREX modem and a time interval counter. The receiver down-converter (LNC) is equipped with a low-noise preamplifier (noise figure lower than 1 dB) installed directly behind the antenna feed. The gain-to-equivalent noise temperature ratio ( $G/T_e$ ) figure of merit of the receiving station is 28 dB/K.

Since it was verified that the LNC free-running ceramic resonator oscillator (CRO) could not meet the frequency stability characteristics required by the MITREX, it showed in fact a frequency variation of some tens of kilohertz over two hours time, as a first solution for the one-way satellite time transfer experiments we used as reference the signal of a microwave synthesizer to which the CRO was locked by injection. An indoor second conversion unit was necessary to reduce the output frequency conversion of the LNC (around 1

GHz) to the 70 MHz center-frequency of the modem input channel. With this set-up, the carrier-to-noise power density ratio ( $C/N_0$ ) with respect to the stations participating in the INTELSAT synchronization experiment, was typically between 60 and 65 dB ( $BW = 1$  Hz). The satellite beacon and carrier frequencies, as measured on June 21 and 26, were found respectively equal to:

$$\nu_b = (12\,500\,975\,100 \pm 20) \text{ Hz}$$

$$\nu_{\text{Europe}} = (12\,543\,410\,220 \pm 30) \text{ Hz}$$

$$\nu_{\text{USA}} = (12\,644\,834\,850 \pm 150) \text{ Hz}$$

To avoid the use of an expensive microwave synthesizer outdoor, a second set-up of the receiving chain was tested. According to the scheme of Fig. 2, the oscillator of the second conversion step in the receiver chain was phase-locked to the nominal down-converted frequency of the beacon satellite signal  $\nu_s$  in such a way to compensate for LNC oscillator frequency drift and to send the desired spread spectrum signal to the modem input. Appendix A describes the mathematical analysis of this phase-lock system. The intensity and the frequency stability of the received beacon signal was verified using the first receiving set-up and it was found that, with a carrier-to-noise power density ratio ( $C/N_0$ ) for the received beacon satellite signal greater than 50 dB ( $BW = 1$  Hz), the new configuration did not give any problem in performing the time synchronization experiments. The reduction of the received beacon signal intensity, due to unfavourable weather conditions (rain, fog, snow), can produce lost of the lock condition and consequently of some synchronization data. The satellite motion creates a frequency Doppler shift both of the beacon and of the spread spectrum signal but this fact has a negligible effect on the frequency converted signal sent to the MITREX modem input.

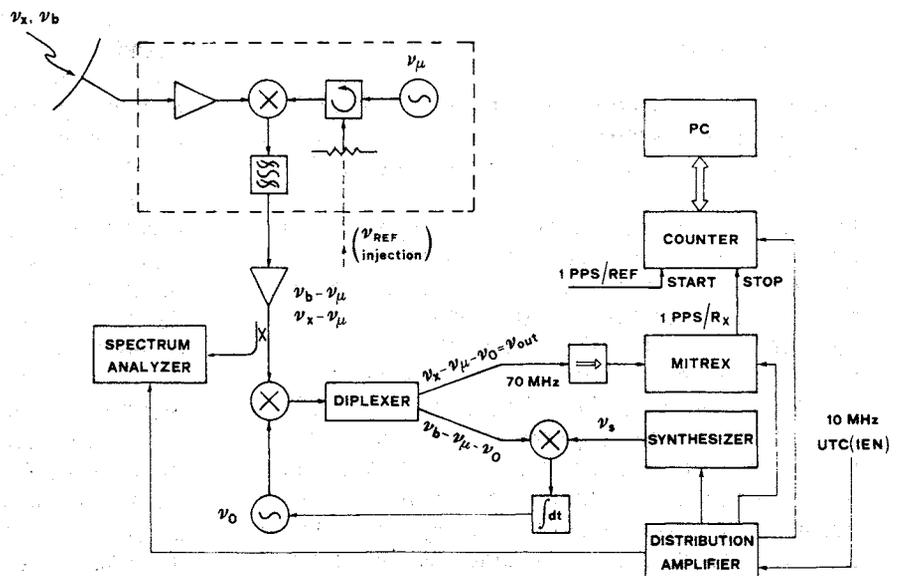


Fig. 2 - The IEN receiving station

### 3 - Characterization of the one-way link

#### 3.1 - Basic concepts

The principle of the one-way synchronization technique used in this experience can be summarized as follows. One station A, operating in the two-way mode, transmits a PRN code synchronized with its reference clock to a geostationary satellite and receives its signals back after a time  $T_A$ , read in its reference time scale. A second station B, operating in a one-way mode, measures the time of arrival of the signals from station A at a time  $T_B$ , relative to its reference time scale. The difference between the time scales of the two laboratories is given by:

$$\Delta T_{AB} = (T_A - T_B) - \frac{r_A - r_B}{c} - (\tau_A^{RX} - \tau_B^{RX}) - (\tau_A^{IO} - \tau_B^{IO}) - (\tau_A^{TR} - \tau_B^{TR}) - (\tau_A^{RE} - \tau_B^{RE}) \quad (1)$$

where,  $(r_A - r_B)/c$  is the differential propagation delay of the signals,  $(\tau_A^{RX} - \tau_B^{RX})$  the differential delay of the receiving stations,  $(\tau_A^{IO} - \tau_B^{IO})$  and  $(\tau_A^{TR} - \tau_B^{TR})$  are the differential corrections for the ionosphere and the troposphere along the propagation paths, and finally  $(\tau_A^{RE} - \tau_B^{RE})$  the relativistic correction for the non reciprocity of the path due to the Earth rotation (Sagnac effect). The first term in the right-side of expression (1) is obtained from the difference of the counters readings at the stations A and B. To evaluate the second term, it is necessary to know precisely the satellite and the stations coordinates in the same reference frame. The third term can be evaluated by measuring the delays of the receiving parts of the stations using a satellite simulator [7]. The ionospheric corrections term can be neglected because at the frequency of 12 GHz transmitted by the satellite used in our experiment, it is at the nanosecond level. Different is the case of the tropospheric correction that shall be computed according to a standard model and is very sensitive to the antenna elevation. The relativistic term can be computed for both stations knowing the stations and satellite coordinates.

#### 3.2 - Evaluation of the station performance

The measurement set-up used at IEN for the synchronization experiment is that of Fig. 2. In addition to the items described in par. 2, a time interval counter with 25 ps resolution and a model 2500A MITREX modem have been used. Several measurements were performed in June and October 1995 on the PRN codes received from FTZ, NIST, NPL, PTB, TUG, USNO and VSL lasting from 300 s up to 12 000 s. From each set

of measurement results, the effect of the satellite movement has been removed by means of a polynomial regression. As a sample of the receiving system performances, in Fig. 3 are reported frequency and time instability estimators  $\text{Mod } \sigma_y(\tau)$  and  $\sigma_x(\tau)$  for integration times between 1s and 2000 s for the link IEN-NPL of October 20, 1995. In this particular case, a fifth order regression has been computed to remove the residual contribution due to the periodical satellite movement. It can be seen that for  $1\text{ s} < \tau < 30\text{ s}$  the noise process has a white phase spectrum with a  $\text{Mod } \sigma_y(\tau)$  equal to  $4 \cdot 10^{-10} \cdot \tau^{-3/2}$  in agreement with the estimation reported in the Appendix.

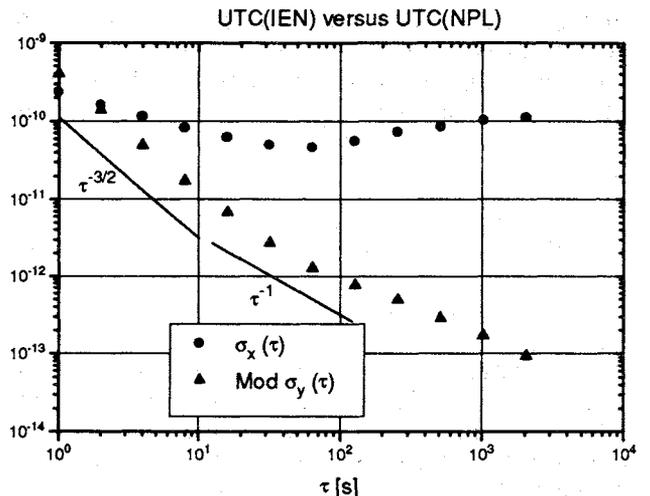


Fig. 3 - IEN/NPL synchronization link instabilities

#### 4 - Experimental results

The determination of the instantaneous satellite position, necessary to get synchronization results from the one-way data, has been performed from the range measurements data supplied by NPL, PTB and TUG because they were available for the whole period of interest and offered the best geometry among the earth stations involved. The range data collected at each laboratory can be expressed as:

$$T_i = \frac{2r_i}{c} + \tau_s + (\tau_i^{TX} + \tau_i^{RX}) + (\tau_i^{IO} + \tau_i^{TR}) \quad (2)$$

where  $2r_i/c$  is the time of flight of the PRN signals,  $\tau_s$  is the satellite transponder delay, and the terms that follow are respectively the total station delay, the ionospheric and tropospheric corrections. The relativistic correction reported in (1) in this case cancels out. Corrections have been only computed for the tropospheric item according to STANAG 4294 model [8] without taking into account the local meteorological parameters, meanwhile the global delays of the stations have been supplied by the laboratories. The ionospheric term has been neglected and no correction for the

Table 1 - Earth station parameters and corrections

Laboratory and Ref. systems	Coordinates				Delays			
	Lat. N [degree]	Long. E [degree]	Height [m]	Elevation [degree]	TX [ns]	RX [ns]	Tropo [ns]	Sagnac [ns]
IEN (WSG 84)	45 00 53,61	7 38 20,05	297,0	11,8	---	---	37,7	134,7
FTZ (WSG 84)	49 52 05,36	8 37 29,30	204,1	9,3	---	---	47,7	124,0
NPL (WSG 84)	51 25 15,82	359 39 37,4	70,2	13,8	429,8	976,9	33,5	108,4
PTB (WSG 84)	52 17 49,79	10 27 37,9	143,4	7,3	1993,7	2512,5	60,0	119,6
TUG (ITRF 88)	47 04 01,58	15 29 36,57	538,1	5,8	717,7	1287,1	68,9	138,5
VSL (ITRF 88)	51 59 59,13	4 22 51,55	65,2	10,9	106,8	1515,3	41,9	113,4

transponder delay has been applied because was not available. In the time of flight term,  $r_i$  is dependent on the differences between the satellite  $(x, y, z)$  and stations  $(x_i, y_i, z_i)$  coordinates in the same geocentric reference frame (WSG 84) as follows:

$$r_i = \sqrt{(x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2} \quad (3)$$

As the  $(x_i, y_i, z_i)$  coordinates are well known, the satellite position can be obtained by a system of three equations and can be used in (1) to compute the synchronization results between IEN and the other stations. Of course, being the satellite transponder delay unknown, the position thus obtained has a bias that could be of the order of hundreds of meters but this effects is consistently reduced by the stations geometry. The parameters of the earth stations and the corrections used to compute the satellite position and the synchronization results, with the exception of the corrections for the differences between the UTC(i) and the 1 PPS/Tx(i), are reported in Table 1. The Tx and Rx delays data differ from those reported in [9] by an amount of 4,2 ns to take into account the average distance between the VSL satellite simulator used and the earth station antennas.

The Rx delay of the IEN station has been determined indirectly considering the synchronization results with TUG via GPS common view and via INTELSAT, over a period of about 20 days, and assuming this delay equal to the average offset. The value that has been adopted for IEN is  $(1823 \pm 4)$  ns. Assuming for the Rx section of the MITREX 2500A modem the average delay of  $(1266 \pm 9)$  ns measured for three devices of the same kind, the delay of the IEN receiving equipment is  $(577 \pm 10)$  ns. The daily GPS synchronization results have been obtained as an average of more than 40 common views, corrected for the difference desumed from the BIPM - Circulars T94 and T95 published results. If we had instead considered the results of the differential time corrections for the IEN and TUG GPS receivers published by BIPM in 1994 and

1995 [10] the difference mentioned above should be diminished by 8 ns.

In Fig. 4 we have reported the comparisons between UTC(IEN) and UTC(TUG) time scales (baseline = 649 km) from November 6 to 29, 1995 by the one-way measurements on INTELSAT and by GPS common-view. Each INTELSAT data point has been obtained performing the quadratic fit over the two set of 300 time interval measurements, one second apart, and computing a value for the midpoint. A typical standard deviation value of the residuals of the quadratic fit is in the range from 0,3 ns to 1 ns. The standard deviation of the residuals between GPS and INTELSAT synchronization data is  $\sigma = 4$  ns. The time interval readings have been corrected for the differences UTC(i) - 1 PPS/Ref supplied by the laboratories.

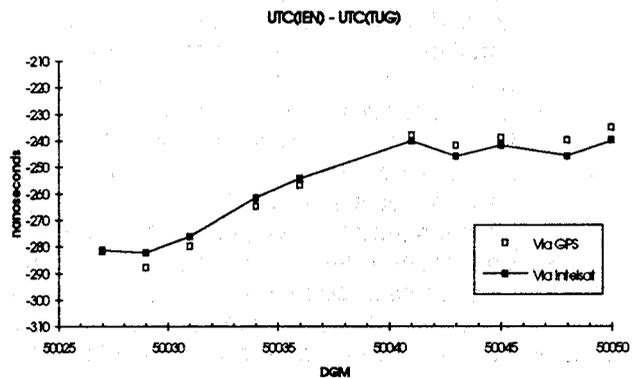


Fig. 4 - IEN/TUG time scale comparisons

Over these one-way results and over one 300 data sample, the modified Allan deviation  $\text{Mod } \sigma_y(\tau)$  and the time deviation  $\sigma_x(\tau)$  have been computed and the results are shown in Fig. 5. In the case of the 300 s sample, the satellite position has been determined every 30 seconds.

As Fig. 5 shows, for  $\tau < 20$  s the  $\text{Mod } \sigma_y(\tau)$  is equal to  $1 \cdot 10^{-9} \cdot \tau^{-3/2}$  (white phase noise) meanwhile for longer observation times the noise process changes into flicker phase. The time and

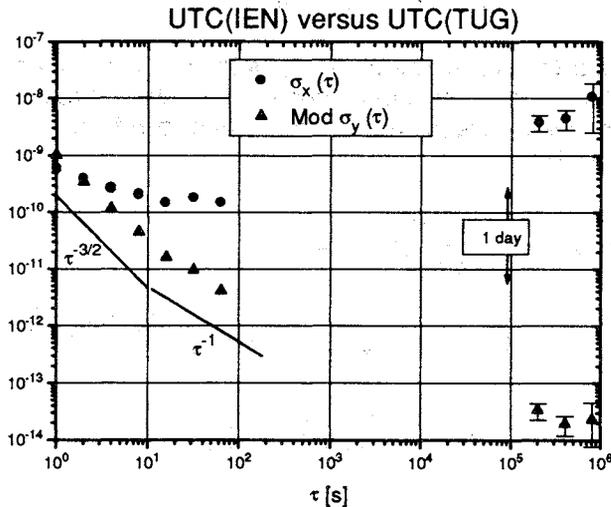


Fig. 5 - IEN/TUG synchronization link instabilities

frequency instabilities for  $\tau > 2$  days have been computed assuming a constant interval between measurements of 2,3 days even if the measurement data are unevenly spaced.

The frequency transfer limit of the one-way system used is of the order of  $3 \cdot 10^{-14}$  for  $\tau > 2$  days. Comparing the results in Fig. 5 with the commercial cesium clocks specifications, it appears that the system features are adequate for their characterization, for measurement time higher than 100 s.

With the same criteria followed in the previous case, from December 1 to 18, 1995 the differences between UTC(IEN) and UTC(NPL) (baseline = 925 km) have been computed from the one-way results, and these have been compared with the data obtained by GPS reported in BIPM Circular T96. The results, reported in Fig. 6, show a larger bias than in the TUG case that can be partly attributed to the worst stations geometry.

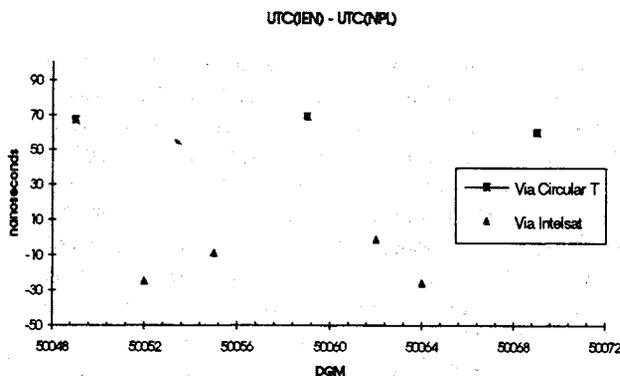


Fig. 6 - IEN/NPL time scale comparisons

### 5 - Conclusions

In the present work, the adaptability of a DBS receiving station to MITREX modem requirements has been demonstrated; the capabilities of the one-way synchronization method with such equipment, as regards the frequency and time transfer, have been tested with some European laboratories in the frame of INTELSAT experiments. Correcting the time measurements for the effect of the satellite movement, using the range data from three laboratories, a frequency transfer capability at the  $3 \cdot 10^{-14}$  level or better for observation times greater than two days has been demonstrated. These results compare very favourably with those obtainable with GPS and are very promising for the establishment of a secondary level network of receiving-only stations integrated in an international TWSTFT system.

Further improvements could be obtained if two-way earth stations with a better geometry become operative, by using a satellite offering higher elevation angles to the ranging stations, adopting a redundant system of equations for the satellite position determination that should also allow for the evaluation of the satellite transponder delay. Another possibility could be offered by the computation of the satellite orbital parameters using a propagation orbit software.

### 6 - Acknowledgments

The authors wish to thank D. Kirchner from TUG, J. A. Davis from NPL and G. de Jong from VSL for the very helpful cooperation. Indeed, the supply of the range data by the European Laboratories (FTZ, NPL, PTB, TUG and VSL) participating in the INTELSAT two-way experiment, has been instrumental to the present work as it has been the loan of the MITREX modem by the Istituto Superiore delle Poste e Telecomunicazioni - Italy.

### Appendix A

The schematic diagram of figure A.1 describes the essential elements of the receiving system: the two-steps conversion and the phase lock loop (PLL). The time synchronization signal  $x$  and the satellite beacon signal  $b$  are down-converted with the LNC from the Ku band to the SHF band, and then, in the indoor unit, from the SHF band to the VHF band. From the analysis of the block diagram of Fig. A.1 we can derive the two equations:

$$\begin{cases} v_x - v_\mu - v_0 = v_{OUT} \\ v_b - v_\mu - v_0 = v_s \end{cases} \quad (A.1)$$

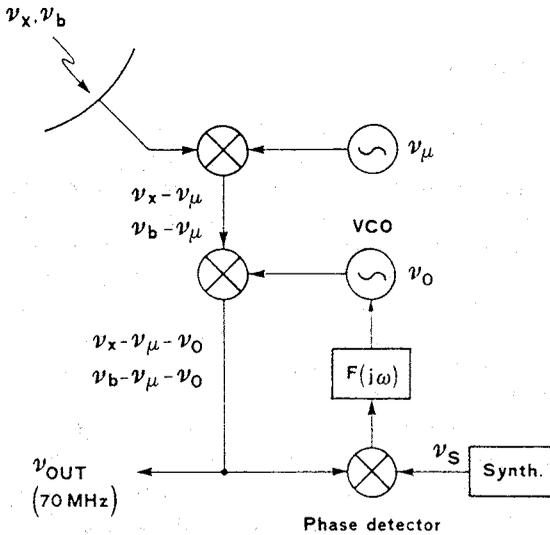


Fig. A.1 - Details of the conversion chain

where  $\nu_x$  is the center frequency of the x signal,  $\nu_b$  the satellite beacon frequency,  $\nu_{OUT}$  the IF input frequency of the MITREX modem at which we must convert the received signal,  $\nu_\mu$  the nominal frequency of the LNC oscillator,  $\nu_0$  the frequency of the Voltage Controlled Oscillator (VCO) used in the second conversion step unit and  $\nu_s$  the frequency of the synthesizer. From this set of equations we derive the nominal frequency values of  $\nu_0$  and  $\nu_s$  where  $\nu_b$ ,  $\nu_x$  and  $\nu_\mu$  values are assumed to be known:

$$\begin{cases} \nu_s = \nu_{OUT} - (\nu_x - \nu_b) \\ \nu_0 = \nu_x - \nu_\mu - \nu_{OUT} \end{cases} \quad (A.2)$$

Analyzing the phase lock loop we can write, in the domain of the Fourier angular frequency  $\omega$ , the equation:

$$\Omega_0 = \Omega_{0,free} + \frac{k_0 k_D}{j\omega} F(j\omega) (\Phi_s - \tilde{\Phi}_b) \quad (A.3)$$

where  $k_0$  is the VCO gain factor,  $k_D$  the phase-detector gain factor,  $F(j\omega)$  the filter transfer function of the PLL,  $\Omega_{0,free}$  the free-running angular frequency of the VCO,  $\Omega_0$  the controlled output angular frequency of the VCO,  $\nu_s$  the synthesizer output phase and  $\tilde{\Phi}_b$  the phase of the down-converted beacon signal, with  $\tilde{\Phi}_b = \Phi_b - \Phi_\mu - \Phi_0$  the right-side terms being respectively the phases of the beacon signal, of the LNC oscillator and of the VCO. To derive equation (A.3) we used the relationship  $\Omega = j\omega\Phi$  between phase and angular frequency in the Fourier space domain. The phase of the x signal down-converted and sent to the modem input is:

$$\tilde{\Phi}_{OUT} = \Phi_x - \Phi_\mu - \Phi_0 \quad (A.4)$$

Substituting equation (A.3) into (A.4) we can derive the spectral density of the relative frequency departure of the signal sent to the MITREX:

$$\begin{aligned} S_{y,OUT}(\omega) = & \left( \frac{\nu_x}{\nu_{OUT}} \right)^2 [S_{y,x} + S_{y,n}] + \\ & + \left| 1 - H(j\omega) \right|^2 \left[ \left( \frac{\nu_0}{\nu_{OUT}} \right)^2 S_{y,0,free} + \left( \frac{\nu_\mu}{\nu_{OUT}} \right)^2 S_{y,\mu} \right] + \\ & + \left| H(j\omega) \right|^2 \left[ \left( \frac{\nu_b}{\nu_{OUT}} \right)^2 S_{y,b} + \left( \frac{\nu_s}{\nu_{OUT}} \right)^2 S_{y,s} \right] \end{aligned} \quad (A.5)$$

with  $H(j\omega) = \frac{k_0 k_D F(j\omega)}{j\omega + k_0 k_D F(j\omega)}$  being the closed

loop transfer function; the subscripts on spectral density of the relative frequency departures terms identify the source. The term  $S_{y,n}$  indicates the additive noise contributions on the channel transmission path. The phase noise contribution  $S_{y,s}$  of a good frequency synthesizer is negligible. The design of a sufficiently large PLL control bandwidth ( $W_B$ ), with respect to the fast-linewidth of the LNC and VCO oscillators, enables to cancel out the contribution of the phase noise of these two oscillators in (A.5).

Within the control bandwidth of the PLL ( $\omega < W_B$ )

$$\left| 1 - H(j\omega) \right|^2 \cong 0 \quad \text{and} \quad \left| H(j\omega) \right|^2 \cong 1 \quad (A.6)$$

so equation (A.5), with the approximation  $\nu_b \cong \nu_x$ , can be reduced into the form:

$$S_{y,OUT} \cong \left( \frac{\nu_x}{\nu_{OUT}} \right)^2 [S_{y,x} + S_{y,n} + S_{y,b}] \quad (A.7)$$

The minimum instability level of the measurement system is determined by the relative frequency departure spectral densities of both the additive noise  $S_{y,n}$  and the beacon signal  $S_{y,b}$ . With a typical signal-to-noise ratio of 45 dB ( $BW = 1$  Hz) for the received spread spectrum signal, in the hypothesis of a white phase noise, an equivalent modem bandwidth of 1 kHz and a negligible contribution of the beacon signal, we obtain a relative frequency departure spectral density  $S_{y,OUT}(f) = 5 \cdot 10^{-25} \cdot f^2$ . The corresponding Modified Allan deviation is  $\text{Mod } \sigma_y(\tau) = 7 \cdot 10^{-10} \cdot \tau^{-1/2}$ . This limiting value is in good agreement with our best experimental results and this confirms the assumption made of the negligibility of  $S_{y,b}$  contribution.

## References

- [1] F. Cordara et al., 1991, Time and frequency comparisons in Europe by means of ECS geostationary satellite, Proc. 22nd PTTI Meeting, 117-130.
- [2] F. Meyer, 1995, One-way time transfer using geostationary satellite TDF2, IEEE Trans. IM, 44, 103-106.
- [3] I. Kardos et al., 1994, Status Report on geostationary satellite positioning by time measurements, Proc. of the Workshop "Accurate Orbit determination and observations of high earth satellites for geodynamics", 11-22.
- [4] D. Kirchner, 1991, Two-way time transfer via communication satellites, Proc. IEEE, 79, 983-990.
- [5] J. Davis et al., 1995, European two-way satellite time transfer experiments using the INTELSAT (VA-F13) satellite at 307 °E, IEEE Trans. IM, 44, 90-93.
- [6] V. Pettiti et al., 1994, Two-way time comparisons via Olympus satellite, Proc. 8th EFTE, 315-328.
- [7] G. de Jong and M.C. Polderman, 1994, Automated delay measurement system for an earth station for two-way satellite time and frequency transfer, Proc. 26th PTTI Meeting, 305-318.
- [8] NATO Standardization Agreement (STANAG) Doc. 4294, Appendix 6 to Annex A, Draft issue 0, pp. A-6-34 to A-6-37, 1991.
- [9] G. de Jong et al., 1995, Results of the Calibration of the delays of earth stations for TWSTFT using the VSL satellite simulator method, to appear on Proc. 27th PTTI Meeting.
- [10] W. Lewandowski et al.' 1995, Rapports BIPM-94/12 and BIPM-95/07.

# TIME TRANSFER WITH GEODETIC GPS RECEIVERS USING CODE AND PHASE OBSERVATIONS

P. Baeriswyl, Th. Schildknecht, T. Springer, G. Beutler

Astronomical Institute  
University of Berne  
Switzerland

## ABSTRACT

Since a couple of years different methods have been developed to transfer time and frequency using the Global Positioning System (GPS). The techniques depend strongly on the institution developing it and/or on the required accuracy. The comparison of high performance frequency standards ( $10^{-14}$ ) asks for a method where frequency transfer can be realised after a reasonable time of observation. Using only GPS-code observations a precision of a few nanoseconds can be achieved over long baselines after one day of observation.

Geodetic applications ask for receivers which are able to acquire all observables of all satellites in view. Because carrier-phase measurements provide a much higher precision than code-observations, it is possible to obtain an accuracy of a few millimetres when processing these data in an interferometric mode. Interpreting this number in the time domain leads to a theoretical accuracy of the time transfer data of some 10 ps.

Two time transfer terminals, each equipped with a geodetic GPS receiver, were built to transfer time between two high-performance clocks. Because we are aiming at sub-nanosecond accuracy it is necessary to take into account influences of a variety of environment variations on the electronics involved in these systems.

We present results of our test series which include temperature dependencies of receiver, receiver antenna, and the electronics involved. The tests were performed on a „zero-baseline“ and a 5km-baseline, respectively. We show that even station clocks of 'poor' quality can be compared with high precision.

## INTRODUCTION

Two time transfer terminals, each equipped with a modified Ashtech Z12 GPS receiver, were built at the Astronomical Institute of the University of Berne (AIUB). The GPS receivers acquire all observables available from all satellites in view, i.e. the C/A-code, both P-codes (P1 and P2), and both carrier-frequencies (L1 and L2).

By processing the data in an interferometric mode with a geodetic software we aim at improving the time and

frequency transfer with respect to current GPS common view techniques. For our project we use *the Bernese GPS Software*, Rothacher et al. (1), which was modified for time transfer purposes.

In order to achieve the mentioned precision of a few millimetres we use precise orbits stemming from the International GPS Service for Geodynamics (IGS), Beutler et al. (2). Preliminary orbits are made available for public use 1.5 days after the observation and have a precision of better than 20 cm. Rapid orbits with a precision of about 10 cm are made available 11 days after the observation.

If we do not use these orbits we run the risk to introduce an error due to the errors in the broadcast orbits, especially under SA. In this case it would not be possible to transfer time over a distance of 500 km with a precision below 1 ns, Bauersima (3).

We assume that in addition to the precise orbits the troposphere delay corrections are also given by the IGS. Because we acquire all observables, ionospheric range corrections may be determined by using both carrier frequencies.

The principle of data processing is to form single differences of all observables between two receivers and to smooth the noisy code (rms of a few nanoseconds) by means of the accurate phase measurements. The code observations determine the initial clock offset and the phase data the drift.

Because the mean error of the phase observation may be neglected compared to the mean error of the code observation, the mean error of the smoothed code goes with  $1/\sqrt{n}$ , where  $n$  is the number of observations. For a detailed discussion see Schildknecht et al. (4).

It is obvious that with our method, the AIUB-method, we have a high level of redundancy which results in a high level of stability and robustness.

## BASIC SETUP AND QUESTIONS

**Basic setup:** A simplified scheme for the time transfer terminal is shown in Figure 1:

- The GPS receiver is driven by the station clock. The

receiver internal reference frequency base is replaced by an external reference (see (5)).

- A time interval counter measures receiver delays (for monitoring and calibration purposes only).
- The temperature of receiver and antenna is acquired and may be stabilised.
- The PC is the controlling unit of the time transfer terminal. It acquires the temperatures and time intervals and transfers the receiver data via internet to a central computer. This link allows to control the time transfer terminal from a remote computer.
- On the central computer all the GPS data are processed with the Bernese GPS Software. This processing is done routinely once per day.

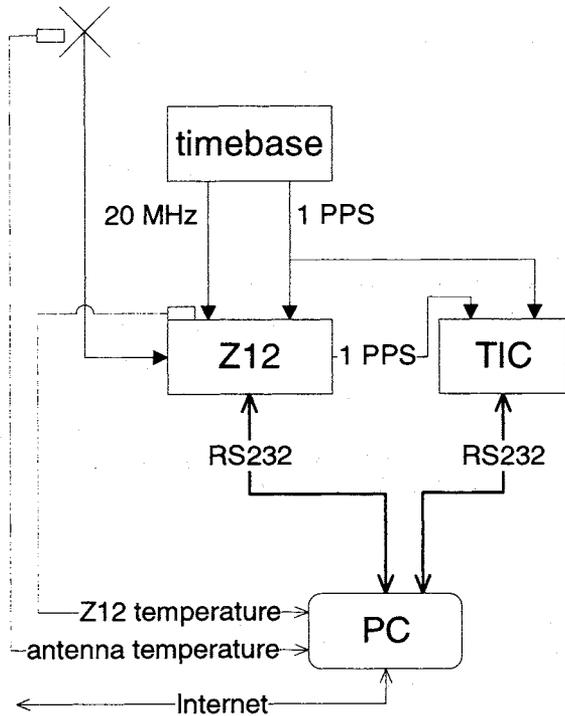


Figure 1: Simplified layout of the *Geodetic GPS time Transfer Terminal*. The temperature of Z12 receiver (including the required electronics) and the antenna may be stabilised and measured.

In this project we were studying the following items:

- The behaviour of the signal path delays and of the time transfer data as a function of the receiver and the antenna temperature, respectively.
- The precision of the synchronisation error on a zero baseline (common antenna) and with a common time base for both receivers.
- The precision of the synchronisation error on a „zero baseline“ (two antennas, 7 m apart) and with a common time base.
- The accuracy of the synchronisation error on a 5 km baseline (compared with an alternative time transfer technique).
- The relation between the synchronisation error computed with different observables.
- The influence of multipath.

- The repeatability of the results.
- The long-term behaviour.

By ‘synchronisation error’ we mean the clock difference between the two time bases which we compare, and by ‘precision of the synchronisation error’ we mean the precision of the time transfer results determined from phase and P-code measurements.

## MEASUREMENTS AND RESULTS

### 7 m baseline

Figure 2 shows the basic experimental setup for this test series. Both receivers are driven by the same time base. With a time interval counter we measure the interval between the receivers’ 1 PPS outputs: The system allows for a maximum rate of one measurement per second. For the tests mentioned, however, a rate of one measurement each 30 seconds was adequate.

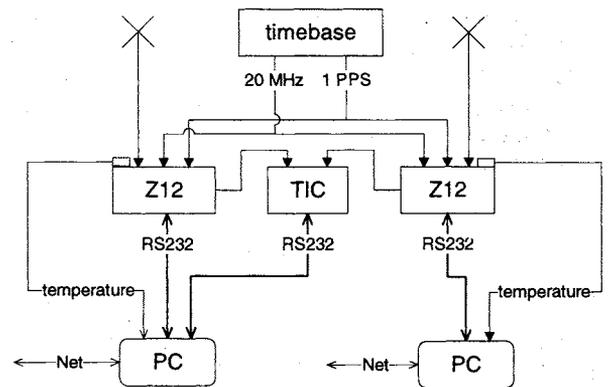


Figure 2: Setup for test series with a common time base. The temperature of both receivers was controlled.

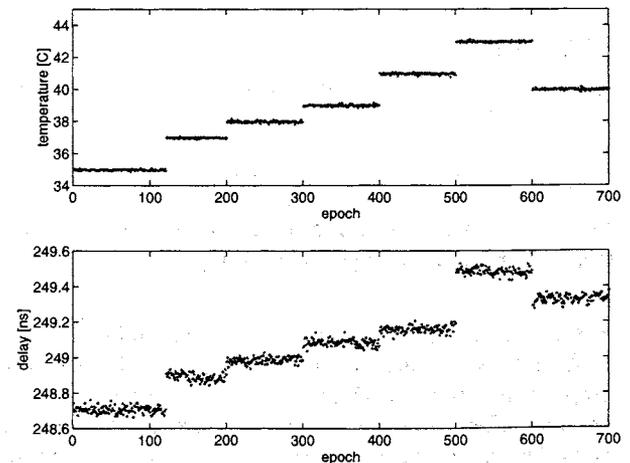


Figure 3: Changes of the receiver internal delays (bottom) with varying receiver temperature (top).

**Z12 signal path delay:** We measured the interval between the 1 PPS output signals at different temperature

levels of one receiver (the other was always kept at the same temperature). Figure 3 shows the influence of the temperature on the signal delays.

The signal delay shows a variation of about 100 ps/K which is not linear over the entire temperature range (compare data between epoch 400 to 500 with data between epoch 600 to 700). The situation is probably complicated by hysteresis effects.

We conclude, however, that the delay is constant with a rms of better than 20 ps as long as the receiver temperature is kept constant within 1 K.

Remark: Figure 3 does not say that all the temperature changes were instantaneous. After increasing the temperature we waited until the temperature was stable again.

**Stability of the synchronisation errors:** Figure 4 shows the synchronisation errors over a 4-day-period. There is no drift between these two data sets. A rms of about 0.8 ns for P1-code and 0.9 ns for P2-code, respectively, was observed.

Figure 5 shows the corresponding result from the phase data. Between day 348.5 and day 349.5 we increased the temperature of one receiver by 4 K. The P-code data are not affected by the temperature change but the synchronisation error computed with the L1 carrier phase changed by 400 ps.

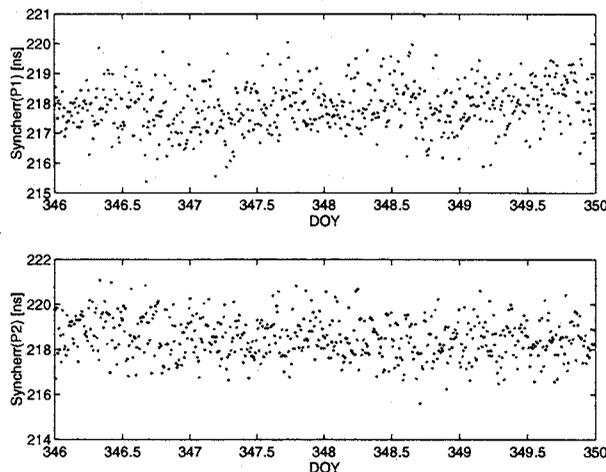


Figure 4: Synchronisation error computed with P1 code (top) and P2 code for a period of 4 days.

The rms of the synchronisation error determined by L1 is about 8 ps (at constant temperature). On L2 the rms is 10 ps. With such a high resolution we are even able to see that after lowering the temperature to its original value the synchronisation error determined by the phase is about 20 ps smaller than at the beginning of the period. This indicates a hysteresis effect of the synchronisation errors.

The difference between the L1 and the L2 synchronisation errors has an rms of about 8 ps even when the re-

ceiver temperature is changing. Therefore we may assume that the relation between these two data sets is a constant offset.

The temperature dependence of the signal path delays as a function of the receiver temperature in Figure 5 suggests that we might also use the external measurements to correct the synchronisation errors instead of keeping the receivers at a constant temperature.

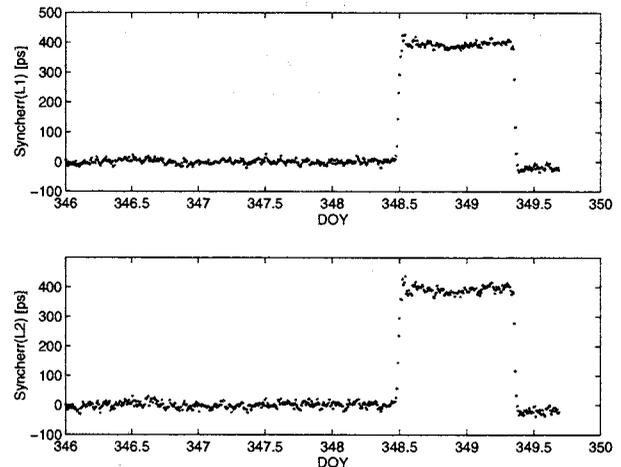


Figure 5: Synchronisation error computed with the L1 (top) and the L2 (bottom) carrier phase. On both data sets an arbitrary offset has been eliminated. The steps are due to increasing and decreasing the receiver temperature by 4 K and -4 K, respectively.

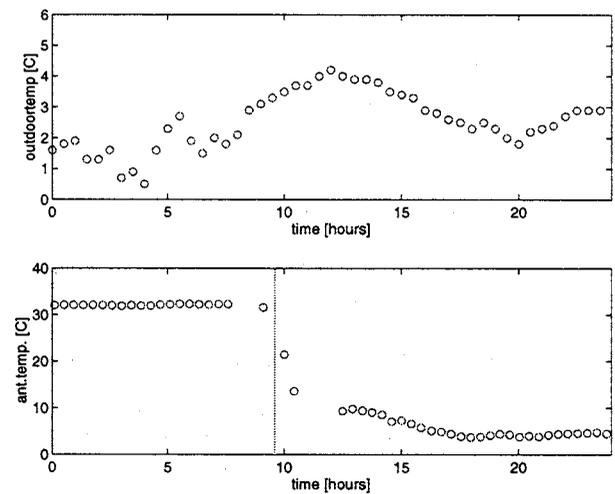


Figure 6: Temperature of the environment (top) and of the antenna (bottom). The vertical line is the moment when the antenna insulation was removed.

**Antenna delays:** If we expect delay variations as a function of the temperature at the receiver we also have to expect delay variations if the antenna temperature is changing. For this purpose we made tests with one antenna temperature stabilised (Figure 6). Until 9:30 a.m. the antenna was stabilised at 32°C. At 9:30 a.m. we removed the insulation box. Although we had a change of the antenna temperature of almost 20 K we could not

observe any change of the synchronisation error computed with the P1-code data (see Figure 7, top).

The second graph of Figure 7 shows the synchronisation error computed with the phase data (L1). An almost sudden change of the synchronisation error can be observed at about 7:30 but this is clearly not due to the temperature change of the antenna. Maybe the change of the slope at 9:30 is due to the decreased antenna temperature but the effect is not alarming.

Comparing these data (Figure 7) with time transfer data of another day (Figure 8) where no antenna was temperature stabilised we realise that the changes of the synchronisation error computed with L1 are related to outdoor temperature changes. The origin of the effect probably resides in the different amount of cable exposed to outdoor temperature for the two antennas. One antenna had about 10 m of cable outside the building (on the roof) whereas the other only 2 m. We believe to see in Figure 8 the differential behaviour of the two antenna cables as a function of the outdoor temperature.

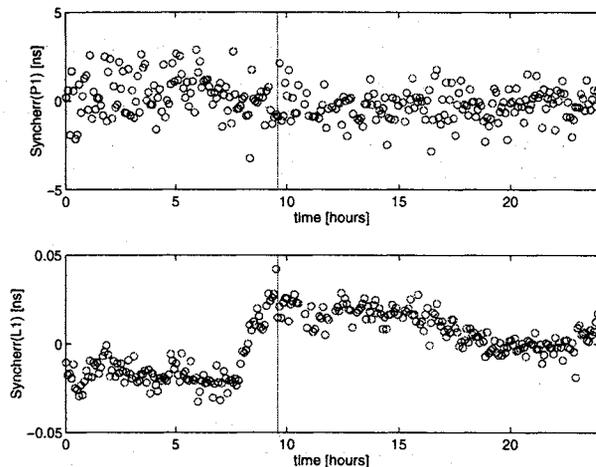


Figure 7: Synchronisation error computed with P1 (top) and L1 (bottom).

The last graph in Figure 8 gives the synchronisation error of L1 as a function of the outdoor temperature. It reveals a linear dependence of about 7 ps/K. If this is really just an effect of the cable means that we have a temperature dependence of the signal delays in the antenna cable of about 1 ps/K/m. This seems to be a small value but if the two sites are separated by a long distance we have to take into account that the temperature variations of the environment on these two sites may not be the same and do not cancel out. If, for instance, one site has 50 m of antenna cable on a roof and we have to expect temperature variations of 50 K over one year (which is not unrealistic) we may have delay variations of about 2.5 ns.

This problem could be solved by installing a second cable of the same type in parallel to the antenna cable and measure the signal path delays periodically in this cable. These measurements could be used to correct the synchronisation errors for the delays in the cable.

## 5 km baseline

**Setup:** It is important in such a project to check whether the results obtained are consistent with those of other methods. For this reason we moved one terminal to the Federal Office of Metrology (OFM) to transfer time between the Zimmerwald observatory and the OFM. On the other hand we installed on both sites a GPS time receiver (TTR5) to compare our results to those obtained with the classical common view technique.

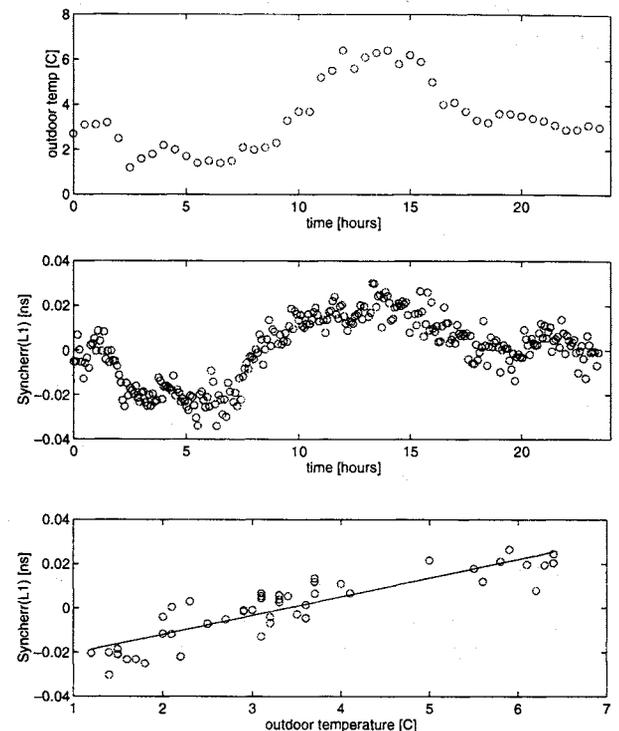


Figure 8: Outdoor temperature (top) and synchronisation error computed with L1 (middle) for a period of one day (antenna not stabilised). Synchronisation error determined with L1 as a function of the outdoor temperature (bottom).

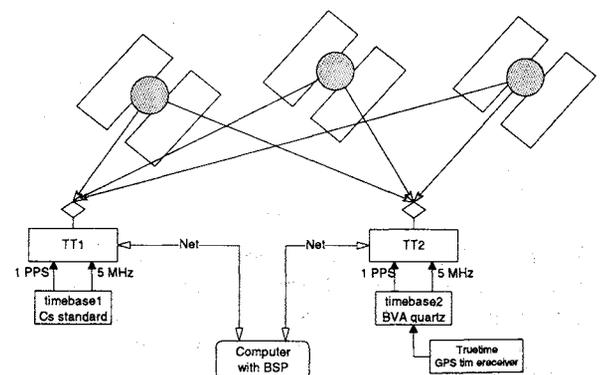


Figure 9: Time transfer between the Swiss Office Federal of Metrology and the observatory in Zimmerwald (ca. 5 km baseline) with the AIUB-method.

The GPS receivers at the OFM were driven by a Cs standard with a stability of about  $1.3 \cdot 10^{-12}$ , in Zimmerwald we used a BVA-quartz. This quartz is disciplined

by a Truetime GPS receiver to GPS time with an accuracy of about 20 ns. The time base in Zimmerwald was therefore a very unstable reference. But our system should be able to deal with clocks of this quality, too.

**Time transfer with AIUB- and Common-View-method:** In Figure 10 we compare the behaviour of the time transfer as realised by the two mentioned methods. The drift is due to the drift of the Cs standard. (with respect to GPS-time). The ripples are caused by the BVA quartz which is synchronised to GPS time by the „Truetime“ unit.

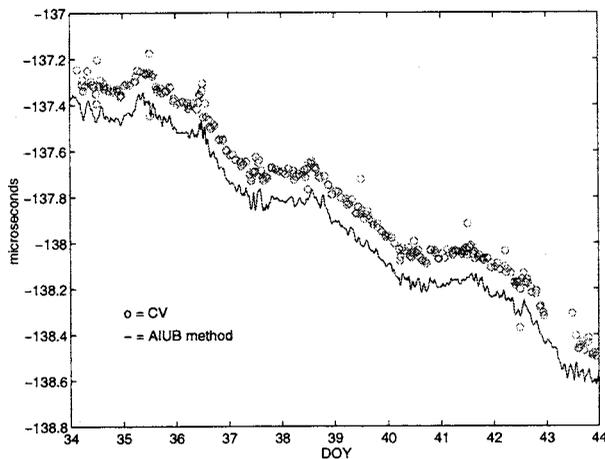


Figure 10: Comparison of the Common-View and AIUB-method (P1-code data) over a period of 10 day.

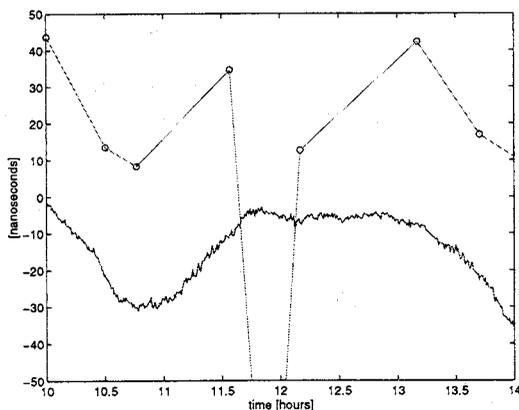


Figure 11: Detail view of Figure 10, day 39. The behaviour of the clocks is well known with the AIUB-method even between the CV observations. For better visibility an arbitrary offset was introduced between CV- and AIUB-data.

Both methods show the same overall behaviour but with an offset between them. This offset may be calibrated, provided it is constant over time.

Figure 11 is a detailed view of a part of Figure 10 with the linear drift eliminated. It illustrates that clock differences between any two epochs may be recovered very

accurately using our method despite of the fluctuations of the reference clock in Zimmerwald.

In both, Figure 10 and Figure 11, we notice some outliers of up to 150 ns in the CV data. A closer look at these data reveals that most of these outliers occur when PRN 2 was used in the CV method. The outliers are clearly due to the CV data. Probably there is a problem with one of the TTR5 receivers.

Even though we computed the synchronisation error only using code observations (the software was not yet adapted to compute the synchronisation error including phase data when such an unstable reference frequency was used) we know that (after smoothing the code data with the phase data) the precision will be well below 100 ps.

**Offset stability between AIUB- and CV-method:** This question may be answered by forming the difference between CV and AIUB results. This difference is given in Figure 12.

A linear fit through these data (outliers > 50 ns removed) results in a slope of about 150 ps/day with a formal error of 120 ps/day. This means that the computed value is consistent with zero or that there is no drift between the two methods. The rms error of the difference is 8 ns. It is mostly due to the CV data.

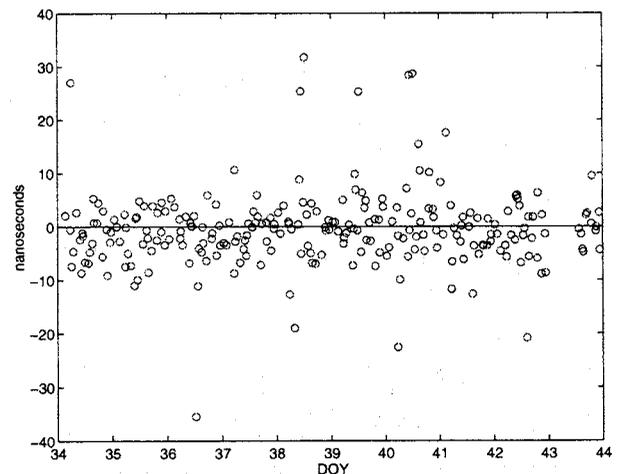


Figure 12: Difference between CV- and AIUB-method (offset eliminated).

## SUMMARY AND CONCLUSIONS

We have made many investigations on a „zero“ and „quasi-zero“ (7 m) baseline with both time transfer terminals installed side by side. Many of these tests were performed first with a common time base for both terminals, later on with separate time bases. The most important goals in these test series were to assess

- the precision achievable with our time transfer method,

- the stability of the synchronisation error computed with the different observables,
- the stability of the signal path delays in the receiver, and
- the influence of antenna and receiver temperature on our time transfer results.

The synchronisation errors on a 7 m baseline computed with the P1 and P2 code data are constant and have a rms of 0.8 ns and 0.9 ns, respectively when both time transfer terminals are driven by the same time base.

Under the same conditions the precision computed with the L1 and L2 carrier phase observation is 8 ps and 10 ps, respectively.

Temperature changes of about 10 K of the receiver do not affect the synchronisation errors computed with the code data. On the other hand the influence of the receiver temperature is about 100 ps / K for the synchronisation errors computed with the phase data and the signal path delays. Therefore if synchronisation errors are affected by temperature variations of the receiver we could correct these results with external measurements of the signal path delays in the receiver.

It seems that temperature changes of the antenna do not influence the synchronisation error computed with the code data. The influence on the carrier phase data seems to be negligible, too.

On the other hand the signal path delays in the antenna cable is about 1 ps/K/m and cannot be neglected if long antenna cables are exposed to big temperature changes.

On a 5-km-baseline we compared our time transfer results to the common view technique. Over a period of ten days we could not see a significant drift between the results of these two methods. In this test we were limited by the precision of the common view technique.

There still is an unexplained offset between the results of the two methods but if we may assume that this offset is constant in time it can be easily taken into account.

In addition to this result we see that even time bases with an unstable behaviour may be compared very accurately using our approach. This requires that the 1 PPS stemming from the time base is synchronised to GPS time to within 1 ms.

Looking at the resolution we obtained on the 5 km baseline (rms  $\cong$  100 ps) a drift of a high performance frequency standard ( $10^{-14}$ ) should be visible after 3 h of observation if receiver and outdoor temperatures are constant within a few degrees. Time transfer techniques with a rms of 2 ns would require at least two days of observation for the same information.

In the near future experiments over long baselines are necessary at sites where an alternative time transfer system is available, i.e. a time transfer system with the same or a better precision than the AIUB-method discussed in this article.

## ACKNOWLEDGMENT

The research presented in this article was funded through a contract issued by the Swiss Federal Office of Metrology. Without this contribution our research work could not have been performed.

## REFERENCE

- (1) Rothacher, M., G. Beutler, W. Gurtner, E. Brockmann, L. Mervart (1993): *Bernese GPS Software, Version 3.4*.
- (2) Beutler G., I.I. Mueller, R.E. Neilan (1996): „The International GPS Service for Geodynamics (IGS): The Story“. In: *GPS Trends in Precise Terrestrial Airborne, and Spaceborne Applications, International Association of Geodesy Symposia*, Symp. No 115, Boulder, CO, July 3-4, 1995, G. Beutler et al., (Eds.) Springer-Verlag, Berlin, Heidelberg, New York, 1996, pp. 3-13.
- (3) Bauersima, I. (1983): „Radiointerferometrische Satellitenbeobachtungen (GPS II).“ *Mitteilung Nr. 10 der Satellitenbeobachtungsstation Zimmerwald*, Astronomisches Institut der Universität Bern, Schweiz, 1983.
- (4) Schildknecht, Th., G. Beutler, W. Gurtner, M. Rothacher (1990): „Towards subnanosecond GPS time transfer using geodetic processing techniques.“ *Proceedings of the 4th European Frequency and Time Forum*, Neuchâtel, 1990, pp. 335-346.
- (5) Baeriswyl, P., Th. Schildknecht, J. Utzinger, G. Beutler (1995): Frequency and Time Transfer with Geodetic GPS Receivers: First Results. *Proceedings of the 9th European Frequency and Time Forum*, Besançon, 1995, pp. 46-51.

# LOW NOISE OSCILLATORS

## Theory and Design

J.K.A. Everard

Department of Electronics

University of York

Heslington, York, YO1-5DD, England

### ABSTRACT

A large number of papers have been published on low noise oscillators, however they are usually very specific to the particular application. This paper will describe a set of design rules which are general and can be used to produce oscillators with very low noise performance where both additive (thermal noise) and Flicker noise are considered. Linear theories will be described which accurately describe the noise performance of resonator type oscillators. The limits set by varactor diodes on the noise performance will be described and the noise degradation caused by open loop phase error will be shown. Seven design examples will be demonstrated which show close correlation with the theory.

### OSCILLATOR NOISE THEORIES

The model chosen to analyse an oscillator is extremely important. It should be simple to enable physical insight and at the same time include all the important parameters. For this reason two models are presented here. Each model can produce different results as well as improving the understanding of the basic model. Both an equivalent circuit model and a block diagram model will be described. We will start with the equivalent circuit model originally used by the author to design oscillators with the potential for high efficiency and easy analysis. These models are used to describe the effects of thermal noise. Flicker noise effects are described later.

### EQUIVALENT CIRCUIT MODEL

The model<sup>1,2</sup> is shown in Figure 1 and consists of an amplifier with two inputs with equal input impedance, one for noise ( $V_{in2}$ ) and one as part of the feedback resonator.

The feedback resonator is modelled as a series inductor capacitor circuit with an equivalent loss resistance  $R_{loss}$  which defines the unloaded Q ( $Q_0$ ) of the resonator as  $\omega/R_{loss}$ .

The operation of the oscillator can best be understood by injecting white noise at the input  $V_{in1}$  and calculating the transfer function while incorporating the usual boundary condition of  $GB_0 = 1$  where G is the limited

gain of the amplifier when the loop is closed and  $B_0$  is the feedback coefficient at resonance where  $f_0 = 1/2\pi\sqrt{LC}$ .

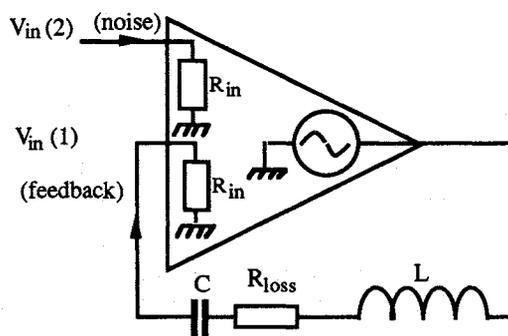


Figure 1. Equivalent circuit model of oscillator

The amplifier model had zero output impedance, a known input impedance and a resonant positive feedback network. The zero output impedance of the amplifier was used because the design of highly power efficient oscillators was of interest. This also reduced the pulling effect of the load. The zero (low) output impedance of the amplifier is achieved by using a switching output stage. In fact the same theory and conclusions can be obtained for oscillators with a finite output impedance so for convenience this will be left at zero.

Input  $V_{in2}$  is used at the input of the amplifier to model the effect of noise. In a practical circuit the noise would come from the amplifier. The noise voltage  $V_{in2}$  is assumed to be added at the input of the amplifier and was dependent on the input impedance of the amplifier, the source resistance presented to the input of the amplifier and the noise figure of the amplifier. In this analysis, the noise figure under operating conditions, which takes into account all these parameters, was defined as F.

The circuit configuration is very similar to an operational amplifier feedback circuit and therefore the voltage transfer characteristic can be derived in a similar way.

If the total RF power dissipated in the complete oscillator is limited (this is useful if minimum sideband noise is required for minimum DC input power) then the following equation can be derived<sup>1,2</sup>.

$$L_{fm} = \frac{GkT}{8Q_L^2 P_{rf}} (f_o/\Delta f)^2$$

where G is the voltage gain, F is the noise Figure, k is Boltzmann's constant, T is the operating temperature, Q<sub>L</sub> is the loaded Q of the resonator, Q<sub>o</sub> is the unloaded Q of the resonator and P<sub>rf</sub> is the power dissipated in the resonator and amplifier.

Eqn. 1 shows that L<sub>fm</sub> is inversely proportional to P<sub>rf</sub> and that better noise is thus obtained for higher feedback power. This is because the absolute value for the sideband power does not vary with the total feedback power. It should be noted that P<sub>rf</sub> is the total power in the system excluding

the losses in the amplifier, from which:

$$P_{rf} = (\text{DC input power to the system}) \times \text{efficiency}$$

For minimum noise the noise figure (F), and the value of G/Q<sub>L</sub><sup>2</sup> should be as small as possible. It should be noted, however, that F, G and Q<sub>L</sub> are directly related to each other and thus cannot be varied independently.

**OPTIMISATION FOR MINIMUM PHASE NOISE**

The first equation is now examined to see which parameters are interrelated so that the equation can be optimised for minimum phase noise. At resonance the gain of the amplifier is 1/β<sub>o</sub> and Δf is 0, then as

$$G = 1/(1-Q_L/Q_o)$$

$$L_{fm} = \frac{FkT}{8Q_o^2 (Q_L/Q_o)^2 (1-Q_L/Q_o) P_{rf}} (f_o/\Delta f)^2$$

This noise equation is minimum when

$$\frac{dL_{fm}}{dQ_L/Q_o} = 0$$

Minimum noise therefore occurs when Q<sub>L</sub>/Q<sub>o</sub> = 2/3. To satisfy Q<sub>L</sub>/Q<sub>o</sub> = 2/3, the voltage insertion loss of the resonator is 1/3 which sets the amplifier voltage gain to 3.

It is extremely important to use the correct definition of power (P), as this affects the values of the parameters required to obtain optimum noise performance.

If the power is defined as the power available at the

input of the amplifier P<sub>avi</sub> then the gain (G) will disappear from the equation. At first glance it would appear that minimum noise occurs when Q<sub>L</sub> is made large and hence tends to Q<sub>o</sub>. However this would require that the amplifier gain and output power both tend to infinity.

If we take the general oscillator model shown in figure 2 where we now define the limited output power as the power available at the output then the following equation can be derived where it is assumed that G = (1-Q<sub>L</sub>/Q<sub>o</sub>)<sup>2</sup> where G is now the transducer power gain.

$$L_{fm} = \frac{FkT}{8Q_o^2 (Q_L/Q_o)^2 (1-Q_L/Q_o)^2 P_{avo}} (f_o/\Delta f)^2$$

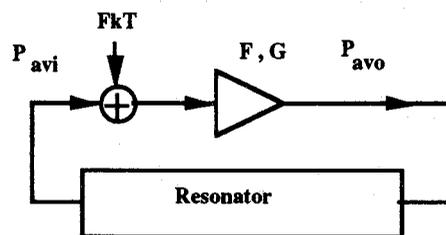


Figure 2.

Note the equivalent circuit for this model is as in figure 3 with the inclusion of an output resistance for the amplifier and the output power being defined as the power available from the output.

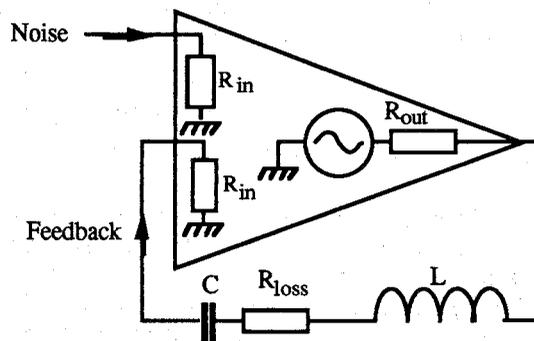


Figure 3

The minimum of the equation occurs when Q<sub>L</sub>/Q<sub>o</sub> = 1/2. It should be noted that P<sub>avo</sub> is constant and not related to Q<sub>L</sub>/Q<sub>o</sub>. The power available at the output of the amplifier is different from the power dissipated in the oscillator, but by chance is close to it. Parker<sup>3</sup> has shown a similar optimum for SAW oscillators. This should be compared with the high efficiency model where the power is defined as the total power dissipated in the resonator and the impedances of the amplifier (P<sub>rf</sub>) which is useful if highly efficient oscillators are

required, then the optimum condition occurs at  $Q_L/Q_0 = 2/3$ . The equation for the noise performance becomes:

$$L(f_m) = \frac{FKT}{8Q_0^2 (Q_L/Q_0)^2 (1-Q_L/Q_0) P_{rf}} (f_0/\Delta f)^2$$

where the last term in the denominator has now changed from  $(1-Q_L/Q_0)^2$  to  $(1-Q_L/Q_0)$

These results are most easily compared graphically as shown in Figure 3. Measurements of noise variation with  $Q_L/Q_0$  have been demonstrated using a low frequency high efficiency oscillator where the power is defined as  $P_{rf}$  and these are also included in Figure 4.

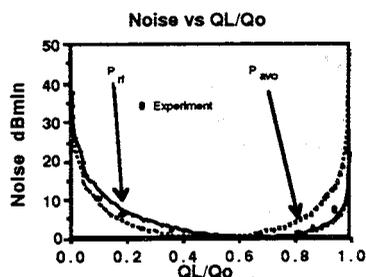


Figure 4. Phase Noise vs  $Q_L/Q_0$

The difference in the noise performance and the optimum operating point predicted by the different definitions of power is small. However care needs to be taken when using the  $P_{avo}$  definition if it is necessary to know the optimum value of the source and load impedance. For example if  $P_{avo}$  is fixed it would appear that optimum noise performance would occur when  $R_{out} = R_{in}$  because  $P_{avo}$  tends to be very large when  $R_{out}$  tends to zero. This is not the case when  $P_{rf}$  is fixed.

It should be noted that the noise factor is dependent on the source impedance presented to the amplifier and that this will change the optimum operating point dependent on the type of active device used. If the variation of noise performance with source impedance is known then this can be incorporated to slightly shift the optimum value of  $Q_L/Q_0$

### FLICKER NOISE TRANSPOSITION

The theory and optima described earlier applied in the region where thermal (additive) noise is the major noise source. This is where the noise in the oscillator falls off at a  $1/\Delta f^2$  rate. In fact for Flicker noise (modulation noise) it is often the case that  $Q_L$  should be as high as possible. However the modulation mechanisms are still not well understood and vary from device to device. Current understanding for GaAs based

devices suggests that pm noise is produced by noise modulation from the gate series noise voltage onto the input non-linear capacitance and that am noise is caused by modulation of the channel width. This group has developed a measurement system, initially presented at the 1990 IEEE MTT conference, capable of measuring the cross correlation between the baseband noise on the drain and the AM and PM components transposed onto the carrier. Further measurements have now been made which include noise measurements on the gate as well. These measurements have shown why low frequency feedback often does not greatly improve the Flicker noise

The system is shown in Figure 5. A low noise reference signal is passed through the amplifier and both the gate and drain noise are measured. The AM and PM components of the noise are measured using the delay line discriminator either in phase for AM or at 90 degrees for PM. By measuring the direct noise at the gate and drain of the device and the demodulated noise simultaneously on a digitising card the cross correlation functions can be derived. From these measurements this group has shown that the internal noise sources and their correlation coefficients can be derived. From these measurements new techniques for Flicker noise reduction can be devised. A typical cross correlation function is shown in Figure 6.

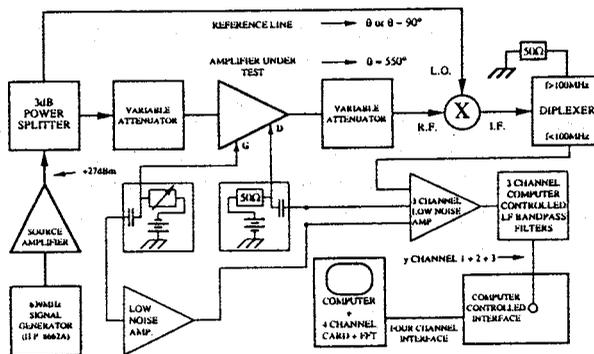


Figure 5. Flicker Noise Measurement System

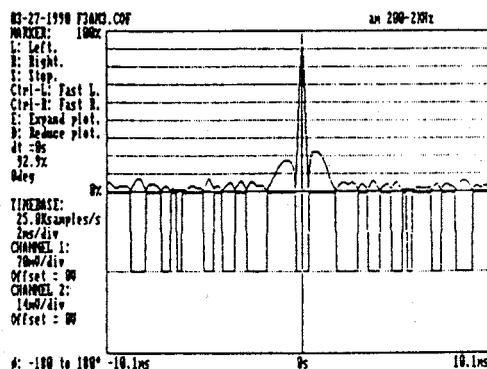


Figure 6. AM/Drain Cross Correlation Function

**Oscillator Designs**

A number of low noise oscillators have been built and these are shown in the following figures.

A 150 MHz Inductor Capacitor oscillator is shown in Figure 7. The resonator consists of a series tuned LC circuit ( $L = 255\text{nH}$ ) with a  $Q_0$  around 300. This sets the series resistance to be  $0.74 \Omega$ . To obtain  $Q_1/Q_0 = 1/2$ , LC matching networks were added at each end to transform the  $50\Omega$  impedances of the amplifier to be  $0.5 \times 50 \Omega$ . Note the series L of the transformer merges with the L of the tuned circuit. To obtain such large transformation ratios high value capacitors were used and therefore the parasitic inductance of these components should be incorporated. The resonator therefore had an insertion loss of 6 dB and a loaded  $Q = 150$ . The noise performance at 25 KHz offset was 136 dBc.

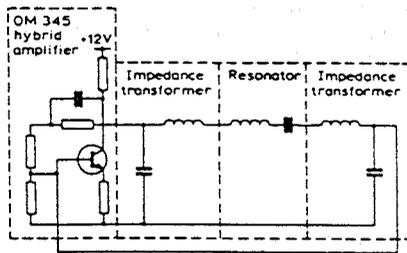


Figure 7.

A 262 MHz SAW oscillator using an STC resonator with an unloaded  $Q$  of 15,000 was built by Curley and Everard in 1987. This oscillator was built using low cost components and the noise performance was measured to be better than -130dBc at 1 KHz where the Flicker noise corner was around 1KHz. The oscillator consists of a resonator with an unloaded  $Q$  of 15,000, impedance transforming and phase shift networks and a hybrid amplifier as shown in Figure 8. The phase shift networks are designed to ensure that the circuit oscillates on the peak of the amplitude response of the resonator and hence at the maximum in the phase slope ( $d\Phi/d\omega$ ). The oscillator will always oscillate at phase shifts of  $N \times 360$  degrees where  $N$  is an integer but if this is not on the peak of the resonator characteristic, the noise performance will degrade with a  $\text{Cos}^4$  relationship.

This noise performance was in fact limited by the measurement system and new measurements of identical oscillators are about to be made. Montress, Parker, Loboda, and Greer<sup>4</sup> have demonstrated some excellent 500 Mhz SAW oscillator designs where he has reduced the Flicker noise in the resonators and operated at high power to obtain -140 dBc at 1Khz offset however the noise appears to be Flicker noise limited over the whole band.

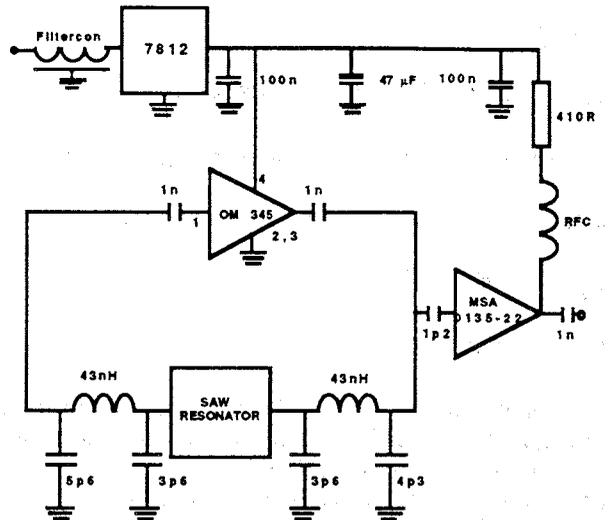


Figure 8 Low noise 262 MHz SAW oscillator

A transmission line oscillator is shown in Figure 9. Here the resonator operation is similar to that of an optical Fabry Perot and the shunt capacitors act as mirrors. The value of the capacitors are adjusted to obtain the correct insertion loss and  $Q_L/Q_0$  calculated from the loss of the transmission line.

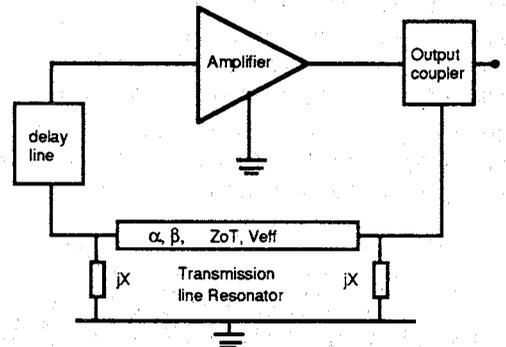


Figure 9. Transmission line oscillator

The resonator consists of a low-loss transmission line (length  $L$ ) and two shunt reactances of normalised susceptance  $jX$ . If the shunt element is a capacitor of value  $C$  then  $X = 2\pi f C Z_0$ . The value of  $X$  should be the effective susceptance of the capacitor as the parasitic series inductance is usually significant. These reactances can also be inductors, an inductor and capacitor, or shunt stubs. If  $Z_0 T = Z_0$ , where  $Z_0 T$  is the resonator line impedance and  $Z_0$  is the terminating impedance, then  $S_{21}$  is given by the following equations

$$S_{21} = 4 \Gamma / ((1+jX) - X^2(1-\Gamma^2))$$

$$\text{where } \Gamma = \exp\{-(\alpha + j\beta)L\}$$

$\alpha$  is the attenuation coefficient of the line  $\beta$  is the phase constant of the line. For small  $\alpha L (< 0.05)$  and  $\Delta f/f_0 \ll 1$ , the following properties can be derived for the first resonant peak ( $f_0$ ) of the resonator where

$$df = f - f_0,$$

$$f_0 = (v_{eff}/2L) \{ 1 + (1/\pi) \tan^{-1}(2/X) \}$$

$$S_{21}(\delta f) = S_{21}(0) / \{ 1 + j2QL(\Delta f/f_0) \}$$

$$S_{21}(0) = (1 - Q_L/Q_0) = 1 / \{ 1 + (\alpha L/2) X^2 \}$$

$$Q_1 = \pi S_{21}(0) X^2 / 4$$

$$Q_0 = \pi / 2\alpha L$$

From these equations it can be seen that the insertion loss and the loaded Q factor of the resonator are interrelated. In fact as the shunt capacitors (assumed to be lossless) are increased the insertion loss approaches infinity and  $Q_1$  increases to a limiting value of  $\pi/2\alpha L$  which we have defined as  $Q_0$ . It is interesting to note that when  $S_{21} = 1/2$ ,  $Q_1 = Q_0/2$ .

A similar design using a copper L band Helical resonator with an unloaded Q of 600 is shown in Figure 10. The helix produces both the central line and the shunt inductors; where the shunt inductors are formed by placing taps 1mm away from the end to achieve the correct  $Q_L/Q_0$ . The equations which describe this resonator are identical to those used for the 'Fabry Perot' resonator described earlier except for the fact that X now becomes  $-Z_0/2\pi fl$  where l is the inductance and L is the effective length of the transmission line. As the Q becomes larger the value of the shunt l becomes smaller eventually becoming rather difficult to realise.

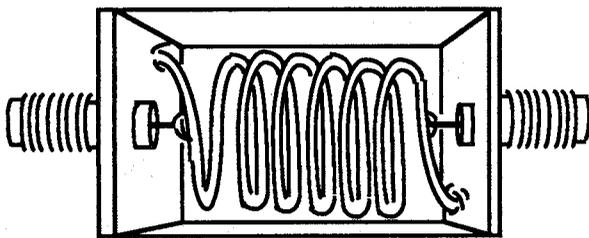


Figure 10. Helical Resonator

Printed transmission line resonators have been developed consisting of a series transmission line with shunt inductors at either end. Unloaded Q's exceeding 500 have been demonstrated at 4.5 GHz. An interesting feature of these resonators is that they do not radiate and therefore do not need to be mounted in a screened box.

#### POWER LIMITATIONS CAUSED BY THE VARACTORS:

The noise performance of a broad tuning range oscillator is usually limited by the Q and the voltage handling capability of the varactor as has been described by Underhill<sup>5</sup>. However this has not been applied to the oscillator under optimum operating conditions. If it is

assumed that the varactor diode limits the unloaded Q of the total circuit, then it is possible to obtain useful information from a simple power calculation. If the varactor is assumed to be a voltage controlled capacitor in series with a loss resistor ( $r_s$ ), The power dissipated in the varactor is  $P = (V_{rs})^2/r_s$ . The voltage across the capacitor  $V_c$  in a resonator is  $V_c = QV_{rs}$ . Therefore the power dissipated in the varactor is  $P_v = V_c^2/Q^2 r_s$ .

The noise power in oscillators is proportional to  $1/PQ_0^2$ . Therefore the figure of merit ( $V_c^2/r_s$ ) should be as high as possible and thus the varactor should have large voltage handling characteristics and small series resistances. However the definition of P and the ratio of loaded to unloaded Q are important and these will alter the effect of the varactor on the noise performance. If we set the value of  $Q_L/Q_0$  to the optimum value where again the varactor defines the unloaded Q of the resonator then the noise performance of such an oscillator can be calculated directly from the voltage handling and series resistance of the varactor. If the value of  $Q_L/Q_0$  is put in as  $2/3$  then:

$$L_{fm} = \frac{9FkT \cdot r_s}{16V_c^2} (f_0/\Delta f)^2$$

If we take a varactor with a series resistance of 1 ohm which can handle an rf voltage of 0.25 volts rms at a frequency of 1 GHz, then the noise performance at 25 KiloHertz offset can be no better than -97 dBc for an amplifier noise figure of 3dB. This can only be improved by reducing the tuning range by coupling the varactor into the tuned circuit more lightly, or by switching in tuning capacitors using PIN diodes, or by improving the varactor. The voltage handling capability can be improved by using two back to back varactors although care needs to be taken to avoid bias line currents.

Two tunable resonators are shown in Figures 11 and 12. Because the transmission line has a low impedance in the middle of the line at the operating frequency the bias resistor can be made low impedance. This means that the low frequency noise which would cause unwanted modulation noise can be kept low ( $e_n^2 = 4KTBr_b$ ). The resistor also suppresses the second harmonic response of the resonator.

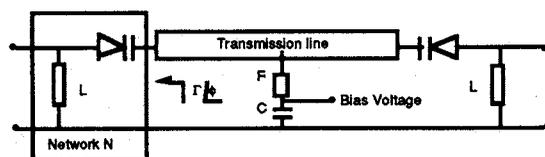
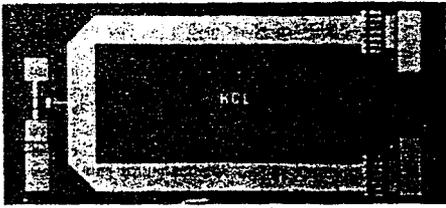


Figure (11) Tunable Octave tuning transmission line resonator with constant  $Q_L/Q_0$ . (3-6 GHz)



Figure( 12). 8 - 10 GHz Tunable MMIC resonator

### DEGRADATION OF PHASE NOISE WITH OPEN LOOP PHASE ERROR

If the open loop phase error of an oscillator is not close to  $N \times 360$  degrees the effective Q is greatly reduced. It can be shown theoretically and experimentally that the noise performance degrades both in the additive noise and the Flicker noise region by  $\cos^4 \theta$  where  $\theta$  is the open loop phase error. For high Q Dielectric resonators with Q's exceeding 4,000 then offset frequencies of 1 MHz at 10 GHz can produce a noise degradation of over 10 dB. A typical plot of the noise degradation with phase error for a high Q oscillator is shown in Figure 13 using both Silicon and GaAs active devices. The black dots are the GaAs devices and the white dots are for silicon bipolar devices.

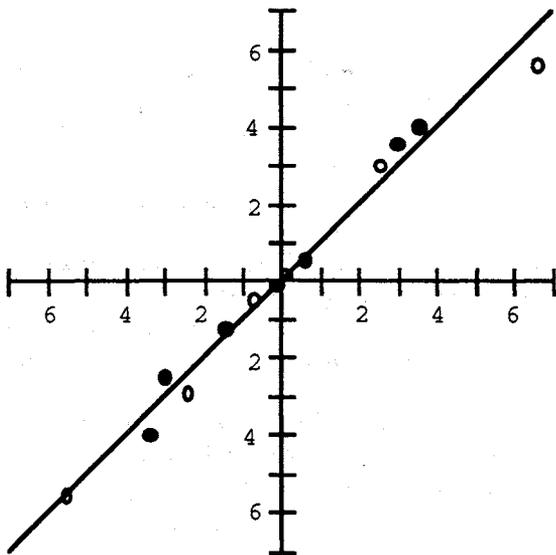


Figure 13. Noise Degradation with Phase Error

**CONCLUSIONS:** The theory required to design low noise oscillators using most of the commonly used resonators has been described and a number of experimental circuits presented. The major parameters which degrade the noise performance have been described.

### REFERENCES

1. J.K.A. Everard "Low Noise Power Efficient Oscillators: Theory and Design" IEE Proceedings Part G, Vol. 133, No.4, August 1986, pp. 172-180
2. Editors D. Haigh and J.K.A. Everard. "Gallium Arsenide Technology and its Impact on Circuits and Systems". Chapter 8 on low noise oscillators Published by Peter Peregrinus Ltd., August 31st 1989.
3. T.E. Parker, " Characteristics and sources of phase noise in stable oscillators". 41st Annual frequency control symposium 1977.pp.99-110.
4. G.Montress, T.E. Parker,M.J. Loboda, J.A. Greer. "Extremely low phase noise SAW resonators and Oscillators: Design and Performance. IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control, Vol 35,No.6, November 1988. pp.657-667.
5. M.J. Underhill,'Oscillator noise limitations', IERE Conf. Proc. 39, 1979, pp.109-118.

# QUARTZ OSCILLATORS: DERIVING OSCILLATION CONDITION BY SYMBOLIC CALCULUS

N. Ratier, R. Brendel, P. Guillemot†

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 - 25044 Besançon Cedex - France

†Centre National d'Etudes Spatiales  
18 avenue E. Belin - 31055 Toulouse Cedex - France

## Abstract

This paper presents the method used to derive the oscillation condition by using symbolic calculus. The program is based on the full nonlinear Barkhausen criterion method. The behaviour of an oscillator is described by a complex polynomial called the characteristic polynomial. This polynomial enables us to calculate the steady state features of the oscillation as well as the differential equation for transient analysis in the time domain.

The literal determination of this characteristic polynomial involves lengthy algebraic calculations and cannot be done by hand as soon as the electronic oscillator circuit involves too many components. We recently developed a formal calculus program allowing to automatically obtain all necessary equations for oscillation analysis. We propose new methods to calculate them in an optimal form.

## Introduction

Accurate analysis of an actual oscillator behaviour requires heavy calculus even for the simplest oscillator structure. Designers need automatic tools to determine oscillation frequency, amplitude and sensitivity to parameters change. We developed such a simulation tool dedicated to the modeling of quartz oscillators.

This paper presents the actual state of a computer program especially designed to accurately simulate the behaviour of quartz crystal oscillators. The plan of the paper is as follows. The first section recalls the method employed to derive the oscillation condition. Reader should consult [3][2][1] where all the explanation are carried out with complete detail. The second section describes the main problem arising when the oscillation condition is computed by symbolic calculus. The next two sections are devoted to the solution of this problem.

## Oscillation condition derivation

The transistor is modelled by a large signal admittance parameter two-port circuit. The  $y$ -parameters of which have a real and an imaginary part both being nonlinear functions of the signal amplitude (Fig. 1)(Eq. 1).

$$y_n = g_n + sC_n \text{ with } n = i, r, f, o \quad (1)$$

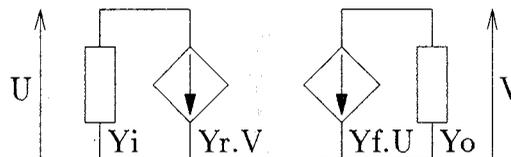


Figure 1:  $y$ -parameter representation of a transistor

The numerical values of the  $y$ -parameter are calculated at run time by using the electrical simulator SPICE [5][7]. These values depend on the amplitude and, of course, on bias and temperature conditions [3].

A quartz crystal oscillator can be reduced to the canonical form represented in Fig. 2 by using a few number of transformations [6].

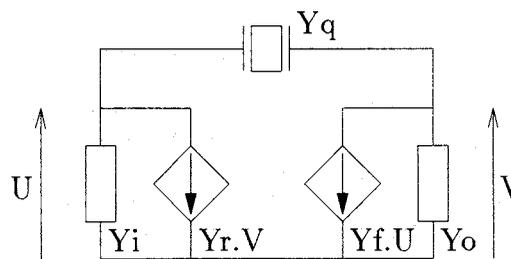


Figure 2: Reduced form of an oscillator

These transformations are divided into three groups. The first one is composed of elementary transforma-

tions on the circuit admittances like series, parallel combinations or Kennelly's transforms. Figure 3 shows one of the elementary transformation.

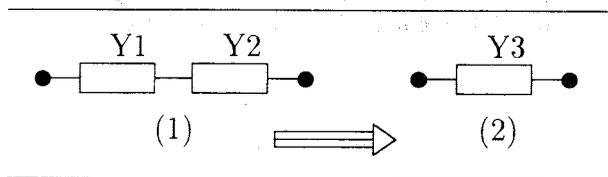


Figure 3: Elementary transformation of the circuit

Equation 2 represents the associated relationship.

$$\begin{aligned}
 Y_1 &\triangleq \frac{N_1}{D_1} \\
 Y_2 &\triangleq \frac{N_2}{D_2} \\
 Y_3 &= \frac{N_1 N_2}{N_1 D_2 + D_1 N_2}
 \end{aligned} \tag{2}$$

The second group is made up by transformations involving the  $y$ -parameter equivalent circuit and the admittances around it. Figure 4 represents an example of these transformations.

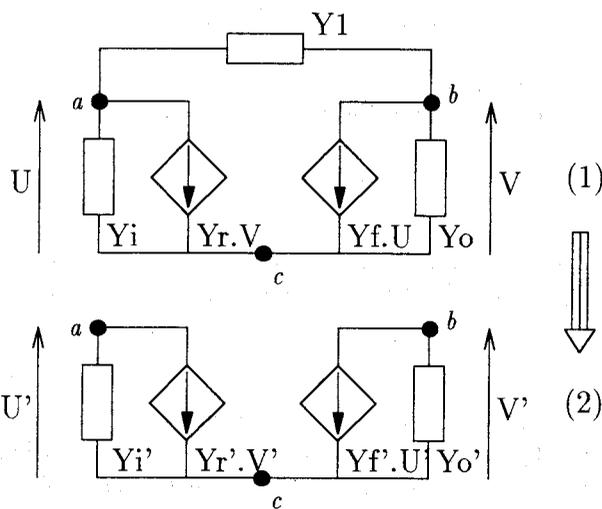


Figure 4: Elementary transformation of the circuit

After such a reduction, the components of the new  $y$ -parameter equivalent circuit are expressed in terms of the initial circuit component, as shown by (3).

$$\begin{aligned}
 Y_i' &\triangleq \frac{N_i'}{D_i'} = \frac{N_i D_1 + D_i N_1}{D_i D_1} \\
 Y_r' &\triangleq \frac{N_r'}{D_r'} = \frac{N_r D_1 - D_r N_1}{D_r D_1} \\
 Y_f' &\triangleq \frac{N_f'}{D_f'} = \frac{N_f D_1 - D_f N_1}{D_f D_1} \\
 Y_o' &\triangleq \frac{N_o'}{D_o'} = \frac{N_o D_1 + D_o N_1}{D_o D_1}
 \end{aligned} \tag{3}$$

$$\begin{aligned}
 U' &= U \\
 V' &= V
 \end{aligned}$$

The third and last group involves only the  $y$ -parameter equivalent circuit. If the quartz is not connected between the base and the collector of the  $y$ -parameter equivalent circuit, it is necessary to swap two of the terminals of the amplifier. Figure 5 illustrates that kind of transformation.

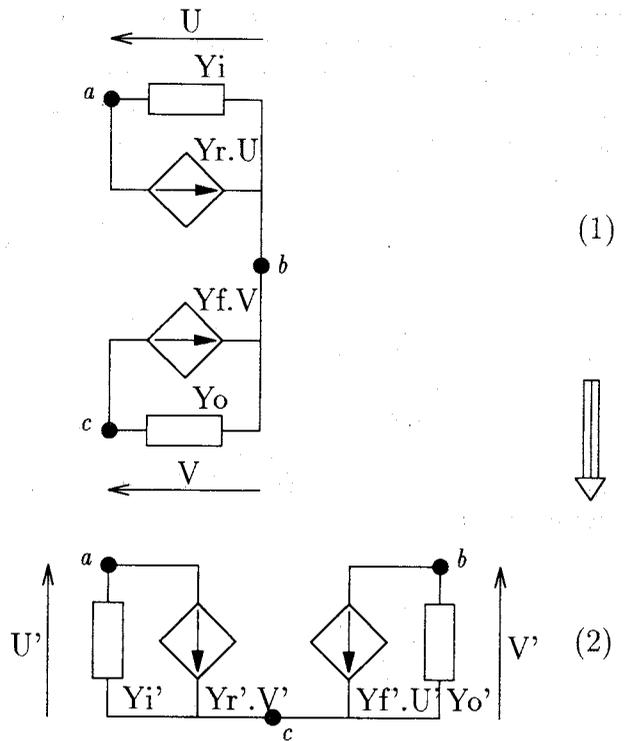


Figure 5: Elementary transformation of the circuit

The representative equations of this transformation are given by (4).

$$\begin{aligned}
 Y_i' &\triangleq \frac{N_i'}{D_i'} = \frac{N_i}{D_i} \\
 Y_r' &\triangleq \frac{N_r'}{D_r'} = \frac{-N_i D_r - D_i N_r}{D_i D_r} \\
 Y_f' &\triangleq \frac{N_f'}{D_f'} = \frac{-N_i D_f - D_i N_f}{D_i D_f} \\
 Y_o' &\triangleq \frac{N_o'}{D_o'} = \frac{N_i D_r D_f D_o + \dots + D_i D_r D_f N_o}{D_i D_r D_f D_o} \\
 U' &= U - V \\
 V' &= -V
 \end{aligned} \tag{4}$$

All the transformations described above are repeated until the final reduced form shown in figure 2 is obtained. Under this form, the oscillation condition can be written:

$$Y_i Y_o - Y_f Y_r + (Y_i + Y_r + Y_f + Y_o) Y_q = 0 \tag{5}$$

In terms of the numerator  $N_\alpha$  and denominator  $D_\alpha$  of the admittance  $Y_\alpha$ , the oscillation condition takes the following form:

$$\begin{aligned} & (N_i D_r D_f N_o - D_i N_r N_f D_o) D_q \\ & + (N_i D_r D_f D_o + D_i N_r D_f D_o \\ & + D_i D_r N_f D_o + D_i D_r D_f N_o) N_q = 0 \end{aligned} \quad (6)$$

Each term of the relation is expressed under the form of a function of previous relation until reaching the first coefficients which are expressed in terms of the circuit elements  $R_\alpha, L_\alpha, C_\alpha$ . The polynomial (6) has the general form (7). Each coefficient  $a_k$  of the polynomial is expressed as a function of the component value of the circuit

$$\sum_{k=0}^K a_k s^k = 0 \quad (7)$$

When the characteristic polynomial coefficients are defined, it is possible to obtain the steady state frequency and amplitude of the oscillation. To this end, the Laplace's variable  $s$  is changed into the harmonic variable  $j\omega$ , this splits equation (7) into two equations, one for the real part (Eq. 8), the other one for the imaginary part (Eq. 9).

$$\sum_{i=0}^N \alpha_i(u) \omega^i = 0 \quad (8)$$

$$\sum_{i=0}^M \beta_i(u) \omega^i = 0 \quad (9)$$

The numerical calculation of the variables  $u$  et  $\omega$  which satisfy both equations (8, 9) determines the frequency and the amplitude of the oscillation. This resolution uses data calculated by the SPICE program and needs the determination of the first derivatives of equations (8, 9). The numerical zero finding algorithm used to solve these equations is described in [1].

Because the characteristic polynomial coefficients are expressed as functions of all the circuit components, it is possible for the program to calculate the influence of a change of any component value on the oscillation amplitude and frequency as well as on the resonator excitation level. In the same way, it is possible to get the sensitivity of these quantities to small component variation.

In addition of the rated value the user can specify the tolerance and the temperature coefficient of each component in the input data file. This allows the program to calculate the induced dispersion and to perform worst case analysis. The temperature behaviour of the resonator is described by the four coefficients of the frequency temperature curve cubic regression so that the frequency variation can be calculated over a given temperature range.

## Setting the problem

After the circuit is read, the numerator and the denominator of each admittance of the circuit is given a name. That define the first coefficients.

$$\begin{aligned} B_1 &= s C_1 \\ B_2 &= 1 \\ \dots &= \dots \\ B_{2l} &= R_4 \end{aligned} \quad (10)$$

During the reduction process, expression of the new admittances are generated as a function of the previous coefficients.

$$\begin{aligned} B_{2l+1} &= \mathbf{B}_{2l+1}(B_1, B_2, \dots, B_{2l}) = B_1 B_4 + B_2 B_3 \\ B_{2l+2} &= \mathbf{B}_{2l+2}(B_1, B_2, \dots, B_{2l+1}) = B_2 B_4 \\ \dots &= \dots \\ B_{2n} &= \mathbf{B}_{2n}(B_1, B_2, \dots, B_{2n-1}) = \dots \end{aligned} \quad (11)$$

When all the reductions have been performed, the equivalent circuit has the reduced form shown in figure 2. The oscillation condition gives the last coefficient:

$$B_{2n+1} = (B_{81} B_{84} B_{86} B_{87} - B_{82} B_{83} B_{85} B_{88}) B_{90} + \dots \quad (12)$$

The equation  $B_{2n+1} = 0$  can be written in the form of a polynomial of the Laplace's variable  $s$ :

$$\sum_{k=0}^K a_k s^k = 0 \quad (13)$$

The main problem is to compute the four following expressions from the coefficients  $B_1, B_2, \dots, B_{2n+1}$ .

$$\begin{aligned} P &= \sum_{i=0}^N \alpha_i(u) \omega^i \quad \left| \quad \frac{\partial P}{\partial \omega} \right. \\ Q &= \sum_{i=0}^M \beta_i(u) \omega^i \quad \left| \quad \frac{\partial Q}{\partial \omega} \right. \end{aligned}$$

The easiest way is to substitute recursively each  $B_i$  by their expression in the last coefficient  $B_{2n+1}$ . By this method, we obtain the polynomial  $\sum_{k=0}^K a_k s^k = 0$  where each coefficient is expressed as a function of the component of the circuit  $R_\alpha, L_\alpha, C_\alpha$ . Unfortunately, the full development of the polynomial coefficient leads to too intricate expression even in simple oscillator structures.

A better way is to express the coefficients of an admittance newly defined as a function of the coefficient previously defined. This method is detailed in the following example. Suppose that the admittance  $Y_3$  is generated after the parallel reduction of the admittance  $Y_1$  and  $Y_2$ .

$$\begin{aligned} Y_3 &= Y_1 + Y_2 \\ &= \frac{\dots}{a_2 s^2 + a_1 s + a_0} + \frac{\dots}{b_3 s^3 + b_2 s^2 + b_1 s + b_0} \\ &= \frac{\dots}{c_5 s^5 + c_4 s^4 + c_3 s^3 + c_2 s^2 + c_1 s + c_0} \end{aligned} \quad (14)$$

This reduction process involve the definition of six new coefficients  $c_0, c_1, \dots, c_5$  (Eq. 15).

$$\begin{aligned} c_0 &= a_0 b_0 \\ c_1 &= a_1 b_0 + a_0 b_1 \\ c_2 &= a_2 b_0 + a_1 b_1 + a_0 b_2 \\ c_3 &= a_2 b_1 + a_1 b_2 + a_0 b_3 \\ c_4 &= a_2 b_2 + a_1 b_3 \\ c_5 &= a_2 b_3 \end{aligned} \quad (15)$$

This method has the following major drawback: The degree of the polynomial increases at each reduction step. The consequence is that the number of new coefficients defined during one step increases with the number of transformations already done. This exponential increase limits the complexity of the oscillator circuit which can be treated.

The method proposed here will compute in the same time the real and the imaginary part of the equation (7 or 13). The first derivatives  $\partial P/\partial\omega, \partial Q/\partial\omega$  are calculated from these coefficients. These values are calculated during the same step. This method will be further explained in the next two sections. The advantages over the previous method are:

- The number of coefficients increases linearly with the number of reduction steps.
- These coefficient are coded in a efficient way.
- In consequence of it, the numerical calculation of the four functions  $P, Q, \partial P/\partial\omega, \partial Q/\partial\omega$  is almost optimal.

Moreover, this method allows one to calculate efficiently the values of  $P, Q, \partial P/\partial\omega, \partial Q/\partial\omega$  without knowing explicitly the expression of the coefficients  $a_k$  in equation (7).

## Determination of the real and imaginary part

The first coefficients defined (see Eq. 10) are function of the Laplace's variable  $s$ , so the splitting into real and imaginary part is straightforward. Equations (11) and (12) show that all the other equations have the specific form of equation (16). In this particular case, an efficient method to calculate symbolically the real and imaginary part of the coefficient can be used.

$$B_{i+1} = \sum \prod_{k \leq i} B_k \quad (16)$$

Suppose that the real and the imaginary part of each coefficient  $B_\alpha$  is known up to the  $i$ th order.

$$\Re(B_1), \Re(B_2), \dots, \Re(B_i), \quad (17)$$

$$\Im(B_1), \Im(B_2), \dots, \Im(B_i) \text{ are known} \quad (18)$$

It is possible to calculate the real and the imaginary part of the coefficient  $B_{i+1}$ . We shall illustrate the method on a concrete example. Consider the expression

$$Exp = B_1 B_2 B_3 + B_4 B_5 \quad (19)$$

The method introduces new coefficient  $C_\alpha$ . These coefficient are calculated so that only one complex multiplication is done in each step. In the case of the previous expression, the result is:

$$\begin{aligned} C_1 &= \Re(B_1)\Re(B_2) - \Im(B_1)\Im(B_2) \\ C_2 &= \Re(B_1)\Im(B_2) + \Im(B_1)\Re(B_2) \\ C_3 &= C_1\Re(B_3) - C_2\Im(B_3) \\ C_4 &= C_1\Im(B_3) + C_2\Re(B_3) \\ C_5 &= \Re(B_4)\Re(B_5) - \Im(B_4)\Im(B_5) \\ C_6 &= \Re(B_4)\Im(B_5) + \Im(B_4)\Re(B_5) \end{aligned} \quad (20)$$

$$\Re(Exp) = C_3 + C_5$$

$$\Im(Exp) = C_4 + C_6$$

Notice that the real part and the imaginary part are computed in the same time. Let us compare this method with the more direct one.

$$\begin{aligned} \Re(Exp) &= \Re(B_1)\Re(B_2)\Re(B_3) - \Re(B_1)\Im(B_2)\Im(B_3) \\ &\quad - \Im(B_1)\Re(B_2)\Im(B_3) - \Im(B_1)\Im(B_2)\Re(B_3) \\ &\quad + \Re(B_4)\Re(B_5) - \Im(B_4)\Im(B_5) \end{aligned} \quad (21)$$

$$\begin{aligned} \Im(Exp) &= \Re(B_1)\Re(B_2)\Im(B_3) + \Re(B_1)\Im(B_2)\Re(B_3) \\ &\quad + \Im(B_1)\Re(B_2)\Re(B_3) - \Im(B_1)\Im(B_2)\Im(B_3) \\ &\quad + \Re(B_4)\Im(B_5) + \Im(B_4)\Re(B_5) \end{aligned} \quad (22)$$

Table 1 indicates the number of arithmetic operation + and \* needed to express the form  $B_1 B_2 \dots B_i$ . The two methods are compared for expressions of different lengths.

	Direct		Proposed	
	Nb +	Nb *	Nb +	Nb *
$B_1$	0	0	0	0
$B_1 B_2$	2	4	2	4
$B_1 B_2 B_3$	6	16	4	8
$B_1 B_2 B_3 B_4$	14	48	6	12
$B_1 B_2 B_3 B_4 B_5$	30	128	8	16

Table 1: Computation cost

## Determination of the first derivatives

The first coefficients are necessarily defined as a function of  $\omega$  and of the admittance of the circuit. So their derivatives respect to  $\omega$  can be easily calculated. Now suppose that the first derivative of each coefficient  $C_\alpha$  is known up to the  $i$ th order.

$$\frac{\partial C_1}{\partial\omega}, \frac{\partial C_2}{\partial\omega}, \dots, \frac{\partial C_i}{\partial\omega} \quad (23)$$

It is possible to calculate the derivative of the coefficient  $C_{i+1}$  by expanding as follows.

$$\frac{\partial C_{i+1}}{\partial \omega} = \frac{\partial C_{i+1}}{\partial C_1} \frac{\partial C_1}{\partial \omega} + \dots + \frac{\partial C_{i+1}}{\partial C_i} \frac{\partial C_i}{\partial \omega} \quad (24)$$

One might think that the form of the  $\partial C_i / \partial \omega$  will be extremely complicated. It is not the case because of the specificity of our problem. From equations (20), it easily follows that each coefficient  $C_\alpha$  depends only on four other coefficients. The consequence is that the expression of the derivative will be quite simple. We showed above an example. The expression of the derivative  $D_{10}$  of the coefficient  $C_{10}$  is given by (26), where all the coefficients  $C_1, \dots, C_9$  and  $D_1, \dots, D_9$  are known.

$$C_{10} = C_3 C_5 - C_4 C_6 \quad (25)$$

$$D_{10} = D_3 C_5 + C_3 D_5 - D_4 C_6 - C_4 D_6 \quad (26)$$

## Conclusion

The works presented here highlights some problems related to the complexity of the expressions to be manipulated. The proposed methods allow an optimization of the calculus at the level of the topological analysis as well as for numerical calculation of the required coefficients to the calculation of the oscillation's condition.

The fact that this optimization is not a simple calculation refinement that one can do without should be stressed. The present method is necessary in order to treat complex oscillator circuits. Indeed, we have shown that with the previous method the complexity of the calculus exponentially increases as a function of the circuit components while the complexity of the new method linearly increases.

## Acknowledgements

This work has been supported by CNES and DGA under contracts: # 962/CNES/91/1476/00, # 962/CNES/94/1230/00.

## References

- [1] R. Brendel, G. Marianneau, T. Blin, M. Brunet, *Computer aided design of quartz crystal oscillators*, Proc. of 1994 IEEE IFCS (1994).
- [2] R. Brendel, T. Blin, G. Marianneau, *Modélisation non linéaire des oscillateurs ultrastables pour applications aérospatiales*, Final report contract 962/CNES/91/1476/00 (1994).
- [3] R. Brendel, F. Djian, E. Robert, *High precision nonlinear computer modelling techniques for quartz crystal oscillators*, Proc. 45th AFCS (1991).

- [4] R. Brendel, F. Djian, *Modélisation des effets d'environnement sur les caractéristiques des oscillateurs à quartz ultrastables*, Final report contract 832/CNES/88/5374/00 (1991).
- [5] L.W. Nagel, *SPICE 2: A computer program to simulate semiconductor circuits*, Memorandum ERL-M520, Univ. of California, Berkeley (1975).
- [6] N. Ratier, R. Brendel, T. Blin, G. Marianneau, P. Guillemot, *Non linear simulation of quartz crystal oscillators*, Proc. of 9th EFTF (1995).
- [7] P. Tuinenga, *SPICE - A guide to circuit simulation and analysis using PSpice*, 2nd edition, Prentice Hall, 1992.

## PHASE NOISE FIGURES COMPARISON IN TRANSISTOR AMPLIFIERS OF DIFFERENT TYPES

R.J. BESSON, M. MOUREY, S. GALLIOU

Laboratoire de Chronométrie, Electronique et Piézoélectricité (ENS2M)  
26 Chemin de l'Épitaphe  
25030 BESANÇON CEDEX - FRANCE

Ph. GUILLEMOT

CNES - 18 avenue Edouard Belin - 31055 TOULOUSE CEDEX - FRANCE

### Abstract

Ultrastable quartz oscillators must be specially designed so as to match resonators parameters and electronic circuitry characteristics. In a recent paper [1] we have pointed out that our results have never been really resonator limited (at least for BVA units) and that further improvements would largely depend on progress in amplifiers. Tremendous advances are possible as pointed out by Fred WALLS et al [2]. In order to improve the output signal stability of ultrastable oscillators it is necessary to reduce oscillating loop phase noise and to minimize phase noise of each stage of output amplifier as well.

In this paper, each type of bipolar transistor amplifier configuration is considered : Common Emitter (C.E.), Common Collector (C.C.) and Common Base (C.B.). Power spectral density (P.S.D.) of phase noise is calculated from the partial derivative of transfer function argument versus transistor and components parameters variations. In such a calculation, those variations are assumed to be uncorrelated (their well known origins will not be discussed here).

After this theoretical calculation, the P.S.D. (Power Spectral Density) of each basic configuration (i.e. C.E., C.C., C.B.) is analysed. Results are compared and discussed taking into account criteria as stage gains, input or output impedances.

Consequences are discussed in terms of amplifiers performances and oscillators stabilities.

### I - Introduction

We start from an usual structure for quartz oscillators, i.e. 3 elements according to scheme of Fig. 1 where amplifiers are transistor amplifiers in usual configuration (common emitter CE, common collector CC, common base CB).

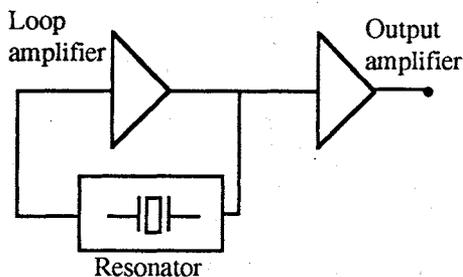


Fig. 1

If an SC-cut resonator is used it is usually necessary to filter out B mode. On the other hand reduction of output harmonic ratio also

needs use of a tuned amplifier. Usually a cascode output amplifier is used. Under those conditions, it is highly desirable to study influence of different transistor internal parameters and tuned circuits elements on frequency stability of oscillators. Some results have already been published [1] [2] [3].

### II - Description of investigation method

Let us consider  $G(x_i, f)$  linear transfer function of a transistor amplifier.  $\{x_i\}$  represents parameters that may fluctuate and  $f$  is frequency.

Phase difference  $\varphi$  between output and input signal is :

$$\varphi = \text{Arg } G(x_i, f) \quad \text{Eq(1)}$$

and

$$d\varphi = \Phi = \sum_i \frac{\partial}{\partial x_i} \text{Arg } G(x_i, f_0) dx_i \quad \text{Eq(2)}$$

for  $f = f_0$  where  $f_0$  is the carrier frequency.  $dx_i$  is variation of parameter  $x_i$  on a 1Hz band width. We assume that the  $dx_i$  are independant of Fourier frequency.

From Eq(2) we derive  $|\Phi_i|^2$  expressed in  $\text{rd}^2/\text{Hz}$ , i.e, homogeneous to a power spectral density (PSD) of phase fluctuations. We consider that only one parameter  $x_i$  varies (in other words :  $dx_j = 0$  if  $j \neq i$ ).  $dx_i$  corresponds to rms values of  $x_i$  fluctuations whose origin is not specified [4] [5] [6].

2.1 - Variable parameters  $x_i$

To represent transistor we use model of PSPICE simulation (called enhanced Gummel-Poon model) represented in Fig. 2 [7].

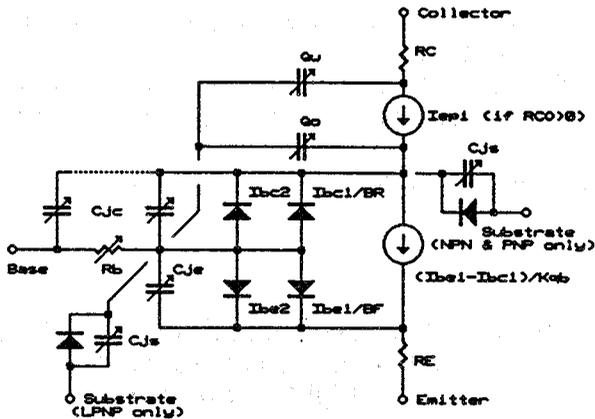


Fig. 2

$x_i$  parameters previously quoted will appear in various tables for results. They represent parasitic capacitance base-emitter  $C_{je}$ , base collector  $C_{jc}$ , the low. frequency gain BF and pure resistances in emitter (RE) and base (Rb).

2.2 - We analysed various configurations using simulation PSPICE software. Numerical values correspond to what is usually met at 10 MHz. Transistor is of the 2N2857 type whose characteristics correspond to following model :

- Model Q2N2857 NPN( $I_s=69.28E-18$   $X_{ti}=3$   $E_g=1.11$   $V_{af}=100$   $B_f=288$   $N_e=1.167$   $I_{se}=69.28E-18$   $I_{kf}=21.59m$   $X_{tb}=1.5$   $Br=1.219$

- $N_c=2$   $I_{sc}=0$   $I_{kr}=0$   $R_c=4$   $C_{jc}=893.1f$   $M_{jc}=.3017$   $V_{jc}=.75$   $F_c=.5$   $C_{je}=939.8j$   $M_{je}=.3453$   $V_{je}=.75$   $Tr=1.607n$   $Tf=115.7p$   $Itf=.27$   $V_{tf}=10$   $X_{tf}=30$   $R_b=10$   $KF=8e-15$   $AF=1$ )

all simulations have been done at 10 MHz using small signal equivalent circuit model around the bias point (analysis called a.c). Following figures Fig. 3a, Fig. 3b, Fig. 3c, Fig. 3d represent performed analysis for common emitter (CE), common collector (CC), common base (CB) and cascode configurations.

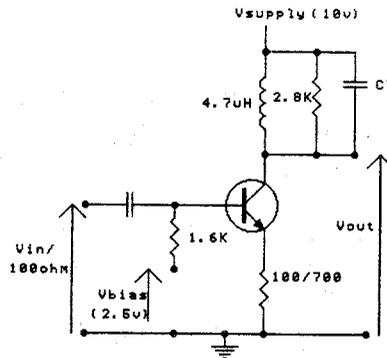


Fig. 3a : C.E.

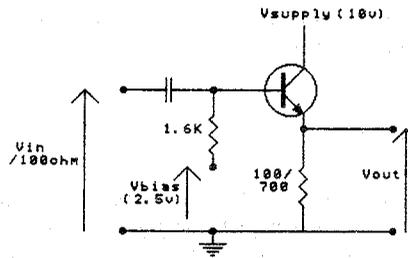


Fig. 3b : C.C.

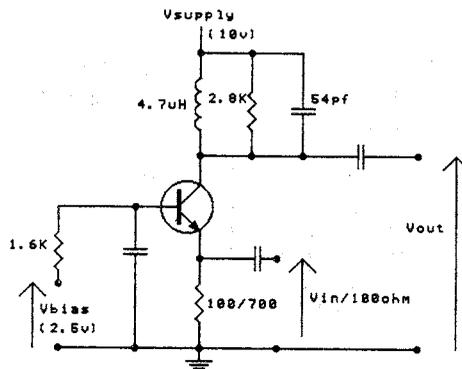


Fig. 3c : C.B.

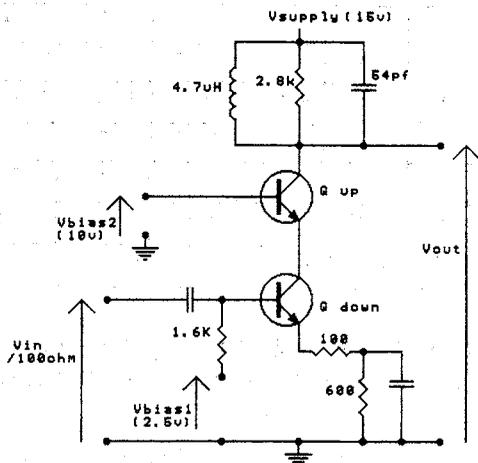


Fig. 3d : cascode configuration

2.3 - Results of simulation

Table 1 gives phase variation (expressed in degrees) caused by variation of one of five  $x_i$  parameters (the four other being constant at nominal values). Variations  $\Delta x_i$  are such that  $\frac{\Delta x_i}{x_i} = 1 p.p.m.$  while polarization current  $I_E$  in transistor is  $I_E = 2.6 \text{ mA}$ . For CE configurations (emitter resistance uncoupled) and for CB configurations voltage gains  $|G|_{dB}$  are respectively 11 dB and 27 dB. Emitter resistance is the same in both cases. Last line in Table 1 gives those results when emitter resistance is partially by-passed so as to obtain a 27 dB gain.

$x_i$	$C_{je}$	$C_{jc}$	$R_b$	$R_c$	$\beta$	$ G _{dB}$
C.C.	$8.6 \cdot 10^{-9}$	$1.6 \cdot 10^{-7}$	$2.15 \cdot 10^{-8}$	$9.6 \cdot 10^{-10}$	$7.1 \cdot 10^{-10}$	
C.E.	$8 \cdot 10^{-8}$	$6.2 \cdot 10^{-6}$	$7.9 \cdot 10^{-8}$	$9.5 \cdot 10^{-9}$	$5.4 \cdot 10^{-9}$	11 dB
C.B.	$6.9 \cdot 10^{-8}$	$4.7 \cdot 10^{-6}$	$4.6 \cdot 10^{-7}$	$8.4 \cdot 10^{-9}$	$7.6 \cdot 10^{-9}$	27 dB
C.E. by passed	$1.25 \cdot 10^{-7}$	$8.4 \cdot 10^{-6}$	$4.8 \cdot 10^{-7}$	$1.47 \cdot 10^{-8}$	$4.5 \cdot 10^{-8}$	27 dB

Table 1 :  $|\Phi_i|$  in degrees for C.C., C.E., C.B. and C.E. by passed configurations when  $I_E = 2.6 \text{ mA}$  and  $\frac{\Delta x_i}{x_i} = 1 \text{ ppm}$ .

In table 2 we express power spectral density (PSD) of phase fluctuation in dBc.

$x_i$	$C_{je}$	$C_{jc}$	$R_b$	$R_c$	$\beta$	$ G _{dB}$
C.C.	-196	-171	-188	-215	-218	
C.E.	-177	-139	-177	-195	-200	11 dB
C.B.	-178	-141	-162	-197	-197,5	27 dB
C.E. by passed	-173	-137	-161.5	-192	-182	27 dB

Table 2 :  $|\Phi_i|^2$  in dBc for  $(\Delta x_i / x_i = 1 \text{ ppm})$ .

Now, let us consider that phase fluctuation  $\phi$  occurs at a Fourier frequency  $\nu$  where  $\nu$  is inside the bandwidth of a 10 MHz quartz resonator ( $f_0 = 10 \text{ MHz}$ ) with a loaded Q of  $10^6$  ( $Q = 10^6$ ). Then, relative frequency fluctuation  $\frac{\Delta f}{f_0}$  of the oscillator built with this resonator and the amplifier is given by Table 3.

$x_i$	$C_{je}$	$C_{jc}$	$R_b$	$R_c$	$\beta$	$ G _{dB}$
C.C.	$9.6 \cdot 10^{-17}$	$1.8 \cdot 10^{-15}$	$2.4 \cdot 10^{-16}$	$1.1 \cdot 10^{-17}$	$7.9 \cdot 10^{-18}$	
C.E.	$8.9 \cdot 10^{-16}$	$6.9 \cdot 10^{-14}$	$8.8 \cdot 10^{-16}$	$1.1 \cdot 10^{-16}$	$6 \cdot 10^{-17}$	11 dB
C.B.	$7.7 \cdot 10^{-16}$	$5.2 \cdot 10^{-14}$	$5.1 \cdot 10^{-15}$	$9.3 \cdot 10^{-17}$	$8.4 \cdot 10^{-17}$	27 dB
C.E. by passed	$1.4 \cdot 10^{-15}$	$9.3 \cdot 10^{-14}$	$5.3 \cdot 10^{-15}$	$1.6 \cdot 10^{-16}$	$5 \cdot 10^{-16}$	27 dB

Table 3 :  $\frac{\Delta f}{f_0}$  for a 10 MHz quartz oscillator with  $Q = 10^6$  and  $\Delta x_i / x_i = 1 p.p.m.$

Table 4 is similar to Table 1 but  $I_E = 10.4 \text{ mA}$ , i.e.  $I_E$  has been multiplied by 4.

$x_i$	$C_{je}$	$C_{jc}$	$R_b$	$R_c$	$\beta$	$ G _{dB}$
C.C.	$1.29 \cdot 10^{-8}$	$1.6 \cdot 10^{-7}$	$3.43 \cdot 10^{-8}$	$4.49 \cdot 10^{-9}$	$1.05 \cdot 10^{-8}$	
E.C.	$4.1 \cdot 10^{-8}$	$7.44 \cdot 10^{-6}$	$2.3 \cdot 10^{-7}$	$2 \cdot 10^{-8}$	$9 \cdot 10^{-8}$	27 dB
B.C.	$2.6 \cdot 10^{-8}$	$4.8 \cdot 10^{-6}$	$5.5 \cdot 10^{-7}$	$1 \cdot 10^{-8}$	$4 \cdot 10^{-8}$	27 dB

Table 4 :  $|\phi_i|$  in degrees  $I_E = 10.4 \text{ mA}$   $\Delta x_i / x_i = 1 p.p.m.$

Results of Table 4 expressed in dBc and in relative frequency fluctuations  $\frac{\Delta f}{f_0}$  lead to Table 5 and Table 6.

$x_i$	$C_{je}$	$C_{jc}$	$R_b$	$R_c$	$\beta$	$ G _{dB}$
C.C.	-193	-151	-164	-202	-214	
C.E.	-183	-137	-168	-189	-176	27 dB
C.B.	-187	-141	-160	-195	-183	27 dB

Table 5 :  $|\phi_i|^2$  in dBc for  $I_E = 10.4$  mA  
 $\Delta x_i / x_i = 1$  p.p.m.

$x_i$	$C_{je}$	$C_{jc}$	$R_b$	$R_c$	$\beta$	$ G _{dB}$
C.C.	$1.4 \cdot 10^{-16}$	$1.8 \cdot 10^{-14}$	$3.8 \cdot 10^{-15}$	$5 \cdot 10^{-17}$	$1.2 \cdot 10^{-17}$	
E.C.	$4.6 \cdot 10^{-16}$	$8.3 \cdot 10^{-14}$	$2.6 \cdot 10^{-15}$	$2.2 \cdot 10^{-16}$	$1 \cdot 10^{-15}$	27 dB
B.C.	$2.9 \cdot 10^{-16}$	$5.3 \cdot 10^{-14}$	$6.1 \cdot 10^{-15}$	$1.1 \cdot 10^{-16}$	$4.4 \cdot 10^{-16}$	27 dB

Table 6 :  $\frac{\Delta f}{f_0}$  for  $I_E = 10.4$  mA.

Since it is very usual to use a cascode configuration in the output amplifier of an ultrastable oscillator (U.S.O.) we have also considered configuration of Fig. 4.

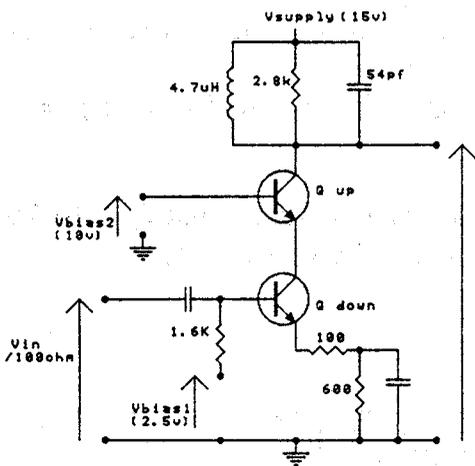


Fig. 4

Here again we assumed a gain of 27 dB. Results that can be derived for such a configuration are given by Table 7.

$x_i$	$C_{je}$	$C_{jc}$	$R_b$	$R_c$	$\beta$	$ G _{dB}$
$Q_{down}$	$1.3 \cdot 10^{-7}$	$4 \cdot 10^{-7}$	$6 \cdot 10^{-8}$	$1.5 \cdot 10^{-8}$	$5 \cdot 10^{-9}$	in degrees
$Q_{up}$	$7.8 \cdot 10^{-8}$	$4.8 \cdot 10^{-6}$	$4 \cdot 10^{-10}$	$9.3 \cdot 10^{-9}$	$3.3 \cdot 10^{-9}$	in degrees
$Q_{down}$	-173	-163	-179.6	-191.6	-201	dBc
$Q_{up}$	-177	-141.5	-223	-196	-205	dBc
$Q_{down}$	$1.4 \cdot 10^{-15}$	$4.4 \cdot 10^{-15}$	$6.7 \cdot 10^{-16}$	$1.7 \cdot 10^{-16}$	$5.6 \cdot 10^{-17}$	$(\Delta f / f_0)$
$Q_{up}$	$8.7 \cdot 10^{-16}$	$5.3 \cdot 10^{-14}$	$4.4 \cdot 10^{-18}$	$1 \cdot 10^{-16}$	$3.16 \cdot 10^{-17}$	$(\Delta f / f_0)$

Table 7 : Cascode amplifier;  $I_E = 2.6$  mA,  $|G| = 27$  dB (emitter resistor partially by-passed).

$Q_{up}$  :

$\Delta x_i / x_i = 1$  p.p.m. for  $Q_{up}$  and  $\Delta x_i = 0$  for  $Q_{down}$   
 $Q_{down} : \Delta x_i = 0$  for  $Q_{up}$  and  $\Delta x_i / x_i = 1$  p.p.m. for  $Q_{down}$

for comparison purposes, an identical fractional variation of 1 p.p.m. of one of the components (L or C) of tuning circuits will provide results expressed in Table 8.

	$ \phi_i $ in degree	$ \phi_i ^2$ in dBc	$\Delta f / f_0$
$Q_t = 25$	$1.2 \cdot 10^{-5}$	-133.6	$1.3 \cdot 10^{-13}$
$Q_t = 50$	$1.7 \cdot 10^{-5}$	-130.6	$1.9 \cdot 10^{-13}$
$Q_t = 75$	$1.8 \cdot 10^{-5}$	-130	$2 \cdot 10^{-13}$

Table 8 : C.E. configuration with  $I_E = 2.6$  mA. Influence of tuning circuit  $\Delta C_T / C_T = 1$  ppm for 3 values of tuning circuit quality factor  $Q_t$ .

### III - Conclusion

From upper results it is clear that :

- C.C. configuration provides **better performances in terms of phase noise**. However, C.C. configuration does not allow (in the contrary of C.E. and C.B. configurations) voltage gain.

- Base collector capacitance is the most influent parameter whatever is the chosen configuration at least for our purpose. Advantage in using a transistor with weak "parasitic" capacitances ( $C_{je}$ ,  $C_{jc}$ ) is obvious at least from that point of view.

In fact, the last remark does not consider the origin of parameters fluctuations. Actually HF transistors with weak capacitances do not always give best results. For a study to be exhaustive it is necessary, in our opinion, to take into account the possible origins of

parameters fluctuations. They may be of three types :

- 1/ - Non linear origin by transfer of low frequency phase noise around carrier.
- 2/ - Amplitude-phase relationship (Miller effect, amplitude frequency effects [8]...).
- 3/ - Pure phase noise.

## REFERENCES

[1] R.J. BESSON, J.J. BOY, M.M. MOUREY, "BVA resonators and oscillators : a review. Relation with space requirements and quartz material characterization". Proceedings of IEEE International Frequency Control Symposium, San Francisco, June 1995.

[2] F.L. WALLS, E.S. FERRE-PIKAL, S.R. JEFFERTS, "The origin of 1/f PM and AM noise in bipolar junction transistor amplifiers", Proc. of the IEEE International Frequency Control Symposium, San Francisco, 1995, pp 294-304.

[3] E.S. FERRE-PIKAL, F.L. WALLS and C.W. NELSON, "Design criteria for BJT amplifiers with low 1/f AM and PM noise". Proc. IEEE International Frequency Control Symposium 1995, pp 305-313.

[4] M. MOUREY, S. GALLIOU, "Electronique et structure thermo-mécanique" ; contrat ENSMM/CNES marché N° 962/94/CNES/1368/00, rapport de phase 0, mars 1995.

[5] M. MOUREY, S. GALLIOU, "Electronique et structure thermo-mécanique" ; contrat ENSMM/CNES marché N° 962/94/CNES/1368/00, rapport final du lot 1, juillet 1995.

[6] E. ORFILA, "Comparaison de la DSP des fluctuations de phase dans des montages à transistors bipolaires EC et CC", rapport de DEA, ENSMM-Besançon, janvier 1996.

[7] PSPICE "Circuit analysis user's guide", Microsim Corporation, 26 Fairbanks Irvine, California 92718.

[8] J.J. GAGNEPAIN, R. BESSON, "Non linear effects in piezoelectric quartz crystals", Physical Acoustics, Vol. XI, pp 245-288, Academic Press (1975).

## SPECTRAL IMPROVEMENT OF DIRECT DIGITAL FREQUENCY SYNTHESISERS AND OTHER FREQUENCY SOURCES

M J Underhill, M J Blewett

Department of Electronic & Electrical Engineering, University of Surrey,  
Guildford, UK

### ABSTRACT

Direct digital synthesis has the major advantage that the low phase noise of the (quartz) reference oscillator can be transferred directly to the signal output. However, limitations on the precision of the high speed Digital to Analogue Converter (DAC) required to generate the output waveform cause a high level of spurious signals to be generated.

The paper describes a simple spurious signal reduction method which can be applied to the output of a high frequency programmable rate multiplier. The advantage of the rate multiplier method of direct frequency synthesis is that no high speed DAC is required. The method extracts a phase error signal in the form of analogue samples directly from the rate multiplier signal. This signal is used in a novel self adjusting delay compensation circuit (SADC) to correct the time positions of the transitions of the square wave output waveform. This has the effect of reducing both the spurious components and the broadband phase noise of the signal.

Also described is a method of converting the waveform from any frequency source into a form suitable for the application of the SADC. This method allows SADCs to be applied in cascade for progressive reduction of phase noise.

### 1. INTRODUCTION

The direct digital frequency synthesiser (DDS) (1,2) has the advantage that it can switch within one clock cycle from one frequency to another. Another advantage is that the DDS in general has very good close in phase noise. Provided that care is taken to re-clock the DDS output as closely as possible with the DDS clock the close-in phase noise can be as good as that of the fixed frequency (usually quartz crystal) clock. A major disadvantage of the DDS is a high level of spurious discrete frequency signals. These signals appear in profusion at frequencies which are always whole number multiples of the lowest common factor of the clock and the output frequencies. The total spurious rms power can be shown to be  $q^2/12$  where  $q$  is the smallest quantisation step of the DDS output. Clearly the more bits that are used in the output digital to analogue converter (DAC) the lower the level of spurious power that can be achieved. The use of higher clock frequencies allows the spurious power to be spread over a proportionately wider frequency band. A further useful technique is to design the DDS so that the unwanted spurious signal

power is concentrated at frequencies well away from the wanted output signal frequency, so that most of it can be easily removed by simple filters. This requires more complicated logic circuitry and also usually requires the use of higher clock frequencies.

With current IC (integrated circuit) technologies higher clock frequencies and/or multi-bit DACs require excessive power consumption for output frequencies in excess of a few megahertz.

Multi-bit DACs give spurious noise levels which are reduced by 6dB for every bit which is added. But for high clock speeds the relative weights of the bits cannot be maintained and the achievable noise reduction becomes limited to only a few bits (eight being typical). DAC power consumption increases with the number of bits and is at least proportional to the clock rate.

With the objective of achieving lower power consumption and lower spurious levels for any given output frequency, the one bit rate multiplier type of DDS was proposed (3). This has delay compensation to reduce time jitter which consequently reduces the

spurious signal levels. In this scheme a DAC is required but it only has to operate at the output frequency, rather than at the clock frequency. In this way the power consumption is reduced.

In a newer type of one-bit delay compensated DDS (4) the compensation signal is derived by analogue circuitry in place of the DAC giving a considerable opportunity for power reduction.

The purpose of this paper is to describe a new proprietary analogue delay compensation method which is simpler. It can either be applied to the one bit rate multiplier DDS directly, or with some further simple circuitry it can be applied to any frequency source. It reduces both broadband phase noise and the spurious signals generated by clocked digital circuits.

In the following sections first the two main types of DDS are described and their performance compared. Then the principle of extracting delay compensation information from a waveform is described. Next to be described is the way in which the required delay modulator can be combined in a simplified way with the analogue delay information generation circuit. Following this the way in which this self adjusting delay compensation circuit can be applied to any frequency source is shown. Finally some first results of the proprietary method applied to noisy frequency sources are given.

**2. A COMPARISON OF TWO TYPES OF DDS**

The traditional type of DDS is shown in Fig. 1. A desired frequency word Y chosen to be a fraction less than unity, is added successively to itself at clock rate  $f_c$  in the accumulator, which overflows if its contents equal or exceeds one. The remainder  $R_n$  from the accumulator corresponds to a phase between 0 and 1 cycle or between 0 and  $2\pi$  radians. The waveform memory maps each remainder a word in this case into sine wave samples which are fed to the DAC. The DAC acts as a sample hold circuit and its output is then filtered in the low pass filter LPF. The output frequency is  $f_o = Yf_c$ .

The total noise power of the spurious signals (normalised to one ohm unit resistance) is the same as the quantiaation noise which can be shown to be  $q^2/12$  where  $q$  is the amplitude of the last significant bit. The assumption is that the error pulses have amplitude which are random and evenly distributed in amplitude between  $-q/2$  and  $+q/2$ .

The noise power of a series of random amplitude impulses with a pulse rate of  $f_c$  is taken to be evenly

spread over the frequency range from zero to  $f_c/2$ . Thus the noise power spectral density (PSD) of the sampler part of the DAC sample hold output circuit is taken to be

$$N(f) = (q^2 / 12) / (f_c / 2) = q^2 / 6f_c \dots\dots\dots 2.1$$

The output quantisation noise pulses are in fact flat topped each with duration  $T_c = 1/f_c$  lasting until the next pulse arrives. It is this finite duration of the pulses which causes the noise spectrum to be divided into a part which causes AM (amplitude) noise and a part which causes PM (phase) noise when added to the digitally generated output carrier signal at  $f_o$ . Further analysis can support the proposition that the sample and hold action which has to exist at the output of DAC divides the sampler noise into in-phase components which cause AM and quadrature components (each corresponding to a  $\pi/2$  phase shifted component) which cause PM noise.

The frequency response of the DAC output zero order hold (ZOH) action is derived by putting  $s=j\omega$  into the ZOH transfer function and it is

$$H(j\omega) = (1 - \exp(-j\omega T_c)) / j\omega T_c \dots\dots\dots 2.2a$$

$$= \exp(-jx) \text{Sin } x/x \dots\dots\dots 2.2b$$

$$= (\text{Cos } x + j \text{Sin } x) \text{Sin } x/x \dots\dots\dots 2.2c$$

where  $x = \omega T_c / 2$ . The proposition is that the  $\text{Cos } x$  term when squared is the proportion of AM noise and the  $j \text{Sin } x$  term after multiplying by its complex conjugate  $-j \text{Sin } x$  is the proportion of PM noise at and around any output frequency  $\omega_o = 2\pi f_o = \omega$ . The proof of this requires equally spaced upper and lower sideband components to be combined with the output carrier at  $\omega_o = 2\pi f_o$ .

Further inspection shows that only the quadrature ( $j \text{Sin } x$ ) components have the correct phase relationship with the carrier for PM modulation to occur. Similarly the in-phase ( $\text{Cos } x$ ) components can be shown only to cause AM.

The ZOH frequency response has an equivalent power frequency response  $H(f) = H(j\omega)H(-j\omega)$  and at low frequencies  $\text{Sin } x = x$ .

The phase noise can be shown to be from equations 2.2 and 2.1.

$$N(f) = (q^2 / 6f_c) \times \text{Sin}^4 x / x^2 \dots\dots\dots 2.3a$$

$$\approx (q^2 / 6f_c) \times x^2 \dots\dots\dots 2.3b$$

$$= q^2 \pi^2 T_c^3 / 6T_o^2 = q^2 \pi^2 f_o^2 / 6f_c^3 \dots\dots\dots 2.3c$$

where the approximation

$$\text{Sin } x \approx x = 2\pi f_o T_c / 2$$

has been used.

The normalised phase noise is defined as  $S_{\phi}(f) = N(f) / P$  where P is the power of the output carrier at  $f_o$ . For a sine wave with peak to peak amplitude just within the n-bit DAC limits the normalised carrier power is

$$P_c = (q \times 2^{n-1})^2 / 8 \dots\dots\dots 2.4$$

giving the normalised phase noise as

$$S_{\phi}(f)_c = N(f) / P$$

$$= (4\pi^2 / 3) \times (f_o^2 / f_c^3) / (2^{n-1})^2 \dots\dots\dots 2.5$$

It should be noted that the phase noise can be improved by about  $16/\pi^2$  or 2.1dB if the carrier signal is programmed to be a square wave. This is because the amplitude of the component at  $f_o$  is increased by  $4/\pi$  in this case. For the square wave we have

$$S_{\phi}(f)_s = (\pi^4 / 12) \times (f_o^2 / f_c^3) / (2^{n-1})^2 \dots\dots\dots 2.6$$

The delay compensated DDS is the second main type. A delay modulator is used to cancel the time jitter of a digitally generated square wave. The delay compensation values are computed either digitally or as will be seen in the next section they can be extracted from the digitally generated square wave itself.

Fig. 2 shows a delay compensated DDS using the successive addition rate multiplier (SARM) technique. This is the same method as used in the accumulator in the traditional DDS shown in Fig. 1 but with the accumulator overflow providing a square wave pulse output signal of the correct nominal frequency. The remainder value from the accumulator, which in the traditional waveform DDS is used to provide phase samples to a read only memory (ROM) to generate the DDS waveform, is used to generate the delay correction values for each output pulse. A disadvantage of this circuit is that the remainder values associated with each pulse have to be scaled in amplitude inversely to the output frequency. This is achieved by DAC + Divider circuitry. Either a digital divider can be used before the DAC (but this is power

consuming) or an analogue divider after the DAC can be used (accurate analogue dividers are not easy to make).

The delay compensated DDS in Fig. 3 uses the rate multiplier in a feedback loop where the pulse subtractor acts as a frequency subtracting summing junction. In this way a fractional digital divider is created. The advantage of this scheme is that no frequency compensation of the delay correction signals is required. It is then possible to maintain an accurate cancellation of the time jitter over a wide frequency range of operation. The disadvantage of this scheme is that frequency steps are proportional to the output frequency as the divider ratio is changed in equal steps.

The desired output frequency value has to be turned into a reciprocal with multi-bit high precision in order to achieve the equivalent operation to a simple synthesiser with a fixed frequency step size.

The noise performance of this second type of DDS can be estimated by converting the residual time jitter rms amplitude  $t_j(f)$  into rms phase jitter  $\phi(f)$  by

$$\phi(f) = 2\pi f_o t_j(f) \dots\dots\dots 2.7$$

where  $f_o = 1/T_o$  is the output frequency. Phase jitter

normalised power  $\phi^2(f)$  can be converted to an equivalent normalised phase noise spectrum by

$$S_{\phi}(f) = N(f) / P_o = \frac{1}{4} \phi^2(f) \cdot (P_{peak} / P_o) \dots\dots\dots 2.8$$

where  $P_o$  is the carrier power at the output frequency  $f_o$  and  $P_{peak}$  is the peak power of the square wave, which is in fact constant.

By application of equations 2.8 and 2.7 it can be shown that the delay compensated DDS with n-bit precision delay compensation has exactly the same phase noise performance as the traditional waveform DDS with an n-bit output DAC and set to generate square waves. The phase noise is defined by equation 2.6. This is 2.1dB better than an equivalent DDS generating a sine-wave output.

In order to simplify the delay compensated type of DDS, techniques have been originated and developed where the delay compensation information is extracted from the square wave time jitter directly by essentially analogue non digital means. These techniques are described in the next section.

### 3. THE SELF ADJUSTING DELAY COMPENSATOR (SADC)

It is reasonable to suppose that all the information

required to compensate the time jitter of any waveform is actually contained within the waveform itself. The "analogue phase predict" method (4) was devised to extract time jitter from pulse waveforms originating from rate multipliers. The important characteristic of these waveforms are that they consist of pulses all of the same height and duration but with variable pulse to pulse spacing. The average pulse repetition rate is the required output frequency. It is easy to show that the mean dc level of such a waveform is directly proportional to frequency.

This is the same principle as is used in the pulse rate type of FM demodulator. Less obvious is the fact that the mean level averaged over any pulse period is a sampled instantaneous measure of the frequency of the waveform. An integrator is then required to convert frequency errors into phase errors. (The definition of frequency is the rate of change of phase). After integration and re-sampling the phase error signals can be fed to the delay modulator for appropriate cancellation of the time jitter.

The self adjusting delay compensator (SADC) as shown in Fig. 4 is an improvement on this analogue phase predict method. Essentially the integrator required to convert frequency errors to phase errors is combined into the same integrator as produces the voltage ramp waveform required to convert voltage to time shift in the time delay modulator.

The dc removal circuit (DCR) is essential to prevent saturation of the integrator. Below the cut-off frequency of the DCR, phase noise is not detected and cannot therefore be cancelled. However, if this cut-off frequency is too low the SADC integrator will take a long time to recover from the saturation caused by a step change of input frequency. In practice it can be useful to provide a dc voltage (from a DAC) which is proportional to frequency and this can be used to cancel the dc level shift in the input to the DCR when the frequency is changed.

Given that the dc input component is correctly removed the operation of the SADC can be seen by reference to Fig. 5. Fig. 5b shows the effect on the integrator output waveform of time jitter occurring on the central pulse of the three pulses shown. In practice all pulses would be expected to be suffering from time jitter with the same probability distribution. From Fig. 5b it can be seen that the times  $t_1$  and  $t_5$  where the comparator is switched on the negative slope of the integrator waveform both are completely unaffected by the time jitter  $t_m$  on the central pulse. This is provided that the start time of the pulse  $t_x$  lies between  $t_1$  and  $t_4$  where by inspection  $t_4 - t_1$  can be seen to be  $T_0 - T_p$  where  $T_0$  is

the average period of  $f_0$  and  $T_p$  is the pulse length. The independence of  $t_4$  and  $t_5$  of the time jitter confirms that the circuit has cancelled the time jitter and phase noise.

The term "self adjusting" reflects the fact that if the dc is removed the up and down slopes of the integrator waveform automatically adjust to the correct value for jitter compensation and cancellation. By the use of the same slope for both detection and compensation of the time jitter the cancellation gain is automatically ensured to be the correct value.

A very important feature of the SADC is that it has a phase noise cancellation bandwidth equal to the output frequency  $f_0$ . Its ability to cancel frequency fluctuations is approximately equal over the whole bandwidth. Broadband phase noise is if anything better cancelled than close in phase noise. Also care must be taken in detailed circuit design and layout because the presence of the integrator means that the circuit is more sensitive to any low frequency circuit noise.

Fig. 6 shows that by the addition of two fixed pulse length monostables the SADC can be used to remove the time jitter and phase noise from any frequency source. The monostable pulse lengths should both be approximately equal to half the period  $T_0$  of the output frequency  $f_0$ . If a wide frequency range is to be covered then the  $T_p$  should be just shorter than the highest frequency period value. In practice the second monostable can replace the comparator of the SADC provided the triggering levels are appropriately adjusted.

From Fig. 6 it should also be obvious that two or more SADCs can be cascaded to provide successive reductions of phase noise up to the limit of capability of the final SADC. This can ease the detailed design requirements of each SADC. This is important because it is difficult to design an integrator which operates well both at the output frequency  $f_0$  and at the lowest phase noise frequency to be cancelled.

#### 4. RESULTS AND CONCLUSIONS

To prove the technique an SADC was constructed to operate at input frequency of up to two megahertz.

Fig. 7a and 7b show the before and after spectra of a noisy 2 MHz input signal. Because a divide by two circuit was used for convenience on the output the Fig. 7b spectrum is at 1MHz. The noise level reduction is actually 6dB less than the difference between the two pictures, because the division of a

frequency source by a factor two gives a 6dB reduction. Thus the actual phase noise reduction shown is of the order of 30dB over a broad band of frequencies.

Fig.8a and 8b show the results on a square-wave modulated source simulating DDS spuri. In this case the before (8a) and after (8b) spectra are on the same 2MHz frequency (taken in both cases at the divide by two output).

These results clearly indicate the potential of the SADC phase noise reduction circuit. Further work is to be undertaken to extend operation to higher frequencies. The expectation is that the power consumption of the SADC will remain considerably lower than a DAC operating at the same frequency and certainly lower than a DAC operating at the system clock frequency.

In conclusion it is felt that the new phase noise reduction technique described here is a potentially very important way of producing direct digital frequency synthesisers with low phase noise and very low spurious levels even when operated at high frequencies.

#### ACKNOWLEDGEMENT

The authors would like to acknowledge the work of Paul Jenkins in his investigations of the analogue phase predict type of synthesiser in his final year undergraduate project in 1993.

#### REFERENCES

1. Manssowitz V. 1980. "Frequency synthesisers, theory and design".  
Second Edition, Wiley.
2. Saul PH, 1990, "Direct frequency synthesis - a review of the technique and potential",  
Radio Receivers and Associated Systems, IEE Conf Pub, 325, 5-9
3. Underhill MJ. Scott RIH, 1983. "Frequency synthesiser", UK Patent No. GB2062315B
4. Underhill MJ. Scott RIH, 1983 "Frequency synthesiser", European Patent application (UK) No. 0089721.
5. Underhill MJ. 1996 "Phase noise reduction circuit", UK Patent application 9603223.0.

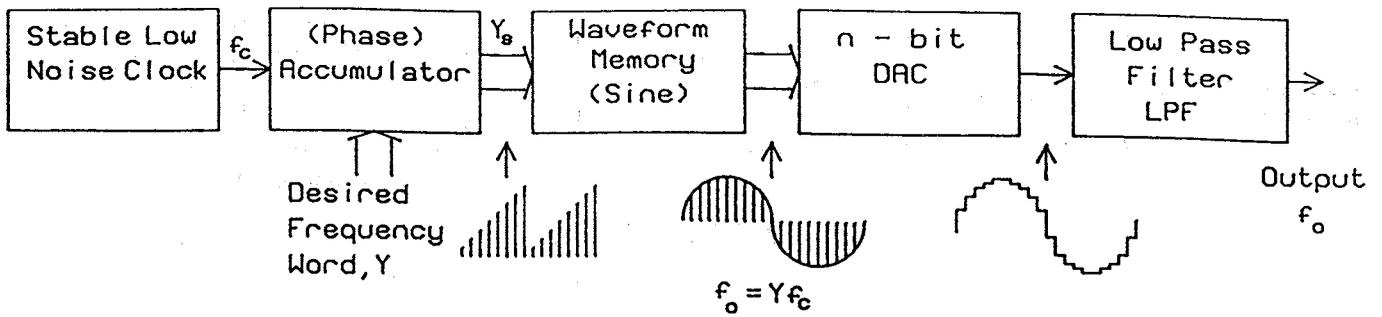


Fig 1. Traditional Direct Digital frequency Synthesiser (DDS)

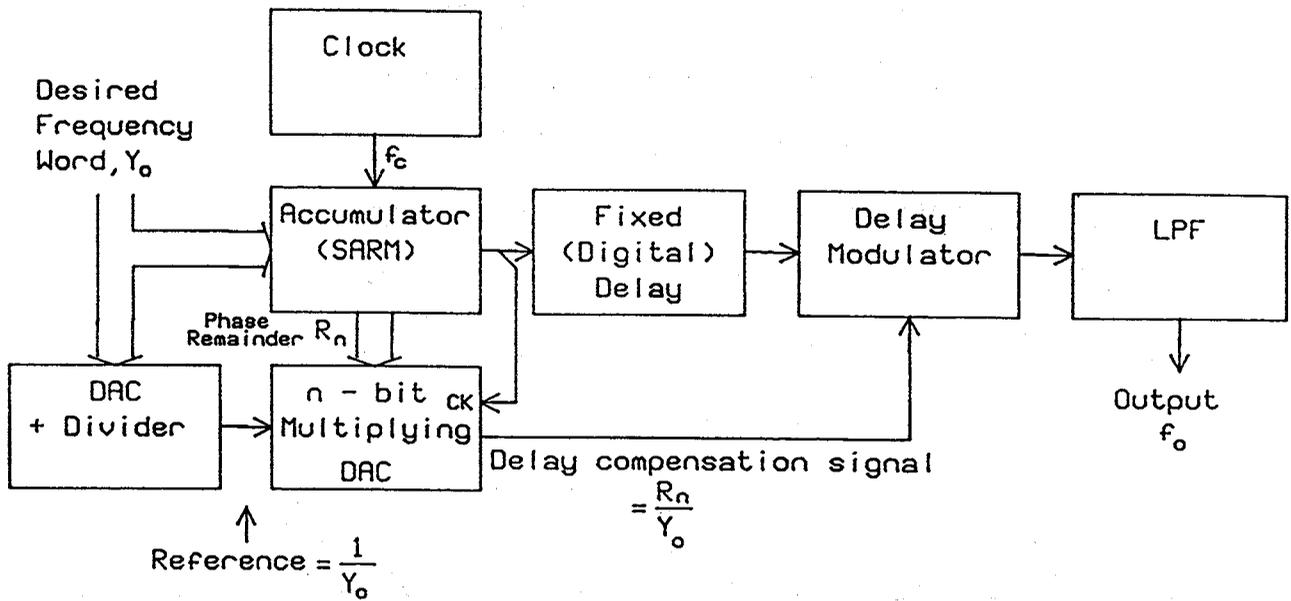


Fig 2 Delay Compensated rate multiplier DDS

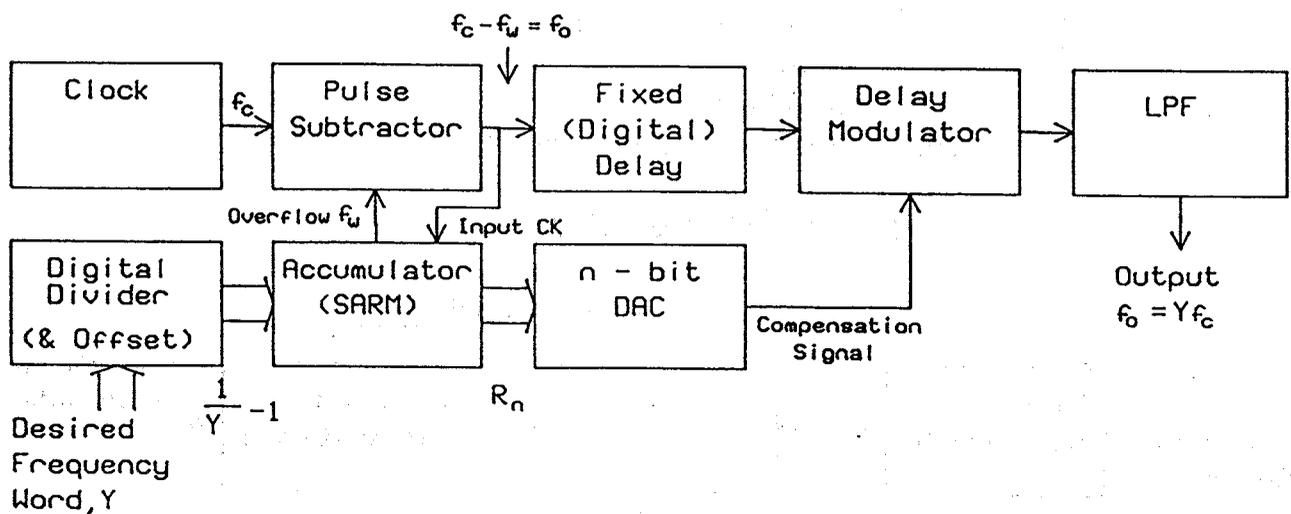
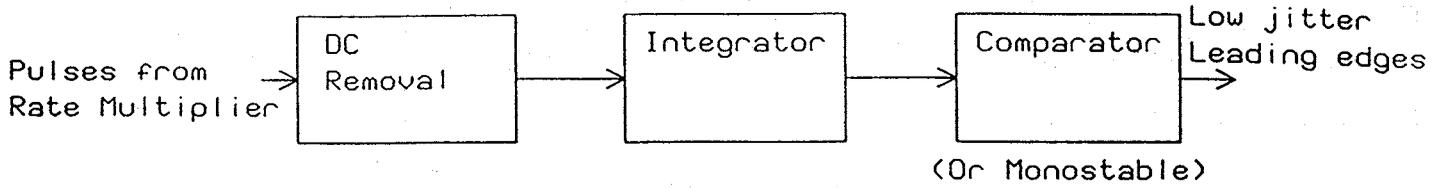
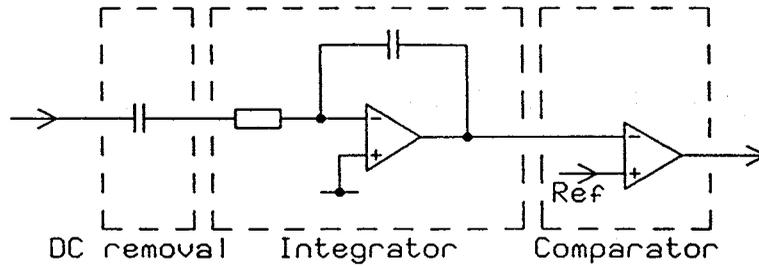


Fig 3 Delay Compensated rate divider DDS

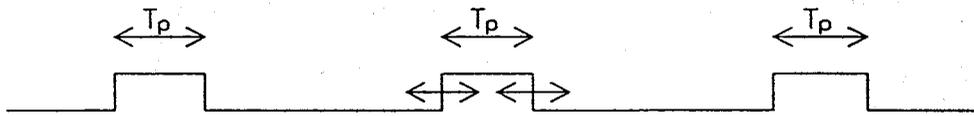


(a) Block diagram

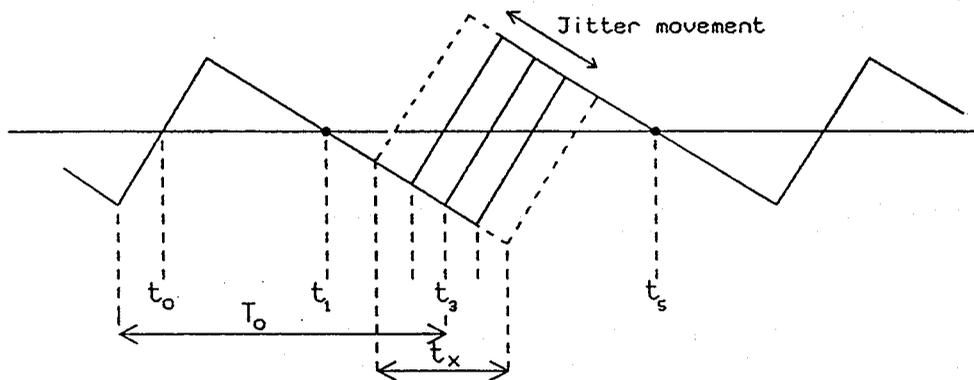


(b) Representative circuit

Fig 4 Self Adjusting Delay Compensator (SADC)



(a) Input waveform



(b) Integrator output waveform

Fig 5 SADC Waveforms

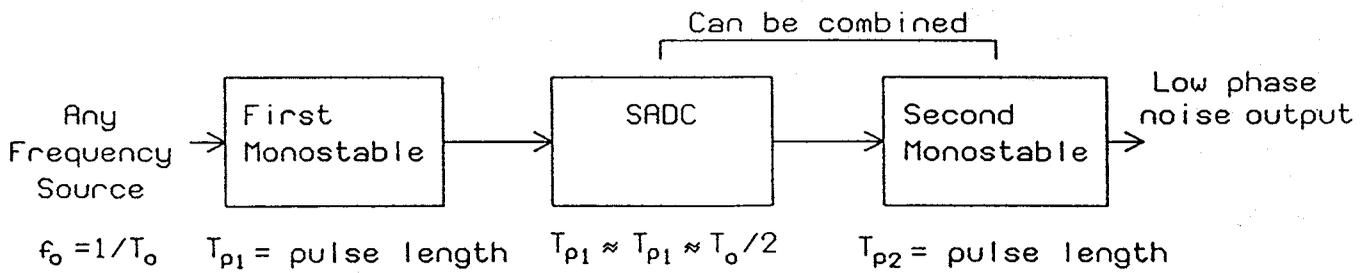
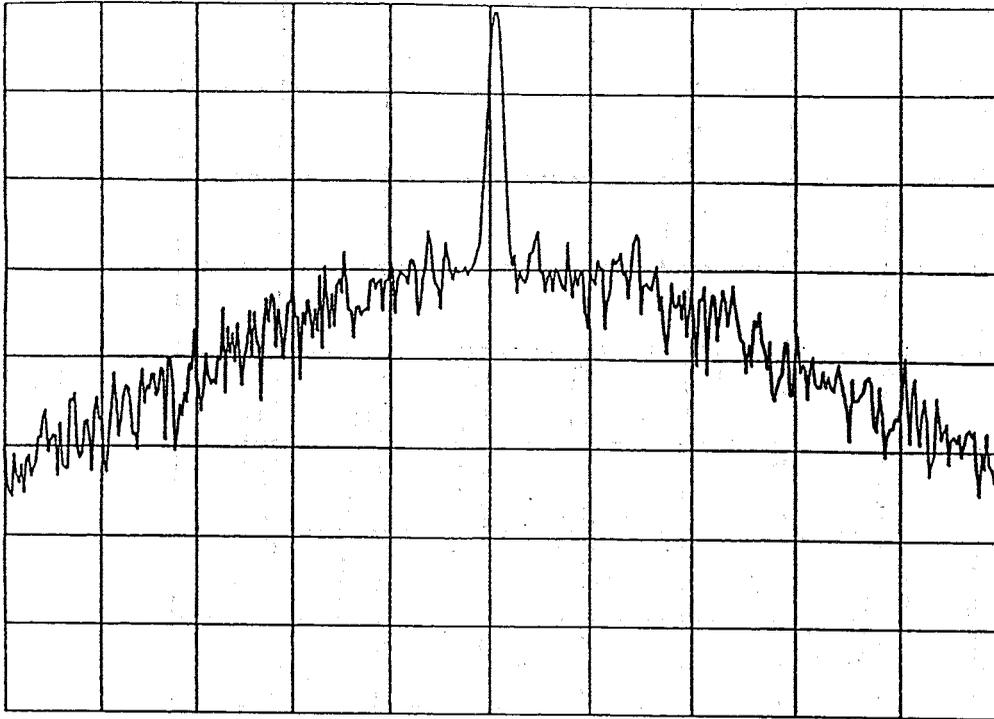


Fig 6 Time jitter and phase noise cancellation on any frequency source.

hp

REF -32.0 dBm AT 10 dB

PEAK  
LOG  
10  
dB/

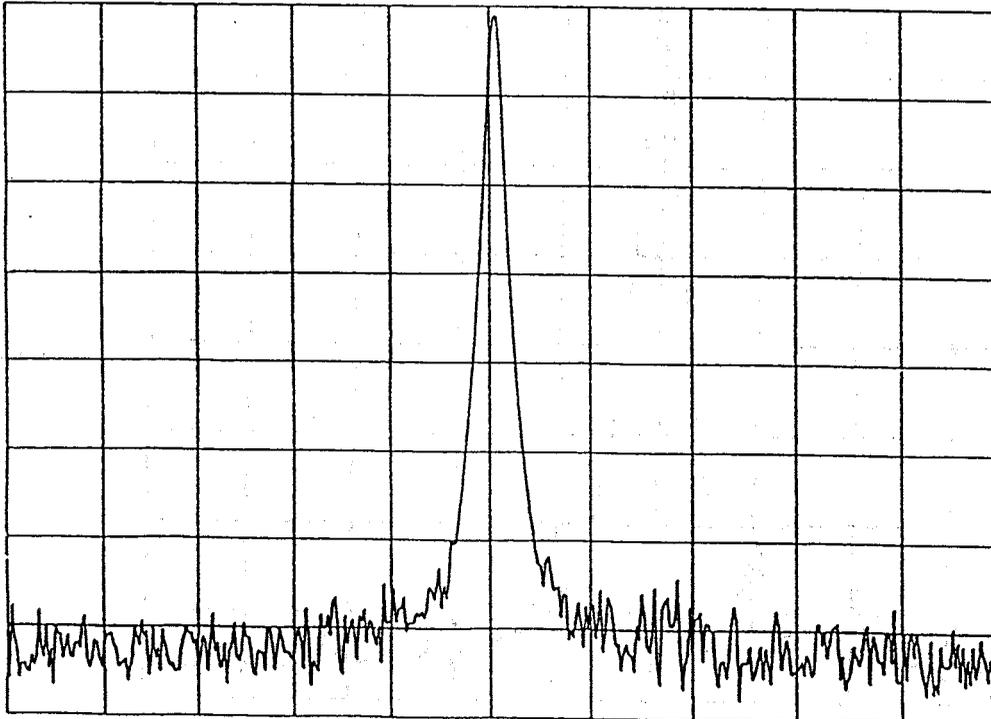


(a) : Input

hp

REF -32.0 dBm AT 10 dB

PEAK  
LOG  
10  
dB/

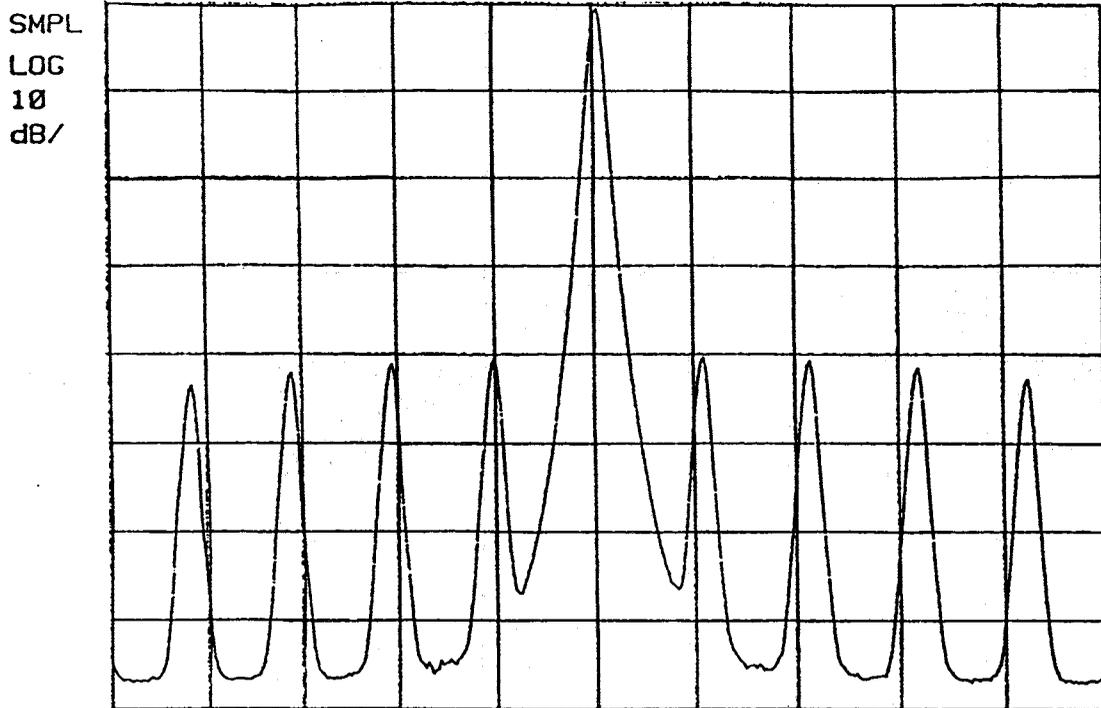


(b) : Output

Figure 7 : Phase noise reduction

hp

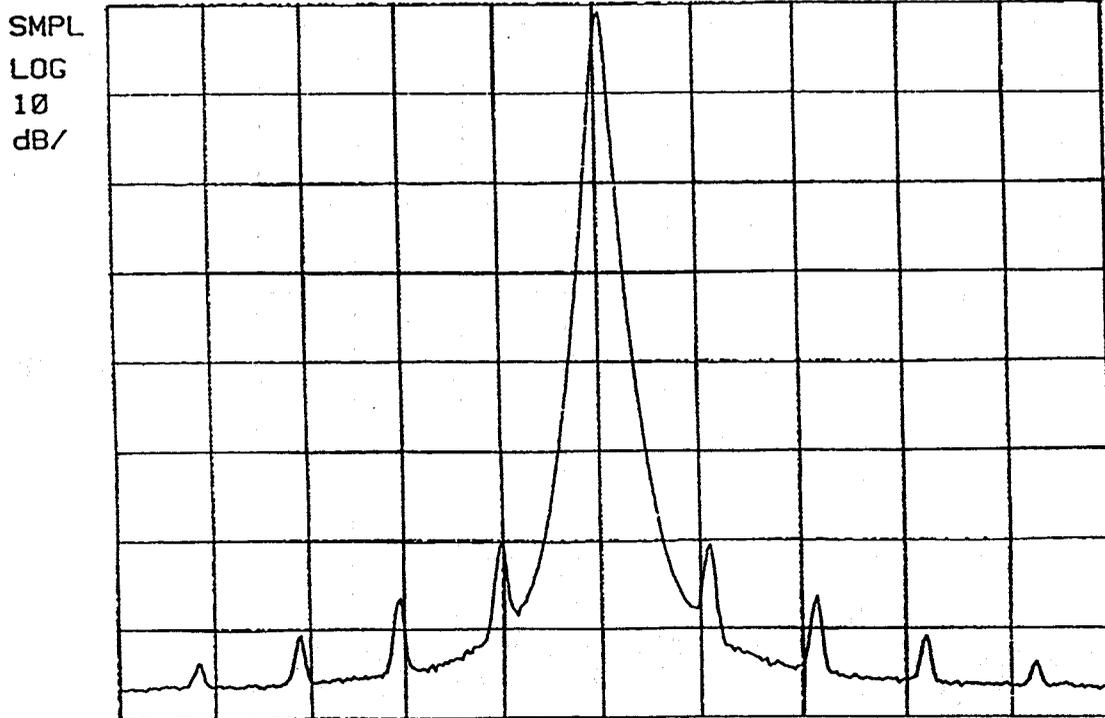
REF -32.0 dBm AT 10 dB



(a) : before

hp

REF -32.0 dBm AT 10 dB



(b) : after

Figure 8 : Spurious reduction

TWO-WAY TIME TRANSFER THROUGH SDH AND SONET SYSTEMS<sup>1</sup>

S. R. Jefferts, Marc A. Weiss, J. Levine, S. Dilla and T. E. Parker  
 N.I.S.T. Time and Frequency Division  
 325 Broadway  
 Boulder, CO 80303  
 U.S.A.

**ABSTRACT**

NIST has built a two-way time transfer device which uses any currently unused byte in the SONET overhead to effect time transfer. The hardware shows stability which allows time transfer at approximately 100 ps. Accuracy at the same level should also be possible.

**INTRODUCTION**

NIST has developed a system which allows us to measure the stability of time transfer over a SONET optical link. The SONET protocol for data transmission is well established and, along with the growing need for improved synchronization in telecommunications systems, promises a vehicle by which improved synchronization can be achieved relatively inexpensively and robustly. Work on this type of system was begun by Kihara [1] who reported on two-way time transfer via SDH/SONET over a greater than 1000 km baseline. The system uses a single overhead byte in each SONET frame to transfer timing data from the remote clock, as well as providing an on time marker (OTM) which is used to transfer the actual time.

**TWO-WAY TIME TRANSFER**

Two-way time and frequency transfer is generally used to compare two geographically separated clocks. The clocks at each end of a link which joins them transmit the time of the local clock and simultaneously receive the time of the remote clock. Each clock then measures the difference between the local clock and the remote clock. If the time difference data from the remote clock are differenced with the data from the local clock, then the path delay effects are removed, assuming that the path from the remote clock to the local clock is reciprocal with the path from the local clock to the remote clock. The time of the remote clock (clock 2) relative to

the local clock (clock 1) can then be written as [2]:

$$\text{Time}(1) - \text{Time}(2) = 1/2\{(\text{TIC}(1) - \text{TIC}(2)) + (\text{T}_{\text{xdelay}}(1) - \text{R}_{\text{xdelay}}(1)) - (\text{T}_{\text{xdelay}}(2) - \text{R}_{\text{xdelay}}(2))\}$$

where

TIC(i) is the Time Interval Counter Reading for System i

T<sub>xdelay</sub>(i) is the transmit delay for system i

R<sub>xdelay</sub>(i) is the receive delay for system i.

Accurate time transfer requires that the absolute magnitudes of the delays, T<sub>xdelay</sub> and R<sub>xdelay</sub>, associated with the hardware on each end of the link be known and that those delays be temporally invariant. Accurate frequency measurements, however, require only that the delays be stable; the magnitudes need not be known. In the present experiment we are attempting to measure the temporal stability of these delays, which in general have instabilities associated with the environment (for example, temperature) as well as delays associated with the digital hardware which may differ from one reset cycle of the equipment to the next.

**RESULTS**

The basic hardware is diagramed in Figure 1. SONET overhead access is provided by the SONET Interface adapter[3]. This device, built around a framer chip [3] provides buffered access to both the received and transmitted SONET overhead. Start and stop commands for the time interval counter

<sup>1</sup> Contribution of U.S. government not subject to copyright

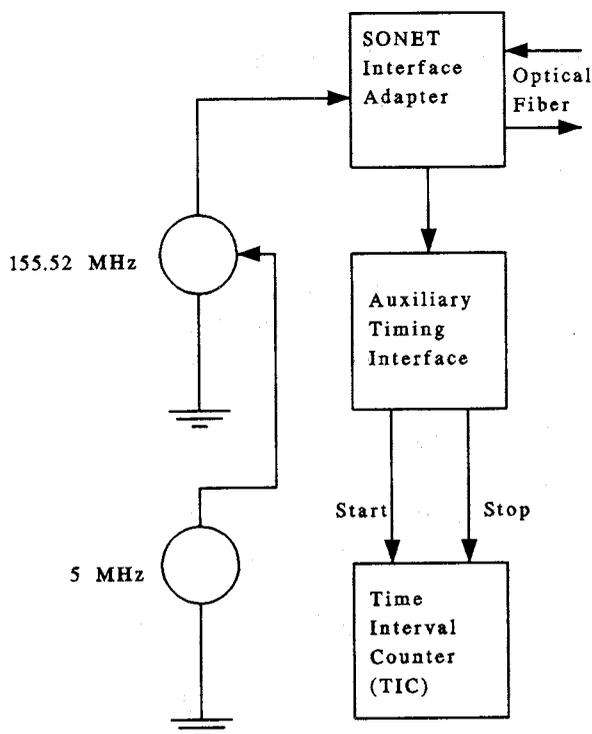


Figure 1. - Simplified block diagram of the SONET two-way time-transfer system.

(TIC) are generated by the auxiliary timing board and overall system control is provided by the controller board.

In the first test, the system was configured for loop-back, as shown in Figure 2. This configuration allows the measurement of the quantity  $T_{x\text{delay}} + R_{x\text{delay}}$

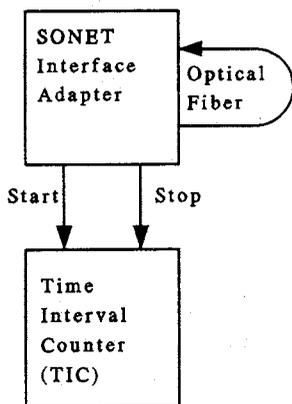


Figure 2. - SONET two-way time transfer system in loop-back mode. In this configuration the system measures the time between transmission and receipt of its own OTM.

combined with the delay associated with the fiber. The fiber used in this test is very short, about 15 cm, and is not expected to be a significant source of instability. The stability achieved using the configuration of Figure 2 is shown in Figure 3; the hardware stability exhibits an approximate flicker phase noise floor less than 10 ps, which is consistent with the flicker floor of the time interval counter used. Both of the two systems used in this experiment are essentially identical and exhibit similar results in this loop-back test.

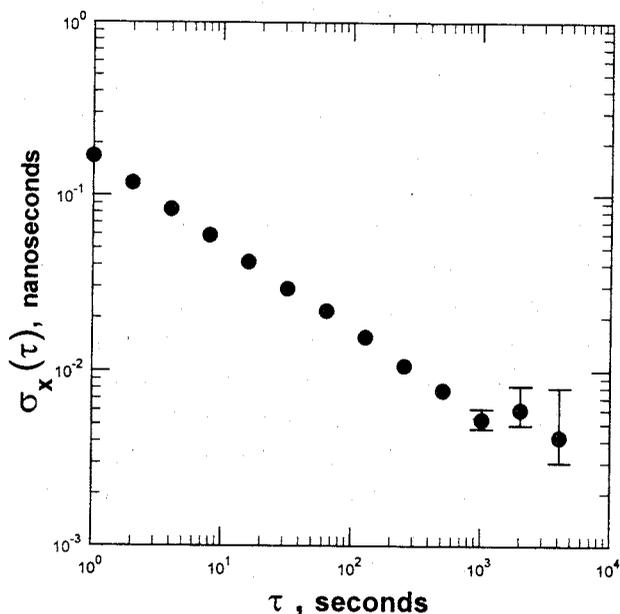


Figure 3. - Stability of the SONET two-way time transfer system in loop-back mode as described in Fig. 2.

A full two-way test using 30 m of twin lead fiber was also conducted. In this test, the two ends of the link were physically situated in the same environment, and the fiber was coiled between them. Further, the two SONET interfaces were driven from a common 155.52 MHz clock. The configuration is shown in Figure 4. The resulting data, Figure 5, while clearly less stable than the data from the loop-back test, still exhibit stability of less than 200 ps at measurement times greater than 10 s, thereby allowing frequency transfer of  $10^{-11}$  at times greater than 100 seconds.

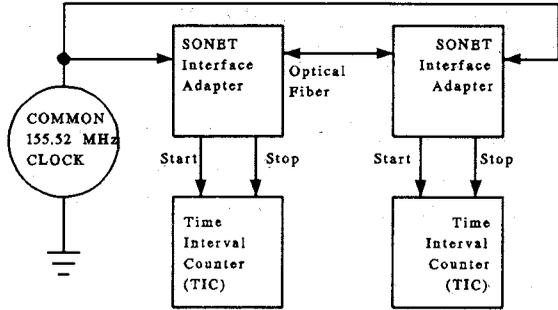


Figure 4. - Common clock two-way time transfer system designed to measure the stability of the hardware delays in the SONET two-way time-transfer system. The common clock arrangement removes clock noise from the measurement.

Assuming the stability is not degraded by increasing distance, the observed stability should allow frequency transfer better than 1 in  $10^{-15}$  at 1 day over distances on the order of 100 km. This assumption will probably not, in general, be valid in a public telecommunications system, but may be possible in a private network, where path reciprocity can be engineered into the system.

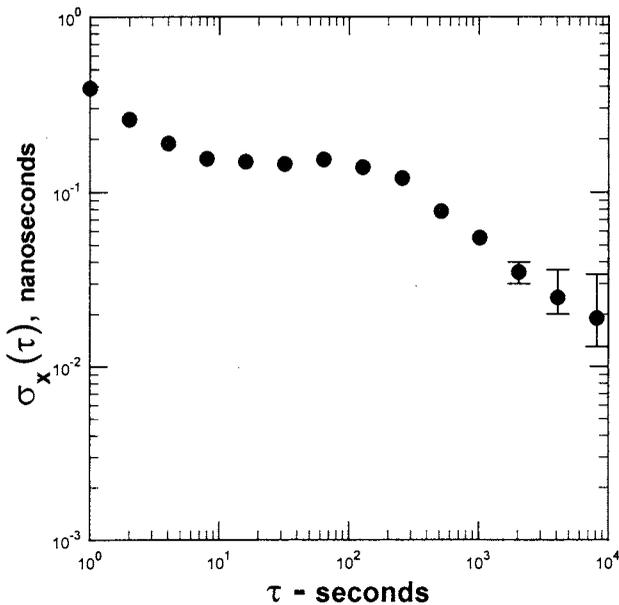


Figure 5. - The stability of the SONET two-way time transfer system in the common clock two-way mode described in Figure 4. The stability is better than 200ps at 10 s.

A preliminary study of the stability of two-way time transfer vs. byte position in the overhead was also undertaken. The data, shown in Figure 6, show the measured stability vs. byte position for the three byte positions 11, 17, and 26 in the overhead frame. The stability does not seem to be a function of the byte position for this particular hardware configuration. A very preliminary study of the stability of the time transfer process vs. temperature was also undertaken.

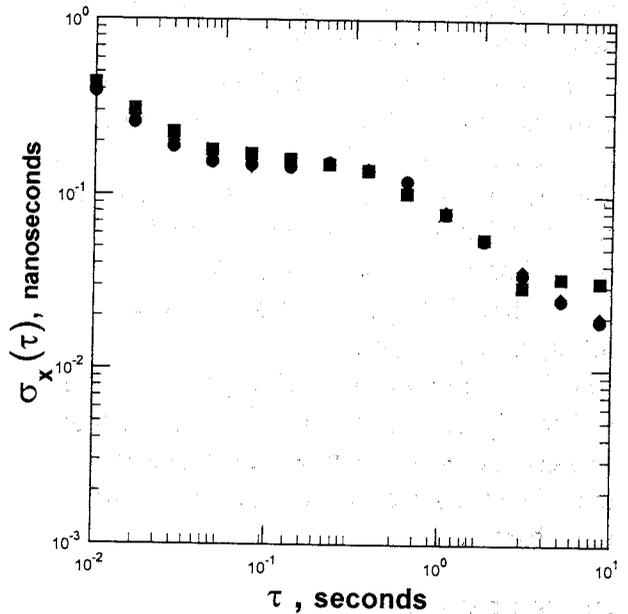


Figure 6. - The stability of the SONET two-way time transfer system as a function of byte position in the overhead. The solid circles are the stability measured for byte position 17, the solid diamonds are the stability measured for byte position 26 and the solid boxes are the stability measured for byte position 11.

Temperature coefficients measured in the loop-back mode of operation measure the quantity  $T_{x\text{delay}} + R_{x\text{delay}}$ , while temperature coefficients measured in the two-way mode are sensitive to the quantity  $T_{x\text{delay}} - R_{x\text{delay}}$ . The temperature coefficients for this equipment using the above technique are  $\delta T_{x\text{delay}}/\delta T \approx 3 \text{ ps}/^\circ\text{C}$  and  $\delta R_{x\text{delay}}/\delta T \approx 44 \text{ ps}/^\circ\text{C}$ .

## MESSAGE FORMAT

The message format developed for this experiment uses only one byte of the overhead and transmits information to the remote end of the link in a byte serial format. The message format supports messages for the transmission of both absolute and relative time and frequency as well as administrative functions. The high-order bit of all bytes is 0, with the exception of the first byte of a block, where the high-order bit is set to 1, thereby identifying the start of a block. The next 7 bits of the first byte of a word are the word identifier, allowing the system to identify the type of data to follow. The first byte signals the start of a block, can act as an on-time-marker (OTM), and can be used as an idle character. The word format is such that all quantities are 32 bit unsigned numbers sent as 5 bytes, with the most significant bit first. The format thus looks like:

$$Scccd_0d_1d_2d_3, Sd_4d_5d_6d_7d_8d_9d_{10}, \dots,$$

$$Sd_{25}d_{26}d_{27}d_{28}d_{29}d_{30}d_{31}.$$

The first 4 bits identify the type of number, and the dynamic range is  $2^{32}$ . A variety of message types are defined, of which this experiment uses only two, the OTM and the idle character.

## CONCLUSION

Preliminary measurement of the stabilities achievable using two-way time transfer over a SONET/SDH OC-3 link at 155.52 MHz are on the order of 100 ps. The measurements presented here are measurements of the stabilities of the delays in the particular SONET terminals used and represent a lower limit on the stabilities with which this hardware could perform two-way time transfer in a real world situation. Further characterization of environmental sensitivities may, possibly, allow time transfer at the 10 ps level demonstrated by a single SONET/SDH device.

## BIBLIOGRAPHY

[1] M Kihara and A. Imaoka, *Proc. of the 9<sup>th</sup> E.F.T.F.*, Besancon 1995.

[2] C. Hackman, S. R. Jefferts and T. E. Parker in *Proc. of the 49<sup>th</sup> Annual Symposium on Frequency Control*, 1995, pp. 275-281.

[3] SONET Interface Adaptor - Odetics Inc. LIMO SONET Interface Adaptor. Frammer Chip PMC- Sierra Inc. The commercial equipment used has been identified for technical completeness only, to allow other researchers to duplicate the results contained herein. Other commercial equipment may perform differently, in particular it may be more or less stable than the equipment described herein. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology.

## SYSTEM CONFIGURATION FOR STANDARDIZING SDH-BASED TIME AND FREQUENCY TRANSFER

Masami Kihara and Atsushi Imaoka

NTT Optical Network Systems Laboratories  
1-2356 Take, Yokosuka-shi, Kanagawa, 238-03 Japan  
Tel:+81-468-59-3164 Fax:+81-468-55-1282  
e-mail:kihara@exa.onlab.ntt.jp

### ABSTRACT

By using the Network Node Interface (NNI) and User-Network Interface (UNI) in the Synchronous Digital Hierarchy (SDH), time and frequency transfer can be realized. The transfer system uses a bi-directional transmission medium instead of the conventional uni-directional medium for reference signal distribution, and can use the information channel called Section OverHead (SOH) to transfer time and frequency without interfering with B-ISDN services. The combination of a single fiber cable and the wavelength division multiplex method, which can be applied between network nodes, permits accurate time dissemination of the order of subnano seconds.

### 1. INTRODUCTION

Distributed time information is a promising tool for timestamping in distributed network operation systems, for fault detection and correction in network systems, and some academic applications[1][2][3]. Telecommunication networks are suitable for transferring time and frequency. The time and frequency transfer network is composed of transmission systems between network nodes and between network nodes and Customer Premises Equipment (CPE). These network systems now use the same SDH-based interface. In the SDH-based time and frequency transfer system, time accuracy and stability depend on transmission system configuration.

ITU-T and ITU-R are now discussing systems for time and frequency transfer[4][5]. Standardization is indispensable since time and frequency are basic elements for global telecommunications. Our purpose is to introduce our new concept and trigger new ITU recommendations. We must coordinate time and frequency functions with those emerging in telecommunication networks and to prevent the adoption of specifications on time and frequency transfer that are incompatible.

This paper describes a basic principle, system configuration, and the transfer format of the new standardized system.

### 2. BASIC PRINCIPLE

In the SDH based transfer system, either the master-slave configuration or the two-way configuration can be selected as the transfer method.

In the two-way system, time information is exchanged between two nodes and the time and frequency sources in the nodes should be carefully configured. Since networks

have their own time and frequency sources, we have to consider the relationship between the network and the two-way sources. This paper however does not consider two-way time transfer.

In the master-slave configuration, time information, as well as a reference clock signal, is distributed node by node from a master to slave nodes and from a higher slave node to lower slave nodes. Time information can be regenerated at a lower node without any time difference being caused by transmission delay, with the processes of transmission delay compensation and time information transfer.

#### 2.1 Transmission delay compensation process

Figures 1 and 2 show the basic configuration for transmission delay compensation and phase relationship among reference pulses, respectively. A reference pulse (Fig.2 (a)), generated by a time information generator at a higher node working as a master, is transmitted to a lower node working as a slave (Fig.2 (b)). The transmitted reference pulse is returned to the higher node (Fig.2 (c)). The returned reference pulse is delayed by round trip transmission through the outgoing and incoming transmission lines. Transmission delay is measured by a time interval counter at the higher node. The practical delay data is one half the measured delay data if transmission lines pass through the same medium. The delay data is transmitted to the lower node (Fig.2 (b) and (c)). The lower node executes delay compensation corresponding to 'pulse period minus delay data', and can regenerate the reference pulse synchronized to the higher node (fig. 2 (d)).

#### 2.2 Time information transfer process

The time information transfer process shares the time information that allows matching to the original time information. This process is executed based on the transmission delay compensation process described in 2.1. There are three steps in this process as shown in Fig.3. First, the original time information is generated with the original reference pulse at the higher node (Fig.3 (a)), and then transmitted to the lower node (Fig.3 (b)). The lower node executes the transmission delay compensation process, and then regenerates the compensated reference pulse from the received reference pulse. The compensated reference pulse is in synchronization with the original reference pulse. If we select the pulse period of one second, we have to correct the time information received with the reference pulse, so that the time information, which is regenerated with the compensated reference pulse, is advanced one second (Fig.3 (c)). The higher node time in-

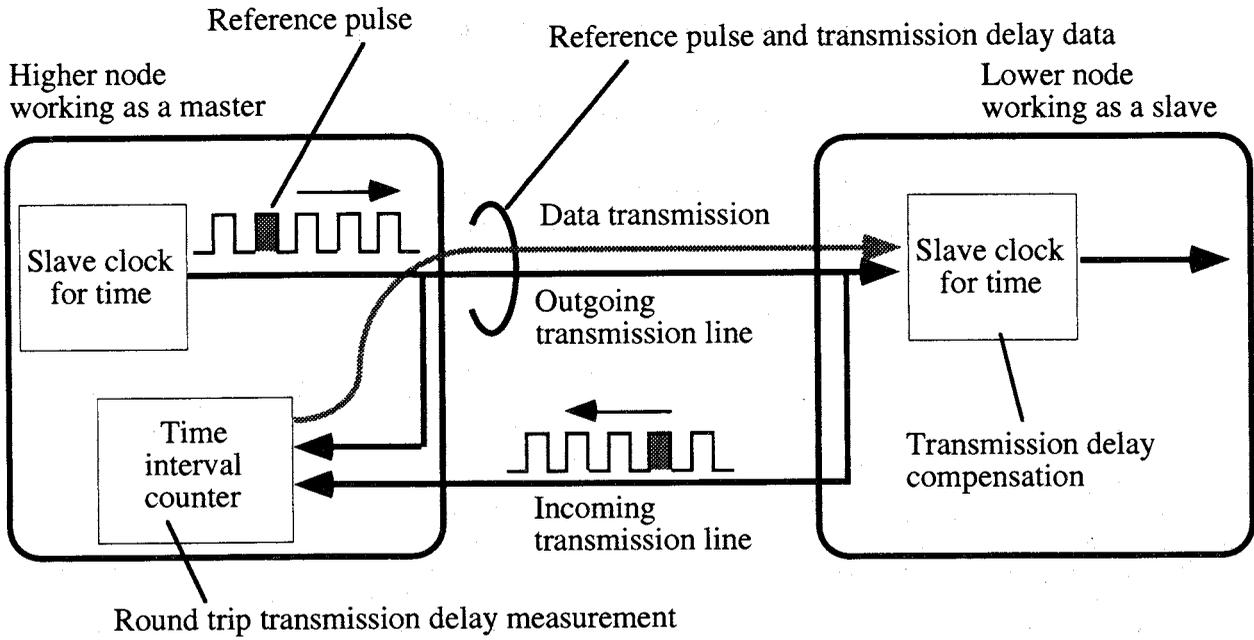


Fig. 1 - Principle of transmission delay compensation

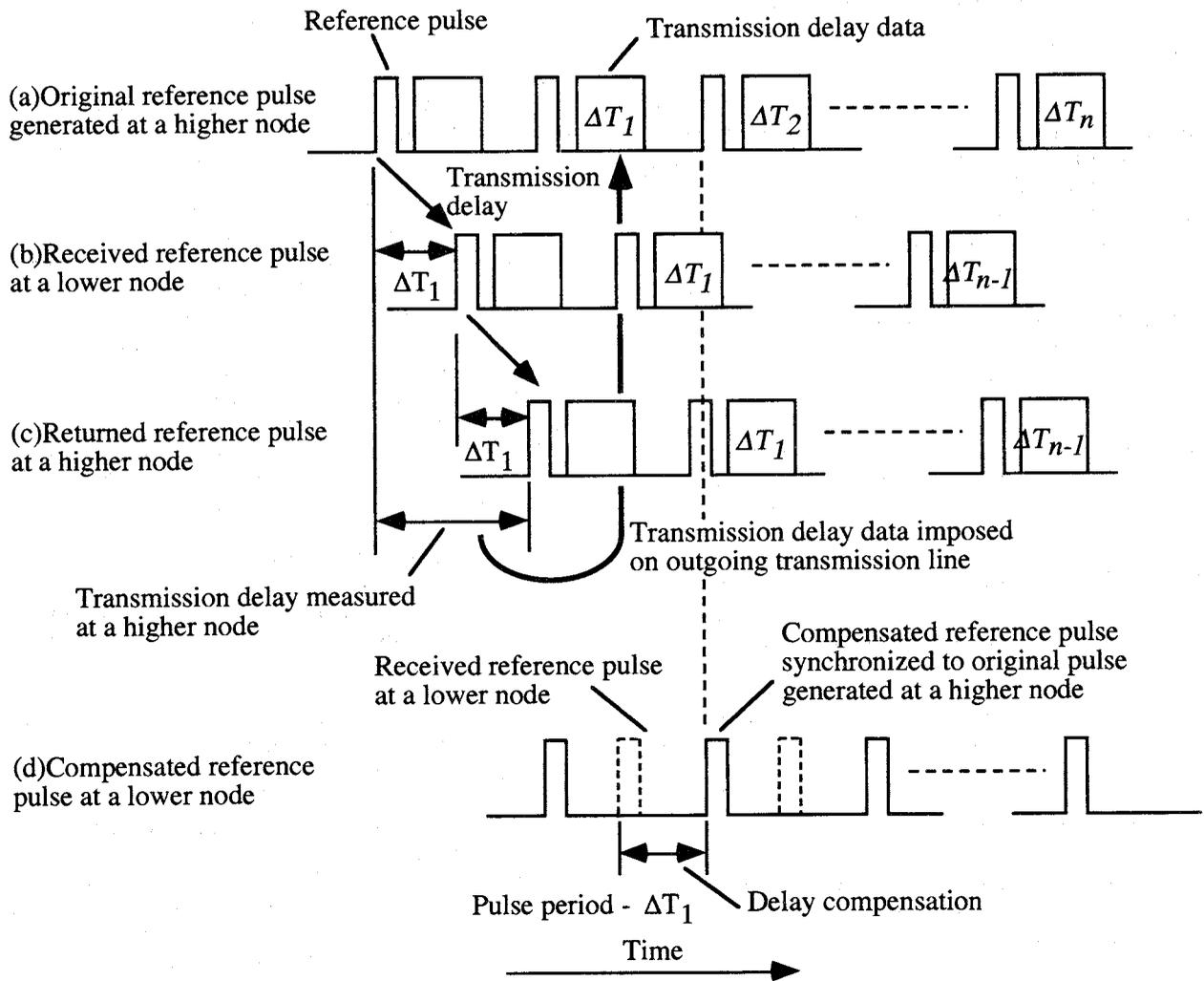


Fig. 2 - Phase relationship among reference pulses for transmission delay compensation

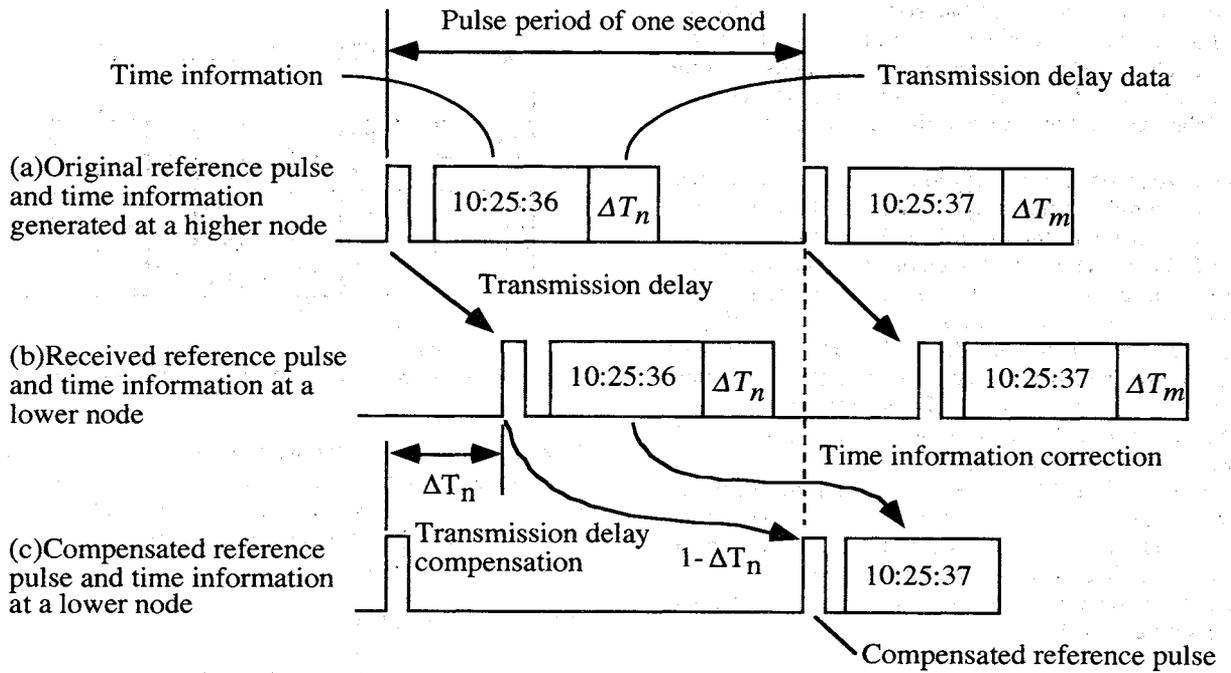


Fig. 3 - Time information transfer

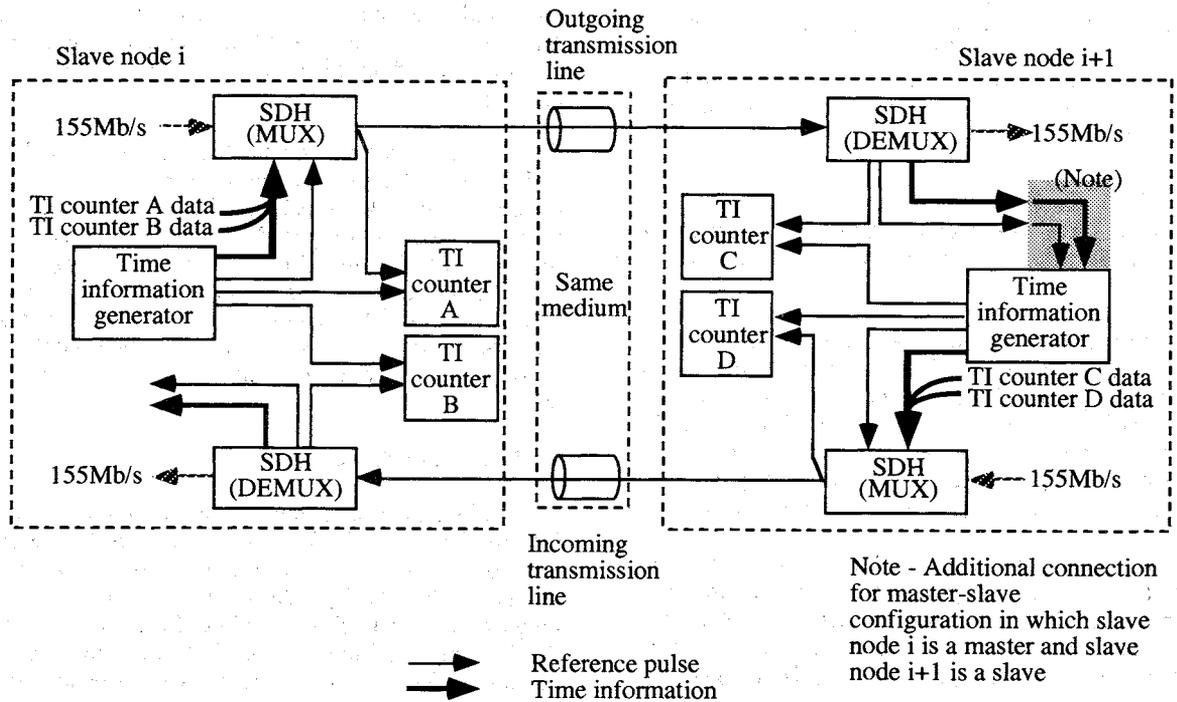


Fig. 4 - The basic configuration of the time information distribution chain

formation can be transferred to the lower node by these steps, and these two nodes can concurrently share the time information as shown in Fig. 3 (a) and (c).

### 3. SYSTEM CONFIGURATION

Figure 4 shows the basic configuration for the new time and frequency transfer system. The system is composed of SDH-based transmission equipment installed in two nodes and the bi-directional transmission medium of outgoing and incoming transmission lines. Time informa-

tion and the reference pulse of one second period generated by time information generator are imposed on the STM-N frame format in the multiplexer (SDH-MUX) at the higher node of slave node i, passed through the outgoing transmission line, and then extracted in the demultiplexer (SDH-DEMUX) at the lower node of slave node i+1. The same process is executed in the direction of the incoming transmission line. While the extracted time information and reference pulse are directly used in slave node i+1 in the two-way time transfer, the extracted signals are additionally fed to the time information gen-

erator in slave node  $i+1$  in the master-slave configuration, the shaded area in Fig. 4. The time information generator in slave node  $i+1$  generates the compensated time information and the reference pulse using the transmission delay data in the extracted time information.

In the master-slave configuration, the higher node measures the round-trip transmission delay. The Time Interval (TI) counter B shown in Fig. 4 measures the time delay ( $D_B$ ) between the reference pulse generated in time information generator of slave node  $i$  and the returned reference pulse from slave node  $i+1$ . Additionally, we have to measure the internal delay of SDH MUX in order to get the true round-trip transmission delay. This measurement is performed by TI counter A ( $D_A$ ) and D ( $D_D$ ) shown in Fig. 4. There is the internal delay of SDH DEMUX as well; however, it is constant. TI counter C measures the time difference ( $D_C$ ) between the extracted reference and the output of the time information generator in the slave node  $i+1$ . Its difference is equal to the transmission delay data included in extracted time information when the time information generator is in the normal tracking condition.

One of the most important points in the new system is that this transfer is achieved by a bi-directional path passing through the same medium. The round-trip path in Fig. 4 is not a loop configuration, but a master-slave configuration. The reference second pulse received by slave node  $i+1$  is influenced by the transmission wander and time delay created within the outgoing transmission line. The delay data ( $TD$ ), which is transferred from slave node  $i$ , is given by

$$TD = (D_B - D_A - D_C - D_D) / 2 + D_A \quad (1)$$

Slave node  $i+1$  can consequently regenerate compensated time information free of the influence of transmission wander.

Time information including delay data is carried in the Section OverHead (SOH) of the SDH frame format in the transmission system of slave node  $i$ , and transferred to the transmission system of slave node  $i+1$ . Since the SOH is used to transfer time information, we can use the STM-N payload area for providing B-ISDN services at the same time; there is no reduction in transmission capacity.

Slave node  $i+1$  regenerates the reference time information and then distributes it to other network systems and to other slave nodes. In customer nodes, transferred reference time information is transferred to customer premises equipment again through the STM-N interface.

#### 4. TRANSFER FORMAT

In SDH, NNI and UNI have SOH bytes independent of the main signal channel for B-ISDN services. We propose that two SOH bytes not defined for any purpose should be assigned to transfer time information (Fig. 5). A byte-wise multi-frame is used to designate the beginning bit of the STM-N frame of 8 kHz period as a second signal. Time and distribution management information are transferred via the 128kb/s data channel. Time information is carried by the first byte, T1 of vector S(9, 4, 1), which includes framing information of one second period, synchronization mode and status, time information

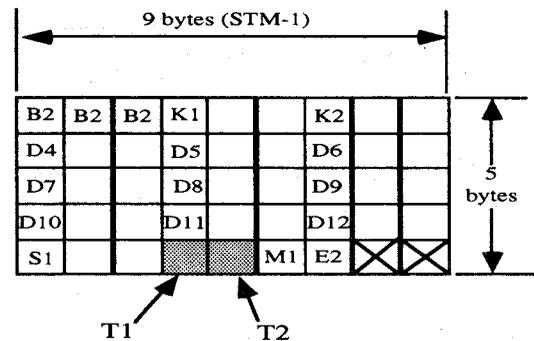


Fig. 5 - Promising MSOH bytes in STM-1

level, and time code category and time itself. Distribution management information is carried by the 2nd byte, T2 of vector S(9, 5, 1), which includes a route map to trace the roots of time[6], a second signal designation and the states of system health.

##### 4.1 Byte T1 of S(9, 4, 1)

Byte T1 is used to transmit time information. The MSB bit of byte T1, b1, is dedicated for framing, which is a multiframe of one second period imposed on the STM-N 8kHz frame. When bit b1 is value '1', Byte T1 including bit b1 indicates the beginning byte of one second frame. The rise time of b1 of byte A1, the beginning of the STM-N frame, including this byte T1, is the reference second (Fig. 6).

Figure 7 shows the relationship between one second frame and blocks. The block length is 20 bytes. Its repetitive period is 2.5ms. 400 blocks can be put into a one second frame. A block is composed of a header of 5 bytes and a time code of 15 bytes. Block content can be varied with time code category. For example, a block using local time category can be followed by a block using UTC category. All blocks can be all '0's except for the first block which must include the beginning information of the one second frame.

##### 4.2 Byte T2 of S(9, 5, 1)

The T2 byte is used to transfer distribution management information. The T2 frame is flexible, and its length is quantized to one second frame multiplied by an integer number. When the T2 status in byte T1 indicates a beginning frame, byte T2 following byte T1 is the beginning of the T2 frame, and the previous byte T2 is the end of the T2 frame. When T2 status indicates a consecutive frame, T2 information is translated together with the previous T2 information. A block of all '0's can follow the end byte of distribution management information until the end byte of T2 frame.

Distribution management information consists of blocks, and includes sender information (16 bytes) and sender status (4 bytes). Sender information includes sender ID's such as sender town, building, floor, equipment, and interface. Sender status includes time information generator conditions such as tracking, holdover, and alarm. The assignment of bit patterns for sender information and sender status is for further study.

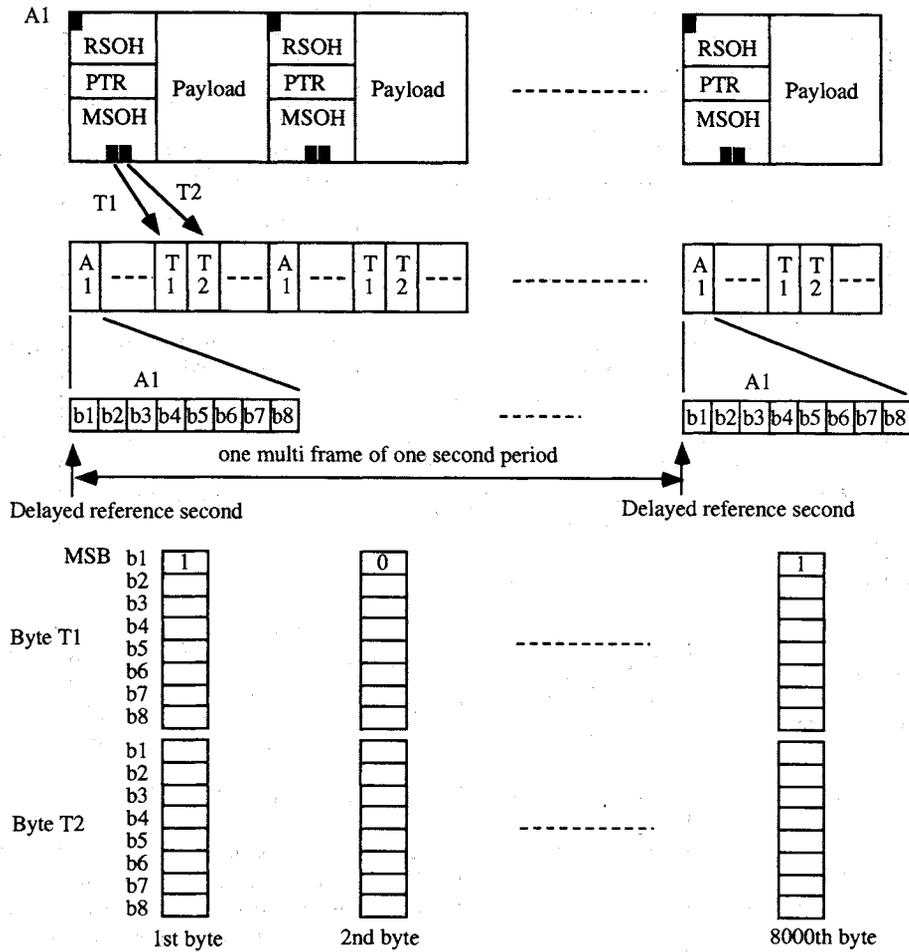


Fig. 6 - Multiframe of one second period

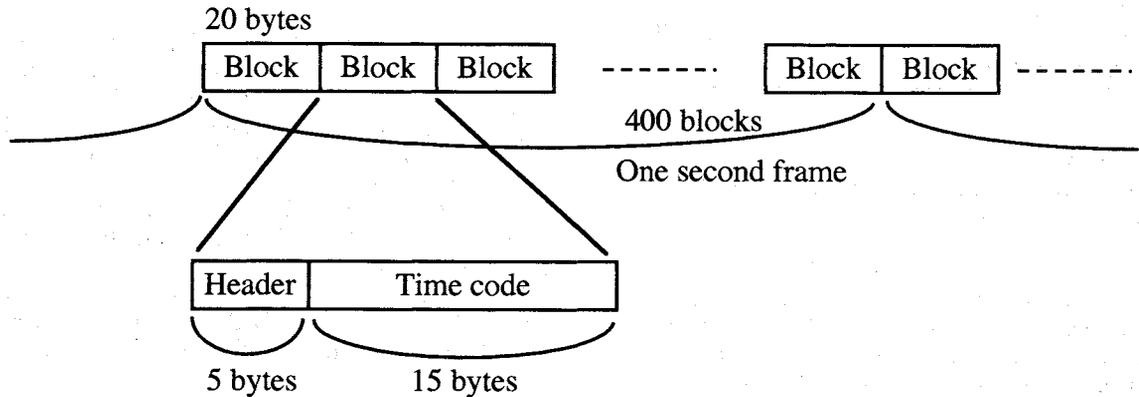


Fig. 7 - One second frame format

**5. EXPECTED PERFORMANCE**

The absolute time error of the new time transfer system is determined by the transmission delay difference between outgoing and incoming transmission lines that is length and characteristic differences between fiber cables. Absolute time error varies with transmission system configurations as listed in Table 1. TYPE I in Table 1 uses two fiber cables. If two fiber cables are installed in the same jacket, absolute time error is typically less than 200ns even for time transfer over 2500km[7]. This error is caused by

fiber cable length asymmetry, dispersion characteristic differences and wavelength differences. TYPE II and III use Wavelength Division Multiplex (WDM) techniques, so one fiber cable can be shared to prevent the biggest error, fiber cable length asymmetry. The transmission delay in TYPE II is influenced by the dispersion difference at two wavelengths. Absolute time error is less than 100ns in the transmission using 1.3μm and 1.5μm over 40km[8]. TYPE III adopts transmission delay compensation. Two wavelength transmission in the same direction enables nodes to measure the difference. The direction of one wavelength is then exchanged, and the usual

Table 1 - Expected absolute time error

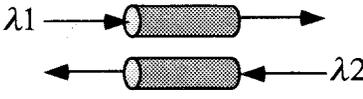
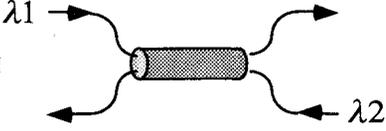
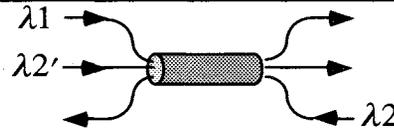
Type		System configuration	Absolute time error factor	Expected absolute time error
TYPE I	Single channel		<ul style="list-style-type: none"> <li>• Optical fiber cable length asymmetry</li> <li>• Dispersion characteristic difference of optical fiber cables</li> <li>• Wavelength difference</li> <li>• Fiber cable length</li> </ul>	$<1\mu\text{s}$ $<200\text{ns}$ (Typical)
TYPE II	Multi channel	WDM 	<ul style="list-style-type: none"> <li>• Wavelength difference</li> <li>• Fiber cable length</li> </ul>	$<100\text{ns}$ (Typical)
TYPE III		WDM with wavelength correction 	<ul style="list-style-type: none"> <li>• slight difference of wavelength</li> </ul>	$<1\text{ns}$

Table 1 - Relative time error

Relative time error factor	Expected relative time error
STM-N frame regeneration Pointer processing	$1\mu\text{s} - 4\mu\text{s}$
Optical module Optical signal power dependence Temperature dependence	$1\text{ns} / 155\text{Mb/s}$
Clock timing generation Device drift Temperature dependence	$<1\text{ns} / 155\text{Mb/s}$
Circuit delay difference Device scatter Temperature dependence	$1\text{ns} - 5\text{ns} / \text{TTL}$

transmission delay and wavelength compensation processes executed. Absolute time error of TYPE III is expected to be less than 1ns[8].

Relative time errors are listed in Table 2. The biggest error is caused by SDH termination equipment. If the SDH signal including time information in SOH is terminated, termination equipment executes pointer processing to convert the input signal into its own frame. This process causes the time error of the order of microseconds. To achieve nanosecond time error, time transfer in a section is required instead of in a path. Other errors in Table 2 depend on SDH bit rates such as STM-1 and STM-4.

The total time error is approximately 10ns with STM-1 (155Mb/s) and is expected to be subnano second with STM-16.

## 6. CONCLUSION

Standardization in ITU-T and ITU-R is indispensable for fundamental and global information such as time and frequency. We have submitted a contribution for the next ITU-T meeting, which proposes this new time transfer

system as a new study item in the next study period of ITU-T[9]. Our contribution also touches on a concrete format and byte reservation in SOH for time transfer. The new recommendation to be made by ITU-T will advance time transfer study and development.

## ACKNOWLEDGMENTS

We are indebted to Judah Levine of NIST for time code information that served as a reference for our own time code determination. We also wish to thank Samuel R. Stein of Timing Solutions Corporation for his discussions (personal communication).

## REFERENCES

- [1] S. R. Stein and A. Gifford; "The Role of Time and Frequency in the Future Systems", Paper 2, the 27th PTTI Applications and Planning Meeting, 1995
- [2] G. Graceffo; "Data and Time-Transfer Using Radio SONET", Paper 26, the 27th PTTI Applications and Planning Meeting, 1995
- [3] M. Weiss and S. Jefferts; "Two-Way Time Transfer through SDH and SONET Systems", the 10th EFTF, 1996 (to be published)
- [4] ITU-T contribution; "Proposal on New Study Items for Clock Characteristics in SDH", COM 13-50-E, 1995
- [5] ITU-R contribution; "Draft New Opinion", Document 7/6-E, 1994
- [6] M. Kihara and A. Imaoka; "Accurate Network Synchronization for Telecommunications Networks", Proc. of the 7th EFTF, pp.83-88, 1993
- [7] M. Kihara and A. Imaoka; "Timing and Time Signal Distribution in Digital Communications Networks", Proc. of the 6th EFTF, pp.489-492, 1992
- [8] A. Imaoka and M. Kihara; "Accurate Time/Frequency Transfer Method Using Bi-directional WDM Transmission", Paper 31, the 27th PTTI Applications and Planning Meeting, 1995
- [9] ITU-T contribution; "Time Information Distribution", SG13, ITU-T 1996

## TEMPERATURE SENSITIVITY OF A FIBRE OPTIC TIMING LINK

Otokar Buzek

Institute of Radio Engineering and Electronics, Academy of Sciences of the Czech Republic

### ABSTRACT

Three single mode 1300 nm fibre optic links are used for a time and frequency transfer between two time laboratories equipped with cesium clocks. These laboratories determine the Czech National Time and Frequency Standard. Out of the total length of 10 km the main part (about 9 km) of the optic cable is laid in a communication cableduct whose depth varies between 2 and 3 m. A smaller part (about 800 m) is buried about 0.5 m underground. The results of routine daily phase time measurements carried out at both ends during nearly 2 years were used to determine the group delay variations of the timing link due to the temperature sensitivity of the fibre.

It has been found that the group delay exhibits seasonal variations strongly correlated with the temperature changes in the cableduct effective depth. As a result, the timing links cause seasonal peak to peak group delay variations of 10 ns and corresponding long-term frequency offsets of the order of  $1 \times 10^{-15}$ .

To correct for the degradation of the time and frequency transfer, the underground temperature variations induced by the mean daily outdoor temperature were modeled for different depths. Using the standard estimation method, the effective depth of the cable was found. The group delay instability corrected for this effect was reduced down to the background white noise PM.

### INTRODUCTION

The factors degrading the group delay stability of various space and terrestrial timing links are widely studied in order to increase the frequency and time transfer accuracy, Čermák (1), Jespersen (2), Primas et al (3).

To minimize the temperature fluctuations, the fibre optic cable is usually buried underground. This helps to suppress the short-term fluctuations of the group delay and, consequently, the short term instability of the time and frequency transfer. Nevertheless, this may not be sufficient to ensure the long-term stability.

The temperature sensitivity of the timing fibre optic links used in Czech National Time and Frequency Standard has already been mentioned by Buzek et al (4) where the preliminary estimation of the factors causing errors in the time transfer was made.

The data available at present have been collected over a period of 700 days which permits the temperature sensitivity of the group delay to be determined more precisely. Based on these robust long-term data, an efficient correction method can be derived.

### EXPERIMENT SPECIFICATIONS

The experiment configuration is shown in Fig.1 where three identical 1300 nm timing links are depicted. All the three links are equipped with the identical electronics devices. The basic technical data of the links are as follows. The output IR LED of each optical transmitter provides the radiated power of  $4 \mu\text{W}$  modulated by 5 MHz signal. The received optical signal is converted to the 5 MHz electric signal by the input photodiode of the optical receiver. This signal is preamplified in an aperiodic preamplifier. For the final amplification a tuned 5 MHz amplifier with a bandwidth of 450 kHz is used. The 5 MHz signal is additionally filtered by a PLL based on a temperature controlled crystal oscillator. The fibre optic cable is 10 050 m long and the total attenuation of the fibre between two end points (including eight optical connectors) reads about 9.6 dB.

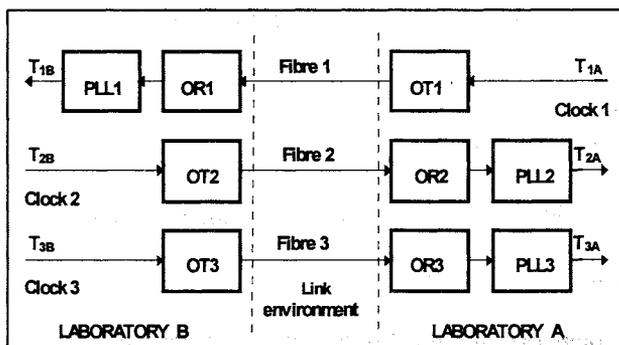


Figure 1: Experiment configuration

The first 5 MHz signal generated by the cesium clock 1 at the laboratory A is transmitted to the laboratory B via the timing link 1 which consists of the optical transmitter OT1, single-mode fibre F1, optical receiver OR1 and finally the phase locked 5 MHz quartz oscillator PLL1. This signal transmitted from A at time  $T_{1A}$  is received at B at time  $T_{1B}$ . The second and third 5 MHz signals derived from two cesium clocks 2 and 3 at B are transmitted to A over two links denoted by 2 and 3. These signals

transmitted at times  $T_{2B}$  and  $T_{3B}$  from B are received at A at times  $T_{2A}$  and  $T_{3A}$ , respectively..

The time difference

$$\Delta T_1 = T_{1B} - T_{1A} = \tau_{10} + \tau_1 \quad (1)$$

where  $\tau_{10}$  is the constant group delay of link 1,  
 $\tau_1$  is the variable group delay component.

In the following analysis, only the variable group delay will be considered. The delay of each link can be represented as a sum of the contributions from the optical receiver (R), fibre (F), optical transmitter (T) and PLL (P). Hence, for the delay  $\tau_1$  in link 1, we can write

$$\tau_1 = \tau_{R1} + \tau_{F1} + \tau_{T1} + \tau_{P1} \quad (2)$$

Analogically, the delays of links 2 and 3 will be denoted by  $\tau_2$  and  $\tau_3$ .

Unfortunately, the resolution of the time interval measurements was limited due to a relatively large counter quantization error representing standard deviation of 2 ns.

From the time interval measurements made at A:

$$\begin{aligned} X_A &= T_{3A} - T_{2A}, \\ Y_A &= T_{1A} - T_{2A}, \\ Z_A &= T_{1A} - T_{3A}, \end{aligned} \quad (3)$$

and from the simultaneous measurements made at B:

$$\begin{aligned} X_B &= T_{3B} - T_{2B}, \\ Y_B &= T_{1B} - T_{2B}, \\ Z_B &= T_{1B} - T_{3B}. \end{aligned} \quad (4)$$

the following relations may be obtained :

$$\begin{aligned} \Delta X &= X_B - X_A, \\ \Delta Y &= Y_B - Y_A, \\ \Delta Z &= Z_B - Z_A, \end{aligned} \quad (5)$$

Substituting from (3) and (4) in (5), we have

$$\Delta X = \tau_3(t) - \tau_2(t) \quad (6)$$

which represents the differential delay of links 3 and 2. In addition, the difference

$$\Delta Y = \tau_1(t) + \tau_2(t) \quad (7)$$

is the summary delay of links 1 and 2.

Similarly, the summary delay  $\Delta Z$  of links 1 and 3 can be expressed as

$$\Delta Z = \tau_1(t) + \tau_3(t) \quad (8)$$

represents the summary delay of links 1 and 3.

The relations (6), (7) and (8) allow the delay to be estimated from the results of the routine daily measurements (3) and (4).

During the experiment, the indoor temperature at both laboratories was daily measured in order to find out a temperature sensitivity of the link electronic equipment. The daily mean temperature reported for Prague by the Czech Hydrometeorological Institute was used to estimate the temperature sensitivity of the fibre.

## EXPERIMENT RESULTS

Records of delays  $\Delta X$ ,  $\Delta Y$ , and  $\Delta Z$  taken over the period of nearly two years, from 29 January 1994 (MJD49381) to 30 December 1995 (MJD 50081), are shown in Fig. 2. Evidently, the summary delay  $\Delta Y$  of links 1 and 2 as well as  $\Delta Z$  of links 1 and 3 exhibit very similar seasonal variations of about  $\pm 10$  ns. These phase time variations result in the periodic long-term relative frequency offset within  $\pm 1.10^{-15}$ . The frequency deviation is negative during the first half of the year and positive during the second half of the year.

In addition, one can see that the differential delay  $\Delta X$  is nearly constant within the limits given by the counter quantization noise. Unfortunately, the resolution of the time interval measurements was limited due to a relatively large counter quantization error representing standard deviation of 2 ns.

The Allan deviation of  $\Delta X$  reads 2.5 parts in  $10^{14}$  for the 1 day averaging interval.

The above results bring in evidence a high degree of conformity of the timing links and, therefore, the delays  $\tau_1$ ,  $\tau_2$ , and  $\tau_3$  may be assumed to be equal.

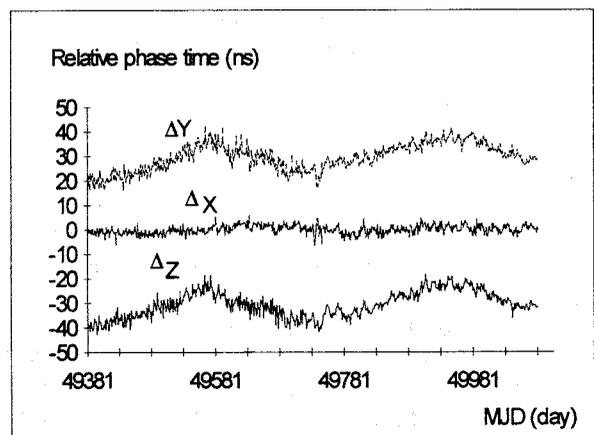


Figure 2: Summary and differential link delays

The records of temperature taken at both laboratories as well as the record of the mean daily temperature in Prague are shown in Fig. 3.

The records are shifted by an arbitrary constant to prevent overlapping.

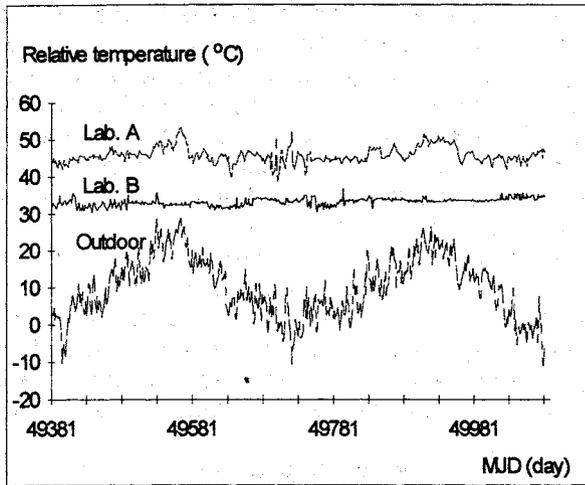


Figure 3: Temperature records

### TEMPERATURE SENSITIVITY

A significant correlation of the summary delays  $\Delta Y$  and  $\Delta Z$  with the outdoor temperature has been found. In contrast, the influence of the indoor temperature can be neglected. In addition, the summary delays  $\Delta Y$  and  $\Delta Z$  when understood as responses to the outdoor temperature are delayed by about 18 days.

This large correlation time can only be explained by the propagation time of the heat over the distance between the ground surface and the buried optic cable.

If the ground surface temperature (2) is given by

$$T = A \cos(\omega t) \quad (9)$$

than the temperature  $T_H$  in the depth  $H$  will be

$$T_H = A e^{-kH} \cos(\omega t - kH) \quad (10)$$

where

$$k = \sqrt{\frac{\omega}{2\kappa}} \quad (11)$$

where  $\kappa$ , denotes the soil conductivity of heat. The formulae (9), (10) and (11) were used to evaluate the temperature  $T_H$  for different values of  $H$  as a function of the outdoor temperature. Firstly, the spectra of the responses  $R_H$  to the unit step of the outdoor temperature were evaluated for  $H$  from 0 to 3 m in 0.25 m steps. Secondly, the temperature dependence was modeled for different  $H$  by evaluating the convolution

$$T_H = R_H \otimes \Delta T \quad (12)$$

where  $\Delta T$  denotes the sequence of daily differences of the outdoor temperature. The typical value  $\kappa = 0.079 \text{ m}^2/\text{day}$  was used.

The curves representing the modeled temperature for different depths are given in Fig.4 (arbitrary constants added).

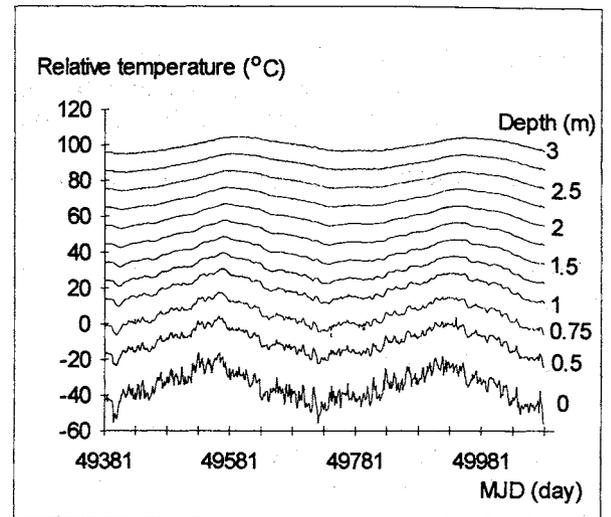


Figure 4: Underground temperature model

The temperature variations shown in Fig.4 are in good agreement with the experimental results obtained by Lednický (6).

Assuming the dominant influence of the underground temperature and of the linear changes in the group delay due to likely aging, the summary delays can be written in the form

$$\Delta Y = K_{Y1} \cdot T_H + K_{Y2} \cdot (t - t_0) \quad (13)$$

$$\Delta Z = K_{Z1} \cdot T_H + K_{Z2} \cdot (t - t_0) \quad (14)$$

where  $K_{Y1}$  and  $K_{Z1}$  are the corresponding thermal coefficients,

$K_{Y2}$  and  $K_{Z2}$  are the ageing coefficients,  
 $T_H$  is the temperature  $H$  meters underground,

$t$  is the time MJD and

$t_0 = 49381$ .

By applying the least-square method to daily measurements for a set of  $N$  equations ( $N = 700$ ) and by using the temperature models  $T_H$ , the coefficients  $K_{Y1}$ ,  $K_{Z1}$ ,  $K_{Y2}$  and  $K_{Z2}$  were found. The optimum estimation was achieved for the depth  $H = 1 \text{ m}$  ( $T_H = T_1$ ) and for the following coefficients:

$$\begin{aligned} K_{Y1} &= 0.81 \text{ ns/grad}, & K_{Y2} &= 0,1 \text{ ns/day} \\ K_{Z1} &= 0,77 \text{ ns/grad}, & K_{Z2} &= 0.1 \text{ ns/day} \end{aligned}$$

For the further computations the mean values will be taken

$$K_1 = 0.79 \text{ ns/grad}, \quad K_2 = 0.1 \text{ ns/day}. \quad (15)$$

The fibre thermal coefficient of dilatation (TCD) can be determined for  $K_1$  as

$$TCD = \frac{K_1 c}{L n} \quad (16)$$

where  $c = 2.99792 \times 10^8$  m/s is the speed of light,  
 $L = 19820$  m is the total length of the fibre  
 under the influence of the temperature  
 $T_H$ ,  
 $n = 1.429$  is the index of refraction of the  
 fibre.

After substitution of the above values to (16) we obtain

$$TCD = 8.5 \text{ ppm} / ^\circ\text{C}$$

which is very near to the values published.

Provided the coefficients representing the temperature sensitivity and the long-term drift of the link group delay are known, the long-term variations in the group delay of the fibre links may be reduced for the current outdoor temperature as well for the ageing

$$\Delta Y_{corr} = \Delta Y - K_1 T_1 - K_2 (t - t_0) \quad (17)$$

$$\Delta Z_{corr} = \Delta Z - K_1 T_1 - K_2 (t - t_0) \quad (18)$$

The resulting summary group delays are shown in Fig. 5.

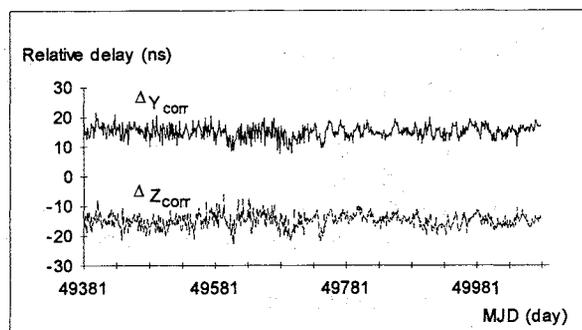


Figure 5: Corrected summary delays

Allan deviation of the two corrected summary delays averaged over 1 day is  $3.3 \times 10^{-14}$  and  $3.5 \times 10^{-14}$ , respectively.

## CONCLUSION

The study of the long-term stability of the group delay in the fibre optic links shows that the temperature variations on the fibre are the dominant degradation factor. In the extremely precise time and frequency distribution systems, the temperature of the ground around the cable should be taken into account.

The study has also shown that an effective depth can be determined (at least in our case) for a fibre

whose parts are buried at different and not exactly known depths. The temperature corresponding to this effective depth is the best approximation of the complex thermal effect on the fibre. Applying the temperature variations of the effective depth to correct for the thermal variations gives acceptable results. The residual instability is at the level of the background noise.

The calculated effective depth of 1 meter is smaller than the mean depth of the cable whose major part is buried 2 to 3 m underground. This can be explained by the non-homogeneity of the ground in the dense urban area and, particularly, by the poor environmental isolation of the cable-duct.

The cause of the ageing phenomenon has not yet been found. This will be the subject of further experiments.

## ACKNOWLEDGEMENTS

The author would like to thank B. Čemusová for the help with data collection and processing, and to J. Čermák for his helpful suggestions.

## REFERENCES

1. Čermák J, 1996, "A GPS parallel common-view experiment", 10th EFTF
2. Jespersen J, 1989, "Impact of atmospheric non-reciprocity on satellite two-way time transfers", 43rd ASFC, 186-192
3. Primas L, Logan R and Lutes G, 1989, "Applications of ultra-stable fiber optic distribution system", 43rd ASFC, 202-211
4. Buzek O, Čermák J and Čemusová B, "Czech National Time and Frequency Group Standard", 9th EFTF, 400-403
5. Carslaw H, Jaeger J, 1959, "Conduction of Heat in Solids", Oxford, At the Clarendon Press,
6. Lednický V, 1983, "The soil temperature in low depths" (in Czech), Meteorologické zprávy 36, 71-73

## INSTRUMENTATION FOR CS-TIME TRANSFER TO CALIBRATION LABORATORY

K. Kalliomäki <sup>(1)</sup>, T. Mansten <sup>(2)</sup>

1) University of Oulu, Department of Electrical Engineering, FIN-90570 Oulu, Finland

2) VTT AUTOMATION, Measurement Technology, P.O. Box 1304, FIN-02044 VTT, Finland

National time of Finland is traced from GPS satellite receiver and from RNS (Russian Loran) due to long NELS test period. Short term stability at National Standards Laboratory (KMNL) is maintained by rubidium clocks. There are several commercial Cs-clocks at 10 km distance from KMNL. During the last years we have studied several methods to build a link to those clocks owned by Finnish TELE. Finally we have build and tested instrumentation based on simultaneous GPS- and TV-time comparison at both sides.

At KMNL the phase of all local clocks is measured every 10 minutes. We decided to use exactly the same data collection procedure at both sides. The instrumentation consists of a GPS-receiver (Navstar XRS-M), TV-receiver (conventional video recorder), time interval counter (HP5334) with 1 ns resolution, a computer (PC) with telephone modem and some additional electronics. In addition several analog variables like temperature and pressure have been measured.

Now the test phase of the new instrumentation is going on. The aim of this project is to have an independent atomic time for KMNL. The reason for two independent comparison methods (GPS and TV) is simply reliability. We are going to describe the instrumentation in detail and actual time transfer results results.

Reference: K. Kalliomäki, T. Mansten, Stability of Cs-clock based 2048 kHz Optical ISDN synchronization signal. 1995. Proc. 9th EFTF, Besancon, France.

## REDUCTION OF PHASE NOISE IN SINGLE TRANSISTOR OSCILLATORS

MJ Underhill

Department of Electronic & Electrical Engineering, University of Surrey,  
Guildford, UK

### ABSTRACT

A single transistor oscillator requires a tuned network which gives 180 degrees of phase shift at the desired frequency of operation. Best phase noise requires that the figure of merit  $PQ^2/F$  be maximised for the oscillator. The product  $PQ^2$  shows that Power  $P$  and circuit  $Q$  can be traded whilst maintaining the same phase noise. However  $PQ^2$  is actually limited by the maximum voltage or current in the tuned circuit together with the associated loss resistance. Optimisation of these automatically minimises the phase noise.

Of the considerable number of suitable tuned networks the pi network class is shown to be preferable over the tee network class because the transistor behaves predominantly as a transconductance device.

Many apparently different oscillator circuits are actually the same for the purpose of analysis of noise or feedback conditions. This is because such analysis is independent of which point in the circuit is regarded as the ground point.

Both the input and output match of a pi network can be adjusted independently over wide ranges in a simple manner. In the Clapp class of oscillator such adjustments can be made practically independently of the circuit overall loaded  $Q$ . Such a circuit allows noise match conditions to be set at the (base) input of the amplifier and at the same time the match to the output (collector) of the transistor can be set independently.

It is shown that noise match conditions are not critical for the optimisation of phase noise particularly if the transistor noise factor  $F$  is near to unity (low added noise). For best overall noise floor it is essential that more power is provided at the oscillator output (for use in the next stage) than is presented as the available power at the input to the oscillator transistor.

There are a number of possible places in the circuit to extract the output power. It is shown that there are two preferred places for the output load which in general should give the best phase noise performance. One of these in fact corresponds to the traditional Colpitts or Clapp oscillator. The other corresponds to an oscillator configuration which is possibly novel. This latter circuit promises better (phase noise) performance at higher frequencies.

### 1. INTRODUCTION

Phase noise in any radio frequency system is inevitable. It limits the achievable selectivity in a channelised communication system and in a radar system it limits the ability to discriminate between a wanted target and an adjacent target or clutter. The oscillator is an essential part of practically all radio frequency systems. It is therefore unfortunate that the generation of phase noise in an oscillator is unavoidable (1). Low cost systems require simple efficient and cost effective oscillator designs.

The single transistor oscillator is a good choice because it is simple and uses a minimum of components.

Originally, the main design objective for an oscillator was that it should oscillate at the desired power level over the required frequency range and at the same time have the desired frequency stability. This paper focuses on the design principles of single transistor oscillators covering the required frequency range at the desired power level and with the best possible phase noise performance. The problem of frequency stability

is not addressed. It is assumed that other means are present for ensuring the frequency stability of the oscillator, such as a phase locked loop (PLL) arrangement and/or reference to a quartz crystal or other stable frequency source.

The noise performance of an oscillator circuit or whether it fulfils the feedback conditions for oscillation over a frequency range is independent of the biasing and power supply circuitry, provided that these are properly designed. For this reason only idealised circuits missing out such details are considered in this paper.

It is also true that noise performance and conditions of oscillation are independent of which single point in the circuit is chosen as rf ground.

Recognition of this fact allow apparently very different circuits to be grouped together and to be treated as the same for the purpose of analysis.

The approach taken starts with the simple phase noise model of any oscillator originally put forward by Leeson. (2) To be able to group all possible single transistor oscillator circuits into a minimum number of classes, standard circuit theory transformations are used. It is then assumed that the transistor can be considered to be split so that input and output matching can be optimised independently. The input matching can then be designed for best noise match conditions, and the output match can then be adjusted independently for the best power efficiency phase noise trade-off. The emphasis is on circuits using the Bipolar Junction Transistor but the principles transfer directly to circuits using Field Effect Transistors (FETs).

The primary aim of this paper is to show that such an idealised approach does nonetheless lead to very practical design guidelines and rules for the minimisation of phase noise in any single transistor oscillator.

## 2. THE Q MULTIPLIER PHASE NOISE MODEL OF AN OSCILLATOR

The essential basic elements of any oscillator are shown in Fig.1. The amplifier provides enough gain so that the small signal positive feedback gain around the loop is greater than unity. By saturating the limiter reduces the loop gain to almost exactly unity when the signal level in the loop becomes large enough.

The original effective  $Q_e$  of the selective element is amplified by a large factor  $k_c$  (typically  $10^{12}$ ) to give

a closed loop  $Q_c = k_c Q_e$ . The input noise has a flat power spectrum of  $FkT$ . Of this  $0.5 FkT$  is phase noise which is amplified by the closed loop gain  $k_c$  and at the same time filtered by the very narrow closed loop bandwidth of  $Q_c/f_o$ .  $f_o$  is the "carrier" output frequency. The resulting power spectrum is shown in Fig.2. The output signal power is actually the area under the power spectral density curve of the output. This curve is so narrow at its peak ( $f_o/Q_c = 3\text{dB}$  bandwidth) that it appears as a pure carrier signal.

The limiter in the loop serves two functions. It both suppresses the amplitude components (also  $0.5 FkT$ ) of the input noise and also sets the output amplitude to a desired value, thereby automatically establishing the closed loop gain at  $k_c = Q_c/Q_o$ . But  $k_c = k_o/(1 - k_o)$  where  $k_o$  is the open loop gain which is therefore very nearly unity.

In Appendix A more details are given showing that this model results in a normalised power spectral density at noise sideband frequency  $f_m = |f_o - f|$  spaced from the carrier  $f_o$  of

$$S_\phi(f) = \left(\frac{1}{2} FkT / 4PQ_o^2\right) \cdot (f_o / f_m)^2 \dots\dots\dots(2.1)$$

where  $P = P_{av}$  = carrier power available at the input of the amplifier. It should be stressed that  $Q_o$  is the operating or loaded  $Q$  of the selective circuit; it includes the effect any external loading appearing on the input or output ports of the selective circuit.

This equation provides the key basis for designing any oscillator to have low phase noise. In summary it shows that maximising the performance index  $PQ^2/FkT$  will automatically minimise phase noise. In practice the operating temperature is rarely taken to be a parameter which can be reduced cost effectively, and so in the following discussions the performance index to be optimised is taken as  $PQ^2/F$ .

## 3. THE CONDITION FOR OSCILLATION

For oscillations to build up at a given frequency in a feedback circuit Barkhausen's conditions are that the phase shift should be zero or a multiple of  $2\pi$  radians or  $360^\circ$ , and the loop gain should be greater than unity.

For the basic oscillator shown in Fig.3 it requires that the gain of the amplifier should be sufficient to

overcome any losses in the selective network and the selective network should provide phase reversal, that is of  $\pi$  radians or  $180^\circ$ . Fig.4 shows a total of 12 basic networks which potentially can be used in a single transistor oscillator circuit.

To provide a phase shift of  $180^\circ$  the six pi network circuits Fig.4(a) to (f) required to be fed from a high impedance source and terminated by a high impedance load. The single transistor is basically a transconductance device (voltage in and current out). It therefore presents the right kind of terminations for the pi networks to operate correctly.

The tee networks in Fig.4(g) to (l) require low impedance terminations which are not naturally provided by the transistor amplifier. For this reason the low impedance loads  $R_1$  and  $R_2$  are included as part of the tee network in each case.

For the purpose of estimating the transistor stage gain the emitter load resistor  $R_E$  and the collector load resistor  $R_C$  are shown in Fig. 5. The resistors represent the total load values presented at the emitter and collector at the oscillation frequency  $f_o$ . The voltage gain  $k_t$  of such a stage is given by

$$k_t = R_c / (R_E + g_m^{-1}) \dots \dots \dots (3.1)$$

where  $g_m = 40 I_C$  to a good engineering approximation for a Bipolar Junction Transistor (BJT) and  $I_C$  is the dc collector current in amps.

If  $R_E$  is smaller than  $g_m^{-1}$  the voltage gain is  $g_m R_C$  and if  $R_E$  is larger than  $g_m^{-1}$  the voltage gain is  $R_C/R_E$ . For  $I_C = 1\text{mA}$ ,  $g_m^{-1} = 25$  ohms.

Provided the pi network is operating with a circulating current which is large with respect to both the input and output currents the ratio of output to input voltages is given by the ratio of the output to input reactances, that is either  $V_2/V_1 = C_1/C_2$  or  $= L_2/L_1$ . Large enough means that resistive loads presented to the input and output are at least 5 times the input and output reactance values respectively.

Thus for circuits a, c and e in Fig. 4 the loop gain condition for oscillation is:

$$(C_1/C_2) \cdot R_c / (R_E + g_m^{-1}) \geq 1 \dots \dots \dots (3.2a)$$

and for b, d and f

$$(L_2/L_1) \cdot R_c / (R_E + g_m^{-1}) \geq 1 \dots \dots \dots (3.2b)$$

In practice a major part of the transistor collector load resistance  $R_C$  is the input resistance of the pi network when at resonance. An additional component of this load or the emitter load  $R_E$  can be the load resistance presented to the circuit by the necessity to couple power out into a subsequent stage (such as a buffer amplifier or rf mixer).

Similar considerations can be used to derive simple equations for the tee network oscillators.

In Figs. 4g to 4l it is the voltage at the tee junction which corresponds (as a dual) to the circulating current in the pi network. This voltage should be at least five times larger than the larger of the voltages across  $R_1$  or  $R_2$ . The implication of this is that the reactances of  $L_1$  or  $L_2$ , or  $C_1$  or  $C_2$  should be higher by five times the respective resistances  $R_1$  and  $R_2$ . For good power efficiency and good match to the transistor  $R_1$  and  $R_2$  should not be too low (not less than about  $1k$ ) and this means that the network reactances become too large to be easily realisable in practical circuits either at low frequencies where the inductors become too large or at high frequencies where the capacitances become too small. For this reason none of the tee networks are in general to be preferred in practical single transistor oscillator circuits.

The condition for oscillation is completely independent of which single point in the oscillator circuit is grounded. Thus depending on which of the three points is chosen then the circuit of Fig. 3 combined with a network from Fig. 4 can be redrawn to reveal any of the popular type of oscillator circuits and also to reveal other arrangements which could be useful in some circumstances. Combining the three choices for ground points with the twelve networks shown indicates a class of thirty six different oscillator types. However as will be seen in section 5 when the oscillator output load is included this number of choices can be expanded further.

#### 4. MATCHING WITH THE RESONANT PI NETWORK

Provided the circulating current is kept high enough as already described the pi network can be said to be "resonant". Under these circumstances it is found that adjusting the values of the input and output reactance provides a very simple and easy way of adjusting the power match oscillator circuit.

Figs. 6a and 6b show the standard parallel to series (and series to parallel) circuit transformations.

The important point to note is that provided the Q is high enough (greater than about 5) then the resistance values are transformed by  $R_p R_s = X^2$  and the reactance values remain the same,  $X_p = X_s = X$ .

Also the current source  $i$  is transformed to a voltage source  $V = iX$ .

Fig 6c show the transformations applied to the pi network in a way which is useful for deriving the total loaded Q of the network. It also illustrates how the choice of  $X_1$  and  $X_2$  can transform  $R_1$  and  $R_2$  to any desired  $R_{11}$  and  $R_{21}$  series equivalents.

Fig. 6d shows how the parallel capacitance across the series arm of a pi network can be split and absorbed into  $X_1$  and  $X_2$  in the same ratio so that the resulting  $X'_1$  to  $X'_2$

ratio and hence the voltage transformation ratio remains the same. This is particularly useful for understanding the design of crystal oscillators. In this case the parallel capacitance  $C_0$  can be transformed to leave the motional arm of the crystal exposed as the sole series element.

## 5. NOISE MATCHING

The noise factor F of an amplifier stage depends on the impedance of the source supplying the signal to the amplifier input. When the source impedance is arranged to be the optimum value for the amplifier the added noise  $N_a = (F-1)kT$  at the input becomes minimised and the signal to noise ratio can be maximised.

The optimum noise match source as seen by the input to the impedance transistor amplifier is in general not entirely resistive. For an oscillator circuit to be able to present a complex impedance to the amplifier input the tuned network needs to be appropriately detuned. This can be achieved by inserting a supplementary phase shift somewhere in the feedback loop. The frequency then alters until the tuned network supplies an equal and opposite phase shift. In these circumstances a suitable complex source impedance can be presented to the amplifier input. Unfortunately a phase shift of  $\phi$  will cause the change

$$Q_c = Q_o \cos \phi$$

Furthermore other power match conditions in the circuit can be disturbed causing the  $PQ^2$  part of the figure of merit to be degraded (3) by a total factor of

$\cos^4 \phi$ . For these reasons optimisation to the reactive part of the noise match resistance is not recommended for oscillator circuits.

It is interesting to note that noise match theory shows that even a gross mismatch of the source resistance can only cause the added noise power to double in the worst case. A gross mismatch to the actual input resistance (not necessarily the same as the noise match resistance) does not affect the available signal power presented to the input; the mismatch simply causes the signal power to be reflected back to the source with a high standing wave ratio.

Table 1 shows the effect of doubling the added noise  $N_a$  for various values of amplifier optimum noise factor. What is of particular note is that the degradation of the noise factor F is actually much smaller than 3dB if the initial added noise is small and F is not much larger than unity.

Table 1

Effect of  $N_a$  (Added Noise) on F (Noise Factor) and FdB (Noise figure)

$N_a/kT$	F	FdB	FdB Increase for +3dB for $F_a$
0.1	1.1	0.41	> 0.38
0.2	1.2	0.79	> 0.67
0.4	1.4	1.46	> 1.09
0.8	1.8	2.55	> 1.8
1	2	3.0	> 2.2
2	3	4.8	> 2.5
4	5	7.0	
8	9	9.5	
9	10	10	
10	11	10.4	

This table shows that a maximum of 3dB improvement in  $S\phi(f)$  can be achieved by noise matching (all else being equal) and if the noise factor is already good the gain is even less.

**6. OPTIMISATION OF THE FIGURE OF MERIT  $PQ^2/F$**

In general it is more difficult to reduce phase noise in tuned oscillators particularly for wide range varactor tuned voltage controlled oscillators (VCOs). For this reason the varactor VCO is now chosen to illustrate how the figure of  $PQ^2/F$  can be optimised.

A good starting point for the optimisation of phase noise in single transistor oscillators is to choose a transistor with sufficient power handling capability, but more importantly the best possible noise factor  $F$  at the operating frequency. (An important further contribution to noise can occur if low frequency transistor noise is converted to phase noise by AM to PM conversion occurring in the transistor. To minimise this it is good practice to choose a transistor with low  $1/f$  noise, and biasing and decoupling arrangements which minimise the generation of low frequency noise).

It will now be shown that the optimisation of  $PQ^2$  does not depend on the actual input impedance of the transistor stage even if this is considerably different from the noise match impedance. This important fact can be seen as follows.

At resonance all the reactances shown in Fig. 7 sum to zero. But note that  $X_3$  is the only inductive reactance so that the  $Q=Q_1$  is

$$Q_t = X_3 / (R'_1 + R_3 + R_4 + R'_a) \dots\dots\dots(6.1)$$

The series current is:

$$i_s = V / (R'_1 + R_3 + R_4 + R'_a) \dots\dots\dots(6.2)$$

The voltage across the inductance

$$V_3 = i_s X_3 = Q_t V \dots\dots\dots(6.3)$$

Equation 6.3 is in fact well known and could have been stated without derivation from equations 6.1 and 6.2.

For  $P=P_{av}$  the power available at the amplifier input we have by definition:

$$P_{av} = V^2 / 4(R'_1 + R_3 + R_4) \dots\dots\dots(6.4)$$

and from equations 6.4 and 6.3 we derive

$$PQ^2 = P_{av} Q_t^2 = V_3^2 / 4(R'_1 + R_3 + R_4) \dots\dots\dots(6.5)$$

The significance of equation 6.5 is that it does not depend on the transformed input impedance  $R_a'$  of the

transistor amplifier. This means that the reactance  $X_2$  in the pi circuits of Fig. 4 can be chosen to transform the noise match impedance  $R_n$  of the amplifier into an equivalent  $R_a'$  which for lowest noise should be chosen so that

$$R'_n \approx R'_1 + R_3 + R_4 \dots\dots\dots(6.6)$$

The  $R_n$  to  $R_n'$  transformation can be chosen in this way without the resulting transistor input impedance  $R_a$  to  $R_a'$  transformation having any effect on  $PQ^2$ . Equation 6.5 also shows what needs to be done to obtain the best  $PQ^2$ : The voltage  $V_3$  across the reactance has to be as high as possible; the varactor and inductor loss resistance,  $R_3$  and  $R_4$ , need to be as low as possible; and the effective transformed input source resistance  $R_1'$  should be as low as possible.

Note that because of the transformation through  $X_1$   $R_1$  being as high as possible corresponds to  $R_1'$  being as low as possible. As will be seen in the next sections, it is easy to achieve this condition because of the naturally high output impedance of a transistor (acting as a transconductor) but only provided that no (output) load is placed across the input ( $X_1$ ) of the pi network.

Usually the major restriction on the voltage  $V_3$  across the inductor is the rf voltage handling capability of the varactor shown as reactance  $X_4$  in Fig. 7. Fig. 8 shows how  $V_3$  can be increased by inserting the capacitance  $X_6$  with a suitable value in series with the varactor.

If the combination ( $X_6 + X_1 + X_2$ ) is made  $n$  times  $X_4$  then the voltage  $V_3$  can be increased  $n$  times.

This allows  $PQ^2$  to be increased by  $n^2$  times according to equation 2.1 corresponding to a  $n^{-2}$  improvement of noise.

The penalty is that the tuning range achievable by the varactor is then reduced by the factor  $1/n$ .

This technique does nonetheless offer considerable phase noise reductions and it is surprising that it is not more used by VCO designers.

In the above example the VCO range can be restored to its original value in absolute frequency terms if the output frequency  $f_0$  is multiplied to  $nf_0$ . However, the phase noise is then increased by  $n^2$  which cancels out the original improvement. As a side issue it

indicates that the phase noise performance of a VCO should always be defined with respect to its absolute tuning range rather than to its actual output frequency.

As an example for VCOs a useful  $S\phi(f)$  criterion for a good oscillator is better than -120dBC in 1Hz at 10kHz for 100MHz range. This is irrespective of whether the tuning range is achieved at 200MHz, 2GHz or 20GHz.

From equation 2.1 it is possible to estimate  $PQ^2$  corresponding to this good VCO criterion. For  $F=2$  (3dB) and taking  $f_0=100\text{MHz}$ ,  $f_m=10\text{kHz}$   $KT=4 \times 10^{-21}$  and  $S\phi(f)=10^{-12}$  we find  $PQ^2=0.1$  watts.

Then assuming circuit losses  $R_5=R_4+R_5(+R_1)$  =2.5ohm then from equation 6.5 we have for the rms rf voltage across the inductor  $V_3 = 1$  volt. If the technique shown in Fig. 8 is not used then

$$V_4 \approx V_3$$

and so the voltage across the varactor is also almost 1 volt.

From this can be seen that the best way to achieve good phase noise is to maximise the voltage  $V$  in the tuned circuit and to minimise the tuned circuit loss resistance  $R_s$  to give the best  $PQ^2 = V^2/4R_s$ . A value greater than 0.1 watts will provide good VCO performance almost automatically.

## 7. THE OPTIMUM OUTPUT LOAD

For the output load optimum means not only the optimum value but also the optimum position within the transistor oscillator circuit.

The first requirement is that in general a minimum of the same power which is made available to the input of the transistor stage should also be made available to the input of the subsequent stage. If this is not so then ratio of the signal power to the noise floor or noise floor  $S\phi(f)$  will become degrade in the second stage. (The noise floor is assumed to be the same in each stage).

The second requirement is that the position of the output load in the circuit should be such to cause the minimum reduction of the loaded  $Q_0$ . A major conclusion of this paper is that because a transistor behaves as a transconductance it is possible to choose an output load position which has little effect on the loaded  $Q_0$ . Fig. 9 shows a total of four different output load  $R_0$  positions in a pi network type transistor oscillator. For the tee networks similar load positions can be chosen in case (a)  $R_0$  can actually

replace  $R_1$  and in case (d)  $R_0$  can replace  $R_2$  referring back to Fig. 4.

Possible earthing points are also shown in Fig. 9 for each load position. Those connected to one or other end of the load are slightly to be preferred for ease of coupling to the next stage. Otherwise an rf transformer output is required to allow the load to remain ungrounded and floating.

An important point is that not only are the conditions for oscillation unaffected by which ground point is chosen, but also the noise match and power match conditions are similarly unaffected.

It can now also be noted that consideration has been given to a total of 14 load position and ground point combinations as in Fig. 9 and 12 tuned networks as in Fig. 4. This implies that a total of  $12 \times 14 = 168$  different oscillator circuits are addressed by the analysis method given in this paper.

It now remains to choose which of the arrangements in Fig. 9 are the best for phase noise performance. In section 6 it was seen that the best phase noise figure of merit could be achieved if the input of the pi network was driven by a current source. In Fig. 9(a) the load  $R_0$  appears across the input to the pi network and therefore degrades  $Q_0$ . Similarly  $Q_0$  is degraded by  $R_0$  at the output of the pi network. In fact by treating the pi network as a perfect transformer case (d) can actually be seen to be the same as case (a).

In both Figs 9(b) and 9(c) the input of the network sees the high impedance of the current source. In Fig. 9(b) because the load is in series with the current source it has no effect on very high impedances of the current source. Both these circuits should theoretically give the best possible phase noise performance.

At high frequencies circuit 9(c) is to be preferred. The collector base capacitance can be transformed into the two shunt arms of the pi network as shown in Fig. 6(d). In circuit 9(b) this cannot be achieved and in fact the combination of this capacitance and the load resistance  $R_0$  will start to give unwanted loop phase shift at the higher frequencies.

Another reason for preferring circuit 9(c) is because the input impedance (resistance) of the amplifier is much higher. Although as shown in section 6 the input resistance was shown not to affect the achievable phase noise performance when circuit voltage ratings were taken as the defining limitation, nonetheless a high input resistance does allow a higher overall dc to rf power efficiency for the oscillator to be achieved.

There is however an additional design requirement for circuit 9(c) to have sufficient loop gain to be able to oscillate. The voltage gain of the transistor stage is in fact  $R_{in}/R_E$ , where  $R_{in}$  is the input resistance of the tuned circuit. Because the emitter rf current through  $R_E$  is essentially the same as the collector rf current into  $R_{in}$  then  $R_E$  should be made slightly larger than  $R_{in}$  to ensure more output power than is made available through the pi network to the amplifier input. The voltage gain of the amplifier is then less than one. Thus the pi network must be made a step up voltage transformer with  $X_2 > X_1$  to keep the overall loop gain greater than unity. This is easily achieved in practice.

For further comparison with commonly used oscillator circuits the arrangements of Fig. 9(b) and 9(c) together with the pi network of Fig. 4(c) are shown redrawn in Fig. 10(a) and 10(b), Fig. 8(a) is clearly the traditional Clapp oscillator with the output load connected to the collector. This provides confirmation that the Clapp oscillator can be considered a preferred arrangement for low phase noise.

The redrawn Fig. 9(c) as shown in Fig. 10(b) indicates the output load to be in a place not normally chosen in the past. With the earth point as shown the load is floating but this could be moved to one or other end of the load without the phase noise performance changing. This arguably novel arrangement should perhaps be further investigated in practice to see if the improved phase noise promised by the above analysis can actually be confirmed.

## 8. CONCLUSIONS

In this paper good phase noise has been taken as the predominant requirement for any oscillator. Many apparently different single transistor oscillator types can be analysed using the same fundamental phase noise model. This model indicates the noise match and power match conditions required to give the best phase noise. Noise matching gives only a small advantage. Phase noise can be optimised almost automatically by minimising tuned circuit losses and maximising tuned circuit voltages. It is shown that the Clapp oscillator is a good design for low phase noise, but there is a modified Clapp design which could give better phase noise performance at high frequencies.

## REFERENCES

1. Underhill MJ. 1992 "Fundamentals of oscillator performance"  
Electron & Commun Eng. J. 4(4), 185-193.

2. Leeson DB, 1966. "A simple model of feedback oscillator noise spectrum"  
Proc IEEE, 54(2), 329-330.
3. Everard JKA, 1986, "Low-noise power efficient oscillators: theory and design"  
Proc IEEE, 54(2), 172-180.

APPENDIX

The phase noise model of section 2 relates to the basic oscillator shown in Fig. 2.1 as shown in this section.

The transfer function for any simple tuned circuit for  $Q=Q_0$  normalised for unity gain at the resonant frequency  $\omega_0=2\pi f_0$  is given by

$$G_o(s) = \left[1 + Q_o \left(\frac{s}{\omega_0} - \frac{\omega_0}{s}\right)\right]^{-1} \dots\dots\dots(A1a)$$

$$= \left[1 + jQ_o \left(\frac{f}{f_0} - \frac{f_0}{f}\right)\right]^{-1} \dots\dots\dots(A1b)$$

$$\approx \left[1 + 2jQ_o \left(\frac{f_m}{f_0}\right)\right]^{-1} \dots\dots\dots(A1c)$$

The closed loop transfer function from input to output in Fig. 2.1 is given by (forward path)/(1-loop path) and is

$$G_c(s) = k_o G_o(s) / (1 - k_o G_o(s)) \dots\dots\dots A1.2a$$

$$= k_o \left[1 + 2jQ_o \cdot f_m / f_0 - k_o\right]^{-1} \dots\dots\dots A1.2b$$

$$= k(Q_c / Q_o) \left[1 + 2jQ_c \cdot f_m / f_0\right]^{-1} \dots\dots\dots A1.2c$$

$$\approx (1/2jQ_o) \cdot (f_o / f_m)$$

$$\text{if } f_o / 2Q_c < |f_m| < f_o / 2Q_o \dots\dots\dots A1.2d$$

and where  $Q_o / Q_c = 1 - k_o$  and  $k_o$  is just less than unity.

The power spectrum transfer function from input to output is

$$\begin{aligned} |G_c(s)|^2 &= G(j\omega) \cdot G(-j\omega) \\ &= \frac{1}{4Q_o^2} \cdot \left(\frac{f_o}{f_m}\right)^2 \dots\dots\dots A1.3 \end{aligned}$$

The sideband phase noise at frequency  $f_m$  originating from  $FkT/2$  input phase noise is thus

$$N(f_m) = \frac{\frac{1}{2}FkT}{4Q_o^2} \cdot \left(\frac{f_o}{f_m}\right)^2 \dots\dots\dots A1.4$$

But  $S\phi(f)$  is the sideband noise normalised by division by the total carrier power  $P$ .

$$S\phi(f) = \frac{N(f_m)}{P} = \frac{\frac{1}{2}FkT}{4Q_o^2} \cdot \left(\frac{f_o}{f_m}\right)^2 \dots\dots\dots A1.5$$

This is equation 2.1.

Note that the value of the closed loop "multiplied"  $Q=Q_c$  is not needed to calculate the normalised phase noise  $S\phi(f)$   $P$  is actually the total phase noise  $N(f_m)$

integrated from  $f_m$ =minus infinity to plus infinity. To do this the more precise values of  $G_c(s)$  given in equation A1.2c is taken but with  $k_o=1$ . This gives

$$P = \frac{1}{2}FkT \cdot \left(\frac{Q_c}{Q_o}\right)^2 \cdot \int_{-\infty}^{+\infty} \frac{1}{1+4Q_c^2 f_m^2 / f_o^2} df_m \dots\dots\dots A1.6$$

By making the substitution

$$ofm = \tan \sigma \text{ where } 2\alpha = f_o / Q_c$$

is the closed loop filter bandwidths the integral is found to be  $\pi\alpha$ . Equation A1.6 thus gives

$$Q_c = \frac{4PQ_o^2}{FkT\pi\alpha} \dots\dots\dots A1.7$$

For the example of  $F=2$ ,  $f_o=100\text{MHz}$

$$\text{and } PQ_o^2 = 0.1 \text{ and } kT = 4 \times 10^{-21}$$

as used in section 6.

$$Q_c = 10^{12} / 2\pi = 1.6 \times 10^{11} \dots\dots\dots A1.8$$

At 100MHz the filter bandwidth is

$$2\alpha = f_o / Q_c = 10^8 / 1.6 \times 10^{11} = 0.63\text{mHz} \dots\dots\dots A1.9$$

With a bandwidth of 0.63 milli-Hertz even pure phase noise appears as a pure carrier signal.

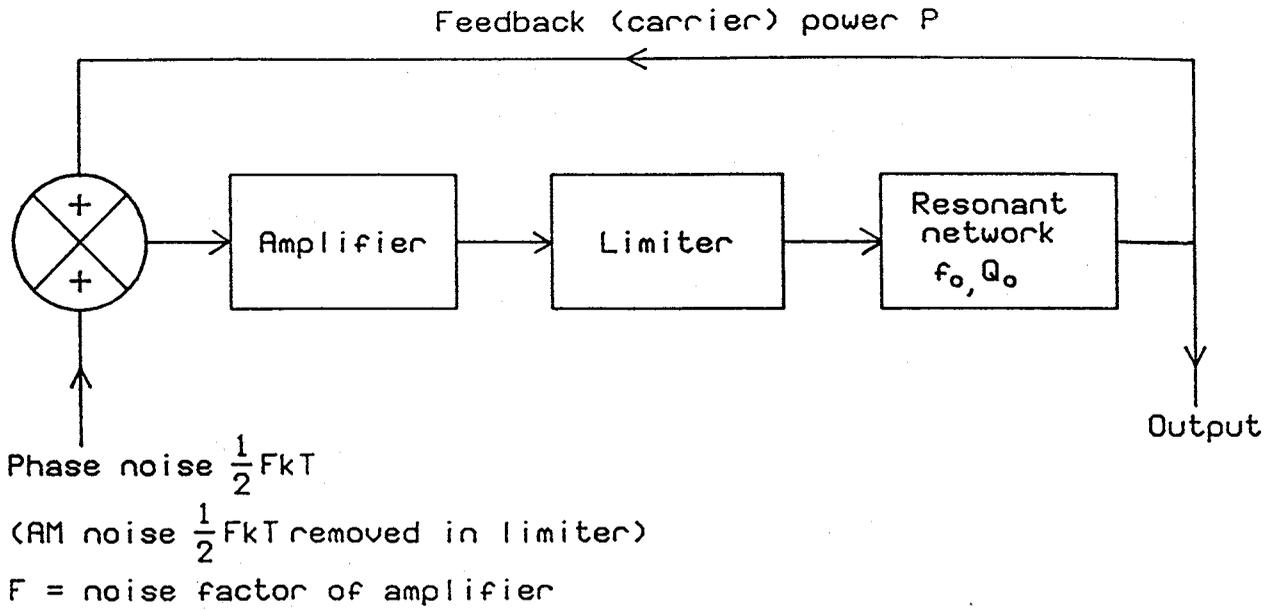


Fig 1 Basic Oscillator model for phase noise

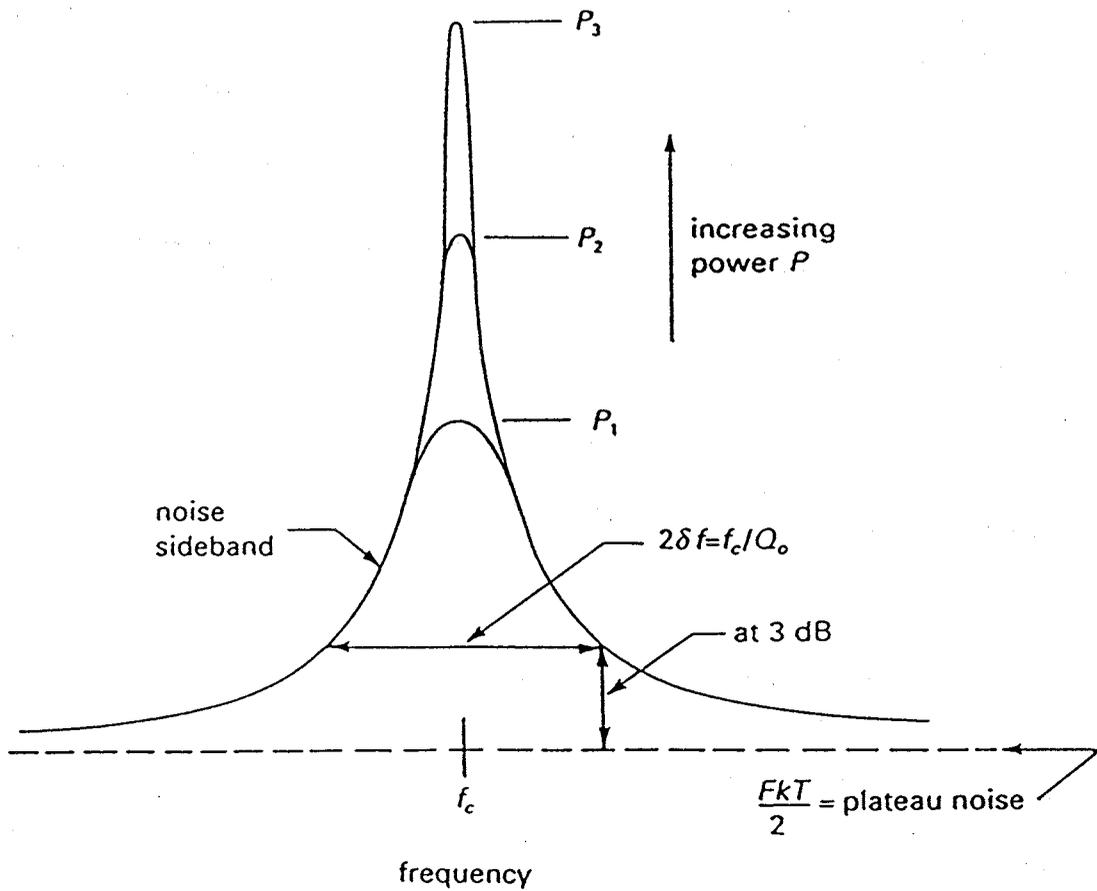


Figure 2 : Oscillator phase noise spectrum

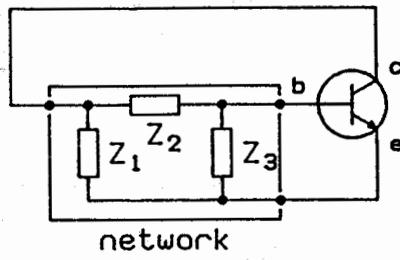


Fig 3 Basic single transistor oscillator

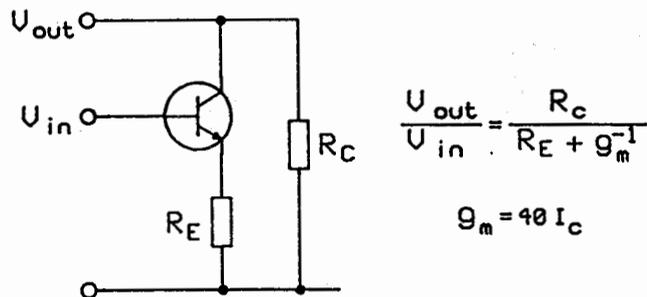


Fig 5 Basic BJT amplifier stage

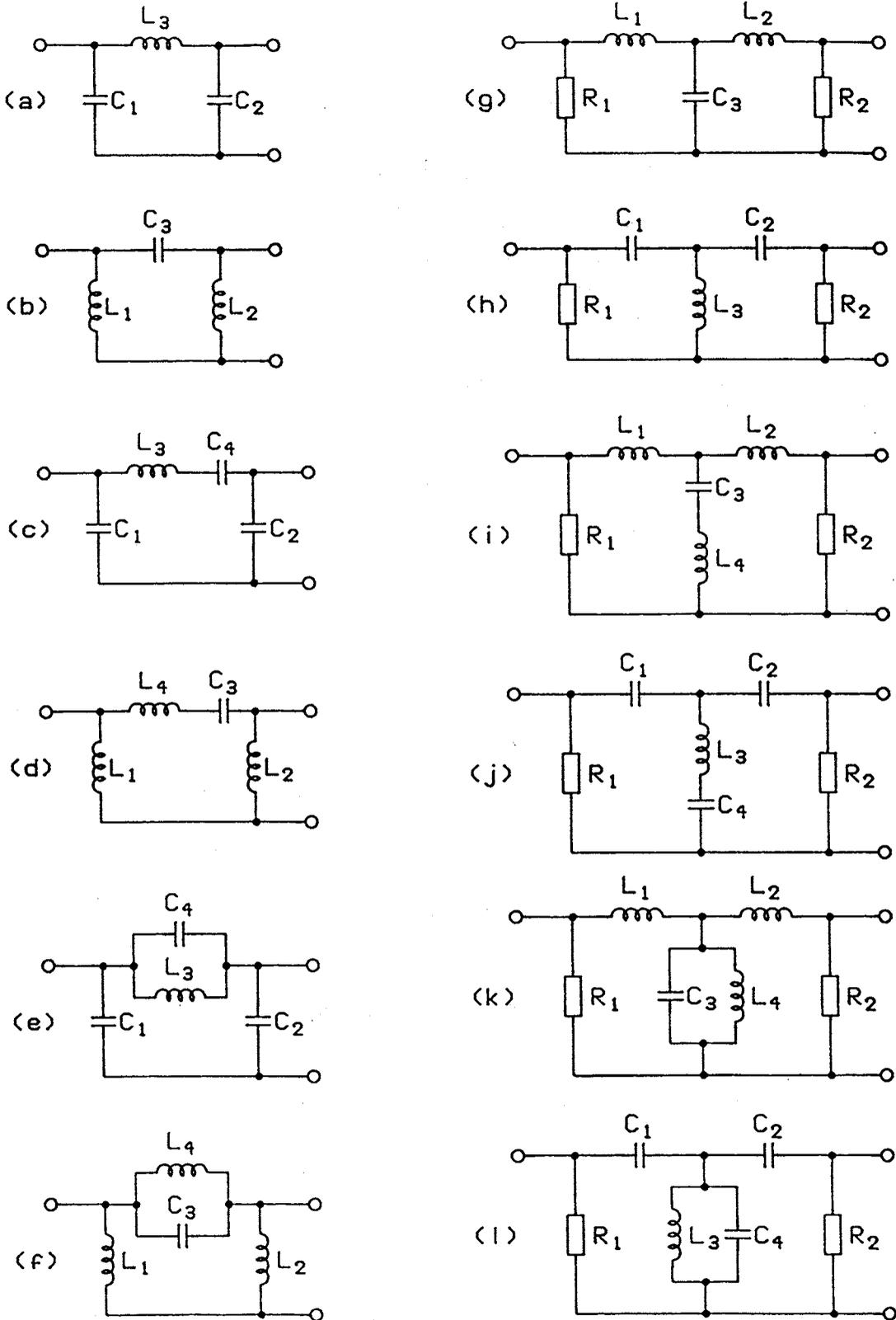


Fig 4 Basic Oscillator Tuned Circuits

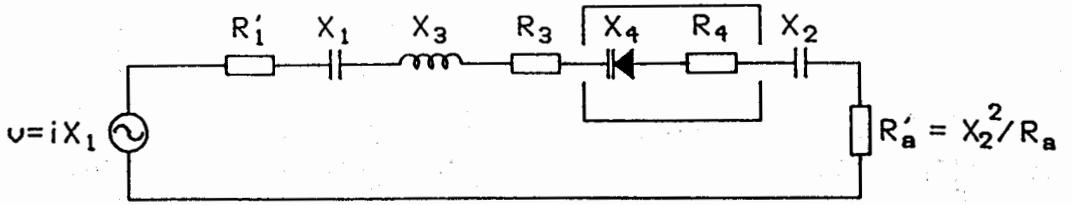
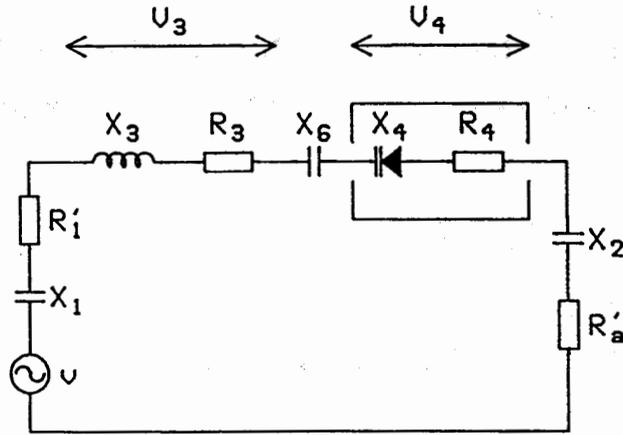


Fig 7 Series equivalent of VCO pi network

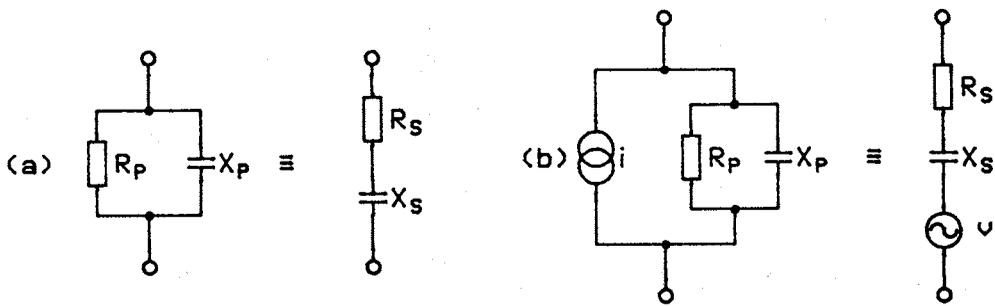


$$PQ^2 = U_3^2 / 4 (R_4' + R_3 + R_1')$$

$$\text{Tuning Range } \frac{\Delta f}{f} \approx \frac{f_{\max} - f_{\min}}{f_{\max} + f_{\min}} = \frac{\Delta C}{2C} = \frac{\Delta X_4}{2X_3}$$

$$\text{Where } \frac{U_4}{U_3} = \left| \frac{X_4}{X_3} \right| = \left| \frac{X_4}{X_6 + X_1 + X_2 + X_4} \right|$$

Fig 8 PQ<sup>2</sup> improvement by restricting VCO range.



$Z_P = Z_S = Z$  and  $Q = R_P/X_P = X_S/R_S$  then  $V = iZ \approx i(jX)$ .

$R_P R_S = X_P X_S \approx X^2$  and  $X_P = X_S (1 + Q^{-2}) \approx X_S = X$  if  $Q$  large for " $\approx$ ".

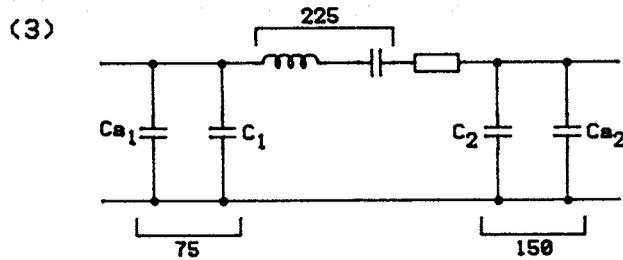
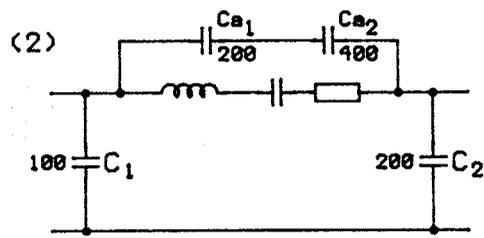
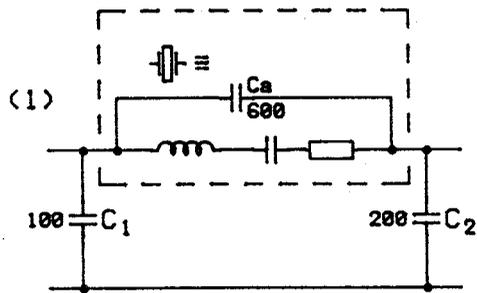
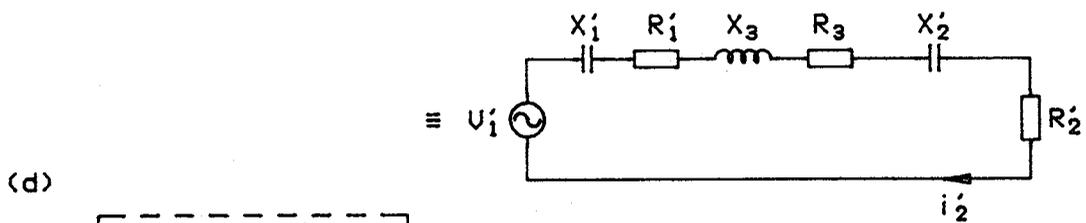
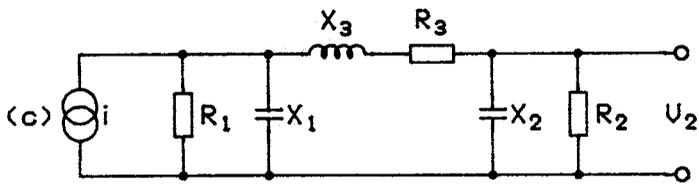


Fig 6 Circuit Transformations in a PI network

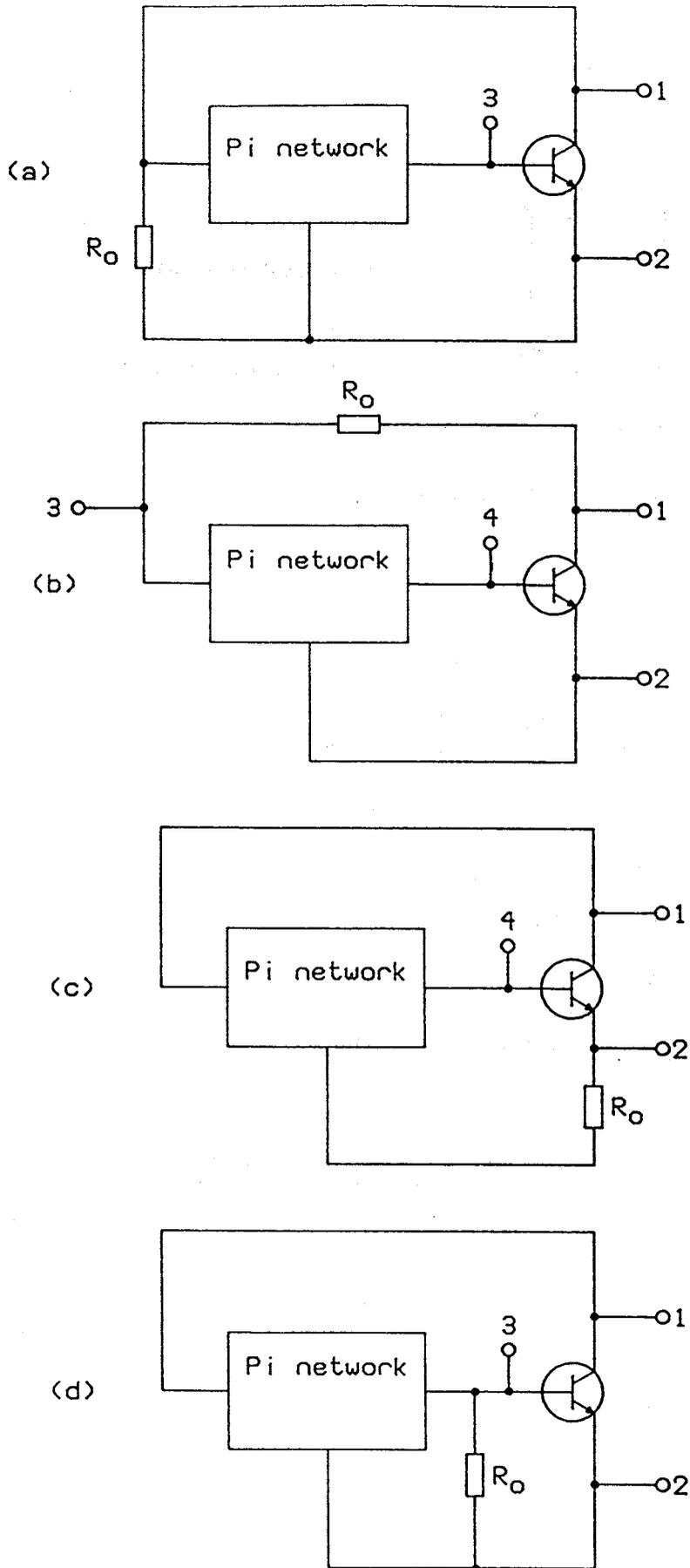
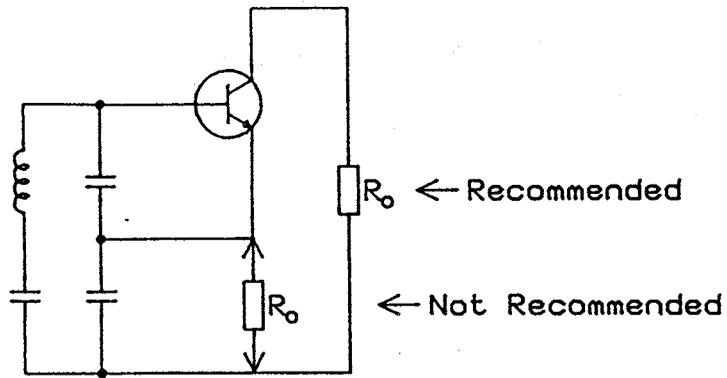
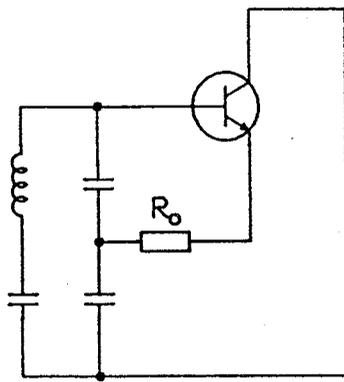


Fig 9 Output load positions and ground points in a single Transistor oscillator



(a) Conventional Clapp Oscillator



(b) "New" Clapp Oscillator

Fig 10 Redrawn Oscillator Circuits 9b and 9c

## ANALYSIS OF MEDIUM-TERM FREQUENCY INSTABILITY OF HIGH STABILITY QUARTZ OSCILLATORS

Barbara Kalinowska, Barbara Gniewińska

*Tele & Radio Research Institute, Ratuszowa 11, 03-450 Warszawa, Poland*

### 1. Introduction

Various types of high stability quartz oscillators have been elaborated and produced on a small scale in Tele & Radio Research Institute over last twenty years.

Systematic investigations of long-term frequency stability of the oscillators have been carried on since several years.

Typical parameters /guaranteed and *practically obtainable*/ of currently manufactured ITR OCXO's are shown in Table 1.

Exemplary measurement results of short-term frequency stability, for average time from 1ms to 100s, are shown in Fig. 1.

According to definition, given in actual standardization documents, long-term instability is the frequency change observed in conditions assuring elimination of other influences on oscillator frequency, especially temperature and supply voltage. During investigation carried on in ITR laboratory there were observed slow, lasting for hours, frequency fluctuations making difficult the determination of true long-term characteristics of measured oscillators.

The results of investigation and calculation procedures, used for data processing in order to increase a confidence level of frequency/time characteristics determination, are presented in this paper.

### 2. Method of measurement

Block diagram of the system for measurement of long-term frequency stability is shown in Fig. 2.

Basic algorithm of measurement procedure is given in Fig. 3.

Investigated oscillators are measured successively each two hours.

Ambient temperature and pressure are measured during each testing cycle.

All measurement data are collected in computer memory in the following files :

- date of measurement (day and hour),
- temperature,
- pressure,
- relative frequency standard deviation for each oscillator.

Frequency resolution of the applied measurement method, for nominal frequency 5MHz and average time 10s, is

- (3-5)E-12 for higher class oscillators,
- (0.5-1)E-11 for standard oscillators ( e.g. OCXO 100).

The measurements are made during continuous operation of investigated oscillators in unconditioned

room. Observed typical temperature changes are not higher than 1K/24h and pressure changes correspond to changes of atmospheric pressure ( in practice they don't exceed 20hPa/24h ).

Taking into account parameters of tested oscillators, it can be expected that, for oscillators of high long-term stability, frequency changes caused by external influences (temperature, pressure etc.) are of the same order like changes resulting from intrinsic oscillator properties.

In connection with that the procedure used for determination of oscillator 24-hours frequency stability, basing on linear regression method, is charged with considerable error.

### 3. Results of investigation

Typical frequency/time characteristics for oscillators of various class, obtained from previously created files of measurement data ( together with temperature and pressure changes existing during measurement ) are presented in Fig. 4 A, B, C.

Data files are characterized by the following values :

- correlation coefficient

$$r_{xy} = \frac{S_{xy}}{S_x \cdot S_y} \quad (1)$$

where:

$$(S_x)^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2;$$

$$(S_y)^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2;$$

$$S_{xy} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y});$$

- slope of linear regression line  $x$  against  $y$

$$b_{x/y} = r_{xy} \cdot S_y / S_x \quad (2)$$

- slopes of linear regression lines  $z$  against  $x$  and  $y$ , when  $x$  and  $y$  are correlated with correlation coefficient  $r_{xy}$

$$b_{z/x} = \frac{S_z}{S_x} \cdot \frac{r_{zx} - r_{xy} \cdot r_{zy}}{1 - (r_{xy})^2};$$

$$b_{z/y} = \frac{S_z}{S_y} \cdot \frac{r_{zy} - r_{xy} \cdot r_{zx}}{1 - (r_{xy})^2}; \quad (3)$$

The procedure used for determination of above mentioned parameters and for introduction of necessary correction is shown in Fig.5.

Corrected frequency/time characteristics, for various oscillator types, are given in Fig.6-8.

Similar characteristics obtained for two oscillator models (of increased sensitivity to temperature and pressure) are presented in Fig.9-10. In one of these oscillators temperature sensitive SC-cut resonator excited on B-mode (Fig.9) was used and the second (Fig.10) with special SC-cut resonator of increased sensitivity to pressure was equipped.

Values of correlation coefficient determined before and after correction process are collected in Table2.

Obtained results show that described correction method gives positive effects only in the case when one explicitly dominating influence exists and for OCXO's of lower class.

For higher class oscillators, frequency/time characteristic after correction process is still correlated with pressure and temperature changes. It shows the presence of so far neglected additional influence, correlated with temperature and pressure.

Such influence on measurement results could have humidity, which effects are of quite different and difficult to predict time dependence and for that are difficult to remove by using the appropriate corrections. The humidity changes can affect negatively evaluation of  $b_T$  and  $b_P$  coefficients, modifying thermal isolation of the oscillator thermostat.

Approaching the residual (after correction process) fluctuations of the frequency/time characteristic of the order of  $1E-11$  it is necessary to consider the Table 1. Technical parameters of ITR OCXO's

Parameter \ Type	OCXO 100	OCXO 80	OCXO 2001
resonator	SC, n1	SC, n3	SC, n3
long-term instability [ $\Delta f/f/24h$ ]	< 5E-9 < 1E-9	< 5E-10 < 1E-10	< 2E-10 < 5E-11
short-term instability [Allan standard dev.]	< 5E-11/1s < 1E-11/(1-100)s	< 5E-12/1s < (1-3)E-12/1s	< 2E-12/1s < 1E-12/(1-100)s
frequency change /in temperature range/	1E-8 (0-50°C)	5E-9 (-10 - +60°C)	5E-10 (-10 - +60°C)

Table 2. Exemplary calculation results

Type	no	File parameters		Correlation coefficients (before/after correction)			Calculated slopes		24h freq. change	stddev from st.line
		n	$r_{TP}$	$r_{fT}$	$r_{fT}$	$r_{fP}$	$b_f / T$ [1/K]	$b_f / P$ [1/hPa]		
OCXO 100	374	53	-0.646	0.952 0.985	0.314 0.512	0.218 -0.03	-8E-11	4E-11	5.3E-10 5.4E-10	2.4E-10 1.4E-10
OCXO 80	3421	29	-0.339	0.989 0.993	-0.324 -0.248	0.64 0.61	-8E-11	6E-12	2.5E-10 2.4E-10	0.5E-10 0.4E-10
OCXO 2001	3410	29	-0.339	-0.372 -0.33	0.424 0.098	-0.318 -0.242	6E-11	-7E-13	-16E-13 -16E-13	7.3E-11 6.7E-11

possibility of slow fluctuation process contribution (of high power exponent;  $S_y \approx F^{-2}$ ).

In Table3 there are shown exemplary measurement results of the short-term frequency stability of OCXO2001 oscillator, made for average times 1-1000s. Allan variance ( $\sigma_A$ ) and three-samples variance ( $\Sigma$ ) are calculated.

Taking into account specific character of variances changes and value of their ratio, it seems that slow fluctuation processes take place in the investigated oscillators.

#### 4. Conclusions

Presented results should be treated as an intermediate stage and further investigation, allowing for better approach to real fluctuation processes existing in tested oscillators, should be continued.

For that purpose it is necessary to conduct investigations of hermetically encapsulated higher class OCXO's and to try new correction methods.

#### REFERENCES

1. J.R. Vig, F.L. Walls, Fundamental limits on the frequency instabilities of quartz crystal oscillators, IEEE, 1994
2. J.J. Gagnepain, Fluctuations and noises in quartz resonators and oscillators, Prace ITR, 99-100, 1985
3. B. Kalinowska, S. Gawor, Automated system for investigation of long term instability of precision quartz oscillators, 8th Conference PIEZO'94, Zakopane, 1994

Table3. Allan variance and three-samples variance for OCXO 2001

$\tau$ [s]	1	10	100	1000
$\sigma_A$	8.5E-13	5.7E-13	6.2E-13	13E-13
$\Sigma$			4.7E-13	5.9E-13
$\Sigma/\sigma_A$			0.76	0.45

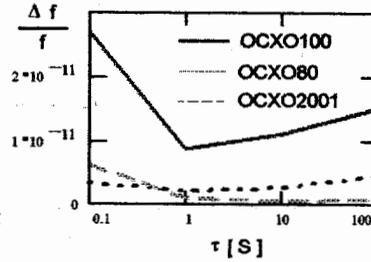


Fig.1 Short-term frequency stability of ITR OCXO oscillators

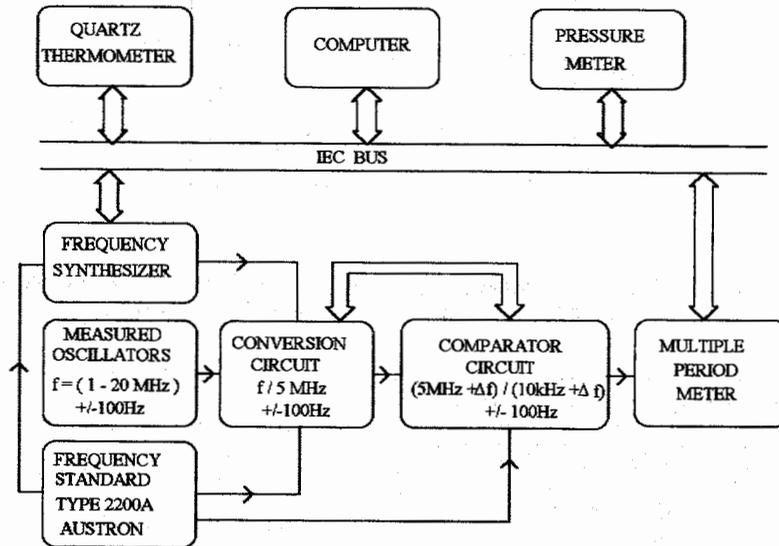


Fig. 2 Block diagram of the system for long-term frequency stability measurement

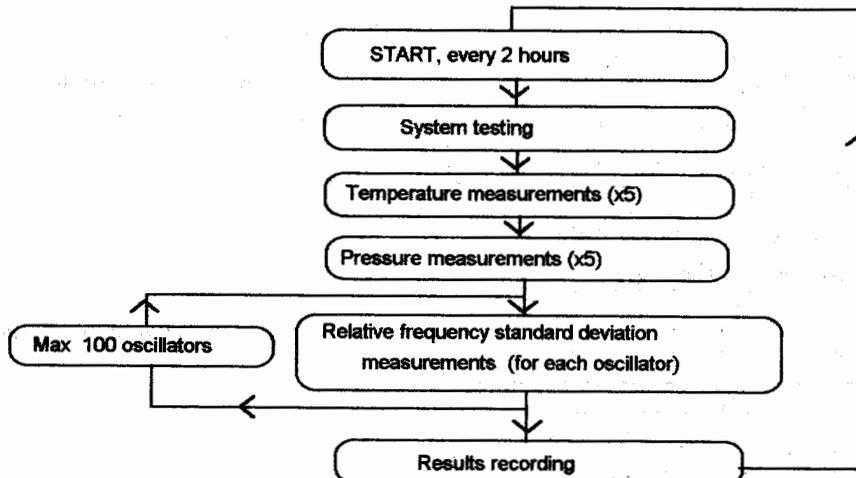


Fig.3 Basic measurement procedure

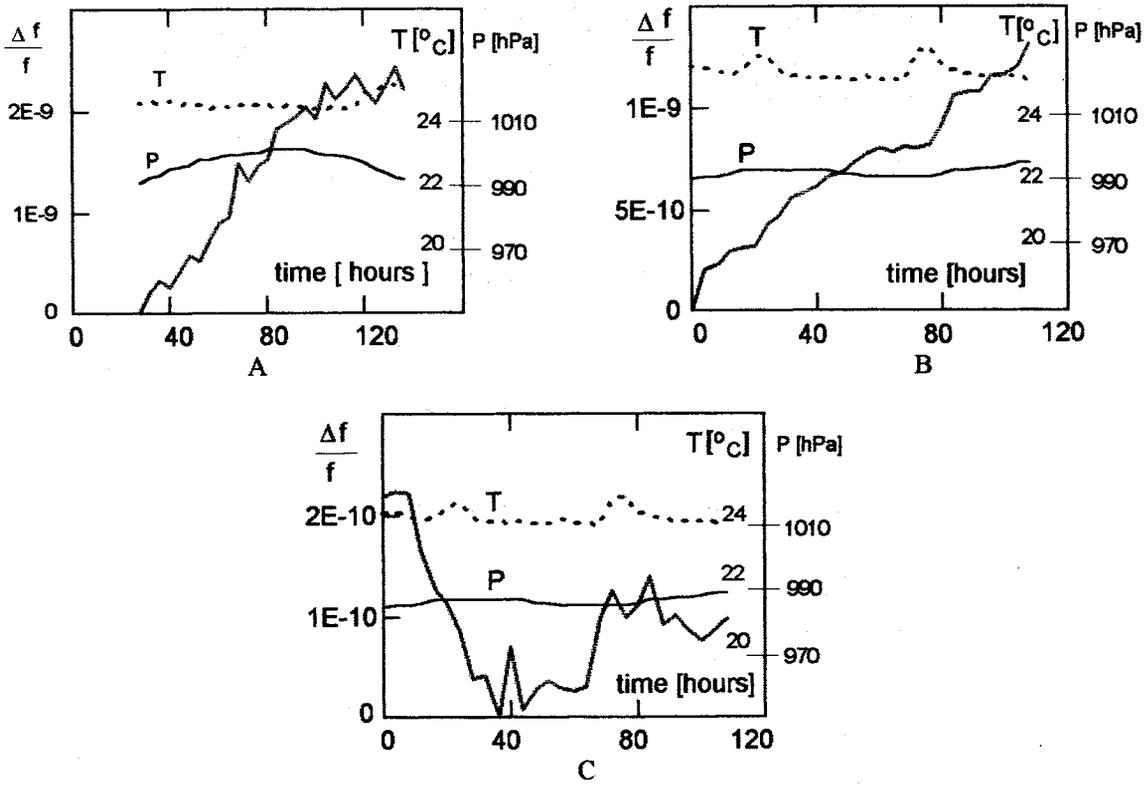


Fig.4 Measured frequency / time characteristics for various OCXO's  
 A - OCXO 100 , B - OCXO 80, C - OCXO 2001

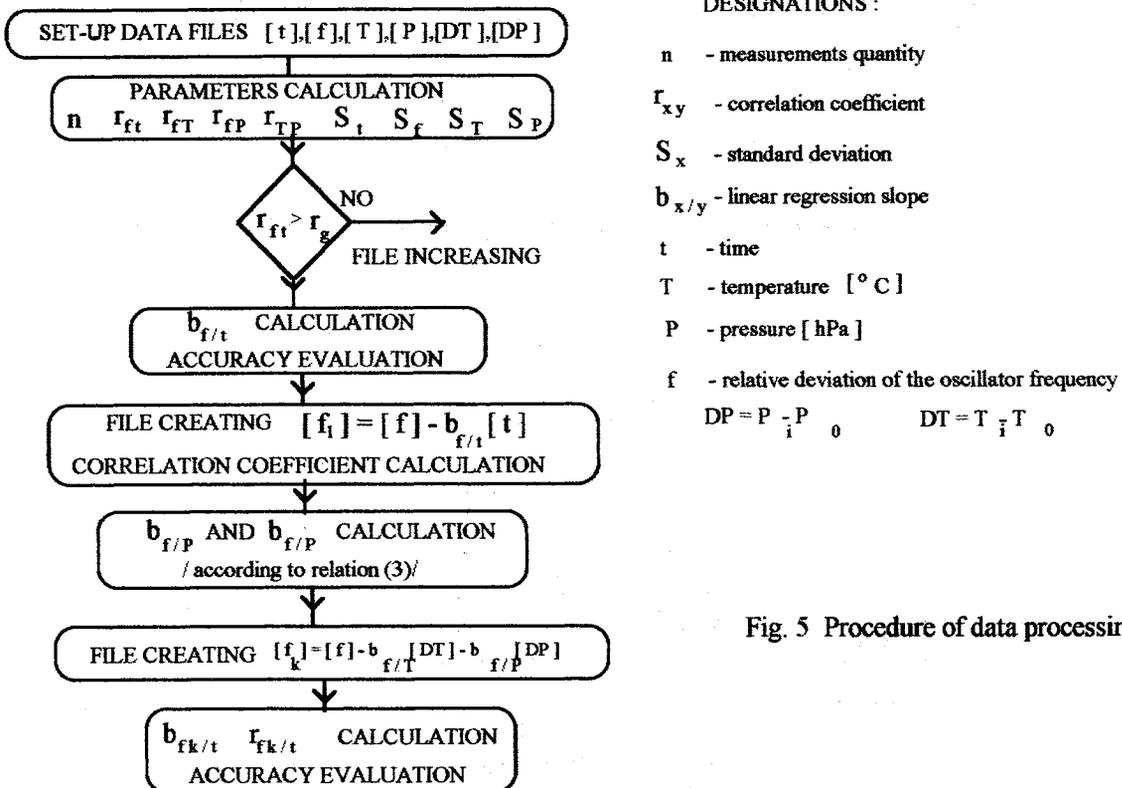


Fig. 5 Procedure of data processing

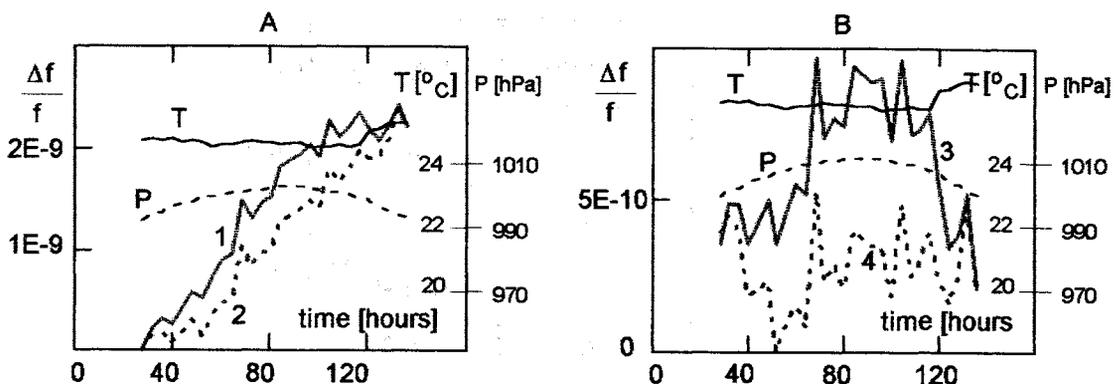


Fig. 6 Frequency/time characteristic of OCXO 100  
 1 - measured, 2 - corrected against temperature (T) and pressure (P),  
 3 - corrected against f/t response line,  
 4 - corrected against temperature (T), pressure (P) and f/t regression line

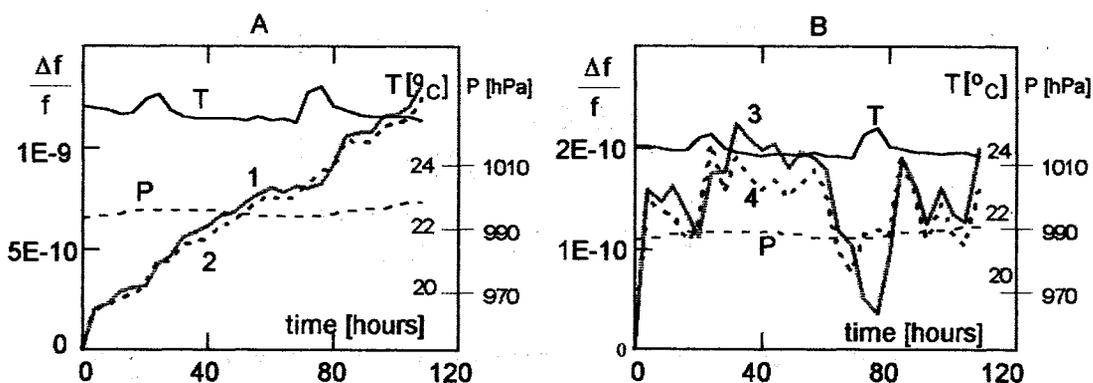


Fig. 7 Frequency/time characteristic of OCXO 80  
 1 - measured, 2 - corrected against temperature (T) and pressure (P),  
 3 - corrected against f/t response regression line,  
 4 - corrected against temperature (T), pressure (P) and f/t response regression line

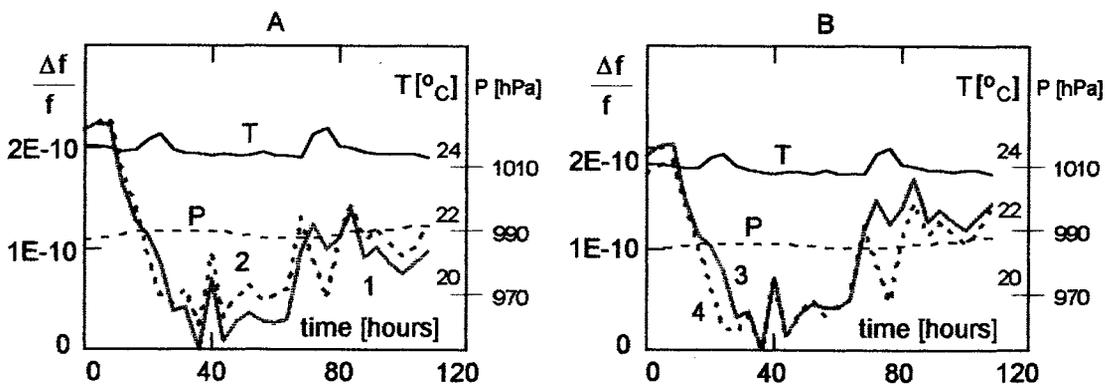


Fig. 8 Frequency/time characteristic of OCXO 2001  
 1 - measured, 2 - corrected against temperature (T) and pressure (P),  
 3 - corrected against f/t response regression line,  
 4 - corrected against temperature (T), pressure (P) and f/t response regression line

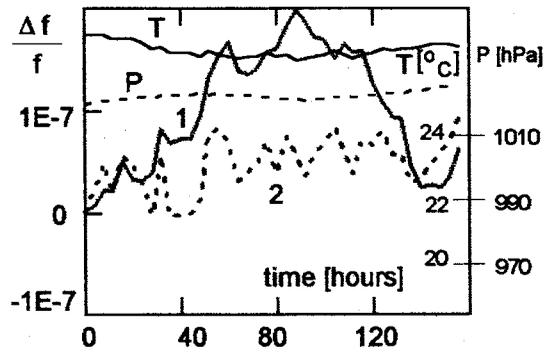


Fig. 9 Frequency/time characteristic of OCXO 80T (sensitive to temperature)  
1 - measured, 2 - corrected against temperature (T) and pressure (P)

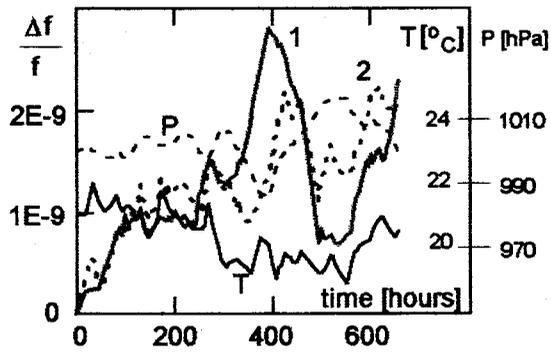


Fig. 10 Frequency/time characteristic of OCXO 2001P (sensitive to temperature)  
1 - measured, 2 - corrected against temperature (T) and pressure (P)

## AN INSITU TECHNIQUE FOR THE RESOLUTION OF AGING CONTRIBUTIONS BETWEEN QUARTZ RESONATORS AND OSCILLATOR CIRCUITS

Gregory Weaver, William Hanson & Timothy Wickard  
Piezo Crystal Company, Carlisle, PA., U.S.A.

### ABSTRACT

Historically, many valid attempts at assessing the aging of quartz crystal oscillator systems have been met with frustration. While the quartz resonator's aging mechanisms, in isolation, have gained due attention, the insitu aging process within an oscillator circuit has received incomplete characterization. As producers of both precision quartz crystals and oscillators, we have been confronted with this uncertainty both internally and with our external resonator customers. Crystals which have been scrutinized for aging drift have yielded poor oscillator performance within well established circuit designs. In a more practical sense, the uncorrelated behavior observed between isolated resonator measurements and actual oscillator performance leads to an uncomfortable segmentation of the aging problem. This has hampered the development of a standard model for precision oscillator aging.

Our method for independently resolving the contributions of the aging drift of the quartz resonator from the contributions of the sustaining oscillator circuit relies on two corollaries of established oscillator theory:

**1.) Resonator aging is a manifestation of a continuous drift in the series resonant frequency of the quartz crystal. However, the phase slope associated with the Q of the resonator remains predominately unchanged.**

**2.) Aging of the oscillator components other than the crystal will result in a net phase shift across the resonator accumulated over time.**

Using these principles, we have established a measurement system which not only records the per unit time frequency drift of a quartz oscillator system, but also records the change of the phase shift across the resonator. By establishing the value of the operating phase slope of the resonator/oscillator system, the contribution of the oscillator circuitry to aging can be inferred and removed. Consequently, the frequency drift attributable to the resonator is separable even though the resonator is within the closed-loop oscillator system.

In practice, our method will facilitate the optimization of oscillator aging through unambiguous experimentation of parameters, such as drive and load. Additionally, the verification of circuit and resonator aging models can be performed under closed-loop conditions. Since the system will allow selective substitution of oscillator components, a more rigorous understanding of the oscillator aging phenomenon will develop. The paper will elaborate, by sharing our method's results on the initialization behavior of a precision SC cut crystal oscillator. In addition, two experiments with OCXO thermal performance will be presented to illustrate the versatility of the insitu technique.

## INTRODUCTION:

Aging in quartz crystal oscillators has been a subject of study in frequency control from almost the beginning. In fact, the first industrial conference relating to frequency control was called on July 11th, 1944 to address the "ageing problem" with military communications.<sup>1</sup> Aging is best described as the time dependent change of frequency with all other external influences assumed to be quasi-static. As the decades have proceeded, aging remains as the principle factor limiting the useful life of a quartz controlled frequency source.

Since it was quickly demonstrated that resonator processing could dramatically effect the aging of quartz oscillators, the quartz crystal manufacturers bore the brunt of the assignment to reduce aging. Two fundamental aids to assist the research efforts of quartz crystal design grew out of this demand. The first was an electrical equivalent model which could map the behavior of a quartz crystal into the oscillator circuit. This electrical model was particularly well established by E. Hafner in 1969.<sup>2</sup> The second was a collection of measurement systems which would allow the quartz crystal designer to estimate the parameters of the quartz crystal's electrically equivalent model.

The evolution of measurement systems was guided by two concerns. The ability to resolve the highly frequency dependent electrical parameters of the quartz crystal in the narrow region around resonance and the ability to determine the electrical parameters when the crystal would experience the effects of a reactive load. The first generation of measurement equipment, known as "Crystal Impedance" meters, were in fact series resonant oscillators. The principle for the use

of the CI meter was that once the drive and phase conditions were established by a substitution resistor, the inserted resonator could be electrically measured. However, as improved resonator designs for aging necessarily increased the Q of the crystals, the limitations of using an active measurement system such as the CI meters became apparent. In particular, the CI meters had no independent source for absolute frequency accuracy and the variability of set-up made precision aging measurements impossible.

The need to include absolute frequency accuracy into resonator measurements led to the invention of the transmission or passive style measurement systems. Specifically, an external frequency source such as a synthesizer is used to provide a stable frequency signal into an arrangement of passive elements. The zero-phase "pi" network transmission test set described by C. Adams and others in 1968 is an example of this approach.<sup>3</sup> The advantages presented by the passive transmission method allowed resonator designers to precisely sweep the impedance of the crystal with a known interrogation frequency. Additionally, the use of the vector-ratio meter provided a direct phase measurement capability.

Nonetheless, as the aging capability of quartz resonators continued to improve by magnitudes, small uncertainties in phase within the transmission systems limited the resolution and repeatability of aging measurements. Recently, there has been the introduction of balanced bridge style systems in which the resonator is unambiguously set to an operating condition. The ability of well designed balanced bridge systems to resolve aging in the ppb has been demonstrated by Hafner and others.<sup>4</sup>

Coordinate with the iterative improvements made in quartz resonators, oscillator designers began considering the stability of the entire oscillator system. However, because of the complex aggregate of components constituent in any oscillator implementation, it became clear that the analysis of non-resonator aging within oscillators would require extensive simulation. Most of these simulation strategies involved some form of reduction technique on the circuit topology to ascertain the relevant oscillator model. An example of a reduction technique is the "negative resistance" model or "oscillator" concept documented by B. Parzen.<sup>5</sup>

Fundamental to all oscillator reduction approaches which consider a quartz resonator is the assertion that  $dX/df$  for all components other than the crystal is practically zero around the narrow region of steady-state operation. Another way of regarding this assertion is that the  $Q$  of the resonator dominates the  $d\phi/df$  characteristic of the steady-state closed loop. With these factors in mind, it becomes convenient to again subdivide the consideration of oscillator circuit behavior into "the crystal and then everything else".

It is our contention that the apparent schism in the consideration of aging mechanisms in the quartz controlled oscillator has been promoted by the natural ability to divide frequency stability behavior at the resonator. More interesting, we feel that the use of the passive transmission techniques to directly measure resonators has presented an obstacle into the resolution of oscillator aging by raising the concern that the measured behavior is somehow different than that applied by the continuous dynamics of the

active oscillator system.<sup>6</sup>

As designers of both quartz crystals and precision oscillator, we are familiar with inconsistencies in aging assessment between resonators and oscillators. Specifically, the short-term or initialization period of a resonator is generally unknown prior to insertion into an oscillator. For our methods at Piezo, the use of  $\pi$  and reflection type transmission systems for resonator measurement can only produce sequenced frequency deltas through conditioning treatments. Therefore, the time record of the resonator aging is unmeasured until installed into the electronic circuit.

The measurement technique discussed in our paper will demonstrate a method to assess the insitu aging of the resonator in an active oscillator system. We will first outline the theory which we believe demonstrates the feasibility of distinguishing the resonator aging contribution from the rest of the oscillator. We will then attempt to validate the technique empirically by examining the results of several experiments. Finally, we will propose a course of further studies in oscillator aging that will be made available through the use of this insitu technique.

## **THEORY OF THE INSITU TECHNIQUE**

The ability to resolve the aging of an operating oscillator into circuitry and resonator contributions would greatly aid the development and verification of theoretical models. We are particularly interested in the effects of initialization and retrace of an operating oscillator during the first several days. These effects can represent a large portion of an oscillators first year aging excursion. Currently, a physical explanation

for the initialization effect of a quartz oscillator is not well understood.

By considering several aspects of the “negative resistance model” oscillators and the electrical equivalent of a crystal, it will be shown that the following corollaries are established:<sup>7</sup>

**1.) Resonator aging is a manifestation of a continuous drift in the series resonant frequency of the quartz crystal. However, the phase slope associated with the Q of the resonator remains predominately unchanged.**

**2.) Aging of the oscillator components other than the crystal will result in a net phase shift across the resonator accumulated over time.**

The causes of aging in crystal oscillators has been investigated extensively from a resonator point of view. Vig and Meeker state that the primary causes of crystal oscillator aging are mass transfer, stress relief in the mounting structure, changes in the electrodes, package leaks and changes in the quartz.<sup>8</sup> Mass transfer is generally described by contamination transfer effects and ascribed to the inductive component of the electrical equivalent circuit. For example, the mass of material diffusing onto a 10 MHz 3rd overtone SC cut resonator aging -  $1 \times 10^{-10}$ /day is  $1.4 \times 10^{-14}$  gm/mm<sup>2</sup> or a  $5.35 \times 10^{-5}$  angstroms increase in thickness per day. This is 53 atoms per million of mass equivalent quartz. Thin film relaxations and mechanical stress relief due to plate distortions are generally attributed to the motional capacitance.<sup>9</sup> All of these causes predominately result in a time related change in the series reactive component of the electrically equivalent resonator.

The leaking of the resonator enclosure is not consistent with this assertion. A substantial change in equivalent resistance is experienced by conductance changes in the thin metal film of the electrode reacting with air. Additionally, the acoustics of the wave will be attenuated if the resonator is a quasi-shear mode. Because the Q of the resonator changes during leaking, the phase slope of the oscillator’s steady state operation cannot be used to calibrate the separation of the oscillator frequency drift. To this extent, the behavior of the insitu technique to leaking must be treated as a special case.

The justification for the “insitu” technique to directly assess the oscillator circuitry aging begins by examining how the phase across the resonator changes with respect to the electrical parameters of an equivalent model. Fig.1.) illustrates a modified model of the standard electrical equivalent for a crystal.<sup>10</sup> In this model, the reactive components associated with the motional series resonant section are reduced to a reactance element  $X_m$ . The admittance of the network of Fig 1.) can be expressed as:

$$Y = -\frac{j}{X_{co}} + \frac{R_s}{R_s^2 + X_m^2} - \frac{j * X_m}{R_s^2 + X_m^2}$$

which can be rewritten to separate the real and imaginary terms:

$$Y = \frac{R_s * X_{co}}{X_{co} * (R_s^2 + X_m^2)} + \frac{j * (R_s^2 + X_m^2 + X_m * X_{co})}{X_{co} * (R_s^2 + X_m^2)}$$

Therefore the change in phase across the crystal resonator is:

$$\phi = \arctan \left[ \frac{R_s^2 + X_m^2 + X_m * X_{co}}{R_s * X_{co}} \right]$$

This expression for  $\phi$  can then be differentiated with respect to  $X_m$  and to  $R_s$  so that the sensitivity of the phase can be analyzed with respect to changes in the equivalent electrical parameters. The sensitivity of the phase change across the crystal with respect to aging of the electrical equivalent parameters must be shown to have a minimal effect. Otherwise it would be impossible to resolve the independent contributions of the oscillator circuitry from the crystal based on phase change of the oscillator closed loop.

First, the sensitivity of  $X_m$  will be examined. The resonance frequency of the motional section of the crystal is:

$$f_{sr} = \frac{1}{2 * \pi * (L_m * C_m)^{1/2}}$$

Where  $L_m$  and  $C_m$  are the inductive and capacitive components of the motional section and the derivative of  $f_{sr}$  with respect to  $C_m$  is:

$$\frac{\delta f_{sr}}{\delta C_m} = \frac{1}{4 * \pi * (L_m^{1/2} * C_m^{3/2})}$$

Likewise the expression for  $X_m$  as a function of  $C_m$  and  $f_{sr}$  is:

$$X_m = \frac{1}{C_m} * \left[ \frac{2 * \pi * f}{2 * \pi * f_{sr}^2} - \frac{1}{2 * \pi * f} \right]$$

Note that the variable  $f$  must now be included as  $X_m$  is a function of the resonator's operating frequency. The derivative of  $X_m$  with respect to  $C_m$  is:

$$\frac{\delta X_m}{\delta C_m} = \frac{1}{C_m^2} * \left[ \frac{2 * \pi * f}{2 * \pi * f_{sr}^2} - \frac{1}{2 * \pi * f} \right]$$

With the substitution of some typical SC cut parameters for  $L_m$ ,  $C_m$  and  $f_{sr}$ ; it is determined:

$$C_m = 182.5 aF$$

$$L_m = 1.38796 H$$

$$f_{sr} = 10.0000014 MHz$$

$$\frac{\delta f_{sr}}{\delta C_m} = -2.74 * 10^4 Hz/aF$$

$$\frac{\delta X_m}{\delta C_m} = -9.55 * 10^{-1} Ohms/aF$$

Which demonstrates that there are four orders of magnitude between the sensitivities. It must also be noted that  $\delta X_m / \delta C_m$  is a function of the operating frequency. Since  $X_m$  increases on either side of  $f_{sr}$ , the sensitivity of  $dX_m / dC_m$  should be examined at the oscillator's operating point. The offset frequency  $f$  from  $f_{sr}$  in the example above was +10 Hz, which represents about a +70 degree of phase shift from the operating point of the series type oscillator used in the experiments.

By reiterating that  $X_m$  reduces as the operating point approaches  $f_{sr}$ , the results of the  $\delta C_m$  comparisons offer strong confidence that the small changes in motional reactance from resonator aging will not skew the phase information across the in-circuit crystal.

The examination of the sensitivity of changes in  $R_s$  to the phase across the crystal require a more developed model of the resonator's behavior in-circuit. Fig. 2.) is a block diagram of the "negative resistance" oscillator model.<sup>11</sup> The model divides the

behavior of the oscillator loop into two sections. The first section (left of the dashed line) are the components which have the narrow band frequency dependence of a resonator. In a crystal oscillator, this would be the real part of the impedance of the crystal's electrical equivalent model,  $R_e$ . The reactive part of the crystal's impedance is represented by  $X_e$ . The balance of the oscillator's behavior is represented in the second section (right of the dashed line) and is considered "frequency independent" over the narrow operational bandwidth of the crystal.

The frequency independent section of the "negative resistance" model has four components, two reactive and two real.  $X_L$  represents the series load reactance often used to tune the oscillator to an absolute frequency.  $X_{LL}$  is a lumped effective equivalent of all other reactive components of the oscillator circuitry. Likewise,  $R_{LLp}$  is the sum of all effective circuit losses, including energy to the output circuit.  $R_{LLg}$  is the "negative resistance" which represents the amplifier gain necessary to support the steady-state oscillator condition.

The fundamental relationships of the model establish the oscillation criteria.<sup>12</sup>

$$\sum R=0; R_{LLg} = -R_{LLp} - R_e$$

$$\sum X=0; X_m = X_L + X_{LL}$$

The first relationship states that at steady-state conditions the closed loop gain will be exactly one. The second allows no accumulated reactance around the loop. This is equivalent to the statement that oscillators have  $n2\pi$  closed loop phase shift, where  $n$  is an integer.

The reduction of the oscillator circuit into the "negative resistance" model allows the sensitivity of the loop to real changes to be analyzed through the use of two parameters;  $d\phi/df$  and  $dX/df$ .  $d\phi/df$  is the phase slope of the oscillator loop, including the crystal.  $dX/df$  is the reactance slope of the frequency dependent component section of the model, the crystal's  $X_e$ .

The sensitivity of the in-circuit phase change across the crystal to  $R_s$  requires an estimation of the other real components of the model. In other words, the  $Q_{op}$  or "loaded Q" of the resonator within the oscillator loop must be determined. The measurement of  $d\phi/df$  in an oscillator under closed loop conditions is the direct way of determining the  $Q_{op}$ .

The measurement of  $d\phi/df$  is obtained by recovering the phase change across the crystal while the  $X_L$  of the oscillator is changed slightly. It is also important to record the absolute operational phase,  $\phi$ . For an example, we will use a 10.0 MHz series resonate type oscillator, employing an SC-cut crystal resonator with the typical electrical parameters discussed previously, and operating at  $\phi = 2^\circ$ . If the measurement of  $d\phi/df$  is:

$$\Delta\phi = 33.3^\circ \quad \Delta f = 5.6 \text{ Hz}$$

$$\frac{\Delta\phi}{\Delta f} = 103.8 * 10^{-3} \text{ rad/Hz}$$

then the "loaded Q" or  $Q_{op}$  is calculated as:<sup>13</sup>

$$Q_{op} = \frac{f}{2} * \left| \frac{\Delta\phi}{\Delta f} \right| * [1 + \tan^2(\phi - \pi/2)]$$

$$Q_{op} = 5.19001 * 10^5$$

The expression for the sum of all real losses around the closed loop of the model, including  $R_e$  can be derived from the general expression for  $d\phi/dX$ :

$$\frac{\delta\phi}{\delta X} = \frac{R}{R^2 + X^2}$$

Since the sum of the reactance around the loop is zero in steady state, then  $d\phi/dX = 1/R_T$  for any steady state condition and  $R_T$  is the sum of all real losses, including  $R_e$ . Therefore, an expression for  $R_T$  can be made using the ratio of  $dX/df$  over  $d\phi/df$ :

$$R_T = \frac{\delta X}{\delta f}; \quad \frac{\delta\phi}{\delta f} = \frac{2 * Qop}{f}$$

$$R_T = \frac{f}{2 * Qop} * \frac{\delta X}{\delta f}; \quad R_T = R_e + R_{LLp}$$

and

$$\frac{\delta X}{\delta f} = \frac{2 * Qx * R_e}{f}; \quad Qx = \frac{1}{2 * \pi * f * C_m * R_s}$$

At an operating phase of approximately zero,  $R_e$  is very close to  $R_s$ . Using the value of  $Qop$  previously determined:

$$\frac{\delta X}{\delta f} = 17.44 \text{ Ohms/Hz}$$

$$R_T = 168.6 \text{ Ohms}$$

$$R_{LLp} = 108.6 \text{ Ohms}$$

With these values established, the sensitivity analysis for changes in  $R_s$  can be completed.

An increase in  $R_T$  will always have the effect of decreasing  $Qop$  at the same operating  $\phi$ . This seems confusing, at first, since  $R_e$  is a summed part of  $R_T$  and in general  $Qop$  actually increases a little as the crystal operates toward the positive  $X_m$  region. However, keep in mind that  $Qop$  is also a function of  $dX/df$  and it can be shown that  $dX/df$  varies at the same rate as  $R_e$ .<sup>14</sup> To calculate the sensitivity of  $R_s$  changes in the oscillator model, the effect on  $d\phi/df$  must be calculated.

From the equation for  $R_T$ ,  $d\phi/df$  can be expressed as:

$$\frac{\delta\phi}{\delta f} = \frac{1}{R_T} * \frac{\delta X}{\delta f}$$

and the derivative with respect to  $R_T$  is:

$$\frac{\delta\phi}{\delta f} * \frac{1}{\delta R_T} = \frac{-1}{R_T^2} * \frac{\delta X}{\delta f}$$

$$\frac{\delta\phi}{\delta f} * \frac{1}{\delta R_T} = \frac{-1}{(R_e + R_{LLp})^2} * \frac{\delta X}{\delta f}$$

which can be simplified, but will be left in this form for analysis purposes. The change of  $d\phi/df$  with incremental changes in  $R_T$  can be evaluated as approximately,  $-6 * 10^{-4}$  for the oscillator and crystal parameters introduced earlier in this discussion.

Since a unit change in  $R_s$  will be nearly equivalent to a unit change in  $R_e$  and  $R_T$  at most operating phases, then the change of  $d\phi/df$  with respect to a unit change in  $R_s$  is nearly equivalent to the value of  $-6 * 10^{-4}$ . Additionally, from the inspection of the equation for  $dX/df$  it can be seen that a unit change in  $R_s$  will result in nearly the same reactance slope. Therefore, the derivative of

$d\phi/df$  with respect to a unit change in  $R_s$  is practically constant.

Likewise, using the general expression for  $d\phi/dR$ :

$$\frac{\delta\phi}{\delta R e} = \frac{-X e}{R e^2 + X e^2}$$

$d\phi/dR e'$  can be evaluated as:

$$\frac{\delta\phi}{\delta R e'} = \frac{-X e}{(R e + 1)^2 + X e^2}$$

for unit changes in  $R_s$  over most of the operating region of the oscillator loop.

By inspection, the equation for  $d\phi/dR e$  will reach its maximum (or minimum) when  $R e$  equals  $|X e|$ , and the operating phase is  $+45^\circ$  ( $-45^\circ$ ). This point in the operating phase can be approximated by letting  $|X e| = R_s$  because both the phase slope of the crystal and  $dX/df$  are still essentially linear in this region.

Therefore, the maximum (minimum) sensitivity of crystal  $d\phi$  to changes in  $R_s$  is:

$$R e = R_s = 60 \text{ Ohms}; |X e| = R e \text{ at } \phi = 0.785 \text{ Rad}$$

$$\frac{\delta\phi}{\delta R e} \text{ max(min)} \approx \pm 8.33 \text{ mRad/Hz}$$

Additionally, the max(min) frequency change for a unit change in  $R_s$  can be estimated by adjusting the  $d\phi/df$  with the sensitivity factor for  $d\phi/df$  with  $R_s$  calculated previously, then multiplying the reciprocal of the iterated  $d\phi/df$  by the max(min) sensitivity of  $d\phi/dR e$ .

$$\frac{\Delta\phi}{\Delta f} = 103.8 \text{ mrad/Hz}; \frac{\delta\phi}{\delta f} = 103.2 \text{ mrad/Hz}$$

$$\frac{\delta f}{\delta\phi} * \frac{\delta\phi}{\delta R e} \text{ max(min)} \approx \pm 0.08 \text{ Hz/Ohm}$$

The results of the sensitivity analysis for changes in  $R_s$  indicate that for a one ohm change in  $R_s$ , the maximum change in  $\phi$  across the crystal is about  $\pm 0.5^\circ$  and the maximum frequency change is about  $\pm 8$  ppb. The polarity of the change with  $R_s$  is dependent on whether the operating point of the oscillator is above or below the zero phase condition of the crystal's equivalent circuit. If  $X e$  is positive, then the operating point of the oscillator is above the zero condition and the changes in  $\phi$  and the operating  $f$  will be negative.

The sensitivity analysis for phase changes across the crystal to  $R_s$  provides the confidence that large phase changes, greater than one degree, can be considered oscillator changes by the insitu technique. The confidence becomes even greater if the oscillator operating point is close to the zero phase condition of the crystal equivalent circuit. This is because  $d\phi/dR e$  approaches zero as  $|X e|$  approaches zero.

Additionally, it is very unlikely that a precision resonator's  $R_s$  will change by more than one ohm over any reasonable measuring period, except in the case of a leaking enclosure. However, because the sensitivity of most passive crystal measuring systems is within one or two ohms, crystals suspected of leaking can easily be rechecked to confirm an enclosure flaw.

## EXPERIMENTAL RESULTS

The insitu technique was used in several experiments to aid in the validation process of the theory just presented. The basic set-up of the technique requires the installation of a Hewlett-Packard 8508 vector voltmeter across the crystal nodes of an oscillator circuit. The oscillator circuit used in the

experiments was a 10 MHz, series resonant OCXO type which uses a 3<sup>rd</sup> overtone, "modified" SC cut as the crystal resonator. In this case, the modified SC has an inflection near 105°C

This circuit was chosen for the following reasons. First, the oscillator is identical to those used in the standard manufacturing line. Consequently, we were familiar with the typical aging patterns of the oscillator. Secondly, because the oscillator operates near series resonance, the ability to resolve the phase change across the resonator into oscillator and crystal components is maximized. Finally, because of the series circuit configuration, the stray impedance associated with probes is generally shunted out.

The procedure for using the insitu technique requires the measurement of the operating Q of the in-circuit crystal. As explained earlier, this is accomplished by temporarily installing a small change to the series load component. Three parameters are then recorded;

- 1.) The starting phase before installing the change component
- 2.) The phase change caused by the component installation
- 3.) The frequency change caused by the component installation. The polarity of the change must also be recorded.

Figure 3.) is a simple diagram of the data collection system. All measurements were made using a IBM PC with a IEEE-488 interface. The computer collects the data via the I/O bus of an Hewlett Packard 8508 vector voltmeter and a Hewlett Packard 5335 frequency counter in parallel. The temperature is monitored using a Keithley

195 digital multimeter with a platinum RTD attachment.

The oscillators were placed into a Delta Design temperature chamber with the ambient temperature maintained at 50 °C. This temperature was chosen to help keep the probes from being mechanically disturbed by cooling blasts of CO<sub>2</sub>.

Data was collected at an interval of 10 minutes beginning at the application of oscillator supply power. Because data is collected at the initial "power-on" state of the oscillator, all data is normalized to 24 hours.

All the data from the experiments are presented in this report using the same method. First, the actual frequency and phase data as collected from the oscillator is displayed (referred to herein as "Raw" data). Using the phase information and the  $d\phi/df$  slope measured during the test set-up, the oscillator component is calculated and displayed (herein called the oscillator component). Finally, the crystal resonator component is extracted by simply subtracted the oscillator component from the raw data.

The first experiment examines the aging characteristic of an unconditioned test oscillator circuit immediately after the application of supply power. It is meant by "unconditioned" that the oscillator did not receive any normally received elevated temperature bake-out. In contrast, the crystal resonator paired into the test oscillator had received an extra degree of thermal conditioning. The purpose of this pairing was to try to emphasize the oscillator aging component and perhaps reveal more time dependent characteristics into the data

Figure 4.) illustrates the "raw" data of the

aging experiment, note that the frequency data has been referenced to time = 24 hours. The data was collected at a rate of once every 10 seconds for the first 140 minutes of testing after which data was collected at ten minute intervals.

A typical OCXO warm-up and aging characteristics was observed in figure 4.). During the first 30 minutes, the unit approached thermal equilibrium. The phase data steadily increases indicating that the oscillator circuitry was changing due to temperature. At the same time, the frequency data shows a quick inversion and then steadily decreases in a similar manner to the phase increase. After about 1 hour into the data, the direction of the phase and frequency data reverse and continue in this manner in a decreasing logarithmic for the rest of the test.

Figure 5.) is the plot of the resolved oscillator and crystal aging of the experiment using the insitu technique. A clear separation of behaviors is noticed. During the warm-up period of the first 30 minutes, the oscillator frequency steadily decreases. It is known that the oscillator circuitry has a negative drift with temperature (as will be demonstrated in the later experiments). Meanwhile, the crystal frequency goes through an inversion in direction, first increasing and then decreasing as would be expected as the oven temperature proceeds past the lower turnover. The crystal frequency is also marked by a rapid overshoot consistent with modified SC warm-up behavior.

After warm-up, both the crystal and the oscillator remain stable for about an hour until a large initialization characteristic of the oscillator appears and continues for the rest of the test. In a comparative sense, the crystal frequency changes very little

indicating that the majority of the first 7 days of frequency aging can be attributed to the oscillator stabilizing. Further analysis is presented in following section on "curve fitting".

The next set of experiments use the insitu method to resolve the contributions of the oscillator and crystal in the frequency versus temperature behavior of an OCXO. Figure 6.) is the "raw" frequency versus temperature behavior of the same oscillator used in the aging experiment at two different slew rates. The slew rates were 1.0 °C/min and 0.05 °C/min respectively.

The oscillator was tested from +70 °C to +100 °C in 1°C increments for both slew rates. Hysteresis loops can be noticed in Figure 6.) of both slew runs. The accentuated divergence of the frequency above +72 °C can be explained as the point where the oven control shuts down and the oscillator then freely drifts with temperature.

Using the insitu method the contributions of the oscillator and crystal are resolved. Figure 7.) and Figure 8.) illustrate the contributions of the oscillator and crystal to the hysteresis loops, respectively. The crystal hysteresis loops of Fig. 8.) emulate the orbital behavior described by Ballato as the "dynamic" or time dependent temperature behavior of a crystal resonator.<sup>15</sup> The frequency loops of Fig. 8.) above +80 °C not only show a difference in the peak to peak excursion but also indicate the probability of higher order coefficients.

The previous experiment examined the frequency versus temperature behavior of the OCXO in its normal operating mode. The next experiment will repeat the analysis using the same oscillator with the oven disabled. Using a slew rate of 0.33

°C/minute, the oscillator is tested again from +70 °C to +100 °C in 1 °C increments.

Using the insitu method, the data is resolved in Figures 9 through 11. The crystal component is seen in Figure 11. The calculated turnover point for the crystal using the recovered crystal frequency is 85.9 °C.

The actual turnover as measured in a passive crystal test system is 87.3 °C. The negative temperature coefficient of the oscillator circuitry reduces the apparent turnover point of the crystal oscillator to 81.7 °C.

### Curve fitting the insitu aging data

As was anticipated, the insitu technique allowed the crystal and oscillator contributions to be separated. Consequently, the individual aging rates associated with each part may be analyzed with curve fitting.

Miljkovic reported on the effectiveness of fitting log, exponential and power functions to aging data.<sup>16</sup> The use of simple log and power functions to the raw data of the aging experiment gave curve fits with correlation coefficients ( $R^2$ ) ranging from 0.974 to 0.996.

Figure 12.) shows the curve fit of the raw frequency data from the aging experiment. The data of the aging experiment was curve fit to a function with a constant term, two exponential functions and a natural log function with a constant. The data used for the fit starts just after the temperature reaches equilibrium at about the beginning of the initialization period. The correlation coefficient for this fit ( $R^2$ ) is 0.9994. The aging shown in Fig. 12.) continues to proceed in a positive logarithmic manner until it levels out to a per day rate of  $+8.45 \times 10^{-10}$ .

After separation, the crystal and oscillator components to the aging characteristic were

fit similar to Fig. 12.). The two component curve fits are presented in figure 13.). The resolved crystal and oscillator components show individual initialization and long term aging rates. The oscillator contribution to the final per day aging is  $+1.816 \times 10^{-9}$  and the final crystal aging rate per day is  $-1.34 \times 10^{-10}$ . It is important to note that not only is there is difference in the magnitude of the aging but in the direction as well.

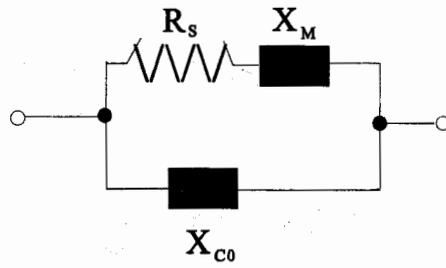
### CONCLUSION

The insitu technique has been validated with theory and demonstrated by experiment. This powerful technique may now be used to investigate both component and parametric behavior in quartz oscillators. Such issues as drive level dependency and short term frequency "pops" may be better understood through the technique. Aging models, particularly those involved with prediction, will benefit from the separation of crystal resonator from circuitry.

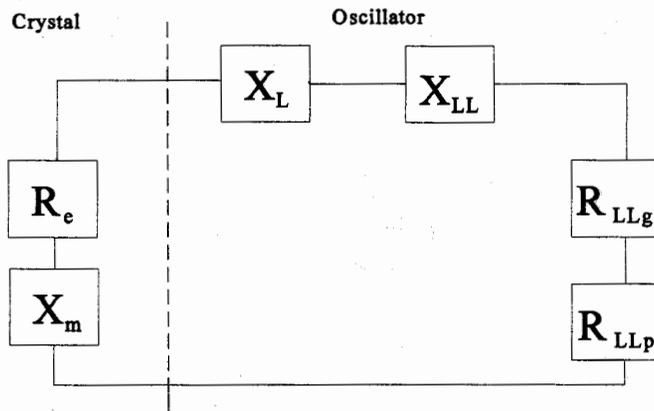
The authors deeply appreciate the support of Arthur Ballato, Bernd Neubig, Eric Hafner and Perkunas Kavolis.

**ENDNOTES**

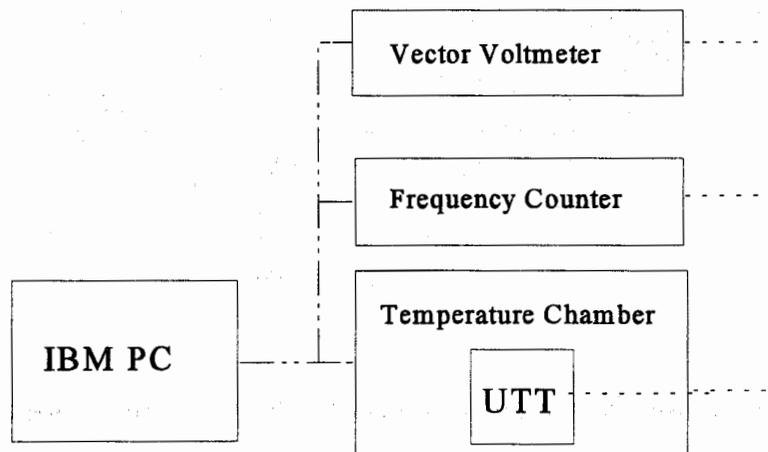
1. As discussed by V. Bottom in the 35th Frequency Control symposium, "A History of the Quartz Crystal Industry"; pg 10.
2. "The piezoelectric crystal unit-definitions and methods of measurement." Proc. IEEE 57, pp. 179-201
3. Adams, C., Kusters and Benjaminson; "Measurement Techniques for Quartz Crystals", Proceedings of the 22nd Freq. Control Symp., pp. 248-258
4. Hafner, E. and Riley, W.; "Implementation of Bridge Measurement Techniques for Quartz Crystal Parameters", Proceedings of the 30th Freq. Control Symp., pgs. 92-102
5. Parzen, B.; Design of Crystal and Other Harmonic Oscillators, New York: Wiley, 1983
6. Neubig, B.; "Comparison of Active and Passive Aging of SC and AT-cut Crystals", 10th European Frequency and Time Forum, to be presented and published
7. Parzen, B.; "Theoretical and Practical Effects of the Resonator Specifications and Characteristics upon Precision Crystal oscillator Design and Performance", Proceedings of the 36th Freq. Control Symp., pp. 480-485
8. Vig J, Meeker T. "The Aging Of Bulk Acoustic Wave Resonators, Filters and Oscillators," 45th Annual Frequency Control Symposium (1991) pp. 77-101.
9. Gerber E. and Ballato A.; Precision Frequency Control: Volume 2, Orlando: Academic Press, 1985 pg.17, contributed by Hafner.
10. [Ref 9] pg.8, contributed by Hafner.
11. [Ref 7] figure 4
12. Parzen B.; "Universal, Computer Facilitated, Steady State Oscillator Analysis Theory", 44th Annual Symposium on Frequency Control, pp 536-541.
13. [Ref 5] pg. 207
14. [Ref 7] pg. 483
15. Ballato A. and Vig J.; "Static and Dynamic Frequency-temperature Behavior of Singly and Doubly Rotated, Oven-controlled Quartz Resonators", 32nd Annual Symposium on Frequency Control (1978), pp 180-188.
16. Miljkovic M., Lj G, Brajovic T, Brajovic V, "AGING PREDICTION OF QUARTZ CRYSTAL UNITS," 42nd Annual Frequency Control Symposium (1988), pp 404-411



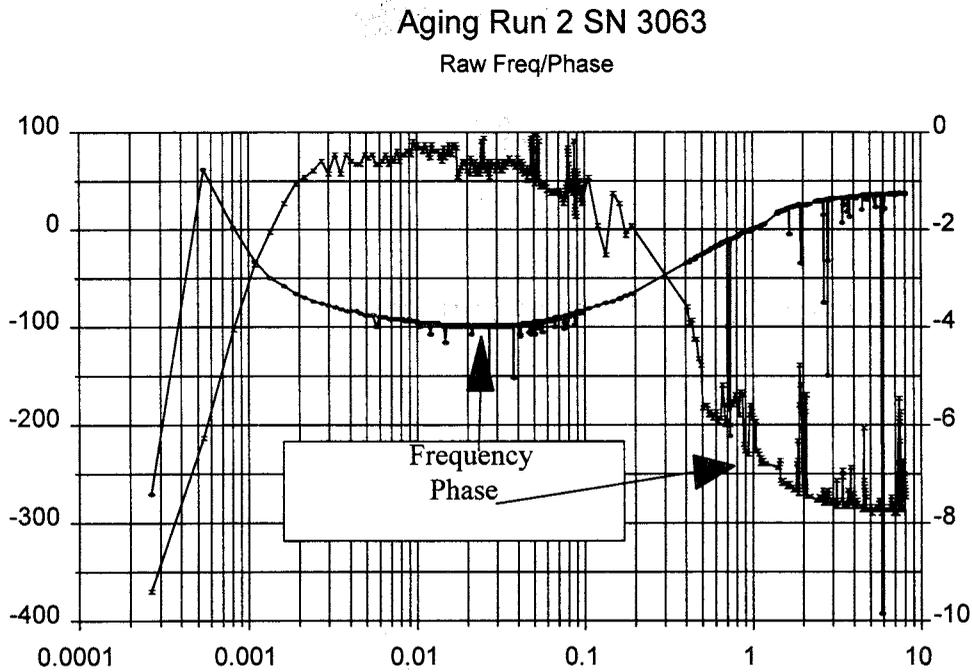
**Figure 1.) Crystal Electrical Equivalent**



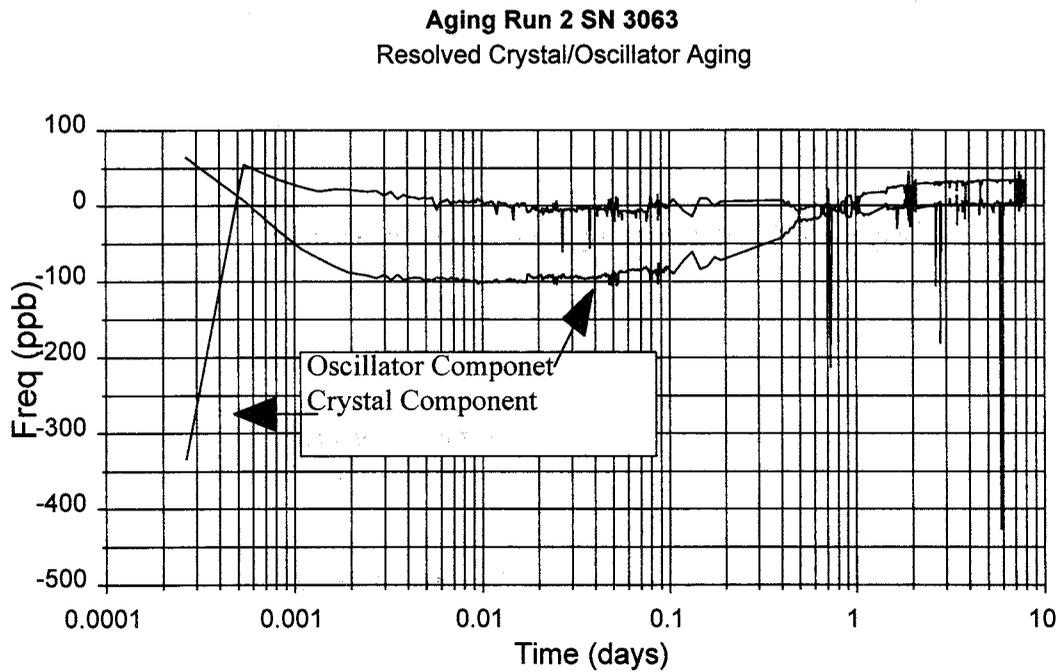
**Figure 2.) Oscillator Negative Resistance Model**



**Figure 3.) Experiment Set-up**

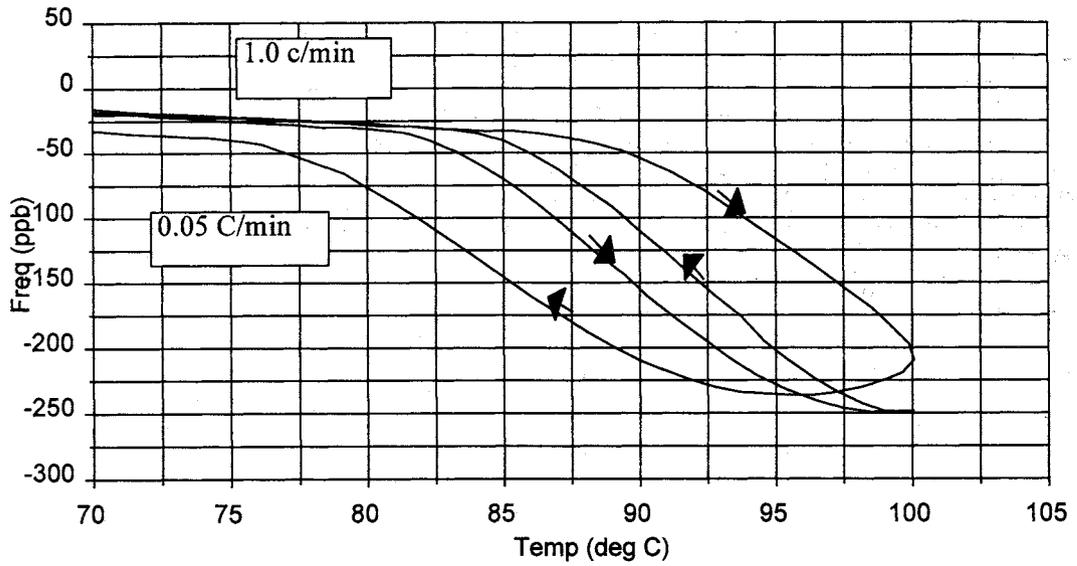


**Figure 4.) Raw Data from Aging Experiment**



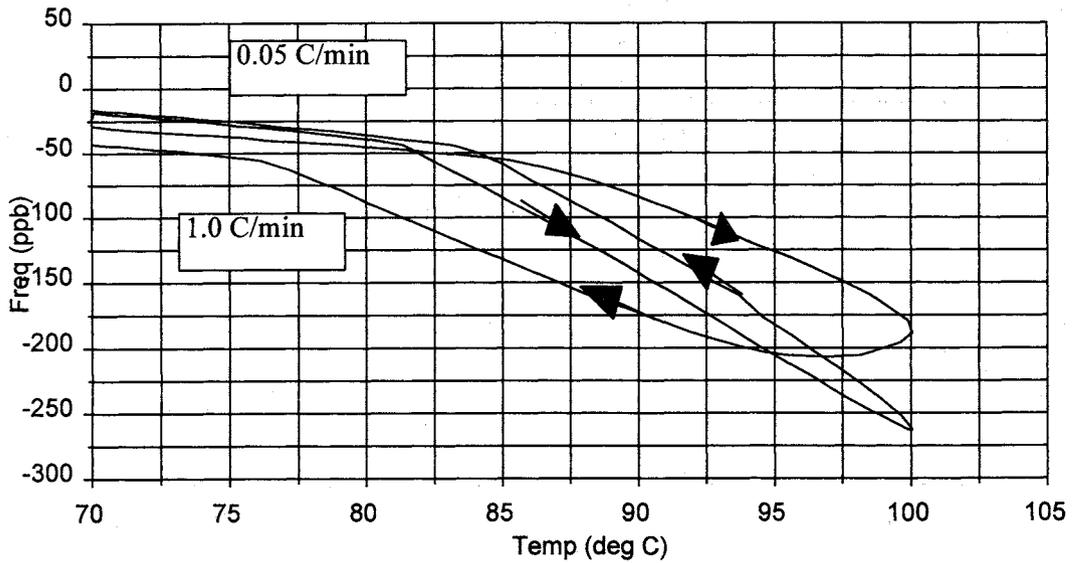
**Figure 5.) Crystal and Oscillator Frequency Resolved by Insitu Technique**

**10.000 MHz 3rd SC**  
Hysteresis Curves Raw Frequency



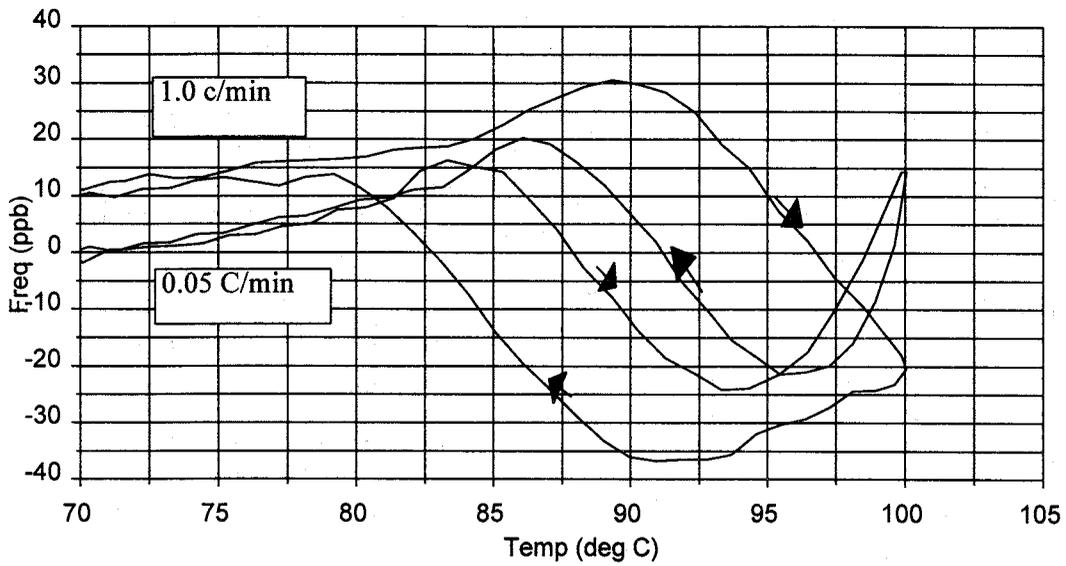
**Figure 6.) Raw Data from Hysteresis Experiment**

**10.000 MHz 3rd SC**  
Hysteresis Curves Oscillator Component



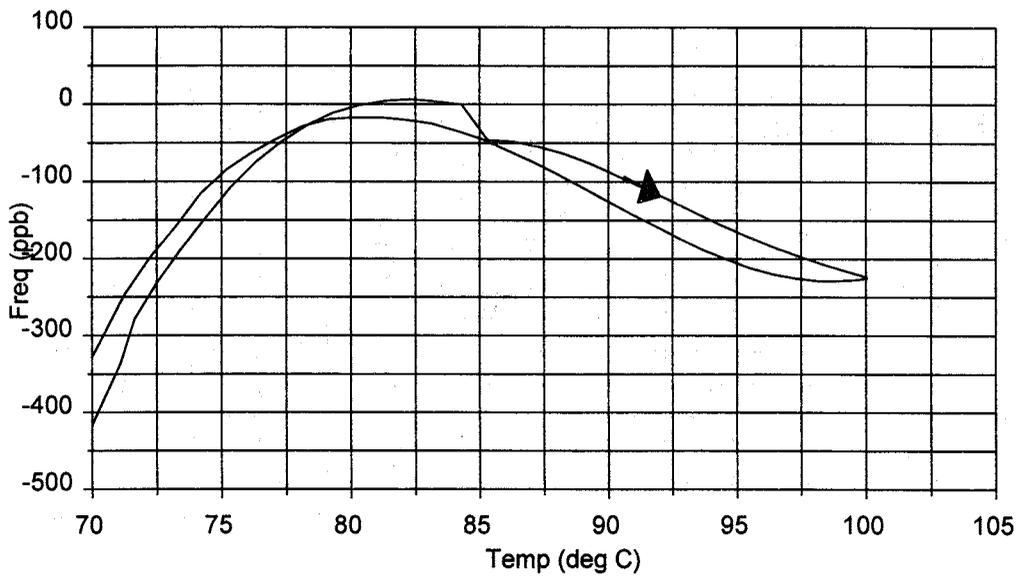
**Figure 7.) Oscillator Hysteresis Resolved by Insitu Technique**

**10.000 MHz 3rd SC**  
**Hysteresis Curves Crystal Component**



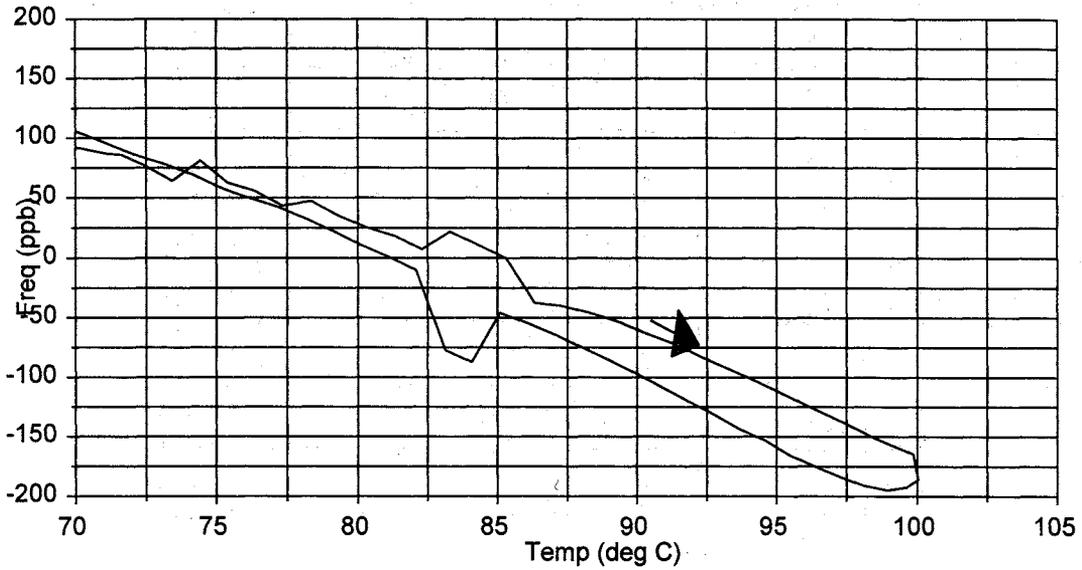
**Figure 8.) Crystal Hysteresis Resolved by Insitu Technique**

**10.000 MHz 3rd SC**  
**Disabled Oven -- Raw Frequency**



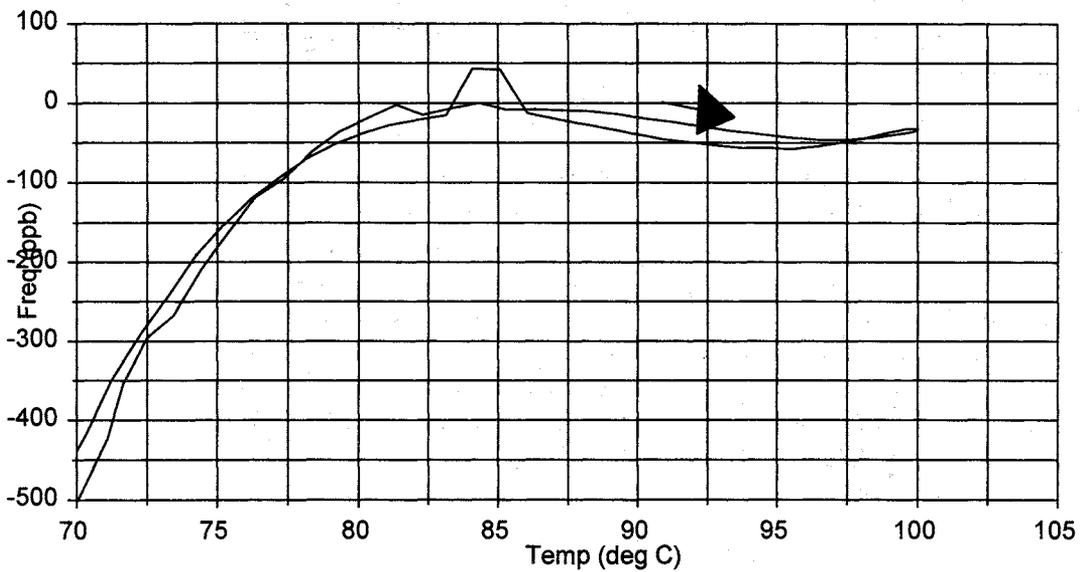
**Figure 9.) Raw Data from Disabled Oven Experiment**

**10.000 MHz 3rd SC**  
 Disabled Oven -- Oscillator Component



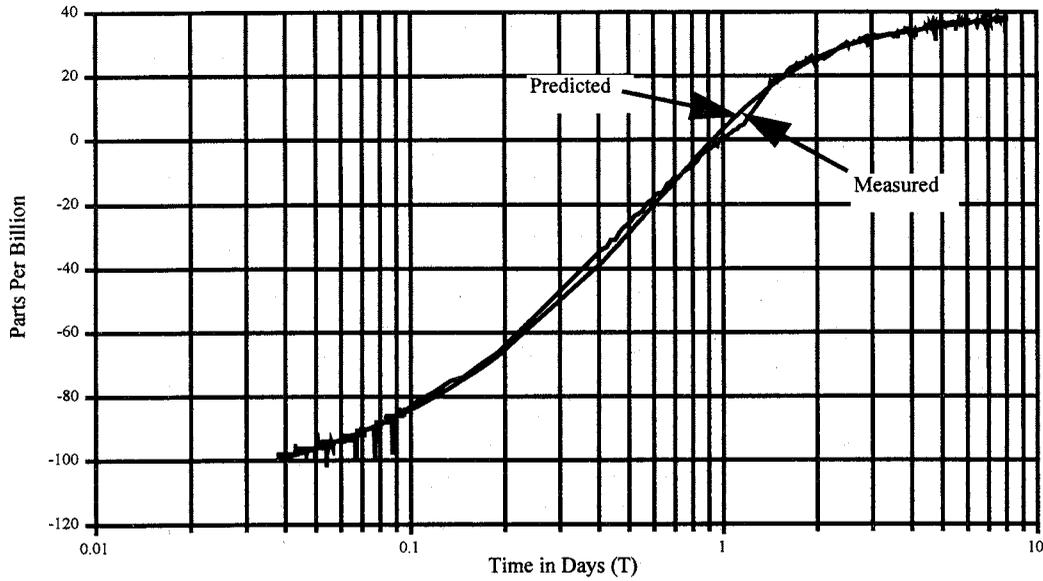
**Figure 10.) Oscillator Temperature Coefficient Resolved by Insitu Technique**

**10.000 MHz 3rd SC**  
 Disabled Oven -- Crystal Component



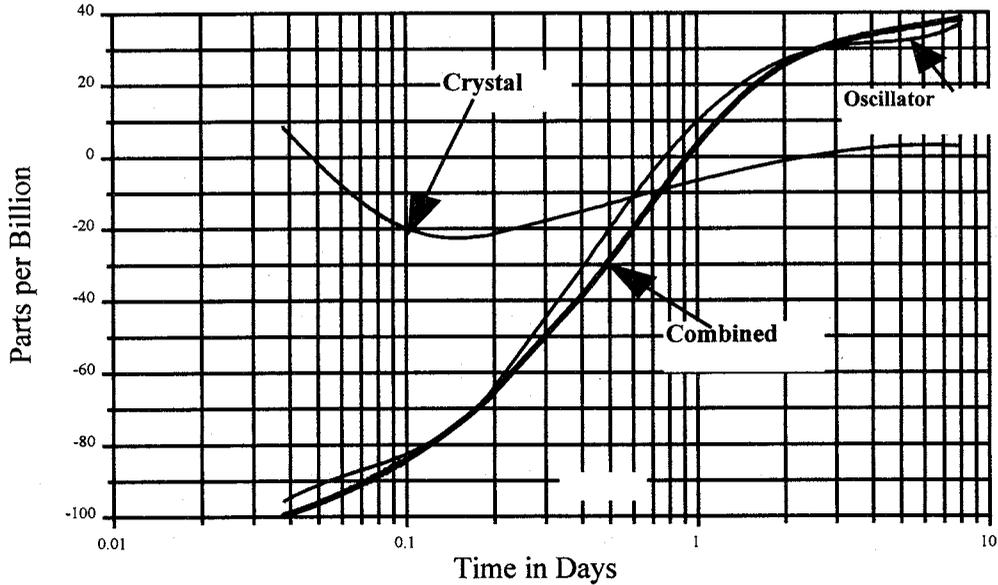
**Figure 11.) Crystal Frequency-Temperature Curve Resolved by Insitu Experiment**

$$y=A + B*EXP(-CT)+D*EXP(-ET)+F*LN(GT)$$

 $r^2=0.99940201$ 


**Figure 12.) Curvefit of Frequency from Raw Data of Aging Experiment**

Raw data reduced to crystal and  
oscillator components



**Figure 13.) Curvefit of Resolved Crystal and Oscillator Data from Aging Experiment**

## PRECISION, LOW POWER, ANALOGUE TCXO USING A SINGLE INTEGRATED CIRCUIT

Hedley Rokos

C-MAC Quartz Crystals Ltd Harlow Essex CM20 2DE England

### SUMMARY

We have designed a single chip ICXO which is capable of temperature compensation to better than  $\pm 0.1$  ppm over  $-50$  to  $+90$  °C. Off-chip components are limited to a decoupling capacitor, the crystal, and a filter capacitor for low-noise applications.

### CIRCUITRY

#### Oscillator Circuitry

The oscillator is a simple Colpitts design, using variable capacitance diodes for frequency pulling. A CMOS buffer capable of output exceeding 50 MHz is provided, as well as a linear buffer with output to at least 15 MHz.

#### Compensation Circuitry

The compensation circuitry generates Chebyshev functions up to 6th order in temperature. The temperature sensor uses transistors run at different current densities ("PTAT" cell) to provide a positive temperature coefficient signal, and a diode potential for the negative coefficient. Orders are generated using buffered Gilbert-cell multipliers, and current subtraction. The amplitude of each order presented to the compensation output is set using multiplying digital to analogue converters.

It has been observed that the sensitivity of different structures on an integrated circuit to temperature cycling induced stress can differ by orders of magnitude. Accordingly, considerable care was taken in selecting appropriate component types and in the layout of the die. We believe this to be a key element in the achievement of the observed performance. To avoid a shift in characteristics following sealing of the outer package, we have arranged for measurements and compensation setting to be controlled digitally through the TCXO pins. We have retained the

ability to use our standard packages by communicating via pins which normally serve other functions.

### MEASUREMENT AND PERFORMANCE

Devices have been built using crystals of standard TCXO design, giving  $\pm 0.1$  ppm performance over the temperature range  $-50$  to  $+90$  °C. The error contributions from band-breaks, crystal retrace, and measurement artefacts were of similar magnitude; semiconductor retrace effects are somewhat smaller. Band-breaks and retrace may be improved by using larger crystals with improved mountings and plating, and we are modifying the measurement program to reduce the effects of the artefacts; degradation caused by semiconductor retrace can be further reduced by suitably selecting the crystal angle. We anticipate that this will compensate to better than  $\pm 50$  ppb.

We present here results from an oscillator, using an AT cut crystal in a glass holder, compensated over  $-50$  to  $+100$  °C, with retrace characteristics measured over  $-50$  to  $+120$  °C. Figure 1 shows the uncompensated oscillator characteristic, the compensation voltage needed to set the oscillator onto nominal frequency, and the predicted compensation error of  $\pm 75$  ppb.

In figure 2, we see the measured results on the compensated oscillator, with an excursion of  $\pm 90$  ppb. The major cause of the differences is almost certainly measurement artefacts, although there is doubtless some contribution from relaxation of the crystal mountings etc.

Figure 3 shows the effect of temperature cycling the oscillator. Measurements are made in  $10$  °C steps, with a 30 minute dwell time at each temperature. The oven ran from  $120$  to  $-50$  °C, followed after the 30 minute dwell time by a run from  $-50$  to  $100$  °C. The unit had previously been exposed to a minimum temperature of  $-50$  °C. The measured frequency excursion here increases to  $\pm 150$  ppb. The dwell times were chosen to allow temperatures to settle within at most 30 m°C. The

main source of the additional error is almost certainly crystal hysteresis.

Figures 4, and 5 show the oscillator tuning characteristic over the compensation range, and the output potentials corresponding to the various orders, respectively. Pulling deviation is about four percent from the best-fit tangent, and we can see that band-gap curvature causes curvature in the output of the temperature sensor. The  $\pm 0.5$  ppm third order fit of figure 6 demonstrates the result of these effects when combined with the temperature dependence of crystal C1, and the imperfectly cubic nature of the AT characteristic.

## CONCLUSION

We have demonstrated an analogue, precision temperature compensation system, based on a single integrated circuit. When used with standard low-aging crystals, we expect routinely to achieve  $\pm 0.25$  ppm over wide temperature ranges. With improved AT crystals, better than  $\pm 50$  ppb should be possible.

## ACKNOWLEDGEMENTS

The staff of the VLSI design group at Nortel, Harlow for skilled simulations and layout. Staff of C-MAC Harlow, particularly D E Bower and K Ward for software and test support.

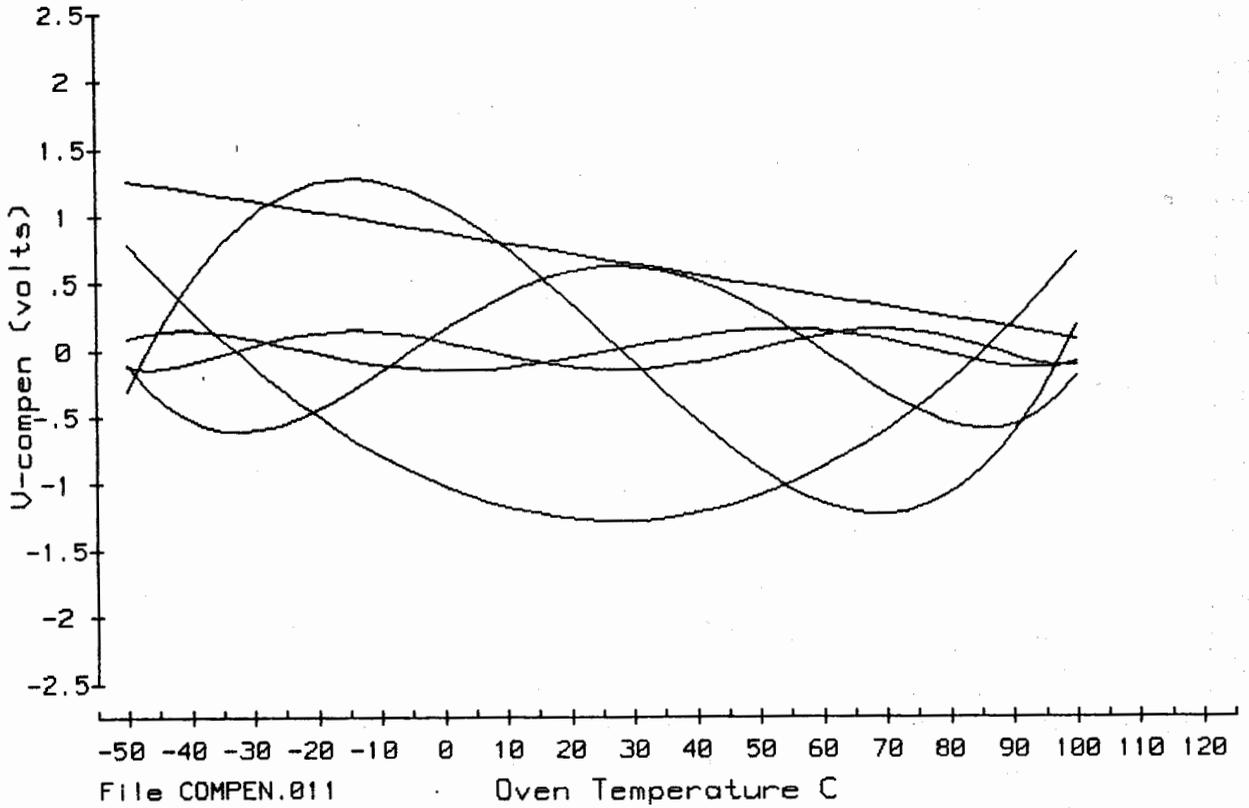


Figure 5: Full-scale order outputs

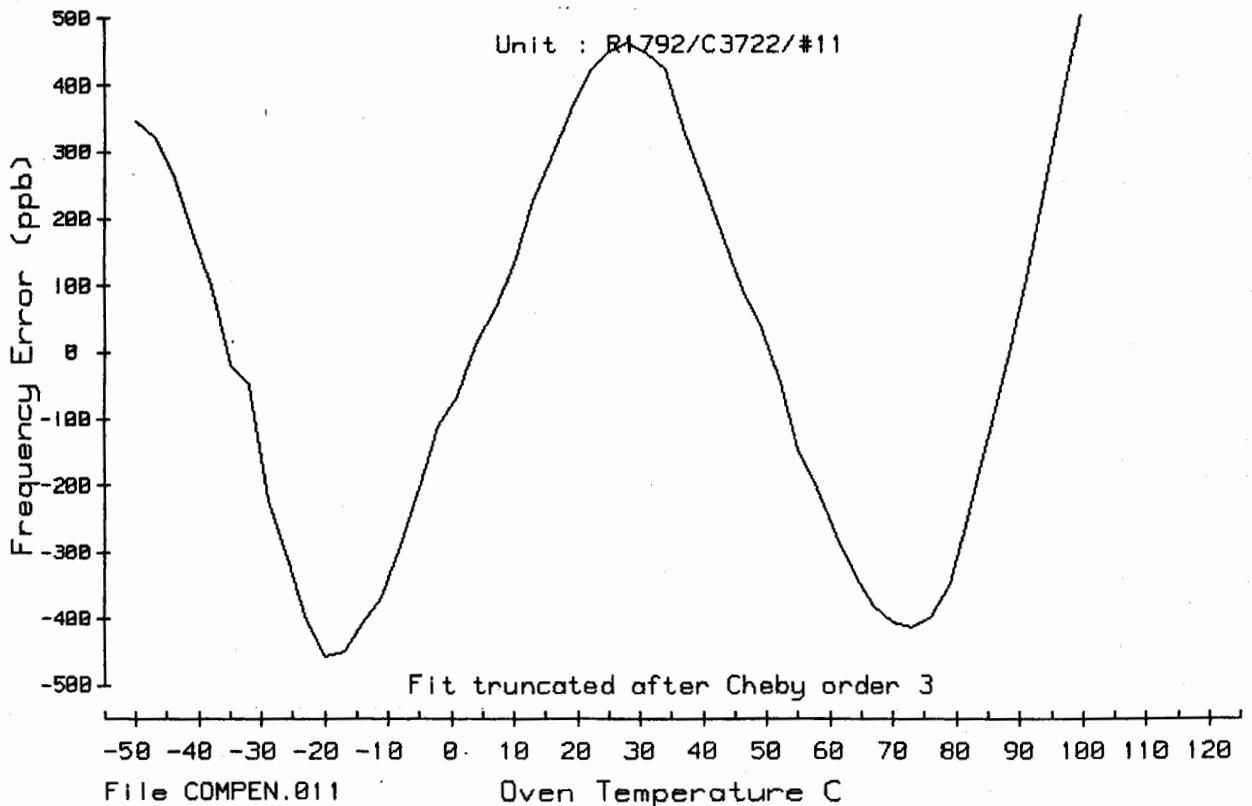


Figure 6: Compensation results truncated at fourth order

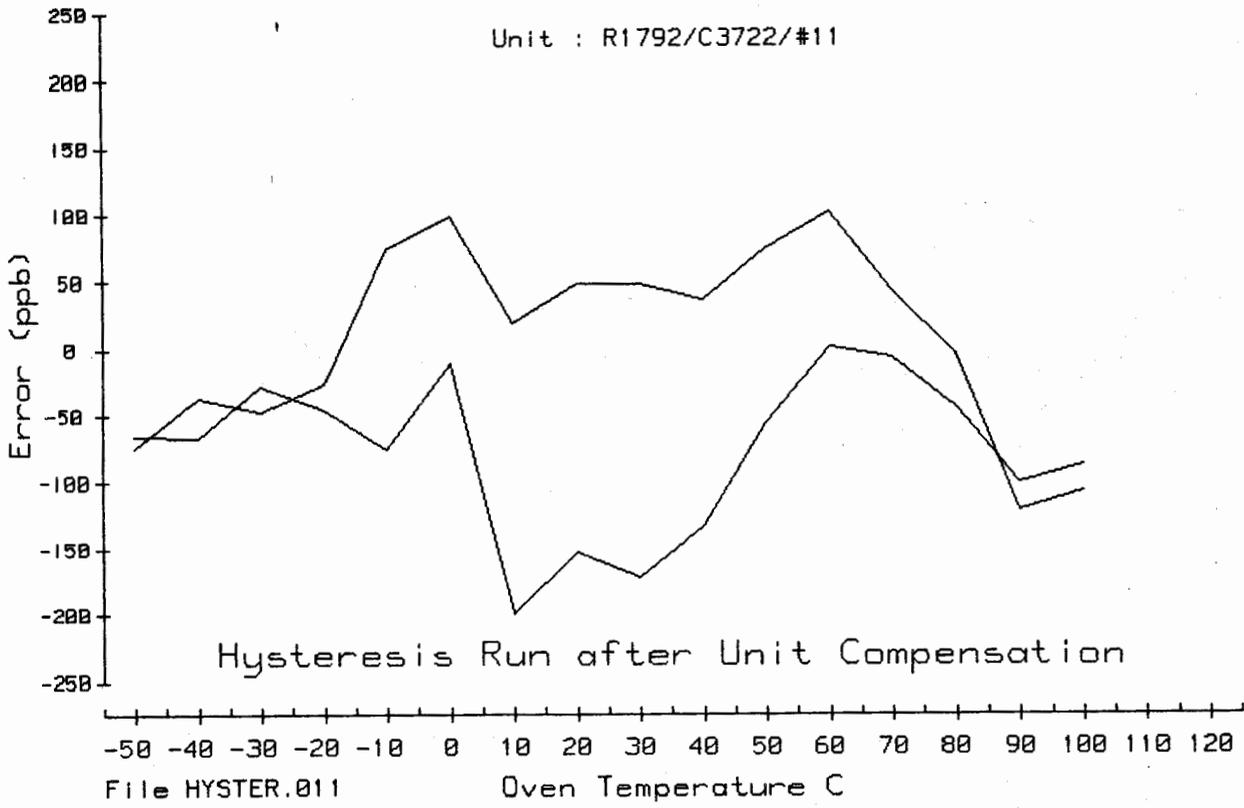


Figure 3: Temperature cycling performance (re-optimised)

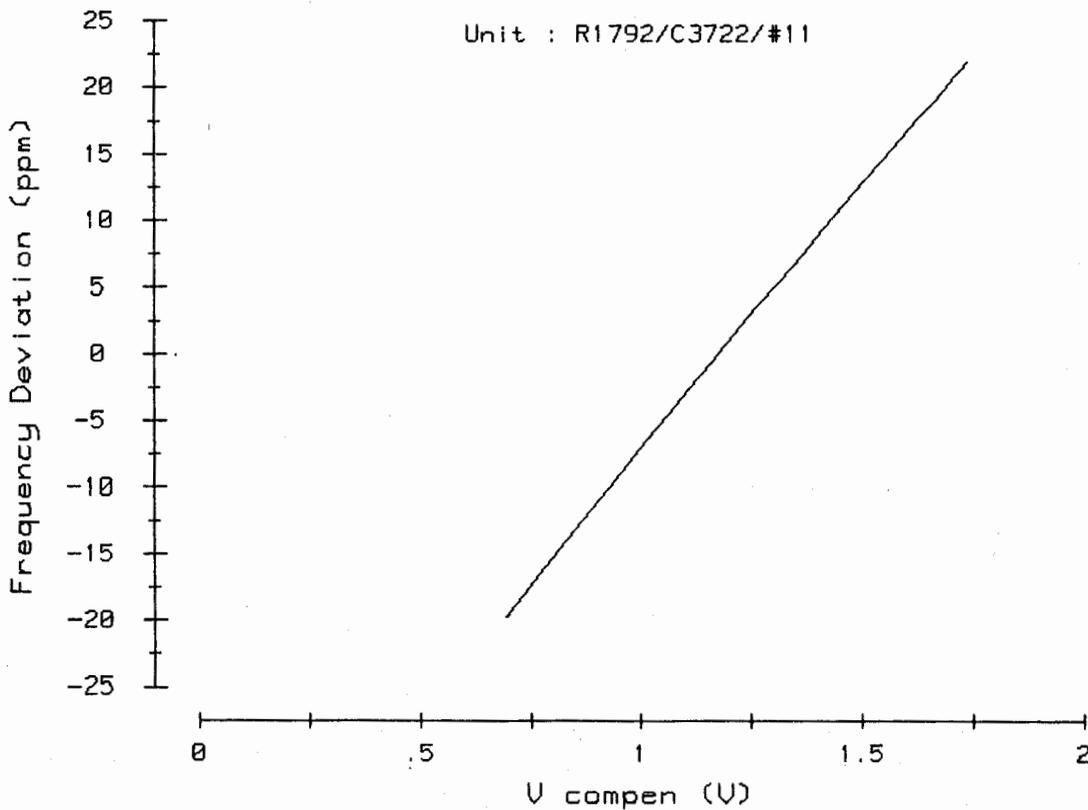


Figure 4: Tuning characteristic over compensation range

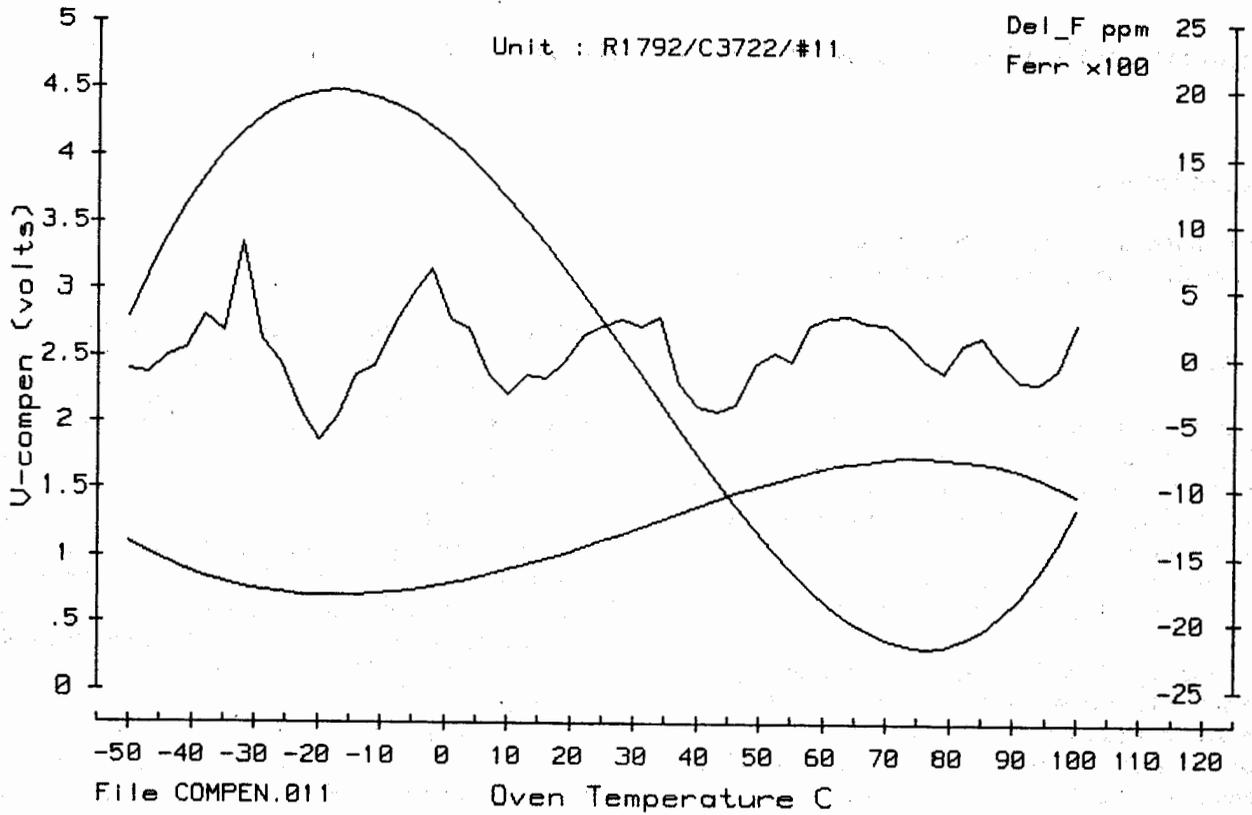


Figure 1: Uncompensated oscillator characteristic, compensation requirement, and predicted compensated performance

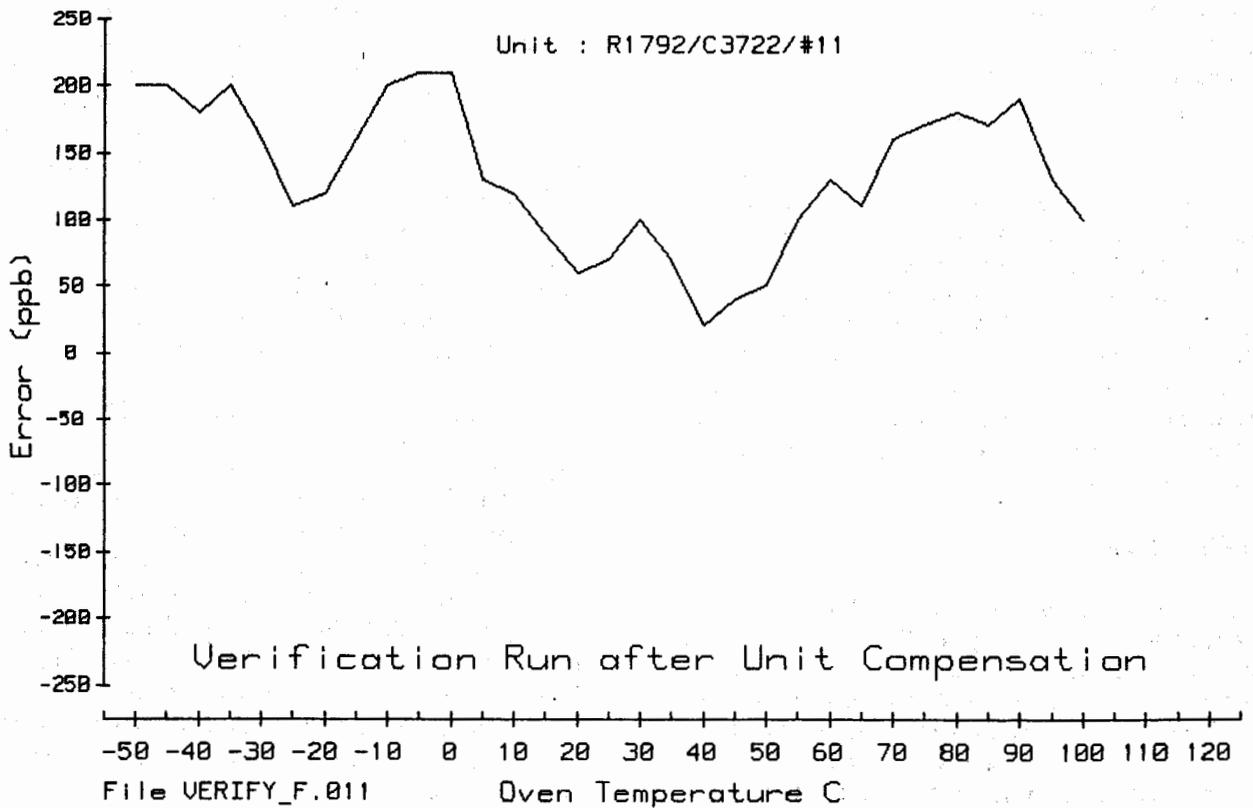


Figure 2: Compensated oscillator performance

## STABILITY AND ACCURACY OF INTERNATIONAL ATOMIC TIME TAI

Claudine Thomas

Bureau International des Poids et Mesures  
Pavillon de Breteuil, F-92312 Sèvres Cedex, France

### ABSTRACT

The BIPM Time Section is in charge of the generation and dissemination of the reference time scale TAI (International Atomic Time) obtained in deferred-time by combining the readings of a number of individual atomic clocks spread world-wide. The algorithms designed for this purpose are optimized for reliability, long-term stability and accuracy of the time scale.

Since the end of 1992, the quality of the timing data received at the BIPM has rapidly evolved thanks to the extensive replacement of older designs of commercial caesium clocks. Consequently, the stability of the reference time scales has improved significantly: it is characterized by an Allan deviation of  $2,6 \times 10^{-15}$  for averaging times  $\tau=40$  d. However, further improvement in stability is still possible without threatening reliability. This was tested by running modified algorithms over the real clock data collected at the BIPM. Results of different studies are shown here, in particular the implementation of an upper relative contribution, chosen equal to 1,37% for any contributing clock, leads to  $\sigma_y(\tau=40 \text{ d}) = 1,8 \times 10^{-15}$ .

The accuracy of TAI is estimated by the difference between the duration of the TAI scale interval and the SI second as produced on the rotating geoid by primary frequency standards. In this paper, TAI accuracy is evaluated from six primary frequency standards LPTF-FO1, PTB CS1, PTB CS2, PTB CS3, NIST-7 and SU MCsR 102 all corrected in a consistent manner for the gravitational shift and the black-body radiation shift. This led to a mean departure of the TAI scale interval of  $1,8 \times 10^{-14}$  s over 1995, known with a relative uncertainty of  $0,5 \times 10^{-14}$  (1  $\sigma$ ).

### INTRODUCTION

The Bureau International des Poids et Mesures, BIPM, is responsible for the generation of world-wide reference time scales, among them International Atomic Time, TAI, and Coordinated Universal Time, UTC. The TAI relies basically on measurements taken from commercial atomic clocks and primary frequency standards maintained in national timing centres. Since 1977, the procedure used for combining these data has been carried out in two steps:

\* The first step of TAI is the computation of the free atomic time scale, EAL (échelle atomique libre),

obtained as a weighted average of a large number of free running and independent atomic clocks spread world-wide. The corresponding algorithm is optimized for long-term stability and post-processes measurements taken over a basic sample period of 60 days [1, 2, 3].

\* In a second step, TAI is derived from EAL by frequency steering with the aim of maintaining the accuracy of its scale unit. The steering corrections, determined by comparing the EAL frequency with primary frequency standards, are of the same order of magnitude as the EAL instability [4, 5].

The aim of this paper is to characterize the stability and the accuracy of the time scales computed by the BIPM. For what concerns stability, one computes Allan deviations  $\sigma_y(\tau)$  characteristic of the free running time scale EAL using the N-cornered hat technique (see Section 2). For what concerns accuracy, one estimates the departure, and its relative uncertainty, of the duration of the TAI scale interval from the SI second as produced on the rotating geoid by primary frequency standards (see Section 3). An important point is that access to TAI is provided by a time series constituted of time differences  $[TAI(t) - T(t)]$  for each standard date  $t$ , MJD ending in 9, between the readings of TAI and any clock or time scale  $T$ . This time series is not stationary and the qualities of stability and accuracy vary with time, especially when the qualities of contributing clocks improve as has been the case since the beginning of 1993 (see Section 1). For this reason, in what follows, we indicate the period of time on which evaluations of stability and accuracy are made.

### 1. DATA USED FOR TAI COMPUTATION

Since the beginning of 1993, the quality of the timing data used for TAI computation has rapidly evolved. Four major changes have been identified:

- \* older designs of commercial clocks have been extensively replaced by the new HP 5071A clocks which present outstanding stability, characterized by a flicker floor level of about  $6 \times 10^{-15}$ , in terms of  $\sigma_y(\tau)$ , for averaging times  $\tau$  between 20 d and 40 d,
- \* the entry into operation, in timing laboratories, of active auto-tuned hydrogen-masers showing drifts of frequency of order  $10^{-17}/\text{d}$ ,
- \* the widespread use of GPS time transfer, which makes it possible to compare the clocks of nearly all laboratories contributing to TAI with an accuracy of a few nanoseconds,

\* evaluations of new primary frequency standards showing type B uncertainties of some parts in  $10^{14}$  ( $1\sigma$ ), and even better (3 parts in  $10^{15}$ ) for the caesium fountain developed at the LPTF (LPTF-FO1).

The first two points are illustrated in Fig. 1, which represents the number of clocks weighted in the EAL computation over the period 1992-1995. All the clocks considered here have operated continuously over at least six consecutive months and have received a weight characteristic of their frequency stability relative to EAL. Among them, some show frequency steps and are intentionally weighted to zero, and some are very stable and are given a fixed weight equal to the upper limit allowed. This particular weight does not correspond to their full stability properties, but is applied to improve the reliability of the time scale [3, 6]. In Fig. 1, three types of clocks are distinguished: HP 5071A clocks, hydrogen masers and other commercial caesium clocks. For each type the number of units reaching the upper limit of weight is also indicated.

Over the four years chosen for Fig. 1 the total number of clocks weighted in the EAL remained nearly constant (about 190), but a progressive replacement of commercial caesium clocks occurred in favour of the new HP 5071 model: 42% of the clocks were of this type at the end of 1995. In addition, nearly all of these clocks reach the upper limit of weight as soon as they enter the ensemble, even after May 1995, the date chosen by the CCDS Working Group on TAI for increasing the upper limit of weight by a factor 2,5 [7]. In contrast, only a few (25% to 30%) of the caesium clocks which are not of the HP 5071A design reach the upper weight. Since May 1995, this number has further decreased, the condition for being weighted at maximum being more stringent [3]. Nearly 50% of the hydrogen masers contributing to EAL are also assigned the upper limit of weight: the weighting procedure was designed so that the detection of a frequency drift automatically deweights the involved clock, keeping at high weight those hydrogen masers which are active and auto-tuned. To conclude, the global stability of the clock ensemble has considerably improved since 1992 with the consequence that the stability of EAL has also improved in both the middle and the long term, a point illustrated in Section 2.

An other important change was the widespread use of GPS common-view time transfer among the national timing centres contributing to TAI. This makes it possible to smooth out measurement noise in time comparisons between distant clocks by averaging over a few days. The averaging period remains shorter than the 10 day interval, and even the 5 day interval used since 1st January 1996 [7], between two TAI updates. It follows that the resulting time scale is no longer affected by white phase noise which results in an improvement in its short-term stability (5 days to 10 days). A possible consequence could be the shortening

of the TAI computation time from 60 days to 30 days or 40 days. However, so far, the results of tests on real data averaged over 30 days have not been completely satisfactory (see Section 2).

Data from primary frequency standards are also used for TAI computation. Over the years 1992-1995 the BIPM Time Section has received data from two types of primary frequency standards:

\* Some operate continuously and thus behave like clocks. Their data are used in the regular EAL computation, and taken into account according to their stability. This is the case for the NRC and PTB standards as shown in Table 1. These clocks are weighted according to the procedure used for commercial clocks and thus cannot individually be assigned a weight greater than the upper limit. The data from the most accurate of these standards, PTB CS1, PTB CS2 and PTB CS3, can also be used for the estimation of the duration of the TAI scale unit (see Section 3) if appropriate corrections, such as those compensating for the gravitational shift and the black-body radiation shift, are consistently applied.

\* Other primary frequency standards operate in a discontinuous mode and are evaluated from time to time. Each evaluation provides a measurement of the TAI frequency with respect to the standard frequency corrected for the gravitational shift and the black-body radiation shift, together with a type B uncertainty characteristic of the standard. These data cannot be used in the EAL computation and the intrinsic stability of the standards brings nothing to the EAL. However, these data provide independent evaluations of the duration of the TAI scale interval and make it possible to estimate the accuracy of TAI (see Section 3) and to decide whether or not to apply a new steering correction. The most accurate of these standards are the optically pumped caesium standard NIST-7, with a type B uncertainty of  $1 \times 10^{-14}$ , and the caesium fountain LPTF-FO1, with a type B uncertainty of  $3 \times 10^{-15}$  [8] (see Table 1).

Table 1 shows that, over the period 1992-1995, the accuracy of TAI can be estimated from measurements provided by six primary frequency standards which are more accurate than others. This is detailed in Section 3. However, it should be noted that the 14 measurements from SU MCsR 102 did not reach the BIPM until December 1995, the first measurement from NIST-7 was received in September 1994, and the totality of measurements from LPTF-FO1 in February 1996. It follows that for several years, the conformity of the TAI scale interval with its definition (the SI second on the rotating geoid) rested almost entirely on the data of a single laboratory, the PTB. In addition, at that time, no corrections for the black-body radiation shift were applied. The situation has now radically changed: six different and accurate standards provide data uniformly corrected for all known frequency shifts: this makes it

possible to determine the length of the TAI scale interval with an improved accuracy (see Section 3).

## 2. STABILITY OF EAL

The stability of TAI is directly linked to the stability of EAL except that, in the very long term, the frequency steering corrections which compensate the drift of EAL with respect to the most accurate frequency standards cause a separation. We thus focus, in this section, on the stability of the free running time scale EAL which is more readily determined.

Since the end of 1992, the stability of the free atomic time scale EAL has improved naturally without significant changes in the algorithm. Values for the stability are estimated by application of the N-cornered hat technique to data obtained from April 1993 to December 1995 in comparisons between EAL and the best independent time scales of the world (maintained at the NIST, the VNIIFTRI, the USNO, the PTB and the LPTF). These lead to the values for the Allan deviation  $\sigma_y(\tau)$  given in the first line of Table 2, especially  $\sigma_y(\tau=40 \text{ d}) = 2,6 \times 10^{-15}$ .

For further improvement, the stability algorithm which produces EAL may need to be revised and several possible changes have been the subject of experiments on real clock data collected at the BIPM. These possibilities mainly concern changes in the upper limit of weights, the use of hydrogen-masers and shortening of the computation time of TAI.

The HP 5071A clocks kept in national laboratories show outstanding long-term stability and generally contribute the maximum weight as soon as their data is used in the TAI computation. Consequently, to take full advantage of the most stable of these clocks calls for an increase in the upper limit of weight. Tests have been conducted with the limits 2500, 5000 and 10000 and have shown an improvement of the stability (lines 2, 3 and 4 of Table 2) [9]. However, the time scales E5000 and E10000 were judged not reliable enough [6] and, following a decision of the CCDS Working Group on TAI, the maximum allowable weight of a clock in EAL was changed from 1000 to 2500 starting from the computation over the two-month interval May - June 1995 [7].

Studies are in hand to assess the advantages of using an upper limit of relative weights, rather than one of absolute weights. Tests show that an upper contribution of 1,37 % for any individual clock would have helped to improve the stability of EAL during the period 1993-1995 (line 5 of Table 2) [3, 6]. This criterion is severely discriminatory even among HP 5071A clocks and continuously operating primary standards, some of these not being stable enough to reach the upper limit.

It has been shown that the introduction of hydrogen maser data in the EAL computation did not degrade its stability for the period 1988 - 1994, though frequency drifts were not taken into account [10]. For averaging times close to the EAL computation time (60 days), the variation of the maser frequencies relative to EAL was dominated by an important drift in only one maser, which consequently received a small weight. However, EAL stability is improving and the frequency drift of some hydrogen masers may become significant when compared with the intrinsic EAL noise. If this proves to be the case, it will be necessary to use a specific weighting procedure and mode of frequency prediction for hydrogen masers, based on estimates of their frequency drifts. Long periods of observation, at least one year, will then be needed before entering hydrogen maser data in the EAL computation.

Reduction of the noise involved in time transfers and the stability characteristics of HP 5071A clocks may make it possible to shorten the computation time of EAL. Tests show that data averaging over 30 day periods rather than 60 day periods improves the middle-term stability of EAL, but does not improve its long-term stability even if it is associated with an increase of the upper limit of weight (line 6 of Table 2) [11, 12]. Nevertheless, the idea of reducing the computation time has not been abandoned: this would reduce the delay of access to UTC and improve its predictability, two features which are important for national timing centres which keep a local representation of UTC [13].

## 3. ACCURACY OF TAI

Primary frequency standards deliver an ensemble of accurate measurements of the TAI frequency, uniformly corrected for all known corrections, especially the gravitational frequency shift (amplitude of about 1 part in  $10^{13}$  for an altitude of 1000 m above the geoid) and the black-body radiation frequency shift (amplitude of about 1,7 parts in  $10^{14}$  for a temperature of 300 K).

Each measurement is carried out over a given time interval centred on a given date. For the primary standards listed in Table 1, these periods of calibration correspond to the usual 60 d intervals defined for TAI computation (January-February, March-April, ...etc), except for LPTF-FO1, LPTF-JPO and NIST-7 the measurements of which are carried on over periods from 10 hours to 3 days, and transferred to TAI over 10 day periods (the shortest time interval between two consecutive TAI updates).

The individual TAI frequency measurements are characterized by an uncertainty which theoretically combines quadratically:

- \* the type B uncertainty of the standard (see Table 1),

- \* the type A uncertainty of the standard, linked to its stability over the period of measurement and always much smaller than the type B uncertainty,
- \* the uncertainty linked to the stability of TAI, and
- \* the uncertainty due to the transfer to TAI of the measure from a local time scale (local representation of UTC or hydrogen maser) kept in the laboratory where the standard is evaluated. This is related to the time link method and to the stability of the local time scale if the calibration interval is shorter than the estimation interval (cases of LPTF-FO1, LPTF-JPO and NIST-7).

In practice, the part coming from the type B uncertainty of the standard predominates for all standards except for LPTF-FO1 for which the parts coming from the stabilities of TAI [ $\sigma_y(\tau=10 \text{ d}) = 3,7 \times 10^{-15}$ ] and of the local hydrogen maser [ $\sigma_y(\tau=10 \text{ d}) \approx 1,5 \times 10^{-15}$ ] are no longer negligible.

Given this ensemble of measurements, the accuracy of TAI can be estimated over a given time interval centred on a given date and is characterized by two numbers: the difference  $d$ , expressed in seconds, between the duration of the TAI scale unit and the SI second as produced by the primary frequency standards on the rotating geoid, and its relative uncertainty  $\sigma$ . It is necessary to specify the length of the estimation interval and its central date when giving the quantities  $d$  and  $\sigma$ .

Figure 2 shows the values of  $d$  obtained from the individual measurements provided over the period 1992-1995 by the six most accurate primary frequency standards:

- \* Points from PTB CS1, PTB CS2, PTB CS3 and SU MCsR 102 are directly those obtained from the measures and are attributed error bars equal to the type B uncertainty of the individual primary standards. All these points are consistent taking into account the error bars though the two points from PTB CS3 are slightly higher than others.
- \* Points from NIST-7 are obtained from the transfer to TAI, over 10 days, of one single measurement carried out at NIST over 3 days included in the 10 day interval of estimation. They are attributed error bars equal to the type B uncertainty of NIST-7. A systematic trend of the NIST-7 frequency can be detected from the six first points. The other points are in good agreement with calibrations from PTB CS2.
- \* Points from LPTF-FO1 are obtained from the transfer to TAI, over 10 days, of several measurements, each carried out at the LPTF over about 10 hours included in the 10 day interval of estimation (there may be as many as 9 calibrations over a given 10 day interval). In Fig. 2 they are attributed error bars equal to the type B uncertainty of LPTF-FO1, though it is certainly too small as explained above. The calibrations from LPTF-FO1 give individual values of  $d$  which are smaller than

those given by NIST-7, PTB CS2 and especially PTB CS3.

The rough average over the period 1992-1995 of these measures gives a value of about  $2 \times 10^{-14}$  s which means that the duration of the TAI scale unit is rather far from the SI second. This can be easily interpreted. In the past, only calibrations of the TAI frequency provided by PTB CS1 and PTB CS2 were available and these were not corrected for the black-body radiation shift. Frequency steering corrections were applied to TAI in order to approximate the duration of its scale interval to the mean SI second provided by these two standards. When it was thought that the black-body correction should be taken into account [7], all available measurements were shifted by about  $1,7 \times 10^{-14}$  in relative frequency and that is the difference which appears here.

The individual measurements can also be treated in a global way in order to deliver a more accurate value of  $d$  for any date of the period under study which takes into account preceding and following calibrations with their respective uncertainties. This is not an easy task because individual measurements are not independent for continuously operating standards and because they are carried out over time intervals of different length. In addition, the system is not stationary since TAI stability improves with the passing of time. This problem was solved in 1977 by Azoubib, Granveaud and Guinot [4]:  $d$  is estimated in post-processing over an interval of a given length centred on a given date through a linear combination of a number of individual measurements which have occurred before or after this central date and during a period of time long with respect to the length of the estimation interval. The coefficients of the combination are determined by minimizing the error of estimation which involves:

- \* the type B and A uncertainties of each standard,
- \* a model for the stability of the time scale,
- \* correlations between measurements which are not independent,
- \* the length of individual calibration intervals,
- \* the length of the estimation interval,
- \* the length of the time interval separating the calibration intervals and the estimation interval.

In practice, those measurements taken at dates close to the central date of the estimation interval provided by the most accurate frequency standards have a large influence on the estimation. Taking into account the correlation between non-independent measurements is also very important: it prevents the estimate being swept along by PTB CS1 or CS2.

The method described in [4] is applied at the BIPM and provides the regular estimations of  $d$  which are published in the successive issues of *Circular T*. Here, estimations of  $d$  have been reprocessed for the period 1992-1995 taking into account all available measurements, especially those from LPTF-FO1 and

SU MCsR 102 made available only recently. The results shown in Fig. 2 were obtained with the following parameters:

- \* the type B uncertainty of each standard is given by the laboratory (see Table 1), the type A uncertainty is linked to the stability of each standard,
- \* the model of stability of the scale includes white frequency noise [ $\sigma_y(\tau=1d) = 2 \times 10^{-14}$ ], flicker frequency noise [ $\sigma_y = 6,3 \times 10^{-15}$ ], and random frequency modulation [ $\sigma_y(\tau=1d) = 0,5 \times 10^{-15}$ ],
- \* the correlation between measurements from PTB CS1, PTB CS2 and PTB CS3 is taken into account,
- \* the length of the individual calibration interval is 60 days except for LPTF-FO1 and NIST-7,
- \* the length of the estimation interval is 60 days (usual two month intervals of EAL computation), and
- \* the length of the maximum time interval separating the central dates of the calibration intervals and of the estimation interval decreases progressively from 24 months at the beginning of 1992 to 8 months at the end of 1995 so that our estimates in 1994 and 1995 are based more strongly on such new standards as LPTF-FO1 and NIST-7.

During 1992-1995, the successive estimations of  $d$  show a stability characterized by an Allan deviation of  $4,0 \times 10^{-15}$  for averaging times of 60 days, a value which is close to the stability of TAI over the same averaging time. The estimates obtained are always positive due to the application of the black-body correction to all measurements.

Values for 1995 are given in Table 3. The LPTF-FO1 calibrations have a dramatic influence on the estimates of  $d$  and  $\sigma$  over the period September-December 1995: the value of  $d$  which had increased from  $1,7 \times 10^{-14}$  to  $2,5 \times 10^{-14}$  abruptly decreases to  $1,2 \times 10^{-14}$  while  $\sigma$  progressively decreases from  $6 \times 10^{-15}$  to  $2 \times 10^{-15}$ . It should be noted that the results for the second half of 1995 are not definitive as reprocessing will be carried out when new calibrations are available in 1996. It seems also that a  $\sigma$  value of  $2 \times 10^{-15}$  is too optimistic. This may come from the fact that the uncertainty due to the transfer to TAI via the LPTF hydrogen maser is not taken into account by our algorithm whereas it is not negligible compared to the LPTF-FO1 type B uncertainty. It remains that the drift of the TAI frequency detected in 1995 when comparing to PTB CS2 and PTB CS3 is largely compensated by the LPTF-FO1 calibrations.

The TAI scale interval shows a global departure from the SI second on the rotating geoid which is due to the uniform application of the black-body radiation correction to primary frequency standard results. Compensation for this discrepancy has already been initiated: it takes the form of cumulative frequency steering corrections, each of relative amplitude  $1 \times 10^{-15}$ , applied on dates separated by 60 day intervals. This

procedure should not degrade the stability of the time scale.

## CONCLUSIONS

Over several years, the stability of the free running time scale EAL computed by the BIPM has considerably improved in the short-term (10 d) and middle-term (up to 80 d) thanks to the wide use of excellent commercial caesium clocks. The accuracy of TAI is estimated by comparison of the duration of its scale unit with the SI second as produced by primary frequency standards on the rotating geoid. In 1995 we faced a radical change with the suggestion to apply a frequency correction to primary standards for compensating the black-body radiation shift. In addition, new primary standards were evaluated: they are more accurate than those operating before, especially the caesium fountain built at the LPTF which is characterized by a type B uncertainty of  $3 \times 10^{-15}$ . The result is that the mean TAI scale unit over 1995 departs from the SI second on the rotating geoid by about  $1,8 \times 10^{-14}$  s, a shift accurately known (relative uncertainty of about  $5 \times 10^{-15}$ ) thanks to numerous accurate calibrations. Cumulative steering corrections have already been applied since several months for decreasing this shift: this makes TAI more accurate without degrading its middle-term stability.

## Acknowledgements

The author is grateful to her colleagues from the BIPM Time Section, J. Azoubib, G. Petit and P. Wolf for helpful discussions.

## References

1. Tavella P and Thomas C, 1991, "Comparative Study of Time Scale Algorithms", *Metrologia*, **28**, 57-63
2. Thomas C, Wolf P and Tavella P, 1994, "Time scales", *Monographie BIPM 94/1*, 52 pages
3. Thomas C and Azoubib J, 1996, "TAI computation: study of an alternative choice for implementing an upper limit of clock weights", *Metrologia*, in press
4. Azoubib J, Granveaud M and Guinot B, 1977, "Estimation of the scale unit duration of time scales", *Metrologia*, **13**, 87-93
5. "Differences between the normalized frequencies of EAL and TAI", 1994, *Annual Report of the BIPM Time Section*, **7**, Table 5, p 35
6. Thomas C and Azoubib J, 1995, "Upper Limit of Weights in TAI computation", *Proc. 27th PTTI*, in press

7. Petit G, 1995, Report on the discussions and decisions of the meeting of the CCDS Working Group on TAI, 7 pages
8. Clairon A, Ghezali S, Santarelli G, Laurent P, Lea SN, Bahoura M, Simon E, Weyers S and Szymaniec K, 1995, "Preliminary accuracy evaluation of a cesium fountain frequency standard", Proc. Symp. on Frequency Standards and Metrology, in press
9. Azoubib J and Thomas C, 1995, "The upper limit of weight for clocks contributing to TAI", Report to the CCDS Working Group on TAI, GT-TAI/95-8, 26 pages
10. Azoubib J and Thomas C, 1995, "The Use of Hydrogen Masers in TAI Computation", Proc. 9th EFTF, 283-287
11. Azoubib J and Thomas C, 1995, "Shortening of the definitive computation time of TAI", Report to the CCDS Working Group on TAI, GT-TAI/95-7, 28 pages
12. Thomas C and Azoubib J, 1995, "A modified version of the TAI algorithm under test", Report to the CCDS Working Group on TAI, GT-TAI/95-9, 15 pages
13. Quinn TJ and Thomas C, 1995, "Role of the BIPM in UTC Dissemination to the Real Time User", Proc. 27th PTII, in press

TABLE 1 -Characteristics of the primary frequency standards which provided measurements of the TAI frequency over the four years 1992-1995.

Laboratory <i>alphabetic order</i>	Standard	Type B uncertainty (1 $\sigma$ )	Mode of operation	Number of TAI frequency measurements over 1992 - 1995
CRL	Cs1	$1,1 \times 10^{-13}$	Discontinuous	1 (Dec 1992)
LPTF	FO1	$0,3 \times 10^{-14}$	Discontinuous	46 (Sep 1995 - Dec 1995)
LPTF	JPO	$1,1 \times 10^{-13}$	Discontinuous	1 (May 1993)
NIST	NIST-7	$1,0 \times 10^{-14}$	Discontinuous	14 (Jun 1994 - Nov 1995)
NRC	CsV	$\approx 1 \times 10^{-13}$	Continuous	5 (1992)
NRC	CsVIA	$\approx 1 \times 10^{-13}$	Continuous	5 (1992) & 5 (1995)
NRC	CsVIC	$\approx 1 \times 10^{-13}$	Continuous	23 (1992 - 1995)
PTB	CS1	$3,0 \times 10^{-14}$	Continuous	21 (Jan 1992 - Jun 1995)
PTB	CS2	$1,5 \times 10^{-14}$	Continuous	24 (1992 - 1995)
PTB	CS3	$1,4 \times 10^{-14}$	Continuous	2 (Sep 1995 - Nov 1995)
SU	MCsR 102	$5,0 \times 10^{-14}$	Discontinuous	2 (1992), 6 (1993), 4 (1994), 2 (1995)

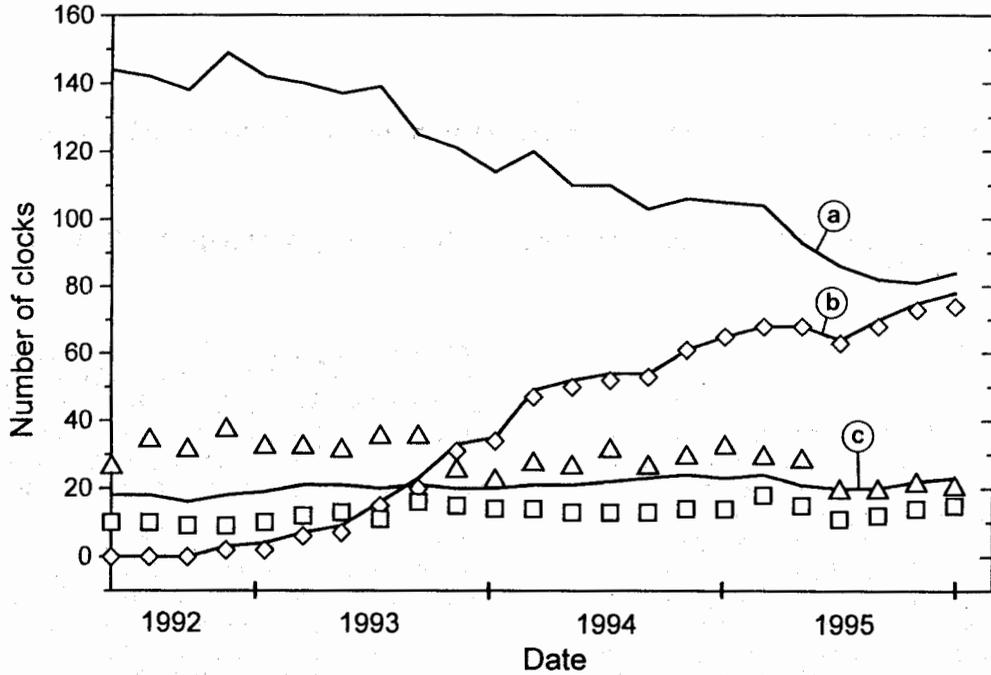
**TABLE 2 -Values of the Allan standard deviation  $\sigma_y(\tau)$  computed for the time scales EAL, E2500, E5000, E10000, ER and E by application of the N-cornered hat technique.**

The data which are used cover the periods: <sup>1</sup> April 1993 - December 1995, <sup>2</sup> January 1993 - December 1995, <sup>3</sup> January 1993 - June 1995, <sup>4</sup> July 1993 - December 1995, <sup>5</sup> January 1993 - April 1995. The time scales E2500, E5000 and E10000 correspond to absolute upper individual weights of 2500, 5000 and 10000 while the value chosen for EAL was 1000 until April 1995. The time scale ER corresponds to a relative upper individual contribution of 1,37%. The time scale E corresponds to a computation time reduced to 30 d, an absolute upper individual weight of 2500 and the introduction of a predicted frequency drift for hydrogen masers.

	$\sigma_y(\tau)$				
	$\tau = 10$ d	$\tau = 20$ d	$\tau = 40$ d	$\tau = 80$ d	$\tau = 160$ d
EAL <sup>1</sup>	$3,7 \times 10^{-15}$	$2,9 \times 10^{-15}$	$2,6 \times 10^{-15}$	$3,0 \times 10^{-15}$	$4,3 \times 10^{-15}$
E2500 <sup>2</sup>	$3,6 \times 10^{-15}$	$2,7 \times 10^{-15}$	$2,3 \times 10^{-15}$	$3,0 \times 10^{-15}$	$4,1 \times 10^{-15}$
E5000 <sup>3</sup>	$3,7 \times 10^{-15}$	$2,7 \times 10^{-15}$	$2,3 \times 10^{-15}$	$3,1 \times 10^{-15}$	$4,4 \times 10^{-15}$
E10000 <sup>3</sup>	$3,4 \times 10^{-15}$	$2,5 \times 10^{-15}$	$2,1 \times 10^{-15}$	$3,1 \times 10^{-15}$	$4,8 \times 10^{-15}$
ER <sup>4</sup>	$2,6 \times 10^{-15}$	$1,6 \times 10^{-15}$	$1,8 \times 10^{-15}$	$2,5 \times 10^{-15}$	$4,5 \times 10^{-15}$
E <sup>5</sup>	$2,7 \times 10^{-15}$	$2,2 \times 10^{-15}$	$2,6 \times 10^{-15}$	$3,9 \times 10^{-15}$	

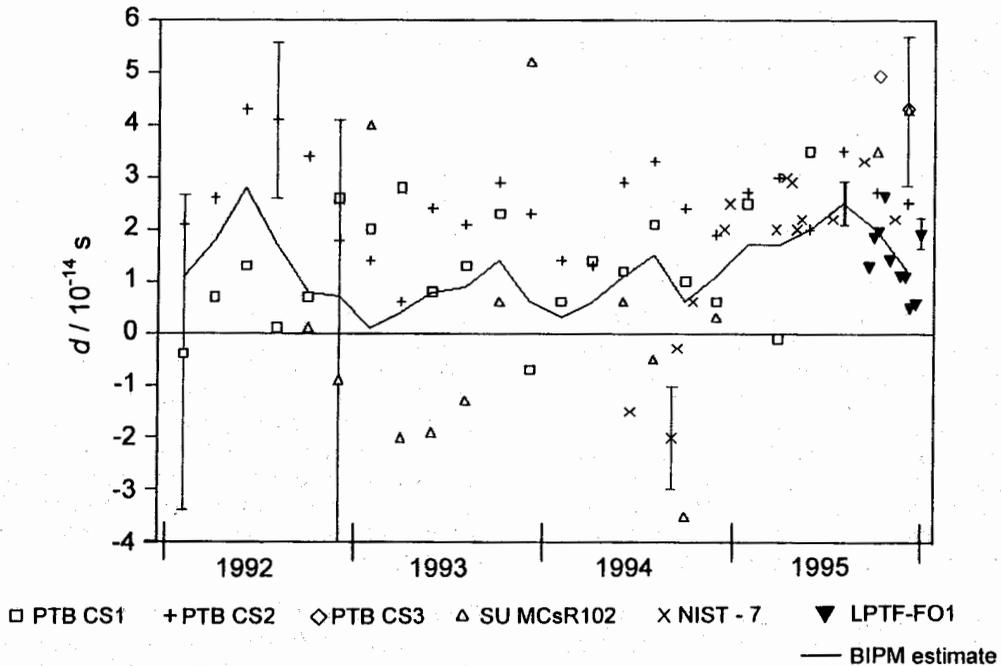
**TABLE 3 -Values of the estimate  $d$ , and of its relative uncertainty  $\sigma$ , of the departure of the TAI scale unit from the SI second as produced by primary frequency standards.**

	$d/s$	$\sigma$
January - February 1995	$1,7 \times 10^{-14}$	$6 \times 10^{-15}$
March - April 1995	$1,7 \times 10^{-14}$	$5 \times 10^{-15}$
May - June 1995	$2,0 \times 10^{-14}$	$5 \times 10^{-15}$
July - August 1995	$2,5 \times 10^{-14}$	$5 \times 10^{-15}$
September - October 1995	$2,0 \times 10^{-14}$	$2 \times 10^{-15}$
November - December 1995	$1,2 \times 10^{-14}$	$2 \times 10^{-15}$



**Figure 1.** Number of clocks weighted in the EAL computation over the years 1992-1995:

- (a) Number of caesium clocks other than HP 5071A units,
- (b) Number of HP 5071A units,
- (c) Number of hydrogen masers,
- $\Delta$  Number of caesium clocks (other than HP 5071A units) at upper weight,
- $\diamond$  Number of HP 5071A units at upper weight,
- $\square$  Number of hydrogen masers at upper weight.



**Figure 2.** Departure  $d$  of the TAI scale unit from the SI second as produced by primary frequency standards and as computed by the BIPM. The error bars given for individual points (thin line) represent the type B uncertainties of the primary standards while the error bar given for the BIPM estimate in 1995 (bold line) is the combined uncertainty resulting from the computation.

# TOWARD 10<sup>-15</sup> and SUB-NANOSECOND INTERNATIONAL FREQUENCY AND TIME METROLOGY

Robin P. Giffard<sup>1</sup>, Leonard S. Cutler<sup>2</sup>, John A. Kusters<sup>3</sup>, Mihran Miranian<sup>4</sup>, and David W. Allan<sup>5</sup>.

<sup>1</sup>, Hewlett Packard Inc., Ca., USA.; <sup>2</sup>, United States Naval Observatory, Washington D.C., USA.;  
<sup>3</sup>, Allan's TIME, Utah, USA.

## 1. INTRODUCTION

The maintenance of the international time scales TAI and UTC serves the needs of the time-keeping community in two important and different ways. The time difference between each of the contributing clocks and UTC becomes known after a calculational interval, and the timing of any event that was calibrated with respect to one of the clocks can therefore subsequently be determined with respect to the best available time-scale. Since the quality of the UTC time-scale surpasses that of the component clocks, it also provides the yardstick by which new frequency standards can be tested, and improvements in time-keeping assessed.

During the last few years, the quality of the time-scales has improved by about an order of magnitude because of improvements in the component clocks, Thomas (1). Recently a large number of improved cesium standards, and a number of very stable automatic cavity-tuned hydrogen masers have begun to contribute to the EAL. The accuracy of primary standards is being aggressively improved, and is now reported to be close to  $1.0 \times 10^{-14}$ .

It has recently been shown that if the standards available now could be ensembled in an optimum manner, the stability of the resulting time-scale would be of the order of  $1E^{-15}$  for averaging times between 1000 seconds and several months, Allan et. al. (2), Ujjanov et.al. (3). In order to realize this performance improvement in UTC, particularly at short times, it will be necessary to improve dramatically the accuracy of intercontinental time transfer. Some developments in this area are discussed in this paper. It would also be necessary to develop the use of adaptive ensembling algorithms which can make optimal use of the stability of different types of contributing clocks. Such algorithms have been discussed in detail elsewhere, Lepek et. al. (4), Weiss et. al.(5).

It is becoming increasingly important, for example in navigation and communication, to provide universal, real-time, time-scales world-wide. It is desirable that these should be closely coordinated with UTC. The stability and accuracy of such a time-scale depends critically on the quality of the local reference clock ensemble used to maintain it,

and the delay involved in time comparison with UTC. Currently UTC is not defined until several weeks after clock data is obtained, and a large ensemble of good clocks is required to estimate UTC in real-time to an accuracy better than 10 nanoseconds, as is currently achieved by USNO.

Improved time transfer techniques and the rapid data communication provided by the Internet would allow an accurate and stable implementation of UTC in real-time using a reliable and stable clock ensemble, Allan et. al. (6). This time-scale could then be disseminated with high accuracy over large distances to similar clock ensembles.

In this paper we will discuss some techniques of time and frequency transfer, and indicate where improvements can be made. Finally, some experimental results obtained using a new method of common-view GPS will be presented.

## 2. BRIEF SURVEY OF METHODS FOR ACCURATE TIME AND FREQUENCY TRANSFER

Several criteria distinguish the methods available for precise time and frequency transfer. These include the precision and accuracy that can be obtained, the level of coordination and communication required between the stations involved, the requirement for post-processing of data, and the overall cost and complexity. The traditional HF, LF, and VLF methods of frequency and time distribution have reached their limits, and have been superseded. All methods currently used over intercontinental distances use microwave, line-of-sight, transmission between ground stations and satellites to transfer time. Time transfer is carried out periodically to compare frequency. The expanding world-wide network of optical fibers has great potential for time coordination, and this is being explored.

Figure 1 summarizes the performance of some of the best methods of time and frequency comparison. The data applies to both distribution and transfer of frequency and time, and the results obtained usually depend on the separation of the clocks being compared. Loran-C is plotted as a convenient reference, although its accuracy and

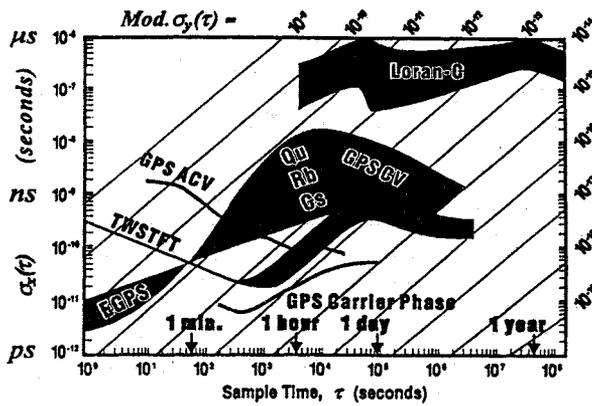


Figure 1. The stability of various significant methods of time comparison between remote clocks as a function of the averaging time. As described in the text, diagonal lines are contours of constant stability for frequency transfer under the assumption of white phase noise. For each method, the highest noise level characterizes typical time transfer between remote sites, and the lowest represents the current instrumentation noise limit. For TWSTFT, frequent observations are assumed. In the case of EGPS, the short-term noise depends on the type of disciplined local oscillator. The GPS advanced common-view (ACV) technique shows the first experimental results for the hardware only. GPS predictions are valid with SA at the current level. The methods are further discussed in the text.

stability are no longer competitive. The stability measure used is the time variance, TVAR, given by  $\sigma_x(\tau)^2 = \langle (\Delta^2 x_t)^2 \rangle / 6$ , where  $\Delta^2$  is the second difference operator,  $x_t$  is the time difference averaged over an interval  $\tau$  and the brackets " $\langle \rangle$ " denotes the expectation operation. What is plotted is  $\sigma_x(\tau)$  for each of the different techniques. Since  $\sigma_x(\tau) = \tau \text{Mod}\sigma_y(\tau) / \sqrt{3}$ , the  $\text{Mod}\sigma_y(\tau)$  values for each decade of stability can be shown on the same graph. This is particularly useful since the confidence on the estimate of the frequency difference measured over an interval  $\tau$  is given by  $2 \text{Mod}\sigma_y(\tau)$  when the residuals are modeled by white-noise PM. If this noise model is not valid, the  $2 \text{Mod}\sigma_y(\tau)$  value still provides a useful estimate of the confidence for using a particular technique for frequency transfer.

Two-way satellite time and frequency transfer (TWSTT or TWSTFT) is a moderately high cost method with high accuracy, Hanson (7). Recent work has shown that instrumental noise levels of about 30 picoseconds can be attained after only 100 seconds of averaging, Hackman et. al. (8). There is evidence that environmental effects make this time transfer performance difficult to extend to longer times.

Figure 1 shows upper and lower estimates for the TWSTT technique representing the stabilities for continuous operation. The bottom curve shows the measured currently attainable instrumentation stability limit. The upper curve is a more typical

stability performance observed between two sites remote from each other. The upper limit is dotted at large times to indicate that TWSTT is not typically used continuously for this range of sample times, but rather three times per week. The frequency transfer uncertainty will not be as good as that shown, but would be approximately given by  $1 \text{ ns}/\tau$ . Distance up to about 9 Mm have been realized.

Time transfer using GPS is now widespread. The common-view technique was a major breakthrough for international time and frequency comparisons and is still today the main means of communicating the times of most of the contributing clocks in the generation of TAI/UTC, Allan et. al. (9). Originally, day-to-day stabilities of as good as 0.8 ns were obtained between stations as far apart as Boulder, Colorado and Ottawa, Canada. Global accuracies of about 4 ns have been demonstrated using careful post-processing. However, since the original GPS common-view receivers were placed in service, there seems to have been a gradual degradation in the performance of this technique. Day-to-day stabilities of from 2 to 8 ns are now more typical, and significant environmental effects have been measured due to antenna and lead-in cable temperature sensitivities. Problems have crept into the common-view technique at about the 10 ns level. Although the introduction of SA at this time as an intentional jitter of the satellite clocks does not affect common view measurements directly, it may have had indirect effects. The reasons for this loss of accuracy are being investigated.

Qualified users with "keyed" receivers can employ dual-frequency measurements of the ionosphere to both avoid SA and reduce the errors due to propagation. Some use of the L2 signal can be made by other users with suitable receivers.

The GPS carrier-phase technique has potentially the smallest uncertainty for frequency comparisons between remote clocks. Experiments have been conducted comparing separated hydrogen masers using geodetic type receivers at both sites. By locking to the carrier-phase of the same GPS satellite at the two sites simultaneously, RMS residuals as small as 30 ps were measured. The data plotted in Figure 1 is from an experiment made between Goldstone, California and Algonquin Park, Canada. The baseline distance is about 3400 km. About 35 monitor stations were involved in determining accurate ephemerides for the satellites. This is a complicated technique involving expensive receivers and a large amount of data processing. For applications in which this is acceptable, this technique is among the best for minimizing remote frequency comparison uncertainty.

The GPS methods described above require extensive data exchange. If post-calculated ephemeris data is used to increase accuracy, the results may not be available for many days after the measurements are complete. Two simpler methods of local synchronization described as "EGPS" and the "melting pot" have been developed which avoid these difficulties. These methods can

take advantage of the compact digital multi-channel receivers which are now available.

In the melting pot method all the data available during one day from a conventional GPS time receiver is averaged to obtain an estimate of difference between a local reference clock and UTC. A schedule is used to ensure that the receiver always tracks a satellite which is high in the sky. Since this is not a common-mode method, the noise level is limited by SA. Although the time estimate is delayed by half of the averaging window, this is a very cost effective technique, and daily stabilities of about 5 ns are achievable with a good clock and receiver.

For the EGPS method, signals from several satellites are simultaneously tracked in a multi-channel receiver. The receiver uses the data in the navigation message to calculate a best-fit time solution referred to UTC. The solution is used to generate an "on-time" 1 pps output pulse. The noise on the solution is reduced if the receiver is operated in the "position-hold" mode with accurately known antenna coordinates. With the current constellation of satellites, 4 to 8 are usually simultaneously visible with a mask angle of 15 degrees at typical latitudes. The noise on the time solution is typically between 30 to 40 nanoseconds with an autocorrelation time between 100 and 200 seconds, and is limited by current levels of SA. When a suitable local oscillator is available, it can be used in a systems approach to filter the time output, reducing the short term rms noise. For example, if a good cesium clock is used as the reference, one-day stabilities of a few nanoseconds may be expected in real time.

The EGPS method is very cost-effective, and generates a local 1 pps tick in real time. The stability obtained depends on the local oscillator, and typical values are shown in figure 1. The use of many satellites results in considerable averaging of ephemeris and satellite clock errors. The long term accuracy of the local on-time signal should depend only on the broadcast GPS/UTC offset which is usually accurate to within 20 nanoseconds. If the observing stations are reasonably close (1000 km for example) it is possible to take advantage of correlations in the data to reduce the effects of SA at the expense of communication and post-processing. Using excellent reference clocks, instrumentation errors have been documented at about 1.5 ns, Kusters et. al. (10).

### 3. DESCRIPTION OF THE GPS-ACV METHOD

We have recently begun to investigate a method of time comparison described as GPS-ACV or advanced common-view GPS. This is a common-view method that uses the largest possible number of observations on all satellites in common view to average the short term noise and reduce the effect of errors. The goal is to make maximum use of the available degrees of freedom provided by the data. The method also takes advantage of the efficient communication provided by the Internet.

Inexpensive modular digital GPS receivers can be used to implement this technique if they provide an on-time output, and have sufficient software flexibility. Environmental stability of the receiver group delay is important.

The ACV method consists of recording, in real time, a pseudorange correction for each satellite in view using the known coordinates of the receiving antenna and the time of the local station clock. The corrections are computed using the ephemerides, the satellite clock models, and the ionospheric corrections in the navigation message. The delays in the antenna, the antenna cable, and the GPS receiver are included in the calculation. The pseudorange corrections are calculated each second and low-pass filtered to reduce the data rate. The filter currently used tracks the first two derivatives of the data with minimal error. The pseudorange correction data for each satellite in view, together with a precise time-stamp, is filed locally at about 10 second intervals. This procedure is analogous to that used in differential position GPS, with the additional feature that the arbitrary receiver clock bias is removed by reference to the local station clock.

In order to perform a time comparison, two stations compare their pseudorange correction data by exchanging files. This can be done in a few seconds using the Internet. An average is then formed by subtracting the data from all satellites that are in common view, as indicated by corrections for the same satellite appearing in both sets of data. The average represents the estimated difference between the station time scales at the time of the common time stamp.

The choice of the low-pass filter time constant is a compromise between the advantage of reducing the raw data rate, and the need to follow the variation of the pseudorange corrections with SA. Slower filters can be used as long as the same filter is used at each station in a time comparison. Very long time constants lead to a loss of observing time while the filters settle after a change in the constellation being tracked. A 10 second filter time allows comparisons with conventional common-view data. Using internet, time comparisons with an accuracy of a few nanoseconds can be made in a few seconds. Alternatively, the data can be averaged for as long as is appropriate. Time differences can be calculated by anyone who has access to both sets of data. Ideally, corrections for all visible satellites are recorded at each location. Data is then available for time comparison with all other stations using the maximum possible number of satellites in common-view.

From a simple "degrees of freedom" argument there is significant advantage to the GPS ACV technique. If the measurement noise for a certain averaging time is white PM, then the confidence on the estimate of the frequency difference between two ideal reference clocks, as determined from a linear regression to the time-difference residuals taken between the two clocks, is:  $\sqrt{12} \cdot \sigma / (\tau_0 \cdot n^{3/2})$ , where  $\sigma$  is the standard deviation of the white-noise residuals,  $\tau_0$  is the measurement interval, and "n"

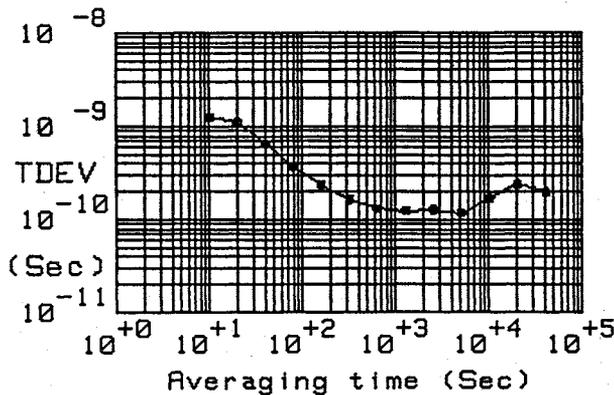


Figure 2. The rms time deviation, TDEV, as a function of averaging time, of time difference data observed between two GPS-ACV systems using a common antenna and a common clock. The time difference is the average using all satellites in common view. Both the vertical and the horizontal scales are in seconds. The systems are described in the text.

is the degrees of freedom (the number of independent measurements). In the method described above,  $n$  is the product of the length of the observations, in units of the averaging time, and the number of satellites in common view.

The errors in this method are the same as those apparent in conventional common-view observations, but it is anticipated that the overall effect of some will be reduced. The use of multiple satellites in the ACV technique dilutes those errors which have a randomly varying nature but which do not introduce bias, Wiess et. al. (11). Such errors should include ephemeris errors and multipath effects. Errors which repeat each sidereal day, including coordinate errors, multipath effects, and some failures of the ionospheric model, can be reduced by averaging or the use of an adaptive filter. Such errors, even if they introduce bias, will not affect frequency comparison. As a result of the use of common-view, satellite clock errors and the direct effect of SA through clock jitter are removed as usual.

Ionosphere and troposphere modelling errors will affect ACV and conventional common view in the same way. The broadcast ionospheric model has been used in experiments so far. Inexpensive receivers that perform ionospheric measurements are not yet available. Ideally, GPS ACV would take advantage of the best available ionospheric and ephemeris measurements.

In order to evaluate the practical magnitudes of these effects some experiments were undertaken. The results of these are described in the next section.

#### 4. EXPERIMENTS AND RESULTS

The measurement system used in the experiments employs an 8-channel, modular, GPS engine (12).

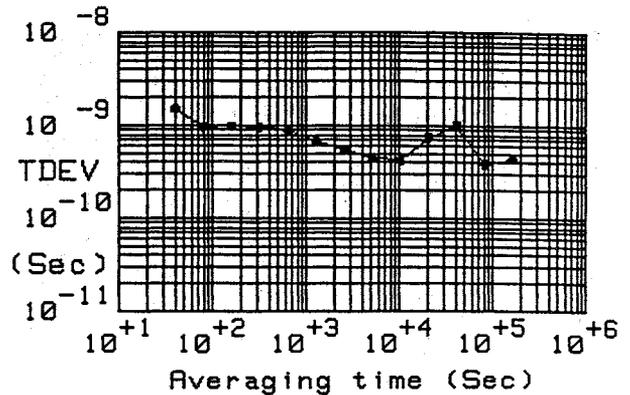


Figure 3. The rms time deviation, TDEV, as a function of averaging time, of time difference data observed between two GPS-ACV systems using a common clock and antennas separated by 17 m. The time difference is the average using all satellites in common view. The ten-second difference data was decimated by a factor of 4. Both the vertical and the horizontal scales are in seconds. The bulge in the TDEV value at 4000 seconds is believed to be due to temperature effects in the system that was in an uncontrolled environment.

The 1 pps output of the receiver is compared with the station clock using a time interval counter with a resolution of better than 1 nanosecond. The data from the receiver and counter is processed by a simple computer, and output via an RS-232 serial interface each 10 seconds. The ASCII output data is stored on a file server or a Unix workstation. The amount of data is typically about 1 Mbyte per day before processing.

The first experiment consisted of operating two identical measurement systems on the same antenna, using the same local clock. With an average of six satellites tracked, the 10 second average time differences showed an rms deviation of 1.4 nanoseconds. Figure 2 shows the TDEV calculated for 2.5 days of data. The TDEV value falls to a value of about 0.13 ns at 1,000 seconds. This experiment indicates the instrumentation limit excluding the effect of the signal to noise ratio at the antenna. The increasing noise at long times is probably due to environmental effects.

This experiment was repeated using two dissimilar antennas separated by 17 m. The relative positions of the antennas were accurately known, and the coordinates used in the receivers were adjusted accordingly. A common clock was used, and the data was processed by identical software at each location. Six satellites were again tracked for most of the time, and the rms of the 10 second averages was found to have increased to 3.0 nanoseconds, presumably due to antenna and amplifier noise. Time differences were calculated for a period of 8 days. The data shows a daily variation of about 4 ns peak to peak. Subsequent experiments have shown that this could have been due to temperature: one system was in an uncontrolled environment. Figure 3 shows the TDEV for this data. The most important difference

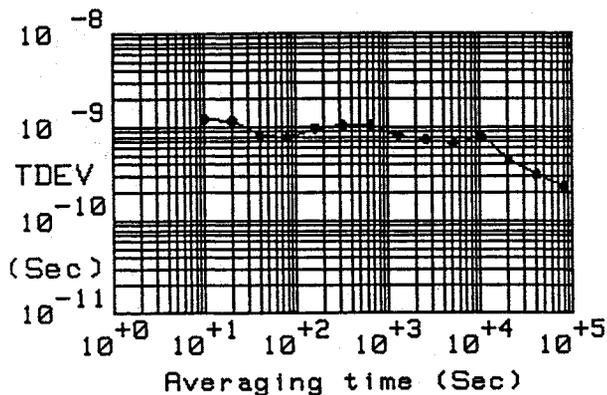


Figure 4. The rms time deviation, TDEV, as a function of averaging time, of time difference data observed between two GPS-ACV systems separated by 15.5 km with different clock ensembles. The time difference is the average using all satellites in common view. Both the vertical and the horizontal scales are in seconds.

from previous data is the increase in noise at about 1,000 seconds which is probably due to multipath effects. The baseline is too small for satellite ephemeris or antenna coordinate errors to be significant. The TDEV value at 1 day is 0.37 ns, and this value is not affected by the diurnal variations.

In a third experiment, data from two laboratories separated by 15.5 km was compared over a continuous period of 4 days. The antennas were dissimilar but identical receivers and software were used at each location. The station clocks in this experiment consisted of ensembles of two and three HP 5071A cesium clocks respectively. The antenna coordinates were determined in advance by averaging the output of the GPS receivers for about 7 days. Six satellites were tracked for most of the time. The rms deviation of the 10 second average differences was 2.4 ns. Figure 4 shows the TDEV for data taken continuously over 4 days. The TDEV value at 10,000 seconds was 0.8 ns. Examination of the data one satellite at a time shows that a relative horizontal position error of about 2 m probably exists, which could give rise to this noise. At long times, the data showed more noise than that from the previous experiment in agreement with the use of independent station clocks. The combined TDEV of the clocks at 1 day is about 0.9 ns. The observed TDEV at 1 day was 0.4 ns, with a large statistical uncertainty.

These experiments are currently being extended to a larger baseline by installing a GPS-ACV station at USNO in Washington D.C. It is hoped to obtain data simultaneously from this station and the existing stations at Hewlett-Packard Labs in Palo Alto, and Hewlett-Packard at Santa Clara. The baseline between USNO and Northern California is about 4 Mm.

## 5. CONCLUSIONS

We have investigated a relatively simple, GPS-based, time comparison system described as "Advanced common-view" using modular GPS receivers. The results of preliminary experiments over short baselines indicate a very low instrumental noise level, and a one-day stability of about 0.3 nanoseconds. The experiments are being extended to longer baselines. If this stability is reproducibly obtained, or even improved, the technique could be used to improve the quality of the data used to determine TAI and UTC. Since the method is easily automated and rapid, it might be useful in implementing real-time UTC.

### Acknowledgment:

We are grateful to Mike King of Motorola for helpful advice on the operation of the GPS receivers.

### References:

1. C. Thomas, 1996, "Stability and accuracy of the international atomic time TAI", presented at the EFTF 1996, see these proceedings.
2. D.W. Allan, A. Lepek, L. Cutler, R. Giffard, and J. Kusters, 1995, "A 10-15 International Frequency Reference Available Now", Proceedings of Fifth Symposium on Frequency Standards and Metrology, Woods Hole, MA, October 1995.
3. A.A. Uljanov, N.A. Demidov, E.M. Mattison, R.F.C. Vessot, D.W. Allan, and G.M.R. Winkler, 1990, "Performance of Soviet and U.S. Hydrogen Masers", 1990, Proceedings of the 22nd Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, NASA Conference Publication 3116, 509 524
4. A. Lepek, A. Shenhar, and D. W. Allan, "INPL Virtual Time Scale, UTC(INPL)", to be published, Metrologia.
5. M.A. Weiss, D.W. Allan and T.K. Pepler, "A Study of the NBS Time Scale Algorithm", IEEE Transactions on Instrumentation and Measurement, **38**, 631 635
6. D.W. Allan, R.P. Giffard, L.S. Cutler, and J.A. Kusters, 1995, "Improved time and frequency opportunities via GPS", Proceedings of the 1995 European Frequency and Time Forum, 71 74
7. D. W. Hanson, "Fundamentals of two-way time transfers by satellite", Proceedings of the 43rd Annual Symposium on Frequency Control, IEEE Catalog No. 89CH2690-6, 174 178

8. C. Hackman, S.R. Jefferts, and T.E. Parker, "Common-Clock Two-Way Satellite Time Transfer Experiments", Proceedings of the 1995 International Frequency Control Symposium, IEEE Catalog No. 95CH35752, 275 281.

9. D.W. Allan, D.D. Davis, M. Weiss, A. Clements, B. Guinot, M. Granveaud, K. Dorenwendt, B. Fischer, P. Hetzel, S. Aoki, M.-K. Fujimoto, L. Charron, and N. Ashby, 1985, "Accuracy of International Time and Frequency Comparisons Via Global Positioning System Satellites in Common-View", IEEE Transactions on Instrumentation and Measurement, IM-34, No. 2, 118 125

10. Results using the EGPS technique were discussed at the 1995 Conference on Precision Electromagnetic Measurements, Boulder, 1995.

11. M.A. Weiss, G. Petit, S. Shattil, T. Tauri, 1993, "A Comparison of GPS and DMA Precise Ephemerides", Proceedings of the 25th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, NASA Conference Publications, 293 305.

12. The receivers used were modular 8-channel highly integrated digital units. During the experiments a temperature-dependant delay was found. The delay was found to increase parabolically as  $(T-25C)^2$ , where T is the temperature in Celsius.

# A NOVEL SLAVE-CLOCK IMPLEMENTATION APPROACH FOR TELECOMMUNICATIONS NETWORK SYNCHRONISATION

R. Urbansky, W. Sturm  
AT&T Network Systems Germany, Nuremberg

**Abstract** - Network synchronisation has gained increasing attention since the introduction of the Synchronous Digital Hierarchy (SDH), as the network synchronisation performance has a major impact on the phase transfer characteristic of SDH-based networks. This is due to the SDH-internal bit rate adaptation technique.

This paper discusses a new approach for the cost-efficient implementation of SDH Equipment Clocks (SEC), which are the basis for improved network synchronisation. The characteristics of the oscillator and the phase detector have a significant effect on the phase error generated by the clock. The paper proposes a synthesiser-based PLL structure employing a fixed frequency highly stable oscillator. A novel approach for an all-digital phase detector provides enhanced resolution, thereby reducing the phase error. Theoretical results obtained from analytical calculations and simulations are complemented by measurement results of a prototype clock to demonstrate the feasibility of this approach.

## I Introduction

In current PDH (Plesiochronous Digital Hierarchy) based networks digital exchanges need a common synchronisation reference signal to prevent byte (frame) slips. The synchronisation links between the exchanges are provided by 2-Mbit/s-PDH-connections. PDH networks are operated plesiochronously and employ bit-stuffing techniques to accommodate phase- and frequency variations between clock rate and data rate, thus minimising phase transfer errors.

The SDH as the new international standard does not provide for timing transparent synchronisation links. In addition, SDH NEs (Network Elements) need a common network synchronisation to prevent excessive jitter and wander of PDH signals carried over the SDH. In [2], [3] an improved phase- and frequency justification technique has been proposed to avoid the need for improved network synchronisation requirements. However, in ITU-T Recommendation G.803 the currently recommended synchronisation technique comprises an SDH-based synchronisation distribution network. The synchronisation signals are carried via the SDH line signal and so-called SDH Equipment Clocks (SECs) in addition to the highly stable switch-internal slave clocks according to ITU-T Rec. G.812. The timing characteristic of the SDH line signal is not impaired by wander due to buffer fill variations or by jitter associated with occasional bit rate justification events. The synchronisation signals are

regenerated at every SDH NE in the data path between exchanges rather than passed transparently through the network. This implies the need for modified equipment clock specifications for normal mode concerning the increased number of clocks in a chain in conjunction with jitter and wander propagation characteristics.

Section II presents a Phase-Locked-Loop (PLL) structures suitable for slave clock applications. It comprises a synthesiser replacing the VCO (Voltage Controlled Oscillator) and it focuses on an efficient digital implementation based on an all-digital phase detector. The phase detector as the predominant source of phase errors for short observation intervals is modelled and analysed in section III. Section IV presents simulation and measurement results of the phase detector performance in the slave clock application. Section V summarises the main results of the paper.

## II Slave-Clocks

Slave-clocks for telecommunications applications are intended for use in synchronisation distribution networks. The basic functions include jitter reduction of the received reference signal, reference signal switching for protection purposes and the generation of a highly stable output signal in holdover mode if all reference signals have failed to provide a minimum quality of service. In normal mode of operation the input jitter is reduced by a low-bandwidth PLL. In this application an oscillator with a sufficiently high short-term stability is required. Additionally, the oscillator has to meet the long-term stability requirements for the holdover mode.

Different PLL structures for slave-clocks have been applied, which include analog, digital or hybrid implementations. SECs are expected to be an integral part of the SDH equipment. Component costs and space are important design criteria. This implies that the amount of analog circuitry should be minimised and an all digital implementation should be preferred, which allows for a higher degree of integration. This results in a reduced number of components associated with reduced cost, power consumption and improved reliability. Furthermore, the parameter sensitivity and internal cross talk effects are minimised and a robust implementation is obtained.

Fig. 1 shows the block diagram of a PLL. It comprises a digital phase detector, a loop filter and a numerical controlled oscillator. In this figure the digital phase detector is represented by an analog phase detector followed by an

analog-to-digital converter (ADC). This combination can be simplified by a counter-based phase detector. The digital loop filter can be implemented by a micro- or signal processor or as a hardwired solution which allows integration into an ASIC (Application Specific Integrated Circuit). The numerical controlled oscillator is represented by a voltage controlled oscillator (VCO) which is fed by a digital-to-analog converter (DAC). The conflicting parameters of the oscillator with respect to low drift (high stability, low noise) and large frequency range may be optimised separately by employing an all-digital implementation, where the stability is provided by a stable fixed-frequency oscillator, and the frequencies are generated by programmable dividers.

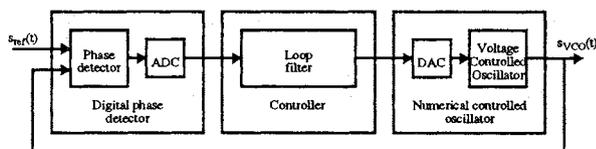


Fig. 1. PLL block diagram

The digital implementation of the controlled oscillator employing a stable fixed-frequency oscillator in combination with a controllable divider can be considered as a frequency synthesiser. The inherent jitter of such a fractional-N divider corresponds to the reference clock period. It can be reduced by a subsequent PLL. If the frequency range of the jitter spectrum is sufficiently high, analog PLLs with high cut-off frequency and simple oscillators may be employed. The fractional-N divider may be included in the feed forward path or in the feedback loop of the jitter suppression PLL. In [4] an accumulator-based fractional-N divider as a numerical controlled oscillator has been presented. In [5] the fractional-N divider is included in the feedback loop and is controlled by a higher order Sigma-Delta Modulator for spectral-jitter shaping purposes.

The error of the output sequence of a higher order Sigma-Delta Modulator equals a high-pass-filtered white-noise spectrum. This output sequence controls the divider stage of the fractional-N divider by selecting appropriate integer values to approximate the desired frequency. The phase error is given by the frequency error due to the coarse integer division approximation in conjunction with the sampling interval, which equals an output signal period. Thanks to the Sigma-Delta Modulation this phase error consists only of high-frequency components. This high frequency phase error is almost fully suppressed by including the fractional-N divider in the PLL feedback loop. The absence of low-frequency phase errors in conjunction with the jitter suppression capabilities of the analog PLL leads to an output frequency with negligible divider induced self-jitter. The frequency range and the resolution can be selected as required by choosing the suitable word length in the digital circuitry without affecting analog components.

Implementations of digital PLLs have been already re-

ported and analysed in [1], [6]. In these widely used implementations the clock input of the phase detector counter is fed by the oscillator of the PLL. The disadvantage of this method is that the operating frequency determines the resolution. Any subsequent averaging process does not improve the resolution for constant frequencies due to the correlation of the oscillator signal with respect to the reference signal, which provides a constant output signal of the phase detector. According to the novel approach the resolution can be improved by resampling the phase detector output signal with an independent clock and subsequent averaging.

Fig. 2 shows the block diagram of the narrowband PLL for use in SECs. Supervision, control and monitoring components are omitted in this diagram. The digital phase detector comprises the original phase detector, the resampling circuit followed by an averaging filter. The averaging filter is required to improve the resolution of the resampled phase values. This ensures a smooth phase output transition of the slave clock if it returns from holdover mode to the locked mode. The Proportional-Integral-controller determines the transfer function of the PLL. It may be realised as a hardwired solution (FPGA, ASIC) or included in the firmware of a microprocessor. The numerical controlled oscillator is implemented as a digital frequency synthesiser. It consists of an analog high-bandwidth PLL (about 1 kHz). The output frequency is varied by encoding the PI-controller output signal into a high frequency binary bit stream by a second order Sigma-Delta modulator, which selects the divider ratio of the PLL feedback loop.

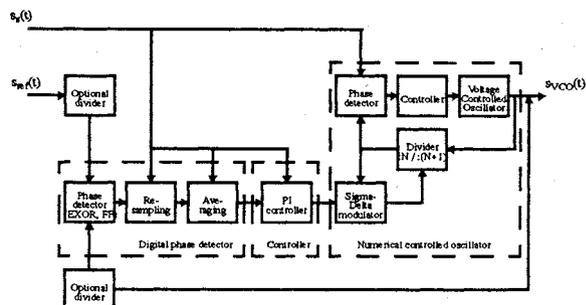


Fig. 2. Block diagram of the slave clock PLL.

### III Phase Detector

The resolution of the previously mentioned counter-based phase detectors is determined by the high-rate clock of the counters. An increase in the counter frequency used improves the resolution of this type of phase detector. This method, which may require a combination of different technologies, complicates the circuitry, causes interface problems and suffers from power consumption. This is not adequate if all components of the phase detector are to be integrated in one technology. The frequency stability of the high-rate clock oscillator determines the accuracy of the phase detector. The bandwidth of the input phase modulation and the bandwidth of the PLL do not require such high sampling rates. In principle, the

sampling frequency must take into account the Nyquist frequency of the input phase modulation as a lower limit in order to avoid aliasing effects regardless of the resolution of this phase detector. Consequently, resampling phase detectors are intended to operate at low frequencies which are easy to handle.

The principle task of a phase detector is not to reconstruct the original high frequency phase difference waveform but to determine a mean phase difference which varies slowly with respect to the PLL time constant. If the sampling frequency  $f_s$  equals an integer multiple of the reference frequency  $f_{ref}$  we would expect that the resolution of the phase detector corresponds to the sampling period  $T_s$ . The quality of the phase detector output signal is determined mainly by the exact value of the sampling frequency. A major consideration for the design of this type of resampling phase detector is the selection of a suitable sampling frequency.

Typically in signal processing applications the signals are bandlimited before AD conversion. This procedure provides an undistorted time-discrete representation of the bandlimited phase difference and prevents aliasing effects. The phase transfer function of a slave clock can be represented by a narrow-band low-pass filter. If the aliasing components in the spectrum are located outside the pass band, they are suppressed. The choice of the sampling frequency with respect to the reference frequency offers a degree of freedom for the shaping of the aliasing components. A suitable choice of the sampling frequency enables the digital filter to process unfiltered, non-band-limited signals. Furthermore, thanks to the binary representation of the phase difference signal, the AD converter can be omitted. This approach simplifies the implementation of this resampling all-digital phase detector and imposes no new requirements with respect to the design technology if appropriate sampling frequencies can be found in the range up to the reference frequency.

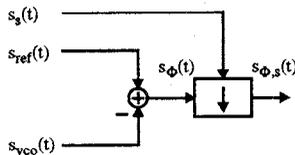


Fig. 3 Block diagram of the resampling phase detector

Fig. 3 shows a block diagram of this resampling phase detector. The phase difference signal  $s_\phi(t)$ , representing the difference between the reference signal  $s_{ref}(t)$  and the VCO signal  $s_{vco}(t)$ , can be considered as being generated by an EXOR type phase detector. Other types of phase detectors may be employed. They may be complemented by preceding divider stages which modify the phase detector gain. The time resolution is not affected, but the overall jitter characteristic of the PLL deteriorates in the case of reduced gain. The phase difference signal is sampled by the periodic signal  $s_s(t)$ . The subsequent averaging filter is inserted to suppress high-frequency error

components resulting from the sampling process.

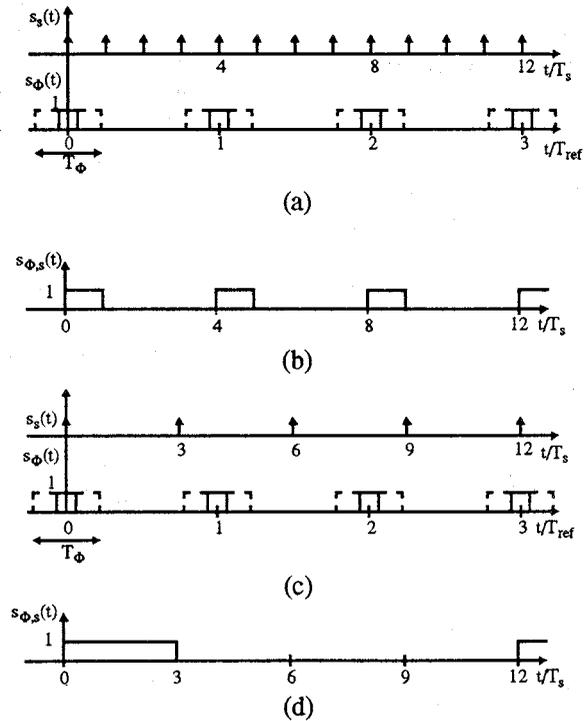


Fig. 4 Sampling signal  $s_s(t)$ , phase difference signal  $s_\phi(t)$  ((a)  $f_s = 4 \cdot f_{ref}$ , (c)  $f_s = 3/4 \cdot f_{ref}$ ; phase difference  $0 < T_\phi < 2T_{ref}$ ); Sampled phase difference signal  $s_{\phi,s}(t)$  ((b)  $f_s = 4 \cdot f_{ref}$ , (d)  $f_s = 3/4 \cdot f_{ref}$ ; phase difference  $0 < T_\phi < 2T_{ref}$ ).

Fig. 4 (a) shows a typical waveform of the phase difference function  $s_\phi(t)$  and the sampling function  $s_s(t)$  versus time. The phase difference signal and the time axis are scaled in radians and time period of the reference signal  $T_{ref}$ , respectively. Fig. 4 (b) shows the sampled phase difference function which is obtained at the output of a hold circuit. The maximum phase detector error is represented by the dashed line of the phase difference signal. Variations of the phase difference pulse width  $T_\phi$  in the interval  $(0, 2 \cdot T_s)$  do not have any effect on the sampled signal.

In Fig. 4 (c) the phase difference function is subsampled with a third of the frequency in Fig. 4 (a) which results in strong aliasing effects. Fig. 4 (d) shows the output signal of a hold circuit as in Fig. 4 (b). In this example after averaging the same result is obtained as in Fig. 4 (b) where the higher sampling frequency has been applied. The shape and the mean value of this phase difference waveform are not altered by the subsampling procedure with this frequency. This result shows that the same phase detector resolution can be obtained by applying lower resampling frequencies. However, for other fractions of the sampling frequency, e.g.  $f_s/(2 \cdot n)$ , the corresponding aliasing effects result in a modified waveform and a different mean value associated with the phase detector error. This demonstrates that the exact value of the sampling frequency has a significant impact on the accuracy of the phase detection.

This time domain example shows some general aspects

of the phase detector error characteristics. The phase detector error can be considered as resulting from aliasing effects due to the sampling process. The frequency domain analysis allows the spectral components of the original phase difference and the sampling error component to be separated.

The combination of the phase detector and the signal post processing results in a phase difference representation of limited resolution which corresponds to a phase quantisation. A constant signal is considered as the worst case input signal for a quantiser with respect to the spectral components of the quantisation error. A constant signal with respect to a phase quantiser refers to a constant phase difference. Therefore, the steady state characteristics of the phase detector error have been analysed in [10]. In addition to the DC error component all other spectral components contribute to the phase detector error. However, this error can be reduced to negligible values by setting the cut-off frequency of the PLL to sufficiently low values.

The spectrum  $S_{\phi,s}(f)$  of the sampled rectangular waveform  $s_{\phi,s}(t)$  is calculated by sampling the original phase difference waveform having the pulse width  $T_\phi$  and period  $T_{ref}$  with the sampling frequency  $f_s$ , where  $\rho = f_{ref}/f_s$ . In the case where  $\rho$  is irrational, the signal  $s_{\phi,s}(t)$  is not periodical and the spectrum is continuous. The phase difference is infinitely fine quantised and the constant phase detector error does not exist. However, the recovery of the ideal phase difference would require an infinite averaging time. The real estimation of the phase difference implies a finite averaging time at cost of increased noise components.

If  $\rho$  is rational, namely  $\rho = p/q$  with  $p$  and  $q$  coprimes, the signal  $s_{\phi,s}(t)$  is a periodical function with period  $q \cdot T_s$ , and its spectrum  $S_{\phi,s}(f)$  in the interval  $[0, f_s)$  can be written in the form [10]:

$$S_{\phi,s}(f) = \sum_{r=0}^{q-1} Y_r \cdot \delta\left(f - \frac{r}{q} f_s + \left\lfloor \frac{r}{q} \right\rfloor f_s\right) \quad 0 \leq f < f_s \quad (1)$$

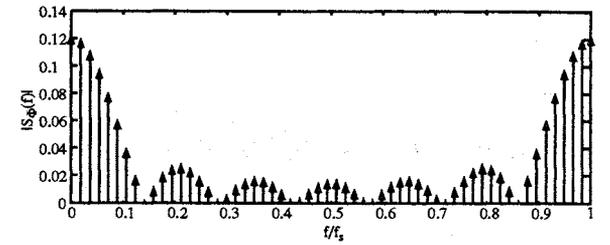
$$Y_r = \begin{cases} \frac{1}{q} \cdot (2m+1) & r = 0 \\ \frac{\sin\left[\frac{r}{q} \cdot \pi \cdot (2m+1)\right]}{q \cdot \sin\left[\frac{r}{q} \cdot \pi\right]} & r = 1 \dots (q-1) \end{cases} \quad (2)$$

$$2m \cdot \frac{T_{ref}}{q} \leq T_\phi < (2m+2) \cdot \frac{T_{ref}}{q}$$

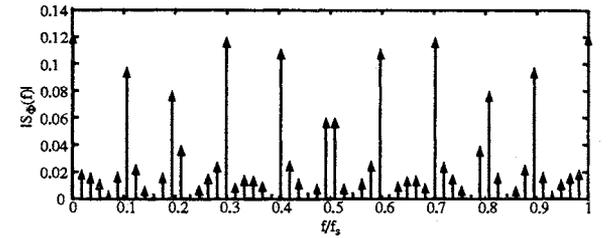
The positive integer coefficient  $m$  describes the phase difference  $T_\phi$  as multiples of  $2 \cdot T_{ref}/q$ . Because of the periodicity of the sampled phase difference function  $s_{\phi,s}(t)$  the spectral lines of the phase difference are located at multiples of the frequency  $f_s/q$ . The amplitudes of the spectral lines are given by  $Y_r$ , which are a function of the positive integer  $m$ , the integer  $q$  and the index of the spectral lines  $r$ .

Fig. 5 shows examples of the spectrum of the sampled phase difference signal  $S_{\phi,s}(f)$  for two different values of  $p$ . For both examples the parameters chosen are:

$q = 57$  and  $m = 3$ . The spectrum of the unsampled rectangular phase difference waveform has an si-characteristic. In the case where the sampling frequency  $f_s$  is a fraction of the reference frequency  $f_{ref}$  corresponding to  $p = 1$ , the rectangular phase difference waveform is sampled with the frequency  $f_s = q f_{ref}$  resulting in an si-type spectrum with small aliasing components. For  $p = 17$  the spectral lines in Fig. 5 b) are reordered compared to Fig 5 a). This is due to the lower sampling frequency associated with stronger aliasing components located at different frequency indices, as the sampling frequency is not an integer multiple of the reference frequency. In this example, significant spectral lines are shifted to higher frequencies. Therefore, the selection of the numerator  $p$  offers a degree of freedom for the shaping of the spectral components, which have to be suppressed by a subsequent filter.



(a)



(b)

Fig. 5 Spectrum of the sampled phase difference signal  $S_{\phi,s}(f)$  ( $m = 3, q = 57$ , (a)  $p = 1$ , (b)  $p = 17$ ).

In addition to the analysis of the low-frequency components, which is known from jitter analysis, the DC component  $Y_0$  will be examined in more detail to establish the resolution of the phase detector.

If  $T_\phi$  is increased from 0 to higher phase difference values associated with an increase of  $m$ , the amplitude  $Y_0$  is altered only every  $2 \cdot T_{ref}/q$ . If the phase detector is considered as a quantiser the parameter  $m$  corresponds to the index of the quantisation steps  $2 \cdot T_{ref}/q$ . Consequently, the resolution of the phase detector is determined by the minimum value  $q_{min}$  within the frequency tolerance range of  $f_{ref}$  and  $f_s$ .

This shows that a continuous increase of the sampling frequency  $f_s$  does not imply a corresponding improvement of the resolution. The resolution is determined by the parameter  $q$  which in turn is not linearly increased with respect to the sampling frequency. As this parameter is determined by the quotient between reference and sampling frequency  $\rho = f_{ref}/f_s = p/q$ , small values for  $q$  can be found even in high frequency ranges. Conversely, there exists a set of suitable sampling frequencies smaller

than the reference frequency which results in sufficiently high values of the parameter  $q$ .

Fig. 6 shows the minimum value for the parameter  $q$  versus the normalised sampling frequency ranging from a quarter to half of the reference frequency under the assumption that  $f_{ref}$  and  $f_s$  vary within  $\pm 50$  ppm. The minimum value  $q_{min}$  denotes the smallest denominator of the quotient  $p$  within the frequency tolerances. If small deviations from  $p/q = 1/3$  are applied, the maximum resolution is obtained corresponding to a minimum denominator  $q_{min} \approx 1500$  within the frequency tolerance. However, for this slave-clock application a value of  $T_{ref}/100$  is sufficient which allows the choice of standard oscillators with commonly used frequencies.

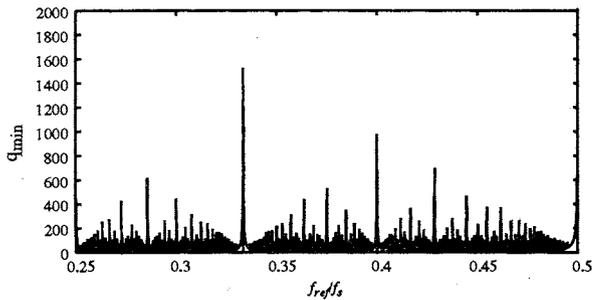


Fig. 6 Smallest denominator  $q_{min}$  for the quotient  $p$  for normalised sampling frequencies  $f_s$ . Reference frequency  $f_{ref}$  and sampling frequency  $f_s$  are varying within  $\pm 50$  ppm.

#### IV Results

The frequency domain analysis of the resampling phase detector resulted in a distinction between different error components. The DC component is associated with the resolution assuming an infinite averaging time. The other spectral error components may be suppressed if they fall within the stopband of the PLL.

The closed loop characteristic of the PLL in the SDH Equipment Clock application employing the resampling phase detector has been examined to demonstrate the feasibility of this approach. The transfer function of these clocks requires a narrowband characteristic for jitter reduction purposes. As the error components can be considered as jitter, the performance of this phase detector can be expected to be excellent in this application.

The oscillator frequency resolution is associated with the controller word length and the digital-to-frequency-conversion process. In the case of the proposed fractional-N synthesiser implementation the accuracy of this conversion process is determined by the Sigma-Delta-Modulator internal word length, which can be chosen accordingly to suppress the frequency quantisation to negligible values without significant cost effect.

The oscillator is the critical component of a slave clock, which is expected to dominate the imperfections of all other components. The characteristic is usually specified in terms of its spectral noise components. This representation supports the optimisation of the PLL transfer characteristic with respect to self-jitter reduction. The

bandwidth range of SECs is limited in ETSI and ITU-T Recommendations to between 1 Hz and 10 Hz, which reduces the degree of freedom in the design optimisation procedure.

A general analysis of the SEC components should be carried out without specific assumptions with respect to the oscillator. The oscillator as the dominating noise source is excluded from the simulation model to provide general results and to distinguish the small phase detector error from other imperfections.

In this example the SEC model has a nominal input frequency of 38.88 MHz and includes the resampling phase detector. A PI-controller is employed, which feeds the oscillator. The PLL has a cut-off frequency of 2.5 Hz with a maximum gain of 0.05 dB. The sampling frequency is chosen as parameter.

The phase detector error in combination with the oscillator frequency resolution results in low-frequency limit cycle oscillations. The performance of the SEC is demonstrated with the Time DEVIATION (TDEV). TDEV is obtained from the sequence of time interval errors by averaging and filtering and subsequent calculation of the standard deviation vs. observation period. The estimation formula for TDEV is given by equation (3), where  $x_i$  represents the time interval error sequence,  $\tau_0$  the sampling period,  $N$  the total number of samples,  $\tau$  the observation interval and  $n$  the index of the observation interval in the range from 1 to  $\lfloor N/3 \rfloor$ .

$$\text{TDEV}(\tau) = \sqrt{\frac{1}{6n^2(N-3n+1)} \sum_{i=1}^{N-3n+1} \left[ \sum_{j=0}^{n-1} (x_{i+2n+j} - 2x_{i+n+j} + x_{i+j}) \right]^2} \quad (3)$$

The characteristics of TDEV have been already discussed in the context of clock stability measurements [7]. The analysis of the SEC has been carried out in the observation interval range of 0.01 s to 100 s. The quantisation effects of the phase detector and the corresponding limit cycles dominate the noise components in this slave clock configuration.

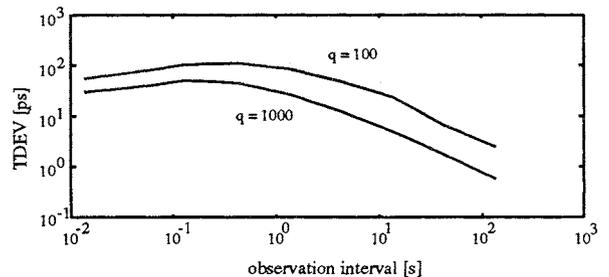


Fig. 7 Simulation results of Time DEVIATION (TDEV) observed at an SDH equipment clock (SEC;  $s_{vco}(t)$  vs.  $s_{ref}(t)$  in synchronised clock configuration) for two different phase detector sampling frequencies (corresponding to  $q = 100$  and  $q = 1000$ ;  $\tau_0 = 125 \times 10^{-6}$  s,  $N = 24000$ ).

Fig. 7 shows the simulation results for two different sampling frequencies which are related to two different values  $q$ . The upper curve shows the TDEV for  $q = 100$  and  $p = 513$ , the lower is associated with  $q = 1000$  and  $p = 3013$ . Both results exhibit a similar shape. The sec-

ond example shows an improved slave clock performance thanks to the enhanced phase detector resolution resulting from the higher value of  $q$  in this observation interval range. Both examples provide sufficient margin for the oscillator noise to comply with the TDEV mask of 3.2 ns as specified in the Draft version of ITU-T Rec. G.81s.

Fig. 8 shows measurement results of a prototype implementation of the SEC. Two different sampling frequencies have been applied which are associated with corresponding values of  $q$ . The upper curve shows the TDEV for  $q = 107$  and  $p = 208$ , the parameters for the lower curve are  $q = 1019$  and  $p = 1981$ . The results demonstrate a significant improvement of the TDEV by applying the higher value of  $q$ .

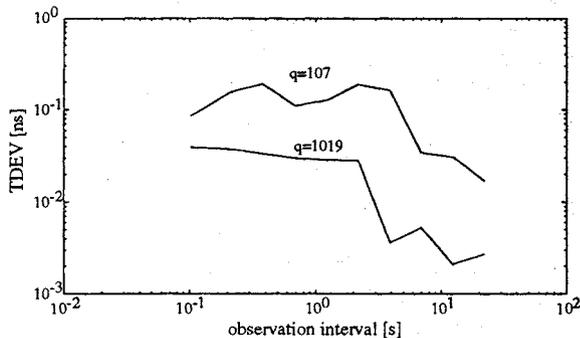


Fig. 8 Measurement result of prototype SDH equipment clock,  $\tau_0 = 0.0124$  s

Fig. 9 shows a measurement result with the same parameter values for  $p$  and  $q$ . The measurement sampling interval  $\tau_0$  has been increased to include a larger observation interval. This result confirms that the clock noise dominates the TDEV at higher observation intervals.

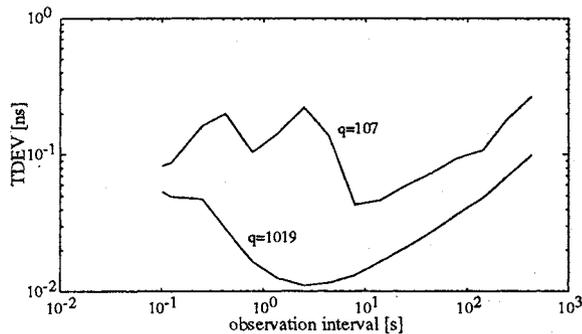


Fig. 9 Measurement result of prototype SDH equipment clock,  $\tau_0 = 0.025$  s

## V Conclusion

The network synchronisation performance in conjunction with the SDH-internal bit rate adaptation technique have a major impact on the phase transfer characteristic of SDH-based networks. In this paper network synchronisation aspects have been discussed with respect to digital exchanges and SDH networks to show the need for cheap slave clocks with small phase error in locked-mode. Slave clocks for telecommunication applications are

preferably implemented as digital PLLs. The phase detector resolution and the oscillator stability have an impact on the phase error of the PLL. The advantages of employing a synthesiser-based clock generator, which allows the optimisation of the oscillator stability regardless of the frequency range, have been discussed. It has been demonstrated that the introduction of the all-digital phase detector with independent subsequent resampling leads to enhanced resolution and full integrability. A model has been developed to determine the limitations of this approach in terms of resolution and spectral noise components and to give guidelines for practical applications with a focus on narrowband slave clocks. Simulation and measurement results of an SEC prototype implementation have demonstrated the the slave clock performance and the feasibility of this approach.

## VI References

- [1] H. L. Hartmann, E. Steiner: *Synchronization Techniques for Digital Networks*. "IEEE Journal on Selected Areas in Communications" Vol. SAC-4, No. 4, July 1986, pp. 506-513.
- [2] M. Klein, R. Urbansky: *Network Synchronization - A Challenge for SDH/SONET?*. "IEEE Communications Magazine", September 1993, pp. 42-50.
- [3] R. Urbansky: *Telecommunications Network Synchronization and Wander Generated by SDH Equipment*. Proc. 9th EFTF. March 1995, pp. 210-216.
- [4] S. M. Walters: *Digital Phase-Locked Loop with Jitter Bounded*, "IEEE Trans. on Circuits and Systems" Vol. 36, No. 7, July 1989, pp. 980-987.
- [5] T. A. D. Riley, M. A. Copeland, T. A. Kwasniewski: *Delta-Sigma Modulation in Fractional-N Frequency Synthesis*. "IEEE Journal of Solid-State Circuits" Vol. 28, No. 5, May 1993, pp. 553-559.
- [6] H. Fukinuki, I. Furukawa: *Intelligent PLL Using Digital Processing for Network Synchronization*. "IEEE Trans. on Communications" Vol. COM-31, No.12, December 1983, pp. 1295-1303.
- [7] S. Bregni, M. Carbonelli, D. De Seta, D. Perucchini, G. Zampilloni: *Clock Stability Measure Dependence on Time Error Sampling Method*. Proc. GLOBECOM 94. December 1994, pp. 1451-1455.
- [8] G. L. Pierobon, R. P. Valussi, *Jitter Analysis of a Double Modulated Threshold Pulse Stuffing Synchronizer*, "IEEE Transactions on Communications" Vol. 39, No. 4, April 1991, pp. 594-602.
- [9] I.S. Gradshteyn, I.M. Ryzhik: *Table of Integrals, Series and Products*. Academic Press, 1990, p. 40.
- [10] R. Urbansky, W. Sturm: *Design Aspects and Analysis of SDH Equipment Clocks*, to be published "ETT" Jan/Feb 1996.

## TIME FROM EUROPEAN TIME KEEPING LABORATORIES VIA TELEPHONE MODEM

Peter Hetzel, Thomas Polewka  
Physikalisch-Technische Bundesanstalt (PTB), Braunschweig

Armin Söring, Ralph Metzger  
Forschungs- und Technologiezentrum der Telekom AG (FTZ), Darmstadt

### ABSTRACT

The time keeping laboratories of Austria, Germany, Italy, Netherlands, Portugal, Sweden, United Kingdom offer telephone time services which allow a computer to receive traceable time information from the national time standards by direct telephone line. The coded time information is sent using a common format, the „European Telephone Code“.

For users requiring high accuracy the uncertainty with which the time information can be received is of particular interest.

Most of the time code generators established for this service in Europe have been developed at the Technical University of Graz (TUG) and provide the feature to measure the round-loop delay in a two-way mode and to correct for the one-way delay. Propagation delays of various telephone connections have been studied. The arrival times of the telephone time signals from the PTB, the TUG and the VSL (Van Swinden Laboratory in Delft) have been measured in the one-way and the two-way mode at the PTB in Braunschweig and the FTZ in Darmstadt. The results show that uncertainties of a few milliseconds are achievable with this technique.

### 1. INTRODUCTION

Time keeping laboratories in many countries operate telephone time services (1), (2), (3), (4). The time information generated by these services is referenced to the national time standards and is accessible to anyone using standard telephone modems.

The time protocols transmitted by telephone services normally use ASCII-character codes which operate with most standard modems and computer systems. In addition to the transmission of coded time information in the one-way mode most telephone services are designed to determine the time delay along the propagation path in the two-way mode. At the receiver end the received signal has to be echoed back to the transmitting time code generator in which the round-loop delay is measured. By advancing the transmitted signal by half the delay time the system thus can

correct the arrival time at the receiver end for signal propagation if the telephone link has approximate reciprocity.

A user who solely is interested in the coded time information normally will not do loop delay measurements. In this case he only needs a passive receiver consisting of a terminal, computer or other display. But if a user wants to make use of the two-way method to get higher timing accuracy the passive receiver must be replaced by an active transponder. This transponder can consist of the same hardware but with the additional feature to echo the received signal back either using a hardware switch or software methods.

In 1990 several time keeping laboratories in Europe agreed upon a common format for their time dissemination via the public telephone network. This „European telephone time code“ is now being distributed in the following countries:

- Austria, Technical University Graz (TUG), Graz
- Germany, Physikalisch-Technische Bundesanstalt (PTB), Braunschweig
- Italy, Istituto Elettrotecnico Nazionale (IEN), Torino
- Netherlands, Van Swinden Laboratory (VSL), Delft
- Portugal, Instituto Portugues da Qualidade (IPQ), Monte de Caparica
- Sweden, Swedish National Testing and Research Institute (SP), Boras
- UK, National Physical Laboratory (NPL), Teddington

Most of the above mentioned European laboratories are operating time code generators developed at the TUG, which have the capability to correct for the propagation delays. A detailed description of the specifications of the TUG timecode generator and the European time code format is given in (5). The aim of this paper is to add some experimental results on delay time measurements of some telephone links in Europe which were made using TUG-type time code generators.

## 2. PROPAGATION DELAY MEASUREMENTS

The block diagram in Fig. 1 shows the principle of the time dissemination via the public telephone network. At the receiving end the PC generates a signal coincident with the time marker of the incoming time protocol and delivers this signal on an I/O-port to the stop-input of a time interval counter TIC. In the European time code the change from „carriage return (CR)“ to „line feed (LF)“ (leading edge of the start bit of the LF) indicates the beginning of each transmitted second. Furtheron the PC has the capability to echo back the incoming CR-LF time markers and to send the commands "// " (two slashes) to stop the code transmission and "GDM" to do generator delay measurements to the time code generator. In the GDM-mode the TUG time code generator measures 8 round-loop delays, calculates mean value and standard deviation and advances the time marker for half the mean value determined. If the GDM has been successful the visible time marker changes from " \* " to " # ".

The modems used are CCITT-V.22-modems with the communication parameters 1200 Baud, 8 ASCII data bits, no parity and one stop bit.

In our experiments we studied the arrival time of the time markers sent via telephone lines from the PTB, the TUG and the VSL and received at the PTB in Braunschweig and at the FTZ in Darmstadt. All of the three time codes generators operated at the PTB, TUG and VSL have been of the same design developed at the TUG. An advance of 50 ms has been adjusted for all transmissions in the one-way mode.

The approximate distances between the transmitting and receiving laboratories are given in table 1. In Braunschweig the tests were performed sending the PTB-signal via the local telephone exchange station back to the PTB.

In Braunschweig and Darmstadt the same receiving equipment consecutively has been used. The tests were started at the PTB. Then the complete equipment inclusively modem was installed at the FTZ and later again at the PTB.

## 3. RESULTS

As the reference time scales UTC(K) of the transmitting laboratories and UTC(L) of the receiving laboratories are in close agreement (time differences < 2.5  $\mu$ s) the results shown in

Fig. 2 and 4 directly represent the timing accuracy. Figure 2 gives the measured time differences between UTC(L) and the arrival time of the received telephone time marker T(TM) in the one-way mode where UTC(L) was either UTC(PTB) or UTC(FTZ). Figure 4 gives the corresponding results in the two-way mode. Each data point is the mean of 25 consecutive one-second time differences measured once per day. The standard deviations belonging to these samples over 25 s are shown in Fig. 3 and 5. Figure 6 gives half the round-loop delays determined by the time code generators.

## 4. DISCUSSION AND CONCLUSION

The results of the delay time measurements show instabilities of up to a few milliseconds from day to day. No systematic delay time changes during the measurement period of about 100 days have been observed.

For the signals from the PTB and the TUG the round-loop method did not exactly correct the received time for the signal propagation delays. But as the same equipment was used and the same reproducibility and offsets were observed at the two receiver ends FTZ and PTB it has to be assumed that the telephone links have approximate reciprocity and that the offsets result from delay differences between the transmit and receive portions of the modems.

## 5. REFERENCES

1. Aida, M., Sato, T., Yamamori, S., 1985, „Time Comparison Using Telephone line“, Rev. Radio Res. (Japan) 31, pp. 115-124
2. Jackson, D., Douglas, R.J., 1986, „A Telephone-based Time Dissemination System“, Proc. 18th PTTI Meeting (USA), pp. 541-552
3. Levine, J., Weiss, M., Davis, D.D., Allan, D.W. and Sullivan, D.B., 1989, „The NIST Digital Time Service“, Proc. 21st PTTI Meeting (USA), pp. 181-190
4. Brännström, E., 1991, „European Computer Time Service“, Proc. 5th European Frequency and Time Forum, pp. 284-287
5. Kirchner, D., 1993, „Zeitübertragung über Telefonmodems“, Verlag Sprache und Technik, Groß-Bieberau, pp. 185-199

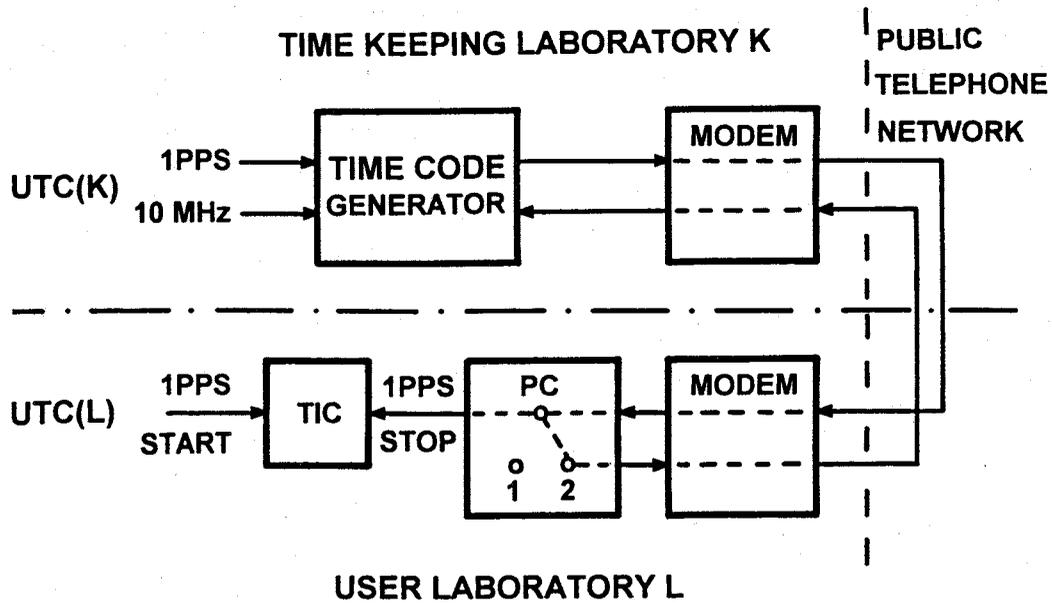


Fig. 1: Principle of the time dissemination via the public telephone network. The system has the capability to measure the round-loop delay and correct for the propagation delay if the received signal is echoed back in the PC (switch position 2). TIC Time interval counter.

RL \ TL	PTB +49531512038	VSL +3115617181	TUG +43316472366
FTZ	300 km	390 km	580 km
PTB	2 km	425 km	675 km

Table 1: Telephone numbers and approximate distances between the transmitting laboratories TL and the receiving laboratories RL.

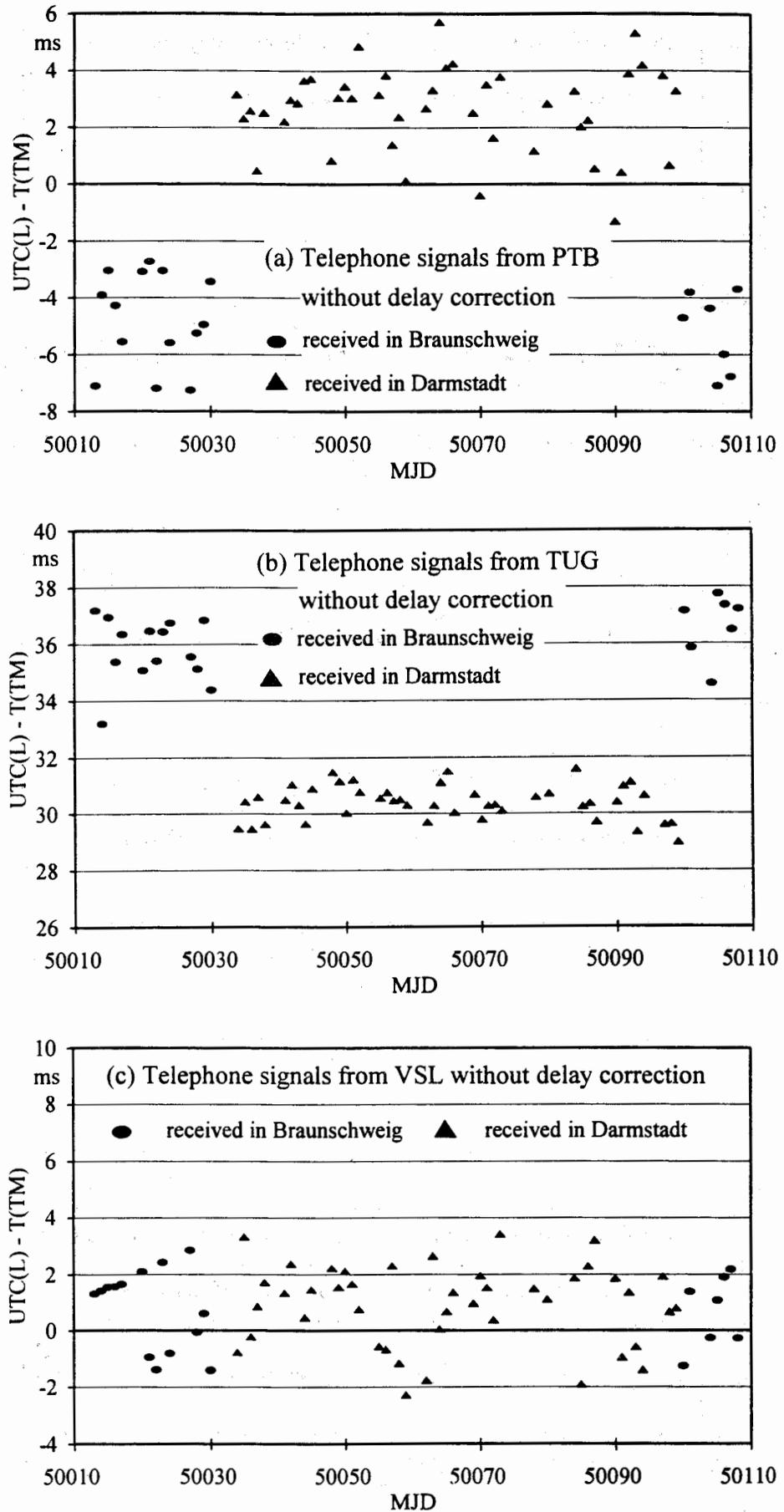


Fig. 2: Time differences  $UTC(L) - T(TM)$  between  $UTC(L)$  and the arrival time of the received telephone time marker  $t(TM)$  without delay correction. Each data point is the mean value of 25 one-second measurements performed once per day.

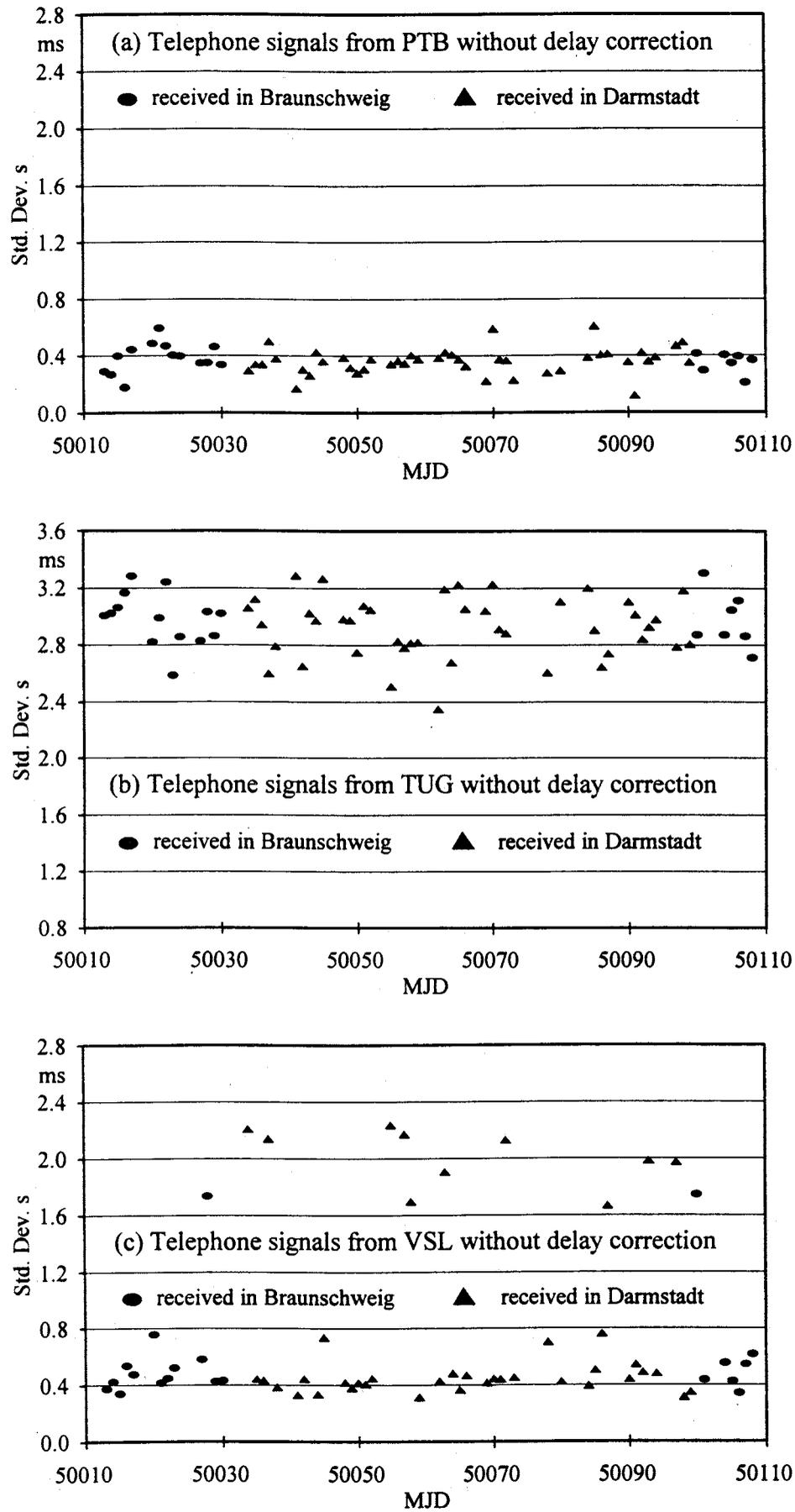


Fig. 3: Standard deviation  $s$  of the one-second measurements belonging to each corresponding mean value in the plots of figure 2.

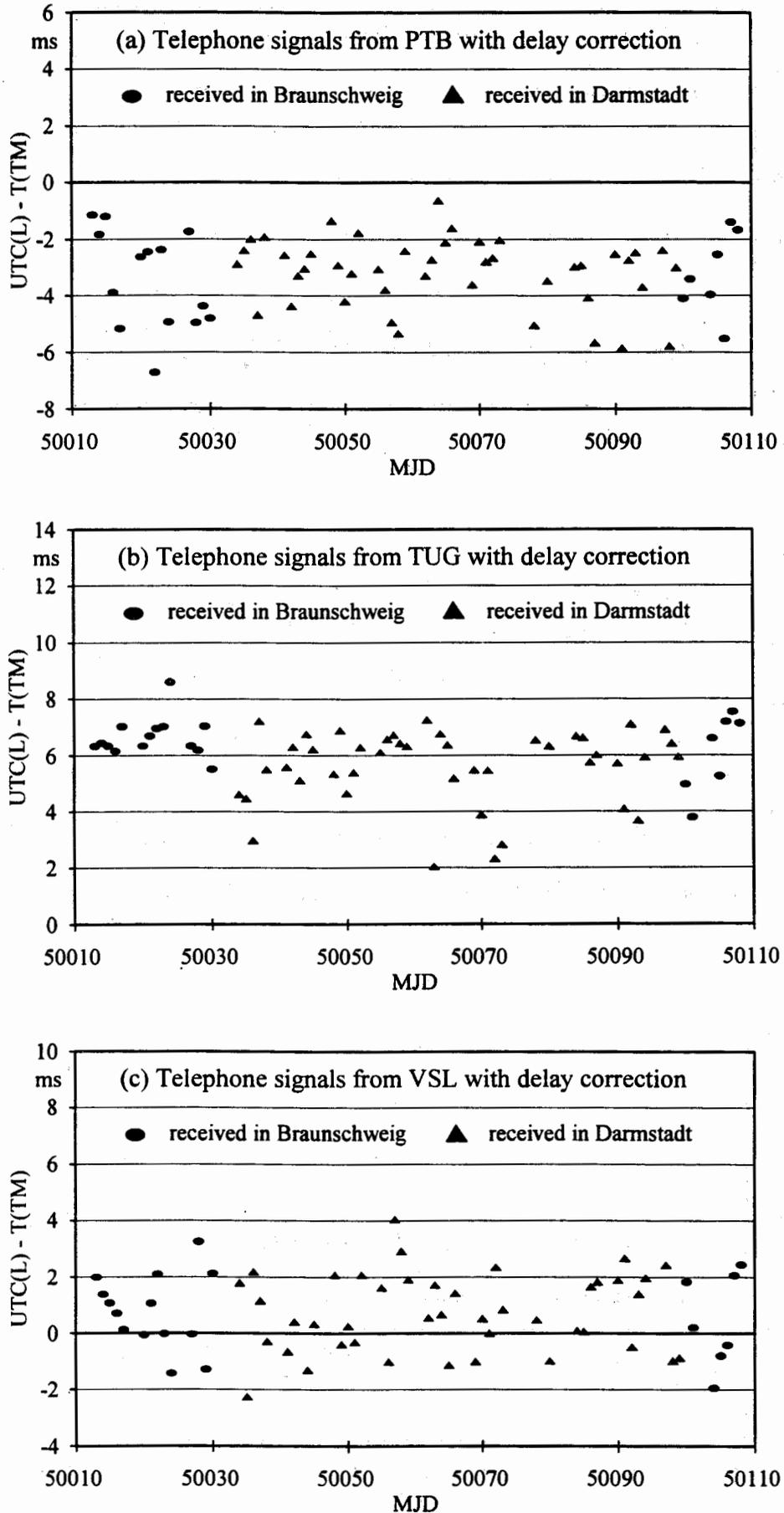


Fig. 4: Time differences  $UTC(L) - T(TM)$  between  $UTC(L)$  and the arrival time of the received telephone time marker  $t(TM)$  with delay correction. Each data point is the mean value of 25 one-second measurements performed once per day.

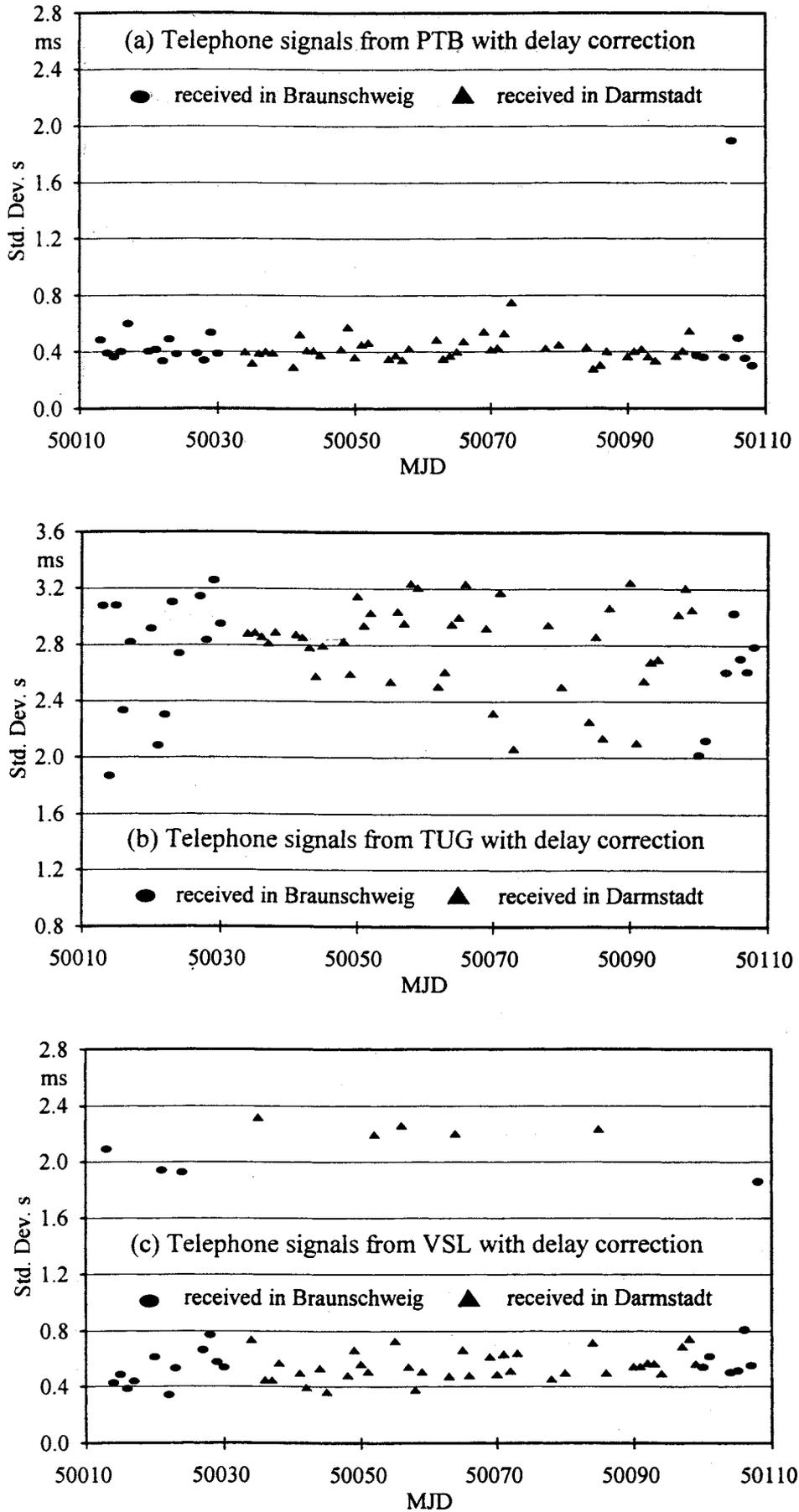


Fig. 5: Standard deviation  $s$  of the one-second measurements belonging to each corresponding mean value in the plots of figure 4.

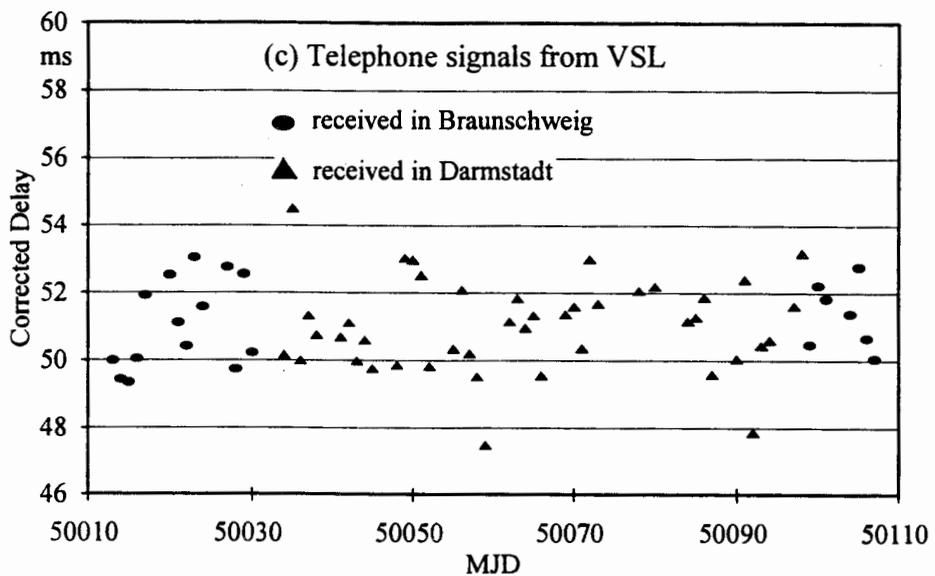
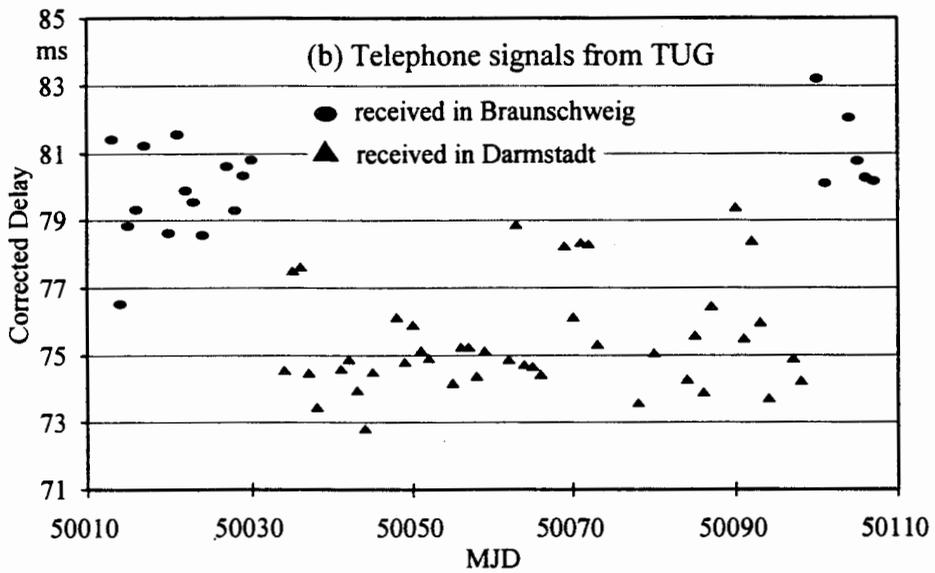
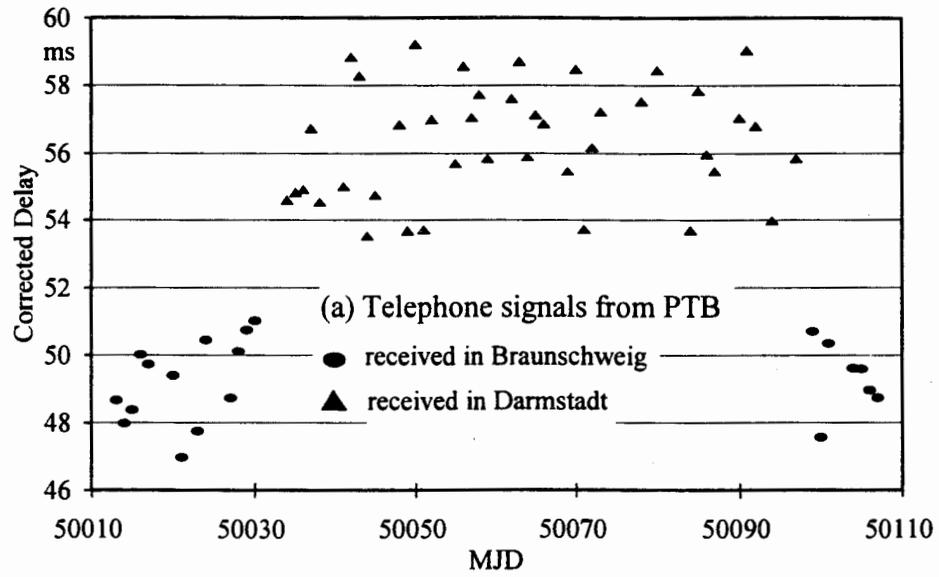


Fig. 6: Half the round-loop delay measured in the time code generator and used in the GDM-mode

